Università degli Studi di Trieste

DIPARTIMENTO DI INGEGNERIA E ARCHITETTURA
XVIII Ciclo del Dottorato di Ricerca in Ingegneria e Architettura
Indirizzo ingegneria Civile e Architettura

Models and algorithms for an efficient market-driven management of European airspace demand

Settore scientifico/disciplinare
ICAR/05 TRASPORTI

Supervisore di Tesi:
Chiar.mo Prof. Lorenzo Castelli
Coordinatore:
Chiar.ma Prof.ssa Paola di Biagi

Dottoranda:
Desirée Rigonat

Anno Accademico 2014 - 2015

UNIVERSITÀ DEGLI STUDI DI TRIESTE

XVIII CICLO DEL DOTTORATO DI RICERCA IN

INGEGNERIA E ARCHITETTURA
INDIRIZZO INGEGNERIA CIVILE E ARCHITETTURA

Models and algorithms for an efficient market-driven management of European airspace demand

Settore scientifico-disciplinare: **ICAR/05 TRASPORTI**

DOTTORANDA
Desirée Rigonat

COORDINATORE
Prof.ssa Paola Di Biagi Paola Di Biagi

SUPERVISORE DI TESI
Prof. Lorenzo Castelli Lorenzo Castelli

ANNO ACCADEMICO 2014 / 2015

Alla mia famiglia

Acknowledgements

This work is co-financed by EUROCONTROL, acting on behalf of the SESAR Joint Undertaking (the SJU) and the European Union as part of Work Package E in the SESAR Programme under project E.02.33 SATURN (Strategic Allocation of Traffic Using Redistribution in the Network). Opinions expressed in this work reflect the authors views only and EUROCONTROL and/or the SJU shall not be considered liable for them or for any use that may be made of the information contained herein.

Esposizione riassuntiva

La gestione del traffico aereo in Europa (Air Traffic Management, ATM) è un' attività complessa, poiché molti attori e componenti vi interagiscono per garantire l'efficienza e la sicurezza delle operazioni. Negli ultimi dieci anni questo sistema ha iniziato a mostrare segni di debolezza in termini di capacità di sostenere in modo efficiente il traffico in costante crescita. Dal 2012, otto Centri di Controllo (Area Control Centre, ACC) hanno riferito sistematicamente gravi ritardi causati da carenza di capacità dello spazio aereo. Tali carenze sono evidenti anche analizzando la distribuzione del traffico attuale e dal momento che le previsioni a lungo termine di crescita del traffico Europeo prospettano un inccremento del 50% dei voli nei prossimi venti anni, il livello di congestione non può che aggravarsi.

Questa tesi di dottorato è stata realizzata nell'ambito di un progetto SESAR WP-E denominato SATURN (Strategic Allocation of Traffic Using Redistribution in the Network), in il cui obiettivo era investigare sulla possibilità di mitigare questi squilibri della domanda di capacità nella fase strategica della pianificazione dei voli, cioè con mesi di anticipo dalla giornata delle operazioni, attraverso la modulazione delle tasse di sorvolo. Queste sono le tasse che le compagnie aeree devono corrispondere ai fornitori di servizi di navigazione (Air Navigation Service Provider, ANSP) per l'utilizzo del rispettivo spazio aereo. Le tasse di sorvolo sono raccolte da Eurocontrol e redistribuite agli ANSP per coprire i costi legati alla fornitura di servizi di comunicazione, navigazione e controllo.

Il lavoro si sviluppa in due parti: la prima comprende un capitolo che descrive il funzionamento del sistema ATM attuale e i criteri con cui le tasse di sorvolo vengono attualmente calcolate. Il capitolo successivo fornisce una panoramica di quali metodi di tariffazione sono applicati in settori diversi dal trasporto aereo per affrontare i problemi di congestione.

La seconda parte della tesi è di carattere sperimentale e presenta un' applicazione di un approccio centralizzato di *peak load pricing, PLP,* al sistema ATM europeo. PLP è una tecnica di tariffazione comunemente applicata a trasporti e utenze i cui picchi di domanda sono periodici e prevedibili. Nel PLP viene assegnato un prezzo più alto per il consumo del servizio quando si prevede forte domanda, e uno più basso per l'utilizzo del servizio nel tempo e/o posizioni che non sono critici. L'applicazione al sistema ATM europeo è modellata come un gioco di Stackelberg che rappresenta l'interazione tra un'autorità centrale che imposta le tasse di sorvolo per tutti gli ANSP europei, e le compagnie aeree che vogliono operare voli al costo minimo. Questa interazione è formulata attraverso un problema di ottimizzazione bilivello, presentato nel terzo capitolo. Il capitolo successivo è dedicato alla preparazione e gestione di dati reali del traffico aereo al fine di ottenere istanze di dati che possano essere utilizzate per testare il modello presentato (e anche per le prove computazionali degli altri modelli studiati in SATURN). L'ultimo capitolo è dedicato agli approcci risolutivi adottati per la risoluzione esatta del modello di peak load pricing sulle istanze di dati ottenute. Questo obiettivo ha presentato diverse sfide dal punto di vista computazionale, sia per la difficoltà intrinseca del modello bilivello, sia per le grandi dimensioni delle istanze di dati. I risultati ottenuti sono realistici nella prospettiva di un'applicazione reale e mostrano che la modulazione delle tasse di sorvolo può indurre una redistribuzione più sostenibile del traffico attraverso la scelta di rotte alternative e moderate modifiche agli orari di partenze e arrivi.

Parole chiave: gestione del traffico aereo, tasse di sorvolo, congestion pricing, peak load pricing, programmazione bilivello, ottimizzazione combinatoria, data engineering, basi di dati geografiche.

Abstract

European Air Traffic Management (ATM) is a complex system where many actors and components interact to ensure the efficiency and safety of operations. In the last ten years this system has started to show signs of weakness in terms of its ability to sustain a growing amount of traffic in an efficient way. Since 2012, eight Area Control Centres have systematically reported severe en route Air Traffic Flow Management delays due to shortages in airspace capacity. Such shortages are evident also in the current and short-term analysis of traffic distribution. Since long term forecasts of European air traffic growth expect the amount of flights to likely increase by 50% in the next twenty years, the level of congestion is only likely to become more severe.

This thesis was carried out within a SESAR WP-E project named SATURN (Strategic Allocation of Traffic Using Redistribution in the Network), which investigated the possibility of mitigating these demand-capacity imbalances at the strategic level of flight planning, that is, months in advance of the day of operations, through the modulation of en-route charges. These are the charges that airspace users have to correspond to Air Navigation Service Providers (ANSPs) when flying through the respective airspace. They are collected by Eurocontrol and redistributed to ANSPs for covering the costs of air navigation and communication service provision.

The work unfolds in two parts: the first comprises a chapter that illustrates how the current ATM system works and how airspace users are charged for air navigation services. Then an overview of how pricing methods are applied in industries other than air transport to tackle congestion issues is also provided. The second part of the thesis is experimental and presents an application of a centralised approach to peak load pricing (PLP) to European ATM. This is a common pricing technique in transports and utilities where peaks in demand are periodic and predictable; PLP assigns a higher price for the consumption of the service when high demand is expected, and a lower one for using the service in time and/or locations that are not critical. The application to European ATM is modelled as a Stackelberg game that represents the interaction between a central authority that sets peak and off peak rates for all European ANSPs and airspace users that want to operate flights at the minimum cost. This interaction is formulated into a bilevel optimisation problem, that is illustrated in the third chapter. The following one is dedicated to preparation and management of actual air traffic data in order to obtain data instances that could be used for testing the presented model (and used for testing other models within SATURN also). The final chapter is dedicated to the exact resolution approaches attempted to solve the centralised peak load pricing model on the obtained data instances, a task that presented several computational challenges due to both the intrinsic difficulty of the bilevel model and to the large size of the data instances. A large size regional example on one country, together with some further insights obtained from a metaheuristic solving approach carried out within SATURN are also provided. Obtained results are realistic with regard to the application. They show that pricing modulation can, indeed, become an effective mean for tackling demand-capacity imbalances by inducing airspace users to apply reroutings and moderate schedule displacements to the flights involved in such imbalances.

Keywords: air traffic management, en route charges, congestion pricing, peak load pricing, bilevel programming, combinatorial optimisation, data engineering, geographic databases.

Contents

List of Figures

List of Tables

Introduction

European Air Traffic Management (ATM) is a complex system where many actors and components interact to ensure the efficiency and safety of operations. In the last ten years this system has started to show signs of weakness in terms of its ability to sustain a growing amount of traffic in an efficient way. After a steady period following the 2008 economic crisis, traffic has started to grow again, and so have delays. Since 2012, eight Area Control Centres have systematically reported severe en route Air Traffic Flow Management delay, that is, more than one minute per flight for more than 30 days in a year. Capacity shortages and associated delays are evident in the current and short-term analysis of traffic distribution and since long term forecasts of European air traffic growth expect the number of flights to increase by 50% in the next twenty years, the level of congestion is only likely to become more severe.

This work investigates the possibility of mitigating these demand-capacity imbalances at the strategic level of flight planning, months in advance of the day of operations, through the modulation of en-route charges, a practice that is already allowed by European regulations, but currently not applied systematically by any Air Navigation Service Provider (ANSP). It is expected that such a modulation could induce airspace users to redistribute traffic in a more sustainable way by means of reroutings and mild schedule displacements.

The first part of the thesis is dedicated to an overview of the problem investigated and of the state of the art, both in practice and in scientific literature.

Chapter 1 describes the organisation of the European ATM system, first by introducing the relevant actors that will be the focus of the remainder of the work; then, by explaining how airspace is organised, how sectorisation works and how routes are structured. To this end, the flight planning process is also illustrated, and the charges that airlines have to pay for Air Navigation Services provision are explained: these comprise charges for navigation and communication services at aerodromes and for the en route portion of the flight. A brief introduction to the Single European Sky regulations is also given, as a legislative background. The end of the chapter presents forecasts of traffic and and en route delay growth, which provide an idea of the magnitude of the problem. To this end, a SESAR WP-E project named SATURN (Strategic Allocation of Traffic Using Redistribution in the Network) was carried out between 2013 and 2015 to investigate the impact of a strategic modulation of en route charges, through different pricing approaches, to redistribute traffic and reduce delays by mitigating demand-capacity imbalances already in the strategic phase (i.e., months in advance of the day of operations). This thesis has been carried out within this project and, where not differently stated, represents the personal contributions brought to it.

The following Chapter 2 investigates how the problem of congestion has been dealt with in other economic sectors, more specifically, in industries whose main business is the sale of services on a network. These comprise sectors that as different as telecommunications, utilities, such as electricity, gas and water, different modes of transport such as road and rail and so on. The chapter provides a classification of different pricing methods applied to four of these, specifically: data networks, electricity wholesale and distribution, road transport and rail transport, with a particular focus on when these pricing techniques were adopted with the aim of reducing imbalances between demand and available capacity. A discussion is the carried out to investigate which of these approaches could be effective also for the European ATM system. To this end, the few existing works in scientific literature dedicated to en route charges modulations are also presented and briefly discussed.

Chapter 3 opens the second, more exploratory, part of the thesis, and is dedicated to peak load pricing (PLP). This is a simplified form of congestion pricing with the fundamental assumptions that peaks in demand are predictable by occurring periodically in time and at some specific locations, and that demand has some degree of elasticity towards time and/or location of service consumption (and therefore is sensitive to its price). A PLP policy assigns a higher price where and when a peak in demand is expected, and a lower price for off-peak areas and/or times. By doing so, it is expected that part of the peak demand will redistribute to cheaper options. From the analysis carried out in Chapter 2, PLP appears to be one of the most common pricing policies applied in transports and utilities for dealing with demand-capacity imbalances, and its assumptions are compatible with the European ATM scenario. The chapter therefore illustrates the application of a peak load pricing policy for en route charges modulation in a centralised approach, where a central authority (or Central Planner, CP) is responsible for setting en-route charges on the network and airspace users (AUs) need to route each flight, based on the set charges. The interaction between CP and AUs is modelled as a Stackelber game, using the bilevel optimisation framework. The remainder of the chapter describes how to reformulate the non-linear bilevel optimisation model into a linear Mixed Integer-Linear Problem formulation that can be implemented in a standard MILP solver software. This linearised formulation is referred to as cPLP (centralised PLP).

Since the objective of the project SATURN was to test the effectiveness of different pricing models on realistic data instances representing actual traffic, Chapter 4 is entirely dedicated to data preparation and management. Data availability is a well known issue for research carried out on European air traffic and air traffic flow management. Sources are scarce, non integrated and non-publicly accessible (differently, for example, from the United States case). Due to these reasons most of the research carried out within the European context relies on randomly generated data instances that likely lack realism; this can potentially limit the reliability of the presented results with regard to actual application of the proposed solutions. Illustrating how to obtain realistic data instances from actual traffic therefore represents one of the major contributions of this work.

The chapter describes the data sources used, the corrections applied for fixing the found inconsistencies, the simplifications applied and their rationales, like the clustering process performed on aircraft types for estimating flight operational costs, and on routes to group together those flying within the same airway. The result of this cleaning and reorganisation process performed on the data is a relational database with geographic capabilities, whose structure is presented in the final section of the chapter (and, in greater detail, in Appendix A).

The following Chapter 5 illustrates the solving approaches attempted to tackle the cPLP model presented in Chapter 3 on the data instances obtained from the process described in Chapter 4 and used also for the project SATURN. These data instances present a challenge in size, since they are at least an order of magnitude larger than those commonly found in test results in scientific literature on bilevel optimisation models. The reason for this lies in the intrinsic difficulty of bilevel problems: in fact, the reformulations carried out to linearise them cause an exponential growth in the number of variables and constraints of the problems, making them hard to solve through standard optimisation techniques such as branch and bound, even for relatively small data samples. The largest instances found in literature have at most a hundred users, while the simplest data instance created for SATURN, that is, one hour of traffic on a single country, counts at least five hundred flights, since no realistic results can be obtained from a traffic sample with less than one hour of traffic. The purpose of the chapter is therefore to explore how far it is possible to go with data of this size and a complex bilevel model with some standard optimisation techniques. The chapter first investigates how to properly set some model parameters that assess the relative weight of schedule displacements, capacity violations at sectors and airports and ANSP revenues. Then, some improvements carried out on the model are described: tighter bounds provided to strengthen the formulation and simple heuristics for providing initial solutions to the branch and bound algorithm. In the following, the resolution of a regional-scale example on French airspace is illustrated. Two pricing variants of cPLP are tested on eight hours of traffic (approximately 4 700 flights) where severe congestion occurs in historical data. Results show the effectiveness of the proposed policy in terms of traffic redistribution through a relatively mild modulation of the en route charges, but the computational time to obtain these results remains high. For comparison purpose, results obtained from a multiobjective implementation of cPLP, solved through a Genetic Algorithm heuristic by a colleague within the project SATURN, are presented in the final section of the chapter. This approach allowed to solve, though not to optimality, much larger data instances, up to one day of traffic on the whole European network (30 000 flights and forty states). Most observations obtained from this larger scale test confirm those drawn from the regional example in terms of traffic redistribution through charges modulations, as well as the existence of unavoidable tradeoffs among schedule shifts and capacity constraints violations when demand is high and capacity is scarce.

Finally, some conclusions are presented to wrap up the main findings and considerations that can be drawn from this thesis and to outline some possible ideas for future work.

Part I

Background and motivations for the study

Chapter 1

Air Traffic Management in Europe

Air traffic management (ATM) is a considerably complex system, with many agents involved. The purpose of the present chapter is to introduce the aspects of European ATM that are relevant to this study, to provide a more complete picture of the problem addressed that is, the delays caused by the congestion of airspace, and the rationale behind the proposed way to alleviate it, the modulation of en-route charges. The most relevant entities that are to be considered when investigating this problem are the Network Manager, the Air Navigation Service Providers and the Airline Operators. These are introduced in Section 1.1. A brief overview of the relevant aspects of European Air Traffic Management, including structure and organisation of airspace and routes, as well as the way flight planning is currently carried out is given in Section 1.2. Note that this introduction focuses on the management of the *en-route* portion of the flight, which is why airports operations and related slot assignment, which alone are subjects as complex as the en-route related operations, fall beyond the scope of this study and are therefore omitted here. An overview of the legislative framework that is regulating the evolution of the European ATM system (EU's Single European Sky regulations) is provided in Section 1.3. Furthermore, Section 1.4 illustrates traffic load and ATC-capacity-related delays in the current European ATM system with some figures on expected traffic growth in the short-to-medium and long term. Finally, an introduction to the project named SATURN (Strategic Allocation of Traffic Using Redistribution in the Network), within which the current study was developed, and whose purpose was to evaluate the impact of en-route charges modulation on the distribution of traffic as a mean of mitigating airspace congestion is provided in Section 1.5.

1.1 Relevant stakeholders

In the present section the main actors that are relevant to this study are briefly introduced, namely, Eurocontrol, Air Navigation Service Providers and Airline Operators.

1.1.1 Eurocontrol (EU designated Network Manager)

Eurocontrol is the European Organisation for the Safety of Air Navigation. It is a civil international organisation working for seamless, pan-European air traffic management. Eurocontrol currently has 40 member states (plus the European Union, which is also a member) out of the 44 belonging to the European Civil Aviation Conference (ECAC) area (see Figures 1.1 and 1.2).

Eurocontrol was originally founded in 1960 as a civil-military organisation to deal with air traffic control for civil and military users in the upper airspace of its six founding European Member States.

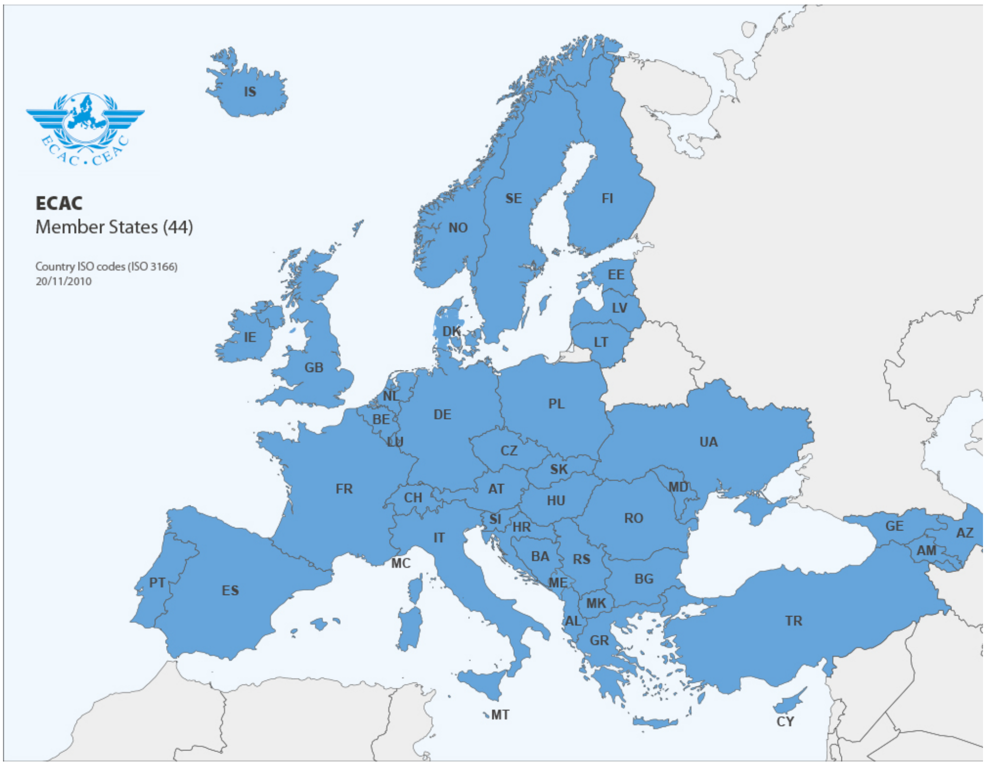


Figure 1.1: ECAC member states (source: Eurocontrol)

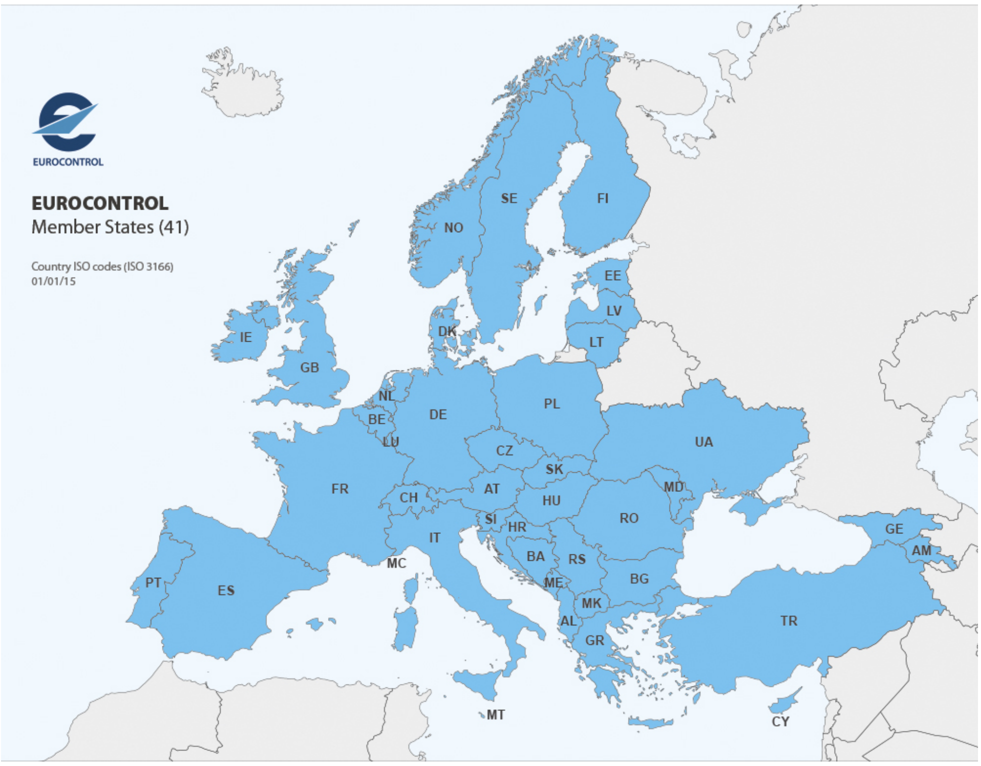


Figure 1.2: Eurocontrol member states (source: Eurocontrol)

Eurocontrol currently coordinates and plans air traffic control for all of Europe. This involves working with national authorities, air navigation service providers, civil and military airspace users, airports, and other organisations. Its activities involve all gate-to-gate air navigation service operations: strategic and tactical flow management, controller training, regional control of airspace, safety-proofed technologies and procedures, and collection of air navigation charges. In addition, it has been delegated parts of the Single European Sky regulations by the European Commission under the role of Network Manager.

Among the many functional units that comprise this organisation, two are particularly relevant to this study and need to be mentioned: the Central Flow Management Unit (CFMU), which manages the flows of the ECAC states to ensure the demand from the airlines does not overload the capacities offered by the infrastructure and the Central Route Charges Office (CRCO), which charges airspace users for providing ATM services on behalf of the Member States through the route charges system (illustrated in detail in Sec. 1.2.4).

1.1.2 Air Navigation Service Providers (ANSPs)

Air navigation service providers (ANSPs) are agencies that manage flight traffic at regional or national level. According to ICAO, Air Navigation Services (ANS) should be provided by independent authorities, entities or companies specifically established to operate these services, rather than by civil aviation authorities, as is sometimes the case today (Cook 2007). ANSPs are commonly administered as a government department depending on the state budget, as autonomous bodies belonging to the public sector (still property of the state but separate from it) or as fully or partially privatised agencies (see Table 1.1). In any case, each European state is responsible for providing air traffic services as a public service and has sovereignty over the airspace above its territories.

European ANSPs are mainly funded through the air navigation charges levied by Eurocontrol's CRCO from airspace users. The funding is carried out according to one of the following two methods (Eurocontrol 2015*d*):

- Full cost Recovery: route charges are calculated based on the estimated costs and traffic for that same year. An adjustment mechanism is applied to ensure that only the actual costs of the service are eventually recovered.
- Determined costs: the costs are determined by the contracting states at the level of the charging zone and are estimated prior to the beginning of each reference period (which covers from three to five years) as part of the performance plan for each calendar year.

Country	ANSP	Regulation
Spain	ENAIRE	State Agency
Germany	DFS	Government corporation
France	DSNA	State Agency
Italy	ENAV	State Agency
UK (continental)	NATS	Public-private partnership
Eurocontrol	MUAC	International organisation

Table 1.1: Type of regulation applied to the largest European ANSPs (source: Eurocontrol 2014*a*)

1.1.3 Airline Operators (AOs)

Airline Operators represent the final users of the ATM system. According to Eurocontrol, every year more than nine million Instrument Flight Rules (IFR) flights are performed by scheduled and business airlines in the ECAC region, corresponding to an average of over 28 thousands flights per day with an average duration of an hour and a half each (see Eurocontrol 2015*e*, and Table 1.2 for a summary). For the purpose of this study, only scheduled IFR flights (both passengers and cargo) are of interest and will be considered in the remainder of it. Concerning the passenger segment, there are currently over 50 airlines that move each more than one million passengers to, from and within European countries (see Table 1.3 for figures about the largest ten).

Breaking down the operational costs of airline operators in order to assess their potential sensitivity to the modulation of air navigation charges is no easy task. In fact airlines are highly heterogeneous with regards to market scope (e.g., regional carrier, international carrier, pure passenger carrier, pure cargo carrier, mixed passenger and cargo etc.) and business models (e.g., traditional full service carrier, low cost carrier, charter etc.).
A rough estimate is provided in IATA Cost Management Group's presentation from 2015 titled "ACMG best year ever" which shows an average cost structure breakdown obtained from a survey on passenger airlines carried out in 2014. The data used for the analysis is the result of the submission of 59 airlines worldwide with European airlines representing 51% of the share and 47% of the passengers carried (Eurocontrol 2015*e*). Figure 1.3 is sourced from the aforementioned presentation and shows that over 50% of the operational costs of the surveyed AOs are due to fleet fuelling, operation and maintenance. Air navigation charges on average make up for a 4.1% of the costs, a figure inferior to the one for airport charges, which sums up to 4.9%.

A more precise estimate can be inferred from the annual investor reports that airlines publish every year (these documents are freely accessible from companies' websites). As an example, and to provide a more detailed insight on the European market, the breakdown percentages of operational costs are illustrated for a full service carrier (Lufthansa Passenger), a low cost carrier (Ryanair) and a cargo carrier (Lufthansa Cargo) in Tables 1.4, 1.5 and 1.6 respectively. These three companies are each the largest carrier in Europe for the respective market segment. From Ryanair's figures it is evident how the cost of fuel and oil has a much more significant weight for low cost carriers, compared to IATA's average (46.03% vs. 33.4% for 2014). The same however also applies to charges. Route charges in particular amount to approximately 12% of the total operating cost, which is almost three times the IATA average figure of 4.1% (a similar proportion applies to airport charges as well).
As for the full service carrier, the figures appear more aligned to IATA's average values with regard to fuel and oil expenses (29.04% vs. 33.4% for 2014). The figures regarding fees on the other hand are much less aligned. Although Lufthansa does not provide separate figures for airport and route charges, the combined percentage of approximately 23% is more than twice the 9% that combining IATA's 4.9% (airport charges) and 4.1% (route charges) gives.
The scenario is again very different for the cargo carrier. The main component of the cost

Operator Segments	IFR Flights (x1000)	Avg. distance	Avg. Fuel Consumption	Avg. Flight Duration
Scheduled Airlines	8 655	652 NM (1 207 km)	7 547 kg	100 minutes (1.66 hours)
Business Aviation	600	485 NM (898 km)	1 136 kg	82 minutes (1.37 hours)
Total Scheduled and Business	9 255	641 NM (1 187 km)	7 132 kg	99 minutes (1.64 hours)

Table 1.2: IFR Flight Information per Operator Segment in 2014 (source: Eurocontrol 2015*e*)

Rank	Country	Airline/Holding	2015	2010	2005	Fleet	Destinations
1	Germany	Lufthansa Group	107.7	90.2	51.3	602	321
2	Republic of Ireland	Ryanair	101.4	72.7	33.4	319	186
3	United Kingdom/Spain	International Airlines Group	88.3	50.6	-	520	248
4	France/Netherlands	Air France-KLM	79.0	70.8	70.0	574	231
5	United Kingdom	easyJet	69.9	49.7	30.3	226	136
6	Turkey	Turkish Airlines	61.2	29.1	14.1	300	280
7	Russia	Aeroflot Group	39.4	14.1	8.1	239	189
8	Germany	Air Berlin Group	30.2	31.8	-	143	145
9	Sweden/Norway/Denmark	SAS Group	30.2	23.1	26.5	137	157
10	Norway	Norwegian Air Shuttle ASA	25.8	13.0	3.3	102	130

Table 1.3: Largest airlines in Europe by total scheduled and chartered passengers carried in millions (source: Wikipedia)

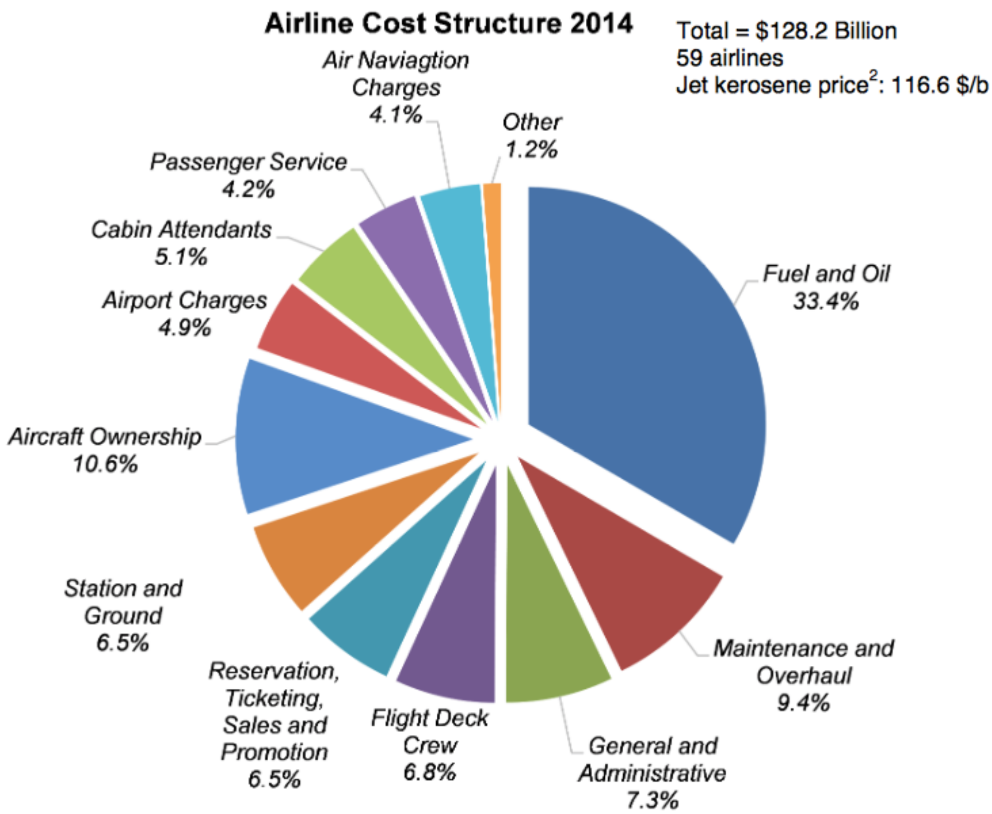


Figure 1.3: Average airline costs breakdown (source: IATA)

Ryanair	% for 2015	% for 2014 %	% for 2013 %
Fuel and oil	43.19	46.03	45.25
Airport and handling charges	15.46	14.04	14.65
Route charges	11.90	11.97	11.72
Staff costs	10.92	10.59	10.43
Depreciation	8.22	8.06	7.85
Maintenance, materials and repairs	2.94	2.65	2.93
Aircraft rentals	2.33	2.30	2.34
Marketing and distribution	5.03	4.37	4.81

Table 1.4: Operational expenses breakdown for a low cost carrier (source: Ryanair 2015)

breakdown in this case is, again quite predictably, the operating lease (since subcontracting is very common in the air cargo business), which in turn drives down the "direct" expenses for fuel. Again, only a combined figure is provided for airport and route charges, corresponding to approximately 12% of the total. Since cargo flights are mostly performed during the night because airport charges are generally lower during those hours, it can be safely assumed that in this case route charges make up for the larger part of these costs.

Lufthansa Passenger	% for 2014	% for 2013
Fuel	29.04	29.81
Fees (airport, handling and route charges)	23.12	22.25
Staff costs	19.55	19.27
Depreciation and amortisation	4.82	6.03
Maintenance and repair services	5.26	6.69
Operating lease	0.15	0.34
Other operating expenses	14.29	13.83

Table 1.5: Operational expenses breakdown for a traditional passenger carrier (source: Lufthansa 2014))

Concerning the relative share of route charges from ATM services in Europe (which cover ANSP infrastructure, staff and other operational costs) per Aircraft Operator segment, Eurocontrol provides the figures illustrated in Table 1.7 for 2014. This data is sourced from Network Manager flight plans, Eurocontrol's Central Route Charges Office (Eurocontrol 2015*e*). According to these figures, traditional, low cost and pure cargo airlines together account for 83.6% of flights and 86.3% of the charges collected.

Lufthansa Cargo	% for 2014	% for 2013
Fuel	19.89	22.86
Fees (airport, handling and route charges)	12.21	12.26
Staff costs	17.00	16.64
Depreciation and amortisation	2.81	1.16
Maintenance and repair services	5.31	5.49
Operating lease	30.83	30.92
Other operating expenses	11.95	10.68

Table 1.6: Operational expenses breakdown for a cargo carrier (source: Lufthansa 2014)

Segment	% of flights	% of KM flown	% of total charges collected
Traditional scheduled airlines	55.20	62.80	74.30
Low cost airlines	24.90	19.20	8.70
Business aviation	7.20	4.80	6.70
Charter airlines	5.40	6.10	4.70
All cargo	3.50	4.90	3.30
Military	1.30	1.30	0.90
Other	2.60	0.90	1.50

Table 1.7: Route Charges Share per Aircraft Operator Segment (source: Eurocontrol 2015*e*)

1.2 Introduction to European Air Traffic Management (ATM)

Air Traffic Management (ATM) is a component of the so-defined Air Navigation Services (ANS), which comprise, besides ATM, also communication, navigation and surveillance services (CNS), meteorological services and further auxiliary aviation services (see Figure 1.4). ATM includes all services related to air navigation, that is, Air Space Management (ASM), Air Traffic Services (ATS) and Air Traffic Flow Management (ATFM). ATS includes Air Traffic Control (ATC), flight information services and alert services (e.g., contact with rescue bodies in case an aircraft is in difficulty). Air Traffic Control services deal with the movement of traffic in the airspace controlled. Civil (commercial and private) flights are referred to as General Air Traffic (GAT), while military flights are known as Operational Air Traffic (OAT). A pilot planning a flight can choose between two different types of flight rules, namely Visual Flight Rules (VFR) and Instrument Flight Rules (IFR). Under VFR, the pilot is responsible for maintaining visual contact with other airspace users and for determining his route with the help of geographical landmarks. Under IFR on-board instruments serve as navigation aids, and together with the indications transmitted by air traffic controllers make up for the lack of visibility and provide the information required for the pilot to be aware of his position at all times (Cook 2007).

