

The Multi-Airport Ground-Holding Problem in Air Traffic Control

by

PETER B. VRANAS

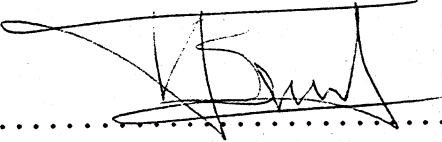
Submitted to the Department of Ocean Engineering
in partial fulfillment of the requirements for the degree of
DOCTOR OF SCIENCE IN OPERATIONS RESEARCH

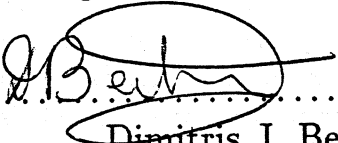
at the

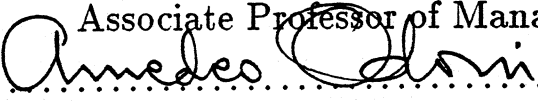
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

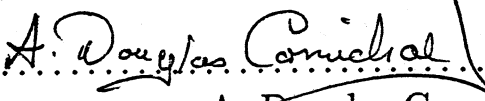
June 1992

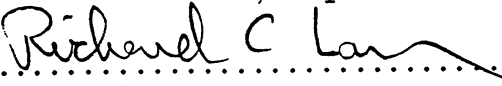
© Massachusetts Institute of Technology 1992. All rights reserved.

Author 
Department of Ocean Engineering and Operations Research Center
May 1, 1992

Certified by 
Dimitris J. Bertsimas, Thesis Supervisor
Associate Professor of Management Science

Certified by 
Amedeo R. Odoni, Thesis Supervisor
Professor of Aeronautics and Astronautics and of Civil Engineering

Accepted by 
A. Douglas Carmichael, Chairman
Graduate Committee, Department of Ocean Engineering

Accepted by 
Richard C. Larson, Chairman
Graduate Committee, Operations Research Center

The Multi-Airport Ground-Holding Problem in Air Traffic Control

by

PETER B. VRANAS

Submitted to the Department of Ocean Engineering
on May 1, 1992, in partial fulfillment of the
requirements for the degree of
DOCTOR OF SCIENCE IN OPERATIONS RESEARCH

Abstract

The Ground-Holding Problem (GHP) is defined as follows. Given a network of airports (including flight schedules with connections), how much must each flight be held on the ground before it takes off so that the total (ground plus airborne) delay cost for all flights is minimized? This problem is of considerable practical importance, given that (a) increasing congestion causes delay costs of the order of the total losses of the U.S. airline industry, and (b) other solutions to the congestion problem are unlikely to be implemented soon. Previous research on the GHP has been restricted to the single-airport case, neglecting thus "network effects" resulting from the propagation of delays between successive flights performed by the same aircraft. In this thesis the multi-airport GHP is dealt with for the first time. Both static and dynamic versions of the problem are examined. Several pure 0-1 integer programming formulations are given. These formulations capture the essential aspects of the problem while being remarkably simple, and are sufficiently flexible to accommodate various degrees of modeling detail, including: imposing airborne delays by means of *en route* speed reductions, cancelling flights, and having interdependent departure and arrival capacities by controlling runway use. The simplicity of the formulations makes large-scale GHPs tractable: extensive numerical results, achieved in reasonable computation times, are presented for networks with as many as 6 airports and 3000 flights. It is found that solving single-airport GHPs for the individual airports results in highly infeasible solutions for the network problem, and that optimal multi-airport ground holds can result in significant cost savings compared to heuristics approximating, to some extent, current ground-holding practice.

Dimitris J. Bertsimas, Thesis Supervisor
Associate Professor of Management Science

Amedeo R. Odoni, Thesis Supervisor
Professor of Aeronautics and Astronautics and of Civil Engineering

Acknowledgments

I feel deeply indebted towards my thesis supervisors, Professors Dimitris J. Bertsimas and Amedeo R. Odoni, not only for their constant academic guidance, but also for creating ideal collaboration conditions and for strongly assisting me in finding a job.

I would like to thank the Charles S. Draper Laboratory for funding my research. I would also like to thank the undergraduate students Mercury Schroepel and Beryl Castello for performing a part of the computations.

This thesis is dedicated to my parents, as a small token of my gratitude. Without their encouragement and support, the thesis would probably not have been written.

Contents

1	Introduction.	11
1.1	Motivation: the congestion problem.	11
1.1.1	The congestion problem is important.	12
1.1.2	Dire predictions.	13
1.2	Possible solutions to the congestion problem.	13
1.2.1	Long-term solutions.	15
1.2.2	Medium-term solutions.	16
1.2.3	Short-term solutions.	17
1.3	The Ground-Holding Problem (GHP).	18
1.3.1	Current ground-holding practice.	19
1.3.2	Previous research on the GHP.	22
1.4	Overview of thesis and preview of main results.	23
1.4.1	Overview of thesis.	24
1.4.2	Preview of main results.	25
2	The static GHP: formulations.	26
2.1	Introduction.	27
2.1.1	Notation.	27
2.1.2	No queueing when capacities are deterministic.	28
2.2	A first pure 0–1 integer programming formulation.	30
2.2.1	Coupling constraints.	30
2.2.2	Assignment variables.	31
2.2.3	The first formulation.	32

2.3	Infinite departure capacities.	34
2.3.1	Infinite departure capacities give zero airborne delays.	35
2.3.2	A second pure 0-1 IP formulation.	36
2.4	Flight cancellations.	38
2.4.1	A third pure 0-1 IP formulation.	39
2.4.2	A fourth pure 0-1 IP formulation.	40
2.5	Extensions.	41
2.5.1	Hub airports: more than one "next" flights.	41
2.5.2	En route speeding.	42
2.5.3	Interdependence between departure and arrival capacities.	42
2.5.4	Similar percentages of continued and noncontinued flights that are delayed.	45
3	The static GHP: results.	47
3.1	Magnitude of network effects.	48
3.1.1	The decomposed problem and its role.	48
3.1.2	Network effects insignificant when cost functions are identical.	49
3.1.3	Network effects significant when cost functions differ.	57
3.2	Importance of network effects.	59
3.2.1	Decomposed solution highly infeasible for network.	60
3.2.2	FCFS heuristic highly inefficient.	60
3.3	Finite departure capacities.	62
3.3.1	Fixed departure capacities have negligible impact.	63
3.3.2	Interdependent departure and arrival capacities have significant impact.	64
3.4	Model with flight cancellations.	64
3.4.1	Impact of cancellation costs.	65
3.4.2	A heuristic.	67
3.4.3	Performance of the heuristic.	70
3.5	Summary of results.	72

4	The dynamic GHP: formulations.	74
4.1	The dynamic deterministic GHP.	74
4.1.1	Notation.	74
4.1.2	A first dynamic deterministic formulation.	77
4.1.3	A second dynamic deterministic formulation.	79
4.1.4	Flight cancellations.	82
4.2	The dynamic probabilistic GHP.	83
4.2.1	A static probabilistic formulation.	84
4.2.2	A dynamic probabilistic formulation.	86
5	The dynamic GHP: results.	89
5.1	Comparing dynamic scenaria.	89
5.1.1	The goal.	89
5.1.2	Results and discussion.	92
5.2	Comparing dynamic policies.	97
5.2.1	The goal.	97
5.2.2	Comparing forecasting methods.	98
5.2.3	Dynamic FCFS heuristic highly inefficient.	99
5.3	Summary of results.	102
6	Conclusions.	105
6.1	Review of main points.	105
6.2	Directions for future research.	108
	Bibliography	110

List of Figures

1-1	Current ground-holding practice: given a calculated arrival profile and a capacity forecast, aircraft falling in a region where demand exceeds capacity are assigned to the first available region where capacity exceeds demand on a FCFS basis. The airborne delays which are calculated in this way are then assigned as ground holds.	21
2-1	Ground and airborne delays.	30
2-2	Modeling of coupling constraints.	31
2-3	An example of a region of possible combinations between the departure capacity d_{kt} and the arrival capacity r_{kt} of airport k at time period t	43
4-1	Modeling additional information over time in the dynamic probabilistic GHP. There are three possible capacity scenaria: $S1$, $S2$, and $S3$. Overall, $S1$ has lower capacities than $S2$, and $S2$ has lower capacities than $S3$. All three scenaria coincide between time periods τ_0 and τ_1 , and scenaria $S2$ and $S3$ coincide between time periods $\tau_1 + 1$ and τ_2 . At τ_1 , $S1$ is realized with probability p_1 and all uncertainty is eliminated. Otherwise (with probability $1 - p_1$), at τ_2 , $S2$ is realized with probability p_2 or $S3$ is realized with probability $1 - p_2$	88

List of Tables

1.1	The 33 U.S. airports which are expected to exceed 20,000 hours of annual delays by 1997 [2, p. 217]. Boldfaced airports are currently the most heavily congested.	14
2.1	Notation used throughout the thesis.	29
3.1	Behaviour of test case W around the capacity border between feasibility and infeasibility.	50
3.2	Results for the test series at the infeasibility border.	51
3.3	Computation times (in CPU seconds) for the results of the test series.	52
3.4	Results above and around the infeasibility border for $F = 1000$	54
3.5	Results above and around the infeasibility border for $F = 2000$	55
3.6	Results above and around the infeasibility border for $F = 3000$	56
3.7	Results for various upper bounds on delays.	56
3.8	Percentages of continued and noncontinued flights that are delayed in the optimal solution of (I_2) for test case W	58
3.9	Number of coupling constraints that the optimal solution of (D_2) violates (test case W).	60
3.10	Comparison of FCFS heuristic values and exact optima.	62
3.11	Results for various cases with finite departure capacities (formulation (I_1)) and with interdependent departure and arrival capacities (formulation (I_5)).	63
3.12	Results for various cases with flight cancellations.	66

3.13	Results for various cases with flight cancellations and nonuniform capacities.	67
3.14	Results for various cases with flight cancellations and three cost classes.	67
3.15	Performance of the heuristic.	71
4.1	Notation for the dynamic deterministic GHP.	76
5.1	The 15 possible dynamic scenaria (cf. Figure 4-1).	91
5.2	Values of the 15 dynamic scenaria for $F'/F = 0.20$ (upper part of the table) and for $F'/F = 0.40$ (lower part).	93
5.3	Values of the 15 dynamic scenaria for $F'/F = 0.60$ (upper part of the table) and for $F'/F = 0.80$ (lower part).	94
5.4	Most probable capacity scenario at τ_0 for various probability combinations (cf. Figure 4-1).	97
5.5	Expected values of dynamic policies and of static random-selection policies for various probability combinations.	100
5.6	Values of (a) the dynamic FCFS heuristic, and (b) (in parentheses) the corresponding exact optimum and the percentage of cost overestimation resulting from applying the heuristic (overestimation = (value of heuristic/exact optimum)-1).	103
5.7	Expected values of dynamic policies and of static random-selection policies for various probability combinations.	104

Chapter 1

Introduction.

This introductory chapter consists of four sections. Section 1.1 deals with the motivation for this thesis, namely the air traffic congestion problem in major American and European airports. It is shown that the problem is very serious and that it is expected to get worse. Section 1.2 describes possible solutions to the congestion problem. It is shown that long- and medium-term solutions are subject to significant implementation difficulties, while short-term solutions, which consist primarily of ground-holding policies, look much more promising. Section 1.3 formulates the Ground-Holding Problem (GHP), which is the topic of this thesis. Current ground-holding practice is described and previous research on the GHP is reviewed. Finally, Section 1.4 gives an overview of the organization and a preview of the main results of the thesis.

1.1 Motivation: the congestion problem.

Congestion is becoming increasingly acute in major American and European airports. This is because demand for airport use has been increasing quite rapidly during recent years while airport capacity has been more or less stagnating. In the United States, for instance, the number of commercial jet operations increased by about 27% between 1978 and 1988, while no new major commercial airport has been developed since 1974 [2, pp. 204, 220].

Congestion results in ground and airborne delays. Ground delays occur, e.g., when an aircraft is ready to depart but is not allowed to do so because there is a departure queue. Airborne delays can take the form of an *en route* speed reduction or of circling in the air in the vicinity of the destination airport because there is an arrival queue. It should be noted that delays arise not only because of limited airport capacities but also because of congestion in terminal area and *en route* airspace [3].

Delays create direct and indirect costs. The direct costs are incurred by the airline companies. Direct costs from ground delays include crew, maintenance, and depreciation costs, while direct costs from airborne delays include, in addition, fuel and safety costs. The indirect costs are incurred primarily by the passengers, and are basically opportunity costs. There are also some indirect costs to the airline companies, due to the fact that excessive delays may cause some passengers to refrain from future air travel.

1.1.1 The congestion problem is important.

The importance of the congestion problem is evidenced by the following figures. In 1991, the total losses of the U.S. airline industry were about \$2 billion; in 1990, they were about \$2.5 billion. On the other hand, the total yearly direct delay costs due to congestion are estimated to be about \$2 billion. This number is a rough estimate, but there is agreement that the costs are of the order of billions of dollars. If one adds to these direct costs the indirect opportunity costs to the passengers, one can easily arrive at a total cost of about \$6 billion. European airlines are in a similar plight [5, p. 8]. It is thus clear that congestion is a problem of undeniable significance.

The problem is exacerbated by the fact that congestion is concentrated in relatively few major airports. Although there are about 400 U.S. airports with towers to provide air traffic control, the 27 largest of them¹ account for 57% of total jet operations and incurred a 26% increase in such traffic between 1978 and 1988. At

¹Defined as those that averaged more than 275 commercial jet operations per day in 1978 [2, p. 205].

present, 21 of these 27 major airports experience more than 20,000 hours of annual flight delays.

1.1.2 Dire predictions.

Not only is the congestion problem severe, it is also expected to get worse. The Federal Aviation Administration (FAA) projects that the number of operations (arrivals plus departures) by major domestic commercial carriers will increase by one-third between 1988 and 2000. This projection is based on an average annual growth rate of 2.5%, which corresponds to the projected growth of the overall economy and is slightly lower than the growth experienced during the last decade [2, p. 215]. Moreover, the FAA estimates that, by 1997, the 33 airports appearing in Table 1.1 will exceed 20,000 hours of annual delays [2, p. 217].

On the other hand, as will be explained in the next section, no appreciable increase in capacity is expected to be realized. What can be done then?

1.2 Possible solutions to the congestion problem.

The measures that can be undertaken to alleviate congestion depend on the time horizon under consideration [3]. *Long-term* approaches have a planning horizon of 5 to 10 years and include the construction of new airports, the construction of new runways at existing airports, and the improvement of Air Traffic Control (ATC) technologies. *Medium-term* approaches have a planning horizon of 6 months to 2 years and include the modification of the temporal pattern of aircraft flow in order to eliminate periods of "peak" demand (e.g., by means of congestion pricing), and the use of larger aircraft. *Short-term* approaches have a planning horizon ranging from a few minutes to a whole day, and include, most importantly, ground-holding policies. The above approaches will be described now in more detail. It will be seen that long- and medium-term solutions are unlikely to be implemented in the U.S. to a significant extent in the foreseeable future, whereas short-term solutions appear to

AREA	AIRPORT
Northeast	Boston Logan New York City: Kennedy, LaGuardia, Newark Pittsburgh Philadelphia
South	Washington: National, Dulles Charlotte Nashville Memphis Atlanta Orlando Miami Dallas-Fort Worth Houston: Hobby and Houston Intercontinental
Midwest	Cleveland Columbus Cincinnati Detroit Chicago O'Hare St. Louis Minneapolis
West	Salt Lake City Las Vegas Phoenix Seattle-Tacoma San Francisco San Jose Los Angeles: Los Angeles International and Ontario Honolulu

Table 1.1: The 33 U.S. airports which are expected to exceed 20,000 hours of annual delays by 1997 [2, p. 217]. Boldfaced airports are currently the most heavily congested.

hold significant potential for handling congestion.

1.2.1 Long-term solutions.

Construction of new airports.

Although the federal government distributes funds for airport development, the FAA has little direct control over decisions to expand airport capacity, which are made at the local level [2, p. 218]. At the local level, however, there is usually significant community opposition to the noise that would result from constructing new airports or from expanding existing ones. Moreover, as far as expansion of major congested airports is concerned, in addition to the community opposition to noise there are the obstacles of limited space and high costs, since most such airports are located in large metropolitan areas.

Given this situation, as was pointed out above, no new major commercial airport has been developed since 1974. The only airport which is expected to provide substantial additional capacity during the 1990s is the one nearing completion at Denver. New airports are being considered in Los Angeles, Chicago, and Boston, but are unlikely to be operational by 2000 [2, p. 220].

Construction of new runways at existing airports.

Building new runways or expanding existing ones is the solution of choice for the FAA. At first glance, this solution seems hopeful: 50 of the top 100 airports have plans for additional runway capacity. A closer look, however, reveals difficulties. First, currently congested airports, such as those in New York, Boston, Washington, D.C. (National), and San Francisco, do not have plans for new runways. Second, and more important, FAA senior officials express considerable scepticism about the realization of the existing plans, again because of community opposition to noise [2, pp. 220-1].

Improvement of Air Traffic Control (ATC) technologies.

The focus now shifts to solutions that aim at improving the use of existing capacity rather than creating new capacity. During poor weather, delays increase quickly because instrument flight rule (IFR) conditions are applied, and these require greater spacing of aircraft than do visual flight rule (VFR) conditions. Use of IFR rather than VFR can reduce airport capacity by as much as 50% [2, p. 221]. Adopting less stringent ATC rules could increase peak capacity during inclement weather by up to 40% at congested airports such as Boston and San Francisco. The challenge, however, is to relax ATC rules without compromising safety. This can be done by introducing new radars with shorter response times and by developing new procedures for aborted landings to minimize the risk of midair collisions.

Since 1981 the FAA has been pursuing the National Airspace System Plan (NAS Plan), with the aim to replace outdated technology in the ATC system. The NAS Plan was originally promulgated as a 10-year, \$12-billion program, but now is a multiyear \$27-billion capital improvement program. The FAA is optimistic that major improvements will begin appearing in the next few years. There are, however, grounds for scepticism. Most components of the NAS Plan have been delayed for a variety of reasons, and the original contract for one subsystem, the Microwave Landing System, was cancelled by the FAA because tests did not meet FAA specifications. It appears that the new technologies may not become available within the near future [2, pp. 222-4].

1.2.2 Medium-term solutions.

Congestion pricing for runways.

Congestion usually occurs only during peak periods (typically in the morning and early evening), rather than throughout the day. In general, however, landing fees do not vary with demand, but are calculated according to aircraft weight. Economic analyses [2, p. 226] indicate that, if the landing fee for a period of the day were to increase, delays would be reduced considerably, although some schedule frequency

would be lost, because some direct service would be substituted for connecting service. The overall effect would be highly beneficial.

Congestion pricing, however, encounters significant legal and practical obstacles. The legal obstacles are illustrated by a recent experiment of Boston Logan Airport with higher fees: the airport was forced to abandon its experiment because the Department of Transportation found that the new fees discriminated against general aviation. In general, federal aid to airports has long been conditioned on an "open access" policy that prohibits airports from discriminating against classes of users. As far as the practical problems with congestion pricing are concerned, one of the most significant ones is the calculation of the optimal fee. In conclusion, congestion pricing is unlikely to be widely implemented in the foreseeable future.

Use of larger aircraft.

Larger aircraft would absorb the expected increase in demand without necessitating additional airport capacity. The problem, however, is that airlines lack incentives to replace their aircraft by larger ones. Congestion pricing would provide such incentives, but this is not likely to happen soon.

1.2.3 Short-term solutions.

The focus now shifts to short-term policies which consider airport capacities and flight schedules as fixed for any given day. These policies consist in adjusting the flow of aircraft on a real-time basis and are referred to as *flow management* policies. They include [3]:

(1) Imposing "ground holds", i.e., delaying the departure of an aircraft by not allowing it to start its engines and leave its gate or parking area even if it is ready to depart.

(2) Imposing *en route* speed controls in order to properly time aircraft arrivals.

(3) Modifying *en route* the flight plans of selected aircraft in order to bypass congested areas.

Ground-holding policies, falling under (1) above, make sense in the following situation. If an aircraft departs on time, it encounters congestion and incurs an airborne delay upon arrival at its destination; but if its departure is delayed, the aircraft arrives at its destination at a later time when no congestion is present and no airborne delay is incurred. Therefore, the objective of ground-holding policies is to absorb airborne delays on the ground.

Ground-holding policies are motivated by the following two fundamental facts. First, airborne delays are much costlier than ground delays, because airborne delays include fuel and safety costs. Typical numbers for ground delay costs range from about \$20 to about \$65 per minute, according to aircraft size, and airborne delays are usually 30–50% higher, according to fuel price. Second, as was pointed out above, airport capacity is highly variable even for a given configuration of runways, because it depends heavily on weather (visibility, wind, precipitation, cloud ceiling). It is not unusual for the capacity of an airport to be reduced by 50% in poor weather, due to the increase in minimum separation between aircraft and the unavailability of some runways when one shifts from VFR to IFR conditions. Given the above two facts, it is seen that there is significant potential for readjusting aircraft flow given weather (hence airport capacity) forecasts, and that such readjustment can result in significant cost reduction if ground delays are substituted for the much costlier airborne delays.

As its title indicates, this thesis deals with the problem of finding optimal ground-holding policies. This problem will be described now in more detail.

1.3 The Ground-Holding Problem (GHP).

The GHP can be described in the following way. We are given a network of airports, a planning horizon typically consisting of a day or a portion of a day, and the schedule of flights to be performed within the planning horizon between the airports in the network as well as from and to the external world. A flight is defined as a trip from one airport to another. The schedule consists of the following data for each

flight: its departure airport, its arrival airport, its scheduled departure time, its scheduled arrival time, and the identification number of the next flight scheduled to be performed by the same aircraft. We are also given the ground delay cost function and the airborne delay cost function of each flight, as well as departure and arrival capacity forecasts for all airports of the network throughout the planning horizon. The objective is to assign a ground hold to each flight so as to minimize the total (ground plus airborne) delay cost for all flights.

The above description is oversimplified because it ignores two essential aspects of the problem: the real-world GHP is *dynamic* and *probabilistic*. It is dynamic because weather (hence capacity) forecasts are updated during the course of the day, and ground holds should be revised whenever new information becomes available. It is probabilistic because weather prediction itself is probabilistic: capacity forecasts are not accurate more than a few hours in advance. The complications arising from the probabilistic nature of the problem become clear when one considers that two types of “errors” may occur. First, if capacity turns out to be lower than expected, significant unforeseen airborne delays will occur. Second, if capacity turns out to be higher than expected, a portion of this capacity may remain unused while aircraft incur unnecessary ground delays. The interplay between overly optimistic and unduly conservative strategies constitutes the challenge of the GHP.

Although the real-world GHP is dynamic and probabilistic, one may want to explore simplified *static* or *deterministic* versions of the problem. This is because the full-scale GHP is a very large problem: typical congested airports have about 500–600 flights per day.

1.3.1 Current ground-holding practice.

The importance of ground-holding policies has long been recognized. The FAA has been operating for several years in Washington, D.C. an Air Traffic Control System Command Center (ATCSCC, formerly called the Central Flow Control Facility), equipped with outstanding information-gathering capabilities, including weather fore-

casts, up-to-the-minute data on the status of airborne traffic throughout the country, and projections of traffic levels over several hours [3]. ATCSCC, however, relies primarily on the judgement of its expert air traffic controllers rather than on any decision-support or optimization models to develop flow management and ground-holding strategies.²

The way in which the ATCSCC assigns ground holds is essentially the following.

(1) The ATCSCC calculates (through what is essentially a deterministic simulation model [5]) the arrival profiles that would result from letting each aircraft depart on time (cf. Figure 1-1).

(2) Given the calculated arrival profiles and the capacity forecasts, the ATCSCC calculates, for each flight f , the airborne delay that f would incur if it were to depart on time.

(3) The ATCSCC assigns to flight f a ground hold equal to the above calculated airborne delay, provided that the delay is not below a certain threshold (typically about 15 minutes).

(4) Whenever new capacity forecasts become available, the ATCSCC updates ground holds by following again steps (1)–(3).