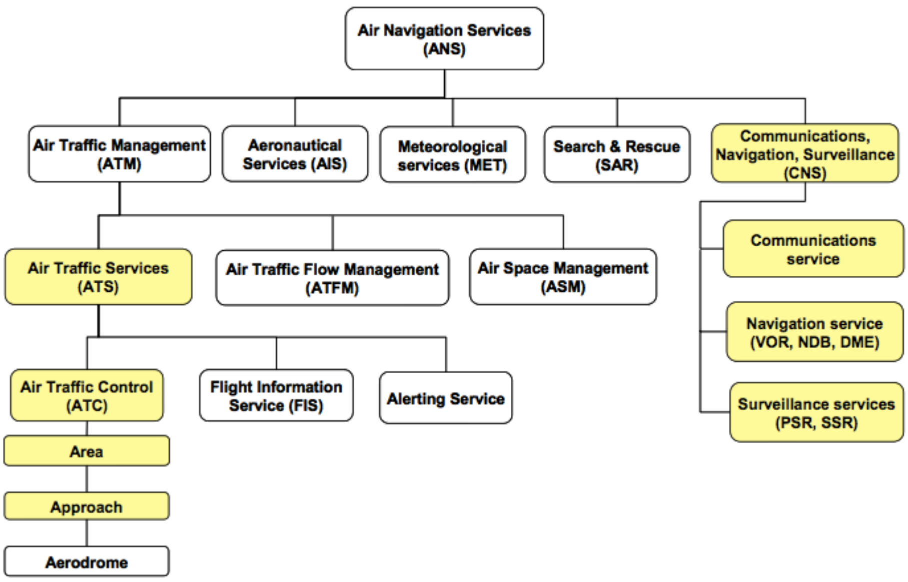


Figure 1.4: Air Navigation Services in Europe (source: Eurocontrol)

1.2.1 Airspace structure and organisation

Since each state is responsible for the airspace above its territories, Flight Information Regions (FIRs) have generally been established according to national borders, with limitations dictated mainly by military and ATC infrastructural needs. Some countries operated a further vertical split by creating an Upper Information Region (UIR). Splits typically occur at Flight Levels 195, 245 and 285 (corresponding to approximately 6, 7.5 and 8.7 Km of altitude respectively); due to a lack of harmonisation, states were initially free to choose the altitude at which to operate the split. One of the mandate of the SES regulations (see Section 1.3) is to create a European UIR. Figures 1.5 and 1.6 illustrate FIRs and UIRs in the ECAC area for lower and upper airspace respectively.

FIRs and UIRs are controlled by Area Control Centres (ACCs); currently there are around 70 ACCs in the ECAC area. According to various factors such as airspace type, quantity of traffic, military requirement and controller availability, ACCs are further divided into smaller sectors. Sectors are defined as the smallest portion of airspace under specific control and their shape and size are ultimately limited by the number of aircraft a controller can handle at any given moment. Based on this empirical figure, a capacity for the sectors is established (a typical figure for European sectors is around 30 aircraft entries per hour) and published to the airspace users and the CFMU. Since aircraft typically cross several sectors in one ACC, the capacity of the ACC is less than the sum of its individual control sectors (Cook 2007).

Sectorisations are not unique: according to controller availability and traffic forecasts an ACC can be organised according to different sector configurations, where sectors can be merged in case of lower capacity demand (for example, during the night) or split when the traffic increases (e.g., during peak hours in the morning and afternoon). ACCs can switch from one configuration to another several times per day. Figure 1.7 shows two example configurations of a German ACC.

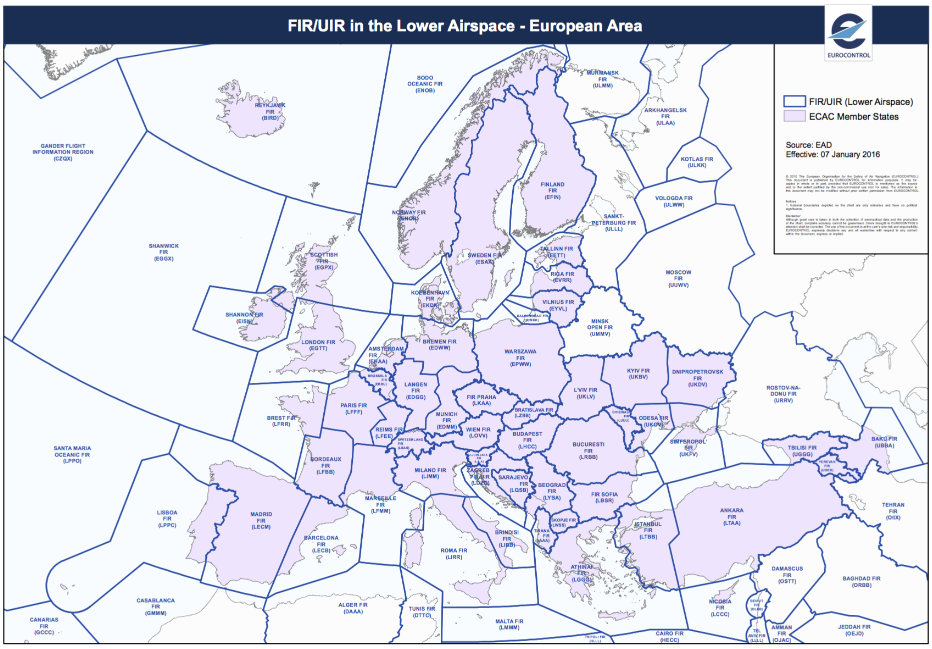


Figure 1.5: Flight Information Regions in ECAC area, lower airspace (source: Eurocontrol)

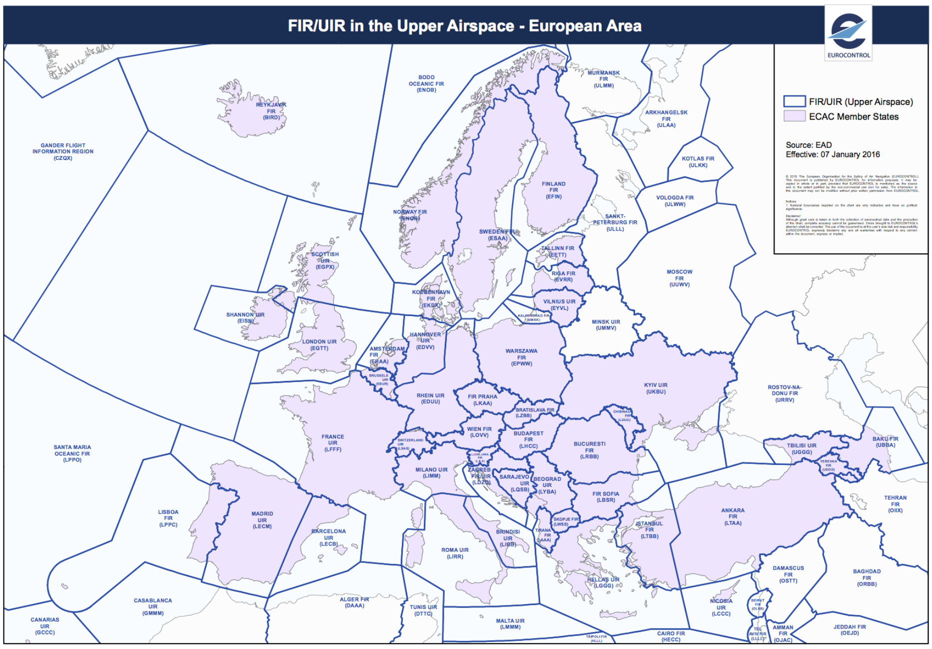


Figure 1.6: Flight Information Regions in ECAC area, upper airspace (source: Eurocontrol)

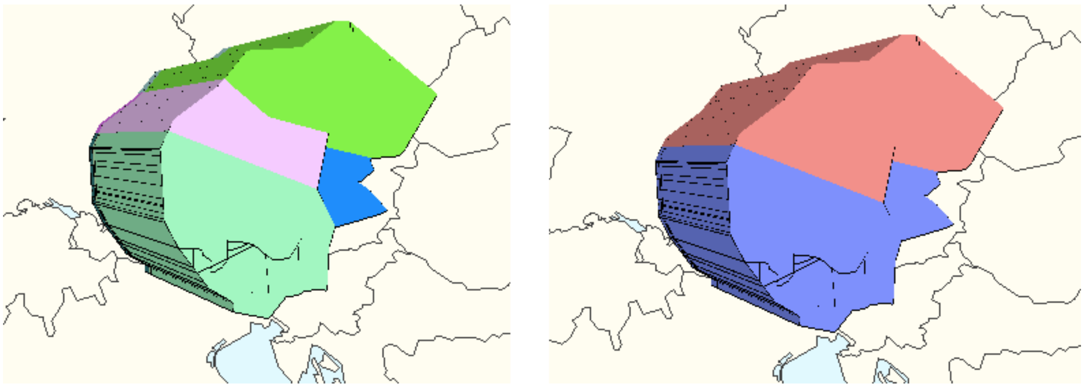


Figure 1.7: Day configuration (left) and night configuration (right) for ACC EDUUUTAS (source: NEST)

A special block of airspace is what is defined as Terminal Control Area or Terminal Manoeuvring Area (TMA), that is, a special type of airspace located above an airport, designed to handle aircraft arriving and departing the airport(s) contained within it. TMAs include SIDs and STARs (Standard Departure and Arrival Routes and procedures defined on some airports, see Sec. 1.2.2) but not aerodrome aprons. Being the most complex type of airspace, only IFR flights are allowed in them.

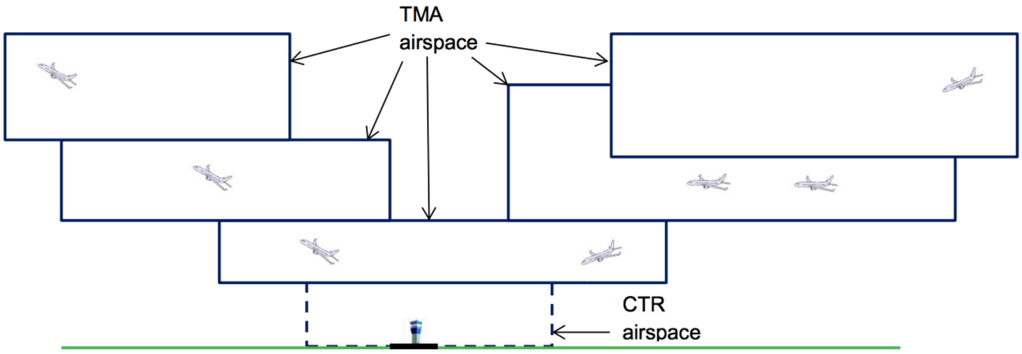


Figure 1.8: Terminal Manoeuvring Area (TMA, source: International Virtual Aviation Organization)

The SES-2 legislative package (see Sec. 1.3) introduced a further functional organisational layer of FIRs, the so-called *Functional Airspace Blocks* (FABs). FABs are defined as airspace blocks based on operational requirements and established regardless of State boundaries, where the provision of air navigation services and related functions are performance-driven and optimised with a view to introducing, in each functional airspace block, enhanced cooperation among air navigation service providers or, where appropriate, an integrated provider. Eurocontrol has been issued by the EC a mandate for supporting the establishment of FABs as part of the Network Manager tasks. The SES-2 regulation requires all EU members to be part of a FAB by 2012. As of today, nine FABs have been declared and established (see Figure 1.9). As of the latest available update (January 2014 according to Eurocontrol (2015*b*)) FIRs and UIRs are grouped in FABs as follows:

- UK-Ireland FAB (Scottish FIR&UIR, London FIR&UIR, Shannon FIR&UIR);

- Danish-Swedish FAB (Copenhagen FIR, Sweden FIR);
- Baltic FAB (Warszawa FIR, Vilnius FIR&UIR);
- BLUE MED FAB (Nicosia FIR&UIR, Athinai FIR&UIR, Brindisi FIR&UIR, Milano FIR&UIR, Roma FIR&UIR, Malta FIR&UIR);
- Danube FAB (Sofia FIR, Bucarest FIR);
- FAB CE (Zagreb FIR, Budapest FIR, Ljubljana FIR, Praha FIR, Wien FIR, Sarajevo FIR&UIR, Bratislava FIR);
- FABEC (Brussels FIR&UIR, Langen FIR, München FIR, Rhein UIR, Hannover UIR, Bremen FIR, Amsterdam FIR, Bordeaux FIR, Reims FIR, Paris FIR, France UIR, Marseille FIR, Brest FIR, Switzerland FIR, Switzerland UIR);
- North European FAB (Tallinn FIR, Finland FIR&UIR, Enor FIR, Riga FIR, Bodo Oceanic FIR);
- South West FAB (Canarias FIR&UIR, Lisboa FIR, Madrid FIR&UIR, Barcelona FIR&UIR).

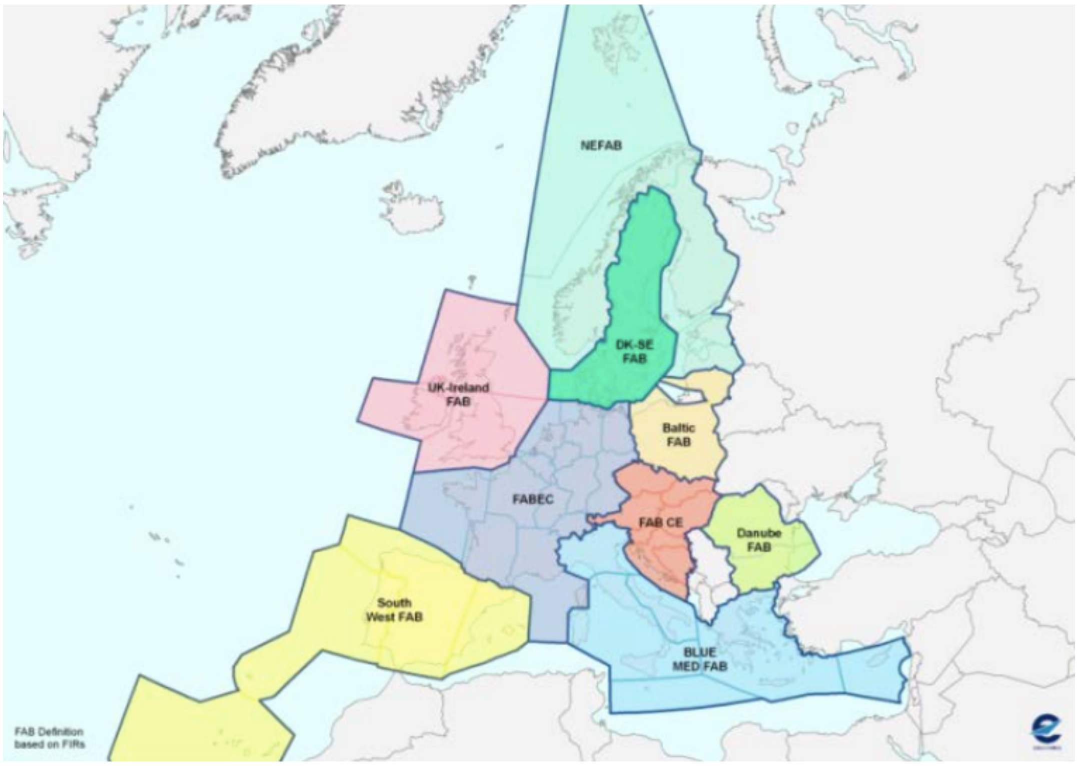


Figure 1.9: Functional Airspace Blocks (source: Eurocontrol)

1.2.2 Routes

In en-route airspace, the main routes for aircraft currently consist of airways (usually with widths of 5 NM either side of the centre-line) and Upper Air Routes (with no defined width). These routes follow straight lines between significant points commonly referred to as "waypoints" or "navpoints". Waypoints are represented by three letter identifiers (e.g., DME) when there is an associated ground-based navigation aid, for example a beacon and by five letters identifiers (e.g., KONAN) when the point is just a geographical coordinate with no ground-based navaid. In addition, airline-defined waypoints can be inserted in flight plans, for example as a temporary routing to avoid weather, or to better navigate a climb or a descent. These ad-hoc waypoints do not appear on official charts.

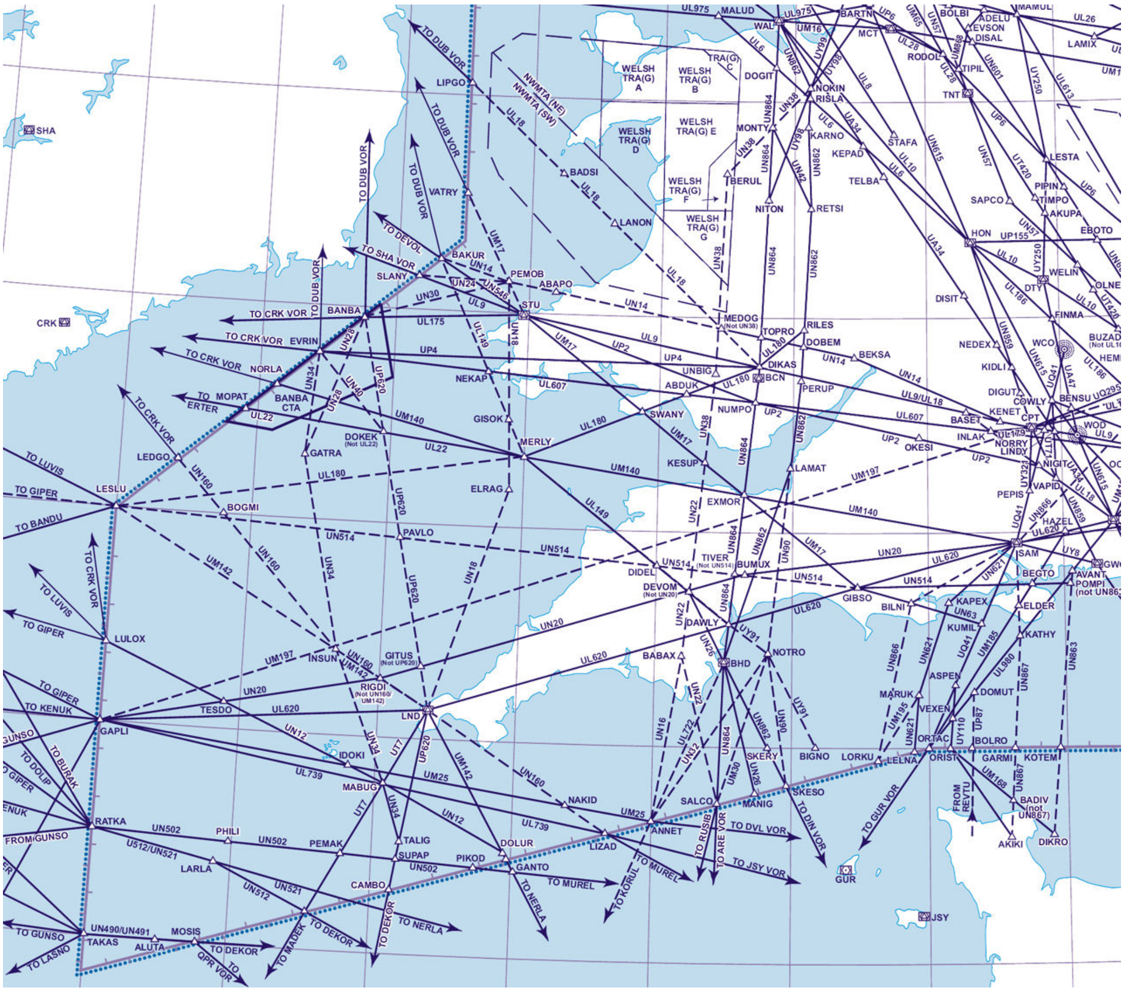


Figure 1.10: Upper air routes and waypoints in the Cardiff area (source: NATS)

To increase the capacity and flexibility of airport operations, Standard Instrument Departure Routes (SIDs) and Standard Terminal Arrival Routes (STARs) have been introduced. Both offer the pilot a pre-planned IFR procedure. SIDs promote highly specific departure procedures, which is very important nowadays with noise abatement and environmental protection being relevant factors around airports. STARs are designed to expedite ATC arrival procedures and to facilitate the transition between en-route and instrument approach segments. Each STAR procedure is presented as a separate chart and may serve a single airport or more than one (Cook 2007).

1.2.3 Air Traffic Flow Management and flight planning

In Europe, Air Traffic Flow Management (ATFM), nowadays often referred to as Air Traffic Flow and Capacity Management (ATFCM) service has been established to utilise the limited capacity available to the maximum extent possible while enabling a safe, orderly and expeditious flow of air traffic. There are three phases to ATFCM (Cook 2007):

- *Strategic Flow Management* which occurs from six months up to seven days in advance of the day of operation, when long-term demand and capacity matching will be planned (e.g., planning for extra traffic due to holidays and seasonal peaks);
- *Pre-Tactical Flow Management* which occurs in the six days before the day of operation during which the strategic plan is fine-tuned in the light of updated demand as filed flight plans are received. An ATFM Daily Plan is published in preparation of the tactical phase;
- *Tactical Flow Management* occurs on the day of operation with the ATFM Daily Plan being updated as actual traffic and capacity are known. Traffic is managed through slot allocation and re-routings.

Since 1996, Eurocontrol's Central Flow Management Unit (CFMU) controls the flight planning and message distribution in the Integrated Initial Flight Plan Processing System (IFPS) zone (IFPZ). IFPS service is split between two functionally identical units located in Haren (Brussels) and Brétigny-sur-Orge (Paris). Flight plans are sent to both sites and the workload is then shared between them. Once received by the IFPS, flight plan data is checked for syntax and semantic correctness in order for the system to build a 4-dimensional profile of every flight within the IFPZ. The window for filing flight plans to IFPS for processing is between 120 and 3 hours before the Estimated Off-Block Time (EOBT). Repetitive flight plans can be filed in advance of the current season (there are two "seasons", summer from the end of March and winter from the end of October) but full IFPS processing occurs only around 20 hours before EOBT. When a flight plan message is accepted by the system, an Acknowledgement message (ACK) is sent back to the airline operator.

The syntax for flight plan messaging is defined by ICAO in a document named *Procedures for Air Navigation Services - Air Traffic Management*, also known as PANS-ATM or "Doc 4444". Flight Plan messages contain all information needed for planning IFR flights, including aircraft type, departure and destination airports, EOBT and route as sequence of waypoints with associated altitude (expressed as the corresponding Flight Level, FL).

When entering the pre-tactical phase, flight plans are processed by the Enhanced Tactical Flow ManagementSystem (ETFMS). This system calculates demand in every airspace sector within the CFMU area using planned information (filed flight plans) received from the IFPS and calculates and allocates capacity slots by the Computer Assisted Slot Allocation (CASA, explained in the following) system. Information is continuously updated until the operational phase using surveillance radar data from ANSPs and position reporting data from aircraft operators. The availability of accurate real-time information allows ETFMS to recalculate the 4-D profile of flights and manage traffic demand in a more accurate way (Cook 2007).
Filing a flight plan to IFPS represents a request for a departure slot. The Estimated Off-Block Time (EOBT) declares the time the aircraft operator expects the aircraft to be ready to depart. Whether the aircraft is permitted to depart at this time depends on the presence of any flow restrictions (called "ATFM regulations") on the aerodromes and the airspace through which the route is planned. A flight affected by an ATFM regulation receives an "ATFM slot"

from CFMU. Slot calculations and assignments are operated according to the “first planned, first served” principle by the CASA system. This system calculates the Estimated Time Over (ETO) for the point of entry at each sector used in the planned route of a flight, and inserts that flight in these sectors’ slot lists. When the number of aircraft in a slot exceeds capacity, a regulation is activated and all flights that cannot be assigned to the scheduled ETO slot are assigned to the next available one and are therefore delayed.
After this process is completed, a Calculated Takeoff Time (CTOT) is determined for the flight, and the corresponding departure slot is assigned. When a flight is subject to multiple regulations, the delay of the most penalising one takes precedence.

1.2.4 ANS charges

The costs of air traffic management services in Europe (infrastructure, staff and other operational costs) are funded through air navigation charges. ATM services are funded through the “user pays” principle, meaning that the airspace users (i.e., the airlines) are directly charged for the operated traffic. There are different sorts of air navigation charges: route charges, terminal navigation charges, and communication charges. Air navigation services (ANS) charges are imposed to recover the cost of providing ANS services in three phases of flight: movements at and around airports (aerodrome control), approach and departure of flights, including initial climb and descent (approach control) and en-route phase (en-route/area control).
Air navigation services provision in Eurocontrol member states is based on the principle of recovering a *a-priori* determined costs for a reference period (the so-called determined cost system, mandatory for SES signatories) or the full cost recovery system (applied by the nine Eurocontrol member states which are not SES signatories).

Route charges represent the remuneration for the costs of en-route ANS provision, including Eurocontrol costs. There is a harmonised route charging system in the Eurocontrol area, the legal basis of which is the Multilateral Agreement relating to Route Charges. This Agreement dates back to 1981, and has set forth common policy on ANS route charges in the Eurocontrol area. Regulation EC 1794/2006 (European Commission 2006), laid down a common charging scheme for ANS services and introduced the notion of *charging zones*. An “en-route charging zone” is a volume of airspace for which the states establish a single cost base and a single unit rate. This “en-route charging” zone extends from the ground up to, and including, upper airspace. A Contracting State is permitted to establish a specific zone for a complex terminal area (a *Terminal Manouvering Area*, TMA, see Sec. 1.2.1), after consultation with airspace user representatives. En-route charging zones are listed in Annex 1 of Conditions of application of the route charges system and conditions of payment (Eurocontrol 2011).

A single en route charge is levied for each flight performed in the Eurocontrol airspace, irrespective of the number of Member States overflown. The system for billing and collection of route charges is common and operated by Eurocontrol’s Central Route Charging Office (CRCO). The billing of route charges is done on a monthly basis and charges income is disbursed weekly to the ANSPs. In principle, all flights are subject to route charges. However, there are several categories of flights exempted from payment in all Eurocontrol states (European Commission 2006), namely:

- Flights performed by aircraft with maximum takeoff weight inferior to 2t;
- Mixed VFR)/IFR flights in the charging zones where they are performed exclusively under VFR and where a charge is not levied for VFR flights;

- State flights performed exclusively for the transport of reigning monarchs, heads of state of government or ministers on official mission;
- search and rescue flights authorised by the appropriate competent body.

In addition, some states may exempt from payment of route charges all military flights, training flights performed for the purpose of obtaining a licence or testing equipment, circular flights, humanitarian flights and flights performed by customs or police officers. Where exemption is granted, the state concerned bears the cost which would otherwise be chargeable to the flights.

How route charges are established

The total charge per flight collected by Eurocontrol (R) equals the sum of the charges generated in the charging zones defined by states:

$$R = \sum_i r_i \tag{1.1}$$

The individual charge (r_i) is equal to the product of the unit rate (t_i) and the number of service units (s_i) in charging zone i for this flight:

$$r_i = t_i \cdot s_i \tag{1.2}$$

The number of service units, s_i, is defined as a product of the distance factor (d_i) and the weight factor (p) for a given flight:

$$s_i = d_i \cdot p \tag{1.3}$$

The distance factor, d_i, is equal to 1/100 of the great circle distance (expressed in km) between the aerodrome of departure within, or the point of entry into, the charging zone (i), and the aerodrome of first destination within, or the point of exit from, that charging zone.
The distance to be taken into account is reduced by 20 km for each take-off and for each landing within a charging zone (i). The entry and exit points are the points at which the lateral limits of the charging zone are crossed by the route described in the flight plan filed by the operator and approved by the CFMU 30 minutes prior to take-off.
The route description per flight is extracted from the flight plan filed by the operator and approved by the Central Flow Management Unit (CFMU) 30 minutes prior to take-off. This enables the CRCO to calculate the distances flown in each State's airspace (Eurocontrol 2014*b*).

The weight factor, p, is proportional to the *maximum takeoff weight* of the aircraft used and is calculated according to following formula:

$$p = \sqrt{MTOW/50} \tag{1.4}$$

where the MTOW is the maximum takeoff weight of the aircraft, expressed in metric tonnes and rounded to one decimal place.
MTOW values are calculated as explained in the following: towards the end of each calendar year, airlines are required to submit to the CRCO a declaration of the composition of their fleet (registration markings, aircraft type, version and certified MTOW). Based on this information, the CRCO calculates the weight factor based on average weight of all aircraft of a basic type. These weight factors then become valid from the 1st January of the following year.

Unit rates (t) are calculated based on required equality of total costs and revenues of an ANSP, as a ratio of cost base (forecast or determined) and forecast number of service units. The cost base includes operating costs, depreciation costs, cost of capital, and a state's share of Eurocontrol costs.

Each Eurocontrol Member State establishes the unit rate of en-route charges (basic unit rate) for the airspace within its responsibility. Each November, the enlarged Commission approves the basic unit rates for the following year.
Basic unit rates are adjusted every month if the national currency of a Member State is not the Euro. The monthly unit rate is recalculated by applying an exchange rate between the Euro and the national currency. This exchange rate is the average of the Closing Cross Rate calculated by Reuters based on the daily Bid rate, for the preceding month.

Terminal charges

Terminal charges are meant to remunerate the cost of aerodrome control services and air traffic services related to the approach and departure of aircraft within a certain distance (at present typically 20 km) of an airport. Although not being further investigated in this study, they are briefly explained here for the sake of completeness. Depending on the country, terminal charges can either be included as part of airport charges (which is the case for example in Belgium and Greece) or paid directly to the ANSP through the CRCO, similarly to the route charges (e.g., Austria, France, Finland and Germany).

Terminal charges (TC) are usually based on aircraft MTOW, but there is no unique formula according to which they are calculated. Regulation EC 1794/2006, as amended by Regulation EC 1191/2010 (European Commission 2010) proposes to harmonise them according to the following formula (that should be mandatory since January 2015):

$$TC = (MTOW/50)^{0.7} \cdot t \tag{1.5}$$

According to Implementing Regulation (EU) No 390/2013, Member States may decide not to apply this Regulation at airports with fewer than 70 000 IFR air transport movements per year. It however must be applied to the airport with the highest number of IFR movements in each Member State, regardless of its traffic volume.

1.3 Legislative framework: the Single European Sky (SES) initiative

As a response to the growth in air travel witnessed in the last two decades, the European Commission passed two Single European Sky (SES) packages to create a legislative framework for European aviation. The main objective is to reform ATM in Europe in order to cope with sustained air traffic growth and operations under the safest, most cost- and flight-efficient and environmentally friendly conditions. This implies de-fragmenting the European airspace, reducing delays, increasing safety standards and flight efficiency to reduce the aviation environmental footprint, and reducing costs related to service provision. The EC states that as a result of the SES policy, average delays for en-route air traffic flow management are now close to 0.5 min per flight, which is a remarkable achievement compared to the heavy delays that occurred in the 1990s and 2000s.

The SES legislative framework consists of four Basic EC Regulations (the SES I Package, n. 549/2004, 550/2004, 551/2004 and 552/2004) covering the provision of air navigation services, the organisation and use of airspace and the interoperability of the European Air Traffic Management Network. The four SES I regulations were revised and extended in 2009 with Regulation (EC) n. 1070/2009 aimed at increasing the overall performance of the air traffic management system in Europe (the SES II Package). SES II changed the SES focus from capacity to performance and environmental impact. The regulation is structured around four key areas, or "pillars" of the ATM reform, namely performance, single safety, new technologies and managing capacity on the ground (Figure 1.11).

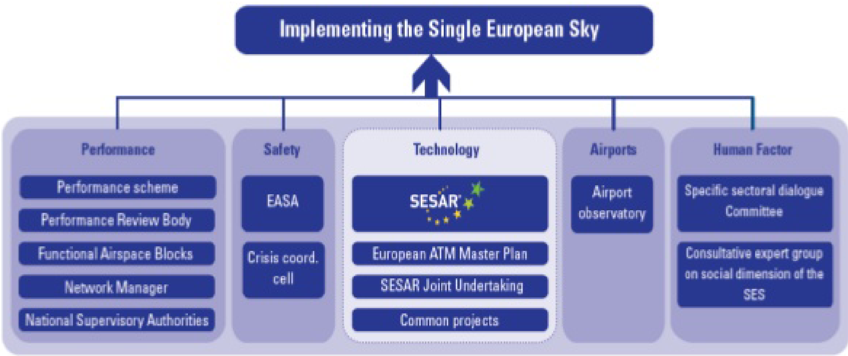


Figure 1.11: SES II pillars (source: SESAR Joint Undertaking)

1. **Regulating performance** Three measures are proposed under this pillar:

    - Driving the performance of the air traffic control system: establishment of an independent Performance Review Body (PRB) to oversee the performance of the system and proposes Communitywide targets for delays, cost reduction and the shortening of routes. These objectives are then approved by the EC and passed on to national supervisory authorities who organise consultations to agree binding national and regional objectives.
    - Facilitating the integration of service provision: establishment of Functional Airspace Blocks (FABs) by the end of 2012 at the latest; FABs are bottom-up initiatives led by the States; they aim at an enhanced cooperation between the air navigation service providers (ANSPs) and the national supervisory authorities (NSAs) to obtain operational efficiency gains through such strategies as common procurement, training and optimisation of air traffic controllers (ATCs) resources. More generally, FABs aim at better coordinating the airspace by arranging it around traffic flows instead of state boundaries.
    - Strengthening the network management function: designation of a European Network Manager to supervise tasks such as European route network design, slot coordination and allocation and management of the deployment of the Single European Sky ATM Research (SESAR) technologies (to be carried out by different actors). This function has been entrusted to Eurocontrol up to 2019.

2. **Single safety framework** The growth in air traffic, the congestion of air space and aerodromes, as well as the use of new technologies justifies a common approach to the development and application of harmonised regulation in order to improve safety levels in

air transport. Following this approach the EC proposes to extend the competence of the European Aviation Safety Agency (EASA) to the remaining key safety fields: aerodromes, air traffic management and air navigation services.

3. **New technologies** The present air traffic control system is being pushed to its limits, working with obsolescent technologies and suffering from fragmentation. SESAR (Single European Sky ATM Research) is the technological pillar of the Single European Sky. It aims to improve ATM performance by modernising and harmonising ATM systems through the definition, development, validation and deployment of innovative technological and operational ATM solutions. These innovative solutions constitute what is known as the SESAR concept of operations. This concept is defined in the European ATM Master Plan, which also defines the operational changes that needed and a roadmap for their implementation; the components of the concept are developed and validated by the SESAR Joint Undertaking (SJU) . The validated essential operational changes are deployed through Common Projects supported by dedicated SESAR deployment governance and incentive mechanisms. All three of these processes (definition, development and deployment) are components of a virtual lifecycle that actively involves the stakeholders and the Commission in different forms of partnerships.

   The SESAR ATM Master Plan has these goals for 2020:

   - improved safety performance by a factor of 10;
   - enabled a three-fold increase in capacity;
   - reduced delays both on the ground and in the air;
   - lowered the cost of air navigation service provision by more than half; n provided a 10% reduction in the environmental impact per flight.

4. **Managing capacity on the ground** Establishment of an Airport Observatory, composed of Member States, relevant authorities and stakeholders, to exchange and monitor data and information on airport capacity as a whole, as well as to provide advice on the development and implementation of Community transport legislation.

Finally, the European Commission proposed an interim update of the SES rules, called Single European Sky 2+ (SES2+). The SES2+ proposal was made in June 2013 and is currently in the process for approval by the European Parliament and Council. The update focuses on seven main areas, namely:

Independence and resources of National Supervisory Authorities (NSAs) Since audits have shown deficiencies in the oversight of Air Navigation Service Providers (ANSPs), SES2+ seeks to improve the system by providing the NSAs with more support through EU-level co-operation and by pooling resources. It also requires more separation of NSAs from the ANSPs they supervise in order to ensure truly independent oversight, such as is found for example in the oversight of airlines.

Support services Traditionally, all ATM services have been bundled into one monopoly provider and designated without use of public procurement rules. SES2+ proposed the application of normal procurement rules to ensure a transparent selection of the provider offering the best cost/benefit ratio.

Performance scheme and the Performance Review Body The performance scheme requires some updates to avoid dilution of targets. The Performance Review Body (PRB) will also be given more independence.

Functional Airspace Blocks FABs have so far been relatively inflexible constructions, focused too much on organisational structures instead of operational benefits. SES2+ aims at making them more flexible, industry led, and more focused on performance. As long as the performance targets are met, the Commission will not try to micromanage the FABs, but rather lets the industry and States devise their own solutions.

Network Manager The network manager mainly provides services to other service providers.

EASA, Eurocontrol and the institutional landscape SES2+ aims at clearing some responsibility overlaps occurred in the past decade by dividing work between the three European-level organisations (EASA, Commission and Eurocontrol) so that Eurocontrol will focus on the operational issues (network manager), EASA on technical rule drafting and oversight authority tasks, and the Commission on economic regulation (performance, charging, institutional issues).

1.4 Air traffic today, congestion of airspace and growth forecasts

In 2014 the European ATM system controlled on average 26 800 flights daily; after the decrease between 2011 and 2013, flights in Europe increased again by 1.7% in 2014 with a positive medium term outlook. According to the latest Eurocontrol Seven Year Forecast (Eurocontrol 2015*b*), flights are expected to grow by 1.5% in 2015 and to continue with an annual average growth rate of 2.5% between 2014 and 2021. Currently, European ATM costs an additional € 2-3 billion every year, compared to other similar systems in the world, due to the inefficiencies in controlling a highly fragmented airspace with an uneven distribution of traffic. In fact five biggest ANSPs (DFS for Germany, DSNA for France, ENAIRE for Spain, ENAV for Italy and NATS for the UK) bear 60% of total European gate-to-gate service provision costs and operate 54% of European traffic.

According to Eurocontrol's Performance Review Report for the year 2014 (Eurocontrol 2015*c*), after a steady improvement between 2010 and 2013, when the lowest level of enroute ATFM delay per flight on record was reached (0.53 minutes per flight), enroute ATFM delays in the Eurocontrol area increased again to 0.61 minutes per flight in 2014 (as shown in Figure 1.12). ATC capacity and staffing related delays increased in 2014. They remain, by far, the main driver of enroute ATFM delays, followed by weather and "ATC Other" which comprises, *inter alia*, ATC industrial actions. The number of flights affected by ATFM enroute delays increased again from 2.7% to 3.2% in 2014. Overall, 1.6% of flights were delayed by more than 15 minutes due to ATFM regulations, compared to 1.3% in 2013.

The performance deterioration in 2014 was therefore mainly attributed to ATC capacity issues: while capacity constraints can occur from time to time, some area control centres generated high delays on a regular basis. In 2014, the most constraining ACCs were Nicosia (due to failure to implement capacity plans), Warsaw (which had issues with the implementation of a new integrated system for ATC), Lisbon (reoccurring issues in November), Canarias (capacity planning and delay classification issues), Athinai/Macedonia (which had insufficient capacity in summer), Reims (capacity planning issues), Brest and Marseille (due to ATC industrial action). These together accounted for more than half of all European enroute ATFM delays (54.6%) but only for 17.8% of total flight hours controlled. Figure 1.13 shows the delay performance of these ACCs in terms of number of days with significant enroute ATFM delay (more than 1 minute per flight for more than 30 days in a year).

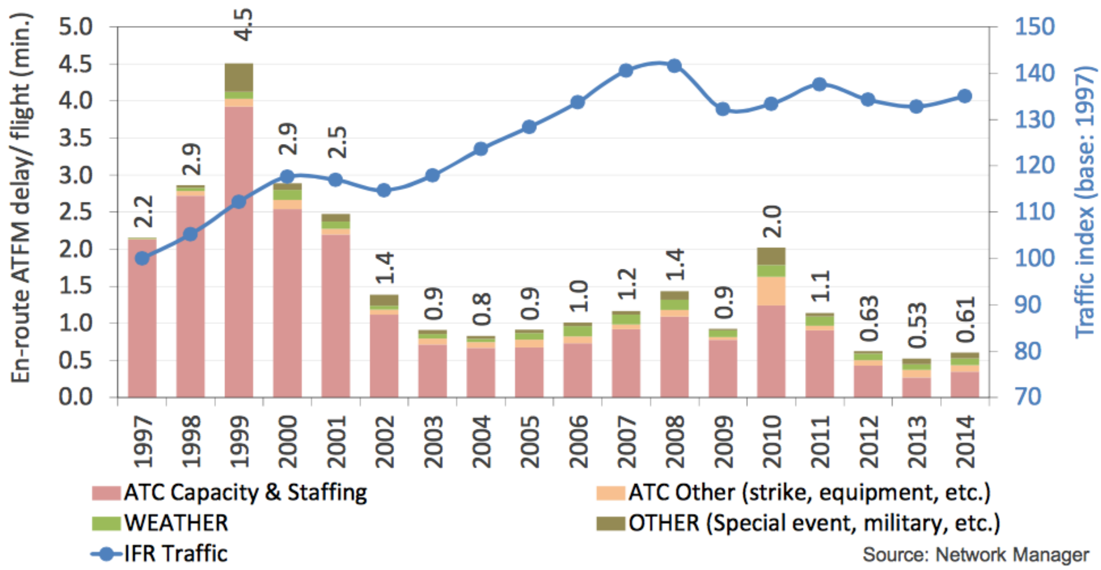


Figure 1.12: Average en-route ATFM delay per flight in the Eurocontrol area between 1997 and 2014 (source: Eurocontrol)

	En-route ATFM delay									Traffic demand		
Most constraining ACCs in 2014	Days en-route ATFM >1 min.	En-route delay /flight (min.)	% of flights delayed >15 min.	En-route delay ('000)	ATC Capacity & Staffing	ATC Other	Weather	Other (special event, military)	% of total en-route delay	Traffic growth vs 2013 (%)	5 Year Annual average growth rate (09-14)	% of total flight hours 2014
Nicosia	191	1.91	5.3%	581	68%	3%	1%	28%	**9.9%**	9.7%	2.7%	**1.0%**
Warsaw	104	0.84	2.5%	547	68%	0%	8%	24%	**9.3%**	1.2%	4.3%	**2.4%**
Lisbon	61	0.53	1.6%	240	96%	2%	2%	0%	**4.1%**	6.9%	3.5%	**2.2%**
Canarias	44	0.42	1.3%	118	76%	0%	23%	1%	**2.0%**	6.9%	1.2%	**1.2%**
Brest	43	0.53	1.2%	496	46%	48%	2%	4%	**8.4%**	4.2%	2.6%	**3.4%**
Athinai+Macedonia	43	0.42	1.3%	275	94%	4%	2%	0%	**4.7%**	8.4%	1.0%	**3.1%**
Reims	40	0.42	1.2%	387	78%	7%	12%	3%	**6.6%**	3.8%	3.0%	**1.8%**
Marseille AC	34	0.57	1.2%	568	33%	59%	6%	2%	**9.6%**	-0.6%	0.3%	**2.7%**

Figure 1.13: Overview of most constraining ACCs in 2014 (source: Eurocontrol)

Capacity shortages and associated delays are evident in the current and short-term analysis of traffic distribution (Sec. 1.4.1) and since long term forecasts of traffic growth (Sec. 1.4.2) expect traffic to likely grow by 50% in the next twenty years (Eurocontrol 2014*d*), the level of congestion is only likely to become more severe.

1.4.1 Short term traffic growth forecasts: 2014-2019

The Network Operations Plan (NOP) is a document developed by Eurocontrol in the context of the Network Management Functions. The latest available version of the Plan (Eurocontrol 2014*d*) covers the activities planned and required to enhance European network operational performance over the period 2014-2018/19, and will be updated on a yearly basis or as deemed necessary. This document addresses the requirements expressed in the EC Regulation No 677/2011 of 7 July 2011 laying down detailed rules for the implementation of air traffic management (ATM) network functions and amending Regulation (EU) No 691/2010, which establishes several ATM network functions to be performed by a Network Manager.

The scope of the NOP covers airspace structure enhancement, ACC capacity enhancement, Air

Traffic Flow and Capacity Management, Airport efficiency enhancement and network integration. The Network Operations Plan 2014-2018/19 includes also the main projects resulting from Functional Airspace Blocks developments. It also gives details of capacity and flight efficiency enhancement measures planned at network level and by each Area Control Centre, as well as a description of the airport performance assessment and improvement measures that are planned at those airports that generate a high level of delay.

Additionally, this document presents forecasts of traffic growth and en-route delay for the reference period. These forecasts are based on the data presented in another document, Eurocontrol's Seven Year Forecast. The latest available version of this document was published in February 2014, covering the period 2014-2020. It indicates that, over the seven year period, the average annual traffic growth in Europe is expected to be between 1.1% and 3.6% with a most likely value of 2.5% per year. Traffic is expected to pass the previous peak of 2008 in 2016. Figure 1.14 illustrates the expected traffic growth at ACC level by 2019 compared to 2013 levels. Turkey is expected to be the main driver of traffic growth, followed by Ukraine and, to a lesser extent, by the Eastern European area.

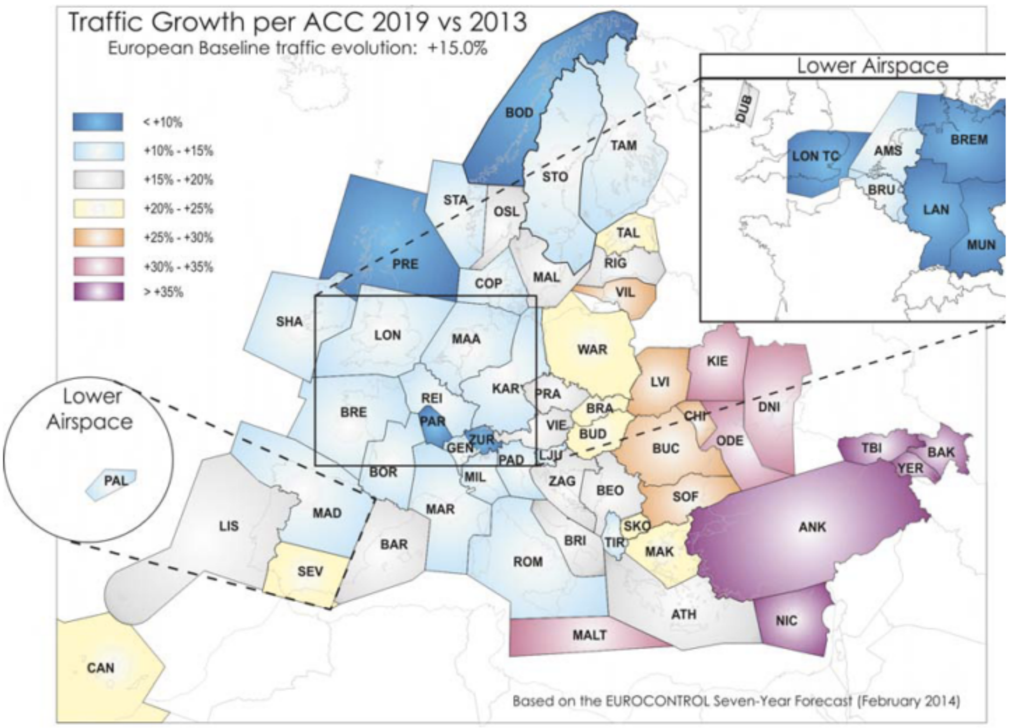


Figure 1.14: Forecast traffic growth for 2019 compared to 2013 (source: Eurocontrol)

According to the NOP (Eurocontrol 2014*d*), at European network level, the delay forecast based on the baseline traffic growth scenario for the period 2014-2019 shows that en-route ATFM delay will be above the target, set at 0.5 minutes per flight, for each year of the period 2014-2019 (see Table 1.8). The delay forecast is based on capacity plans agreed with all ANSPs during the period November 2013 April 2014, and on the baseline scenario of the 2014 Eurocontrol Seven Year Forecast of traffic demand (Eurocontrol 2015*b*).

Year	Delay target	Delay forecast (annual mean)	Delay forecast (summer season)
2014	0.5	0.51	0.78
2015	0.5	0.55	0.81
2016	0.5	0.64	0.98
2017	0.5	0.67	1.01
2018	0.5	0.62	0.94
2019	0.5	0.57	0.81

Table 1.8: Delay forecast (min/flight) at network level for the 2014-2018/19 period (source: Eurocontrol 2014*a*)

Delays are likely to be above the target at some ACCs in Europe throughout the period, in Cyprus, France, Greece, Poland, Portugal and Spain. In most cases this is due to inflexible use of staff, shortage of qualified controllers in some areas and unresolved staff management issues, despite the plans for capacity increase formulated in response to the capacity requirement profiles 2014-2019 (illustrated in the NOP and calculated according to the September 2013 Eurocontrol Seven Year Forecast). In spite of significant efforts to ensure a more effective and flexible resource utilisation, these ACCs are expected to continue to open less-than-optimum sector configurations during peak hours.

Figures 1.15 and 1.16 are taken from the 2014 NOP and illustrate, as an example, the forecast increase in en-route delay at ACC level for 2019 compared with 2013 levels despite the planned increase in capacity availability for the same period. Capacity increases are planned on the vast majority of European territory, with the exception of Scandinavian, Italian and part of Eastern European areas, where current capacity is expected to be sufficient to meet the demand.

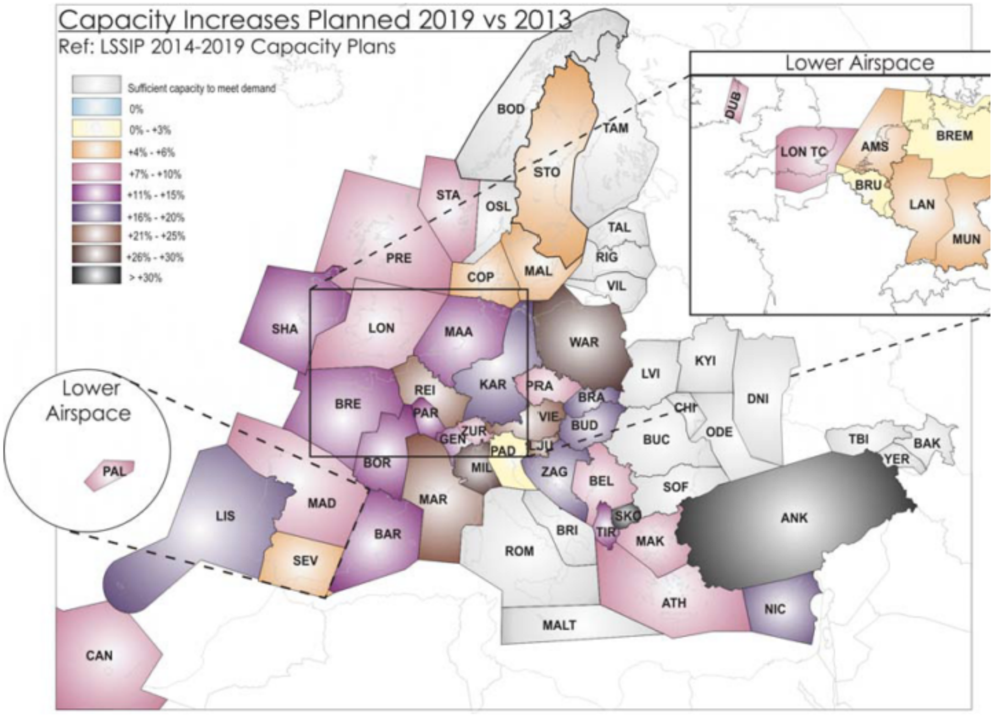


Figure 1.15: Planned capacity increases for 2019 compared to 2013 (source: Eurocontrol)

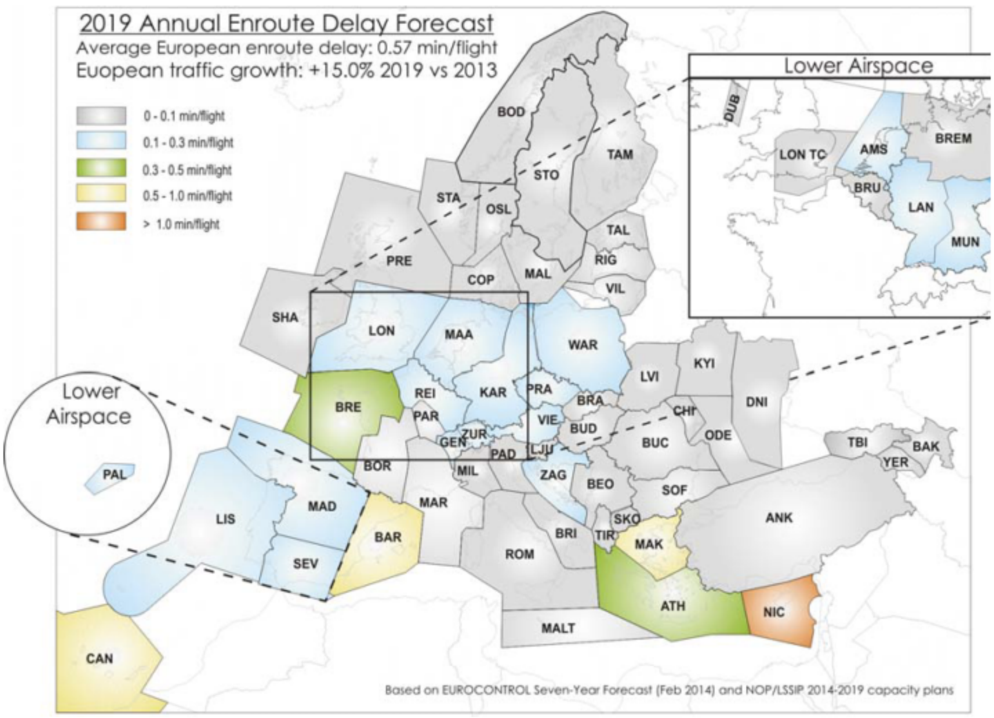


Figure 1.16: Forecast en-route delay for 2019 compared to 2013 (source: Eurocontrol)

1.4.2 Long term traffic growth forecasts: 2035

Eurocontrol's Challenges of Growth series of studies aims to deliver information to support long-term planning decisions, analysing what are the challenges of growth for commercial aviation in Europe between now and both 2035 and 2050. Four studies have been completed so far, in 2001, 2004, 2008 and 2013. The information and forecasts illustrated in the present section are sourced from the latter (Eurocontrol 2014*d*).

The five principal challenges identified for European aviation in 2035 and beyond are summarised as:

- The continuing difficulty of delivering airport capacity when, where and at the price it is needed;
- The difficulty of delivering the required level of performance on a congested network, when airport delay increases on an average busy day by a factor of 5 or 6 to become a frequent, major contributor to overall delay;
- Keeping the industry financially viable in an era of slower growth and more interesting opportunities for investors away from Europe, where aviation will be growing more quickly;
- The amount of emissions from aviation, which are likely to increase even with a slow traffic growth. Development competitively-priced low-carbon fuels may become a priority in this regard;
- Building resilience to climate change. A growing number of organisations are making resilience to climate change a routine part of their business or operational planning. However, more needs to be done to build local and network climate resilience. Some of the solutions are relatively low-cost (training and procedures), or happy side-effects of other investment. An early start should save money in the long run.

For the sake of robustness of the presented forecasts, four potential scenarios for how air transport in Europe and the factors influencing it might develop have been elaborated in the study:

- Scenario A: *Global Growth* (Technological Growth). Strong economic growth in an increasingly globalised World, with technology used successfully to mitigate the effects of sustainability challenges such as the environment or resource availability.
- Scenario C: *Regulated Growth* (Most-Likely). Moderate economic growth, with regulation reconciling the environmental, social and economic demands to address the growing global sustainability concerns. This scenario has been constructed as the "most- likely" of the four, most closely following the current trends.
- Scenario C': *Happy Localism.* With European economies being more and more fragile, increasing pressure on costs, stricter environmental constraints, air travel in Europe would adapt to new global environment but taking an inwards perspective. There would be less globalisation, more trade inside EU; also slow growth of leisure travel to outside Europe, however certainly more inside EU; more point-to- point traffic within Europe.
- Scenario D: *Fragmenting World.* A World of increasing tensions between regions, with more security threats, higher fuel prices, reduced trade and transport integration and knock-on effects of weaker economies.

Scenario	IFR Mvts (million) 2035	Traffic Multiple 2035/2012	Avg. Annual Growth 2035/2012	Extra flights/day
A: Global Growth	17.3	1.8	2.6%	21 000
C: Regulated Growth	14.4	1.5	1.8%	13 000
C': Happy Localism	13.8	1.4	1.6%	12 000
D: Fragmenting World	11.2	1.2	0.7%	5 000

Table 1.9: Summary of the key traffic values expected in the 4 scenarios for Europe by 2035 (source: Eurocontrol)

Each scenario has different input assumptions: economic growth, fuel prices, load factors, hub-and-spoke versus point-to-point etc. This leads to different volumes of traffic and different underlying patterns of growth: long- versus short-haul, rates of up-gauging of aircraft and so on.

In all four scenarios, forecast traffic to 2035 grows more slowly than historical rates (see Figure 1.17); the most-likely scenario C has 50% more flights in 2035 than in 2012. Some flows and some parts of Europe are expected to see faster growth, but overall slower growth will make it harder to deliver increasing cost-effectiveness for air traffic management.

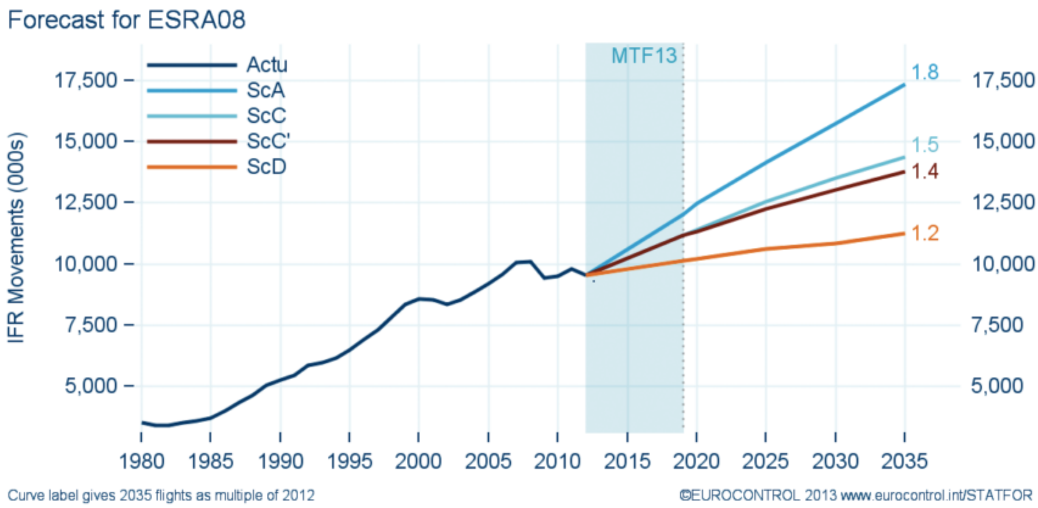


Figure 1.17: Traffic growth patterns forecast to 2035 (source: Eurocontrol)

Each scenario paints a picture of a different future, with different pattern of traffic growth. The results are shown in Figure 1.17 and Table 1.9 and are:

- Scenario A, presents the most challenging traffic situation for Europe supported by quite strong economic growth, slower fuel price growth, and a wide range of open skies agreement (compared to other scenarios). There will be 17.3 million flights in 2035 in Europe, corresponding to 1.8 times the 2012 traffic levels. The average annual growth of 2.6% is the highest of the four scenarios. However, relatively rapid growth rates (around 3.5%) during the first 8 years will then slow down to around 2%. This deceleration is explained by increasing market maturity and especially capacity constraints at airports.
- In the most-likely scenario (scenario C) there will be 14.4 million flights in Europe in 2035, 1.5 times the 2012 volume. That is an average of 1.8% increase per year, around half the historic rate from the 1960s to the peak of 2008. Traffic growth will slow down from 2025

as markets mature, economic growth decelerates and as the capacity limits at airports increasingly become an issue.

- Scenario C' follows almost the same pattern in growth as scenario C, partly because it starts at the same point. However the traffic growth develops less rapidly in scenario C' from 2020 as a result of slower economic growth, higher fuel prices and higher load factors (compared to scenario C). The growth rates slacken from 2025, resulting in a difference of 0.6 million fewer movements in scenario C' compared to scenario C in 2035.

- Scenario D starts from a low-growth forecast, in which Europe has struggled for much of the decade to get back into growth. This weak growth is compounded by high oil prices, fragile economic growth, no population migration, no free trade agreements with extra European partners, high price of travel etc an accumulation of factors that lowers the demand not only for international flights but also for intra-European ones. This scenario has just 11.2 million flights for Europe in 2035, an annual growth rate of 0.7%.

1.5 Route charges modulation as a mean to mitigate airspace congestion (project SATURN)

The possibility of using pricing mechanisms to redistribute part of the demand for airspace capacity and therefore making the expected growth in traffic more sustainable was thoroughly investigated in SESAR WP-E project SATURN (Strategic Allocation of Traffic Using Redistribution in the Network). The current study was developed within this project. The idea of modulating air navigation charges to reduce airspace congestion in Europe is not entirely new but the scientific literature on the subject is scarce (the most relevant contributions on this subject are presented in Section 2.7.2). Project SATURN represents the largest study carried out to date on this subject.

SATURN was motivated by Article 16 of EC Regulation 391/2013, which states: Member States [...] may [...] reduce the overall costs of air navigation services and increase their efficiency, in particular by modulating charges according to the level of congestion of the network in a specific area or on a specific route at specific times. [...] The modulation of charges shall not result in any overall change in revenue for the air navigation service provider [...]. This feature already gives nowadays Member States and hence, ANSPs, the opportunity to use pricing as an instrument to reduce recurring congestion problems.

The project showed how economic signals could be provided to airspace users and ANSPs to improve capacity-demand balancing by anticipating part of the flight planning process to the strategic phase, as opposed to the currently applied measures (like re-sectorisation and ATFM regulations in the tactical phase). Furthermore the benefits of having a centralised planner compared with those of decentralised maximisation of self- interests by the ANSPs and/or airspace users were also investigated.

Different pricing approaches have been investigated within the project, both monetary and non-monetary (specifically, *permit-based*); among the monetary based approaches, charges modulated according to expected traffic peaks (*peak-load pricing*) or according to the time of flight plan submission (*rewarding predictability*). These are summarised in Figure 1.18.

The effects of these three main mechanisms were tested on one full day of European air traffic, 12 September 2014 (the fourth most trafficked day of 2014 not affected by major non-ATC related regulations), including additional tests on regional instances. The results showed that the modulation of en-route charges, as advocated by EC Regulation 391/2013, is indeed a viable option to redistribute traffic in the European air network. These modulations may produce changes in the airlines' operational costs that may incentivise the rerouting of some flights, or to request different departure times, to avoid expensive areas or to take advantage of reduced charges. Results predictably also pointed out the existence of trade-offs between total delay, satisfaction of network capacity constraints and ANSP operational cost recovery constraints.

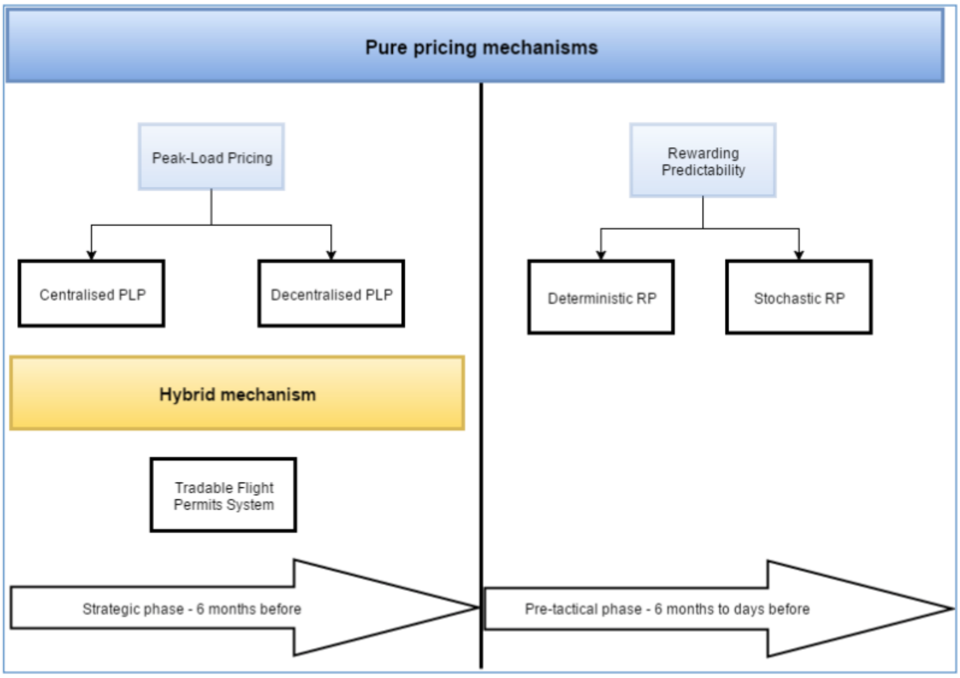


Figure 1.18: SATURN mechanisms and the time-line of their application (source: SATURN D.6.5)

Chapter 2

Pricing in network-based industries

The present chapter analyses how the issue of traffic congestion due to capacity scarceness is dealt with through market-based measures (i.e., pricing) in industries that share some common characteristics with air transport, one above all, being service-based and operating on a network. The purpose of this analysis is to evaluate which pricing measures could be successfully applied to the European ATM system, specifically as a mean for modulating en-route charges and, by doing so, obtaining a more sustainable distribution of traffic by reducing airspace congestion. In the beginning of the chapter, a brief recall on elements of economic theory about pricing that are relevant to this study is provided (Section 2.1). In Section 2.2 a set of criteria is introduced for classifying pricing techniques applied to the analysed markets according to a common framework. The subsequent sections are dedicated to illustrating pricing principles applied (or described in the scientific literature) to network-based industries other than air transport, specifically data transmission networks (Section 2.3), electricity wholesale and distribution (Section 2.4), road and rail transport (Sections 2.5 and 2.6 respectively). The remainder of the chapter (Section 1.2.4) is dedicated to European ATM; first, the current system of route charges based on country-specific unit rates is explained (Sec. 1.2.4). This charging system is then analysed according to the classification criteria defined in Section 2.2 and some considerations are drawn by comparison with the other analysed markets in order to identify which pricing principles could potentially be applied with success also to the European ATM system (Section 2.7). Finally, the proposals for alternative pricing schemes available up to date in the literature are also classified and analysed according to the common laid down framework (Section 2.7.2).

2.1 Relevant elements of theory of pricing

Network industries are generally priced according to the models and methods of the market for services. The market is generally composed of several actors, namely an infrastructure manager/owner, who may also act as a regulator, one or more service providers and consumers. This structure is more similar to a theoretic monopolistic market model when the infrastructure manager also acts as a regulator (e.g., state-owned services), or more similar to a competitive market when these entities are distinct, in which case there is usually more than one infrastructure manager or service provider (e.g., de-regulated services). In either case the objective of the regulator is generally to maximise network efficiency, the objective of service providers is maximise revenues while minimising costs and the objective of the consumers is to obtain the service at minimum cost.

When dealing with networks, efficiency and costs are generally related to the level of congestion of the network itself. Congestion is a negative externality representing an imbalance between demand and supply (i.e., available network capacity) which deteriorates the quality of the service by means of delays and hence costs, and if not properly tackled, may lead to denial of the service for some or all users.

Since users are willing to minimise their cost for using the network (in terms of time and monetary), and congestion generates externalities, marginal cost pricing is generally regarded as the way to internalise the cost of congestion. Marginal cost is the variation in the total cost that arises when the quantity produced is incremented by one unit, that is, it is the cost of producing one more unit of a good. In network terms, it represents the cost for providing an additional unit of capacity on a link or path or, equivalently, the externality caused by each user of the network.

The so-called first-best pricing principle states that a toll equal to the user's externality (a Pigovian tax) should be charged on each link in order to obtain the optimal network traffic flow configuration (see de Palma & Lindsey 2011). This scheme represents theoretical optimality, where marginal revenues (revenue obtained by the service provider from each users) equal marginal costs.

However, estimating marginal costs for actual implementation is, in general, difficult. Required information includes customers' demand elasticity and cross demand. Customers are usually very reluctant to reveal their willingness to pay (which is, in general, higher when demand elasticity is low), as it is subject to strategic behaviour. Moreover, if marginal costs are considered only in short term, they don't cover the costs of upgrading the infrastructure, potentially affecting the development of the industry (deficit). As a consequence, optimal marginal-cost pricing schemes are rarely implementable in reality and second-best pricing regimes (i.e., models that deliver a solution that is sub-optimal but workable in reality) are generally preferred for pricing actual networks.

The theory of second best (see R. G. Lipsey 1956) in general states that in a system where conditions are such that a Pareto optimum exists, if one condition is changed so that it is no longer at its optimum state (or if one or more condition are subject to uncertainty), a second best optimum (because the first best optimum cannot be reached) cannot be obtained by simply re-evaluating the uncertain condition and leaving the others to their first-best values; instead, the theory of second best states that all the other conditions must be changed from their original first best optimum states. In general, the direction and magnitude of the changes necessary are not known.

One relevant example of a second-best pricing scheme is Ramsey pricing (RP). The aim of RP is deficit coverage, obtained by means of pricing markups to be added to marginal costs proportionally to customers' elasticities; the price markup should be inverse to the price elasticity of demand: the more elastic demand for the product, the smaller the price markup. This markup, generally a percentage on the marginal costs, is inversely proportional to the price elasticity of the demand of the customers at zero profit (inverse elasticity rule). RP is in general hard to implement: since it builds upon marginal costs, and therefore faces the same information restraints.

Another very common second-best pricing scheme is Peak load pricing (PLP), typically applied in utilities and public transports. PLP is a simplified case of RP where users are charged for marginal and capacity costs in an environment where demand peaks (and therefore capacity shortages) are easy to predict. The resulting pricing scheme is usually divided in fixed time periods priced differently, namely off-peak, peak time and (possibly) shoulder periods.

A different approach is represented by Fully Distributed Costs (FDC). FDC take (estimated) marginal costs as a starting point and cover the economic deficit by allocating the remaining costs proportionally to selected parameters (e.g., utilisation by a certain user, revenue, or marginal costs themselves). Since FDC do not take demand elasticity into account, on one hand they are easier to implement than Ramsey prices but are in general Pareto-inferior (i.e., they deliver lower theoretical revenues).

Finally, it is generally argued that all types of congestion charging schemes are, in fact, discriminatory since they tend to penalise users with lower incomes. Hence, non-monetary pricing (NMP) schemes have been proposed as an alternative to first-best and second-best charging schemes in order to grant equal rights to all users. NMP schemes generally charge a certain amount of freely distributed credits or travel permits for travelling during peak times. The equity issue is then transferred to the initial endowment of credits or permits among users. An exhaustive overview of non-monetary pricing schemes proposed for road transport can be found in Fan & Jiang (2013).