This approach amounts to assigning flights to available capacity on a first-come-first-served (FCFS) basis. Even if the problem is deterministic, i.e., if capacity forecasts are perfectly accurate, the approach may not be optimal, because it neglects the so-called “network” (or “down-the-road”) effects. Network effects are due to the fact that each of a large number of aircraft typically performs more than one flight on any given day, so that, when a flight is delayed, the next flight scheduled to be performed by the same aircraft may also have to be delayed. Moreover, at a “hub” airport, a late arriving aircraft may delay the departure of several flights, given current airline scheduling practices which emphasize passenger transfers. Therefore, not all flights are equal: delaying flights that are continued will typically result in a propagation of

²There is no analogous system in operation anywhere else in the world.

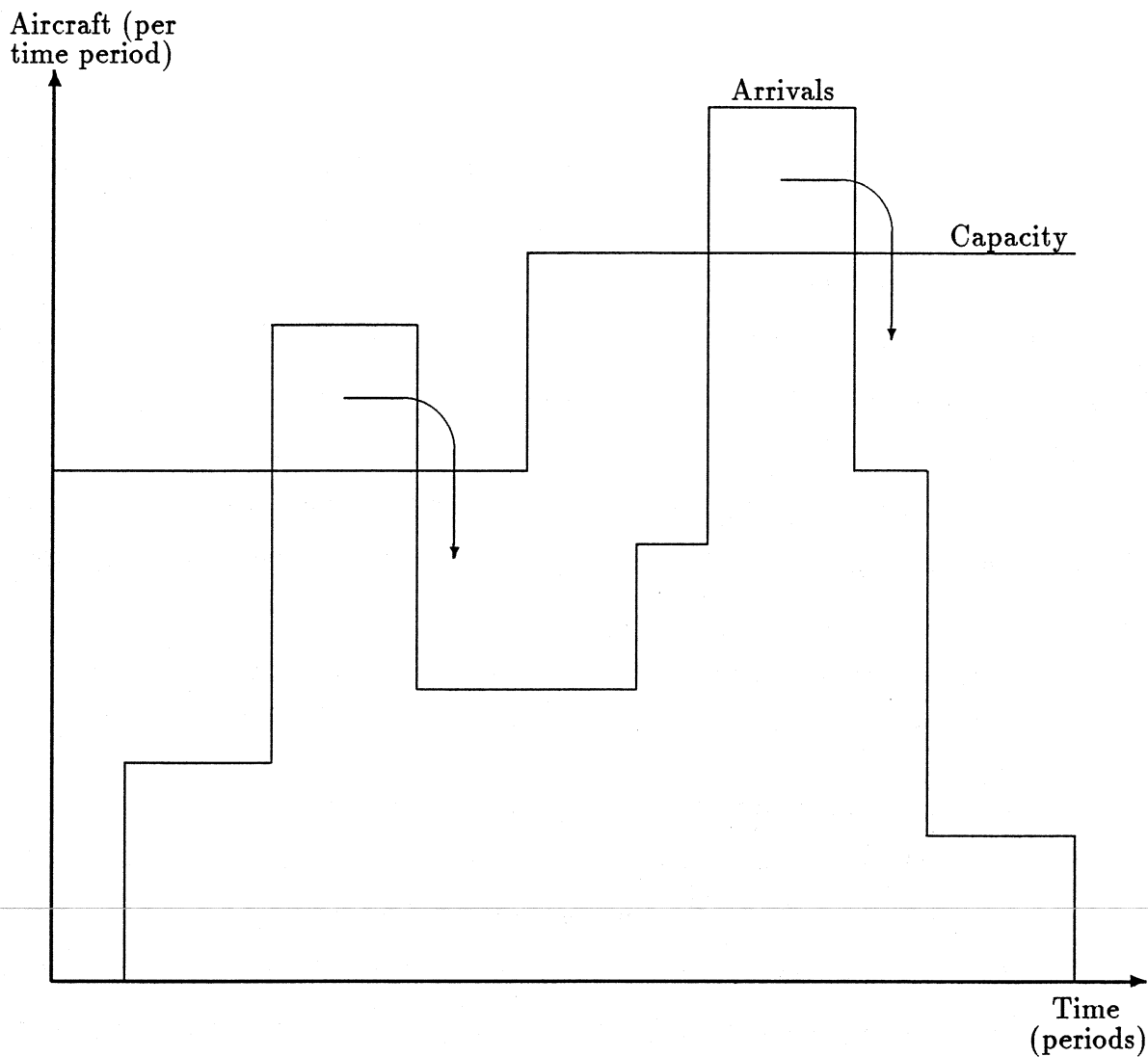


Figure 1-1: Current ground-holding practice: given a calculated arrival profile and a capacity forecast, aircraft falling in a region where demand exceeds capacity are assigned to the first available region where capacity exceeds demand on a FCFS basis. The airborne delays which are calculated in this way are then assigned as ground holds.

delays throughout the network. By neglecting this fact, current practice may result in a suboptimal assignment of ground holds.³

1.3.2 Previous research on the GHP.

Previous research on the GHP has been restricted to the single-airport problem and has neglected network effects. The single-airport GHP is the following special case of the multi-airport problem: all airports of the network except one have infinite departure and arrival capacities.⁴ Although this is a simplified version of the GHP, it is of interest in its own right and it provided significant insights for the multi-airport case.

The first systematic description of the GHP is given by Odoni [3]. Terrab [5] gave the following pure 0–1 integer programming formulation for the static deterministic single-airport GHP:

$$(I_0) \min \sum_{f=1}^F \sum_{t=r_f}^T c_{ft}^g v_{ft}$$

$$\text{s.t. } \sum_{f=1}^F v_{ft} \leq R(t), \quad t \in \{1, \dots, T\}; \quad (1.1)$$

$$\sum_{t=r_f}^T v_{ft} = 1, \quad f \in \{1, \dots, F\}; \quad (1.2)$$

$$v_{ft} \in \{0, 1\}, \quad f \in \{1, \dots, F\}, t \in \{r_f, \dots, T\}. \quad (1.3)$$

There are F flights and T time periods. The scheduled arrival time of flight f is r_f , and c_{ft}^g is the ground delay cost incurred if flight f is assigned to land at time period t . Decision variable v_{ft} is 1 if flight f is assigned to land at time period t and 0 otherwise. Constraints (1.2) ensure that, for a given flight f , exactly one v_{ft} will be equal to 1 and the rest will be 0. Constraints (1.1) are the capacity constraints,

³The performance of FCFS heuristics approximating, to some extent, current ground-holding practice will be examined, by means of computational results, in Subsection 3.2.2 for the static case and in Subsection 5.2.3 for the dynamic case.

⁴Call the airport with finite arrival and departure capacities Z . Strictly speaking, it is also presupposed in the single-airport problem that the following possibility never occurs: a delayed flight leaves from Z and causes, through a chain of continued flights, the delay of another flight arriving later on to Z .

$R(t)$ being the landing capacity at period t .⁵

Terrab gave a reformulation of (I_0) as a minimum cost flow program, as well as a second reformulation as an assignment problem. Then he proposed an efficient algorithm for solving (I_0) when cost functions belong to a special class. He also suggested a two-airport formulation and a closed-network three-airport formulation for the static deterministic GHP.

As far as the probabilistic GHP is concerned, Andreatta and Romanin-Jacur [1] proposed a dynamic programming algorithm for the single-airport static probabilistic GHP with one time period. Terrab extended their algorithm to the multi-period single-airport case and gave several heuristics. Finally, Richetta [4] dealt with the single-airport probabilistic and dynamic GHP by means of stochastic programming with recourse. It seems that no significant research has been done to date concerning the effects of ground-holding policies on an entire network of airports.

1.4 Overview of thesis and preview of main results.

This thesis constitutes the first attempt to deal with the *multi*-airport GHP. It focuses on both static and dynamic versions of the deterministic multi-airport GHP, and has the following two objectives. First, to show that the multi-airport GHP is tractable: to propose efficient mathematical programming formulations which can be used in order to assign ground holds on a real-time basis for a network of airports. These formulations must be sufficiently flexible to accommodate various degrees of modeling detail. Second, to investigate the extent of the cost reductions that can be achieved by applying optimal ground-holding policies. This second objective falls into two parts. First, to compare the cost of optimal ground-holding policies with the cost arising from heuristics approximating, to some extent, current ground-holding practice (cf.

⁵Formulation (I_0) was described in some detail here because it is similar in spirit to some of the formulations in this thesis (cf., especially, formulation (I_2) in Subsection 2.3.2).

Subsection 1.3.1). Second, to find the magnitude of the network effects, i.e., to compare the optimal values of multi-airport with single-airport GHPs.

1.4.1 Overview of thesis.

The bulk of the thesis consists of Chapters 2 through 5. Chapters 2 and 3 deal with the static GHP (where decisions are taken once for all at the beginning of the day), while Chapters 4 and 5 deal with the dynamic GHP (where decisions are updated during the course of the day as new capacity forecasts become available).

Chapter 2 proposes several pure 0–1 integer programming formulations for the static GHP. Some of these formulations take into account the possibility of cancelling flights; others take advantage of the fact that departure capacities are typically larger than arrival capacities; yet others take into account that the split between departure and arrival capacities can be changed by modifying runway use. All formulations are quite compact: they have numbers of constraints and numbers of variables which are small linear multiples of the number of flights. This compactness is responsible for the tractability of multi-airport GHPs.

Chapter 3 presents extensive computational results based on the formulations of Chapter 2. These results are used to investigate the behaviour of the multi-airport GHP under various combinations of input parameters, and to ascertain the extent of cost reductions resulting from ground-holding policies. Chapter 3 also describes and evaluates a heuristic finding a feasible solution of the integer program based on the optimal solution of the linear programming relaxation.

Chapter 4 puts forward formulations for the dynamic GHP. These formulations are basically extensions of the static formulations of Chapter 2. A formulation for the probabilistic GHP is also presented. Chapter 5 describes a variety of possible dynamic ground-holding policies and compares them on the basis of computational results corresponding to the formulations of Chapter 4.

Finally, Chapter 6 concludes by reviewing the main points and by proposing di-

rections for future research.

1.4.2 Preview of main results.

The thesis accomplishes both of its aforementioned objectives.

First, the multi-airport GHP is demonstrated to be tractable: problems with as many as 6 airports and 3000 flights are routinely solved in reasonable computation times. Moreover, it is seen that one does not have to solve the integer programs exactly, because the optimal solutions of the linear programming relaxations typically have few noninteger values, so that they can be rounded off (by means of a heuristic presented in Subsection 3.4.2) to give good feasible solutions of the integer programs.

Second, the extent of cost reductions achieved by optimal ground-holding policies is investigated. It is shown that one must distinguish the *magnitude* of network effects from their *importance*. Network effects can be of large magnitude in the general case, but they are very small in the special case where all flights have identical cost functions. This special case is of practical interest, because different cost functions might be considered a form of discrimination against classes of users. Even with identical cost functions, however, network effects are of considerable importance: optimal ground holding policies can result in significant cost reductions with respect to FCFS heuristics approximating, to some extent, current ground-holding practice. Moreover, even with identical cost functions, ignoring network effects (i.e., solving a single-airport GHP for each airport of the network) results in highly infeasible solutions of the multi-airport GHP.

The bottom line is that optimal ground-holding policies seem to be implementable with relative ease and would probably result in a significant alleviation of the congestion problem.

Chapter 2

The static GHP: formulations.

This chapter presents several pure 0–1 integer programming (IP) formulations for the static deterministic multi-airport GHP and is divided into five sections. Section 2.1 introduces the notation that will be used throughout the thesis and explains why the assumption that the problem is deterministic enables one to dispense with departure and arrival queueing without loss of generality. Section 2.2 describes a first IP formulation which also allows for the possibility of imposing airborne delays, e.g., by means of *en route* speed reductions. Section 2.3 discards this additional possibility and shows that, when cost functions are linear and departure capacities are infinite, if the problem has an optimal solution, it also has an optimal solution in which there are no airborne delays. This result allows the construction of a compact second IP formulation for the case of infinite departure capacities. Section 2.4 presents two further IP formulations which correspond to the formulations of Sections 2.2 and 2.3 but have the additional possibility of cancelling flights. Finally, Section 2.5 proposes some extensions of the above formulations. For instance, it is explained how some of them can be modified to handle the case where departure and arrival capacities are interdependent (e.g., because some take-off runways can be used for landings when arrivals accumulate, or vice-versa).

2.1 Introduction.

2.1.1 Notation.

Consider a set of airports $\mathcal{K} = \{1, \dots, K\}$ and an ordered set of time periods $\mathcal{T} = \{1, \dots, T\}$. For instance, \mathcal{K} might be the set of the 20 or so busiest U.S. airports, and \mathcal{T} might be a set of 64 time periods of 15 minutes each, amounting to a time horizon of 16 hours, i.e., the portion of a day from 7am to 11pm (when most flights take place). Consider also a set of flights $\mathcal{F} = \{1, \dots, F\}$. (Note that a single aircraft may perform several of these flights.) \mathcal{F} is the set of all flights of interest, e.g., all flights departing from an airport in \mathcal{K} and arriving to another airport in \mathcal{K} . This interpretation of \mathcal{F} corresponds to a *closed* network of airports, for which departures from and arrivals to the external world are not considered important. If an *open* network of airports is to be considered, then \mathcal{F} will be the set of all flights departing from an airport in \mathcal{K} or arriving to an airport in \mathcal{K} (or both).

For each flight $f \in \mathcal{F}$, the following data are assumed to be known: $k_f^d \in \mathcal{K}$, the airport from which f is scheduled to depart; $k_f^a \in \mathcal{K}$, the airport to which f is scheduled to arrive; $d_f \in \mathcal{T}$, the scheduled departure time of f ; $r_f \in \mathcal{T}$, the scheduled arrival time of f (so the scheduled travel time is $t_f = r_f - d_f$); $c_f^g(\cdot)$, the ground delay cost function of f (whose argument is the ground delay of f in time periods); and $c_f^a(\cdot)$, the airborne delay cost function of f (whose argument is the airborne delay of f in time periods). For each $(k, t) \in \mathcal{K} \times \mathcal{T}$, the departure capacity $D_k(t)$ and the arrival capacity $R_k(t)$ (in number of aircraft) are also given. Since this chapter deals with *deterministic* versions of the GHP, these capacities are considered fixed numbers rather than random variables. Typical values, e.g., for Boston Logan Airport in good weather are 15–16 landings per time period of 15 minutes; departure capacities are 15–30% higher.

Consider finally the set $\mathcal{F}' \subset \mathcal{F}$ of those flights that are “continued”. A flight is said to be continued if the aircraft which is scheduled to perform it is also scheduled to perform at least one more flight later on in the day. For each flight $f' \in \mathcal{F}'$, we are

given the next flight f scheduled to be performed by the same aircraft.

Define the decision variables $g_f, f \in \mathcal{F}$, equal to the number of time periods that flight f is held on the ground before being allowed to take off, and the decision variables $a_f, f \in \mathcal{F}$, equal to the number of time periods that flight f is further held in the air (e.g., by means of an *en route* speed reduction) before being allowed to land. Since this chapter deals with *static* versions of the GHP, it is assumed that these ground and airborne holds are decided once for all at the beginning of the day for all flights.

Table 2.1 summarizes the above notation for reference purposes. Table 2.1 also includes some symbols which will be defined in the sequel.

2.1.2 No queueing when capacities are deterministic.

Consider now the following description of the real-world situation. If a flight f is scheduled to depart at period d_f and is delayed on the ground for g_f periods, then it will be *available to depart* at period $d_f + g_f$. Will it actually depart at that period? This will depend on whether the total number of aircraft *available to depart* from airport k_f^d at that time period will exceed or not the available departure capacity. If it does exceed it, then the aircraft performing flight f will have to wait q_f^d time periods in the *departure queue*. q_f^d will depend on the particular *service discipline* adopted for the departure queue. So flight f will actually take off at period $d_f + g_f + q_f^d$. Since flight f will be further delayed in the air for a_f time periods, it will arrive at its destination, airport k_f^a , and will be *available to land* at period $r_f + g_f + q_f^d + a_f$. Will it actually land at that period? This will depend on whether the total number of aircraft *available to land* at airport k_f^a at that period will exceed or not the available landing capacity. If it does exceed it, then the aircraft performing flight f will have to wait q_f^a time periods in the *arrival queue*, and will actually land at period $r_f + g_f + q_f^d + a_f + q_f^a$. The total cost corresponding to flight f will be the sum of $c_f^g(g_f + q_f^d)$ (the ground

Symbol	Denotation
$\mathcal{K} = \{1, \dots, K\}$ $\mathcal{T} = \{1, \dots, T\}$ $\mathcal{F} = \{1, \dots, F\}$ $\mathcal{F}' \subset \mathcal{F}$	Set of airports. Set of time periods. Set of flights. Set of continued flights.
K T F F'	Number of airports. Number of time periods. Number of flights. Number of continued flights.
$k_f^d \in \mathcal{K}$ $k_f^a \in \mathcal{K}$	Departure airport of flight f . Arrival airport of flight f .
$d_f \in \mathcal{T}$ $r_f \in \mathcal{T}$ $t_f = r_f - d_f$	Scheduled departure time of flight f . Scheduled arrival time of flight f . Scheduled travel time of flight f .
$c_f^g(\cdot)$ $c_f^a(\cdot)$	Ground delay cost function of flight f . Airborne delay cost function of flight f .
$s_{f'}$	Slack time of continued flight f' .
$D_k(t)$ $R_k(t)$	Departure capacity of airport k at period t . Arrival capacity of airport k at period t .
g_f a_f	Ground delay decision variable for flight f . Airborne delay decision variable for flight f .
u_{ft} v_{ft}	Departure assignment decision variables for flight f . Arrival assignment decision variables for flight f .
q_f^d q_f^a	Departure queueing of flight f . Arrival queueing of flight f .
G_f A_f	Upper bound on the ground delay of flight f . Upper bound on the airborne delay of flight f .
\mathcal{T}_f^d \mathcal{T}_f^a	Set of allowable departure periods for flight f . Set of allowable arrival periods for flight f .
z_f M_f	Cancellation decision variable for flight f . Cancellation cost of flight f .

Table 2.1: Notation used throughout the thesis.

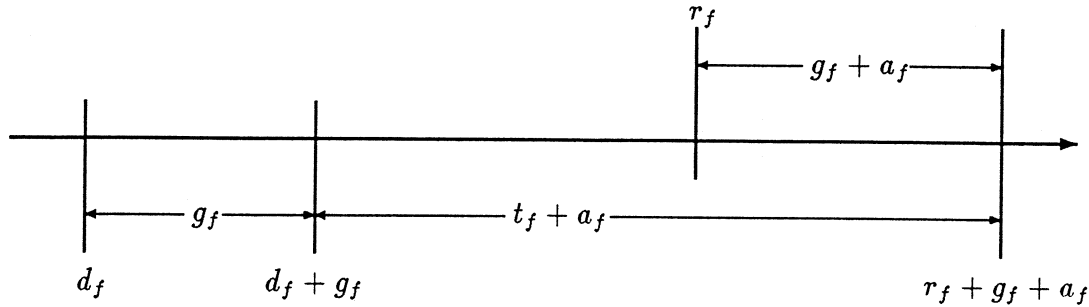


Figure 2-1: Ground and airborne delays.

delay cost) and $c_f^a(a_f + q_f^a)$ (the airborne delay cost).¹

Because we are examining the *deterministic* case, the above description can be considerably simplified. In fact, it makes little sense to assign to a flight f a ground hold of g_f time periods such that f will have to further wait q_f^d time periods in the departure queue: one might as well assign to f a total ground hold of $g_f + q_f^d$ time periods such that f will not have to wait at all in the departure queue. Similar remarks hold for airborne delays. Given this simplification, the total ground delay of flight f will be g_f , and its total airborne delay will be a_f , resulting in a cost of $c_f^g(g_f) + c_f^a(a_f)$. The situation is depicted in Figure 2-1.

2.2 A first pure 0–1 integer programming formulation.

2.2.1 Coupling constraints.

Network effects will be taken into account in the following way. For each continued flight $f' \in \mathcal{F}'$, we are given the “slack” or “absorption” time $s_{f'}$. The slack is defined as the number of time periods such that, if f' arrives at its destination at most $s_{f'}$

¹ Actually the ground delay cost components corresponding to g_f and to q_f^d are different, because q_f^d , unlike g_f , is incurred with the engines running, so that it includes fuel costs. This complication, however, makes no difference for the argument that follows.

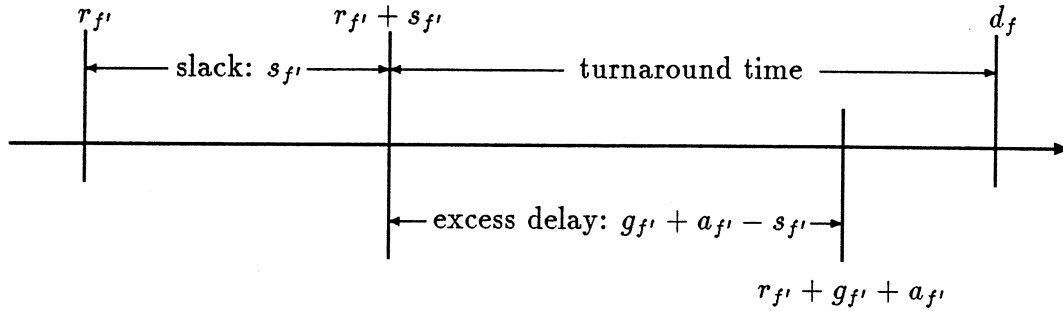


Figure 2-2: Modeling of coupling constraints.

time periods late, the departure of the next flight f is not affected, whereas if f' lands with a delay greater than the slack, the “excess delay” of f' (i.e., the delay minus the slack) is transferred to the next flight f . In the latter case, the next flight f will incur a ground delay at least equal to the excess delay of f' . The situation is depicted in Figure 2-2, where it can be seen that the slack $s_{f'}$ is equal to the difference between (i) the time interval between the scheduled departure time of f and the scheduled arrival time of f' , and (ii) the minimum “turnaround” time of the aircraft performing both flights.

2.2.2 Assignment variables.

The *delay decision variables* g_f and a_f were introduced above. Now we introduce the *assignment decision variables* u_{ft} , defined to be 1 if flight f finally takes off at period t (i.e., if $r_f + g_f = t$) and 0 otherwise, and v_{ft} , defined to be 1 if flight f finally lands at period t (i.e., if $r_f + g_f + a_f = t$) and 0 otherwise. These new decision variables are introduced because the capacity constraints cannot be expressed in a simple linear way in terms of the more natural delay decision variables.

Moreover, since we don't want to have excessive ground or airborne delays, We also introduce upper bounds on those delays. G_f is the maximum number of time periods that flight f may be held on the ground, and A_f is the maximum number of time periods that flight f may be held in the air. Introduction of these bounds results in

no loss of generality, since they can be arbitrarily large. In practice, however, typical values are $G_f = 4-5$ and $A_f = 2-3$, corresponding to maximum ground and airborne delays of about one hour and half an hour, respectively.

Given the above setup, the set \mathcal{T}_f^d of time periods to which flight f may be assigned to take off is given by:

$$\mathcal{T}_f^d = \{t \in \mathcal{T} : d_f \leq t \leq \min(d_f + G_f, T)\}. \quad (2.1)$$

Similarly, the set \mathcal{T}_f^a of time periods to which flight f may be assigned to land is given by:

$$\mathcal{T}_f^a = \{t \in \mathcal{T} : r_f \leq t \leq \min(r_f + G_f + A_f, T)\}. \quad (2.2)$$

For every flight f , exactly one of the variables u_{ft} must be equal to 1 and the others must be equal to zero, and similarly for the variables v_{ft} . Given this fact, the delay variables g_f and a_f can be expressed in terms of the assignment variables u_{ft} and v_{ft} in the following way:

$$g_f = \sum_{t \in \mathcal{T}_f^d} tu_{ft} - d_f, \quad f \in \mathcal{F}; \quad (2.3)$$

$$a_f = \sum_{t \in \mathcal{T}_f^a} tv_{ft} - r_f - g_f, \quad f \in \mathcal{F}. \quad (2.4)$$

We are now ready to give a first pure 0–1 integer programming formulation of the static deterministic multi-airport GHP.