2.2 Classification framework for pricing techniques

In order to better analyse and compare the pricing approaches adopted in, or proposed for, network based industries, a common framework was necessary. Based on the classifications proposed in de Groot et al. (2002) and Avlonitis & Indounas (2005), a simplified set of criteria was built for classifying pricing techniques across network industries, specifically data transmission networks, electricity generation, distribution and retail, road, air and rail transport. These criteria are illustrated in Table 2.1. Criteria 1 and 2 are market-related and can thus be discussed across the whole industry; they cover the type of control applied to that industry and the objective of the pricing strategy. Criteria 3 to 8 are specific to each pricing technique and cover the type of tariff, time or location-based modulation, presence or absence of user discrimination and other relevant characteristics. All options per criterion are mutually exclusive with the exception of criterion 8.

Environment-related	
1. Control	a. Fully centralised
	b. Fully market-based
	c. Market-based with a regulator
2. Pricing strategy objective	a. Revenue/costoriented
	b. Resource consumption oriented
	c. Both 2.a and 2.b
Pricing-related	
3. Type of tariff	a. Flat: a fixed fee gives unrestricted access to the network
	b. First-best: based on exact marginal costs i.e., users pay proportionally to the load they impose to the network
	c. Second-best: not based on exact marginal costs, i.e., average tariff for all users
	d. Multi-part: any combination of the previous
4. Modulation of the tariff	a. Time/space invariant: the network is tariffed in the same way all the time
	b. Time-dependent, space invariant: prices can vary according to time
	c. Time-invariant, space dependent: prices can vary according to location in the network
	d. Time/space dependent: the network is tariffed according to location and time
5. Users classification	a. No differentiation: all users are equal
	b. Users are differentiated: e.g., in classes
6. Price setting strategy	a. Customer-perceived value: willingness to pay determines the price
	b. Resource-estimated value
	c. Both 6.a and 6.b
7. Payment	a. Monetary
	b. Non-monetary: e.g., credits or permits
	c. Hybrid monetary/non-monetary
8. Quality of Service (QoS)	a.i. Best effort a.ii. Guaranteed service c.iii. Variable
	b. Capped service: e.g., capacity-constrained
	c. Compensation for denial of service

Table 2.1: Classification framework for pricing techniques

2.3 Data networks (telecommunications)

Up to current days, no regulation in Europe imposes separation between infrastructure owner and operator for telecommunications networks (although a process of liberalisation of telephone companies did take place across European member states between 1996 and 2006). Hence it is not uncommon to have a combination of unified infrastructure owners and service providers as well as virtual operators that rent other companies' lines and offer competing services. This holds true for both landline and mobile telecommunications. In general, the industry is configured as a free market where competitors aim at maximising revenues.

In telecommunications, and specifically in the Internet, congestion control through pricing approaches is a well-explored field of research. Since the network was designed as a best-effort service, congestion due to bandwidth demand of heavy Internet applications, such as video streaming services and Voice Over IP, seemed as an actual, inevitable problem. The currently used IPv4 protocol allows for congestion control only at the single node level and not providing any form of network-wide congestion control. Since network access is granted to most users by Internet Service Providers, each controlling its own network, the Internet gained the shape of a hierarchical interconnection of semi-isolated networks, where connecting points with other Internet Service Providers' network can indeed become bottlenecks under heavy traffic.

This problem was thoroughly investigated in scientific literature in the late 1990s and several solutions were proposed for reducing network congestion by means of an appropriate pricing policy. Up to that moment, and still nowadays, the most common pricing option for the Internet has been flat pricing. The issue that is most commonly raised with flat pricing is fairness. In fact it is generally pointed out that flat pricing is unfair since it does not allow any form of service differentiation, nor does it promote a virtuous behaviour within users by means of traffic distribution or bandwidth release when not needed. Falkner et al. (2000) give an extensive overview of literature for pricing approaches in best effort services. The most relevant are described in the following.

Flat pricing Flat pricing (e.g., Anania & Solomon 1997) is a fixed-rate price, and is typically found with best-effort services such as Internet service provided by an Internet Service Provider. The user is charged a fixed amount per time unit (e.g., month), irrespective of usage. This pricing scheme is simple and convenient. Flat pricing makes no assumptions about the underlying network technology that is already deployed. Since charges are unrelated to usage, no measurements are required for billing and accounting. This leads to social fairness in the sense that no distinction is made between poor and rich users. Provided that the flat price can be paid, anybody can access the network while receiving the same service level. However, this scheme does not allow the network to influence the users' transmission behaviour when network congestion occurs. Users are deterred from being adaptive. They have no incentive to alter their transmission behaviour to support the network operation. Flat pricing is therefore unsuitable for congestion control or traffic management.

Paris-metro pricing A Paris-Metro Pricing-based network (e.g., Odlyzko 1999) is a set of logical networks, where the total bandwidth capacity is divided into several subnetworks, each priced differently according to the offered service. Each logical network operates on a best-effort basis and is priced differently. Users choose one of these logical networks for the transmission of their traffic according to their respective budget, and this implicitly defines the service level. The main advantage of the scheme is its simplicity. Since each logical network implements flat pricing, the scheme is compliant with existing technologies. A technical disadvantage of Paris-Metro Pricing networks is their potential for instability; during periods of congestion, price-insensitive users may choose a higher-priced network in expectation of receiving better service. This may lead to congestion in the higher priced networks and thus cause instability.

Priority pricing Priority pricing (e.g., Cocchi et al. 1993, Gupta et al. 1997) can produce relative service classes of best-effort service depending on the priority class selected by the user. Under this policy users can send appropriate signals to the network to facilitate traffic management, even at short time frames. During periods of congestion, traffic is transmitted by priority level. Low-priority traffic is delayed or even dropped. The pricing scheme should deter each user from using the highest priority class as a default. Users are forced to indicate the value of their traffic by selecting a priority level. During periods of congestion the network can then carry the traffic by the indicated level. Users are able to attain a higher relative positioning (with respect to traffic from other users) by paying an increasing amount. The impact of priority pricing on Quality of Service (QoS) for a typical user and congestion is studied by Gupta et al. (1997) for the Internet. They show that if prices are calculated such that congestion is prevented, then the service level improves.

Smart-market pricing Smart-market pricing (e.g., Mackie-mason & Varian 1995) is similar in concept to priority pricing but is based on an auction mechanism for priority levels. In addition to a fixed charge to cover the connection costs and a (possibly small) charge per packet to cover the incremental cost of sending a packet, Mackie-mason & Varian (1995) introduce a usage charge when the network is congested. This charge is determined through an auction. The user associates a price with each packet, carried in the packet's header, communicating the user's willingness to pay for transmission. The network collects and sorts all the bids. It then determines a threshold value and transmits all the packets whose bid exceeds the threshold value. The threshold value is determined by the network's capacity and represents the marginal cost of congestion. Each transmitted packet is then charged this marginal congestion cost, not the value of the bid. The introduction of an auctioning mechanism leads to non-compliance with existing technologies and this complicates implementation. This schema has the advantage of encouraging both network and economic efficiency. The bids carried in each packet could also be used to facilitate routing decisions. Packets with high bids could be routed over shorter paths, whereas packets with low bids may be routed through longer paths. On the other hand however prices may deter poor users from using the service, thus creating a case for government regulation.

The second issue, beyond fairness, that arises commonly in telecommunication networks is Quality of Service. In telephony it comprises requirements on all the aspects of a connection, such as service response time, loss, signal-to-noise ratio, cross-talk, echo, interrupts, frequency response, loudness levels, and so on; in computer networks it denotes the ability of the service to provide different priority to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow. The Internet, because of its early design, does not provide any QoS mechanisms; it is in fact defined as a best effort network. The necessity of implementing QoS emerged when multimedia Internet applications began to gain momentum. A newer, improved version of the Internet Protocol, IPv6 was developed since 1996; it introduces the possibility of negotiating Quality of Service before data transmission. As of today, no more than 2% of Internet traffic flows on IPv6 but it is expected that many high bandwidth demanding application will switch to the new standard in the nearby future. A lot of research was carried out on how to exploit QoS for designing pricing mechanisms that are fairer with regard to the amount of traffic that each user generates and that are able to keep network congestion under control. An overview of pricing schemes that offer more potential for use with guaranteed QoS services is given, again, in Falkner et al. (2000). These are described in the following.

Responsive pricing Responsive pricing (e.g., Masuda & Whang 1999, Murphy & Murphy 1994) uses dynamic price-setting schemes to encourage users to adapt to congestion, using prices as the congestion- control mechanism. Similar to smart-market pricing, the charging mechanism only comes into operation during periods of congestion. Responsive pricing is based on the realisation that users are adaptive and respond to price signals. In case of high network utilisation, resources are stressed and the network increases the prices for the resources. Adaptive users then, by definition, reduce the traffic offered to the network. Similarly, in case of low network utilisation, the network decreases the price and the community of adaptive users increases their offered traffic. In this way, adaptive users do not just increase the network efficiency, but also economic efficiency.
Adaptive users fall into two classes: elastic users and inelastic users. Elastic users cannot tolerate losses but are able to delay transmission. Inelastic users require strict delay guarantees, but can tolerate some degree of losses.
The network sets the prices using either a closed-loop or a smart-market approach. In the former scheme, the network measures its resource utilisation, for example the buffer occupancy at the user-network interface, and then determines the price per packet. This ensures a desired level of network utilisation.

Edge pricing Edge pricing (e.g., Shenker et al. 1996, Keon & Anandalingam 2003) provides a conceptual contribution to shift the focus to locally computed charges based on simple expected values of congestion and route. An example of such an expected congestion cost would be time-of-day charges. Basing the charge on the expected distance between source and destination allows the charge to be applied at the edge of the network, either at the source or the destination. Prices may be communicated from any place in the network. Such a pricing scheme is much simpler and facilitates receiver payments. Any pricing scheme should be flexible enough to enable charges to be billed to either sender or receiver. This implies that prices need to be determined locally. Receiver-charging is important for multimedia applications, such as video-on-demand. It is designed to be used over short to medium time frames to modify user transmissions. It has been associated with Asynchronous Transfer Mode (ATM) and Resource Reservation Protocol (RSVP) protocols and is therefore proposed with guaranteed QoS in mind.

Expected capacity pricing Expected capacity pricing (e.g., Clark 1997) sets a long-term price based on expectations of transmissions rather than actual use. This scheme centres around the user specifying the required expected capacity. The user is then charged according to the expected capacity that the network provisions, not on actual usage, based on a long- term contract with the network. This indicates the capacity he or she expects to use when the network is congested. Expected-capacity pricing again seems to be compatible with ATM or RSVP. This implies support for congestion control or traffic management by encouraging users to determine the service level and then charging accordingly. Individual QoS guarantees can therefore be given. One of the main advantages of expected-capacity pricing is that charging is not related to actual traffic volume, but rather to expected traffic volume. Measurements are not required, saving significant overhead for the network.

Note that expected-capacity pricing should be distinguished from resource-based pricing, where the user is charged according to the measured resource usage. Such schemes are typically computationally more demanding. A resource-based pricing policy is discussed in Parris et al. (1992).

Effective bandwidth pricing Effective bandwidth pricing (Jiang & Jordan 1995, Jordan & Jiang 1995, Kelly et al. 1998) is designed to induce disclosure of traffic characteristics by the user during a call-admission process; it then charges according to a function defined in relation to the effective bandwidth of the stream of traffic. This pricing scheme ensures that the network can deduce the anticipated load generated by a user. The scheme allows the network to infer the actual (parameterised) function from the user's declaration. The user is assumed to act rationally in the sense of wanting to minimise the economic cost of the connection.

Note that the user is only required to provide the expected value of the traffic stream. Any additional information about the joint distribution of the mean rate and the peak rate is not used in the scheme, and it would thus be needless for the user to describe the distribution's higher moments, or for the net- work to require such higher moments. A disadvantage of this scheme is that the functional form of the effective bandwidth is assumed to be known in advance. This pricing scheme can also be extended to allow for time-varying prices and to deter the user from splitting the source traffic. A user could be discouraged from transmitting a single stream over several connections by imposing a connection charge.

Proportional fairness pricing Proportional fairness pricing (e.g., Gibbens & Kelly 1999, Kelly et al. 1998), concentrates on matching users' utilities to the networks available resources. The users optimally choose their prices and the network optimally assigns rates to the users.

Congestion in this scheme is avoided by allocating resources to the users. During periods of high demand, each user would get a proportionally smaller amount of bandwidth, for example, and thus users need to throttle their transmission at the edge of the network. Note that billing and measurements are not required, since each user is charged according to the indicated willingness to pay. The network only needs to keep track of the willingness to pay and allocate resources accordingly. This pricing scheme is motivated by the desire to incorporate the notion of fairness into the allocation of network resources. In proportional fairness pricing a resource allocation is fair if it is in proportion to the users willingness to pay. Note that such an allocation of resources guarantees economic efficiency, since users' utilities are maximised. It is also fair in the sense that every user who is willing to pay is allocated some bandwidth.

In recent years (from 2005 on) Internet traffic pricing and congestion issues have lost a lot of interest as research topics. This is most likely due to the fact that Internet users

since then have incremented exponentially. This might appear as non-intuitive at first, but since all new users are paying customers of some Internet Service Providers, the very same providers had larger incomes and thus were able to make bigger investments on network upgrades.
At the same time, political events in recent years have largely changed the public opinion on congestion pricing and class-based or QoS-based pricing for the Internet. The feelings for the importance of Internet neutrality today are very strong; today, policies that may inhibit access to certain services to some users because of their willingness to pay are regarded as discriminatory and potentially a mean for censorship.

Time-Dependent usage Pricing Time-Dependent usage Pricing (TDP) (e.g., Joe-Wong et al. 2011), is a pricing practice borrowed from the electricity industry that is similar to time-of-day pricing. TDP does not differentiate based on traffic type, protocol, or user class, and thus sits lower on the radar screen of network neutrality scrutiny. TDP charges a user based on not just how much bandwidth is consumed but also when it is consumed, as opposed to time-independent usage pricing (TIP), which only considers monthly consumption amounts. TDP has the potential to even out time- of-the-day fluctuations in bandwidth consumption. Given the time inelasticity of bandwidth demand in different demographics and applications, it is not clear how much TDP can reduce Internet Service Providers' costs, due to either impatient users or time-sensitive applications. Even TDP's feasibility needs examination. Research on integrating traffic measurement, optimal price determination, and user interface design is necessary for TDP to become feasible. Furthermore, it is unclear if time-dependent prices could be optimised in a computationally efficient way for near real- time control. Nevertheless, time-dependent pricing today seems to be the only approach to network congestion pricing that is both technically feasible and socially acceptable for the users.

Table 2.2 summarises the characteristics of the pricing techniques described in the current section according to the framework introduced in 2.2. Criteria 1 and 2 (control and pricing strategy objective) are omitted, being valid for the whole industry and already discussed.

Note that no pricing scheme is space (or location) dependent. This is not due to technical infeasibility but rather to the fact that routing decisions on the Internet are not necessarily shortest-route nor user-decision bound. Hence, location dependent pricing may introduce an unwelcome degree of unpredictability in the price of transmissions.

2.4 Electricity

For most of the last century, electricity was sold in regulated environments in which the retail price did not vary based on the time it was used. Customers faced a constant price for electricity regardless of the supply/demand balance in the grid. In reality, of course, the excess demand at peak times is not allowed to cause blackouts. Instead, capacity is expanded to meet the high demand that results at peak times. The question then is whether this is a good use of resources. In the real world, this inefficiency shows up in the form of excess capacity that is under-utilised, but still must be built in order to accommodate the peak demand. The value customers get out of this capacity is not great enough to justify the capital investment. With time-varying pricing of electricity, this excess capacity is not necessary because higher prices at peak times encourage customers to consume less at those times, either by shifting peak consumption to off-peak or by simply reducing consumption at peak times.

Pricing	*3. Type of tariff*	*4. Modulation of tariff*	*5. Users classification*	*6. Price setting strategy*	*7. Payment*	*8. QoS*
Flat p.	a. Flat	a. T/s inv.	a. No diff.	b. Res. Val.	a. Monetary	a. Best eff.
Paris-Metro p.	a. Flat	a. T/s inv.	b. Classes	a. Cust. Val.	a. Monetary	a. Class-dep.
Priority p.	a. Flat	b. T.dep / S.inv	b. Classes	c. Both	a. Monetary	a. Class-dep.
Smart market p.	d. Multi part	b. T.dep / S.inv	a. No diff.	c. Both	a. Monetary	a. Best eff.
Responsive p.	c. 2nd best	b. T.dep / S.inv	a. No diff.	a. Cust. Val.	a. Monetary	a. Best eff.
Edge p.	b. 1st best	b. T.dep / S.inv	a. No diff.	b. Res. Val.	a. Monetary	a. Best eff.
Expected capacity p.	c. 2nd best	b. T.dep / S.inv	b. Classes	b. Res. Val.	a. Monetary	a. Variable b. Capped
Effective bandwidth p.	b. 1st best	a. T/s inv. extendib. to b.	a. No diff.	b. Res. Val.	a. Monetary	a. Best eff.
Proportional fairness p.	b. 1st best	a. T/s inv.	b. Classes	a. Cust. Val.	a. Monetary	a. Best eff. b. Capped
Time-Dependent usage p.	c. 2nd best	b. T.dep / S.inv	a. No diff.	b. Res. Val.	a. Monetary	a. Best eff.

Table 2.2: Classification of pricing techniques in data networks

In a deregulated electricity market (such as most European countries and the US) there can be different operators that handle, and compete for, power generation, transmission, management, operations and finally retailing. The market is usually differentiated between wholesale, which include generators, transmitters, managers and impartial regulators, and retail, which includes all companies that sell electricity to home or business consumers. These two markets generally have different pricing mechanisms that may not work synchronously. One typical example is the hourly fluctuation of power cost in the wholesale market, which is not reflected in many consumer tariffs, like Uniform Pricing, where the unit cost of electricity varies two or three times per year. This can cause serious imbalance between demand and offer, that can ultimately lead to massive service disruption like state-wide blackouts.

2.4.1 Wholesale market

Similarly to telecommunications, electricity industry in the EU underwent a process of liberalisation between 1996 and 2006 (it was introduced much earlier in the US market; deregulation started in 1982 with the implementation of Public Utilities Regulatory Policies Act of 1978). Currently, each member state in Europe is responsible for the maintenance of its power grid and acts as a regulator authority (generally through a dedicated public body), and there can be competing operators for both power generation and distribution (wholesale market) and power retailing to customers. Wholesale market is therefore configured as a monopsony (multiple producers for a single customer) where competitors aim at maximising revenues, while retail is a free market.

A wholesale electricity market exists when competing generators offer their electricity output to retailers. The retailers then re-price the electricity and take it to market. While wholesale pricing used to be the exclusive domain of large retail suppliers, in recent years increasing markets are beginning to open up to end-users.
Today in most markets electricity storage costs are prohibitive and production is subject to rigid short-term capacity constraints. Since demand is highly variable, there will be times when there is plenty of capacity and the only incremental costs of producing electricity are fuel and some operating and maintenance costs. At other times, the capacity constraint will be binding, causing the incremental cost to increase greatly, and wholesale market prices to rise. (Borenstein et al. 2002). As a result the wholesale price of electricity varies continuously (generally every hour or half an hour), reflecting the supply/demand interaction. The end-use customer, however, sees the retail price, which typically is constant for months at a time and does not reflect the hour-by-hour variation in the underlying wholesale cost of electricity.
A detailed economical analysis for a self-sustainable wholesale electricity market is explained in Hogan (1998), where several aspects, from involved actors to transmission constraints and contracts are taken into consideration. The author identifies several characteristic that a market should meet in order to flourish, which can be summarised as a "coordinated spot market that has bid-based, security-constrained, economic dispatch with nodal prices". Meeus et al. (2005) give an exhaustive overview of the structure of the Internal Electricity Market in Europe. They describe regulation, market architecture, current state of development and improvements that are needed. Finally, the authors identify two possible stages for the further development of the market, requiring an increasing degree of coordination and harmonisation. Further insight on the European market, and a comparison, from the institutional point of view, with flow and capacity management in air traffic, is given by Duthaler & Finger (2011).
Differently from others network-based infrastructures, such as road transports or telecommunications, equilibrium in a power grid is system-wide, since electricity flows across all available paths from generator to consumer. Location based congestion in the electricity system may occur when specific transmission lines approach their capacity to transmit power. In order to maintain system stability, generation plants are typically re-dispatched (i.e. their power output is adjusted). Doing so ultimately increases the costs of electricity since the output of less efficient plants must be increased and that of more efficient plants reduced.
In some markets, these costs are internalised in the overall electricity price, while in others there are explicit location based congestion charges known as locational marginal pricing, LMP (Black & Larson 2007). A variant of LMP, called Continuous LMP, was proposed in Li (2007). Such charges roughly correspond to the marginal cost for generating additional electricity and are calculated by the regulator authority on a regular basis. Power generation is then allotted to generators through an auctioning mechanism. Kahn et al. (2001) compare two pricing mechanisms for the generated electricity.

Uniform Pricing Under Uniform Pricing, the price of power is set as equal to the LMP and all generators bid according to the capacity they are able to deliver. Since the cost of producing electricity rises exponentially when approaching full capacity (i.e., when approaching congestion), generators must place a bid that takes into consideration the trade-off between producing more and the increasing generation cost. When the auction is closed the winners are all paid the same unit price. This pricing technique has drawn a lot of critiques since it does not prevent generators from collaborating to withdraw power (i.e., declare they are able to produce less electricity than their actual capacity) in order for the market price to soar. Kahn et al. (2001) describe a notorious case that happened in California in 2001.

Pay-as-bid Considering the fact that the cost of electricity results from power generation and ancillary services, in a pay-as-bid pricing approach, generators are asked to bid not only for the amount of electricity they are willing to produce, but also for its price. In other words, they are asked to estimate their own marginal cost. Advocates of pay-as-bid state that since all the bids below the highest marginal cost output (necessary for the sum total of accepted bids to satisfy market demand) will under uniform pricing receive more than their bid prices. Thus the change in the rules would simply wipe out those mark-ups (Kahn et al. 2001). Detractors, on the contrary, assert that once bidders are able to predict the market-clearing price, the savings from the change in the rules for consumers would prove to be zero. Moreover, they state that small firms would be disadvantaged in such a market. In fact, under this approach competitors prosper or fail on the basis of their relative generating efficiencies alone; therefore the small firm would have to mount the same kind of effort efforts to gather the requisite information and make such forecast on a continuing hour-by-hour and day-by-day basis, with the same dimensions, as a large one (Kahn et al. 2001).

Table 2.3 summarises the characteristics of the pricing techniques described in the current section according to the framework introduced in 2.2. Criteria 1 and 2 (control and pricing strategy objective) are omitted, being valid on a market-wide scope and already discussed.

Pricing	***3. Type of tariff***	***4. Modulation of tariff***	***5. Users classification***	***6. Price setting strategy***	***7. Payment***	***8. QoS***
Uniform p.	b. 1st best	a. T/s dep.	a. No diff.	b. Res. Val.	a. Monetary	a. Guaranteed
Pay-as-bid p.	b. 1st best	a. T/s dep.	a. No diff.	c. Both	a. Monetary	a. Guaranteed

Table 2.3: Classification of pricing techniques in electricity wholesale

2.4.2 Retail market

Retail market presents several variants of the basic Time Of Usage (TOU) tariff. The difference lies in how the unit price for KW/h consumption adheres to generation cost. In fact the wholesale market has, generally, more price fluctuation than retail. Retail tariffs range from TOU (unit price adjusted one or two times a year) to Real Time Pricing (unit price adjusted every few minutes). Surcharges can be added for power consumption during peak times, incentives can be given for not consuming during peak loads, power hedging is also implementable, either with or without cash-back for unused amounts and several other variants.
Many programs have been implemented or proposed to make the economic incentives of customers more accurately reflect the time-varying wholesale cost, in order to mitigate price volatility in wholesale electricity spot markets. This approach is usually referred to as Demand-Response Pricing (Borenstein et al. 2002) or Demand-Side Management (Zarnikau 2010). Black & Larson (2007) estimate that if congestion costs are proportional to peak demand, they can be reduced by at least 10% as a result of an 11% reduction in peak demand (percentages are referred to the US market). Borenstein et al. (2002) give an overview of several such programs proposed and later implemented in the US market (deregulated across all states). However, similar pricing proposals are implemented in Europe as well.

Time of Usage (TOU) Under TOU, the retail price varies in a preset way within certain blocks of time. The rates for each time block (usually called peak, shoulder, and off-peak) are adjusted infrequently, typically only two or three times per year. As a result, the price is the same at a given time of day (on a weekday) throughout the month or season for which the prices are set. Typically, the weekend and holiday rates are equal to the off-peak weekday rate.

TOU programs set prices months in advance and therefore logically cannot capture any of the shorter-term variation in supply/demand balance and in prices in the wholesale market.

Technology plays a role in this tradeoff: in fact until recent years, the cost of TOU metering was substantially less than real-time metering and the ability to send real-time price information to customers was limited. Technology changes like electronic meters have virtually eliminated these issues.

TOU with Demand Charges Because Time Of Usage rates don't capture the price variation within a price block, TOU pricing is often combined with a separate charge for peak usage. These demand charges are a price per kilowatt for the customer's highest usage during the billing period (usually a month). Demand charges are based on the customer's maximum usage (during a short, i.e.15 minute interval) regardless of whether that usage occurs at a time when the system as a whole has a tight supply/demand balance or not. Most of the meters that register maximum usage for demand charge billing are not capable of storing information indicating the precise date and time at which that maximum usage occurred. Some criticism to Demand Charges is pointed out in Borenstein et al. (2002). They state that the economic incentives that they establish are an imperfect proxy for the real economic cost imposed on the system. First, demand charges are not synchronised to the usage on the system as a whole, so they charge as much for a peak usage that occurs at a lower demand time as at a higher demand time. As a consequence, they give no incentive for a customer to conserve until usage is near the peak level for the period. Second, Demand Charges make no adjustment for the supply side of the market, ignoring that variations in supply availability and prices can be as important as variations in demand in explaining fluctuating wholesale prices.

Interruptible Electricity Rates (IER) IER are constant nearly all of the time. When the system operator declares certain potential shortages, however, these customers are called upon to cease electricity consumption. Despite the name, service to these customers is generally not actually physically interrupted. Rather, the price that they face increases dramatically. Thus, the customer on an interruptible program retains an option to continue to consume after being instructed to discontinue usage, but at a greatly increased price. If for some reason economic incentives fail to equilibrate supply and demand for even a brief period, the system operator must have the ability to curtail usage by some customers. IER contracts offer a certain amount of insurance to customers by telling them they won't be called more than a pre-specified number of times during a year.

Critical Peak Pricing (CPP) CPP programs usually start with a TOU rate structure, but then they add one more rate that applies to critical peak hours, which the utility can call on short notice. CPP programs typically limit the utility to call no more that 50 or 100 critical peak hours per year. CPP is a clear improvement on TOU with demand charges, because the additional charges are based on consumption when the system is actually constrained, rather than when the particular customer's demand peaks. In fact the additional charges are based on consumption when the system is actually constrained, rather than when the particular customer's demand peaks.

CPP Programs are the natural evolution of demand charges when more sophisticated metering is available. Charges increase at critical system peaks rather than at the individual customer's demand peak, which is much more consistent with the true costs of consumption. CPP still has two economic weaknesses, though they may actually be strengths in terms of customer acceptance. First, the prices are limited and levels are pre-set for the critical peak periods, therefore they can't be calibrated to move with the actual prices in the wholesale market. Second, the number of critical peak hours that can be called in a year is limited. As a result, the utility protects customers against seeing very high prices, even only on marginal purchases, for more than a fixed number of hours.

CPP with Incentive Preserving (IP) Rebates As observed by Letzler (2010), customer resist signing up for CPP and other programs with dynamic price change, but are in general more receptive to baseline-rebate programs that create similar incentives by rewarding customers whose critical-period power use is below a baseline level. Analysing the psychology behind such behaviour, the author states that the key to success of CPP programs lies in how the dynamic pricing is presented to the customer. Incentive preserving rebates transform the presentation of CPP while preserving CPP's revenue streams and marginal incentives. IP rebates make critical events into opportunities for customers to gain by selling each customer rights to a block of power at the regular price during critical events and rebating the value of unused rights. Setting the right rebate value per kWh gives customers the right incentives to choose between using their rights (obtained by paying a monthly fee) and cashing them in.

Demand Reduction Programs (DRP) DRP pay a customer to reduce their consumption at certain times. They are activated by the system operator when grid conditions meet certain pre-determined criteria that indicate that the supply/demand balance is likely to be very tight over some ensuing period of time. The operator then offers to pay participating customers to cut back their usage.

These programs must first determine a baseline from which demand reduction can be measured. Once the baseline is set, the price offered for demand reduction determines the level of economic incentive to reduce demand when the system operator calls. The price offered is usually pre-determined and does not vary with the tightness of supply.

The fundamental weakness with demand-reduction programs is that there is no reliable baseline from which to pay for reduction. Baselines are usually set from the past behaviour of the customer, but, it has been observed (Borenstein et al. 2002), this can lead to opportunistic behaviour from the customer.

Real Time Pricing (RTP) RTP describes a system that charges different retail electricity prices for different hours of the day and for different days. Obviously, a longer lag time between the price announcement and the price implementation will result in prices that less accurately reflect the actual real-time supply/demand situation in the market (TOU is an extreme case of such scenario). On the other hand, the longer lag time means that the prices will be less volatile than the real-time wholesale electricity price but, on the other hand, lower volatility may appear more reassuring to customers. Designers of RTP programs face this trade-off between greater advanced price notification and more accurate price signals.

Under the simplest RTP system, the customer is billed the real-time price of electricity on an hourly basis (or some other frequent interval) for all power the customer consumes during that hour.

Roozbehani et al. (2012) make a very critical analysis of RTP from a control-oriented perspective. They state that directly linking price sensitive consumers to the wholesale electricity markets fundamentally changes the architecture of the system from an open-loop system in which demand is an exogenous input, to a closed-loop feedback dynamical system. The authors assert that in the absence of a well-designed control law, such direct feedback may lead to increased volatility, decreased robustness to disturbances, and new fragilities that increase the risk of a systemic failure.

Furthermore, because the real-time or day-ahead price can be quite volatile, many customers have balked at such a program for fear that they could find themselves paying extremely high prices for their consumption during any given hour (Borenstein et al. 2002).

In order to make RTP more appealing to customers, buying options such as hedging-purchasing some power through a long-term contract, before a period of system stress is evident - are generally offered in actual RTP implementations. Such measures allow customers to stabilise their overall bill while still facing the real-time price for incremental consumption. Two variants of this solution are described in the following.

RTP with customer baseline load (CBL) Many RTP programs assign a baseline consumption that they purchase at the regulated rate during each hour. The price for purchase of power at the baseline consumption is usually the TOU rate that a customer would otherwise face. Then the customer pays the real-time price for any consumption above its baseline level and receives a rebate based on the real-time price if its consumption falls below its baseline level. In financial terms, the baseline is just a forward contract for a quantity equal to the customer's baseline level, which the customer has purchased at a price set by the regulatory process.

RTP with build-your-own (BYO) baseline In BYO rather than assigning the customer a baseline level, the customer can purchase a baseline, i.e., a forward contract, to hedge as much price risk as it wants. The key would be to offer the baseline or forward contract at a price that equals the best forecast of the future spot price. Then, for incremental consumption decisions, the customer still faces the real-time price as its cost and thus has strong incentives to conserve at peak times. The BYO baseline cannot create a perfect hedge since the customer won't know in advance exactly what quantity it will consume in each hour.

Credit-based Pricing (CBP) CBP has been proposed as a viable solution for markets where consumers are also producers. In areas where renewable sources have gained, or are expected to gain, a consistent share in the market, proposals such as the one by Sgouridis & Kennedy (2010) are becoming more and more frequent. In credit-based pricing, the consumer/producer is issued a certain quantity of a virtual currency that can be spent solely on the energy market. A limited market can be allowed for such currency, allowing the users to trade their unused credits. Opportunistic behaviours and credit capitalisation are avoided by giving the virtual currency a limited lifespan (i.e., credits are issued and expire every month). The ultimate purpose of a credit-based scheme is for the system to be self-contained and sustainable on one hand, to instigate price-responsiveness and a saving-oriented-attitude in the users on the other.

The described pricing schemes are summarised in the following table. Note that due to the physics of power distribution (i.e., electricity flowing through all available paths from generator to consumer), no space-dependent pricing is possible in the retail electricity market. Additionally, in these pricing schemes hybrid payment, i.e., monetary and credit payment, is generally associated with power production (Credit-based pricing) or power rights release (Critical Peak Pricing with Incentive Preserving Rebates).

Pricing	*3. Type of t.*	*4. Modul. of t.*	*5. Users class.*	*6. Price s. str.*	*7. Payment*	*8. QoS*
Time of Usage (TOU)	c. 2nd best	b. T.dep / S.inv	a. No diff.	b. Res. Val.	a. Monetary	a. Guaranteed
TOU with Demand Charges	d. Multi part	b. T.dep / S.inv	a. No diff.	b. Res. Val.	a. Monetary	a. Guaranteed
Interruptible Electricity Rates	c. 2nd best	b. T.dep / S.inv	b. Classes	a. Cust. Val.	a. Monetary	a. Best eff. b. Capped
Critical Peak Pricing (CPP)	d. Multi part	b. T.dep / S.inv	a. No diff.	b. Res. Val.	a. Monetary	a. Guaranteed
CPP with Incentive Preserving Rebates	d. Multi part	b. T.dep / S.inv	a. No diff.	b. Res. Val.	c. Hybrid	a. Guaranteed
Demand Reduction Programs	c. 2nd best	b. T.dep / S.inv	a. No diff.	b. Res. Val.	a. Monetary	a. Guaranteed c. Incentive for not consuming
Real Time Pricing (RTP)	c. 2nd best	b. T.dep / S.inv	a. No diff.	b. Res. Val.	a. Monetary	a. Guaranteed
Credit-based Pricing	c. 2nd best	b. T.dep / S.inv	a. No diff.	b. Res. Val.	c. Hybrid	a. Guaranteed

Table 2.4: Classification of pricing techniques in electricity retail.

2.5 Road transport in Europe

Nowadays, state concession to one or more private (or partially private) firms is the most common operational paradigm for motorways in Europe (Albalate et al. 2009). In a concession agreement the State, as concession grantor, entrusts concessionaires with all the responsibilities and risks relating to the construction and operation of motorways. Currently, no EU framework for regulating motorway concessions exists. The market is configured as a monopolistic competition (as concessionaires' road facilities don't usually overlap), where revenue maximisation is constrained by the availability of public free roads.

Pricing roads, in general, involves studying the way traffic is distributed and travellers' behaviour. The subject of traffic equilibrium (also referred to as traffic assignment) is the description, through analytical tools, of the stationary distribution of vehicles in a transportation network. The network is generally represented by a directed graph, users and vehicles as commodities travelling between an origin and a destination on the graph. Assuming that travellers seek to minimise their individual travel cost, equilibrium is reached when no traveller has an incentive to modify its travel decision (Marcotte & Patriksson 2007). This type of equilibrium is known as user equilibrium and is formally defined by Wardrop's first principle (Wardrop 1952). A second type of equilibrium, known as system optimum and defined by Wardrop's second principle represents the situation where the overall travel cost on the network is minimised. User equilibrium is the configuration we will refer to from now on.