2.2.3 The first formulation.

$$(I_1) \quad \min \quad \sum_{f=1}^F c_f^g g_f + c_f^a a_f$$

$$\text{s.t.} \quad \sum_{f:k_f^d=k} u_{ft} \leq D_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}; \quad (2.5)$$

$$\sum_{f:k_f^a=k} v_{ft} \leq R_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}; \quad (2.6)$$

$$\sum_{t \in T_f^d} u_{ft} = 1, \quad f \in \mathcal{F}; \quad (2.7)$$

$$\sum_{t \in T_f^a} v_{ft} = 1, \quad f \in \mathcal{F}; \quad (2.8)$$

$$g_{f'} + a_{f'} - s_{f'} \leq g_f, \quad f' \in \mathcal{F}'; \quad (2.9)$$

$$a_f \geq 0, \quad f \in \mathcal{F}; \quad (2.10)$$

$$u_{ft}, v_{ft} \in \{0, 1\}. \quad (2.11)$$

In the objective function of (I_1) , the cost functions $c_f^g(t), c_f^a(t)$ were replaced by their linear counterparts $c_f^g t, c_f^a t$ (c_f^g, c_f^a being the constant marginal costs). Constraints (2.5) and (2.6) are the departure and arrival capacity constraints, respectively. Recall that these have to be satisfied because we choose g_f and a_f such that the queueing delays q_f^d, q_f^a are 0 (and that we can do this because the problem is deterministic). (Strictly speaking, we also need the condition that G_f and A_f be sufficiently large.) Constraints (2.7) (together with (2.11)) ensure that, for a given f , exactly one u_{ft} will be 1 and the rest will be 0. Similarly for constraints (2.8).

Constraints (2.9) are the coupling constraints: they transfer any excess delay of flight f' to its next flight f . In fact, constraints (2.9) say that, if flight f' arrives at its destination with a total delay $g_{f'} + a_{f'}$ which is greater than $s_{f'}$ (the "slack" defined above), then the next flight f will have to be delayed on the ground at least $g_{f'} + a_{f'} - s_{f'}$ time periods; otherwise, the departure of the next flight f will not be affected. Note that the existence of these coupling constraints allows us to have a separable objective function: the cost of delaying flight f because of an excessive delay of its previous flight f' is taken into account via the term of the objective function corresponding to f (i.e., $c_f^g g_f$), and so need not be included in the term corresponding to f' . Note also that, if the coupling constraints did not exist, the problem would decompose into K subproblems concerning one airport each, so that one could use the already existing techniques to solve for each of the K airports separately.

Note that nonnegativity of g_f is guaranteed by (2.3) (given (2.1)), whereas nonnegativity of a_f is not guaranteed, this is why constraints (2.10) are needed.

For simplicity of exposition, variables g_f and a_f were kept in formulation (I_1) , but it should be clear that they can be eliminated by mere substitution through (2.3) and (2.4), so that u_{ft} and v_{ft} are the only decision variables. We give now the result of this substitution: formulation (I'_1) , where only u_{ft} and v_{ft} appear.

$$\begin{aligned}
(I'_1) \quad \min \quad & \sum_{f=1}^F c_f^a (\sum_{t \in \mathcal{T}_f^a} t v_{ft} - r_f) - (c_f^a - c_f^g) (\sum_{t \in \mathcal{T}_f^d} t u_{ft} - d_f) \\
\text{s.t.} \quad & \sum_{f: k_f^d = k} u_{ft} \leq D_k(t), & (k, t) \in \mathcal{K} \times \mathcal{T}; \\
& \sum_{f: k_f^a = k} v_{ft} \leq R_k(t), & (k, t) \in \mathcal{K} \times \mathcal{T}; \\
& \sum_{t \in \mathcal{T}_f^d} u_{ft} = 1, & f \in \mathcal{F}; \\
& \sum_{t \in \mathcal{T}_f^a} v_{ft} = 1, & f \in \mathcal{F}; \\
& \sum_{t \in \mathcal{T}_{f'}} t v_{f't} - r_{f'} - s_{f'} \leq \sum_{t \in \mathcal{T}_f^d} t u_{ft} - d_f, & f' \in \mathcal{F}'; \\
& \sum_{t \in \mathcal{T}_f^a} t v_{ft} - \sum_{t \in \mathcal{T}_f^d} t u_{ft} \geq r_f - d_f, & f \in \mathcal{F}; \\
& u_{ft}, v_{ft} \in \{0, 1\}.
\end{aligned}$$

Note the compactness of the above formulation. The number of constraints is $3F + F' + 2KT$, and the number of variables is at most $\sum_{f \in \mathcal{F}} (2G_f + A_f + 2)$ (we have at most $G_f + 1$ u_{ft} variables and $G_f + A_f + 1$ v_{ft} variables for a given f). For the typical values $G_f = 4$ and $A_f = 2$, the number of variables becomes at most $12F$, which is a small linear multiple of the number of flights. Although this formulation is quite compact, one can do even better, as the next section will show.

2.3 Infinite departure capacities.

Formulation (I_1) is sufficiently general for the static deterministic case, but it can be simplified considerably without significant loss of practical applicability. Note, first, that it is usually undesirable to delay aircraft in the air. In fact, the fundamental goal of ground holding policies is to avoid this kind of delay. Therefore, we may eliminate airborne delays *as decision variables*; i.e., we may discard the possibility of *en route* speed reduction. We will be left then with airborne delays resulting only from arrival

queueing (denoted earlier by q_f^a), and our only decision variables will be g_f . (Note that, since the problem is deterministic, q_f^a are determined if g_f and service disciplines for the arrival queues are given.)

2.3.1 Infinite departure capacities give zero airborne delays.

As pointed out in Chapter 1, departure capacities are typically higher than landing capacities. This is due to the fact that the minimum separation between landings is greater than the minimum separation between take-offs. Motivated by this fact, we examined what happens if departure capacities are very large, theoretically infinite.

It will be shown now that, if departure capacities are infinite, ground and airborne delay cost functions are linear, and $c_f^a > c_f^g$, then, if formulation (I_1) without airborne delays as decision variables has an optimal solution, then it also has an optimal solution in which no flight incurs an airborne delay.

Consider a feasible solution $\{g_f, f \in \mathcal{F}\}$ and the associated arrival queueing delays $\{q_f^a, f \in \mathcal{F}\}$, and compare its cost with the cost of the new solution $\{g_f + q_f^a, f \in \mathcal{F}\}$, in which all airborne delays are incorporated into ground holds. Given that the cost functions are linear, and given that airborne delays are costlier than ground delays, the new solution will have a lower cost than the previous solution. In fact, $c_f^g(g_f + q_f^a) = c_f^g g_f + c_f^g q_f^a < c_f^g g_f + c_f^a q_f^a$. Moreover, the new solution $\{g_f + q_f^a, f \in \mathcal{F}\}$ is feasible (assuming sufficiently large G_f), since there are no departure capacity constraints.

Now are we entitled to assume that departure capacities are infinite? Computational experience reported in Subsection 3.3.1 shows that the impact of finite departure capacities is negligible (when departure capacities are higher than arrival capacities by realistic amounts). This *a posteriori* argument justifies the assumption of infinite departure capacities. Note that, in the single-airport case, which is the only case considered so far in the literature, no departure capacities are involved, so that

one is in fact rigorously justified to consider only feasible solutions with zero airborne delays (provided the problem is deterministic and the cost functions are linear).

Assuming infinite departure capacities eliminates thus airborne delays and gives a second pure 0–1 integer programming formulation of the static deterministic multi-airport GHP.

2.3.2 A second pure 0–1 IP formulation.

The second formulation is in some sense a special case of (I_1) but requires some manipulations in order to be derived from (I_1) . Given (2.4), by setting $a_f = 0$, one gets the following expression for g_f :

$$g_f = \sum_{t \in \mathcal{T}_f^a} tv_{ft} - r_f, \quad f \in \mathcal{F}. \quad (2.12)$$

By comparing then (2.12) with (2.3), it is seen that²:

$$\sum_{t \in \mathcal{T}_f^a} tv_{ft} - \sum_{t \in \mathcal{T}_f^d} tu_{ft} = r_f - d_f, \quad f \in \mathcal{F}, \quad (2.13)$$

so that (given (2.7) and (2.8)) one of the two sets of variables is now redundant. We choose to discard u_{ft} and to keep v_{ft} , since v_{ft} appears in the arrival capacity constraints (2.6), which must be kept. The departure capacity constraints (2.5) are discarded, as are the assignment constraints (2.7). We are left with the following formulation:

$$(I_2) \quad \min \quad \sum_{f=1}^F c_f^g g_f$$

$$\text{s.t.} \quad \sum_{f: k_f^a = k} v_{ft} \leq R_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}; \quad (2.14)$$

$$\sum_{t \in \mathcal{T}_f^a} v_{ft} = 1, \quad f \in \mathcal{F}; \quad (2.15)$$

$$g_{f'} - s_{f'} \leq g_f, \quad f' \in \mathcal{F}'; \quad (2.16)$$

²By comparing (2.1) and (2.2), one can see that A_f must be equal to 0 in the case of infinite departure capacities without airborne delays as decision variables: if flight f takes off at $d_f + t$, it will land at $r_f + t$.

$$v_{ft} \in \{0, 1\}, \quad f \in \mathcal{F}, t \in \mathcal{T}_f^a. \quad (2.17)$$

Direct substitution of (2.12) into (I_2) gives:

$$\begin{aligned} (I'_2) \quad \min \quad & \sum_{f=1}^F c_f^g (\sum_{t \in \mathcal{T}_f^a} t v_{ft} - r_f) \\ \text{s.t.} \quad & \sum_{f: k_f^a = k} v_{ft} \leq R_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}; \\ & \sum_{t \in \mathcal{T}_f^a} v_{ft} = 1, \quad f \in \mathcal{F}; \\ & \sum_{t \in \mathcal{T}_{f'}^a} t v_{f't} - r_{f'} - s_{f'} \leq \sum_{t \in \mathcal{T}_f^a} t v_{ft} - r_f, \quad f' \in \mathcal{F}'; \\ & v_{ft} \in \{0, 1\}, \quad f \in \mathcal{F}, t \in \mathcal{T}_f^a. \end{aligned}$$

Note the simplicity of formulation (I_2) . The number of constraints is $F + F' + KT$, and the number of variables is at most $\sum_{f \in \mathcal{F}} (G_f + 1)$ which, if all G_f are equal to 4 (corresponding to a maximum ground hold of one hour), becomes $5F$. Therefore, the total number of flights F is the major determinant of the size of the problem. The number of time periods T has almost no influence on the size of the problem, and the same holds for the number of airports K . Of course, the number of airports has an indirect influence on the size of the problem, since it influences the number of flights to be considered. Typically, a major U.S. airport has 1000–1200 operations (landings plus take-offs) each day, corresponding to 500–600 flights per day. But still, the fact that the problem is insensitive as to how the total number of flights is distributed among airports and time periods is quite welcome. This becomes clear in dynamic versions of the ground-holding problem in which the time horizon is limited to a portion of a day, so that fewer flights per airport have to be considered, and it becomes possible to solve the problem for a large number of airports.³

Note, finally, that, if the coupling constraints (2.16) are omitted from the formulation, what is left is essentially the single-airport formulation given, e.g., in [5] (cf. formulation (I_0) of Subsection 1.3.2). It follows that the coupling constraints (2.16) are the gist of the model. It is indeed surprising that the network effects can be taken

³Cf., however, p. 75, footnote 2.

into account in such a simple way without loss of generality. *Simplex sigillum veri* (“the simple is the sign of the true”).

2.4 Flight cancellations.

In situations where delays become excessive, it is common airline practice to cancel some flights, especially at hub airports. Motivated from this fact, we developed formulations which take into account the possibility of cancelling flights. These formulations have the additional advantage that they escape infeasibility problems which might arise with formulations (I_1) and (I_2) . Infeasibility occurs when airport capacities are low: even though the total daily capacity of an airport may be sufficient to accommodate the total number of flights scheduled to depart from or arrive at that airport, the problem may still be infeasible if excessive congestion appears during some portion of the day. This is mainly due to the requirement that there be upper bounds, G_f and A_f , to the delays of flight f . In order to grasp this point with respect, e.g., to formulation (I_2) , take the extreme case where the landing capacity of an airport is reduced to zero for $G_f + 1$ successive time periods. Then, if a flight was scheduled to arrive exactly before the zero capacity interval, it will be impossible to reassign this flight and the problem will become infeasible. Similar remarks hold for formulation (I_1) .

We will give two new formulations, (I_3) and (I_4) , corresponding to (I_2) and to (I_1) , respectively. Keep the old decision variables u_{ft} and v_{ft} , and define the new decision variables $z_f, f \in \mathcal{F}$, to be 1 if flight f is cancelled and 0 otherwise. Denote by M_f the *cancellation cost* of flight f . When a flight in \mathcal{F}' (i.e., a flight that is “continued”) is cancelled, there are two possibilities concerning the next flight initially scheduled to be performed by the same aircraft: either it is performed by a replacement (or a “spare”) aircraft, or it is also cancelled. The first case is more common in practice, especially in hub airports where most cancellations take place, but the formulations are general enough to incorporate a combination of both cases. Partition \mathcal{F}' into \mathcal{F}'_1 , the set of those flights in \mathcal{F}' whose cancellation will not affect their next flight, and

\mathcal{F}'_2 , the set of those flights in \mathcal{F}' whose cancellation will entail the cancellation of their next flight. We give now first the new formulation corresponding to (I_2) .

2.4.1 A third pure 0–1 IP formulation.

$$(I_3) \quad \min \quad \sum_{f=1}^F (c_f^g g_f + (M_f + c_f^g r_f) z_f)$$

$$\text{s.t.} \quad \sum_{f:k_f^g=k} v_{ft} \leq R_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}; \quad (2.18)$$

$$z_f + \sum_{t \in \mathcal{T}_f^g} v_{ft} = 1, \quad f \in \mathcal{F}; \quad (2.19)$$

$$g_{f'} - s_{f'} + (s_{f'} + r_{f'} - r_f) z_{f'} \leq g_f, \quad f' \in \mathcal{F}'_1; \quad (2.20)$$

$$g_{f'} - s_{f'} + (s_{f'} + r_{f'} + G_f + 1) z_{f'} \leq g_f + (r_f + G_f + 1) z_f, \quad f' \in \mathcal{F}'_2; \quad (2.21)$$

$$v_{ft}, z_f \in \{0, 1\}. \quad (2.22)$$

The above formulation incorporates some technical tricks which are necessitated by the fact that, when a flight f is cancelled (i.e., $z_f = 1$), then all v_{ft} corresponding to f are 0 (by (2.19)), so that (2.12) gives $g_f = -r_f$. Keeping this fact in mind, it can be seen immediately that, when $z_f = 1$, the objective function term corresponding to f is M_f . It is also clear that, when $z_{f'} = 1$, (2.20) becomes $-r_f \leq g_f$, which holds even if flight f is cancelled (so that cancellation of f' leaves f unaffected). Finally, if $z_{f'} = 1$, (2.21) becomes $G_f + 1 \leq g_f + (r_f + G_f + 1) z_f$, entailing $z_f = 1$ (since $g_f \leq G_f$ always), which is precisely what is wanted: if f' is cancelled, then f is also cancelled.

The variables g_f were again left in the formulation, but it should be clear that they can be eliminated by mere substitution through (2.12). Now an important point is that the variables z_f can also be eliminated through (2.19), provided that (2.19) is replaced by $\sum_{t \in \mathcal{T}_f^g} v_{ft} \leq 1$. The outcome of effecting all these substitutions is:

$$(I'_3) \quad \min \quad \sum_{f=1}^F [M_f + \sum_{t \in \mathcal{T}_f^g} v_{ft} (c_f^g (t - r_f) - M_f)]$$

$$\text{s.t.} \quad \sum_{f:k_f^g=k} v_{ft} \leq R_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T};$$

$$\sum_{t \in \mathcal{T}_f^g} v_{ft} \leq 1 \quad f \in \mathcal{F};$$

$$\begin{aligned} \sum_{t \in \mathcal{T}_{f'}^a} v_{ft}(t - s_{f'} - r_{f'} + r_f) &\leq \sum_{t \in \mathcal{T}_f^a} tv_{ft}, \quad f' \in \mathcal{F}'_1; \\ \sum_{t \in \mathcal{T}_{f'}^a} v_{ft}(t - s_{f'} - r_{f'} - G_f - 1) &\leq \sum_{t \in \mathcal{T}_f^a} v_{ft}(t - r_f - G_f - 1), \quad f' \in \mathcal{F}'_2; \\ v_{ft}, z_f &\in \{0, 1\}. \end{aligned}$$

The fact that the new formulation (I_3) has exactly the same number of variables and of constraints as the previous corresponding formulation (I_2) is particularly significant, since (I_3) enjoys considerable advantages both in terms of generality (the real-world problem is better approximated) and in terms of flexibility (infeasibility problems are eliminated).

2.4.2 A fourth pure 0–1 IP formulation.

We give now the formulation with flight cancellations corresponding to (I_1):

$$(I_4) \quad \min \quad \sum_{f=1}^F [c_f^g g_f + c_f^a a_f + (M_f + c_f^g d_f + c_f^a (r_f - d_f)) z_f] \quad (2.23)$$

s.t.

$$\sum_{f: k_f^d = k} u_{ft} \leq D_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}; \quad (2.23)$$

$$\sum_{f: k_f^a = k} v_{ft} \leq R_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}; \quad (2.24)$$

$$z_f + \sum_{t \in \mathcal{T}_f^d} u_{ft} = 1, \quad f \in \mathcal{F}; \quad (2.25)$$

$$z_f + \sum_{t \in \mathcal{T}_f^a} v_{ft} = 1, \quad f \in \mathcal{F}; \quad (2.26)$$

$$g_{f'} + a_{f'} - s_{f'} + (s_{f'} + r_{f'} - r_f) z_{f'} \leq g_f, \quad f' \in \mathcal{F}'_1; \quad (2.27)$$

$$g_{f'} + a_{f'} - s_{f'} + (s_{f'} + r_{f'} + G_f + 1) z_{f'} \leq g_f + (r_f + G_f + 1) z_f, \quad f' \in \mathcal{F}'_2; \quad (2.28)$$

$$a_f \geq 0, \quad f \in \mathcal{F}; \quad (2.29)$$

$$u_{ft}, v_{ft}, z_f \in \{0, 1\}. \quad (2.30)$$

The new technical tricks in the objective function are necessitated by the fact that, when flight f is cancelled (i.e., $z_f = 1$), then all u_{ft} and v_{ft} corresponding to f are 0 (by (2.25) and (2.26)), so (2.3) and (2.4) give $g_f = -d_f$ and $a_f = d_f - r_f$. Note that in that case we have $g_f + a_f = -r_f$, this is why the coupling constraints of (I_4) are essentially the same as those of (I_3).

We conclude this section by giving the result of eliminating g_f , a_f , and z_f from (I_4) :

$$(I'_4) \quad \min \quad \sum_{f=1}^F [M_f + c_f^a \sum_{t \in \mathcal{T}_f^a} v_{ft}(t - r_f) - \sum_{t \in \mathcal{T}_f^d} u_{ft}((t - d_f)(c_f^a - c_f^g) + M_f)]$$

$$\text{s.t.} \quad \sum_{f: k_f^d = k} u_{ft} \leq D_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T};$$

$$\sum_{f: k_f^a = k} v_{ft} \leq R_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T};$$

$$\sum_{t \in \mathcal{T}_f^d} u_{ft} \leq 1, \quad f \in \mathcal{F}; \tag{2.31}$$

$$\sum_{t \in \mathcal{T}_f^d} u_{ft} = \sum_{t \in \mathcal{T}_f^a} v_{ft}, \quad f \in \mathcal{F}; \tag{2.32}$$

$$\sum_{t \in \mathcal{T}_{f'}^a} v_{ft}(t - s_{f'} - r_{f'} + r_f) \leq \sum_{t \in \mathcal{T}_f^d} t u_{ft} + r_f - d_f, \quad f' \in \mathcal{F}'_1;$$

$$\sum_{t \in \mathcal{T}_{f'}^a} v_{ft}(t - s_{f'} - r_{f'} - G_f - 1) \leq \sum_{t \in \mathcal{T}_f^d} u_{ft}(t - r_f - G_f - 1) + r_f - d_f, \quad f' \in \mathcal{F}'_2;$$

$$\sum_{t \in \mathcal{T}_f^a} t v_{ft} - \sum_{t \in \mathcal{T}_f^d} t u_{ft} \geq r_f - d_f, \quad f \in \mathcal{F};$$

$$u_{ft}, v_{ft} \in \{0, 1\}.$$

Note that, given (2.32), (2.31) could be replaced by $\sum_{t \in \mathcal{T}_f^a} v_{ft} \leq 1$.

2.5 Extensions.

A particularly attractive feature of the formulations developed in this chapter is their flexibility: they can be extended in several ways in order to accommodate various degrees of modeling detail. Four examples of such extensions will be examined in this section.

2.5.1 Hub airports: more than one “next” flights.

In hub airports, an arriving flight typically has passengers connecting to several departing flights. This can be taken into account in any of the formulations (I_1) – (I_4) by means of an easy extension. It suffices to *reinterpret* the coupling constraints as linking not only a pair of flights scheduled to be performed by the same aircraft, but also any pair of flights f' and f such that f is not allowed to leave before f' lands, because passengers in f' connect to f . With this reinterpretation of the coupling

constraints, a continued flight may have more than one next flights. Therefore, the formulations remain unchanged and a number of new coupling constraints is simply added to them.

Note that the slack in a coupling constraint of the new kind (connecting flights) will typically be different from the slack in a coupling constraint of the old kind (continued flights). This is because the turnaround time involved in connecting flights is restricted to moving passengers and their luggage, while the turnaround time involved in continued flights also includes cleaning and refuelling the aircraft.

2.5.2 En route speeding.

Sometimes there is the possibility to speed up an aircraft *en route*, so that the aircraft may arrive even before its scheduled arrival time. This possibility can be easily taken into account in any of the formulations with airborne delays presented in this chapter. It suffices to take r_f to be not the scheduled arrival time, but the earliest possible arrival time. If, for instance, an aircraft is scheduled to arrive at time period 28, but may be speeded up so as to arrive up to two periods earlier, r_f for the corresponding f will be equal to 26. An airborne “delay” a_f equal, e.g., to 1 will correspond to a speeding up of one time period, whereas an a_f equal to 3 will correspond to a slowing down of one time period. The actual arrival time of flight f will of course also depend on its ground delay: if f departs with a ground delay g_f equal to 3 and is speeded up by one time period, it will arrive with a total delay of two time periods, i.e., at period 30 (which is $r_f + g_f + a_f = 26 + 3 + 1$). Note that the upper bound A_f will have to be increased by 2 in this example.

2.5.3 Interdependence between departure and arrival capacities.

Usually the departure and arrival capacities of a given airport at a given time period are not independent, because they are determined by the way in which runway use is assigned to departing or arriving aircraft. If all runways are exclusively used for

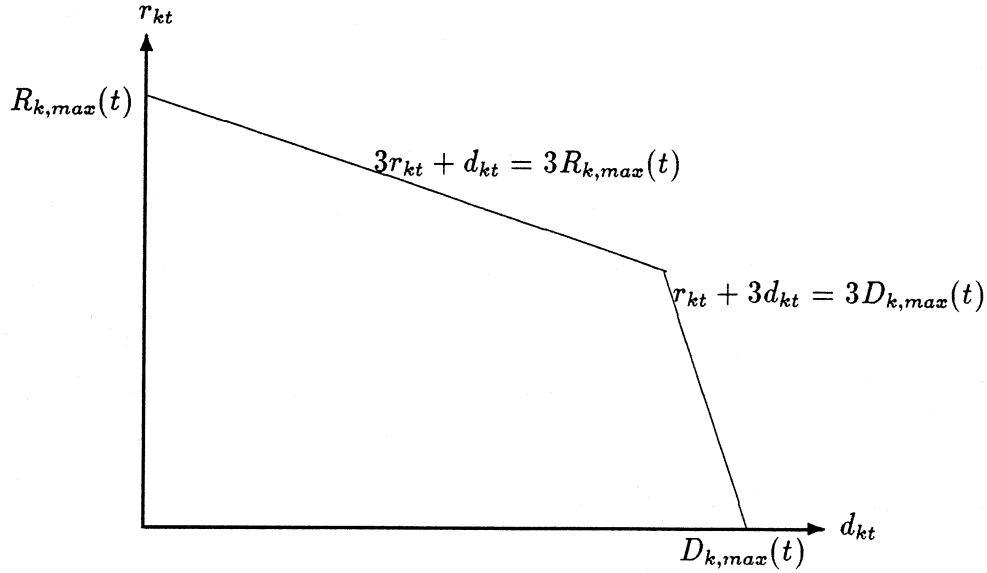


Figure 2-3: An example of a region of possible combinations between the departure capacity d_{kt} and the arrival capacity r_{kt} of airport k at time period t .

landings, arrival capacity reaches a maximum value $R_{k,max}(t)$ determined by the minimum separation between successive landings, while departure capacity is 0. If all runways are exclusively used for take-offs, departure capacity reaches a maximum value $D_{k,max}(t)$ determined by the minimum separation between successive take-offs (which is less than the minimum separation between landings, so that $D_{k,max}(t) > R_{k,max}(t)$), while arrival capacity is 0. Intermediate cases give departure and arrival capacities belonging to a region with the general shape of a two-dimensional convex polytope (Figure 2-3). Note that this region differs among airports and, for a given airport, it can change with time (because weather can change).

The above situation can be easily taken into account in any of the formulations with finite departure capacities presented in this chapter, i.e., (I_1) or (I_4) . This is achieved by introducing the new integer (not 0–1) decision variables d_{kt} and r_{kt} standing, respectively, for the departure and the arrival capacities of airport k at time t . These new decision variables will replace the constants $D_k(t)$ and $R_k(t)$ in the right-hand sides of the capacity constraints. In addition, new sets of constraints will be needed, one set for each time period, ensuring that d_{kt} and r_{kt} fall within

the region of their possible combinations. These constraints will be of the following general form:

$$\alpha_{kt}^i d_{kt} + \beta_{kt}^i r_{kt} \leq \gamma_{kt}^i, \quad (k, t) \in \mathcal{K} \times \mathcal{T}, i \in \{1, \dots, I_{kt}\},$$

where $\alpha_{kt}^i, \beta_{kt}^i, \gamma_{kt}^i$ are constants and I_{kt} is the number of linear constraints describing the departure-arrival capacity region of airport k at time period t . For instance, for the region shown in Figure 2-3, the following two constraints are needed for period t :

$$3r_{kt} + d_{kt} \leq 3R_{k,max}(t); \quad (2.33)$$

$$r_{kt} + 3d_{kt} \leq 3D_{k,max}(t). \quad (2.34)$$

As an example, we give the modification of formulation (I_1) :

$$(I_5) \quad \min \quad \sum_{f=1}^F c_f^g g_f + c_f^a a_f$$

$$\text{s.t.} \quad \sum_{f:k_f^d=k} u_{ft} \leq d_{kt}, \quad (k, t) \in \mathcal{K} \times \mathcal{T}; \quad (2.35)$$

$$\sum_{f:k_f^a=k} v_{ft} \leq r_{kt}, \quad (k, t) \in \mathcal{K} \times \mathcal{T}; \quad (2.36)$$

$$\alpha_{kt}^i d_{kt} + \beta_{kt}^i r_{kt} \leq \gamma_{kt}^i, \quad (k, t) \in \mathcal{K} \times \mathcal{T}, i \in \{1, \dots, I_{kt}\}; \quad (2.37)$$

$$\sum_{t \in \mathcal{T}_f^d} u_{ft} = 1, \quad f \in \mathcal{F};$$

$$\sum_{t \in \mathcal{T}_f^a} v_{ft} = 1, \quad f \in \mathcal{F};$$

$$g_{f'} + a_{f'} - s_{f'} \leq g_f, \quad f' \in \mathcal{F}';$$

$$a_f \geq 0, \quad f \in \mathcal{F};$$

$$u_{ft}, v_{ft} \in \{0, 1\};$$

$$d_{kt}, r_{kt} \text{ integers} \geq 0. \quad (2.38)$$

Note that formulation (I_5) has $2KT$ variables more than (I_1) (a negligible number compared with the remaining about $12F$ variables), and $\sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} I_{kt}$ more constraints. Thus the additional number of constraints will be a small multiple of KT , which is again negligible compared with the remaining $3F + F' + 2KT$ constraints.

Variables d_{kt} and r_{kt} were introduced in order to make formulation (I_5) clearly understandable, but these variables can be eliminated. In fact, constraints (2.35) and (2.36) can be replaced by equalities since constraints (2.37) are inequalities. Then d_{kt} and r_{kt} will represent the *used* capacities of airport k at time period t . We are thus left with the following formulation (I'_5) , which has exactly the same number of variables as (I_1) , but has $2KT$ variables and $2KT$ constraints *fewer* than (I_5) !

$$\begin{aligned}
(I'_5) \quad \min \quad & \sum_{f=1}^F c_f^g g_f + c_f^a a_f \\
\text{s.t.} \quad & \alpha_{kt}^i (\sum_{f:k_j^d=k} u_{ft}) + \beta_{kt}^i (\sum_{f:k_j^a=k} v_{ft}) \leq \gamma_{kt}^i, \quad (k,t) \in \mathcal{K} \times \mathcal{T}, i \in \{1, \dots, I_{kt}\}; \\
& \sum_{t \in \mathcal{T}_f^d} u_{ft} = 1, \quad f \in \mathcal{F}; \\
& \sum_{t \in \mathcal{T}_f^a} v_{ft} = 1, \quad f \in \mathcal{F}; \\
& g_{f'} + a_{f'} - s_{f'} \leq g_f, \quad f' \in \mathcal{F}'; \\
& a_f \geq 0, \quad f \in \mathcal{F}; \\
& u_{ft}, v_{ft} \in \{0, 1\}.
\end{aligned}$$

2.5.4 Similar percentages of continued and noncontinued flights that are delayed.

As will be shown in Subsection 3.1.3, the optimal solution of the formulations presented in this chapter may result in a large discrepancy between the percentages of continued and noncontinued flights that are delayed, continued flights being typically delayed much less often than noncontinued flights. This might be considered a form of discrimination against classes of users, because some airlines may have a larger percentage of continued flights in their schedules than other airlines. The situation can be easily remedied by adding the following single constraint to any of the formulations (I_1) – (I_4) :

$$\frac{\sum_{f \in \mathcal{F}'} \sum_{t \in \mathcal{T}_f^a \setminus \{r_f\}} v_{ft}}{F'} \geq \frac{\sum_{f \in \mathcal{F} \setminus \mathcal{F}'} \sum_{t \in \mathcal{T}_f^a \setminus \{r_f\}} v_{ft}}{F - F'} - \alpha, \quad (2.39)$$

where α is a number between 0 and 1. If, e.g., one wants the percentage of continued flights which are delayed to be within 10% of the percentage of noncontinued flights

which are delayed, then $\alpha = 0.1$. Alternatively, or in addition, one could want the mean delay per flight (rather than the percentage of delayed flights) to be about the same for continued and noncontinued flights. This could be achieved by adopting the following single constraint:

$$\frac{\sum_{f \in \mathcal{F}'} g_f}{F'} \geq \frac{\sum_{f \in \mathcal{F} \setminus \mathcal{F}'} g_f}{F - F'} - \beta, \quad (2.40)$$

where β is not restricted to be between 0 and 1.

Further concerns about discrimination among various classes of users could be handled in similar ways.

Chapter 3

The static GHP: results.

This chapter presents a systematic investigation of the static deterministic multi-airport GHP by means of extensive computational results based on the formulations of Chapter 2. A major theme of the chapter is the examination of the *network effects* differentiating the multi-airport from the single-airport GHP. A fundamental conclusion of the chapter is that the *magnitude* of network effects must be distinguished from their *importance*.

The chapter is divided into five sections. Section 3.1 examines the *magnitude* of network effects, measured as cost savings of optimal multi-airport ground-holding policies compared to optimal single-airport ground-holding policies. It is found that cost savings are small when all flights have identical cost functions but can be large when cost functions differ. Section 3.2 shows that network effects are *important* even when all cost functions are identical, because optimal single-airport ground-holding policies are highly infeasible for the multi-airport problem, and because the optimal value of the multi-airport problem is much lower than the cost of a heuristic approximating, to some extent, current ground-holding practice. Section 3.3 shows that the impact of finite but fixed departure capacities is negligible, while finite departure capacities interdependent with arrival capacities can result in significant cost savings. Section 3.4 examines the model with flight cancellations. It proposes a heuristic finding a feasible solution of the integer program on the basis of the optimal solution

of the linear programming relaxation and it shows that this feasible solution is good when cancellation costs are reasonable. Finally, Section 3.5 summarizes the results of the chapter.

3.1 Magnitude of network effects.

3.1.1 The decomposed problem and its role.

Given any of the formulations (I_i) of Chapter 2, the *decomposed* problem (D_i) corresponding to that formulation is defined as the linear programming (LP) relaxation of that formulation without the coupling constraints. For instance, the decomposed problem (D_2) corresponding to formulation (I_2) is the LP relaxation of (I_2) without constraints (2.16), namely:

$$(D_2) \min \quad \sum_{f=1}^F c_f^g g_f$$

$$\text{s.t.} \quad \sum_{f: k_f^a = k} v_{ft} \leq R_k(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}; \quad (3.1)$$

$$\sum_{t \in \mathcal{T}_f^a} v_{ft} = 1, \quad f \in \mathcal{F}; \quad (3.2)$$

$$0 \leq v_{ft} \leq 1, \quad f \in \mathcal{F}, t \in \mathcal{T}_f^a. \quad (3.3)$$

Note that (I_i) without the coupling constraints has a totally unimodular constraint matrix [5]. This is why (D_i) was defined as a linear program: its optimal solution will always be integer.

It is easy to see that solving the decomposed problem is equivalent to solving the single-airport GHP (cf. formulation (I_0) of Subsection 1.3.2) for each airport separately and then adding the optimal objective function values corresponding to the various airports. Therefore, by comparing the optimal values of (I_i) and (D_i) , one gets a measure of the magnitude of network effects. A large gap between these optimal values (denoted by v_{I_i} and v_{D_i} , respectively) presumably justifies one in pursuing the application of algorithms pertaining to the multi-airport (coupled) GHP rather than solving for each airport separately by means of the existing methods for the single-

airport GHP.

Although the comparison between v_{I_i} and v_{D_i} is of interest, it must be noted that (I_i) and (D_i) are not directly comparable, because, as will be shown in Subsection 3.2.1, the optimal solution of (D_i) can be highly infeasible for (I_i) .

We will also include in the comparisons the optimal value, v_{L_i} , of the LP relaxation, (L_i) , of (I_i) . Obviously we always have:

$$v_{D_i} \leq v_{L_i} \leq v_{I_i}, \quad i \in \{1, \dots, 5\}. \quad (3.4)$$

It should be noted that the objective of this chapter is to investigate the behaviour of the problem under various combinations of the input parameters, *not* to demonstrate the efficiency of any particular algorithm. In fact, the various instances of the problem were solved by using the well-known commercial package *MPSX*, rather than any custom-tailored algorithm. Some computation times (in CPU seconds) are given simply in order to indicate whether the problem can be solved in reasonable time, rather than in order to provide any “good” bounds on computation times.

3.1.2 Network effects insignificant when cost functions are identical.

This subsection starts with examination of a representative test case and then moves on to a systematic series of test cases showing that network effects are small in the case of identical cost functions. The systematic series of test cases is then used to examine the behaviour of computation times, as well as the impact of input parameters (such as capacities and upper bounds on delays) on the optimal objective function value.

Test case W.

We start with a test case referring to formulation (I_2) . Since this test case will be used a number of times in the sequel, let us give it a name: call it W. Test case W has $K = 3$ airports, $T = 100$ time periods, $F = 1800$ flights (600 flights per airport),

Capacities	v_{D_2}	v_{L_2}	v_{I_2}
(10,10,10)	43,550	43,550	43,550
(9,10,10)	51,900	52,800	52,900
(10,10,9)	48,500	49,000	50,600
(9,10,9)	56,850	57,450	57,950
(10,10,8)	55,650	56,700	58,000

Table 3.1: Behaviour of test case W around the capacity border between feasibility and infeasibility.

and $F' = 600$ flights. With the exception of capacities, all parameters will be kept fixed in this test case: the cost function slopes are 50, the slacks are 0, and the upper bounds on the delays are 4 time periods. The scheduled arrival times were arbitrarily chosen.

As mentioned in Section 2.4, if arrival capacities are very low, the problem becomes infeasible. On the other hand, if capacities are very high, there is little need to delay aircraft. It follows that the capacity region of interest is around the *infeasibility border*: capacities high enough for the problem to be feasible but low enough so that, if they are marginally decreased, the problem becomes infeasible.

Let us consider for the moment only cases in which the arrival capacity of any given airport is constant over the whole time horizon: $R_k(t) = R_k$. Then it was found (by trial and error) that, for the particular test case W under consideration, for $(R_1, R_2, R_3) = (10, 10, 10)$ the problem is feasible, while for $(9, 9, 9)$ the problem is infeasible. Furthermore, for $(9, 10, 10)$, $(10, 10, 9)$, $(9, 10, 9)$, and $(10, 10, 8)$ the problem is feasible, while for $(10, 9, 10)$, $(8, 10, 10)$, and $(10, 10, 7)$ the problem is infeasible. These results give a fairly good picture of the boundary between capacity regions corresponding to feasibility and to infeasibility for test case W.

Table 3.1 gives the optimal objective function values of (D_2) , (L_2) , and (I_2) for the various capacity cases. It is seen that these values always turn out to be very close.

F	F'/F	Capacities	v_{D_2}	v_{L_2}	v_{I_2}	Nonint.
1000	0.20	(12,14)	71,000	71,000	71,000	63
1000	0.40	(10,10)	56,000	56,000	56,000	84
1000	0.60	(11,11)	84,200	84,300	84,700	168
1000	0.80	(10,10)	65,000	65,000	65,500	128
2000	0.20	all 14	96,300	96,300	99,000	117
2000	0.40	all 14	88,400	89,933	93,200	195
2000	0.60	all 12	71,600	71,600	71,800	252
2000	0.80	all 17	53,250	57,387	65,500	355
3000	0.20	all 12	128,000	129,200	129,400	110
3000	0.40	all 18	55,800	55,800	57,300	119
3000	0.60	all 17	90,200	96,550	99,650	232
3000	0.80	all 18	80,500	84,250	87,050	414

Table 3.2: Results for the test series at the infeasibility border.

This is a surprising result which needs explanation, but it must first be ascertained that it is a general phenomenon rather than a peculiar feature of the particular test case under consideration.

Systematic series of test cases.

We examined a series of cases with 2, 4, and 6 airports and 500 flights per airport (so the three cases had 1000, 2000, and 3000 flights, respectively). For each of the three cases, four values of F' were examined, corresponding to a ratio F'/F equal to 0.20, 0.40, 0.60, and 0.80. In all the 12 resulting cases, T was kept fixed and equal to 64 (corresponding to a 16-hour time horizon with 15-minute periods). The cost function slopes were always 50, all slacks were 1, and all upper bounds on delays were 4.

Table 3.2 summarizes the results. The capacities appearing in the table for any particular case are at the infeasibility borders and were found by trial and error. It is seen that the gap between v_{D_2} and v_{I_2} is always small. The last column of Table 3.2 gives the number of flights for which the optimal solution of (L_2) had noninteger values. It can be seen that this number is usually small, around 10% of F . This observation provided the motivation for the development of a heuristic which would

F	F'/F	Capacities	t_{D_2}	t_{L_2}	t_{I_2}
1000	0.20	(12,14)	218	258	371
1000	0.40	(10,10)	235	327	894
1000	0.60	(11,11)	242	377	6958
1000	0.80	(10,10)	235	453	9512
2000	0.20	all 14	664	731	5126
2000	0.40	all 14	652	973	9522
2000	0.60	all 12	644	1148	13607
2000	0.80	all 17	617	1603	18093
3000	0.20	all 12	1188	1453	11360
3000	0.40	all 18	1208	1808	13291
3000	0.60	all 17	1166	2547	17980
3000	0.80	all 18	1180	3072	25021

Table 3.3: Computation times (in CPU seconds) for the results of the test series.

take as input the optimal solution of (L_i) and would round it off in order to give as output a feasible solution of (I_i) . Such a heuristic was developed for the model with flight cancellations, and will be presented in Subsection 3.4.2.

Computation times.

Table 3.3 gives the computation times, in CPU seconds, for the 12 cases of Table 3.2. The following remarks can be made. First, the computation times t_{D_2} and t_{L_2} are quite reasonable, but t_{I_2} can become excessive. Second, as one would expect, the computation times increase as F increases, because the number of constraints and the number of variables increase. Third, for any given F , t_{D_2} does not vary significantly with F' , while t_{L_2} and t_{I_2} increase as F' increases. This is due to the fact that an increase in F' increases the number of constraints of (L_2) and (I_2) (which have $KT + F + F'$ constraints), while it leaves unaffected the number of constraints of (D_2) ($KT + F$).¹

¹For any given F , the four problems corresponding to the four values of F'/F don't simply differ in the connections between flights, but have different arrival schedules; therefore, the corresponding four decomposed problems are not identical.

Optimal value as a function of capacity.

Tables 3.4, 3.5, and 3.6 give results for some of the cases of Table 3.2 but for capacities higher than and around the infeasibility borders. Computation times are also given. (For purposes of comparison, the relevant numbers from Table 3.2 are reproduced in Tables 3.4, 3.5, and 3.6.) As expected, v_{L_2} is an increasing function of capacity, and its increments increase for given capacity increments as the infeasibility border is approached. In other words, if v_{L_2} were differentiable, its second derivative would be positive.

The fact that, in several cases, v_{L_2} is 0 (i.e., no delays are incurred) when all airport capacities are equal to 20 is due to the particular way in which arrival schedules were generated for those cases: aircraft were scheduled to arrive in batches of 20 or less at various time periods. Note also that the four problems corresponding to each of the three values of F in the systematic series are unrelated¹; this explains why, in Table 3.2, the optimal values are not increasing functions of F'/F .

Impact of upper bounds on delays.

The reason why there is an infeasibility border is that upper bounds on delays were imposed (the G_f of Chapter 2). These bounds were kept fixed and equal to 4 time periods (corresponding to 1 hour) in all the above cases, but one may wonder what happens if greater delays are allowed. Table 3.7 gives results for $F = 3000$ and G_f going up to 8 time periods (2 hours). As expected, increasing G_f decreases the infeasibility border, and this results in an increase of the optimal value. However, once the problem is feasible for given capacities, increasing G_f further has little impact on the optimal value. This is important, because it shows that, once the problem is feasible for desirable capacities, one need not worry about increasing the upper bounds on delays in order to decrease costs. This is also welcome from a computational point of view. It will be remembered, from Subsection 2.3.2, that the number of variables in (I_2) is at most $\sum_{f \in \mathcal{F}} (G_f + 1)$. This number almost doubles when G_f increases from 4 to 8, and this is reflected in the computation times shown in Table 3.7.

F	F'/F	Capacities	v_{L_2}	t_{L_2}	Nonint.
1000	0.20	all 30	17,000	265	0
1000	0.20	all 20	34,000	268	18
1000	0.20	all 15	55,750	267	15
1000	0.20	all 14	61,600	263	47
1000	0.20	(13,14)	65,850	264	70
1000	0.20	(12,14)	71,000	258	63
1000	0.40	all 30	4,000	295	0
1000	0.40	all 15	22,500	312	10
1000	0.40	all 11	44,500	326	65
1000	0.40	all 10	56,000	327	84
1000	0.60	all 15	15,500	409	21
1000	0.60	all 13	23,300	402	37
1000	0.60	all 11	84,300	377	128
1000	0.80	all 15	16,500	432	14
1000	0.80	all 11	46,500	468	126
1000	0.80	all 10	65,000	453	168

Table 3.4: Results above and around the infeasibility border for $F = 1000$.

F	F'/F	Capacities	v_{L_2}	t_{L_2}
2000	0.20	all 30	0	680
2000	0.20	all 15	63,500	740
2000	0.20	all 14	96,300	731
2000	0.40	all 20	0	846
2000	0.40	all 18	17,600	837
2000	0.40	all 15	58,750	965
2000	0.40	all 14	89,933	973
2000	0.60	all 20	0	1028
2000	0.60	all 15	35,750	1056
2000	0.60	all 14	45,900	1067
2000	0.60	all 13	57,250	1064
2000	0.60	all 12	71,600	1148
2000	0.80	all 20	0	1411
2000	0.80	all 18	32,700	1433
2000	0.80	all 17	57,387	1603
2000	0.80	(16,17,17,17)	58,938	1537
2000	0.80	(15,17,17,17)	60,488	1666
2000	0.80	(17,17,17,16)	61,713	1708
2000	0.80	(16,17,17,16)	63,263	1551
2000	0.80	(17,17,17,15)	67,994	1657

Table 3.5: Results above and around the infeasibility border for $F = 2000$.

F	F'/F	Capacities	v_{L_2}	t_{L_2}
3000	0.20	all 20	0	1360
3000	0.20	all 19	10,450	1376
3000	0.20	all 18	20,900	1401
3000	0.20	all 15	55,000	1430
3000	0.20	all 12	129,200	1453
3000	0.40	all 20	0	1751
3000	0.40	all 18	55,800	1808
3000	0.40	(18,17,18,18,18,18)	57,300	1785
3000	0.40	(18,18,18,18,17,18)	57,900	1790
3000	0.40	(18,18,18,18,16,18)	60,000	1792
3000	0.60	all 20	3,000	1969
3000	0.60	all 18	57,300	2204
3000	0.60	all 17	96,550	2547
3000	0.80	all 30	5,000	2495
3000	0.80	all 20	27,000	2927
3000	0.80	all 18	84,250	3072

Table 3.6: Results above and around the infeasibility border for $F = 3000$.

F	F'/F	Capacities	G_f	v_{L_2}	t_{L_2}
3000	0.20	all 12	4	128,000	1453
3000	0.20	all 11	5	Infeasible	—
3000	0.20	all 11	6	187,650	2045
3000	0.20	all 11	7	186,900	2232
3000	0.20	all 11	8	186,900	2484
3000	0.60	all 17	4	96,550	2547
3000	0.60	all 16	4	Infeasible	—
3000	0.60	all 16	5	147,253	3627
3000	0.60	all 16	6	140,944	5298
3000	0.60	all 15	6	Infeasible	—
3000	0.60	all 15	7	210,440	6855
3000	0.60	all 14	8	Infeasible	—

Table 3.7: Results for various upper bounds on delays.

3.1.3 Network effects significant when cost functions differ.

Now we must explain the surprising fact that network effects were found to be insignificant. Our conclusion will be that this is because all cost functions were identical. Before we argue for this conclusion, let us examine two other possible explanations that might be adduced. A first explanation might be that the capacities at the border between feasibility and infeasibility, although they cannot be lowered in the context of the present model, are still too high for network effects to have a severe impact. This explanation, if true, would undermine the utility of formulation (I_2) (though not of (I_1)) as a representation of the real-world situation. This explanation, however, is not true. First, v_{D_2} and v_{L_2} are very close even with quite low capacities (see the second and the fourth rows of Table 3.2). Second, in Subsection 3.4.1, where formulation (I_3), which is immune to infeasibility, is examined, it will be seen (cf. fifth row of Table 3.12) that v_{L_3} and v_{D_3} are very close even with capacities as low as 256 aircraft per airport per day (4 per period) (with 500 aircraft scheduled to land, so that the remaining flights are cancelled).

A second possible explanation is that arrival capacities were taken to be uniform (i.e., constant over the whole time horizon). As pointed out in Subsection 1.2.3, ground-holding policies make sense when one delays aircraft on the ground now because one expects less congestion later on at the destination airports of the delayed aircraft. But when airport capacities are uniform throughout the day, how can one expect less congestion later on? The answer is that less congestion can be expected when fewer aircraft are scheduled to arrive later on, even if arrival capacities are uniform. Nevertheless, this second possible explanation has some validity, as shown by computational results reported in Subsection 3.4.1 (cf. Table 3.13), where it is seen that nonuniform capacities can give significant network effects.