In general, the traffic assignment problem is a convex optimisation problem with linear constraints; congestion is generally taken into account, implying that the travel choice of one user is influenced by other users' travel choices. In the particular case where congestion is not considered, the problem becomes a shortest path problem with multiple users. A complete mathematical analysis and overview of solving algorithms for the Traffic Assignment problem can be found in the work by Patriksson (1994); for a more generic introduction on optimisation problems on transportation networks we refer to the work by Sheffi (1985).
Setting tariffs on a road or a network of roads implies studying the way the traffic would respond to the new pricing policy. Hence, the interaction between users and regulator (or whoever owns the roads or is in charge for setting the tariffs) must be taken into consideration. We grouped the most relevant pricing approaches for the tariff-setting problem into three categories, namely pure pricing, congestion pricing and hybrid pricing.

2.5.1 Pure pricing

From a mathematical point of view, the tariff-setting problem in a network involves two non-cooperative groups, tariff-setting agents and tariff-following clients (Bouhtou et al. 2007). Each arc in the network is owned by (at most) one agent, who can set a tariff for transiting on the arc, in order to maximise his revenues. The clients wish to route a certain demand for flow capacity on a path connecting an origin and a destination. A selected route can involve connections belonging to different agents. It is generally assumed that each client will select a route with minimum cost. This set-up is also referred to as Stackelberg Pricing Problem (Van Hoesel 2008) or Network Pricing Problem (Labbé et al. 1998).
The tarification problem has a wide range of applications such as tariff setting in freight transportation and highway toll optimisation; see Brotcorne et al. (2000, 2001) and Labbé et al. (1998), respectively. These works formulate the tarification problems as a linear bilevel model. The upper level relates to the arc owner and fixes the tariffs, the lower level belongs to the clients who, given the tariffs set by the leader, can determine their best (shortest) path. Bouhtou et al. (2007) build upon these works by expanding them with network-reduction methods and solving algorithms. Van Hoesel (2008) introduces a related network simplification method based on shortest paths.

2.5.2 Congestion Pricing

Congestion on a transport network occurs when a link or node is carrying more flow than its capacity is able to accommodate. As a consequence, the quality of service deteriorates, leading to various disadvantages for the users of the network. These are commonly quantified as delay, increase in transportation costs and, ultimately, denial of service.
Since users are willing to minimise their travel cost (in terms of time and monetary), and congestion generates externalities, marginal cost pricing is generally regarded as the way to internalise the cost of network congestion. Most of the costs of traffic congestion are borne by travellers collectively but, because individual travellers impose delays on others, they do not pay the full marginal social cost of their trips and therefore create a negative externality (de Palma & Lindsey 2011). Such externality is the difference between marginal social cost and marginal private cost and is generally referred to as user externality. The standard economic prescription to internalise the costs of a negative externality is a Pigovian tax. This was first formalised in Pigou (1920) where the author argued for a tax on congestion and thereby launched the literature on congestion pricing. Most economists have supported congestion pricing although many have been concerned about the details of implementation (Lindsey 2006). Congestion pricing has a big advantage over other transportation demand management policies in that it

encourages travellers to adjust all aspects of their behaviour: number of trips, destination, mode of transport, time of day, route, and so on, as well as their long-run decisions on where to live, work and set up business (de Palma & Lindsey 2011).
The first-best pricing principle assumes that a toll equal to the user externality is charged on each link so that the optimal network traffic flow condition can be obtained. Based on this marginal-cost pricing principle, Walters (1961) estimated an efficient system of taxation for a network of highways using actual data of traffic flow and velocity. Subsequent studies of marginal-cost pricing on networks were made by Dafermos & Sparrow (1971) and several others. In spite of its perfect theoretical basis, the marginal-cost pricing scheme is of little practical interest (Zhang & Yang 2004). The problem stems from the fact that it is impractical to charge users on each network link in view of the operating cost and public acceptance. As a result, various second-best pricing regimes have been considered.
A detailed overview of these pricing schemes, which are the methods by which congestion pricing is actually implemented today, can be found in de Palma & Lindsey (2011). A summary is given in the following section, following their proposed classification.

Types of congestion pricing schemes

Facility-based schemes Facility-based schemes include all tolls that are imposed on roads, bridges and tunnels. These are the most common form of road pricing although tolls designed to price congestion have only been implemented on a few facilities. Tolls can be levied either on all lanes of a facility or on designated toll lanes; tolls can also be levied either at a single point on a facility or at multiple points with the total amount paid determined by distance travelled.

Cordons Toll cordons are a form of area-based charging in which vehicles pay a toll to cross a cordon in the inbound direction, in the outbound direction, or possibly in both directions. A cordon scheme can encompass multiple cordons, and it can include radial screen lines to control orbital movements. All existing schemes are single cordons.

Zhang & Yang (2004) state that a systematic way for determination of toll cordons and toll levels on a general road network is not available in previous literature. Therefore, they investigate the cordon-based second-best congestion-pricing problems on road networks, developing models and algorithms for optimal selection of both toll levels and toll locations.

Examples: Norwegian rings (Bergen (1986), Oslo (1990), and Trondheim (1991)), Stockholm congestion charge (2002), Milan EcoPass (2008), Singapore Electronic Road Pricing (1998)

Zonal schemes With a zonal scheme (sometimes called an area charge), vehicles pay a fee to enter or exit a zone, or to travel within the zone without crossing its boundary. Zone boundaries can be defined by natural features such as rivers, lakes, oceans, and mountains, as well as by elements of the built environment such as roads, tunnels, bridges, residential neighbourhoods, and jurisdictions (states or provinces can define zones).

Examples: London congestion charge (2003)

Distance-based schemes With distance-based schemes, charges vary with distance travelled, either linearly or nonlinearly. As noted above, some facilities charge on the basis of distance. Networks of truck-only toll lanes and networks of High-Occupancy-Toll lanes are under consideration, and tolls on these networks are likely to be distance-based as well. For schemes that encompass multiple roads or regions the charge rate can depend on type of road. Examples: Switzerland, Austria, Germany, the Czech Republic and the Slovak Republic.

Degree of time differentiation

Flat tolls Flat tolls are constant over time. Historically, tolls on most facilities were flat because of technological or administrative difficulties in changing the toll. For some schemes the toll prevails 24 h a day. For others, such as the London congestion charge, the toll is levied at a constant rate during daytime on weekdays and not levied at other times.

Time-of-day tolls Time-of-day tolls vary by time of day, day of week, and season according to a predetermined schedule. The time intervals between toll adjustments vary across schemes, and in some cases the interval in a given scheme varies by time of day. Examples: some HOT lane facilities in the US, Singapore's Electronic Road Pricing, and Stockholm's congestion charge.

Responsive tolls Responsive tolls vary in real time (or near real time) as a function of prevailing traffic conditions. Responsive tolls are reactive" in the sense that they are set, with a short time lag, as a function of current congestion levels.

Lou et al. (2011) develop two approaches for setting tolls on toll lanes. One (a feedback-control approach) increases the toll if lane occupancy exceeds a target level. The other (a reactive self-learning approach) learns motorists' willingness to pay a toll and iteratively adjusts the toll to maintain free-flow conditions on the tolled lanes while maximising throughput. Lou et al. (2011) build on Yin & Lou (2009) by using a more realistic representation of traffic dynamics and an explicit formulation for toll optimisation. A step beyond reactive pricing is an anticipatory, or predictive, scheme in which tolls are based on forecast congestion.

Dong et al. (2011) develop an algorithm to implement predictive pricing on a High-Occupancy Toll lane facility and show that it can anticipate breakdowns in flow and maintain higher throughput than reactive pricing. Predictive pricing has long been envisaged as a tool for traffic management, but the information, communications, and computational requirements are challenging.

Examples: some High-Occupancy Toll lane facilities where tolls are adjusted to maintain free-flow speeds; the congestion pricing trial conducted in Cambridge (UK) in the 1990s in which drivers paid a charge when travel speed dropped below a threshold value. The scheme applied to all roads within the central city zone.

Other dimensions of differentiation

Differentiation by vehicle type, number of axles, and weight is common practice although these vehicle characteristics are imperfectly correlated with the congestion externality that a vehicle imposes since the externality also depends on road characteristics and the mix of users on the road. Toll differentiation according to speed and other correlates of dangerous driving behaviour has been precluded in the past by lack of information but is now technologically feasible. Toll discounts and exemptions for certain categories of vehicles and drivers are also common. (Examples: discounts for residents in London and Stockholm). A number of toll roads and High-Occupancy Toll lane facilities also offer quantity discounts in the form of reduced prices for advance purchase of multiple trips, or discounts based on cumulative usage over an accounting period (Wang et al. 2012). Quantity discounts are commonly used in transportation markets as well as other sectors of the economy, but they are inconsistent with congestion pricing according to marginal social cost pricing principles (de Palma & Lindsey 2011).

Implementations

For decades congestion pricing remained largely an ivory-tower idea, but interest gradually spread outside academia and congestion pricing has come into limited practice (de Palma & Lindsey 2011). The main operating schemes are High-Occupancy Toll lane facilities in the US, the London congestion charge, the Stockholm cordon charge and Singapore's Electronic Road Pricing system. Few benefit-cost analyses of these (or other) congestion pricing systems have been undertaken. However, the limited evidence suggests that well-designed schemes can yield significant net economic benefits. For a brief overview of existing congestion pricing systems see Black & Larson (2007) and de Palma & Lindsey (2011).

In London, drivers pay a fee to enter the central London zone during the hours from 7 a.m. until 6:30 p.m. The fee is static and does not vary by hour or road conditions (Litman 2005). In the first year of implementation, traffic in the central London zone was reduced by approximately 15%. Even though the program significantly reduced traffic in the central zone, the majority of the reduction is in the intermediate hours. The evening peak actually increases and the morning peak reduction is less than the overall reduction. This is the limitation of having a single fee for all of the commuting hours. The overall system capacity requirement was not reduced as much as it would be with differential pricing for the two peak hours.

Singapore introduced zonal pricing in 1975 and as increased the sophistication of the system over time. Singapore now uses electronic charging with 30-minute time increments and locational differentiation. Also, the transitions from one charging level to another are graduated to prevent sudden, dramatic changes that can lead to adverse consumer behaviour (such as speeding to avoid a higher priced time period). There are also different prices for cars, motorcycles, and light and heavy trucks. Congestion prices are reviewed and updated quarterly. The congestion pricing system has enabled average highway and arterial road speeds to be maintained. (LTA (2006)).

The Stockholm congestion charge is the only cordon scheme designed to manage congestion. The cordon surrounds the city centre and has 18 control points. Tolls are paid on each inbound passage up to a daily maximum. Pricing is in effect on weekdays from 6:30 to 18:30. The toll varies depending on time of day. There is no charge on weekends, holidays, or the day before holidays (Eliasson et al. 2009).

According to de Palma & Lindsey (2011) several other countries have considered regional or national road-pricing schemes to internalise congestion and other traffic externalities. However, despite the apparent success of existing schemes, and plans to establish more, congestion pricing continues to be a hard sell. Several major proposals have been scuttled by public or political opposition. The authors point out some relevant examples. Cordon tolling schemes for Edinburgh and Manchester were rejected by public referenda in 2005 and 2008, respectively. An online petition to the UK government in early 2007 attracted more than 1.8 million signatures against road pricing, and effectively put an end to plans for a national scheme in the UK for the time being. A cordon toll plan for New York City was stopped by the New York state legislature in April 2008 when it declined to vote on the proposal. And a plan to introduce a national distance-based charge in the Netherlands has been put on hold after the Dutch government collapsed in February 2010. These setbacks illustrate the difficulties of designing congestion pricing schemes that are both efficient and publicly acceptable.

Summarising, pricing schemes for road transport are, in general, either bound to a specific place or facility, or travelled-distance based, as illustrated in table 2.5.

Pricing	***3. Type of tariff***	***4. Modulation of tariff***	***5. Users classification***	***6. Price s. str.***	***7. Payment***	***8. QoS***
Facility-based/ Cordon/ Zonal p.	Generally a. flat; can be c. 2nd best	Space dep. Either c. or d.	a. No diff.	b. Res. Val.	a. Monetary	a. Best eff. b. Can be capped
Distance-based p.	Can be c. 2nd best or d. Multi part	Space dep. Either c. or d.	a. No diff.	b. Res. Val.	a. Monetary	a. Best eff.

Table 2.5: Classification of congestion pricing techniques for road transport.

Road pricing schemes are illustrated from both a theoretical and practical point of view in de Groot et al. (2002). As the authors point out, there is a lot of flexibility in the type of tariff and modulation of tariff combinations in the road transport industry: nearly all the possible combinations according to the classification proposed in the present work have been implemented in reality.

2.5.3 Hybrid (credit-based) pricing

Congestion pricing, from a theoretical point of view, is a perfectly reasonable way to internalise marginal costs of road transport. Its existing implementations have proved successful in reducing congestion but, still, as some authors point out, its application raises several equity issues. It is commonly pointed out that that, under CP, average commuters" are worse off without special revenue redistribution policies and that CP policies tend to advantage users that have a high value of travel time while penalising the others Kockelman & Kalmanje (2005) give an exhaustive overview of the literature on the subject of CP equity.
Moreover, given the general political resistance to congestion charges, some researchers and planners have turned to quantity instruments to curtail the unrestricted use of private vehicles. In quantity control, the government determines the travel demand to be served, and then assigns fixed mobility rights equally to all individual travellers or inhabitants so that fairness is explicitly demonstrated. The simplest quantity control method is plate-number-based road space rationing, widely used in many countries nowadays.
A more sophisticated quantity control method, termed the *tradable driving permit* (also called *right* or *credit*) or the *emission cap- and-trade scheme*, has been proposed in the transportation literature (for references and literature review, see Yang & Wang 2011). In a tradable credit scheme, a policy target is defined in terms of quantity and the associated equilibrium price of credits is determined by the market through free trading.

Kockelman & Kalmanje (2005) introduce an equitable CP policy called Credit Based Congestion Pricing (CBCP). They describe the policy as follows:

Under a CBCP policy, drivers receive a monthly allowance of monetary travel credits, to use on the roads. Time- and link-varying prices recognise variable demands and their associ-

ated negative externalities. Drivers do not pay money out of pocket" unless they exceed their allowance. Drivers spending less than their limit can use the credits later or exchange them for cash. For drivers with special, socially desirable travel needs (e.g., welfare-to-work participants, and single parent low-income household heads), extra credits may be allotted. CBCP has the potential to be an equitable and effective CP policy since drivers can choose to save credits or expend allowance based on their travel needs. CBCP is also revenue neutral with monthly revenues (after discounting for administrative costs) being returned as the next month's credit allowances. . . . A CBCP policy has the potential to achieve optimal network use while addressing the primary impediments to congestion pricing policies, namely equity, welfare, and revenue-distribution. Following CBCP we can also expect transit use to increase, peak travel to decrease, and total vehicle emissions to decrease.

Kockelman & Kalmanje (2005) introduce CBCP and illustrate the results of a survey on public response to it. The survey was taken by a sample of 500 individuals from Austin, TX. In Gulipalli & Kockelman (2008) hypothetic traffic and travel-welfare impacts of CBCP on Austin are analysed. In Gulipalli et al. (2011) expert opinions and system cost prediction are presented. Finally in Kockelman & Lemp (2011) the revenue- generation opportunities and welfare impacts of flat-tolling schemes, standard CP, and CBCP policies are compared.

Yang & Wang (2011) present a similar tradable credit distribution and charging scheme, where credits are universal for all tariffed links but link-specific in the amount of credit charge. The credit scheme works as follows:

1. The government initially issues credits to all eligible travellers.
2. The government predetermines the charge or the consumption rate for each roadway link.
3. Credits can then be traded freely in a competitive market without government intervention.

The scenario is modelled into a convex optimisation problem and the workings of the scheme are illustrated for a few simple example graphs. The authors demonstrate that if there is an initial uniform distribution of credits to all registered travellers, the scheme involves at least the same equity as a strict rationing policy. Furthermore, they state that the scheme is revenue-neutral, in contrast with conventional CP that involves a financial transfer from motorists to the government; in fact the scheme confines transfers to within a predefined group of travellers. In Yang & Wang (2011) fixed and elastic demand scenarios are taken into consideration. In Wang et al. (2012) the model is expanded by considering heterogeneous users with different VOTT and the formulation of the model as variational inequalities problem is introduced.

Garcia et al. (2012) introduce a "cap and trade" scheme for congestion control, pointing out that it may be easier for a central planner to set acceptable limits for aggregate utilisation on certain critical links (thus, a quantity-based approach) than identifying congestion prices which requires the accurate estimation of the derivatives of cost or latency functions associated to each link (the traditional approach to congestion charges). In their proposed schema, the planner sets constraints for aggregate utilisation on certain critical links in the network. Dynamic trading of usage rights in a secondary market is expected to identify prices clearing demand for the utilisation of the constrained links. The authors point out that if prices in a cap and trade scheme stabilise relatively quickly, a social planner can fine-tune the caps for aggregate utilisation on critical links. Their model is formulated as a variational inequalities problem that incorporates price updates, user (re)distribution on the network and travel permits trade.

Among the most relevant findings in this work is that prices stabilise and flows (or routes) converge to an equilibrium if the pace at which prices are updated is faster than that at which users adjust their decisions.

Wu et al. (2012) propose a modelling framework to design more equitable yet efficient congestion pricing and tradable credit schemes. They consider the effect of income on travellers' choices of trip generation, mode and route on multimodal transportation networks and capture the distributional impacts of congestion-mitigation policies on different income and geographic groups. The proposed mathematical program determines a pricing scheme that improves both the equity and social welfare. As these two objectives are often conflicting, they seek for a balance between them. Some suggestion is also given for actual algorithmic implementations. From the numerical examples the authors realise that it becomes difficult to design an equitable pricing policy unless a shortfall in the total toll revenue is allowed. In contrast, a more equitable and progressive tradable credit scheme can be obtained. Hence, the numerical examples also highlight the difference between these two instruments in the mechanism of achieving better equity.

Tian et al. (2013) examine the efficiency of a revenue-neutral tradable travel credit scheme for managing bottleneck congestion and modal split in a competitive highway or transit network. A continuous function represents the heterogeneity in the individuals' value of time. Each user is initially endowed with a certain amount of travel credits and can sell or buy additional credits in a free trading market. Time-dependent credit charge is implemented only for usage of the road bottleneck. Compared with the traditional time-dependent tolls, the authors prove that the proposed credit scheme is competent in resolving both the efficiency and equity problem without government participation, because the credits are freely traded in a competitive market and the revenue redistribution problem is creatively resolved. Numerical results show that the optimal-credit scheme is always unique and Pareto-improving if the system optimum is achieved and the total social cost is reduced for all levels of transit fare considered. This implies that the credit scheme can make everyone better off and simultaneously reduce the total social cost by reallocating the departure time and mode choices in an efficient manner.

Non-monetary and hybrid pricing schemes have been proposed in order to overcome the criticism that is commonly raised against congestion charges (like time-varying facility-based pricing or non-linear distance-based pricing), especially in road transport. An overview of non-monetary pricing schemes for this market can be found in Fan & Jiang (2013) and is summarised in table 2.6.

2.6 Rail transport in Europe

Due to EU regulation 2001/14/EC and subsequent 2001/12/EC, 2001/13/EC and 2001/14/EC (known as first railway package) each member state had to establish a free market. Hence, on a national scale, there generally is a Network Manager (NM)/ Infrastructure owner who owns the rails and one or more competing operators that own the trains and provide transport services for passengers and freight. NM's objective is generally revenue-oriented in countries where the role is taken over by a private company and cost-recovery-oriented where it is state-owned.
Pricing schemes for rail transport are extremely heterogeneous across countries. The main issue lies in the fact that not all customers use the same set of infrastructure services (e.g., diesel trains do not consume electricity). Hence, some Network Managers may choose to split the tariff according to a fixed part that all operators have to pay, plus a variable part that is dependent on

Pricing	*3. Type of tariff*	*4. Modulation of tariff*	*5. Users classification*	*6. Price setting strategy*	*7. Payment*	*8. QoS*
Credit Based Congestion Pricing (CBCP)	b. 1st best	Can be either a., b., c., or d.	a. No diff.	c. Both	c. Hybrid	a. Best eff.
Tradable Fuel Permits (TFP)	c. 2nd best	a. T/S inv	a. No diff	b. Res. Val.	c. Hybrid	a. Best eff. b. Can be capped
Vehicle Miles Travelled (VMT)	c. 2nd best	Can be either a., b., c., or d.	b. differentiated on vehicle	b. Res. Val.	c. Hybrid	a. Best eff. b. Can be capped
Time-place specific Tradable Mobility Permits (TMP)	a. Flat (permit is bound to a specific facility)	a. T/S inv	a. No diff but restrictions are also applicable	c. Both	c. Hybrid	a. Best eff. b. Can be capped
Tradable Driving Day Rights (TDDR)	a. Flat (one permit per day)	b. T.dep / S.inv (limited days validity)	b. No diff	c. Both	c. Hybrid	a. Best eff. b. Can be capped
Genoa Mobility Rights (GMR)	c. 2nd best	a. T/S inv (number of trips)	a. No diff but restrictions are also applicable	c. Both	c. Hybrid	a. Best eff. b. Can be capped

Table 2.6: Classification of non-monetary pricing schemes for road transport.

the service used; others prefer not to make such a distinction. Detailed descriptions of pricing schemes applied to rail transport in EU countries can be found in EU (2005) and Peter (2003).

Rail infrastructure has long been a regulated sector. The public influence on the rail network is usually of a severe kind, leading to a public ownership of the infrastructure in nearly all European countries (Peter 2003). The legislation of the European Commission refuses to require a specific a organisational structure and ownership of the infrastructure manager (IM). However, some conditions concerning open access, the price setting and slot allocation procedures have been laid down in directive 2001/14/EC (and subsequent 2001/12/EC, 2001/13/EC and 2001/14/EC, also known as First Railway Package), but leave some degree of freedom for the national governments and the respective IMs.

2.6.1 Costs of a rail infrastructure

Railway infrastructure is used as an input for different services. Freight trains and passenger trains operate on it and further differentiations can be made within these market segments, as for instance single wagon load transport incurs costs and attracts demand different from trainload transport. These services - whether they are provided inside an integrated company or over the borders of two enterprises - partly share the same infrastructure, e.g. the trackbed. Certain features of the infrastructure might be shared by one or more, but not all, services, thus generating block-wise variable costs. E.g., only electric trains make use of the power supply facilities, diesel trains don't account for the costs generated by the wires etc.

Other costs, which are entirely common to all operators, cannot be traced to any particular service or group of services. Thus, the costs of the slot provision depend not only on the traffic volume, but also on the characteristics of the infrastructure and the superstructure. Rothengatter (2003) formally includes all these costs into a mathematical formulation for a cost function composed of additively separable parts.

The difficulties of charging systems result from allocating the common costs and the block-wise variable costs to the operators, as their nature prevents them from being distributed in an impartial way. Once the block-wise variable costs are identified, the problem is reduced, as they are to be distributed only between the member of the user group at stake, which is still difficult. The problem is aggravated by the fact that these costs account for up to 80 - 90% of the total social costs of the rail infrastructure (Hylen 2001).
The remaining short run marginal costs (SRMC) change with every further movement and can be attributed directly to a particular operator. Their determination requires detailed cost studies, which can comprise a variety of elements:

- Operating costs, that can be traced to a particular train movement, e.g. for personnel and signalling;
- Wear and tear costs for maintenance and renewal of the infrastructure;
- Costs for energy consumption (electricity or diesel);
- Additional timetable planning and administration costs.

If SRMC consider ecological costs, impacts on congestion, on the noise level and accident costs of other parties, they are referred to as short run marginal social costs.
Scientific attention has also being paid to the capacity costs (Nash & Matthews 2003): since rail infrastructure managers control access to the network on a planned basis, shortages in capacity on the rail network manifest themselves in a different form from roads and other unplanned transport infrastructures. Two effects of increased capacity utilisation - congestion and scarcity - are usually considered.

Congestion represents the expected delays resulting from the transmission of delays from one train to another. The introduction of an additional rail service onto the network reduces the infrastructure manager's ability to recover from an incident and increases the probability of delays. This becomes worse at high levels of capacity utilisation, since there is a lack of spare capacity to recover from any delays (Nash & Matthews 2003). Congestion costs only occur on track sections with dense traffic, where it is more difficult for the IM to manage reactionary delays. They consist of the costs of time and energy imposed on other users of the network (Peter 2003). If the infrastructure investment is done optimally, the revenues from an optimal congestion charge will cover the deficit that is otherwise incurred (Mohring & Harwitz 1962).
Also, congestion costs, being additional delays to other operators' trains resulting from higher capacity utilisation, are borne directly by train operating companies, rather than the infrastructure manager, so they are not part of the marginal private cost of the infrastructure manager (EU 2005). But they may directly affect the demand for track access by affecting the quality of service provided, and also there may be conditions requiring the infrastructure manager to compensate train operators for delays. In either case there will be a cost, or a loss of revenue, to the infrastructure manager.

Scarcity represents the inability of a train operator to obtain the path they want, in terms of departure time, stopping pattern or speed (Nash & Matthews 2003). For example, a train operator may identify a demand for a train service from A to B at a particular time and at a particular speed but may be unable to operate the service because of the presence of other train operators' services on that part of the network at that time. This then results in the operator running its service at a time different to that which it would wish to and/or at a slower speed than it would wish to; in the extreme, the operator may decide not to, or be unable to, introduce the service at all. The inability of the train operator to provide the service it estimates will best meet its customers' demands represents a cost to society equal to the social value of that train service (Nash & Matthews 2003). Scarcity costs are external marginal costs and as such not to be confounded with opportunity costs. Pricing of scarcity ensures that the service with the highest value gets the slot and is therefore most important for the timetabling and the slot allocation during operation. It has to be answered as well for the adjustment of schedules in long-term franchises. The problem is independent of the organisation structure as integrated railways have to decide as well, how to distribute capacities (freight trains, local trains) in bottlenecks (Peter 2003).

It should also be pointed out that there are two competing philosophies of how to calculate scarcity costs. In the pure short run, when selling slots on a spot market, it is the cost of pushing another service off the tracks, or into an inferior slot, that is relevant. In a longer term track access agreement which grants specific access rights, it may be more appropriate to think in terms of the costs of providing capacity for those additional slots. These two approaches are known to economists as short run and long run marginal cost pricing respectively. When capacity is optimally adjusted, and in the absence of indivisibilities, the marginal cost of additional capacity is equal to the value of the marginal additional train, so the two costs are equal (EU 2005).

2.6.2 Pricing principles

Peter (2003) provides an overview of the pricing principles that are used or been proposed to build the tariff system in liberalised railway markets.

Short Run Marginal Cost Pricing (SRMC) Marginal costs are the costs which are incurred by an additional train run. They include the above mentioned components. Applying this pricing principle, it is ensured, that every train operator, whose willingness to pay covers or exceeds the marginal costs, can run their train. Each slot allocation will lead to a net benefit. As external costs are substantial (Nash & Matthews 2003), they should be included in the infrastructure charge.

A number of examples shows that the implementation is, at least in a rough way, possible, although the definition of the components of marginal costs may differ from country to country. Moreover, it finds acceptance among operators, due to the low costs it generates. However,if marginal costs are considered only in the short perspective, they don't cover the costs of upgrading and new investments in infrastructure, leaving this as a serious problem for the development of the rail industry as a whole. The IM will not have the necessary funds for investments, nor will he have the incentives, as new lines would only increase the deficit in the regime of SRMC-pricing. The problem is enforced by the lack of incentives for the IM to develop new cost-saving technologies, if he is regulated on the basis of marginal costs. He has no means to adjust prices to the demand of the operators, as the prices are set irrespective of this demand. He cannot gain information for investment decisions through price variations.

Ramsey-Pricing (RP) The aim of Ramsey-pricing is to maximise social welfare under the constraint of deficit coverage. It considers the fact, that the IM supplies different products. They can be defined from the demand-side and the supply-side. Rail infrastructure slots can be differentiated according to different regions, different times and different customers. Ramsey-pricing tries to find mark-ups for these products to cover the deficit that results from SRMC-pricing. The inverse elasticity rule is applied to define these mark-ups. According to this rule, the mark-up (as a percentage) on the marginal costs is reciprocally proportional to the price elasticity of the demand of the operators (while the profit is zero). A rough example in the railway sector is peak-load-pricing. Assuming, that the elasticity of operators' demand is lower at peak- times, the infrastructure tariffs can be raised during these periods.

Ramsey-pricing in the textbook form can hardly be implemented. The information requirements impose a restriction on every trial, notably the need of demand elasticity and cross demand elasticity for a variety of market segments. Operators are usually very reluctant to reveal their real willingness to pay, as it is subject to strategic behaviour. Demand curves are not easy to estimate, because of the interactions with other trains. The same holds for cost curves. Therefore, a rule of thump should be applied, following the principle to "charge, what the market can bear". This is a rough but intuitive implementation of Ramsey pricing. It has to consider the marginal costs as minimal price and Ramsey-prices are a second best solution, as they deviate from welfare maximisation. A set of second best prices is generated for the products of the IM. Prices are higher than marginal costs and the traffic volume is consequently lower than in a marginal cost pricing regime. Ramsey prices build upon marginal costs and therefore face the same information restraints as SRMC-pricing and the same lack of incentives for infrastructure upgrade.

Fully-Distributed Cost Pricing (FDC) FDC take the SRMC as a starting point. They cover the deficit by allocating the remaining costs according to selected parameters. Usual parameters are: track-km, revenues, or the SRMC themselves (Rodi, 1996). The decision, which parameters are to choose, usually doesn't consider block-wise variable costs and is therefore purely arbitrary. This makes the implementation of FDC for railway infrastructure fairly easy and is tempting for decision makers. FDC-pricing is Pareto-inferior to Ramsey-Pricing, as it doesn't take the demand elasticity into account. This is of course assuming, that these elasticity are known. If the common costs of the rail infrastructure are distributed according to the SRMC or the track-km, slots for feeder-lines and other parts of the secondary network will become very expensive. If the respective operators are priced off the network, all remaining services will have to bear a higher share of the common costs. In this way, particularly the FDC-pricing scheme can cause negative chain reactions. FDC-pricing usually does not differentiate the demand according to different train products, regions or times of the day.

Non-linear Tariffs Non-linear tariffs charge different prices per unit for different amounts of slots. The basic idea is to charge every slot with its marginal costs and to cover the resulting deficit with a fixed fee, that the operator has to pay for a certain period of time (entrance fee). A huge variation of non-linear tariffs exists, including block tariffs. The simplest form is a two-tier tariff, consisting only of one fixed fee (no differentiation between users) and one variable component. The difficulty is to define the fix part in such a way that it doesn't influence the demand of the operators. Therefore, the fixed component must not be higher than the surplus of the marginal operator.

There are significant problems in meeting this condition, if the demands of the operators differ. This is for instance the case in a market with a state owned incumbent and some small competitors. If the deficit covering fee is spread evenly across the operators, competitors are likely to be priced off the rails or the fixed fee can indeed establish a market entry barrier. A possible solution is the adjustment of the fixed fee for each operator or group of operators, leaving the variable unit-price unchanged. This approach might not to meet the competition legislation in many countries because of discrimination. If there are detailed information about the demand curve of each group of operators towards the fixed part and the variable part, customised tariffs can be assigned. Depending on the elasticity, they lie between the basic multi part concept (identical fixed and variable parts for all users) and the Ramsey-prices (no fixed part). As the regulation authority will find it very difficult to generate the necessary information, a system of self-selecting tariffs is a variation, which can be used in practice. It leads to a volume discount. It can be observed in several end consumer markets in network industries, e.g. in the electricity and the telecommunication sector.

A serious caveat of self-selecting tariffs is its reliance on the operators demand. Users have to know their consumption pattern when choosing a tariff. If the users are uncertain, the danger of selecting a wrong tariff is increased with the number of tariffs on offer and a new tariff might lead to a welfare loss (Train, 1989).

2.6.3 Implementation

The tools available to implement an access charge regime are basically of two types, simple charges and two-part charges (EU 2005).

Simple charges are directly variable with measures of use: gross tonne-km, net tonne- km, passenger-km, train-km, kW and kWh of electric traction used, per cent of revenue, etc. These can be weighted by: speed, axle loadings, types of rolling stock, the specific route (including the geometry requirements of the route), time of day, and freight commodity, among many others. Simple charges are probably more effective in collecting marginal (direct) costs, and they may be more effective in charging for social costs and externalities. They are more distorting in collecting allocated shares of fixed costs and they may not give effective signals to encourage the financing of added capacity. If used to collect fixed costs they may no longer give the right signals to use existing capacity to the full. Simple charges might be most appropriate for a relatively simple network, with few users and where traffic is not approaching network capacity.

Two-part charging systems have one or more parameters related directly to use, such as the variable parameters above. In addition, two-part systems have a second component based on the capacity forecast to be used or on some estimate of the fixed costs of the system to be recovered. This second component, sometimes called the fixed component, can be based on scheduled path-km, or scheduled train-km, among other options. The second component can also be weighted by factors such as path quality, scheduled speed, particular line, time of day, etc. It is significant that most of the second component factors tend to be passenger service-driven (particularly by commuter traffic) rather than freight-related: that is most freight users can adjust their usage to avoid peak time use (and thus do not have to burden capacity) whereas most passenger traffic must travel at times and at speeds that increase the need for capacity. Two-part regimes are more efficient at relating use to economic cost, but they raise an issue of potential discrimination among users. Two-part regimes also tend to be more complex and expensive to implement.

Two-part systems are often said to have a variable part and a fixed part. This may be somewhat misleading in the current context, since the so-called fixed part is often related to some measure of expected system use. In practice, the difference is that the variable part tends to be related to actual, measured wear and tear usage, whereas the fixed part tends to be related to the planned use of capacity. Where this simply reflects elements of marginal cost, in terms of train planning costs or use of scarce capacity, this cannot be regarded as discriminatory. However, a heavy loading of costs on to these elements combined with a requirement to reserve paths may discourage small operators, particularly freight operators, where actual capacity requirements are particularly uncertain.

2.6.4 Tariff systems currently in use

In Europe, directive 2001/14/EC claims for transparency and requires railway tariff data to be published. Peter (2003) gives a detailed description of the tariff system of European countries with the exemptions of Greece and the Republic of Ireland, which had, to that date, not implemented tariff systems. This work, although out-dated, gives a very exhaustive and complete overview of the relevant tariff parameters for each country, and of how tariff are finally calculated. Updated data can be found in EU (2005) and Macário et al. (2007).
Laurits R. Christensen Associates (2009) is part of a detailed study of competition in the US rail industry. Hence, a detailed description of the way the infrastructure is charged is given. Costs are described, the cost function illustrated and a complete economical analysis is provided.
Carlson & Nolan (2005) give a historical overview of the evolution of regulated market of railways in Canada and introduce the basic formulation for Regulated Access Pricing, comparing its Canadian implementation with others from the US, and specifically from the state of New York.

The main characteristics of the pricing approaches described so far for rail transport are illustrated in table 2.7.

Pricing	***3. Type of tariff***	***4. Modulation of tariff***	***5. Users classification***	***6. Price s. str.***	***7. Payment***	***8. QoS***
Simple charges	b. 1st best or c. 2nd best	Space dep. either c. or d.	a. No diff.	b. Res. Val.	a. Monetary	a. Guaranteed
Two-part charges	d. Multi part	Space dep. either c. or d.	a. No diff.	b. Res. Val.	a. Monetary	a. Guaranteed

Table 2.7: Classification of pricing schemes in use for rail transport in Europe

2.6.5 Further relevant research

Harrod (2013) evaluates auctions as a method of pricing and selling train paths on a shared "open access" network in order to maintain transparency in the contracting process. The author demonstrates that for multiple random samples of auction pricing for a single track railway line in the US the infrastructure entity will receive approximately 15.6% less than the true value of the contracted train paths. This loss of revenue threatens the objective of reducing government subsidy for the railway network. Hence the question of pricing train paths for open access railway networks in North America is discussed.