The main explanation, however, is the identity of cost functions, as will now be argued.

Capacities	% of $f \in \mathcal{F}'$ delayed in (I_2)	% of $f \in \mathcal{F} \setminus \mathcal{F}'$ delayed in (I_2)
(10,10,10)	12%	30%
(9,10,10)	18%	36%
(10,10, 9)	17%	34%
(9,10, 9)	20%	40%
(10,10, 8)	19%	37%

Table 3.8: Percentages of continued and noncontinued flights that are delayed in the optimal solution of (I_2) for test case W.

Few continued flights delayed with identical cost functions.

Small network effects mean that few continued flights are delayed. This is expected to be the case if continued flights have the same cost functions as noncontinued flights. In fact, if there is a choice (in (I_i)) between delaying a continued flight f' and a noncontinued flight f , it will usually be preferable to delay the latter, since delaying the former would probably result in a greater total cost (because the next flight of f' might also have to be delayed). It follows that, in the optimal solution of (I_i) , few flights in \mathcal{F}' will be delayed if all cost functions are identical. This effect would be particularly noticeable for small slacks.

A look at Table 3.8 corroborates the above hypothesis. Table 3.8 refers to test case W (cf. Table 3.1), and gives the percentages of continued and of noncontinued flights that are delayed in the optimal solution of (I_2) . It can be seen that the percentage of continued flights that are delayed is half (or even less) the percentage of noncontinued flights that are delayed.²

Significant cost differences with different cost functions.

A second way to confirm the above hypothesis is by varying the cost function slopes so as to disadvantage continued flights. If continued flights have much lower marginal

²Recall, from Subsection 2.5.3, that this asymmetry between continued and noncontinued flights might be considered a form of discrimination but can be remedied by adopting constraint (2.39) or (2.40).

delay costs than noncontinued flights, then it may often be preferable to delay a continued rather than a noncontinued flight when a choice is available, with the consequence that network effects may be significant. Test case W was run with capacities equal to 10 and with cost function slopes equal to 10 for the continued flights and equal to 100 for the noncontinued flights, and the results were: $v_D = 13,950$ and $v_L = 22,811$, a very significant gap. Other results with different cost functions reported in Subsection 3.4.1 (cf. Table 3.14) also show significant network effects.

3.2 Importance of network effects.

The previous section showed that network effects can be large in the general case, but are typically small in the special case of identical cost functions. This would not be worrisome if identical cost functions were of little practical interest, as it seems that they should be. In fact, as pointed out in Subsection 1.2.3, delay costs vary according to aircraft size. Political considerations, however, make the case of identical cost functions of significant practical interest, because nonidentical cost functions would constitute a form of discrimination among classes of users. This is why almost all computations were performed with identical cost functions. But now one may begin to worry about the insignificance of network effects in this case. If, in a case of considerable practical interest, solving the multi-airport GHP is about the same as solving one single-airport GHP for each airport of the network, then why pursue research on the multi-airport problem?

We will now give two decisive arguments showing that, even in the case of identical cost functions, although network effects as defined in the previous section are of *small magnitude*, they are of *great importance*. The two arguments will be: first, that solving the single-airport problems gives solutions which are highly infeasible for the multi-airport problem; second, that the optimal value of the multi-airport problem can be much lower than the cost incurred by nonoptimal FCFS ground-holding practices in certain instances.

Capacities	# of violated coupling constraints
(10,10,10)	179
(9,10,10)	204
(10,10, 9)	183
(9,10, 9)	238
(10,10, 8)	235

Table 3.9: Number of coupling constraints that the optimal solution of (D_2) violates (test case W).

3.2.1 Decomposed solution highly infeasible for network.

The purpose of the GHP is to assign ground holds to all flights in a network of airports so as to minimize total cost. It goes without saying that these ground holds must be *feasible*; in particular, they must respect the coupling constraints. If an aircraft is scheduled to perform flights f_1 and f_2 in succession, with a slack of 1 time period, we cannot assign a ground hold of 3 time periods to f_1 and of 0 time periods to f_2 : the aircraft will simply not be ready to depart on time for f_2 . Therefore, if solving the decomposed problem results in solutions which violate a large number of coupling constraints, then the decomposed problem is useless for practical purposes as far as the multi-airport GHP is concerned.

The above is indeed the case. Table 3.9 refers to test case W and shows the number of coupling constraints violated by the optimal solution of (I_2) . Since there are 600 coupling constraints, about one third of them is typically violated, so the solution of (D_2) is highly infeasible for (I_2) .

3.2.2 FCFS heuristic highly inefficient.

The fact that the optimal solution of (D_i) is infeasible for (I_i) suggests that v_{I_i} should not be compared to v_{D_i} , but rather to the cost arising from some other way of assigning feasible ground holds to the flights in a network of airports. Such a way is suggested by the current ground-holding practice described in Subsection 1.3.1. Recall that this

consists in calculating the excesses of scheduled arrivals over forecasted capacities and in assigning ground holds on a FCFS basis. This FCFS heuristic gives a feasible solution (provided the upper bounds on delays are sufficiently large) and is formally described by the following algorithm.

BEGIN

Initialize: $g_f = 0$.

FOR $t = 1$ TO T DO:

FOR $k = 1$ TO K DO:

Define $\mathcal{S}_k(t) := \{f : (k_f^a = k)(r_f + g_f = t)\}$.

Define $S_k(t) := |\mathcal{S}_k(t)|$.

IF $S_k(t) > R_k(t)$ THEN:

Choose $S_k(t) - R_k(t)$ flights in $\mathcal{S}_k(t)$.

FOR $f = 1$ TO $S_k(t) - R_k(t)$ DO:

Set $g_f = g_f + 1$.

IF f has a next flight \hat{f} THEN:

IF $g_f > s_f$ THEN:

Set $g_{\hat{f}} = g_{\hat{f}} + 1$, and similarly if \hat{f} has a next flight and so on.

END IF

END IF

CONTINUE f

END IF

CONTINUE k

CONTINUE t

END

Table 3.10 compares the cost of the FCFS heuristic with the optimal value of the multi-airport GHP. Some results for the model with flight cancellations are also

F	F'/F	Capac.	v_{FCFS}	v_{L_2}	v_{D_3}	FCFS overestimation
1000	0.20	(14,14)	77,200	61,600		25.3%
1000	0.40	(13,14)	80,700	65,850		22.6%
1000	0.60	(12,14)	84,200	71,000		18.6%
1000	0.40	all 11	61,050	44,500		37.2%
1000	0.40	all 10	92,000	56,000		64.3%
1000	0.60	all 13	29,000	23,300		24.5%
1000	0.60	all 11	101,350	84,300		20.2%
1000	0.60	all 10	191,500		117,000	63.7%
1000	0.60	all 08	431,600		240,700	79.3%
1000	0.60	all 06	722,500		402,600	79.5%
1000	0.60	all 04	1,035,800		582,300	77.9%
1000	0.80	all 11	86,050	46,500		85.1%
1000	0.80	all 10	122,500	65,000		88.5%

Table 3.10: Comparison of FCFS heuristic values and exact optima.

included. All results are for identical cost functions. It is seen that the FCFS heuristic always results in a cost much higher than what one can achieve by solving the multi-airport GHP optimally. Equivalently, assigning optimal ground holds can reduce the cost of the myopic procedure by a significant percentage, as much as about 50%. In this sense network effects are quite important, and one is justified in pursuing research on the multi-airport GHP.³

3.3 Finite departure capacities.

In this section it is shown that finite but fixed departure capacities have negligible impact, whereas finite departure capacities which are interdependent with arrival capacities can result in very significant cost savings. In other words, the optimal value of formulation (I_1) is very close to the value of formulation (I_2), but significantly

³Note that the FCFS heuristic does not correspond exactly to current ground-holding practice because, as pointed out in Subsection 1.3.1, the FAA *dynamically* updates ground holds during the course of the day. The dynamic FCFS heuristic will be examined in Subsection 5.2.3.

F	F'/F	$R_k(t)$	$D_k(t)$	Flight times	v_{L_2}	v_{L_1}	v_{L_5}
1000	0.20	(12,14)	∞	—	71,000	—	
1000	0.20	(12,14)	(12,14)	Uniform: 2	—	71,000	
1000	0.20	(12,14)	(15,17)	Nonuniform: 1 or 2	—	71,500	
1000	0.40	(10,10)	∞	—	56,000	—	
1000	0.40	(10,10)	(10,10)	Uniform: 2	—	56,000	
1000	0.40	(10,10)	(14,14)	Nonuniform: 1 to 30	—	Infeas.	52,779
1000	0.40	(10,10)	(15,15)	Nonuniform: 1 to 30	—	62,083	43,268
1000	0.40	(10,10)	(16,16)	Nonuniform: 1 to 30	—	57,250	37,586

Table 3.11: Results for various cases with finite departure capacities (formulation (I_1)) and with interdependent departure and arrival capacities (formulation (I_5)).

higher than the value of formulation (I_5) .

3.3.1 Fixed departure capacities have negligible impact.

In order to check the impact of finite departure capacities and to demonstrate that formulation (I_1) , which has more than twice as many variables and three times as many constraints as formulation (I_2) , can be also solved in reasonable computation times, we examined the problems of the first two rows of Table 3.2 with various departure capacities. In order to make meaningful comparisons, the scheduled arrival times were kept unchanged. The new data, besides the departure capacities, were the scheduled departure times or, equivalently, the flight times. Table 3.11 gives results for various combinations of departure capacities and flight times. Airborne marginal delay costs were taken to be 75 versus ground marginal delay costs of 50.

It can be seen from Table 3.11 that, when flight times are uniform (e.g., all equal to 2 time periods) or slightly nonuniform, the differences between finite and infinite departure capacities are negligible. It is only with strongly nonuniform flight times that some minor differences appear. (The nonuniform flight times of Table 3.11 were 1 or 2 time periods for $F'/F = 0.20$ and varied from 1 to 30 time periods for $F'/F = 0.40$.) These results justify one in pursuing the investigation of the multi-

airport GHP with the more manageable formulation (I_2). In any event, however, formulation (I_1) is also manageable (running times for the cases of Table 3.11 were about 2000 CPU seconds).⁴

3.3.2 Interdependent departure and arrival capacities have significant impact.

The last row of Table 3.11 shows results for formulation (I_5), with interdependent departure and arrival capacities (Subsection 2.5.3). The feasible region for departure and arrival capacities was modeled by equations (2.33) and (2.34) (cf. also Figure 2-3). In order to effect a meaningful comparison, the values of $D_{k,max}(t)$ and $R_{k,max}(t)$ were determined by the condition that the intersection point of the two lines in Figure 2-3 correspond to the $D_k(t)$ and $R_k(t)$ of the corresponding formulation (I_2). For instance, in the row of Table 3.11 with $D_k(t) = 15$ and $R_k(t) = 10$, (2.33) and (2.34) give: $3 \times 10 + 15 = 3R_{k,max}(t)$, $10 + 3 \times 15 = 3D_{k,max}(t)$, so that $D_{k,max}(t) = 18.33$ and $R_{k,max}(t) = 15$.

Table 3.11 shows that the cost savings which can be achieved by formulation (I_5) over formulation (I_2) are very significant, about 35–40%. The extra degree of freedom that one can achieve by modifying, in a real-time basis, the mix between departure and arrival capacities, seems to be well worth pursuing.

3.4 Model with flight cancellations.

This section presents results pertaining to formulation (I_3) of Subsection 2.4.1.

⁴The justification in pursuing formulation (I_2) rather than (I_1) is not in the fact that the optimal values v_{I_1} and v_{I_2} are very close (this would be open to the objection concerning the proximity of v_{D_i} and v_{I_i}), but in the fact that the optimal solution of (I_2) violates only a few departure capacity constraints.

3.4.1 Impact of cancellation costs.

Table 3.12 gives results for selected cases from Table 3.2, but for formulation (L_3), and for various capacities and cancellation costs M . The rows with "infinite" cancellation costs correspond to formulation (L_2) and are taken from Table 3.2. All marginal delay costs were equal to 50.

These results strongly support the conclusion that, for cancellation costs greater than 100 times the marginal delay cost (i.e., here, $M > 5000$), no flight is ever cancelled, so that models (I_2) and (I_3)⁵ give the same results. For cancellation costs greater than 20 times the marginal delay costs ($M > 1000$), few flights are cancelled, so that the optimal values of (I_2) and (I_3) are very close. Finally, for cancellation costs less than 10 times the marginal delay cost ($M < 500$), more flights are cancelled and significant differences between (I_2) and (I_3) emerge. Note also that, in that last region of cancellation costs, the slope of the optimal value as a function of the cancellation cost becomes quite abrupt.

Nonuniform capacities.

Table 3.13 gives results concerning cases with nonuniform arrival capacities. (Cf. the discussion in Subsection 3.1.3 explaining why some network effects are expected in this case.) It can be seen that gaps between v_D and v_I are significant.

Three cost classes.

As explained in Subsection 3.1.3, the main reason why network effects were found to be insignificant was the assumption that all cost functions are identical. In order to check this, we ran some cases with three classes of costs: 40% of all flights had cost 100, 40% had cost 50, and 20% had cost 20, corresponding to the relative costs of large, medium-sized, and small aircraft, respectively. Aircraft performing continued flights

⁵Although the results in Table 3.12 are for (L_3) rather than (I_3), one can draw conclusions concerning (I_3) since the investigation of Subsections 3.1.2–3.1.3 showed that network effects (hence, *a fortiori*, gaps between v_{L_i} and v_{I_i}) are insignificant when all cost functions are identical.

F	F'/F	Capacities	M	v_{D_3}	t_{D_3}	v_{L_3}	t_{L_3}
1000	0.60	all 11	1000	70,300	297	70,300	479
1000	0.60	all 10	1000	117,000	286	117,000	475
1000	0.60	all 08	1000	240,700	280	241,805	524
1000	0.60	all 06	1000	402,600	274	403,476	513
1000	0.60	all 04	1000	582,300	272	583,417	484
1000	0.60	all 11	100	28,700	283	28,700	498
1000	0.60	all 11	1000	70,300	297	70,300	473
1000	0.60	all 11	10000	84,200	276	84,300	444
1000	0.60	all 11	∞	84,200	242	84,300	377
2000	0.20	all 14	500	77,500	652	77,500	803
2000	0.20	all 14	1000	94,000	691	94,000	922
2000	0.20	all 14	5000	96,300	717	96,300	931
2000	0.20	all 14	∞	96,300	664	96,300	731
2000	0.40	all 14	500	73,100	815	74,983	1020
2000	0.40	all 14	1000	86,100	690	86,372	1102
2000	0.40	all 14	5000	88,400	675	89,933	1176
2000	0.40	all 14	∞	88,400	652	89,933	973
3000	0.60	all 17	100	38,250	1119	38,693	1911
3000	0.60	all 17	500	71,800	1128	72,240	1708
3000	0.60	all 17	750	81,000	1148	81,338	1931
3000	0.60	all 17	1000	87,000	1187	87,156	2114
3000	0.60	all 17	10000	90,200	1248	96,550	3767
3000	0.60	all 17	∞	90,200	1166	96,550	2547
3000	0.80	all 18	100	36,600	1114	38,042	1846
3000	0.80	all 18	500	71,500	1140	71,559	2320
3000	0.80	all 18	750	78,700	1128	78,707	2693
3000	0.80	all 18	1000	80,500	1235	82,214	2900
3000	0.80	all 18	10000	80,500	1230	84,250	3227
3000	0.80	all 18	∞	80,500	1180	84,250	3072

Table 3.12: Results for various cases with flight cancellations.

F	F'/F	Cap.	M	v_D	t_D	v_L	t_L
3000	0.80	nonun.	500	232,800	1142	252,045	1973
3000	0.80	nonun.	750	302,700	1200	330,040	2217
3000	0.80	nonun.	1000	366,200	1215	403,127	2228

Table 3.13: Results for various cases with flight cancellations and nonuniform capacities.

F	F'/F	Cap.	M	v_{D_3}	t_{D_3}	v_{L_3}	t_{L_3}
3000	0.60	nonun.	500	305,690	1236	373,271	2099
3000	0.60	nonun.	750	385,830	1253	491,791	2219
3000	0.60	nonun.	1000	460,230	1305	601,331	2332

Table 3.14: Results for various cases with flight cancellations and three cost classes.

were generally assigned to the large- or medium-cost category. Results are shown in Table 3.14. It can be seen that the differences are quite significant (22–27%).

3.4.2 A heuristic.

As noted in Subsection 3.1.2, the optimal solution of (L_i) typically has only a small number of flights with noninteger values. This observation provided the motivation for the development of a heuristic which finds a feasible solution of the integer program (I_3) starting from a feasible solution of (L_3) . The previous subsection (3.4.1) showed, on the basis of computational experience, that it is easy to solve (L_3) (cf. the computation times t_{L_3} in Table 3.12). The next subsection (3.4.3) will show that, when one applies the heuristic to the optimal solution of (L_3) , one gets a “good” feasible solution of (I_3) .

The heuristic will be described first verbally in rough outline; then an algorithmic presentation will be given.

Consider a feasible solution $\{v_{ft} : f \in \mathcal{F}, t \in \mathcal{T}_f^a\} \cup \{z_f : f \in \mathcal{F}\}$ of (L_3) and denote by Φ the set of “problematic” flights $f \in \mathcal{F}$, i.e., the set of flights for which some integrality constraint is violated. The heuristic gives a “rounding” scheme for flights in Φ which leaves undisturbed, as far as possible, the remaining flights (which already satisfy integrality). The basic idea of the heuristic is to treat each flight in Φ once for all.

The heuristic starts by partitioning Φ into classes, each class corresponding to an aircraft and containing all and only the flights of Φ scheduled to be performed by that aircraft. The heuristic treats each class separately; the order in which the classes are treated is more or less arbitrary.

Each class is treated in the following way. The flights in the class are examined one at a time, in the order in which they are scheduled to be performed by the aircraft defining the class. For each specific flight ϕ , the heuristic takes the following actions. (It will help the reader at this point to refer to formulation (I_3) in Subsection 2.4.1.) For each time period t at which ϕ can be allowed to land, the heuristic computes the available “capacity slacks” (i.e., the slacks of constraints (2.18)) $R_{k_\phi^a}(t) - \sum_{f:k_f^a=k_\phi^a} v_{ft}$, which will be denoted by $S_\phi(t)$. (If some v_{ft} have already been updated by new values, then the new values are used in the computation of the capacity slacks.) It can be seen that, if $S_\phi(t) \geq 1 - v_{\phi t}$, then it is possible to assign flight ϕ to period t without violating the corresponding capacity constraint. If this is possible for no t , then flight ϕ is cancelled and we are done with it. Otherwise, i.e., when there are time periods to which it is possible to assign flight ϕ without violating the corresponding capacity constraint, flight ϕ is assigned to the earliest such period, τ . (Recall that this assignment is made once for all.) After this assignment, all constraints involving flight ϕ are satisfied, with the possible exception of the coupling constraints.

In order to deal with the coupling constraint linking flight ϕ with its next flight $\hat{\phi}$ (if such a next flight exists), the heuristic removes certain time periods from the set of time periods at which $\hat{\phi}$ can be allowed to land, and proceeds to examine $\hat{\phi}$. The removed time periods are those which would violate the coupling constraint in

question if $\hat{\phi}$ were assigned to them (given that ϕ has already been assigned to τ). It can be seen that, if flight ϕ has a *previous* flight ϕ' , the coupling constraint linking ϕ' and ϕ need not be dealt with while examining flight ϕ , because it has already been dealt with when examining flight ϕ' (since ϕ is the *next* flight to ϕ' .)

As pointed out above, this was only a rough outline, and a more rigorous presentation will be given now. The presentation will be given for the case in which, when a flight is cancelled, the next flight scheduled to be performed by the same aircraft is not affected. The other case, in which the next flight is also cancelled, can be treated *mutatis mutandis*.

BEGIN

Define $\Phi := \{\phi \in \mathcal{F} : (z_\phi \notin \{0, 1\}) \vee (\exists t)(v_{\phi t} \notin \{0, 1\})\}$.

Partition Φ into its equivalence classes corresponding to the equivalence relation “is performed by the same aircraft as”: $\Phi = \bigcup_{\psi=1}^{\Psi} \Phi_\psi$.

Order each class according to the order in which the flights in the class are scheduled to be performed by the aircraft defining the class: $\Phi_\psi = \{\phi_{\psi 1}, \dots, \phi_{\psi, \Xi(\psi)}\}$.

Order the classes, e.g., in decreasing order of the cost of their first flight, and break ties, e.g., according to the increasing order of scheduled arrival times for first flights.

FOR $\psi = 1$ TO Ψ DO:

FOR $\xi = 1$ TO $\Xi(\psi)$ DO:

Set $\phi = \phi_{\psi\xi}$.

IF $\xi = 1$ THEN:

Define $\bar{T}_\phi := T_\phi^a$.

IF ϕ has a previous noncancelled flight ϕ' THEN:

Remove from \bar{T}_ϕ those t that are smaller than $r_\phi + g_{\phi'} - s_{\phi'}$ (because, if ϕ were assigned to such a t , then the coupling constraint linking ϕ and ϕ' would be violated).

END IF

END IF

Define the Capacity slacks $S_\phi(t) := R_{k_\phi^a(t)} - \sum_{f:k_f^a=k_\phi^a} v_{ft}$, $t \in \bar{T}_\phi$.

Define $\hat{T}_\phi := \{t \in \bar{T}_\phi : S_\phi(t) \geq 1 - v_{\phi t}\}$.

IF $\hat{T}_\phi = \emptyset$ THEN

Cancel ϕ : Put $z_\phi = 1, v_{\phi t} = 0, t \in T_\phi^a$.

CONTINUE ξ

END IF

Assign current flight to τ , the smallest element of \hat{T}_ϕ : set $z_\phi = 0, v_{\phi\tau} = 1, v_{\phi t} = 0, t \in T_\phi^a \setminus \{\tau\}$.

IF ϕ has a next flight $\hat{\phi}$ THEN:

IF $\tau - r_\phi - s_\phi > g_{\hat{\phi}}$ AND $\hat{\phi} \notin \Phi$ THEN

Include $\hat{\phi}$ in Φ_ψ as $\phi_{\psi, \xi+1}$ and modify subsequent indices ξ accordingly.

END IF

Define $\bar{T}_{\hat{\phi}} = \{t \in T_{\hat{\phi}}^a : t - r_{\hat{\phi}} \geq \tau - r_\phi - s_\phi\}$.

END IF

CONTINUE ξ

CONTINUE ψ

END

3.4.3 Performance of the heuristic.

Table 3.15 repeats the main part of Table 3.12 but also has a column giving the objective function value v_{H_3} of the feasible solution found by the heuristic. Of course, one always has:

$$v_{D_3} \leq v_{L_3} \leq v_{I_3} \leq v_{H_3}. \quad (3.5)$$

F	F'/F	Capacities	M	v_{D_3}	v_{L_3}	v_{H_3}
1000	0.60	all 11	1000	70,300	70,300	78,500
1000	0.60	all 10	1000	117,000	117,000	125,450
1000	0.60	all 08	1000	240,700	241,805	253,750
1000	0.60	all 06	1000	402,600	403,476	411,500
1000	0.60	all 04	1000	582,300	583,417	586,700
1000	0.60	all 11	100	28,700	28,700	30,250
1000	0.60	all 11	1000	70,300	70,300	78,500
1000	0.60	all 11	10000	84,200	84,300	240,700
1000	0.60	all 11	∞	84,200	84,300	—
2000	0.20	all 14	500	77,500	77,500	82,000
2000	0.20	all 14	1000	94,000	94,000	103,300
2000	0.20	all 14	5000	96,300	96,300	165,200
2000	0.20	all 14	∞	96,300	96,300	—
2000	0.40	all 14	500	73,100	74,983	75,800
2000	0.40	all 14	1000	86,100	86,372	93,650
2000	0.40	all 14	5000	88,400	89,933	168,900
2000	0.40	all 14	∞	88,400	89,933	—
3000	0.60	all 17	100	38,250	38,693	42,350
3000	0.60	all 17	500	71,800	72,240	84,600
3000	0.60	all 17	750	81,000	81,338	95,000
3000	0.60	all 17	1000	87,000	87,156	130,300
3000	0.60	all 17	10000	90,200	96,550	667,750
3000	0.60	all 17	∞	90,200	96,550	—
3000	0.80	all 18	100	36,600	38,042	58,900
3000	0.80	all 18	500	71,500	71,559	83,350
3000	0.80	all 18	750	78,700	78,707	106,800
3000	0.80	all 18	1000	80,500	82,214	111,350
3000	0.80	all 18	10000	80,500	84,250	509,900
3000	0.80	all 18	∞	80,500	84,250	—

Table 3.15: Performance of the heuristic.

It can be seen that v_{H_3} is quite close to v_{L_3} (hence to v_{I_3}) for small cancellation costs. For large cancellation costs, however, the heuristic performs poorly. This was to be expected, because the heuristic will inevitably cancel some flights, and these will inflate the objective function value if the cancellation cost is excessive. This is not worrisome, however, since, as was pointed out in Subsection 3.4.1, for cancellation costs above 1,000 few flights are cancelled, so that for such high cancellation costs neither formulation (I_3) nor the heuristic have much practical use compared to formulation (I_2).

3.5 Summary of results.

This chapter has reached several results pertaining to the static deterministic multi-airport GHP, and it is useful to put these results in perspective in a compact summary.

(1) In the general case (when cost functions differ), network effects, defined as the difference between the optimal objective function values of the integer and the decomposed problems, can be large (Table 3.14). Network effects can also be large when airport capacities are not uniform (Table 3.13).

(2) In the special case where all cost functions are identical, network effects are of small magnitude (Tables 3.1, 3.2, 3.4, 3.5, 3.6).

(3) However, even when all cost functions are identical, the optimal solution of the decomposed problem typically violates a large number of coupling constraints and is thus useless for practical purposes (Table 3.9).

(4) The optimal ground holds resulting from the integer program entail a cost much lower than that resulting from applying a FCFS heuristic (Table 3.10).

(5) In view of (2), on the one hand, and of (3) and (4), on the other hand, one must carefully distinguish the *magnitude* of network effects from their *importance*: network effects can be small but are usually important.

(6) Increasing the allowable upper bounds on delays has little effect on the optimal

objective function value (Table 3.7).

(7) Finite but fixed departure capacities have negligible impact (Table 3.11).

(8) The possibility of having interdependent departure and arrival capacities can result in very significant cost savings (Table 3.11).

(9) As far as the model with flight cancellations is concerned, high cancellation costs are impractical because they result in no flights ever being cancelled (Table 3.12).

(10) The heuristic which finds a feasible solution of the IP with cancellations on the basis of the optimal solution of the LP relaxation performs quite well for low cancellation costs (Table 3.15).

Chapter 4

The dynamic GHP: formulations.

This chapter presents four formulations for the dynamic *deterministic* multi-airport GHP and one formulation for the dynamic *probabilistic* multi-airport GHP. The chapter is divided into two sections, dealing with the deterministic and the probabilistic problem, respectively. Section 4.1 gives four dynamic deterministic formulations which are generalizations of formulations (I_1) – (I_4) given for the static deterministic GHP in Chapter 2. There is a complication, however, arising from the fact that airborne delays cannot be completely avoided in the dynamic case, even if departure capacities are infinite. Section 4.2 presents a formulation for the static probabilistic GHP and extends it to the dynamic case.

4.1 The dynamic deterministic GHP.

4.1.1 Notation.

In the dynamic GHP, decisions are not taken once for all at the beginning of the day; decisions are instead taken at¹ various time periods $\tau \in \mathcal{T}' \subset \mathcal{T}$. \mathcal{T}' is the set of decision time periods. At each decision time period τ , denote by $\mathcal{F}_\tau \subset \mathcal{F}$ the set of flights for which decisions can be taken; i.e., the set of flights not having yet landed

¹By “at” I will understand “just before”. Thus initial decisions are taken just before time period 1, and final decisions may be taken as late as just before time period T .

at¹ τ . Partition \mathcal{F}_τ into \mathcal{F}_τ^g , the set of flights not having yet taken off at τ , and \mathcal{F}_τ^a , the set of flights in the air at τ .

The reason why it makes sense to take new decisions at τ is that one has new departure and arrival capacity forecasts $D_k^\tau(t)$, $R_k^\tau(t)$, for $t \in \mathcal{T}_\tau := \{\tau, \dots, T\}$. Note that these capacity forecasts extend over the remainder of the time horizon, until time period T , rather than just until the next decision time period $\hat{\tau}$.² It will be assumed that these capacity forecasts are perfectly accurate between decision time periods. In other words, the actual airport capacities $D_k(t)$, $R_k(t)$ will be equal to $D_k^\tau(t)$, $R_k^\tau(t)$ for $t \in \{\tau, \dots, \hat{\tau} - 1\}$, $\tau \in \mathcal{T}'$.

In analogy with the notation of Chapter 2, we introduce the following decision variables: g_f^τ , $f \in \mathcal{F}_\tau^g$, is the ground delay of flight f as decided at τ ; a_f^τ , $f \in \mathcal{F}_\tau$, is the airborne delay of flight f as decided at τ ; u_{ft}^τ , $f \in \mathcal{F}_\tau^g$, $t \in \mathcal{T}_f^d \cap \mathcal{T}_\tau$, is 1 if flight f is decided at τ to depart at³ t , and 0 otherwise; and v_{ft}^τ , $f \in \mathcal{F}_\tau$, $t \in \mathcal{T}_f^a \cap \mathcal{T}_\tau$, is 1 if flight f is decided at τ to land at³ t , and 0 otherwise.

Finally, let \hat{G}_f , $f \in \mathcal{F}$, be the actual number of time periods (not a variable) that flight f was held on the ground, and \hat{A}_f , $f \in \mathcal{F}$, be the actual number of time periods (not a variable) that flight f was held in the air. The value of \hat{G}_f is determined at the time f departs, and the value of \hat{A}_f is determined at the time f lands. The final cost of a dynamic policy will be $\sum_{f \in \mathcal{F}} (c_f^g \hat{G}_f + c_f^a \hat{A}_f)$.

Table 4.1 summarizes the above notation for reference purposes. Some symbols from Table 2.1 will also be freely used.

²An alternative way of modeling the dynamic deterministic problem would be to assume that the capacity forecasts at a decision time period extend only until the next decision time period (and are perfectly accurate). This way would enable one to solve larger problems, as explained in Subsection 2.3.2 (p. 37), but would probably often create infeasibility, since the rest of the time horizon (from the next decision time until the end of the day) would not be taken into account. Moreover, one often does not know in advance when the next information update will occur. For these reasons, this alternative way will not be pursued in the sequel.

³This “at”, of course, is not understood as “just before”!

Symbol	Denotation
τ'	Decision time period previous to τ .
$\hat{\tau}$	Decision time period next to τ .
$T' \subset T$	Set of decision time periods τ .
$\mathcal{T}_\tau = \{\tau, \dots, T\}$	Set of remaining time periods.
$\mathcal{F}_\tau \subset \mathcal{F}$	Set of flights not having yet landed at ¹ τ .
\mathcal{F}_τ^g	Set of flights on the ground at τ .
\mathcal{F}_τ^a	Set of flights in the air at τ .
$D_k^\tau(t)$	Departure capacity forecast at ¹ τ for airport k at ³ period t .
$R_k^\tau(t)$	Arrival capacity forecast at τ for airport k at period t .
g_f^τ	Ground delay decision variable at τ for flight f .
a_f^τ	Airborne delay decision variable at τ for flight f .
u_{ft}^τ	Departure assignment decision variable at τ for flight f .
v_{ft}^τ	Arrival assignment decision variable at τ for flight f .
\hat{G}_f	Actual ground delay of flight f (not a variable).
\hat{A}_f	Actual airborne delay of flight f (not a variable).
E_{kt}^τ	Excess at τ of airport k at period t (p. 79).
α_f^τ	Unavoidable airborne delay of flight f in the air at τ (p. 80).

Table 4.1: Notation for the dynamic deterministic GHP.

4.1.2 A first dynamic deterministic formulation.

We present now a dynamic formulation which extends the static formulation (I_1) of Subsection 2.2.3. The dynamic formulation consists in solving, at each decision time period τ , the following pure 0–1 integer program:

$$(I_1^\tau) \min \sum_{f \in \mathcal{F}_\tau^g} c_f^g g_f^\tau + \sum_{f \in \mathcal{F}_\tau} c_f^a a_f^\tau$$

$$\text{s.t. } \sum_{f \in \mathcal{F}_\tau^g: k_f^g = k} u_{ft}^\tau \leq D_k^\tau(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}_\tau; \quad (4.1)$$

$$\sum_{f \in \mathcal{F}_\tau: k_f^a = k} v_{ft}^\tau \leq R_k^\tau(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}_\tau; \quad (4.2)$$

$$\sum_{t \in \mathcal{T}_f^g \cap \mathcal{T}_\tau} u_{ft}^\tau = 1, \quad f \in \mathcal{F}_\tau^g; \quad (4.3)$$

$$\sum_{t \in \mathcal{T}_f^a \cap \mathcal{T}_\tau} v_{ft}^\tau = 1, \quad f \in \mathcal{F}_\tau; \quad (4.4)$$

$$g_{f'}^\tau + a_{f'}^\tau - s_{f'} \leq g_f^\tau, \quad f' \in \mathcal{F}' \cap \mathcal{F}_\tau^g; \quad (4.5)$$

$$\hat{G}_{f'} + a_{f'}^\tau - s_{f'} \leq g_f^\tau, \quad f' \in \mathcal{F}' \cap \mathcal{F}_\tau^a; \quad (4.6)$$

$$a_f^\tau \geq 0, \quad f \in \mathcal{F}_\tau; \quad (4.7)$$

$$u_{ft}^\tau, v_{ft}^\tau \in \{0, 1\}, \quad (4.8)$$

where:

$$g_f^\tau = \sum_{t \in \mathcal{T}_f^g \cap \mathcal{T}_\tau} t u_{ft}^\tau - d_f, \quad f \in \mathcal{F}_\tau^g; \quad (4.9)$$

$$a_f^\tau = \sum_{t \in \mathcal{T}_f^a \cap \mathcal{T}_\tau} t v_{ft}^\tau - r_f - g_f^\tau, \quad f \in \mathcal{F}_\tau^g; \quad (4.10)$$

$$a_f^\tau = \sum_{t \in \mathcal{T}_f^a \cap \mathcal{T}_\tau} t v_{ft}^\tau - r_f - \hat{G}_f, \quad f \in \mathcal{F}_\tau^a. \quad (4.11)$$

Several comments are in order. The whole formulation presents a dichotomy necessitated by the fact that flights on the ground at τ and flights in the air at τ must be treated differently. The objective function is a sum of two terms, corresponding to the ground delay costs of flights on the ground and to airborne delay costs of all flights in \mathcal{F}_τ . Similarly, the departure capacity constraints (4.1) and the departure assignment constraints (4.3) refer only to flights on the ground, while the arrival capacity constraints (4.2) and the arrival assignment constraints (4.4) refer to all flights in \mathcal{F}_τ . The coupling constraints are also divided into two categories because,

for the continued flights f' which are already in the air at τ , the ground hold $\hat{G}_{f'}$ has already been determined: it is a number, not a decision variable (constraints (4.6)). Similar remarks hold for the expression of the delay decision variables a_f^τ in terms of the assignment decision variables (constraints (4.10) and (4.11)).

The integer program (I_1^1) is identical with the static formulation (I_1) , because $\mathcal{F}_1 = \mathcal{F}_1^g = \mathcal{F}$ and $\mathcal{T}_1 = \mathcal{T}$. Having solved the program (I_1^τ) , one solves the program $(I_1^{\hat{\tau}})$ corresponding to the next decision period $\hat{\tau}$ by taking as inputs $D_k^{\hat{\tau}}(t)$, $R_k^{\hat{\tau}}(t)$, $\mathcal{T}_{\hat{\tau}}$, and by updating the flight sets as follows:

$$\mathcal{F}_{\hat{\tau}}^g = \mathcal{F}_{\tau}^g \setminus \{f \in \mathcal{F}_{\tau}^g : d_f + \hat{g}_f^\tau < \hat{\tau}\}; \quad (4.12)$$

$$\mathcal{F}_{\hat{\tau}}^a = (\mathcal{F}_{\tau}^a \setminus \{f \in \mathcal{F}_{\tau}^a : r_f + \hat{G}_f + \hat{a}_f^\tau < \hat{\tau}\}) \cup \{f \in \mathcal{F}_{\tau}^g : (d_f + \hat{g}_f^\tau < \hat{\tau})(r_f + \hat{g}_f^\tau + \hat{a}_f^\tau \geq \hat{\tau})\}, \quad (4.13)$$

where \hat{g}_f^τ and \hat{a}_f^τ are the optimal values returned by (I_1^τ) . In words, (4.12) simply says that the new set of flights on the ground is equal to the previous set of flights on the ground minus the flights that were assigned by (I_1^τ) to leave before the new decision period $\hat{\tau}$. Similarly, (4.13) says that the new set of flights in the air is equal to the old set of flights in the air minus any flights in that set that were assigned to land before the new decision period, plus any flights that were previously on the ground, were assigned to depart before the new decision period, and were assigned to land at or after it.

The final cost resulting from the above dynamic formulation is $\sum_{f \in \mathcal{F}} (c_f^g \hat{G}_f + c_f^a \hat{A}_f)$, where \hat{A}_f is the final airborne delay of flight f . At $\hat{\tau}$ one updates the cost by adding to it the sum of $c_f^g \hat{g}_f^\tau$ for $f \in \{f \in \mathcal{F}_{\tau}^g : d_f + \hat{g}_f^\tau < \hat{\tau}\}$ (i.e., flights which were on the ground at τ but left before $\hat{\tau}$), and the sum of $c_f^a \hat{a}_f^\tau$ for $f \in \{f \in \mathcal{F}_{\tau}^a : r_f + \hat{G}_f + \hat{a}_f^\tau < \hat{\tau}\} \cup \{f \in \mathcal{F}_{\tau}^g : r_f + \hat{g}_f^\tau + \hat{a}_f^\tau < \hat{\tau}\}$ (i.e., flights that either were in the air or were on the ground at τ and landed before $\hat{\tau}$). One also sets $\hat{G}_f = \hat{g}_f^\tau$ and $\hat{A}_f = \hat{a}_f^\tau$ for flights in the above sets.

Note that the size of formulation (I_1^τ) decreases as τ increases, because the sizes of the flight sets and of the time set decrease. An interesting feature of formulation

(I_1^r) is that decisions can be updated or overruled as long as they have not been implemented. For instance, aircraft can be assigned ground or airborne delays smaller or larger than what they had been previously assigned, and aircraft not having yet taken off can be given priority over aircraft already in the air. The last possibility is not expected to appear often in the optimal solution, given that c_f^a is higher than c_f^g . Nevertheless, in practice one would almost never want to deal with this possibility, so formulation (I_1^r) may be too general. The second dynamic formulation, presented below, always gives aircraft in the air priority over aircraft not having yet taken off.

4.1.3 A second dynamic deterministic formulation.

The second dynamic formulation is a generalization of the static formulation (I_2) (Subsection 2.3.2), which assumed infinite departure capacities and eliminated airborne delays. In the dynamic case, however, airborne delays cannot be completely eliminated even if departure capacities are infinite. This is because, at a given decision time, the new arrival capacity forecasts may be significantly reduced with respect to the forecasts of the previous decision time. Then it can happen that even the number of aircraft already in the air exceeds the new capacity forecasts, so that some of these aircraft may have to wait in the air when they arrive at their destination.

For a given decision time period τ , define the *excess* at airport k and time period t , denoted by E_{kt}^τ , as the number of aircraft currently in the air which will arrive at k at t minus the currently forecasted capacity of k at t . At each decision time period τ , one needs to do the following preliminary calculations in order to find what the excesses E_{kt}^τ are:

BEGIN

FOR $k = 1$ TO K DO:

$$E_{k,\tau-1}^\tau = E_{k,\tau-1}^{\tau'}$$

FOR $t = \tau$ TO T DO:

$$E_{kt}^\tau = \max(E_{k,t-1}^\tau, 0) + |\{f \in \mathcal{F}_\tau^a : (k_f^a = k)(r_f + \hat{g}_f^{\tau'} = t)\}| - R_k^\tau(t).$$

CONTINUE t

CONTINUE k

END

$E_{k,\tau-1}^{\tau'}$ are the excesses calculated at τ' ,⁴ the decision time previous to τ . These previously calculated excesses are the actually realized excesses, since capacity forecasts at τ' are accurate until $\tau - 1$. Any positive excess at a time period $t - 1$ is transferred to the next time period t . As a result of the above preliminary calculations, if $E_{kt}^{\tau} < 0$, then $-E_{kt}^{\tau}$ is the *available arrival capacity* of airport k at time period t , i.e., the currently forecasted capacity minus the number of aircraft that are already in the air and will arrive at k at t . On the other hand, if $E_{kt}^{\tau} \geq 0$, then there is no available capacity at k at t , and the best one can do, supposing that aircraft in the air have priority over aircraft on the ground, is to not assign any of the aircraft currently on the ground to arrive at k at t .

Another preliminary calculation which needs to be carried out at each decision time period concerns the unavoidable airborne delays of flights already in the air, arising when the new capacity forecasts are not sufficient to accommodate these flights. These delays, which will be denoted by α_f^{τ} (numbers, not variables), will be needed for the coupling constraints and for the calculation of the total cost. They are calculated by means of the excesses E_{kt}^{τ} in the following way:

BEGIN

FOR $f \in \mathcal{F}_{\tau}^{\alpha}$ DO:

$$\alpha_f^{\tau} = \alpha_f^{\tau'}.^5$$

CONTINUE f

FOR $k = 1$ TO K DO:

FOR $t = \tau$ TO T DO:

⁴At decision time 1, one begins with $E_{k0}^1 = 0$.

⁵At $\tau = 1$, one begins with $\alpha_f^1 = 0$.

IF $E_{kt}^\tau > 0$ THEN

Select E_{kt}^τ flights in $\{f \in \mathcal{F}_\tau^a : (k_f^a = k)(r_f + \hat{g}_f^\tau + \alpha_f^\tau = t)\}$
and set $\alpha_f^\tau = \alpha_f^\tau + 1$ for them.

END IF

CONTINUE t

CONTINUE k

END

The selection of the flights to be delayed depends on the arrival queuing discipline, usually FCFS.

We are now in the position to give the second dynamic deterministic formulation, with infinite departure capacities:

$$(I_2^T) \min \quad \sum_{f \in \mathcal{F}_\tau^g} c_f^g g_f^\tau$$

$$\text{s.t.} \quad \sum_{f \in \mathcal{F}_\tau^g: k_f^a = k} v_{ft}^\tau \leq \max(-E_{kt}^\tau, 0), \quad (k, t) \in \mathcal{K} \times \mathcal{T}_\tau; \quad (4.14)$$

$$\sum_{t \in \mathcal{T}_f^a \cap \mathcal{T}_\tau} v_{ft}^\tau = 1, \quad f \in \mathcal{F}_\tau^g; \quad (4.15)$$

$$g_{f'}^\tau - s_{f'} \leq g_f^\tau, \quad f' \in \mathcal{F}' \cap \mathcal{F}_\tau^g; \quad (4.16)$$

$$\hat{G}_{f'} + \alpha_{f'}^\tau - s_{f'} \leq g_f^\tau, \quad f' \in \mathcal{F}' \cap \mathcal{F}_\tau^a; \quad (4.17)$$

$$v_{ft}^\tau \in \{0, 1\}, \quad (4.18)$$

where:

$$g_f^\tau = \sum_{t \in \mathcal{T}_f^a \cap \mathcal{T}_\tau} t v_{ft}^\tau - r_f, \quad f \in \mathcal{F}_\tau^g. \quad (4.19)$$

Constraints (4.14) say that, if $E_{kt}^\tau < 0$, then the number of aircraft assigned to arrive at airport k at time period t must not exceed the available excess capacity $-E_{kt}^\tau$; whereas, if $E_{kt}^\tau \geq 0$ (i.e., no available capacity at k at t), then v_{ft}^τ will be 0 for all f that could arrive at k at t , so that no new aircraft will be assigned to arrive at k at t . Note that decisions are taken only for flights on the ground at τ . Flights in the air at τ influence the decisions by means of (4.14) (by determining the excesses)

and of (4.17) (their airborne delays enter into the coupling constraints).

The total cost is calculated and updated at each decision time period in a way analogous to that explained in the previous subsection for formulation (I_1^τ) . Note also that, by an argument completely analogous to that presented in Subsection 2.3.1 for the static case, one can show that, if departure capacities are infinite and $c_f^a > c_f^g$, then, if formulation (I_1^τ) without airborne delays as decision variables has an optimal solution, then it has an optimal solution in which, at every decision period τ , $a_f^\tau = 0$ for all flights on the ground at τ .

4.1.4 Flight cancellations.

For the sake of completeness, we will give now the straightforward extensions of formulations (I_2^τ) and (I_1^τ) for the case where flight cancellations are allowed. The extensions are completely analogous to the way in which the static formulations (I_3) and (I_4) extended (I_2) and (I_1) , respectively (cf. Section 2.4). Of course, only flights on the ground may be cancelled; moreover, cancellations are irreversible: once z_f^τ , the cancellation variable corresponding to flight f , is 1, flight f is removed from \mathcal{F}_τ .

A third dynamic deterministic formulation.

Formulation (I_3^τ) extends (I_2^τ) .

$$\begin{aligned}
(I_3^\tau) \quad \min \quad & \sum_{f \in \mathcal{F}_\tau^g} (c_f^g g_f^\tau + (M_f + c_f^g r_f) z_f^\tau) \\
\text{s.t.} \quad & \sum_{f \in \mathcal{F}_\tau^g: k_f^g = k} v_{ft}^\tau \leq \max(-E_{kt}^\tau, 0), & (k, t) \in \mathcal{K} \times \mathcal{T}_\tau; \\
& z_f^\tau + \sum_{t \in \mathcal{T}_f^g \cap \mathcal{T}_\tau} v_{ft}^\tau = 1, & f \in \mathcal{F}_\tau^g; \\
& g_{f'}^\tau - s_{f'} + (s_{f'} + r_{f'} - r_f) z_{f'}^\tau \leq g_f^\tau, & f' \in \mathcal{F}'_1 \cap \mathcal{F}_\tau^g; \\
& g_{f'}^\tau - s_{f'} + (s_{f'} + r_{f'} + G_f + 1) z_{f'}^\tau \leq g_f^\tau + (r_f + G_f + 1) z_f^\tau, & f' \in \mathcal{F}'_2 \cap \mathcal{F}_\tau^g; \\
& \hat{G}_{f'} + \alpha_{f'}^\tau - s_{f'} \leq g_f^\tau, & f' \in \mathcal{F}' \cap \mathcal{F}_\tau^a; \\
& v_{ft}^\tau, z_f^\tau \in \{0, 1\},
\end{aligned}$$

where g_f^τ are given by (4.19).

A fourth dynamic deterministic formulation.