Van den Berg & Verhoef (2014) propose a model for congestion pricing in a road and rail network with heterogeneous values of time and schedule delay. The authors point out that the two modes are imperfect substitutes: on the road there is bottleneck congestion; in each train there is crowding congestion. They model two dimensions of preference heterogeneity that have opposite effects on the welfare impact of congestion pricing and lead to different distributional effects.

2.7 Air transport in Europe

According to the description provided in the previous section, in the current configuration European airspace has one Network Manager (Eurocontrol), who is responsible for collecting en-route charges and redistributing them to national Air Navigation Services Providers. Charges are set by national ANSPs with the aim of recovering operational costs of providing air navigation services to the airlines. From the economic point of view, the current environment is a monopolistic competition, where competitors (ANSPs) are differentiated on a location basis (country boundaries) and competitors' pricing policies are not taken into account.

Environment	
1. Control	b. Decentralised control: unit rates are set by ANSPs and collected by Eurocontrol.
2. Pricing strategy objective	a. Objective of cost recovery: en-route charges are collected to recover operational costs of national ANSP for ATC services.
Pricing	
3. Type of tariff	c. Consumption-proportional: monthly adjusted unit rates.
4. Modulation of the tariff	c. Space dependent and time independent: unit rates vary on a country bases (although EU reg. 391/2013 art. 16 allows unit rate modulation).
5. Users classification	a. No differentiation among customers: all airlines are equal;
6. Price setting strategy	b. Prices are set according to resource value: cost of ATC services
7. Payment	a. Monetary payment
8. Quality of Service (QoS)	a. and b. Guaranteed service, capped by ATC sector capacity by imposing ATC delay (but this is applied on Day of Operation).

Table 2.8: Current configuration of the route charges system

The pricing schema described above is conceptually similar to distance-based pricing for road transport (de Palma & Lindsey 2011) and simple charges for rail transport (Peter 2003). Edge pricing (telecommunications, see Falkner et al. 2000) is also similar, with the notable difference of being location independent.

2.7.1 Alternative scenario: centralised control

The present section introduces an environment where prices are set and controlled by a central authority whose objectives are cost recovery of ANS expenses and reduction of network congestion. This alternative scenario is introduced as background for the centralised pricing models described in Chapter 3.

This scenario represents a monopolistic environment having a Network Manager, ANSPs as operators and airlines as customers; this configuration is similar to the the way rail transport is organised in most European countries.

Let us now consider pricing techniques developed within other network industries that are compatible with the characteristics delineated for this alternative scenario and let us evaluate their applicability to European airspace.

Flat pricing and user-class based options should be excluded a priori: the former because it does not incentive sustainable traffic distribution, the latter because it clashes with the requirement of equity among airlines stated by EU Reg. 391/2013 art. 16: Member States, [...] may, at national or functional airspace block level and on a non-discriminatory and transparent basis, modulate air navigation charges. Since pricing is considered at a strategic level, real-time

Environment	
1. Control	a. Centralised control: tariffs will be set and collected by a Central Planner.
2. Pricing strategy objective	c. Objective of cost recovery & congestion reduction.
Pricing	
3. Type of tariff	b. or c. Proportional to travelled distance, sector forecast capacity or both; either an exact marginal cost or a second-best charging schema is plausible.
4. Modulation of the tariff	c. or d. Always space dependent, could be either time dependent or invariant.
5. Users classification	a. Equity is a priority; hence, user classes among airlines would not be welcomed.
6. Price setting strategy	b. Prices are set according to resource value: cost of ATC services; a combination of resource and user-perceived value is also acceptable.
7. Payment	a. or c. Payment could be either monetary or hybrid.
8. Quality of Service (QoS)	a. and b. Guaranteed service, capped by ATC sector capacity (this could be applied in advance of Day of Operation).

Table 2.9: Centralised control with modulated charges configuration

pricing should also be excluded.
Time dependent usage pricing (telecommunications, see Falkner et al. 2000) and Time of Usage (electricity, see Borenstein et al. 2002) are of the peak-load pricing type. Users are charged proportionally to resource consumption and the tariff varies according to the time at which the service is provided. Tariffs are set according to congestion level forecasts obtained from historical data and are adjusted periodically (e.g., once a month). There are no remarkable issues in making such a pricing scheme also location dependent, as proved by other transport modes that use time-and-place dependent peak load pricing, so it is a viable option for European airspace as well. Leal de Matos et al. (2001) investigates the issues and applicability of peak-load pricing and yield management techniques to European airspace.
Ramsey pricing based mechanisms, such as responsive pricing for telecommunications (Falkner et al. 2000), and simple charges for rail transport (Peter 2003) set the congestion charge, or the congestion dependent component of the tariff, as inversely proportional to demand elasticity of the users. Critical peak pricing for electricity retail (Borenstein et al. 2002) is conceptually similar but maximum price for peak times is generally capped. It is legitimate to assume that demand elasticity varies among users also in the air transport industry. For example an airline operating on a hub-and-spoke paradigm is likely to have a less elastic demand than one operating on a point-to-point basis due to the constraints imposed by connecting flights.
Critical Peak pricing with Incentive Rebates and Demand Reduction Programs (both from electricity retail, see Borenstein et al. 2002) are variants of the Ramsey pricing scheme where the user pays a baseline charge and is offered an incentive for not consuming during peaks and/or cash-back for consuming less than what is provided by the baseline. From a model point of view, they are equivalent to Critical Peak pricing, the only difference lying in the way the scheme is presented in the contract to the customer. For this reason, it is reasonable to discard them from our analysis.
Time of Usage with demand charges (electricity retail, Borenstein et al. 2002) is a peak-load pricing scheme where the consumer is charged an additional tax for his own peak usage The tax is independent of the level of network congestion at that time. Since generation costs in the electricity industry raise exponentially when imbalances between demand and supply occur, such a scheme is designed to deter the customer from generating peaks of usage at all. This type of pricing is therefore tightly related to a peculiar characteristic of electric grids that is

not shared by airspace and therefore has no meaningful applicability in air traffic.
Bid-price and auction-based mechanisms, such as smart-market pricing for telecommunications (Falkner et al. 2000) and pay-as-bid pricing for electricity wholesale (Kahn et al. 2001) may give rise to equity issues, since they favour economically sounder airlines that are likely to pay more. However, should the auctioning process be carried out without the use of real money, but rather with freely distributed non-monetary credit, it would be acceptable and still effective for reducing congestion. Such a system could be combined with the current distance-based en-route charges so that it would also guarantee ANS operational costs recovery.

As for other non-monetary schemes, both day permits and vehicle miles travelled (Fan & Jiang 2013) are not suitable for managing congestion in air transport. Day permits would constrain airlines in scheduling flights on allowed days only; such a limitation would be unwelcome by both airlines and their customers, and would raise severe issues on equity. Vehicle miles travelled is a fair option for reducing pollution through emission charges; in fact it shares some traits with the already implemented EU Emissions Trading Scheme (ETS), whose aim is to mitigate the climate impacts of aviation. Such a system, however, charging only on travelled distance (as with current ANS charges), is ineffective for managing congestion.
The permit system that better fits the needs of air traffic congestion management is time-place specific allowances (Fan & Jiang 2013). Under such schemes, the capacity of a resource determines the total quota of available permits. Since permits are defined for use in a specific period, their validity in time is also clearly defined and limited. In order to enforce capacity constraints for air routes or sectors, it is therefore sufficient to issue a limited number of permits for each resource and time period and distributing them with a strategy that grants equity among airlines (e.g., first-come-first-served or auctioning with credits).
Credit allowances (Fan & Jiang 2013), represent a very flexible option for dealing with demand/capacity imbalances in air traffic. Credits may be used in different ways, as a currency for congestion charges or auctions for resources, as previously suggested for bid-price and auction-based mechanisms. Provided that the initial endowment is fair, a charging system which uses credits instead of real money may partially mitigate the advantage that larger and economically sounder airlines may have under several pricing schemes.

2.7.2 Alternative pricing strategies for European ATM from literature

The possibility of modulating ANS charges by modifying the unit rate has been previously investigated in some scientific literature. The present section illustrates the most relevant contributions.

The seminal work on this subject can be found in Andreatta & Odoni (2001), who first envisaged that en route charges could be adjusted to reflect the presence of airspace congestion. The authors conceive a set of differential charges - higher in congested sectors and lower in less congested ones, which would be set in such a way that the total amount of collected charges corresponds to the cumulative cost incurred by ANSPs. This work broadly corresponds to the idea of modulated charges. However, no detailed elaboration of the proposed concept was presented.

An approach similar to peak-load pricing was proposed in Deschinkel et al. (2002); here, an assignment of price levels from a pre-determined set to airspace sector/hour pairs is formulated as a Logit model and solved through a simulated annealing algorithm. Experimental results are encouraging, however the problem of how to determine the set of price levels in the first place is not investigated thoroughly.

In Raffarin (2004) an alternative pricing rule for air navigation charges is proposed, based on the idea of giving airlines economic incentives to modify their behaviour, so that the resulting routing choices are optimal from both social and individual points of view. The author points out the limitations of the pricing rule in use today for air navigation charges in Europe, observing that it normally penalises larger aircraft, when in fact smaller ones are the ones more likely to cause congestion, due to the limited number of seats. Starting from this observation, a new pricing rule is proposed where demand and frequency of flights are re-defined as functions of customers' utility; finally, optimal tariffs are identified by equalling marginal utility to marginal costs. Despite the sound observations on which this pricing model is built, no conclusion can be drawn on its effectiveness due to the lacking of computational tests.

In accordance with the rationale of Andreatta & Odoni (2001) who envisage that en-route charges could be adjusted to reflect the presence of airspace congestion, on a small-scale real-world example, Castelli & Ranieri (2007) demonstrate the cost consequences of the trade-offs airlines are facing in situations when able to choose between the (ground) ATFM delay and rerouting. The preliminary results shown in the paper suggest that there is some scope to improve the global performance of the considered system (aggregate delay) by giving incentives to a relatively minor number of users to reroute from an overloaded to its neighbouring sector with spare capacity. However, the approach based exclusively on the general unit-rate modulation (not allowing for negative charges) led to rather high unit rate values, along with some associated unfair effects, as judged by the authors themselves. An alternative approach, with targeted unit-rate reduction (limited solely to carriers who decide to reroute) yields more acceptable unit rate values, but also a consequent loss of income to the ANSP in question, i.e., the violation of the full cost recovery principle.

An anticipatory, time-dependent modulation of air navigation charges to bring the traffic demand more in line with available network capacities is proposed by Jovanović et al. (2014). The charges are modulated so as to minimise the total cost to airspace users: a System Optimum assignment is first performed on the network to determine an optimal traffic distribution; then the network is priced in order to obtain a User Equilibrium (i.e., minimisation of users' travel costs) with the same traffic distribution. This paper puts into effect a desirable feature: revenue neutrality of the pricing scheme, meaning that no additional revenue should be generated by applying the modulation. This is achieved by introducing a toll for the use of a premium resource (overloaded network segment), and in providing economically reasonable alternatives for users who cannot get access to it due to capacity constraints. The collected toll revenues are used to subsidise the use of alternative, under-utilised network segments. The results of a medium-scale case study indicate that the use of revenue neutral tolls and rebates on a congested airspace network may yield a fairly equitable assignment. Furthermore, obtained assignment seems more efficient compared to current administrative flow management practices, capacity-wise.

Adequate modelling of route charges modulation (through modification of the unit rate) needs to address the impact on airspace users. That is to say, the obvious impact on the route choice stage that the modulation would bring along, even in a non-congested setting. In this context, Castelli et al. (2013) propose a bi-level programming pricing model for the maximisation of ANSP revenues through en route charge modulation. Results from a small-scale real-world test case suggest that the ANSP's revenue depends highly on the unit rate value. Therefore, the unit rate can be an effective instrument for modification of the route choice of flights. It follows that this approach can be a starting point to develop a pricing model with modulated en route charges, with the aim to alleviate route congestion.

The pricing rule proposed by Raffarin (2004) defines demand and frequency of flights as functions of customers' utility; optimal tariffs are identified by equalling marginal utility to marginal costs. In Deschinkel et al. (2002) optimal price levels are identified as a function of marginal utility variation. Being first-best approaches, both suffer from the implementation difficulties already pointed out for this type of approaches. Calculating actual values of marginal costs and utility poses a severe challenge, as airlines are unlikely to reveal strategic data and estimates may not lead to reliable results. Furthermore, neither work takes into account operational costs other than en-route charges (which usually amount to up to 10% of the total cost of a flight).
Hoffman et al. (2013) hints at the possibility of measuring the impact that any flight has on the overall congestion at a pre-tactical level and assessing adequate tolls accordingly, as a variation to a base tariff. No model is actually presented in this work, however the proposed scheme is most likely a real time marginal cost approach and therefore difficult to adapt to the strategic phase. Real time pricing introduces non-determinism on tolls, which is not a desirable feature when airlines still have to choose the route of a flight.

Finally, the study prepared for the European Commission by Steer Davies Gleeve (Steer Davies Gleave 2015), although not containing mathematical models nor simulation results, provides a valuable overview the applicability of congestion charges to the European airspace. This study considers several approaches to congestion pricing (proportional charges, a fixed congestion toll) and several levels at which they could be applied (ANSP level, FAB level) and analyse how they could be integrated in the current ATM system, how it would affect the airspace users and which issues might arise from each proposed modulation.

Pricing	***1. Control***	***2. Objective***	***3. Type of tariff***	***4. Modulation of tariff***	***5. Users classification***	***6. Price setting strategy***	***7. Payment***	***8. QoS***
Castelli et al. (2013)	a. cent.	a. max rev; rev. neutrality also workab.	c. 2.nd best	c. space dep, time inv.	a. No diff.	c. both	a. Monetary	a. Guaranteed
Deschinkel et al. (2002)	Not specified. Assumed a. cent.	b. resource consumption	b. 1st best (marginal utility)	d. space / time dep.	a. No diff.	a. Res. value	a. Monetary	a. Guaranteed
Hoffman et al. (2013)	a. cent.	b. resource consumption	c. multi-part (base tariff + 1st best cong. Charge)	d. space / time dep.	a. No diff.	a. Res. value	a. Monetary	a. Guaranteed b. can be capped
Jovanović et al. (2014)	a. cent.	c. resource consumption & revenue neutrality	c. 2nd best	Space dep. Compatible with both c. and d.	a. No diff	a. Res. value	a. Monetary with subsidies	a. Guaranteed
Raffarin (2004)	a. cent.	b. resource consumption	b. 1.st best	c. space dep, time inv. (ext. to d.)	a. No diff.	a. Res. value	a. Monetary	a. Guaranteed

Table 2.10: Classification of pricing rules from scientific literature for European en-route charges.

Part II

Alternative pricing schemes for the European airspace

Chapter 3

Centralised Peak-Load Pricing

From the analysis of pricing methods carried out in Chapter 2, Peak Load Pricing (PLP) stands out as one of the policies worth evaluating in regard to its applicability to European ATM. The objective is to obtain a pricing policy that guarantees a more balanced load on network capacity by redistributing part of the traffic that would otherwise converge on high demand-capacity imbalance risk areas. Such a policy is expected to reduce this imbalance and consequent delay in the tactical phase by inducing airspace users to apply re-routing and strategic delay.

Peak-load pricing is a pricing mechanism commonly used in transport and utilities. It is a simplified form of congestion pricing with the fundamental assumptions that peaks in demand are occurring periodically, in both time and location (and are therefore predictable), and that demand has some degree of elasticity towards time and/or location of service consumption (and therefore is sensitive to its price). Under these assumptions the PLP policy is to assign a higher rate for times and/or locations where a peak in demand is expected, and a lower rate for off-peak areas and times. By doing so, it is expected that part of the peak demand will deviate their travel/consumption to a cheaper option. It is therefore essential that peak and off-peak prices are set adequately so that the pricing policy is effective with regard to both business sustainability and efficient capacity management. The first can be achieved by imposing that total revenues are not lower than total marginal costs; the second can be obtained by setting peak rates greater than the willingness to pay by the users in excess, which will therefore prefer the cheaper off-peak option (in Borenstein et al. (2002) this principle is exemplified for the electricity retail market). PLP is widely used also in scheduled transport (public urban transport, rails, see for example Peter (2003)) and is, in general, transparent and predictable to users, since peak times and prices are known in advance.

In the context of European ATM, demand peaks in time are generally easy to predict, both at daily and seasonal level. Daily peaks usually occur in the morning between 7:00 and 10:00 and afternoon between 13:00 and 15:00 (UTC times); seasonal peaks occur close to holidays and special events, with the last week of June generally being the most trafficked in the year. As for price sensitivity, very few studies have been carried out on the sensitivity of airlines to route charges with regard to flight planning. Delgado (2015) carried out one such analysis on the traffic from September 12, 2014. The research presented in this paper shows that in the selection of a route by an aircraft operator the cost of the charging zones can be a factor. This impact is maximised around adjacent areas where differences in price are significant (e.g., Italy vs. the western Balcanic area) and alternative routes are similar with regard to other characteristics (e.g., route length, cruise altitude etc.). Therefore, some degree of sensitivity to price exists.

PLP is also compatible with the current unit rate setting guidelines by Eurocontrol and with EU Regulation No 391/2013 of 3 May 2013 *Laying down a common charging scheme for air navigation services*, Article 16, stating that *"Member States [...] may [...] reduce the overall costs of air navigation services and increase their efficiency, in particular by modulating charges according to the level of congestion of the network in a specific area or on a specific route at specific times. [...] The modulation of charges shall not result in any overall change in revenue for the air navigation service provider. Over- or under recoveries shall be passed on to the following period. The modulation of air navigation charges means a variation of the en route charge and/or the terminal charge [...]"*. This distinctive feature of Regulation 391/2013 provides Member States and therefore ANSPs with an operational instrument to practice demand management which could help in dealing with the recurring demand-capacity imbalance problems.

This chapter illustrates a centralised approach to PLP (which we will refer to as cPLP), where a central authority (or Central Planner, CP) is responsible for setting en route charges on the whole network. cPLP consists of two phases. In the first phase congested airspace sectors and related peak and off-peak hours are identified. The identification can be done by analysing past traffic and route choice data (as previously mentioned, air traffic shows seasonal periodicity throughout the year), or by analysing forecast Origin/Destination (O/D) demand. Specifically, in either method, the traffic demand is counted for all the sectors, taking into account all the flights and their routes, along a chosen time horizon (for example one hour). The ratio between hourly traffic count and nominal hourly capacity gives us the hourly load factor which can be used to assign the peak or off-peak label to a specific region for a specific hour. cPLP could in fact be applied at regional (i.e., FAB) level, national (i.e., ANSP) level or local (i.e., ACC or sector) level, on the whole network or on selected regions only.
In the second phase, CP has to set en route charges on the network, and the airspace users (AUs) need to route each flight, based on the set charges. These charges should guarantee that ANSPs are able to recover their operating costs and that AUs are able to perform flights, while at the same time mitigate the imbalance between demand and available airspace capacity. En route charges are set by the CP to achieve a network level objective, such as to reduce the amount of delay on the network. However, there is often a trade-off to consider between the system level (CP) and the user level (AU) objectives. In general, allowing users to minimise, for instance, their individual delays (the so-called *User Optimum* assignment) does not lead to a solution where the global network delay is also minimised. On the contrary, optimising delay at system level only (known as *System Optimum* assignment), would most likely penalise certain users more than others, which is also not ideal from an equity point of view.

Since in the current system en route charges are set by the States (ANSPs) and the AUs can only react to them by choosing alternative and cheaper routes, we model the relationship between the CP and the AUs as a Stackelberg game where a leader (CP) makes his decision first, with the complete knowledge on how the follower(s) (AUs) would react to it. The Stackelberg equilibrium can be obtained by means of an optimisation problem formulated as a bilevel linear programming model (a brief introduction to bilevel models is provided in section 3.1) where the CP sets the rates for each sector/period and the AUs make their routing choice.
Further constraints can be imposed to define the pricing scheme, ranging from allowing a different price for each sector/hour pair to imposing a single peak/off peak couple of unit rates per ANSP.

3.1 Introduction to bilevel programming

Bilevel models represent a sequential game between two decision levels, commonly referred to as *Leader* (upper level) and *Follower* (lower level), where the leader has complete knowledge of the follower's strategy, and is therefore able to anticipate it, while the opposite does not hold. The follower can therefore only react to the leader's action. This type of model is known as *Stackelberg game* (see Von Stackelberg 1952) and its solution is referred to as *Stackelberg equilibrium*, as opposed for example to *Nash's equilibrium* (introduced in Nash et al. 1950) where both players have complete knowledge of each other's strategy.

Formally, a bilevel model is a hierarchical optimisation problem in which part of the constraints translate the fact that some of the variables constitute an optimal solution to a second optimisation problem (Labbé & Violin 2013). By denoting x and y respectively the leader's and follower's decision variables vectors, this can be described mathematically as:

$$\min_{x,y} F(x,y) \tag{3.1}$$

$$\text{s.t. } G(x,y) \leq 0 \tag{3.2}$$

$$y \in \arg\min_{y} f(x,y) \tag{3.3}$$

$$s.t. g(x,y) \leq 0 \tag{3.4}$$

Note that the formulation above assumes that if there are multiple optimal solutions for the lower level problem, the solution that is most profitable for the leader is selected; this is an optimistic approach, in opposition to a pessimistic approach where the leader chooses the solution that protects himself against the follower's worst possible reaction. Both scenarios have been investigated in literature (for a detailed bibliography on this, see Heilporn 2008).

As mentioned by Labbé & Violin (2013), this type of model may represent a decision-maker, representing the upper level of the problem and whose actions lead to some reaction within a particular market or entity, corresponding to the lower level of the problem. Transportation planning is a typical domain in which examples of such hierarchical structures appear: the upper level corresponds to the network manager or owner seeking to improve the performance of the network, and the lower level corresponds to network users making their travel choices. Two common bilevel programming problems that arise in this context and are of interest to this thesis for sharing several elements with the cPLP model presented in Sec. 3.2, are the *Network Design Problem* (NDP, introduced in LeBlanc (1975) and Abdulaal & LeBlanc (1979) in its discrete and continuous forms, respectively) and the *Network Pricing Problem* (NPP, introduced in Labbé et al. 1998). In the Network Design Problem, the leader's objective is to minimise total cost on the network by choosing the links to activate and the capacity associated. The second level problem is the so-called *network equilibrium problem*, where users distribute onto the network to obtain a final equilibrium configuration where no one can further improve their objective (the so-called *User Optimum* configuration, proved to be equivalent to a Nash Equilibrium). In the Network Pricing Problem, the leader seeks to maximise his or her revenue by setting prices on a subset of the network, while the followers seek to minimise their travel cost. The followers' problem in this case is a shortest path problem, which is a linear-cost case of the network equilibrium problem.
Further domains where this type of hierarchical relationship appears include management (e.g., facility location, environmental regulation, credit allocation, energy policy, hazardous material), economic planning (social and agricultural policies, electric power pricing, oil production) and

engineering (optimal design, structures and shape).

Bilevel programs were first introduced in 1973 by Bracken & McGill (1973), while the complete formulation, as described above, was first introduced by Shimizu & Aiyoshi (1981). Generically non differentiable and non convex, bilevel problems are, by nature, hard. Even the linear bilevel problem, where the objective functions and the constraints are linear, was proved to be $\mathcal{NP}$-hard by Jeroslow (1985), while Hansen et al. (1992) proved strong $\mathcal{NP}$-hardness. Vicente et al. (1994) strengthen these results and prove that merely checking strict or local optimality is strongly $\mathcal{NP}$-hard. Due to the complexity of bilevel models, solution approaches generally focus on particular cases where functions have convenient properties, such as linearity or convexity, in order to exploit their structure to develop efficient solution methods. An annotated bibliography containing more than one hundred references on the subject can be found in Vicente & Calamai (1994), while the books by Shimizu et al. (1997) and Luo et al. (1996) are devoted, in full or in part, to this subject.

3.2 Centralised Peak-Load Pricing (cPLP) model

This section illustrates the proposed Centralised Peak Load Pricing model: Section 3.2.1 describes the assumption made, the used notation is presented in Sec. 3.2.2. Sections 3.2.3 and 3.2.4 present the formulation and constraints for the Central Planner problem and for the Airspace Users respectively, while Sec. 3.2.5 combines the two into a bilevel problem. Section 3.2.6 illustrates the reformulations carried out on the bilevel problem to obtain an equivalent mixed integer linear formulation.

3.2.1 Modelling assumptions

1. *Fixed demand matrix.* Fixed number of flights between any airport pair in the network: the intention of the proposed pricing mechanism is to modify its spatial/temporal pattern to bring it in line with available capacities, not to scale down the total demand.

2. *Heterogeneous demand*, in terms of different aircraft types and associated cost coefficients.

3. The *infrastructure capacity constraints* are known in advance, in terms of airspace sectorisation and maximum number of aircraft that can enter sectors per given period of time (i.e., nominal capacity). Since the mechanism is applied strategically, only nominal sector and airport capacities are considered, without variations introduced by regulations (which are caused by weather and other less predictable reasons, tactically).

4. *Finite set of possible (reasonable) 4D routes* for each Origin/Destination/Aircraft triple: users can select a route from a set of pre-determined routes (derived from actual traffic). Duration and profile of each route is assumed to be constant, for each aircraft type (i.e., speed profiles are assumed constant for each route/aircraft pair).

5. *Users are rational decision makers.* All AUs are assumed to choose the least-cost 4D route available. Flight cost components are attached to each route. AUs' routing decisions are therefore sensitive to modulations of en-route charges.

6. *Revenue neutrality* is established as a desired principle, meaning ANSPs' revenues are to be kept as close as possible to their operating costs: the adjustment of charges should not generate additional revenue (on top of the cost of ANS provision), nor deficit.

7. *Peak times and locations are known in advance.* The expected load on a sector, during a specific time is estimated by analysing initially submitted flight plans.

3.2.2 Model notation

F set of all flights, indexed by f

N set of all ANSPs, indexed by n

B set of all aircraft types, indexed by b

b_f aircraft type b used to operate flight f

W_{b_f} weight factor of aircraft type b used by flight f (calculated according to eq. 1.4)

A set of all airports, indexed by a

R set of all routes, indexed by r

R_f set of all routes that can be flown by flight f

S set of all sectors, indexed by s

S_n set of all sectors controlled by ANSP n

S_r set of all sectors crossed by route r

H time periods (hours), indexed by h

M time instants (minutes), indexed by m

MGS maximum ground shift (in minutes) allowed for a flight, i.e., the maximum difference between requested and allocated departure time (the formula applied is illustrated by eq. 3.7)

M_f possible departure time instants for flight f (i.e., $m \in [dt_f - MGS, dt_f + MGS]$)

T_h set of minutes m belonging to time period (hour) h

$Q_s^{(h)}$ capacity of sector s during time period (hour) h

$Q_{a,dep}^{(h)}$ departure capacity of airport a during time period (hour) h

$Q_{a,arr}^{(h)}$ arrival capacity of airport a during time period (hour) h

$Q_{a,gl}^{(h)}$ total (departures + arrivals) capacity of airport a during time period (hour) h

$D_{s,r}$ one hundredth of the great-circle distance flown in sector s if route r intersects sector s (see Sec. 1.2.4); 0 otherwise

$e_{s,r}$ estimated entry time since departure of route r in element (sector or airport) s

dt_f requested departure time for flight f

at_f requested arrival time for flight f

$adep_f$ departure airport for flight f

$ades_f$ destination airport for flight f

U_n actual (historical) en-route unit rate for ANSP n (€/service unit)

$GS^{(m)}_{f,r}$ global shift for flight f using route r and departing at time m

$OC^{(m)}_{f,r}$ airline operational costs for flight f using route r and departing at time m

$RC^{(m)}_{f,r}$ modulated en-route charges for flight f using route r and departing at time m

$TC^{(m)}_{f,r}$ total cost for operating flight f using route r and departing at time m.

3.2.3 Model formulation: Central Planner (CP) problem (upper level problem)

Main decision variables: the decision process of the CP is illustrated by an optimisation problem that identifies optimal rates for the considered airspace such that a user-optimal assignment of routes to flights will minimise the metric of network inefficiency represented by the CP objective function.

CPs decision variables are defined as:

$$RC^{(m)}_{f,r} = \text{en-route charges for flight } f \text{ departing at minute } m \text{ using route } r \quad \forall f \in F, r \in R_f, m \in M_f \tag{3.5}$$

While AU's decision variables are:

$$x^{(m)}_{f,r} = \begin{cases} 1 & \text{if flight } f \text{ departs at minute } m \text{ using route } r \\ 0 & \text{otherwise} \end{cases} \quad \forall f \in F, r \in R_f, m \in M_f \tag{3.6}$$

Route charges $RC^{(m)}_{f,r}$ can be defined in several ways, and their definition greatly affects the complexity of the model. The different pricing schemes analysed in this work and the complexity of the resulting problem are described in Sec. 3.2.7

CP objective function: in general, the objective of the Central Planner is to minimise network inefficiency. Several metrics can be used to define efficiency (or lack thereof). Delay minimisation is usually a common choice in ATFM models (see for example Bertsimas et al. 2011) but since the cPLP models deals with strategic planning, the closer equivalent to tactical delay is in fact global shift minimisation (meant as a measure of temporal displacement from the intentions stated by the airlines).

The global shift for flight f using route r and departing at time m is defined as the sum of minutes of later-than-requested departure plus the number of minutes of earlier-than-requested arrival.

$$GS^{(m)}_{f,r} = \max\{0, m - dt_f\} + \max\{0, at_f - (m + e_{ades_f,r})\} \quad \forall f \in F, r \in R_f, m \in M_f \tag{3.7}$$

A second and third components to be minimised in the CP objective are respectively the violation of the revenue neutrality constraints, ϵ and the violations of capacity constraints α, explained in the following. The complete CP objective function is therefore the following:

$$\min_x \left(\sum_{\substack{f\in F,\\ r\in R_f,\\ m\in M_f}} GS_{f,r}^{(m)} \cdot x_{f,r}^{(m)} + K_1 \cdot |\epsilon| + K_2 \cdot \sum_{s\in S, h\in H} \alpha_s^{(h)} + K_3 \cdot \sum_{a\in A, h\in H} \left(\alpha_{a,dep}^{(h)} + \alpha_{a,arr}^{(h)} + \alpha_{a,gl}^{(h)} \right) \right) \tag{3.8}$$

where K_1, K_2, K_3 are the weights to be assigned to the respective objective components.

Revenue neutrality constraint: the CP should optimise the rates so that ANSPs are guaranteed to recover their operational costs for providing air navigation services to AUs, as established by the cost recovery/determined costs system through which ANSPs are funded (see Sec.1.1.2). Therefore, the rates chosen by the CP should be revenue-neutral, meaning that they should not generate additional revenues. By definition, the currently applied country-based unit rates represent the marginal cost for providing one ANS service unit. As illustrated in Sec. 1.2.4, an ANS service unit is defined as the product of a distance-related term (specifically, one hundredth of the Great Circle Distance between entry and exit point in the national airspace) and an aircraft-weight-related term (namely, the square root of one fiftieth of the maximum takeoff weight of the aircraft). The revenue neutrality constraint therefore states that the revenues levied from the modulated route charges should not differ (in absolute value) from the revenues that would be obtained from the historical unit rates by more than a variable term ϵ, which is to be minimised in the CP objective function through the $K_1 \cdot |\epsilon|$ term.

$$\sum_{\substack{f\in F,\\ r\in R_f,\\ m\in M_f}} RC_{f,r}^{(m)} \cdot x_{f,r}^{(m)} - \sum_{\substack{n\in N, f\in F,\\ r\in R_f, m\in M_f\\ s\in(S_n\cap S_r)}} U_n \cdot D_{s,r} \cdot W_{b_f} \cdot x_{f,r}^{(m)} = \epsilon \tag{3.9}$$

Sector and airport capacity constraints: under an effective peak load pricing policy, the number of capacity breaches will be minimum. A breach occurs whenever the number of aircraft entering a sector or an airport during a certain time period (e.g., one hour) exceeds the declared capacity for that resource. Capacity constraints are therefore defined by stating that the number or aircraft entering a resource during a certain hour minus the declared capacity should not exceed a variable quantity α. These are defined for each capacitated resource and every hour when the capacity constraint is applied. In the model, they are represented by the following variables:

$\alpha_s^{(h)}$ = Number of flights exceeding capacity of sector s during hour h

$\alpha_{a,dep}^{(h)}$ = Number of flights exceeding departure capacity of airport a during hour h

$\alpha_{a,arr}^{(h)}$ = Number of flights exceeding arrival capacity of airport a during hour h

$\alpha_{a,gl}^{(h)}$ = Number of flights exceeding global (departure and arrival) capacity of airport a during hour h

The value of α variables could either be forced to zero (*hard capacity constraint*) or penalised in the objective (*soft capacity constraint*). While hard constraints would be a better choice in terms of formulation and could be used for cutting planes generation, soft capacity constraints are more realistic with regard to the application, and are therefore preferred in this case. In fact they better represent actual ATC practice, where a mild violation of nominal capacity is tolerated and handled by controllers. The formulation of the capacity constraints is therefore the following:

$$\sum_{f\in F, r\in R_f, h\in H, m\in M_f | m\in T_h} x_{f,r}^{(m)} - Q_s^{(h)} \leq \alpha_s^{(h)} \qquad \forall s\in S, h\in H \tag{3.10}$$

$$\sum_{\substack{f\in F | adep_f=a, r\in R_f,\\ h\in H, m\in M_f | m\in T_h,}} x_{f,r}^{(m)} - Q_{a,dep}^{(h)} \leq \alpha_{a,dep}^{(h)} \qquad \forall a\in A, h\in H \tag{3.11}$$

$$\sum_{\substack{f\in F | ades_f=a, r\in R_f, h\in H,\\ m\in M_f | (m+e_{a,r})\in T_h; e_{a,r}\neq 0}} x_{f,r}^{(m)} - Q_{a,arr}^{(h)} \leq \alpha_{a,arr}^{(h)} \qquad \forall a\in A, h\in H \tag{3.12}$$

$$\sum_{\substack{f\in F | adep_f=a \vee ades_f=a, r\in R_f,\\ h\in H, m\in M_f | (m+e_{a,r})\in T_h,}} x_{f,r}^{(m)} - Q_{a,gl}^{(h)} \leq \alpha_{a,gl}^{(h)} \qquad \forall a\in A, h\in H \tag{3.13}$$

$$\alpha_s^{(h)} \geq 0 \qquad \forall s\in S, h\in H \tag{3.14}$$

$$\alpha_{a,dep}^{(h)}, \alpha_{a,arr}^{(h)}, \alpha_{a,gl}^{(h)} \geq 0 \qquad \forall a\in A, h\in H \tag{3.15}$$

and their violation is penalised in the objective function by the following term: $K_2 \cdot \sum_{s\in S, h\in H} \alpha_s^{(h)} + K_3 \cdot \sum_{a\in A, h\in H} \left(\alpha_{a,dep}^{(h)} + \alpha_{a,arr}^{(h)} + \alpha_{a,gl}^{(h)}\right)$

3.2.4 Model formulation: Airspace Users' (AUs) problem (lower level problem)

The decision process of the AU is modelled as an optimisation problem where each AU aims at choosing the routes that minimise costs for each of its flights. AUs decision variables are introduced in eq. 3.6.