Formulation (I_4^T) extends (I_1^T) .

$$\begin{aligned}
(I_4^T) \quad \min \quad & \sum_{f \in \mathcal{F}_f^g} [c_f^g g_f^T + (M_f + c_f^g d_f + c_f^a (r_f - d_f)) z_f^T] + \sum_{f \in \mathcal{F}_\tau} c_f^a a_f^T \\
\text{s.t.} \quad & \sum_{f \in \mathcal{F}_f^g: k_f^d = k} u_{ft}^T \leq D_k^T(t), (k, t) \in \mathcal{K} \times \mathcal{T}_\tau; \\
& \sum_{f \in \mathcal{F}_\tau: k_f^a = k} v_{ft}^T \leq R_k^T(t), (k, t) \in \mathcal{K} \times \mathcal{T}_\tau; \\
& z_f^T + \sum_{t \in \mathcal{T}_f^d \cap \mathcal{T}_\tau} u_{ft}^T = 1, f \in \mathcal{F}_\tau^g; \\
& z_f^T + \sum_{t \in \mathcal{T}_f^a \cap \mathcal{T}_\tau} v_{ft}^T = 1, f \in \mathcal{F}_\tau^g; \\
& \sum_{t \in \mathcal{T}_f^a \cap \mathcal{T}_\tau} v_{ft}^T = 1, f \in \mathcal{F}_\tau^a; \\
& g_{f'}^T + a_{f'}^T - s_{f'} + (s_{f'} + r_{f'} - r_f) z_{f'}^T \leq g_f^T, f' \in \mathcal{F}'_1 \cap \mathcal{F}_\tau^g; \\
& g_{f'}^T + a_{f'}^T - s_{f'} + (s_{f'} + r_{f'} + G_f + 1) z_{f'}^T \leq g_f^T + (r_f + G_f + 1) z_f^T, f' \in \mathcal{F}'_2 \cap \mathcal{F}_\tau^g; \\
& \hat{G}_{f'} + a_{f'}^T - s_{f'} \leq g_f^T, f' \in \mathcal{F}' \cap \mathcal{F}_\tau^a; \\
& a_f^T \geq 0, f \in \mathcal{F}_\tau; \\
& u_{ft}^T, v_{ft}^T, z_f^T \in \{0, 1\},
\end{aligned}$$

where g_f^T, a_f^T are given by (4.9), (4.10), and (4.11).

4.2 The dynamic probabilistic GHP.

One way to model the case of probabilistic capacities is by considering that capacity forecasts take the forms of various scenaria, each scenario coming with a given probability of realization. In symbols, there are L possible capacity scenaria, and scenario l , having probability p_l ($\sum_{l=1}^L p_l = 1$), is denoted by $R_k^l(t)$, $(k, t) \in \mathcal{K} \times \mathcal{T}$. Note that a capacity scenario involves capacity forecasts for all airports of the network. In other words, p_l is the probability that airport 1 will have capacities $R_1^l(t)$, $t \in \mathcal{T}$, and that airport 2 will have capacities $R_2^l(t)$, $t \in \mathcal{T}$, and so on. This is because capacities at various airports may not be independent, especially for airports close enough to have similar weather.

In the probabilistic GHP, static policies are subject to a “paradox” which does not appear in the deterministic GHP. At the beginning of the day, one knows the possible capacity scenarios and their probabilities of realization, but of course one does not know which scenario will be realized. If the problem is static, one must make irrevocable decisions concerning ground holds at the beginning of the day. But sooner or later some scenario will be realized, and at that point one should normally take into account the new information and update ground holds. So the paradox of the static probabilistic GHP is that new information will inevitably become available but will not be taken into account. The static *deterministic* GHP encounters no similar problem because, by assumption, capacity forecasts are perfectly accurate and no new information will become available.

The above considerations show that static probabilistic formulations may be of somewhat limited practical interest in themselves. Nevertheless, they can be used as building blocks for dynamic probabilistic formulations. This is entirely analogous to the way in which the static deterministic formulations of Chapter 2 were used as building blocks for the dynamic deterministic formulations of Section 4.1. As an example, we will present below, for the case of infinite departure capacities, a static probabilistic formulation and its dynamic extension.

4.2.1 A static probabilistic formulation.

Define the decision variables g_f , equal to the ground delay of flight f , and v_{ft}^l , equal to 1 if scenario l is realized and flight f lands at t , and equal to 0 otherwise. Denote also by a_f^l the airborne delay of flight f if scenario l is realized. Under scenario l , the total delay of flight f , $g_f + a_f^l$, is equal to the difference between the actual arrival time, $\sum_{t \in \mathcal{T}_f^a} t v_{ft}^l$, and the scheduled arrival time, r_f , so that:

$$a_f^l = \sum_{t \in \mathcal{T}_f^a} t v_{ft}^l - r_f - g_f, \quad f \in \mathcal{F}, \quad l \in \{1, \dots, L\}. \quad (4.20)$$

Assuming infinite departure capacities, we have now the following static probabilistic IP formulation:

$$(I_P) \min \sum_{f \in \mathcal{F}} (c_f^g g_f + c_f^a \sum_{l=1}^L p_l a_f^l)$$

$$\text{s.t.} \quad \sum_{f: k_f^a = k} v_{ft}^l \leq R_k^l(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}, l \in \{1, \dots, L\}; \quad (4.21)$$

$$\sum_{t \in \mathcal{T}_f^a} v_{ft}^l = 1, \quad f \in \mathcal{F}, l \in \{1, \dots, L\}; \quad (4.22)$$

$$g_{f'} + a_{f'}^l - s_{f'} \leq g_f, \quad f' \in \mathcal{F}', l \in \{1, \dots, L\}; \quad (4.23)$$

$$g_f \in \{0, 1, \dots, G_f\}, \quad f \in \mathcal{F}; \quad (4.24)$$

$$v_{ft}^l \in \{0, 1\}. \quad (4.25)$$

Although formulation (I_P) looks superficially similar to previously presented formulations, it has several peculiarities which need mentioning. By solving (I_P) , one gets values for g_f and v_{ft}^l . The values for g_f are the ground holds which will be implemented no matter which capacity scenario is ultimately realized, since we are examining the static case. On the contrary, which values v_{ft}^l will be implemented will depend on the capacity scenario that will be realized. If, for instance, scenario 3 is realized, then flight f will land at the time period t for which v_{ft}^3 is equal to 1. Therefore, it can be seen that g_f *cannot* be expressed in terms of v_{ft}^l : there are two independent sets of decision variables. Moreover, (I_P) is not a *pure* 0–1 IP formulation, since g_f are not binary variables.

Another important comment concerns the coupling constraints (4.23). These must ensure that the ground holds g_f , which are irrevocably decided at the beginning of the day, will be feasible no matter which capacity scenario is ultimately realized. But the capacity scenario which will be realized may affect the airborne delays of continued flights, which may in turn affect the ground delays of their next flights. This problem is solved, in (4.23), by having the ground delay of a continuing flight be at least equal to the *maximum* excess delay of its previous flight over all possible capacity scenarios.⁶

⁶ $g_f \geq g_{f'} + a_{f'}^l - s_{f'}, l \in \{1, \dots, L\}$, is equivalent to: $g_f \geq \max(g_{f'} + a_{f'}^l - s_{f'}, l \in \{1, \dots, L\})$.

A final remark concerns the size of formulation (I_P). There are $LKT + (L+1)F + LF'$ constraints, and the number of decision variables is at most $F + \sum_{l=1}^L \sum_{f \in \mathcal{F}} (A_f^l + 1)$, where A_f^l is an upper bound on the airborne delay of flight f under scenario l . Such an upper bound cannot be arbitrarily imposed, but can be calculated for given arrival profiles and capacity scenario. In the worst case, $A_f^l = T - r_f$ will do. Therefore, the number of constraints and the number of variables can become excessive, especially if L is large, but may remain manageable for small L .

The static probabilistic formulation (I_P) generalizes the static deterministic formulation (I_1) but has infinite departure capacities (note that airborne delays as decision variables are allowed). Extensions of the other static deterministic formulations presented in Chapter 2 (e.g., with flight cancellations) are straightforward.

4.2.2 A dynamic probabilistic formulation.

Extending the static probabilistic formulation of the previous subsection to the dynamic probabilistic case, in analogy with the exposition of Section 4.1, is rather straightforward, but there is a minor complication. The complication concerns the way of modeling the additional information that emerges as time goes on. In an extreme case, one of the possible capacity scenario is realized at a time δ and all uncertainty is eliminated after that time. In a more realistic case, at various *realization time periods* $\delta \in \Delta \subset \mathcal{T}$ the probabilities of the various scenario are updated to p_t^δ . In this case the reasonable thing to do is to identify the set of decision time periods \mathcal{T}' with the set of realization time periods Δ , since it is exactly at realization time periods that new information becomes available.

The situation can be explained with the help of Figure 4-1, which refers to a case with $L = 3$ capacity scenario and is also the basis of the computational results presented in Chapter 5.⁷ At time $\tau_0 = 1$, one knows that scenario $S1, S2, S3$ will be

⁷As was pointed out above, a capacity scenario includes capacity forecasts for all airports of the network. Figure 4-1 gives only the parts of scenario $S1, S2$, and $S3$ that correspond to a given airport.

ultimately realized with probabilities $p_1^{\tau_0} = p_1$, $p_2^{\tau_0} = (1-p_1)p_2$, $p_3^{\tau_0} = (1-p_1)(1-p_2)$, respectively. (Moreover, one knows the capacities with certainty until time τ_1 , since all three scenaria coincide until that time.) So at τ_0 one solves formulation (I_P) with $p_l^{\tau_0}$ as above. Now at time τ_1 new information is obtained: either scenario $S1$ is realized or it is not. This new information gives $p_l^{\tau_1}$. If $S1$ is realized, then $p_1^{\tau_1} = 1$ and $p_2^{\tau_1} = p_3^{\tau_1} = 0$. If $S1$ is not realized, then $p_1^{\tau_1} = 0$, $p_2^{\tau_1} = p_2$, and $p_3^{\tau_1} = 1 - p_2$. So at τ_1 one solves formulation (I_P) with $p_l^{\tau_1}$ as above.⁸ Similarly at τ_2 .

We give now the dynamic probabilistic formulation corresponding to (I_P) . The notation generalizes that of Subsection 4.1.1 (cf. Table 4.1).

$$(I_P^\tau) \min \sum_{f \in \mathcal{F}_\tau^g} c_f^g g_f^\tau + \sum_{f \in \mathcal{F}_\tau} c_f^a \sum_{l=1}^L p_l^\tau a_f^{l\tau} \quad (4.26)$$

$$\text{s.t.} \quad \sum_{f \in \mathcal{F}_\tau: k_f^g = k} v_{ft}^{l\tau} \leq R_k^{l\tau}(t), \quad (k, t) \in \mathcal{K} \times \mathcal{T}_\tau, l \in \{1, \dots, L\}; \quad (4.27)$$

$$\sum_{t \in \mathcal{T}_f^g \cap \mathcal{T}_\tau} v_{ft}^{l\tau} = 1, \quad f \in \mathcal{F}_\tau, l \in \{1, \dots, L\}; \quad (4.28)$$

$$g_{f'}^\tau + a_{f'}^{l\tau} - s_{f'} \leq g_f^\tau, \quad f' \in \mathcal{F}' \cap \mathcal{F}_\tau^g, l \in \{1, \dots, L\}; \quad (4.29)$$

$$\hat{G}_{f'} + a_{f'}^{l\tau} - s_{f'} \leq g_f^\tau, \quad f' \in \mathcal{F}' \cap \mathcal{F}_\tau^a, l \in \{1, \dots, L\}; \quad (4.30)$$

$$g_f^\tau \in \{0, 1, \dots, G_f\}, \quad f \in \mathcal{F}; \quad (4.31)$$

$$v_{ft}^{l\tau} \in \{0, 1\}, \quad (4.32)$$

where:

$$a_f^{l\tau} = \sum_{t \in \mathcal{T}_f^g \cap \mathcal{T}_\tau} t v_{ft}^{l\tau} - r_f - g_f^\tau, \quad f \in \mathcal{F}_\tau^g, l \in \{1, \dots, L\}; \quad (4.33)$$

$$a_f^{l\tau} = \sum_{t \in \mathcal{T}_f^a \cap \mathcal{T}_\tau} t v_{ft}^{l\tau} - r_f - \hat{G}_f, \quad f \in \mathcal{F}_\tau^a, l \in \{1, \dots, L\}. \quad (4.34)$$

This concludes the presentation of the dynamic formulations. The next chapter will present computational results based on the above framework.

⁸Of course at τ_1 the new scenaria extend from τ_1 to T (rather than from τ_0 to T as they did at τ_0).

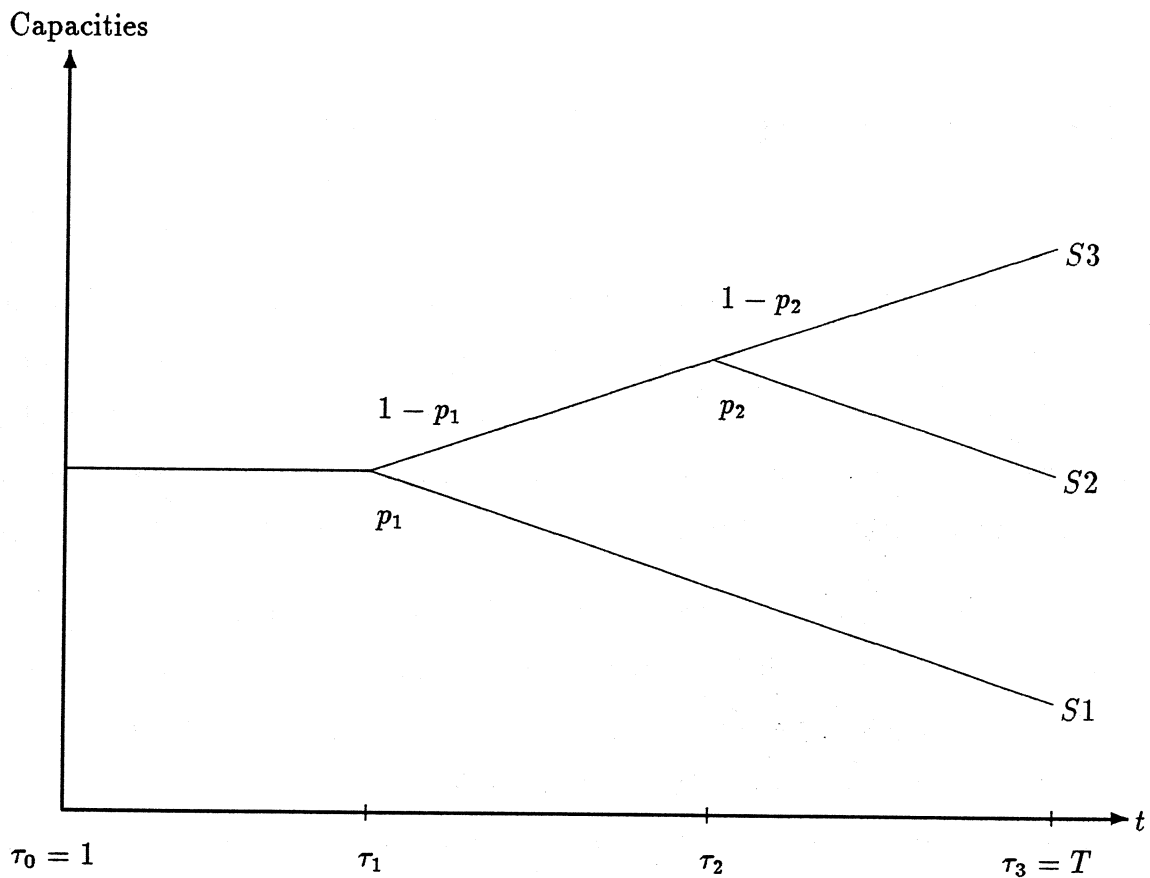


Figure 4-1: Modeling additional information over time in the dynamic probabilistic GHP. There are three possible capacity scenarios: S_1 , S_2 , and S_3 . Overall, S_1 has lower capacities than S_2 , and S_2 has lower capacities than S_3 . All three scenarios coincide between time periods τ_0 and τ_1 , and scenarios S_2 and S_3 coincide between time periods $\tau_1 + 1$ and τ_2 . At τ_1 , S_1 is realized with probability p_1 and all uncertainty is eliminated. Otherwise (with probability $1 - p_1$), at τ_2 , S_2 is realized with probability p_2 or S_3 is realized with probability $1 - p_2$.

Chapter 5

The dynamic GHP: results.

This chapter investigates the behaviour of the dynamic multi-airport GHP by means of computational results based on formulation (I_2^r) of Subsection 4.1.3 and on a dynamic heuristic generalizing the static heuristic presented in Subsection 3.2.2. The framework is a case with three capacity scenaria and two realization times, as in the example of Subsection 4.2.2 (Figure 4-1).

The chapter is divided into three sections. Section 5.1 explains that 15 different *dynamic scenaria* can result from the framework of the three *capacity scenaria*, and examines the relationships between these 15 dynamic scenaria on the basis of a series of test cases with 3 airports and 1500 flights. Section 5.2 introduces *forecasting methods* (most-probable, worst-case), and compares the expected cost of *dynamic ground-holding policies* based on these methods. The behaviour of the dynamic heuristic is also examined. Finally, Section 5.3 summarizes the results of the chapter.

5.1 Comparing dynamic scenaria.

5.1.1 The goal.

The goal of this section is to gain insight on the behaviour of the dynamic multi-airport GHP by examining a relatively realistic case with three capacity scenaria and two realization times (identical with the two decision times), on the model of

Figure 4-1. In the most general case (see Figure 4-1), one solves first $(I_2^{\tau_0})$ with capacity forecasts equal to $S1$ or $S2$ or $S3$, for $t \in \{1, \dots, T\}$. Then one solves $(I_2^{\tau_1})$ with new capacity forecasts (e.g., $S2$ or $S3$ if the previous forecast was $S1$), but now for $t \in \{\tau_1, \dots, T\}$. Finally one solves $(I_2^{\tau_2})$ with yet new capacity forecasts and for $t \in \{\tau_2, \dots, T\}$. In special cases, one may not need to solve $(I_2^{\tau_1})$, $(I_2^{\tau_2})$, or both, if the capacity forecast does not change. Suppose, for instance, that the forecast at τ_0 is $S2$ and that $S1$ is realized at τ_1 . Then one needs to solve $(I_2^{\tau_1})$ with forecast $S1$, but no new problem needs to be solved at τ_2 , since the forecast will inevitably remain $S1$.

A particular combination of (at most three) problems to be solved in the dynamic case will be referred to as a (*dynamic*) *scenario* (not to be confused with a *capacity scenario*, which is one of $S1$, $S2$, $S3$). In the example at the end of the last paragraph, the scenario will be referred to as 2-1, since one solves $(I_2^{\tau_0})$ with forecasts $S2$ and then $(I_2^{\tau_1})$ with forecasts $S1$. Some reflection should convince the reader that, assuming all branches in Figure 4-1 have nonzero probabilities, there are 15 possible dynamic scenaria that one may have to solve, depending on which capacity scenaria are forecasted and which are realized. These 15 dynamic scenaria are given in Table 5.1.

To be quite explicit, the relationship between the “forecast” and the “realization” columns of Table 5.1 is the following: only a capacity scenario which can be realized may be forecasted, and *any* capacity scenario which can be realized may be forecasted. At τ_0 , for instance, all three capacity scenaria are possible, so any of them may be forecasted. At τ_1 , if $S1$ is realized, then only $S1$ can be forecasted, whereas, if $S1$ is not realized, then either $S2$ or $S3$ may be forecasted. In practice, of course, which capacity scenario will be forecasted will depend on the probabilities of the branches in Figure 4-1 and on the forecasting method (e.g., most-probable, worst-case, etc). Probabilities and forecasting methods will be introduced in the next section; for the moment we are just examining all possible cases.

We want to examine the possible dynamic scenaria before introducing probabilities

Dynamic Scenario	Forecast at τ_0	Realization at τ_1	Forecast at τ_1	Realization at τ_2	Forecast at τ_2
1	$S1$	$S1$	$S1$	—	$S1$
1-2	$S1$	not $S1$	$S2$	$S2$	$S2$
1-3	$S1$	not $S1$	$S3$	$S3$	$S3$
1-2-3	$S1$	not $S1$	$S2$	$S3$	$S3$
1-3-2	$S1$	not $S1$	$S3$	$S2$	$S2$
2	$S2$	not $S1$	$S2$	$S2$	$S2$
2-1	$S2$	$S1$	$S1$	—	$S1$
2-3	$S2$	not $S1$	$S3$	$S3$	$S3$
2-2-3	$S2$	not $S1$	$S2$	$S3$	$S3$
2-3-2	$S2$	not $S1$	$S3$	$S2$	$S2$
3	$S3$	not $S1$	$S3$	$S3$	$S2$
3-1	$S3$	$S1$	$S1$	—	$S1$
3-2	$S3$	not $S1$	$S2$	$S2$	$S2$
3-2-3	$S3$	not $S1$	$S2$	$S3$	$S3$
3-3-2	$S3$	not $S1$	$S3$	$S2$	$S2$

Table 5.1: The 15 possible dynamic scenaria (cf. Figure 4-1).

of realization because there are interesting and insightful comparisons to be made. As an example, the cost of scenario 3 must be lower than (or equal to) the cost of scenario 3-3-2, because in 3-3-2 the new capacity forecasts at τ_2 are lower than the previous forecasts, so some already departed aircraft may have to wait in the air. In other cases the comparison is less clear. Compare, for instance, the costs of scenarios 1-2-3 and 1-3-2. At τ_2 some departed aircraft in 1-3-2 may have to wait in the air, while this is not the case in 1-2-3. But the departed aircraft in 1-3-2 have probably had lower ground holds than in 1-2-3, since the previous capacity forecasts were more optimistic in 1-3-2 than in 1-2-3. There is thus a trade-off between ground and airborne delay costs, and it is interesting to pursue the investigation further.

5.1.2 Results and discussion.

Computational experiments were performed for four cases with $K = 3$ airports and $F = 1500$ flights. All cases have the same scheduled arrival profile but different percentages of continued flights, F'/F , ranging from 0.20 to 0.80. The four cases are comparable, in the sense that cases with lower F'/F are obtained from cases with higher F'/F by eliminating some connections between flights. All four cases have slacks equal to 0 and identical arrival capacity profiles in the spirit of Figure 4-1, with the difference that S_2 has a positive slope and S_1 is constant (rather than S_1 and S_2 having negative slopes, as they have in Figure 4-1). S_1 is at the infeasibility limit, and is equal to 11, 10, and 10 aircraft per time period for airports 1, 2, and 3, respectively. The two realization and decision times are $\tau_1 = 21$ and $\tau_2 = 41$; the time horizon has $T = 64$ time periods. Ground delay costs are 50 and airborne delay costs are 75.

The results of the computations are shown in Tables 5.2 and 5.3. It should be noted that the computations were performed with the LP relaxation (L_2^r) rather than (I_2^r), because the values of (L_2^r) and (I_2^r) are very close for identical cost functions (cf. Chapter 3). Nevertheless, rounding mistakes are expected.

After some reflection, one can make the following observations on the basis of

Dynamic Scenario	No capacity updates	Cap. updated at $\tau_1 = 21$ only	Cap. updated at $\tau_2 = 41$ only	Cap. updated at both τ_1 and τ_2
1	50,550			
1-2		35,700		
1-3		32,250		
1-2-3				32,500
1-3-2				35,850
2	35,700			
2-1		50,550		
2-3		32,250		
2-2-3				32,500
2-3-2				35,850
3	32,250			
3-1		50,550		
3-2		35,700		
3-2-3				32,500
3-3-2				35,850
1	50,550			
1-2		36,100		
1-3		32,800		
1-2-3				32,400
1-3-2				35,200
2	36,100			
2-1		50,550		
2-3		32,800		
2-2-3				32,667
2-3-2				35,775
3	32,800			
3-1		50,550		
3-2		36,100		
3-2-3				32,869
3-3-2				34,950

Table 5.2: Values of the 15 dynamic scenarios for $F'/F = 0.20$ (upper part of the table) and for $F'/F = 0.40$ (lower part).

Dynamic Scenario	No capacity updates	Cap. updated at $\tau_1 = 21$ only	Cap. updated at $\tau_2 = 41$ only	Cap. updated at both τ_1 and τ_2
1	63,188			
1-2		40,250		
1-3		35,750		
1-2-3				37,895
1-3-2				38,574
2	40,250			
2-1		63,188		
2-3		35,750		
2-2-3			36,020	
2-3-2				40,092
3	35,750			
3-1		63,188		
3-2		40,250		
3-2-3				37,013
3-3-2			38,447	
1	70,500			
1-2		46,100		
1-3		40,150		
1-2-3				42,974
1-3-2				45,922
2	46,100			
2-1		70,500		
2-3		40,150		
2-2-3			41,315	
2-3-2				44,755
3	40,150			
3-1		70,500		
3-2		46,100		
3-2-3				41,337
3-3-2			46,700	

Table 5.3: Values of the 15 dynamic scenarios for $F'/F = 0.60$ (upper part of the table) and for $F'/F = 0.80$ (lower part).