Flight operational costs: the cost for operating a flight typically includes route and terminal charges, aircraft fuel and maintenance costs, staff costs. Most airline operators release an annual report with aggregated cost figures for each category (see for example Ryanair 2015). Cook & Tanner (2015) identify and calculate the relevant cost coefficients per minute for fifteen reference aircraft types (estimated to cover 90% of European air traffic).The following coefficients are defined accordingly to the latter:

cam_b strategic cost of airborne maintenance for aircraft type b (€/min)

cgm_b strategic cost of ground maintenance for aircraft type b (€/min)

cf_b strategic cost of fleet utilisation for aircraft type b (€/min)

cc_b strategic cost of crew utilisation for aircraft type b (€/min)

afb_b average fuel burn for aircraft type b (Kg/min)

fc fuel cost (€/Kg)

Airborne operations costs include aircraft maintenance, fleet and crew utilisation costs plus fuel costs. These costs are accounted for the duration of a flight with the chosen route. Ground operations costs include aircraft maintenance, fleet and crew utilisation costs. These costs are accounted for every minute of strategic ground shift assigned to a flight. Note that terminal charges are excluded from this calculation because demand (i.e., origin and destination airports) is assumed as fixed. Aggregated strategic cost coefficients for airborne operations (ca_b) and ground operations (cg_b) are defined as follows:

$$ca_b = cam_b + cf_b + cc_b + fc \cdot afb_b \qquad \forall b \in B \tag{3.16}$$

$$cg_b = cgm_b + cf_b + cc_b \qquad \forall b \in B \tag{3.17}$$

(3.18)

Operational costs for operating flight f using route r and departing at time m is therefore given by the ground shift cost coefficient times the ground shift (i.e., minutes of later departure), plus the airborne cost coefficient, times the route duration.

$$OC_{f,r}^{(m)} = cg_b \cdot \left(GS_{f,r}^{(m)} - (e_{ades_f,r} - \min_{r' \in R_f} e_{ades_f,r'}) \right) + ca_b \cdot e_{ades_f,r} \quad \forall f \in F, r \in R_f, m \in M_f \tag{3.19}$$

AU Objective function: the total cost an airline has to endure for operating a flight f using route r and departing at time m is the sum of operational costs and en-route charges. Minimising total cost for operating flights is the objective function proper to each airspace user.

$$\begin{aligned} \min_x \quad & TC_{f,r}^{(m)} \cdot x_{f,r}^{(m)} \\ & = (OC_{f,r}^{(m)} + RC_{f,r}^{(m)}) \cdot x_{f,r}^{(m)} \end{aligned} \qquad \forall f \in F, r \in R_f, m \in M_f \tag{3.20}$$

Route uniqueness constraints: for every flight, exactly one route r and one departure time m are to be chosen.

$$\sum_{r \in R_f, m \in M_f} x_{f,r}^{(m)} = 1 \qquad \forall f \in F \tag{3.21}$$

3.2.5 Model formulation: bilevel cPLP

The CP problem can be combined with the AU problem into a bilevel formulation representing a Stackelberg game between the two agents, with the CP acting as leader and the AUs as followers. In such a configuration, the CP is able to anticipate the followers' reaction (in terms of route choice) to his/her pricing strategies and can therefore choose a set of rates that will optimise his objective by anticipating the associated followers' optimal route choice.

The formulation of the bilevel optimisation problem is the following:

$$\min_x \left(\sum_{\substack{f\in F,\\ r\in R_f,\\ m\in M_f}} GS_{f,r}^{(m)} \cdot x_{f,r}^{(m)} + K_1 \cdot |\epsilon| + K_2 \cdot \sum_{s\in S, h\in H} \alpha_s^{(h)} + K_3 \cdot \sum_{a\in A, h\in H} \left(\alpha_{a,dep}^{(h)} + \alpha_{a,arr}^{(h)} + \alpha_{a,gl}^{(h)} \right) \right) \tag{3.22}$$

$$\text{s.t.} \sum_{\substack{f\in F,\\ r\in R_f,\\ m\in M_f}} RC_{f,r}^{(m)} \cdot x_{f,r}^{(m)} - \sum_{\substack{n\in N, f\in F,\\ r\in R_f, m\in M_f\\ s\in(S_n\cap S_r)}} U_n \cdot D_{s,r} \cdot W_{b_f} \cdot x_{f,r}^{(m)} = \epsilon \tag{3.23}$$

$$\min_x \sum_{\substack{f\in F,\\ r\in R_f,\\ m\in M_f}} (OC_{f,r}^{(m)} + RC_{f,r}^{(m)}) \cdot x_{f,r}^{(m)} \tag{3.24}$$

$$\sum_{f\in F, r\in R_f, h\in H, m\in M_f | m\in T_h} x_{f,r}^{(m)} - Q_s^{(h)} \le \alpha_s^{(h)} \qquad \forall s\in S, h\in H \tag{3.25}$$

$$\sum_{\substack{f\in F | adep_f = a, r\in R_f,\\ h\in H, m\in M_f | m\in T_h,}} x_{f,r}^{(m)} - Q_{a,dep}^{(h)} \le \alpha_{a,dep}^{(h)} \qquad \forall a\in A, h\in H \tag{3.26}$$

$$\sum_{\substack{f\in F | ades_f = a, r\in R_f, h\in H,\\ m\in M_f | (m+e_{a,r})\in T_h; e_{a,r}\neq 0}} x_{f,r}^{(m)} - Q_{a,arr}^{(h)} \le \alpha_{a,arr}^{(h)} \qquad \forall a\in A, h\in H \tag{3.27}$$

$$\sum_{\substack{f\in F | adep_f = a \vee ades_f = a, r\in R_f,\\ h\in H, m\in M_f | (m+e_{a,r})\in T_h,}} x_{f,r}^{(m)} - Q_{a,gl}^{(h)} \le \alpha_{a,gl}^{(h)} \qquad \forall a\in A, h\in H \tag{3.28}$$

$$\sum_{r\in R_f, m\in M_f} x_{f,r}^{(m)} = 1 \qquad \forall f\in F \tag{3.29}$$

$$x_{f,r}^{(m)} \in \{0,1\}, RC_{f,r}^{(m)} \ge 0 \qquad \forall f\in F, r\in R_f, m\in M_f \tag{3.30}$$

$$\alpha_s^{(h)} \ge 0 \qquad \forall s\in S, h\in H \tag{3.31}$$

$$\alpha_{a,dep}^{(h)}, \alpha_{a,arr}^{(h)}, \alpha_{a,gl}^{(h)} \ge 0 \qquad \forall a\in A, h\in H \tag{3.32}$$

3.2.6 Linearisation of the bilevel problem

The bilevel cPLP, in the form illustrated in section 3.2.5 can be reformulated into a mixed-integer problem in two steps: first by reformulating the bilevel problem into a single level problem, then by substituting bilinear terms (i.e., products of two variables $RC \cdot x$) with continuous variables.

From bilevel to single level cPLP

The second level objective (eq. 3.20) can be reformulated into a set of equivalent constraints. In fact, since no flight can bear a negative cost, the function is separable with respect to flights, being thus equivalent to:

$$\min_{x} \sum_{\substack{r\in R_f,\\ m\in M_f}} \left(RC_{f,r}^{(m)} + OC_{f,r}^{(m)}\right) \cdot x_{f,r}^{(m)} \quad \forall f \in F \tag{3.33}$$

The meaning of this objective is that for each flight the chosen route and departure time (i.e., $x_{f,r}^{(m)}$ s.t. $x_{f,r}^{(m)} = 1$) must be the least costly one among all alternatives available for that flight. This can be equivalently expressed through a set of constraints by introducing a continuous variable c_f that represents the minimum cost of flight f:

$$c_f \leq \left(RC_{f,r}^{(m)} + OC_{f,r}^{(m)}\right) \quad \forall f \in F, r \in R_f, m \in M_f \tag{3.34}$$

$$\left(RC_{f,r}^{(m)} + OC_{f,r}^{(m)}\right) \cdot x_{f,r}^{(m)} = c_f \quad \forall f \in F, r \in R_f, m \in M_f \tag{3.35}$$

Route uniqueness constraints 3.21 impose that for each flight, only one x will be one while the others will be zero. Therefore, eq. 3.34-3.35 are equivalent to:

$$\sum_{\substack{o\in R_f,\\ i\in M_f}} \left(RC_{f,o}^{(i)} + OC_{f,o}^{(i)}\right) \cdot x_{f,o}^{(i)} \leq \left(RC_{f,r}^{(m)} + OC_{f,r}^{(m)}\right) \quad \forall f \in F, r \in R_f, m \in M_f \tag{3.36}$$

By replacing the AU's objective function (eq.3.20) with equation 3.36 an equivalent single level formulation is obtained for the cPLP model.

Linearisation of bilinear terms

The formulation obtained so far is still non linear because of the terms containing the product of two variables (i.e., $RC_{f,r}^{(m)} \cdot x_{f,r}^{(m)}$. However, since $x_{f,r}^{(m)}$ variables are binary, it is possible to linearise these terms through variable substitution. Let us introduce a new set of variables $y_{f,r}^{(m)}$, defined as follows:

$$y_{f,r}^{(m)} = \begin{cases} RC_{f,r}^{(m)} \cdot x_{f,r}^{(m)} & \text{if } x_{f,r}^{(m)} = 1 \\ 0 & \text{if } x_{f,r}^{(m)} = 0 \end{cases} \quad \forall f \in F, r \in R_f, m \in M_f$$

The $y_{f,r}^{(m)}$ variables represent the route charges component for the chosen route assignment and zero for all non-chosen alternatives. This is enforced by adding the following constraints to the model:

$$y_{f,r}^{(m)} - RC_{f,r}^{(m)} \leq 0 \qquad \forall f \in F, r \in R_f, m \in M_f \tag{3.37}$$

$$RC_{f,r}^{(m)} - y_{f,r}^{(m)} - N_{f,r}^{(m)} \cdot (1 - x_{f,r}^{(m)}) \leq 0 \qquad \forall f \in F, r \in R_f, m \in M_f \tag{3.38}$$

$$y_{f,r}^{(m)} - N_{f,r}^{(m)} \cdot x_{f,r}^{(m)} \leq 0 \qquad \forall f \in F, r \in R_f, m \in M_f \tag{3.39}$$

Eq. 3.37 imposes equality between $y_{f,r}^{(m)}$ and $RC_{f,r}^{(m)}$ when $y_{f,r}^{(m)}$ is not zero; Eq. 3.38 binds $RC_{f,r}^{(m)}$ and $y_{f,r}^{(m)}$ to be equal when $x_{f,r}^{(m)}$ is 1; the third constraint (eq. 3.39) forces $y_{f,r}^{(m)}$ to be zero when $x_{f,r}^{(m)}$ is zero.
$N_{f,r}^{(m)}$ terms are *Big-M* values.

The resulting single level linear cPLP formulation is the following:

$$\min_x \left(\sum_{\substack{f\in F,\\ r\in R_f,\\ m\in M_f}} GS_{f,r}^{(m)} \cdot x_{f,r}^{(m)} + K_1 \cdot |\epsilon| + K_2 \cdot \sum_{s\in S,h\in H} \alpha_s^{(h)} + K_3 \cdot \sum_{a\in A,h\in H} \left(\alpha_{a,dep}^{(h)} + \alpha_{a,arr}^{(h)} + \alpha_{a,gl}^{(h)} \right) \right) \tag{3.40}$$

$$\text{s.t.} \sum_{\substack{f\in F,\\ r\in R_f,\\ m\in M_f}} y_{f,r}^{(m)} - \sum_{\substack{n\in N, f\in F,\\ r\in R_f, m\in M_f\\ s\in(S_n\cap S_r)}} U_n \cdot D_{s,r} \cdot W_{b_f} \cdot x_{f,r}^{(m)} = \epsilon \tag{3.41}$$

$$\sum_{\substack{o\in R_f,\\ i\in M_f}} \left(y_{f,o}^{(i)} + OC_{f,o}^{(i)} \cdot x_{f,o}^{(i)} \right) \leq RC_{f,r}^{(m)} + OC_{f,r}^{(m)} \qquad \forall f \in F, r \in R_f, m \in M_f \tag{3.42}$$

$$y_{f,r}^{(m)} - RC_{f,r}^{(m)} \leq 0 \qquad \forall f \in F, r \in R_f, m \in M_f \tag{3.43}$$

$$RC_{f,r}^{(m)} - y_{f,r}^{(m)} - N_{f,r}^{(m)} \cdot (1 - x_{f,r}^{(m)}) \leq 0 \qquad \forall f \in F, r \in R_f, m \in M_f \tag{3.44}$$

$$y_{f,r}^{(m)} - N_{f,r}^{(m)} \cdot x_{f,r}^{(m)} \leq 0 \qquad \forall f \in F, r \in R_f, m \in M_f \tag{3.45}$$

$$\sum_{f\in F, r\in R_f, h\in H, m\in M_f | m\in T_h} x_{f,r}^{(m)} - Q_s^{(h)} \leq \alpha_s^{(h)} \qquad \forall s \in S, h \in H \tag{3.46}$$

$$\sum_{\substack{f\in F | adep_f = a, r\in R_f,\\ h\in H, m\in M_f | m\in T_h,}} x_{f,r}^{(m)} - Q_{a,dep}^{(h)} \leq \alpha_{a,dep}^{(h)} \qquad \forall a \in A, h \in H \tag{3.47}$$

$$\sum_{\substack{f\in F | ades_f = a, r\in R_f, h\in H,\\ m\in M_f | (m+e_{a,r})\in T_h; e_{a,r}\neq 0}} x_{f,r}^{(m)} - Q_{a,arr}^{(h)} \leq \alpha_{a,arr}^{(h)} \qquad \forall a \in A, h \in H \tag{3.48}$$

$$\sum_{\substack{f\in F | adep_f = a \vee ades_f = a, r\in R_f,\\ h\in H, m\in M_f | (m+e_{a,r})\in T_h,}} x_{f,r}^{(m)} - Q_{a,gl}^{(h)} \leq \alpha_{a,gl}^{(h)} \qquad \forall a \in A, h \in H \tag{3.49}$$

$$\sum_{r\in R_f, m\in M_f} x_{f,r}^{(m)} = 1 \qquad \forall f \in F \tag{3.50}$$

$$x_{f,r}^{(m)} \in \{0,1\}, RC_{f,r}^{(m)} \geq 0, y_{f,r}^{(m)} \geq 0 \qquad \forall f \in F, r \in R_f, m \in M \tag{3.51}$$

$$\alpha_s^{(h)} \geq 0 \qquad \forall s \in S, h \in H \tag{3.52}$$

$$\alpha_{a,dep}^{(h)}, \alpha_{a,arr}^{(h)}, \alpha_{a,gl}^{(h)} \geq 0 \qquad \forall a \in A, h \in H \tag{3.53}$$

3.2.7 Pricing schemes for cPLP

According to the way the en-route charges variables $RC_{f,r}^{(m)}$ are defined, different pricing schemes can be derived. This section illustrates some relevant examples.

Trajectory-based pricing

In a trajectory-based pricing scheme rates are set on a 4-dimensional route basis, meaning that a price is assigned to the entire route and the modulation depends on the time chosen for departure.

In terms of modelling, each $RC_{f,r}^{(m)}$ variable is independent of the others and therefore no further constraint is added to the model.

Segment-based pricing

In a segment-based pricing scheme different rates are set for each airspace sector. En-route charges for a route are therefore calculated as the sum of the charges associated with the segment flown in each airspace sector of that route. Modulation depends on the entry time in each sector, thus, indirectly, on the time chosen for departure.

Let us therefore introduce the $P_s^{(h)}$ rates as the unit rates applied on sector s during hour h. Resulting en-route charges for a flight f flying on route r and departing at time m are defined as follows:

$$RC_{f,r}^{(m)} = \sum_{s \in S_r} P_s^{(h|(m+e_{s,r}) \in T_h)} \cdot D_{s,r} \cdot W_{b_f} \tag{3.54}$$

ANSP-based peak load pricing

In an ANSP-based pricing scheme, rates are applied at ANSP level, meaning that all sectors included in the airspace of a certain ANSP share the same rates. En-route charges modulation is accomplished by imposing a higher rate when and where traffic peaks are expected (*peak rate*) and a lower rate for less trafficked periods and areas (*off peak rate*). A couple of peak and off peak rates is therefore defined for each ANSP and their values are assigned to each sector/hour pair, as follows:

$$P_s^{(h)} = \begin{cases} Pp_n & \text{if } h \text{ is peak time for sector } s \\ Po_n & \text{otherwise} \end{cases} \qquad \forall n \in N, s \in Sn_n, h \in H \tag{3.55}$$

Where P_{p_n} and P_{o_n} represent, respectively, the peak and off-peak rate for ANSP n. Resulting en-route charges for a flight f flying on route r and departing at time m are calculated as in the general segment-based pricing scheme (eq. 3.54).

A variant of the ANSP based peak load pricing scheme is obtained if the peak and off-peak rates are parametrised, meaning that one is defined proportionally to the other. In this case the parametrised set of tariffs is defined as follows:

$$P_s^{(h)} = \begin{cases} Po_n & \text{if } h \text{ is off-peak time for sector } s \\ \delta \cdot Po_n & \text{if } h \text{ is peak time for sector } s \end{cases} \qquad \forall n \in N, s \in Sn_n, h \in H, \delta \geq 1 \tag{3.56}$$

where δ is a term greater than 1 denoting the proportionality ratio between the peak and off-peak rates. Note that in this case either Po_n or δ can be decision variables of the problem. ANSP-based peak load pricing schemes have the desirable feature of being simple and transparent, being the most similar to the schemes already applied successfully in several other industries (as illustrated in Chapter 2). Therefore only this type of scheme will be further analysed in the remainder of the work, in two variants: a two-modulated-rates per ANSP scheme, referred to as *cPLP-two*, obtained by adding eq. 3.54 and 3.55 to the model, and a one-rate per ANSP scheme referred to as *cPLP-one*, obtained by adding eq. 3.54 and 3.56 to the model, with variable δ and a constant value of Po_n set equal to the historical unit rate U_n.

Chapter 4

Air traffic data management

Real data availability is a well known issue for research on European aviation. Differently, for example, from the United States case, where statistics and data sets on airport schedules, routing, flight delays are freely available for research purposes from the Department of Transport's and from the Federal Aviation Administration's websites, no such service is provided in Europe. Therefore, accordingly to the nature and the specificities of the research carried on, *ad-hoc* agreements have to be made with Eurocontrol, ANSPs, airlines or other authorities to obtain the needed data. Further complications then arise from the technical difficulties in cleaning and harmonising data obtained from different sources.
Due to these difficulties, scientific literature on European Air Traffic Flow Management, Airport scheduling and slot assignment often relies on randomly, self-generated data, which unavoidably biases the plausibility of the results obtained with regard to actual implementability of the proposed solutions.
The present chapter illustrates the process through which the data files used as testing instances for the pricing models introduced in Chapter 3 (and used throughout the SATURN project) were built, from real traffic data sets licensed by the Network Manager. The data sources used are presented in Section 4.1, while Section 4.2 illustrates the cleaning, filtering and harmonising processes carried out to obtain data files that formalise actual airspace configuration, traffic and route structure being at the same time computationally viable and adherent to reality. The support tool built within this process is a relational database with geographic capabilities. The general structure of this tool is illustrated in Section 4.4, while a detailed documentation of the database structure and use-case examples are provided in Appendices A and B respectively.

4.1 Data sources

Traffic, network and airspace structure data are obtained from several repositories owned by Eurocontrol. The different sources and tools used are briefly described in the following.

4.1.1 Eurocontrol Demand Data Repository (DDR2) data sets

As stated in the Network Operations Plan (Eurocontrol 2014*d*), the Demand Data Repository, phase 2 (DDR2) was initially created in response to planners who needed access to an integrated pan-European strategic view of traffic demand and distribution. Access to DDR2 is therefore restricted to ANSPs and Airline Operators within Europe. The DDR2 provides, through a web application, an interface that allows generation and download of future traffic forecasts (for planning purposes) and past traffic (for post-operations analysis of traffic trends and statistics) data sets. Past traffic demand data is available for any day, for all Europe, from January 2006 and grouped chronologically according to the Aeronautical Information Regulation And Control (AIRAC) cycle of reference. AIRAC cycles (as set by ICAO) have 28 days, the starting date of a cycle is therefore always on a Thursday (e.g. 4th August 2005, 1st September 2005 etc.). Approximately one week after the end of each AIRAC cycle, once all data has been finalised, the data files for the 28 days are generated and then published on the DDR2 website.

Figure 4.1 describes the AIRAC data file preparation process.

Figure 4.1: DDR data sources. (source: Eurocontrol)

The data format and conventions used throughout the over 40 file types available through DDR2 website is briefly described in the DDR2 Reference Manual (for the latest updated version, see Eurocontrol 2014*c*).

4.1.2 Eurocontrol NEtwork Strategic Tool (NEST) software

NEST is an airspace design and capacity planning tool that can be used by the Network Manager and national ANSPs for airspace structure design and development, capacity planning and post operations analysis, strategic traffic flow organisation, scenario preparation for simulations and studies at local or network level (Eurocontrol 2014*d*). The main functionalities of the software include the following:

- 3-dimensional visualisation of traffic demand and airspace;
- Evaluation of traffic loads;
- Identification of capacity bottlenecks;
- Simulation of future traffic demand;
- Simulation of different routing options;
- Simulation of delays, including network effects;
- Optimisation of configurations and opening scheme;
- Editing of airspace and route network, including free route airspace;
- Evaluation of flight efficiency indicators.

NEST compatible data are compressed archives with a .nest extension that include airspace/network and traffic data and are available at the end of each AIRAC cycle through the DDR2 data repository web interface. The 28 days of data stored in an AIRAC data file can contain either official data validated by Eurocontrol or custom data from an alternative source (Figure 4.2 illustrates the data sources used for .nest archives). It is possible to export data from .nest archives to other file formats (namely, formatted text files and Microsoft EXCEL spreadsheets) which offers some degree of compatibility with other software.

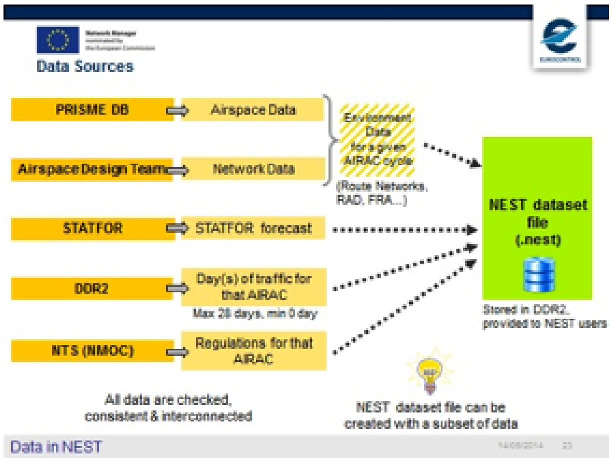


Figure 4.2: NEST data sources. (source: Eurocontrol)

In particular, NEST datasets comprise:

1. **Airspace data:** files containing 2-dimensional coordinates and the corresponding altitudes (expressed as upper and lower Flight Levels) of airblocks, definitions of how elementary sectors are built from airblocks and how elementary sectors can be collapsed. Airport, sector and traffic flows capacity figures are also provided. Further data provides the description of bigger airspace volumes (e.g., ACCs) together with the different sector configurations used by ACCs, and the start and end time when these configurations are activated.

2. **Network data:** files related to the structure of the network, like waypoints and airports coordinates, route segments, rule files (e.g., for SIDs and STARs), Flight Level constraints.

3. **Traffic data**: files normally provided by the Network Manager and referred to the following three phases of flight operations:

   Initial (M1): the last filed flight plan from the airline;

   Regulated (M2): same as the last filed flight plan for non-delayed flights and with a constant time offset added to the ATFM delayed flights. These offsets correspond to the CASA-calculated ATFM delay;

   Actual (M3): corresponds to the last filed flight plan data updated with available radar information whenever a flight deviates from its filed flight plan by more than any of the pre-determined thresholds of five minutes, seven flight levels or 20 NM at any point from the planned trajectory. This represents the closest estimate available in NEST data files of the flight trajectories actually handled by ATC controllers on the day of operations.

The Extended IFPZ Area is the geographical area used by NEST to define a daily flight list. This area is based on the operational ECAC Area for which the Network Manager receives data, but is extended to include additional areas in order to ensure stability of data. The extended area includes, but are not limited to, Tunisia, Senegal, Egypt, Libya, Israel, Iran, Jordan, Lebanon and Syria, as well as parts of Iceland and from the UK border to Africa and Cape Verde (Figure 4.3).

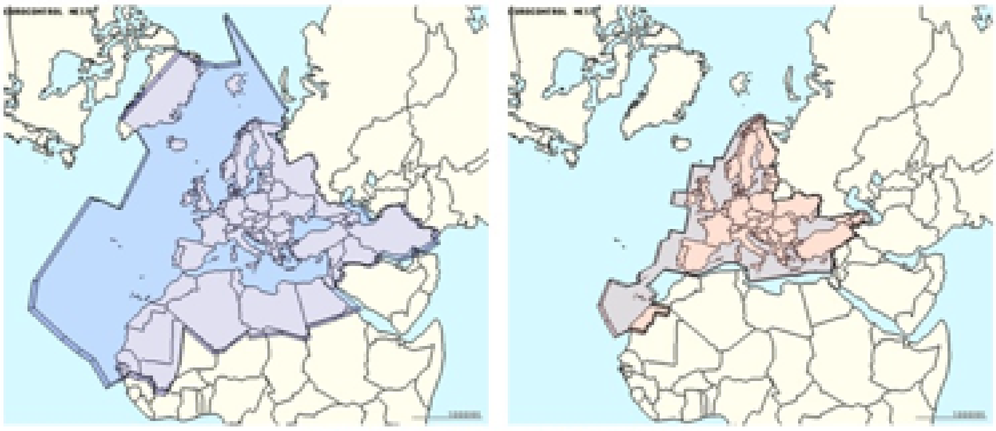


Figure 4.3: Left: IFPZ Area (left) and ECAC Area (right) (source: NEST)

4.1.3 European airline delay cost reference values report (Cook & Tanner 2015)

This report is designed as a reference document for European delay costs incurred by airlines; the cost of delay is calculated separately for strategic delays (those accounted for in advance) and tactical delays (those incurred on the day of operations and not accounted for in advance).

Costs are assigned under three cost scenarios: *low*, *base* and *high*. These scenarios are designed to cover the likely range of costs for European operators. The base cost scenario is, to the greatest extent possible, designed to reflect the typical case. All calculations are undertaken for fifteen reference aircraft (that together make up for more than 90% of European air traffic): B733, B734, B735, B738, B752, A319, A320, A321, A332, AT43, AT72, B744, B763, DH8D and E190.

The following cost components are considered:

- **Fuel:** the cost of fuel burned per minute (only for off-gate phases);
- **Maintenance:** relates to factors such as the mechanical attrition of aircraft waiting at gates (strategically or tactically) or aircraft accepting longer re-routes in order to obtain a better departure slot (tactically). Both strategic and tactical at-gate costs are relatively low (compared to the other phases) because relatively little wear and tear on the airframe is experienced at-gate and the engines are off for the vast majority of this time.
- **Fleet:** refers to the full cost of fleet financing, such as depreciation, rentals and leases of flight equipment. These costs are determined by service hours. Since utilisation has only a very small effect on these costs, they are wholly allocated to the strategic phase and the corresponding tactical delay costs are thus taken to be zero.

- **Crew:** typical pilot and flight attendant salaries were calculated in 2008 for various European airlines, using their corresponding payment schemes with realistic annual block/flight duty hours, sectors flown and overnight stopovers. To update the 2008 costs to 2010 values, pay deals since 2008 for ten European airlines are considered. Pilots salaries generally increase by size of aircraft. Flight attendants salaries are more consistent across all aircraft types. Tactically, in certain cases, delays may not generate additional crew costs, and the low cost scenario is set at zero cost. The high cost scenario is based on overtime rates. The base cost scenario is based on typical time-based costs. The crew costs commonly apply to ground and airborne phases.

- **Passengers** only the costs that impact on airlines business are included in the report. *Hard* costs are due to such factors as passenger rebooking, compensation and care. *Soft* costs represent revenue losses for the airline and manifest themselves in several ways, which are generally hard to quantify. For example, due to a delay on one occasion, a passenger may defect from an unpunctual airline as a result of dissatisfaction (and maybe later come back). A passenger with a flexible ticket may arrive at an airport and decide to take a competitors on-time flight instead of a delayed flight, on which they were originally booked.

The report presents costs of delay by four flight phases (as illustrated by Figure 4.4): at-gate, taxi, cruise extension and arrival management (tactical only). Block hours are defined as the time spent off- blocks (aircraft utilisation). Service hours are defined as the total time spent in service during the operational day.

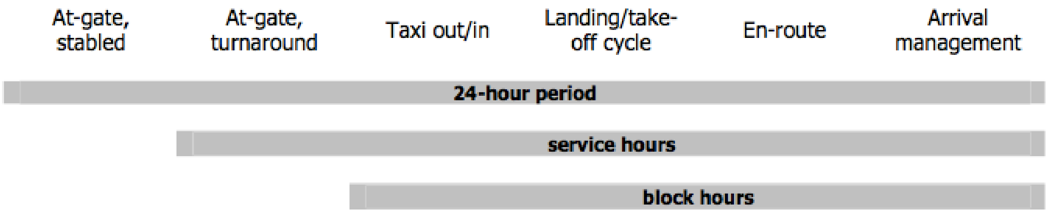


Figure 4.4: Flight phases and basic definitions (source: Cook 2007)

Figure 4.5 shows which cost types are assessed at each level. Strategic costs and tactical costs are not independent: reactionary delays depend on the airlines ability to recover from the delay, due to the amount of schedule buffer, for example. If no buffers were used, the reactionary costs would increase markedly and the tactical costs would be significantly higher.

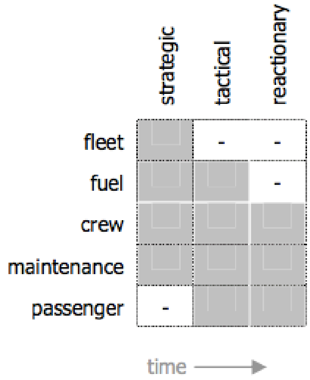


Figure 4.5: Cost types assessed at each level (source: Cook 2007)

4.1.4 Eurocontrol Base of Aircraft DAta (BADA) data repository

The Base of Aircraft Data (BADA) provides a set of ASCII files containing performance and operating procedure coefficients for 338 different aircraft types. For 117 of these aircraft types, data is provided directly in files. For the other 221 aircraft types, the data is specified to be the same as one of the directly supported 117 aircraft types: in this case they are declared as equivalent to one of the other 117 aircraft model. The coefficients include those used to calculate thrust, drag and fuel flow and those used to specify nominal cruise, climb and descent speeds (although, for the purpose of the present thesis, only basic technical data about aircraft, like maximum takeoff weight and aircraft category, was sourced from BADA). The BADA User Manual (the latest version available from Eurocontrol's website is Eurocontrol 2011) describes in detail how each coefficient is calculated. Access to BADA is given through a license agreement to Air Navigation Service Providers, research and development organisations, universities and commercial entities working on Air Traffic Management related projects.

4.1.5 Eurocontrol ATM Cost-Effectiveness (ACE) 2013 Benchmarking Report with 2014-2018 outlook (Eurocontrol 2015*a*)

This report (updated annually) presents a review and comparison of ATM cost-effectiveness for 37 Air ANSPs in Europe. The benchmarking work is carried out by Eurocontrol's Performance Review Commission supported by the Performance Review Unit and is based on information provided by ANSPs themselves.
ACE 2013 presents information on performance indicators relating to cost-effectiveness and productivity for the year 2013, and how they changed over time (over the 2009-2013 time arc). It examines both individual ANSPs and the Pan-European ATM system as a whole. In addition, it analyses forward-looking information covering the 2014-2018 period based on information provided by ANSPs in November 2014.
The ACE reports ANSP's costs figures from ANS provision, which were used to validate the results obtained from the pricing models developed within the SATURN project. The ACE report is freely available from Eurocontrol's website.

4.1.6 Eurocontrol Slot Coordinator tool

The Slot Coordinator tool is a web-based interface that offers different post-analysis table views of traffic at slot coordinated airports: flight plans without slot detail, slots without flight plan detail, slot adherence, aircraft type discrepancy, flight detail list, slot detail list, airport slot granularity, and airport slot distribution. Two tables are of interest to the current study:

1. **Flight detail list table:** flights can be filtered according to the set of airports, and the set of aircraft operators and the specific time frame of interest. The data fields provided include origin and destination airports, aircraft type and tail number, slot time (slot assigned to the aircraft operator for that flight, either during the primary slot allocation process, or in the ad-hoc procedure from the pool of unassigned slots), requested slot time (slot aircraft operator requested for that flight, either before the primary slot allocation, or in the ad-hoc procedure), the type of service (e.g., scheduled, non-scheduled, etc.)

2. **Airport slot distribution table:** this table gives an overview of the arrival and departure slots. For each time interval or slot width, for which the slot coordinator assigns the slots, the table lists how many arrival (or departure) slots are assigned to each airline operating to/from that airport. The table gives both the exact slot time assigned and the slot starting and ending points, as they are not always the same. For example, the slot width can be 10 minutes. An airline can get a slot at minute 1, that is to say at the beginning of the slot width, or at minute 5 (in the middle of slot width interval). Different coordinators use different slot widths, as well as the method for slot assignment (i.e., at the beginning, or in the middle of slot, or a combination). The slot width is most often 5-10 minutes long, but can be 20 minutes, or even 60 minutes in extreme cases. The slot width can be assigned to one or more airlines, and can be for different types of operations (arrival and/or departure).

The Slot Coordinator tool is licensed and accessed under similar terms to the DDR2 data sets.

4.2 Data selection

This section illustrates how the data about airspace structure, traffic and cost was encoded into the data instances used for the computational tests of the models described in Chapter 3. For the vast majority of the tests, September 12, 2014 was chosen as reference day for building the data instances, being a highly trafficked day (33 810 flights, the fourth most trafficked day of the year) with no major disruption that affected ATC capacity (i.e., weather, strikes etc). The original data files, as provided by the Network Manager, were highly heterogeneous in terms of format, both regarding file structure and data itself (e.g., different coordinates system, different formats used for dates and time etc.). Some initial effort was therefore put in data harmonisation and cleaning, since some incoherences were also found in the reported flight trajectories and trajectory/sectors intersection files (although this appears to be a known issue within these data sets, see for example Zanin et al. (2014)). Cleaning of traffic data was therefore performed to filter out all flights that were beyond the scope of the SATURN project (i.e., non-scheduled IFR, flights, circular flights etc.).
Finally, due to the complexity of the system and modelling requirements it was necessary to introduce some degree of simplification and to perform data clustering operations (described in the following Section 4.14).

4.2.1 Airspace structure data

Sectors

In order to obtain a representation of airspace as realistic as possible, several aspects had to be taken into consideration:

Sector configurations and openings The configuration of open sectors for each Area Control Centre can change multiple times per day. Two options were tested with regard to this: first, a configuration where the maximum available capacity is considered. This can be justified by the fact that no decision on sectorisation is taken strategically; therefore it makes sense to consider the maximum capacity that the network is able to sustain. Such a scenario can be approximated with the configuration that is active at the most trafficked hour of a day (the 7:00-8:00 AM period is typically a good choice in this regard). The second option is to define, for each time period (i.e., one hour) the set of sectors that were actually active at that time. This can be justified by the fact that a realistic airspace configuration is more likely subject to congestion-related issues and therefore is better suited for testing models that aim at effective traffic redistribution (for these reason this option was preferred over the less realistic "maximum capacity" to test the pricing models).