Tables 5.2 and 5.3. First, the 15 dynamic scenarios fall into two groups, call them group *A* and group *B*. Group *A* has three subgroups of three scenarios each. Subgroup *A1* consists of scenarios 1, 2-1, and 3-1; subgroup *A2* consists of scenarios 2, 1-2, and 3-2; and subgroup *A3* consists of scenarios 3, 1-3, and 2-3. It can be seen that, for any of the four values of F'/F , all three scenarios within any of the above three subgroups have equal values. In symbols, we always have:

$$v_1 = v_{2-1} = v_{3-1} > v_2 = v_{1-2} = v_{3-2} > v_3 = v_{1-3} = v_{2-3}. \quad (5.1)$$

Why should this be so? The reason is probably that all scenarios within the above subgroups (with the exception of the static scenarios 1, 2, 3) have capacity updates only once, and at a rather early time period ($\tau_1 = 21$ with $T = 64$). Recall that all three capacity scenarios coincide until τ_1 . In conclusion, (5.1) seems to say that, *if one can get the correct capacity forecasts early enough in the day, one can almost completely compensate for incorrect capacity forecasts made at the beginning of the day*. It should be noted that one does not always expect the equalities (5.1) to hold exactly. For instance, scenario 3-2 can sometimes have a higher value than scenario 2, because some aircraft in the air at τ_1 may have to incur airborne delays when the forecast shifts from $S3$ to $S2$.¹

The second group, group *B*, consists of two subgroups of three scenarios each. Subgroup *B2* consists of scenarios 1-3-2, 2-3-2, and 3-3-2; subgroup *B3* consists of scenarios 1-2-3, 2-2-3, and 3-2-3.² These subgroups are detected by means of the case $F'/F = 0.20$ (upper part of Table 5.2), where all scenarios within each of these two subgroups have equal values. In symbols:

$$v_{1-2-3} = v_{2-2-3} = v_{3-2-3} < v_{1-3-2} = v_{2-3-2} = v_{3-3-2}, \text{ for } F'/F = 0.20. \quad (5.2)$$

¹For an example of a case where (5.1) do not hold exactly, see Table 5.6.

²The reason why the two subgroups of group *B* were named *B2* and *B3*, rather than *B1* and *B2*, is that subgroups *B2* and *A2* (and, similarly, subgroups *B3* and *A3*) share an important feature: all scenarios within *B2* and *A2* end in 2. In other words, all six scenarios in these two subgroups correspond to the case where capacity scenario $S2$ is ultimately realized.

Examination of the other three cases for F'/F supports (5.2), although the equalities become approximate.

The equalities in group B have presumably the same origin as those in group A : incorrect forecasts are corrected early enough, so that their influence is minimized. Scenarios within each subgroup of group B (e.g., scenarios 1-2-3 and 3-2-3) differ only up to time $\tau_1 = 21$.

One can also compare scenario 3 with 2-2-3, or scenario 2 with 3-3-2. Here we have incorrect predictions which are corrected late in the day (at $\tau_2 = 41$ with $T = 64$). It is seen that only minor differences appear. In other words, the values of scenarios within subgroup $B2$ are quite close to the values of scenarios within subgroup $A2$ (similarly for subgroups $B3$ and $A3$). Therefore, it seems that getting the correct capacity forecasts even relatively late in the day suffices to minimize the impact of incorrect capacity forecasts made at the beginning of the day.

The main result of this section is that, if one gets the correct capacity forecasts before the end of the day, then one can, for the most part, compensate for incorrect capacity forecasts made earlier on. Referring again to Figure 4-1, one can understand the reason. The difference between different dynamic scenarios is mainly due to aircraft which may have to wait in the air when the new forecast is lower than the previous one. Such aircraft are in the air at the current decision time period, so they will arrive at their destination soon. So the only differences between capacity scenarios that really matter are the differences in the vicinity of decision time periods. But such differences are usually small, because the capacity scenarios usually diverge smoothly rather than abruptly. This is why dynamic scenarios ending in i will have values quite close to v_i . What mostly matters is to get the correct forecasts, not *when* to get them.

Range of p_1	Range of p_2	Most probable scenario at τ_0
$p_1 > 1/2$	$0 \leq p_2 \leq 1$	$S1$
$1/3 < p_1 < 1/2$	$p_2 > p_1/(1 - p_1)$	$S2$
	$p_1/(1 - p_1) > p_2 > (1 - 2p_1)/(1 - p_1)$	$S1$
	$p_2 < (1 - 2p_1)/(1 - p_1)$	$S3$
$p_1 < 1/3$	$p_2 > 1/2$	$S2$
	$p_2 < 1/2$	$S3$

Table 5.4: Most probable capacity scenario at τ_0 for various probability combinations (cf. Figure 4-1).

5.2 Comparing dynamic policies.

5.2.1 The goal.

A *dynamic policy* is defined as a pair of: (i) a method of forecasting capacity given probabilistic information (e.g., most-probable, worst-case) and (ii) a method of assigning ground holds given a capacity forecast (e.g., by solving one of the formulations in this thesis). The aim of this section is to investigate the relative efficiency of various dynamic policies. Four policies will be examined, corresponding to two methods of forecasting capacity and to two methods of assigning ground holds.

The two methods of forecasting capacity will be the most-probable forecast and the worst-case forecast. Referring again to Figure 4-1, at time τ_0 the *most-probable* forecast is $S1$ if $p_1 > 1/2$, while it is $S2$ if $p_1 < 1/3$ and $p_2 > 1/2$. All possible cases are given in Table 5.4. Similarly, at τ_1 , supposing $S1$ is not realized, $S2$ is most probable when $p_2 \geq 1/2$. The *worst-case* forecast, on the other hand, is obviously $S1$ at τ_0 and, if $S1$ is not realized, $S2$ at τ_1 .

The two methods of assigning ground holds given a capacity forecast will be the dynamic deterministic³ formulation (I_2^T) of Subsection 4.1.3 and a dynamic FCFS

³It should be clear that we are solving the dynamic deterministic formulation of Subsection 4.1.3 rather than the dynamic probabilistic formulation of Subsection 4.2.2: we are using probabilities

heuristic (given in Subsection 5.2.3) generalizing the static FCFS heuristic of Subsection 3.2.2.

The way of assessing these four dynamic policies will be by comparing their expected values. Subsection 5.2.2 compares the expected values of forecasting methods, and Subsection 5.2.3 examines the performance of the dynamic FCFS heuristic.

5.2.2 Comparing forecasting methods.

This subsection deals with the two dynamic policies in which capacity is forecasted by the most-probable or the worst-case method and then ground holds are assigned by means of the dynamic deterministic formulation (I_2^r). The two policies based on the dynamic FCFS heuristic will be examined in the next subsection.

For given values of the probabilities p_1 and p_2 , the expected value of the most-probable policy is:

$$MP = p_1 v_{MP|1} + (1 - p_1)p_2 v_{MP|2} + (1 - p_1)(1 - p_2)v_{MP|3}, \quad (5.3)$$

where $v(MP|i)$ is the value of the dynamic policy given that capacity scenario i is ultimately realized. For instance, if $S2$ is realized and $p_1 > 1/2$, $p_2 < 1/2$, the most-probable forecasts will be $S1$ at τ_0 and $S3$ at τ_1 , so that $v_{MP|2} = v_{1-3-2}$. It is seen that $v_{MP|i}$ depends not only on i but also on p_1 and p_2 .

For given values of the probabilities p_1 and p_2 , the expected value of the worst-case policy is:

$$WC = p_1 v_{WC|1} + (1 - p_1)p_2 v_{WC|2} + (1 - p_1)(1 - p_2)v_{WC|3}. \quad (5.4)$$

It is easily seen that, for the case of Figure 4-1, $v_{WC|1} = v_1$, $v_{WC|2} = v_{1-2}$, and $v_{WC|3} = v_{1-2-3}$, regardless of the values of p_1 and p_2 .

only to generate the deterministic capacity forecasts.

In order to compare the dynamic with the static GHP, we also included in the comparison the expected value of a *random-selection* static policy:

$$RS = p_1 v_1 + (1 - p_1) p_2 v_2 + (1 - p_1)(1 - p_2) v_3, \quad (5.5)$$

where v_i is the value of scenario i . The random-selection static forecast is defined by means of a probabilistic event: at τ_0 , one performs an experiment which yields outcomes 1, 2, and 3, with probabilities p_1 , $(1-p_1)p_2$, and $(1-p_1)(1-p_2)$, respectively. If outcome $i \in \{1, 2, 3\}$ occurs, then the random-selection forecast is Si , and is not updated at later decision times.

Table 5.5 gives the expected values of the two dynamic policies under consideration and of the static random-selection policy for typical values of p_1 and p_2 corresponding to the combinations of Table 5.4. It is seen that the expected values of both dynamic policies are always very close to the expected value of the random-selection policy and to each other. It seems, therefore, that both forecasting methods perform equally well. On the other hand, the dynamic policies seem to result in no significant cost savings over the static random-selection policy.

These results were expected, given the conclusion of Section 5.1. In fact, $v_{MP|i}$ and $v_{WC|i}$ are always close to v_i , so (5.3), (5.4), and (5.5) entail the approximate equality of MP , WC , and RS .

5.2.3 Dynamic FCFS heuristic highly inefficient.

The static FCFS heuristic presented in Subsection 3.2.2 can be easily generalized to the dynamic case. The only (minor) complication is that, at each decision time period, one must first take care of the next flights of the aircraft currently in the air, in order to satisfy the coupling constraints. The result is as follows.

BEGIN

Initialize: $g_f = 0$.

FOR $t = \tau$ TO T DO:

F'/F	p_1	p_2	Random-selection	Most-probable	Worst-case
0.20	0.8	0.8	47,442	47,452	47,452
0.20	0.8	0.2	47,028	47,586	47,068
0.20	0.4	0.8	41,226	41,256	41,256
0.20	0.4	0.6	40,812	40,872	40,872
0.20	0.4	0.4	40,398	40,434	40,488
0.20	0.4	0.2	39,984	40,002	40,104
0.20	0.2	0.8	38,118	38,158	38,158
0.20	0.2	0.2	36,462	36,486	36,622
0.40	0.8	0.8	47,528	47,512	47,512
0.40	0.8	0.2	47,132	47,096	47,068
0.40	0.4	0.8	41,484	41,276	41,436
0.40	0.4	0.6	41,088	40,992	40,992
0.40	0.4	0.4	40,692	40,476	40,548
0.40	0.4	0.2	40,296	40,158	40,104
0.40	0.2	0.8	38,462	38,441	38,398
0.40	0.2	0.2	36,878	36,694	36,622
0.60	0.8	0.8	58,420	58,506	58,506
0.60	0.8	0.2	57,880	57,813	58,224
0.60	0.4	0.8	48,885	48,918	49,143
0.60	0.4	0.6	48,345	48,860	48,860
0.60	0.4	0.4	47,805	47,403	48,577
0.60	0.4	0.2	47,265	47,049	48,295
0.60	0.2	0.8	44,117	44,161	44,461
0.60	0.2	0.2	41,958	41,669	43,330
0.80	0.8	0.8	65,382	65,495	65,495
0.80	0.8	0.2	64,668	64,661	65,160
0.80	0.4	0.8	55,146	55,286	55,485
0.80	0.4	0.6	54,432	55,110	55,110
0.80	0.4	0.4	53,718	53,675	54,735
0.80	0.4	0.2	53,004	53,076	54,360
0.80	0.2	0.8	50,028	50,214	50,480
0.80	0.2	0.2	47,172	47,268	48,979

Table 5.5: Expected values of dynamic policies and of static random-selection policies for various probability combinations.

FOR $k = 1$ TO K DO:

FOR $f \in \mathcal{F}_\tau^a$ DO:

IF $k_f^a = k$ AND $r_f + \hat{G}_f + \alpha_f = t$ AND f has a next flight \hat{f} THEN $g_f = \max(\hat{G}_f + \alpha_f - s_f, 0)$. Similarly if \hat{f} has a next flight and so on.

CONTINUE f

Define $\mathcal{S}_k(t) := \{f \in \mathcal{F}_\tau^g : (k_f^a = k)(r_f + g_f = t)\}$.

Define $S_k(t) := |\mathcal{S}_k(t)|$.

IF $S_k(t) > -E_k(t)$ THEN:

Choose $Q_k(t) := S_k(t) + \min(E_k(t), 0)$ flights in $\mathcal{S}_k(t)$.

FOR $f = 1$ TO $Q_k(t)$ DO:

Set $g_f = g_f + 1$.

IF f has a next flight \hat{f} THEN:

IF $g_f > s_f$ THEN:

Set $g_{\hat{f}} = g_{\hat{f}} + 1$, and similarly if \hat{f} has a next flight and so on.

END IF

END IF

CONTINUE f

END IF

CONTINUE k

CONTINUE t

END

Computations were performed for the same case as in Subsection 5.1.2 (3 airports and 1500 flights), but only for $F'/F = 0.80$, and with slacks equal to 1 instead of 0. The new capacity scenaria are lower, since the infeasibility limit is lower (due to the

increase of the slack). The new $S1$ is equal to 10, 9, and 9 aircraft per time period for airports 1, 2, and 3, respectively.

The values of the 15 dynamic scenarios, both for formulation (I_2^r) and for the dynamic FCFS heuristic, are given in Table 5.6. It is seen that, as was the case in the static problem, the FCFS heuristic always results in a cost much higher than what one can achieve by solving the multi-airport GHP optimally. This is a significant result, since the dynamic FCFS heuristic is, to some extent, a reasonable approximation of the essential aspects of current ground-holding practice (cf. Subsection 1.3.1).

Finally, Table 5.7 gives the expected values of all four dynamic policies and of the corresponding two static random-selection policies for the case under consideration. It is seen that the two dynamic policies which use the FCFS heuristic perform about equally well, but perform much more poorly than the two policies which use formulation (I_2^r). Moreover, in accordance with the results of the previous subsection, it is seen that the dynamic policies result in no significant cost savings over the static random-selection policies.

5.3 Summary of results.

This chapter has reached the following main conclusions on the dynamic GHP.

(1) If incorrect capacity forecasts made at the beginning of the day are corrected early enough, then their influence on the total cost of the dynamic problem can be minimized.

(2) The most-probable and the worst-case methods of forecasting capacities given probabilistic information perform about equally well, if their performance is measured by their expected values.

(3) The dynamic policies using either a most-probable or a worst-case forecasting method perform, in terms of their expected values, about as well as a static random-selection policy. (Note, however, that the static random-selection policy will have a

Dynamic Scenario	No capacity updates	Cap. updated at $\tau_1 = 21$ only	Cap. updated at $\tau_2 = 41$ only	Cap. updated at both τ_1 and τ_2
1	151,850 (97,163; 56.3%)			
1-2		108,600 (62,814; 72.9%)		
1-3		104,500 (60,765; 72.0%)		
1-2-3				99,750 (63,805; 56.34%)
1-3-2				103,825 (64,373; 61.3%)
2	108,550 (63,050; 72.2%)			
2-1		151,900 (98,034; 54.9%)		
2-3		104,450 (61,975; 68.5%)		
2-2-3			99,700 (65,194; 52.9%)	
2-3-2				103,775 (66,550; 55.9%)
3	104,450 (61,000; 71.2%)			
3-1		151,900 (97,347; 56.0%)		
3-2		108,550 (63,287; 71.5%)		
3-2-3				99,700 (63,307; 57.5%)
3-3-2			103,775 (66,650; 55.7%)	

Table 5.6: Values of (a) the dynamic FCFS heuristic, and (b) (in parentheses) the corresponding exact optimum and the percentage of cost overestimation resulting from applying the heuristic (overestimation = (value of heuristic/exact optimum)-1).

F'/F	p_1	p_2	RS	MP	WC	RS_{FCFS}	MP_{FCFS}	WC_{FCFS}
0.80	0.8	0.8	90,258	90,333	90,333	143,026	142,846	142,846
0.80	0.8	0.2	90,012	90,028	90,452	142,534	142,353	141,784
0.80	0.4	0.8	76,449	77,301	76,673	125,378	124,828	124,838
0.80	0.4	0.6	76,203	76,791	76,791	124,886	123,776	123,776
0.80	0.4	0.4	75,957	76,928	76,910	124,394	123,278	122,714
0.80	0.4	0.2	75,711	76,217	77,029	123,902	123,349	121,652
0.80	0.2	0.8	69,545	70,390	69,842	116,554	115,804	115,834
0.80	0.2	0.2	68,561	69,173	70,318	114,586	113,832	111,586

Table 5.7: Expected values of dynamic policies and of static random-selection policies for various probability combinations.

greater variance than the dynamic policies, since the latter are designed to minimize the effect of initial incorrect predictions and succeed in so doing, given (1) above.)

(4) Finally, and most importantly, assigning ground holds by means of the dynamic FCFS heuristic is highly inefficient compared with optimal ground holds based on the formulations of Chapter 4. This result is particularly important because the dynamic FCFS heuristic may approximate, to some extent, current ground-holding practice (cf. Subsection 1.3.1).

Chapter 6

Conclusions.

6.1 Review of main points.

The fundamental contribution of this thesis is an *efficient method* for modeling various versions of the multi-airport GHP. The thesis is organized naturally into two parts: *presentation* of the method and *demonstration* of its efficiency. These two parts correspond, respectively, to the *formulations* (Chapters 2 and 4) and to the *computational results* (Chapters 3 and 5). The computational results serve not only to demonstrate the efficiency of the method, but also to provide insight on the behaviour of the multi-airport GHP, and this is a second important contribution of the thesis.

As far as the formulations are concerned, it should be clear by now that their sum constitutes a *method* rather than a mere bag of tricks. By this we mean that they are sufficiently flexible to accommodate various degrees of modeling detail, rather than being rigidly limited to some particular cases.

In order to appreciate the flexibility of the method, recall that, starting with the simplest formulations (I_1) and (I_2) for the static deterministic case, the following extensions were presented:

(1) Flights may be cancelled. When a continued flight is cancelled, the next flight scheduled to be performed by the same aircraft can be either also cancelled or not

affected at all (because, e.g., it is performed by a spare aircraft).

(2) Continued flights may have more than one “next” flights, e.g., because passengers in the continued flight connect to several other flights (especially in hub airports).

(3) Aircraft in the air may be speeded up, even to the point of arriving at their destination before their scheduled landing time.

(4) Departure and arrival capacities may be interdependent; their mix may be modified from time to time by controlling runway use.

(5) In order to avoid discrimination, one may adopt constraints ensuring that the optimal ground-holding policy will delay a similar percentage of continued and of noncontinued flights. Similarly, one might adopt additional constraints in order to avoid discrimination against short-duration flights or other forms of discrimination among airlines or other classes of users.

(6) One may update dynamically ground (and airborne) holds whenever better weather (and hence capacity) forecasts become available.

(7) One may introduce probabilistic capacity forecasts.

Two points are worth making here. First, the above list is an open one. New generalizations may be devised according to the demands of the real-world situation. Second, and most important, these generalizations are not mutually exclusive: they can be combined to yield even more general formulations. It is easy, for instance, to write a dynamic probabilistic formulation with flight cancellations, interdependent departure and arrival capacities, continued flights having more than one next flights, the possibility of speeding aircraft in the air, and so on. This is the power of the *method*.

As far as the computational results are concerned, these perform three functions: showing that the formulations can be solved for large-scale problems in reasonable computation times, comparing optimal solutions with alternative ways of assigning ground holds for a network of airports, and providing insight on the behaviour of the

multi-airport GHP for various combinations of the input parameters.

Chapter 3 presented results for networks with as many as 6 airports and 3000 flights. But this is not in any sense an upper limit; we simply chose not to perform computations with larger networks. Moreover, the computations were not performed by means of some custom-tailored algorithm, but by solving the linear programs exactly (and then using the heuristic of Subsection 3.4.2). These facts show that, by using special algorithms and supercomputers, one may well be able to assign ground holds, on a real time basis, for the whole network of the major U.S. congested airports. Using supercomputers is not unrealistic, given the importance of the problem, explained in Chapter 1.

The tractability of large-scale GHPs is due to the simplicity and the *compactness* of the formulations. The number of constraints and the number of variables are typically small linear multiples of the number of flights. In sum, the method presented in this thesis is both flexible and efficient.

Concerning now the cost savings that can be achieved by applying the formulations, these depend primarily not on the formulations themselves, but rather on the real-world data and on what the alternative ways of assigning ground holds are. Nevertheless, the results presented in Chapters 3 and 5 are encouraging. For a variety of cases, optimal ground-holding policies result in costs significantly lower than the costs of FCFS heuristics that may approximate, to some extent, current ground-holding practice. No one pretends, of course, that these simplified cases adequately model every aspect of the true ATC environment. However, the computational results are sufficiently encouraging to suggest that the problem deserves further research, with a view towards applying the algorithms in a realistic simulation environment and, if the results continue to be positive, eventually introducing them to ATC practice.

Concerning, finally, the insights obtained on the behaviour of the multi-airport GHP, the reader is referred to the detailed summaries of Sections 3.5 and 5.3.

6.2 Directions for future research.

The following suggestions for future research on the multi-airport GHP arise naturally from the remarks of the previous section.

(1) The real-world GHP should be further investigated in order to find out whether any interesting aspects have been left out in the formulations of this thesis. If such aspects are found, it is possible that they could be incorporated in the formulations by means of additional constraints or other modifications. As an example, one could try to extend the formulations in order to take into account that delays sometimes occur not only because of limited airport capacities but also because of congestion in terminal area and *en route* airspace.

(2) The optimal ground-holding policies arising from the formulations should be checked for acceptability on non-technical grounds. For instance, it was found in Subsection 3.1.3 that optimal solutions discriminate against noncontinued flights. (It was also shown how the formulations could be modified to avoid such discrimination.) Other problems of this type should be checked for, and, if found to exist, they should be eliminated by appropriate modifications of the formulations.

(3) In order to further reduce running time and to make it feasible to assign ground holds on a real-time basis even for a very large network of airports, special-purpose algorithms for solving the formulations should be sought. Such algorithms could take into account the special structure of the constraint matrix: as pointed out in Subsection 3.1.1, if the coupling constraints did not exist, the constraint matrix would be totally unimodular.

(4) Reduction of computation times could also be achieved by looking for heuristics that would provide good feasible solutions of the formulations, rather than insisting on optimal solutions.

(5) Finally, the most important task for the future is to apply in practice the ground holds arising from the formulations, at first on a limited scale or in a simulation

environment, in order to obtain realistic feedback which would then spur further research.

We conclude by expressing the earnest wish that the work presented in this thesis will ultimately find its way into actual practice and will result in a more efficient way of operating the air travel network. This wish is based on the philosophical conviction that real-time dynamic control is essential to the efficient functioning of any sufficiently complex system. Any student of Biology and Physiology cannot help being impressed by the way in which these most complex of machines, the living beings, operate: they are constantly in a state of dynamic equilibrium, with hundreds of thousands of molecules being broken down and resynthesized at every second. No matter how many new airports or runways are built, it is essential that optimal real-time dynamic control become an integral part of the air traffic system.

Bibliography

- [1] G. Andreatta and G. Romanin-Jacur. Aircraft flow management under congestion. *Transportation Science*, 21:269–288, 1987.
- [2] Transportation Research Board. *Winds of Change: Domestic Air Transport since Deregulation*. National Research Council, Washington, D.C., 1991.
- [3] Amedeo R. Odoni. *The Flow Management Problem in Air Traffic Control*. In A. R. Odoni, L. Bianco, G. Szego (eds), *Flow Control of Congested Networks*, Springer-Verlag, Berlin, 1987.
- [4] Octavio Richetta. *Ground Holding Strategies for Air Traffic Control under Uncertainty*. Ph.D. Thesis, Massachusetts Institute of Technology, June 1991.
- [5] Mostafa Terrab. *Ground Holding Strategies for Air Traffic Control*. Ph.D. Thesis, Massachusetts Institute of Technology, February 1990.