Sector capacity Sector capacity is considered with a granularity of one hour, meaning that some approximations were made for capacities that were defined for an inferior period of time or whose capacity figure changed during the hour. In these specific cases, a weighted average was calculated and the resulting figure was applied as maximum capacity for that hour. For example, a sector active from 8:30 (and previously inactive) on with a declared capacity of 30 entries per hour would be assigned a capacity of 15 for the 8:00-9:00 period. Similarly, a sector whose capacity changed from a figure of 20 until 8:30 to 30 from 8:31 on, will be assigned a capacity equal to 25. All sectors that did not have a capacity figure specified when active were assigned a 999 figure (corresponding, in practice, to an infinite hourly capacity).

Traffic volumes (TVs) and TV capacities Traffic volumes extend the definition of airspace sector by means of rules for the exclusion or inclusion of traffic flows. This allows for a finer-grained and more detailed definition of airspace capacity (e.g., "the maximum number of flights coming from waypoint W that can enter sector S during hour H is equal to N."). In order to keep the set of capacity constraints as compact as possible, instead of adding new ones for TVs, it was preferred to incorporate TV capacity into sector capacity constraints, in the following way:

Sectors marked as active with no TV associated: nominal hourly capacity of the sector is taken;

Sectors with an associated TV with no inclusion/exclusion rules: either nominal sector or TV capacity is taken, depending on which one is marked as active in the data;

Sectors with one or more associated TV with inclusion/exclusion rules: either nominal sector or TV capacity is taken, depending on which one is marked as active in the data.
Often, only a subset of the flights entering the specified sector is counted by the TV. For many TVs, applying the TV capacity to constrain the number of total entries in a sector is not possible because the two sets may largely differ. An example is shown in Figure 4.6. TV EHTRANS regulates sector EHAACOD, but traffic coming from

Shanwick oceanic region sectors					
EGATLANT	EGGXDUIVEL	EGGXL60	EGGXM59	EGGXSECORN	EGGXUBEGAS
EGBRAND	EGGXFIR	EGGXLFRR	EGGXM60	EGGXU45	EGGXUPOR
EGGX45N	EGGXL45	EGGXLPOR	EGGXMBEGAS	EGGXU55	EGOCEAN
EGGX45N30W	EGGXL55	EGGXM45	EGGXMPOR	EGGXU56	EGSOYUZ
EGGXALL	EGGXL56	EGGXM55	EGGXOCA340	EGGXU57	EGTRAGNY
EGGXBEGAS	EGGXL57	EGGXM56	EGGXPOR	EGGXU58	
EGGXBUF	EGGXL58	EGGXM57	EGGXSCO	EGGXU59	
EGGXCCC	EGGXL59	EGGXM58	EGGXSE4945	EGGXU60	

Table 4.1: Sectors and traffic volumes belonging to the Shanwick oceanic region

or going to the Netherlands is not included in the TV count. TV capacity is never exceeded by the traffic in EHAACOD that satisfies the exclusion rule. Without considering the exclusion rule, the TV capacity would be violated from 03:00 to 21:59. A total of 51 flights that satisfy the rules enter the TV during the whole day, while 985 flights enter the sector attached to this TV.
Therefore, to obtain a realistic capacity figure for sectors where TV with inclusion/exclusion rules are active, the sector entry count from actual flown data (i.e., M3, radar data) is compared with the capacity of the TV, to determine the total number of flights exceeding TV capacity in each hour. The average hourly capacity exceedance is then summed to the nominal TV capacity, to obtain the adjusted TV capacity to be used in models.

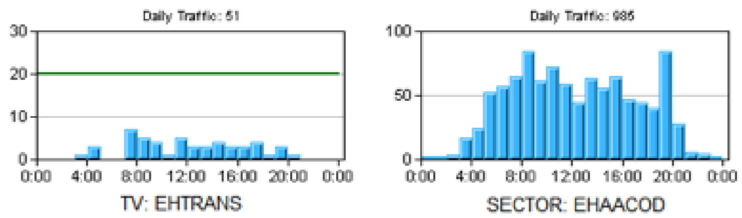


Figure 4.6: Comparison of traffic volume count and sector entry count (source: NEST)

Air Navigation Service Providers (ANSPs)

In order to calculate en-route charges correctly, each sector must be associated to the correct ANSP. The European airspace we consider comprises 41 states, which are memebers of Eurocontrol's CRCO system plus Estonia and Ukraine, for geographic reasons. Normally, the first two letters of a sector name identify the state to which that sector belongs; there are however a few exceptions, that have to be considered on a case-by-case basis and hard-coded in the instance generation process. These are the following:

Shanwick Oceanic region: Shanwick area is an oceanic region comprising a total of 45 sectors (listed in Table 4.1) whose control is delegated to UK and Ireland by ICAO. This area is subject to a different charging system than the usual applied to all other Eurocontrol areas. Specifically, flights are subject to a single rate for entering the region instead of a distance-proportional charge.

Santa Maria oceanic region: Santa Maria is an oceanic region belonging to the Eurocontrol area whose control is delegated to Portugal. Sectors belonging to this region normally have names starting with "AZ", which identifies the Azores ACC, which has a unit rate different from the one applied to Portuguese airspace (names starting with “LP”). There are however two sectors belonging to the Santa Maria region whose names start with “LP” (LPPOALL and LPPOFIR) that need to be identified in order to assign them to the correct charging region.

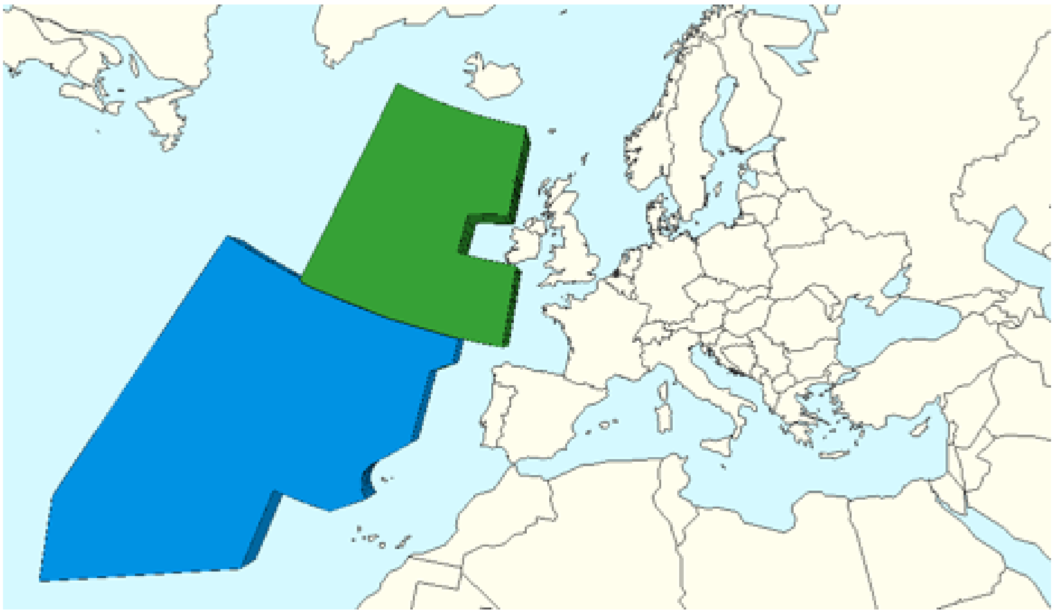


Figure 4.7: Shanwick (green) and Santa Maria (blue) oceanic regions (source: NEST)

Gran Canaria ACC: the Canary Island region is controlled by Enaire, the Spanish ANSP, but has a different unit rate from the one applied to continental Spain. When evaluating ANSP revenue from en-route charges then, both charges levied from Spanish airspace (sector names starting with “LE”) and from the Gran Canaria ACC (sector names starting with “GC”) should be considered for Enaire.

Bosnian airspace: between November 2014 and November 2015, the ANSP from Bosnia and Herzegovina, Bhansa, is progressively taking over control over its national airspace (sector names starting with “LQ”). Due to delays in establishing a proper system of air traffic control in Bosnia, control was previously delegated to the Serbian and Croatian ANSPs, with quotas of 55% of upper airspace and 45% of upper plus 100% of lower respectively (Barta & Kazda 2014). Since it was not possible to retrieve exact information on which sectors were controlled by which of the two ANSPs, this issue was discarded and in the process of data instances generation it is assumed that Bhansa controls the entire set of sectors that constitute Bosnian airspace.

Maastricht Upper Area Control Centre (MUAC): Eurocontrol’s Maastricht Upper Area Control Centre (MUAC), located at Maastricht Aachen Airport, provides air traffic control for traffic above 24,500 ft over Belgium, Luxembourg, the Netherlands, and north-west Germany. This area covers the Brussels UIR, the Hannover UIR and the Amsterdam FIR. The lower airspace (from level 0 to 24,500 feet) over this area is managed by Belgocontrol, Deutsche Flugsicherung and Luchtverkeersleiding Nederland respectively. Sectors controlled by MUAC have names that start with “EDYY”.

Airports and airport capacities

Although this study focuses on en-route capacity management, ignoring airports and associated capacity constraints altogether would lead to completely unrealistic results. Therefore it was preferred to keep allowed departure and arrival times fixed within the assigned slots whenever possible, and to include airport capacity constraints in the pricing models.
Airports can have associated traffic volumes, just as sectors, but differently from these latter, they can have up to three distinct constraints on hourly capacity, namely for departure, arrival and global movements (departures plus arrivals).
From an analysis of actual traffic it emerged that for some airports the capacity figures specified in the data are systematically broken. Moreover, it turned out that these capacity constraints could never be satisfied by actual traffic figures, even imposing very high values for strategic shift.
Table 4.2 lists the airports where this discrepancy between capacity data and actual traffic occurs: the capacity figures reported by the NEST software are compared with the maximum movement counts in radar data. (Movement counts are taken for September 12th, 2014). London Gatwick (EGKK) has a departure count that is constantly above capacity. Nice Cote DAzur (LFMN) has a total capacity that is even lower than departure and arrival capacity at peak hours, and therefore it appears an incorrect value that is largely exceeded. Dusseldorf (EDDL) has a similar issue with total capacity, which is equal to arrival capacity. At Istanbul Ataturk (LTBA) total, departure, and arrival capacity is constantly exceeded over the day. Finally, at Bergamo Orio al Serio (LIME) departure capacity can be largely exceeded by performed departures (8 vs 18).

Airport	Total capacity		Departure capacity		Arrival capacity	
	NEST	Max count	NEST	Max count	NEST	Max count
EGKK	999	56	24	36	30	30
LFMN	22	44	14-27	20	28	24
EDDL	33	52	999	32	33	33
LTBA	58	70	29	38	29	38
LIME	24	24	8	18	16	12

Table 4.2: Declared capacities and maximum hourly movement counts on 12 September 2014

It was therefore necessary to adjust the capacity figures from the data for these airports. The following rationales were applied: since London Gatwick (EGKK), Nice Cote dAzur (LFMN), and Dusseldorf (EDDL) are coordinated airports, meaning that they have airport slots assigned, capacity information was corrected with slot figures. Since London Gatwick (EGKK) has a different number of allowed arrivals/departures for each hour, we used maximum values for summer 2014. Istanbul Ataturk (LTBA) and Bergamo Orio al Serio (LIME) are non-coordinated airports, therefore needing alternative solutions. For Istanbul Ataturk (LTBA), since the maximum number of actual departures is 70, this value is set as total capacity. Departure and arrival capacities are then set to half this value to equally distribute them. For Bergamo Orio al Serio (LIME), since no limit seems to be applied on departure capacity in reality, the departure capacity constraint was simply removed. Table 4.3 summarises these corrections.

Airport	Total capacity	Departure capacity	Arrival capacity
EGKK	55	35	35
LFMN	40	24	20
EDDL	47	36	33
LTBA	70	35	35
LIME	24	999	16

Table 4.3: Corrected capacities applied

4.2.2 Traffic data

Airlines and airline categories

Matching each flight to the corresponding airline and therefore airline type allows to associate different cost profiles to different flights. This allows to make more realistic assumptions regarding flight costs (e.g., it is reasonable to assume that operational costs for a flight operated by a full service carrier airline are different from those endured by a low cost carrier) and also to differentiate the traffic in terms of costs, otherwise, if estimated operational costs are the same for all flights, it is like assuming that all traffic is operated by one single airline.
It is not always clear in the DDR2 dataset as to which airline is operating the flight. Traffic (EXP2) files contain multiple data fields with airline information (although not all are populated): the callsign, the company, the operating company and the aircraft registration number for which the operator can be checked (although there are potential further complications when codeshare flights are considered).

The University of Westminster, within the project SATURN, developed in-house algorithms to assign airlines to flights, with Table 4.4 providing a range of examples. Many cases are straightforward, such as the first example, where field contents concur. In cases where the operating airline is known (examples two and three), often the operating airline is a subsidiary whereas the callsign airline is a more useful indicator. Similarly, the operating airline may be operating via a wet lease contract (ACMI), hence the callsign airline is more appropriate (example four). Clarifying the airline also helps to determine flights for exclusion, such as military and cargo operators. The fifth example shows a typical passenger aircraft (B752) that is in fact operating a military flight.

Callsign	**Registr.**	**Aircraft type**	**Callsign (first 3 chr.)**	**Company**	**Operat. company**	**Assigned airline**	**Comments**
BAW58L	GCIVC	B744	BAW	BAW	BAW	BAW	All fields show British Airways (BAW)
AUA26	OELPD	B772	AUA	AUA	TYR	AUA	Marketed as Austrian Airlines (AUA), operated by Tyrolean Airways/Austrian Arrows (TYR, a subsidiary)
AF303RV	missing	CRJX	AF3	AFR	BZH	AFR	Marketed as Air France (AFR), operated by Brit Air/HOP! (BZH, a subsidiary)
BER334U	DANFE	AT72	BER	BER	ATV	BER	Marketed as Air Berlin (BER), operated by Avanti Air (ATV, an ACMI airline)
SAM106	980001	B752	SAM	missing	DOD	DOD	USAF Air Mobility Operations (DOD)

Table 4.4: Assigning airline name to flights

Each airline is assigned one of four categories that broadly describe their primary type of operation: full-service, regional, low-cost carrier (LCC) or charter (seeTable 4.5). Airlines providing services covering more than one category (e.g. charter airlines offering full scheduled services; regional services provided by full-service carriers) are assigned the most appropriate type. All remaining uncategorised airlines are assigned to the regional category since this has a fairly cost-neutral effect on the model. Other non-commercial IFR passenger operators such as all-cargo, military transport and private/business aviation are flagged for exclusion, being out of scope for strategic traffic management (as discussed in more detail in the following section about flight filtering).

Airline type	**Description**
CHT	Charter
FSC	Full-service
LCC	Low-cost carrier
REG	Regional and all remaining uncategorised operators
XXX	Non-passenger operator, including all-cargo, military transport and private/business aviation (flagged for exclusion)

Table 4.5: Assigning airline type to flights

Flights filtering

The DDR2 traffic dataset includes all IFR flights, however not all were in scope for the project, and were therefore excluded from the data files. Exclusions are based on the following criteria (note that some flights may be excluded for multiple reasons):

- Operator types: all-cargo, military, etc. (e.g. AirBridge Cargo, German Air Force) and unknown operators (i.e. coded ZZZ);
- Aircraft types: all-cargo, military, helicopter, etc. (e.g. A124, C130, EC45) aircraft;
- Individual airframes: all-cargo, military, etc. identified by their registration numbers (e.g. HSHRH a Royal Thai Air Force B734, FGIUC an Air France B744 freighter);
- Circular flights: departure airport (ADEP) is the same as destination airport (ADES);
- Unknown airports: either airport (ADEP or ADES) is coded as ZZZZ (i.e. includes Air Filed Flight Plans).

This filtering resulted in 29 539 flights (out of the total 33 810) for the reference day chosen, 12 September 2014. 28 355 of these were passenger flights that were also assigned fine-tuned cost profiles by the University of Westminster within the scope of project SATURN. Cost profiles were assigned as illustrated in Table 4.6.

The 14 ECAC airports (plus the two non-ECAC Moscow airports) listed in Table 4.7, which had over 25 million passengers in 2014, were considered as "hub" airports.

Airline type	Cost scenario	Flights
Charter	base	1806
Charter to hub	base	356
Full Service Carrier	base	8904
Full Service Carrier to hub	high	5063
Low Cost Carrier	low	7080
Low Cost Carrier to hub	low	1584
Regional Carrier	base	3061
Regional Carrier to hub	high	501

Table 4.6: Cost profile assignment to passenger IFR flights on Sep 12, 2014

ICAO Code	Airport Name
EGLL	London/Heathrow
LFPG	Paris Charles De Gaulle
EDDF	Frankfurt Main
LTBA	Istanbul-Atatürk
EHAM	Amsterdam/Schiphol
LEMD	Madrid Barajas
EDDM	München
LIRF	Roma/Fiumicino
EGKK	London/Gatwick
LEBL	Barcelona
LFPO	Paris Orly
LTAI	Antalya
EKCH	Copenhagen Kastrup
LSZH	Zürich

Table 4.7: Hub airports

4.3 Data clustering

The present section illustrates the clustering operations carried out on some classes of data to make them suitable to the modelling requirements. Specifically, clustering was operated on the aircraft model used to perform the flights and on 4-dimensional trajectories. The first operation was necessary for estimating flight operational costs (since reference values were available only for the 15 reference aircraft from the Cook & Tanner (2015) study, see Sec. 4.3.3). Route clustering was necessary to narrow the number of options available for each origin/destination pair down to a reduced number of routing alternatives which significantly differ one from another.

4.3.1 Aircraft clustering

The purpose of aircraft clustering is to ease the assessment of operational strategic costs for each flight by associating each aircraft found in traffic files to the estimated costs for the fifteen most-used aircraft from the Cook & Tanner (2015) report. The report states that the square root of the maximum takeoff weight (MTOW) of the aircraft can be taken as a sound proxy for two aircraft having similar size and operational costs.
All aircraft used in traffic data from the reference period are therefore grouped into fifteen clusters that use the aircraft types from Cook & Tanner (2015) as centroids. Since the number of clusters to expect is known a priori, the K-means algorithm is the preferred clustering algorithm, with the square root of aircraft MTOW as clustering criterion.
The K-means algorithm works as follows. Suppose to have n sample vectors $x_1, x_2, \ldots, x_n$ that should be grouped into k clusters (with k given, $k \leq n$). Let μ_i be the mean of cluster i. Then, x_j belongs to cluster i if the squared distance $||x_j - \mu_i||^2$ is the minimum over all i. In other words, the objective of K-means is to minimise the distances between vectors x_j and cluster centroids μ_i:

$$\min \sum_{j=1}^{n} \sum_{i=1}^{k} ||x_j - \mu_i||^2 \tag{4.1}$$

Basic algorithm steps are described in the following and illustrated in Figure 4.8.

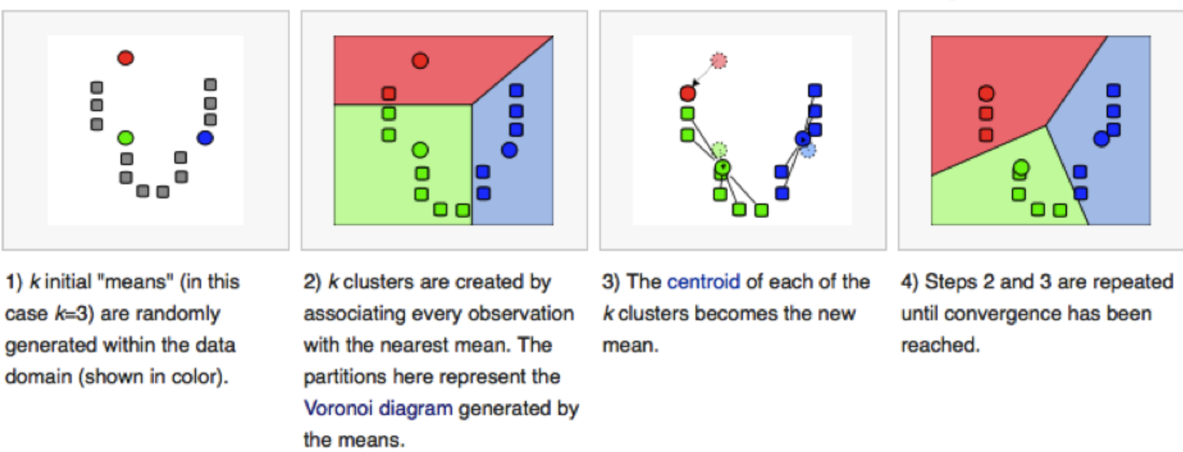


Figure 4.8: K-means algorithm steps. Source: Wikipedia

Table 4.8 exemplifies some cluster affiliation resulting from the application of the K-means algorithm to the square root of aircraft MTOW keeping the fifteen reference aircraft as cluster centroids.

Reference Aircraft (ICAO designator)	**Reference MTOW (source: NEST)**	**Other aircraft in the same cluster ICAO designator (MTOW)**
A319	68.98	B737 (67.88)
A320	74.48	MD83 (72.03); MD88 (72.6); MD90 (73.96)
A321	86.47	B722 (86.27)
A332	229.51	B788 (227.9), B773 (299.37), IL96 (251.33)
AT43	16.83	AT44 (17.8); AT45 (18.6)
AT72	22.15	AT75 (22.8); E135 (22); E145 (21.02)
B733	61.6	B736 (59.29); MD87 (62.36)
B734	65.63	A318 (63.6); MD81 (63.95)
B735	56.55	B732 (53.5), YK42 (56.77)
B738	76.47	B721 (77.27); B739 (78.79)
B744	392.09	A346 (370.78), A388 (548.83)
B752	111.17	B720 (99)
B763	181.81	A332 (229.51); B764 (204.3)
DH8D	29.14	E170 (36.77), B461 (37.49)
E190	49.07	B462 (41.66), RJ1H (45.14)

Table 4.8: Aircraft clustering results (example)

In the actual implementation, the final cluster centroids were desired not to significantly deviate from the initial set (since cost data were available only for the initial fifteen). Therefore, the algorithm was stopped after the first iteration, i.e., before centroids were recalculated. Furthermore, all aircraft with MTOW inferior to 10T were excluded from the clustering, since the operational costs of the smallest of the fifteen, the ATR-43 (AT43), would not have been realistic for these. These aircraft were grouped into a sixteenth cluster denominated OTHR.

Aircraft clustering could be further improved by subdividing aircraft by category (i.e., turboprop, heavy jet etc.). At the moment, for example, clustering on the square root of MTOW only groups medium business jets and turboprops in the same cluster. However, these two aircraft types are known to have very different operational costs in reality. They also have different speed and altitude capabilities turboprops can fly up to FL250, while business jets can fly above FL400 and this might induce non-realistic solutions in route choice models. Aircraft categories and additional data could be sourced from Eurocontrol BADA (Base of Aircraft Data) for improving aircraft clustering.

4.3.2 Route clustering

The use of custom-defined navigation points is quite common in actual air traffic data. While on one hand this allows for greater flexibility in flight planning, on the other hand it can cause issues when trying to identify the available routes between an origin-destination pair. This situation is illustrated by the example described in Figure 4.9. This example shows traffic data from November 18th, 2013 between Paris Orly (LFPO) and Toulouse-Blagnac (LFBO). Two clearly distinct routes can be identified. However, looking at route segments data (Figure 4.10 shows a detailed visualisation obtained in a Geographic Information System, GIS, software), several almost-overlapping routes become visible. Notice that although this example is 2-dimensional, it still applies when altitude is considered.

Figure 4.9: Flights from LFPO to LFBO (scale 1:1440439 dec. deg.)

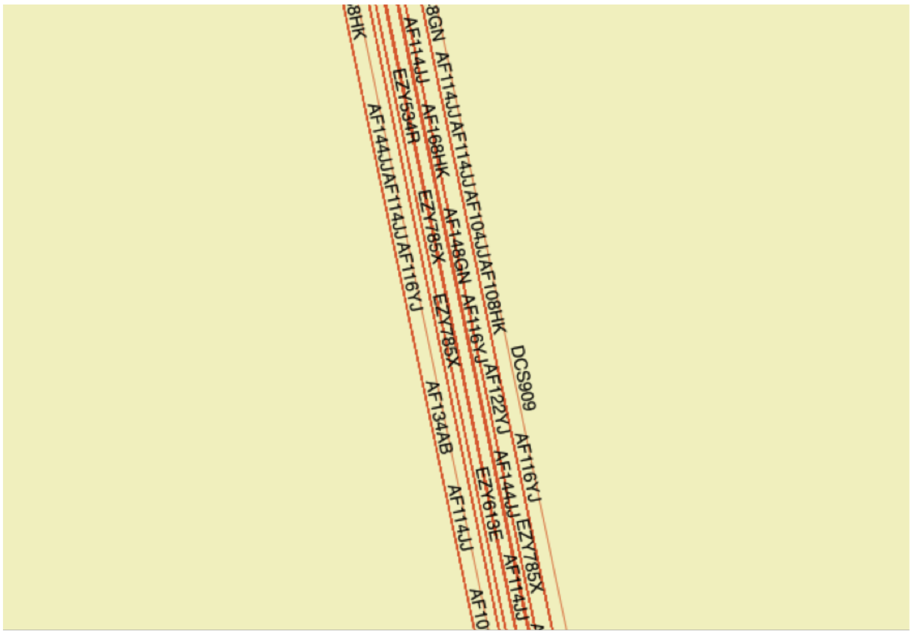


Figure 4.10: Flights from LFPO to LFBO (detail, scale 1:624 dec. deg.)

This is a key issue for using this data to test optimisation models, since these overlapping routes are mostly equivalent in terms of sector entries, flight time, flown distance, and route charges. The rationale for clustering routes is therefore to reduce the complexity of data instances by grouping routes that are geographically close. By doing so, only significantly different routing alternatives are considered.

A metric that is commonly used in literature for measuring the distance between two routes is the so-called *Hausdorff Distance*. Given two sets $S_1, S_2 \in R^3$ (i.e., the points belonging to route 1 and route 2, respectively), their Hausdorff distance is defined as:

$$d_H(S_1, S_2) = \max \left\{ \sup_{p_1 \in S_1} d(p_1, S_2), \sup_{p_2 \in S_2} d(p_2, S_1) \right\} \quad (4.2)$$

That is, the Hausdorff distance is the maximum of the distance between all elements of S_1 and set S_2 and the distance between all elements of S_2 and set S_1. A distance matrix can be built by calculating the Hausdorff distance between all routes connecting an origin-destination pair. Such a matrix can be given as input to a clustering algorithm that will put into the same cluster all routes whose distance is lower than a threshold value T. Not knowing in advance how many clusters will be obtained, we relied on a hierarchical clustering algorithm instead of the K-means algorithm used for the clustering of aircraft.
Hierarchical clustering is typically carried out in two phases:

1. **Linkage:** a distance tree is obtained from the distance matrix of a given origin-destination pair. The final height of the tree represents the maximum distance between two elements of the original set, i.e., the maximum distance between two routes. The result can be visualised as a dendrogram plot, see Figure 4.11;

2. **Clustering:** the threshold value T is used for sorting the elements into clusters. Graphically, it represents the height at which the dendrogram is cut: all elements (i.e., routes) that are still attached to the same sub-tree after the cut belong to the same cluster.

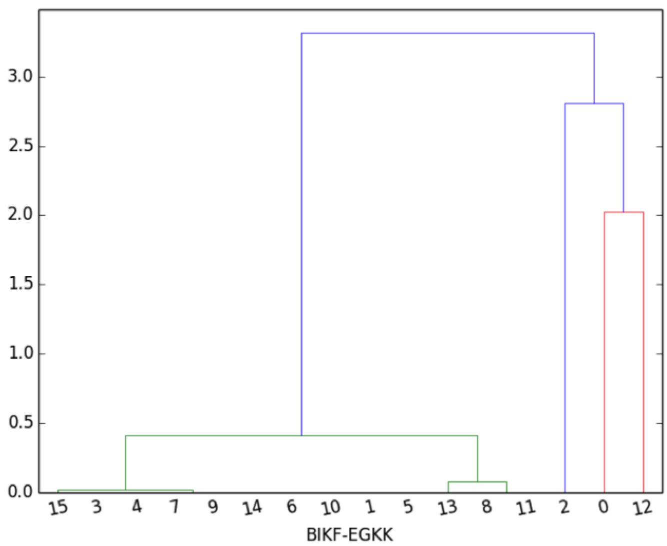


Figure 4.11: Dendrogram of routes between Kevlavik Int. (BIKF) and London Gatwick (EGKK). Routes are listed on the horizontal axis, relative distance (measured in hundreds of kilometres on the vertical axis).

Hierarchical clustering is performed on every origin-destination pair, excluding reserved ICAO codes ZZZZ (i.e., no ICAO code exists for that airport) and AFIL (i.e., flight plan received from an airborne flight).

In the current implementation, a total of 102 504 routes for two weeks of traffic (from September 1st to September 14th, 2014) were obtained. A cut threshold of 0.3 (i.e., routes that are less than 0.3 degrees far from each other are grouped into the same cluster; at the latitude of the ECAC area this corresponds to 20-25 Km approximately) on 3-dimensional geographic representation of routes was used. From previous clustering experiments on larger data sets (up to 3 weeks) and different cut threshold values (0.5 and 0.1), it could be noticed that the number of routes stabilised around 110,000 from a 2-week dataset on, with the added routes being mostly oceanic (hence, heavily dependent on weather changes). Route clustering, in combination with aircraft clustering, allows for further useful data aggregations to be performed, such as:

- Average flight time in sectors (per route/aircraft cluster pair);
- Average entry time in sector after flight departure (per route/aircraft cluster pair);
- Average entry flight level in sector (per route/aircraft cluster pair);
- Average strategic route costs (per route/aircraft cluster pair).

4.3.3 Cost coefficients

The present section illustrates the coefficients used to assess the operational costs of flights: national unit rates for estimating en-route charges and ground and airborne shift cost coefficients for the respective components of each 4-dimensional trajectory.

Unit rates

The country-specific unit rates used for calculating route charges and to estimate ANSP revenue in the pricing models were the values applied in September 2014, sourced from Eurocontrol's website (with the exception of Estonia and Ukraine, that were included for geographic reasons and whose unit rates were sourced from the respective ANSPs' website) and illustrated in Table 4.9.

Country	ICAO code	Unit rate (€)	Country	ICAO code	Unit rate (€)
Portugal S M	AZ	10.6	France	LF	65.92
Belg.-Luxembourg	EB	72.19	Greece	LG	34.68
Germany	ED	77.47	Hungary	LH	40.14
Estonia	EE	23.97	Italy	LI	78.98
Finland	EF	52.21	Slovenia	LJ	67.61
United Kingdom	EG	88.54	Czech Republic	LK	43.26
Netherlands	EH	66.62	Malta	LM	27.76
Ireland	EI	30.77	Austria	LO	73.54
Denmark	EK	71.43	Portugal Lisboa	LP	38.89
Norway	EN	52.24	Bosnia Herzeg.	LQ	51.46
Poland	EP	35.33	Romania	LR	38.34
Sweden	ES	69.71	Switzerland	LS	100.41
Latvia	EV	28.59	Turkey	LT	32.12
Lithuania	EY	45.92	Moldova	LU	38.05
Spain Canarias	GC	58.51	FYROM	LW	60.1
Albania	LA	45.58	Serb.-Montenegro-KFOR	LY	46.83
Bulgaria	LB	37.53	Slovak Republic	LZ	61.08
Cyprus	LC	38.56	Armenia	UD	34.24
Croatia	LD	43.11	Georgia	UG	24.85
Spain Continent.	LE	71.84	Ukraine	UK	45.56

Table 4.9: Route charges unit rates from September 2014 (Source: Eurocontrol)

Aircraft ground and airborne shift costs coefficients

In order to estimate the operational costs of flights, the reference values illustrated in Cook & Tanner (2015) were adopted and applied to the clustered aircraft and routes. The report provides per-minute coefficients for the costs of fuel, aircraft maintenance, staff and fleet, for each of the reference aircraft and according to three cost profiles, (low, base and high, see Sec. 4.1.3).

The reference values for strategic delay were used to calculate the cost of ground shift (summing up the components for fleet, crew and maintenance and multiplying by the amount of minutes of ground shift assigned to the flight), while the values given for tactical delay were applied for the airborne portion of the flight (summing up fuel, crew and maintenance costs and multiplying by the duration of the airborne portion of the flight). The profile was assigned on a per-flight basis where these data were available (see Sec.4.2.2); where this was not the case, the Low profile was assigned as default.

The applied coefficients are illustrated in Table 4.10.

Reference aircraft	Weight factor	Cost profile	Ground shift cost coef. (€/min)	Airborne shift cost coef. (€/min)
A319	1.154	L	7.42	39.97
		B	12.91	57.17
		H	22.35	80.7
A320	1.213	L	8.06	41.32
		B	14.26	58
		H	24.99	84.84
A321	1.315	L	9.96	49.29
		B	16.96	68.92
		H	29.09	96.31
A332	2.142	L	16.07	88.92
		B	28.15	120.29
		H	52.54	174.5
AT43	0.58	L	2.18	9.69
		B	3.87	15.02
		H	6.7	22.42
AT72	0.665	L	3.14	13.81
		B	5.51	21.25
		H	9.7	31.59
B733	1.099	L	4.53	38.1
		B	8.58	53.28
		H	15.01	72.98
B734	1.145	L	4.98	39.15
		B	9.58	54.88
		H	16.94	75.68
B735	1.051	L	4.14	35.31
		B	8.14	49.16
		H	14.51	67.4
B738	1.205	L	8.53	43.12
		B	15.97	60.81
		H	28.68	88.82
B744	2.802	L	14.3	148.12
		B	24.65	182.46
		H	44.76	233.75
B752	1.464	L	6.71	51.88
		B	11.75	70.12
		H	19.95	94.07
B763	1.901	L	12.07	76.72
		B	21.45	103.78
		H	41.13	151.16
DH8D	0.763	L	5.2	18.49
		B	8.71	27.24
		H	15.23	40.22
E190	0.991	L	7	31.93
		B	12.34	45.79
		H	21.99	65.81
OTHR	0.522	L	1.97	8.72
		B	3.49	13.52
		H	6.03	20.18

Table 4.10: Cost coefficients for reference aircraft (Source: Cook, 2015)

4.4 Development of a bespoke geographic database for air traffic data

NEST is a network analysis and design tool exclusively meant to support non-automated tasks. Currently, it cannot be interfaced to other software for either data input or output, and it offers no option for simulating different pricing rules or policies. It is a valuable tool for analysing specific issues or items and for producing related reports but is of little use for both testing pricing models and for automated knowledge discovery tasks. In order to overcome such limitations a geographical database was built from DDR2 and NEST generated data as a support tool for the SATURN project. The rationale for building such a database, which only uses openly available software, was threefold.

1. Due to the heterogeneity in file format of NEST exported and/or DDR2 files, it is easier to obtain the data instances needed for the tests by querying a database than by parsing multiple input files per model.

2. Data analysis or statistics tasks can be performed with a high level of automation. Tasks such as ranking the busiest days or sectors over a whole year, for example, can be performed automatically with a single query. The same tasks require manual count or other operations if carried out within NEST. The geographical extension allows all operations concerning geography and intersections to be easily performed, similarly to NEST, with perhaps a higher level of flexibility.

3. The expected output from SATURN project was evidence that the proposed models can improve traffic distribution within the European airspace. Therefore both statistics and data visualisations need to be provided. Again, NEST is a valuable tool for visualisations and reporting but may not be the most straightforward solution for visualising models output. Within this context, Geographic Information System (GIS) software like QGIS 2, using coordinate system WGS-84 and map projection EPSG:4326 (Figure 4.12), proved to be a simpler and more flexible solution for visualising the output of the pricing models tested. Most GIS software can directly access and retrieve data from geographical databases and can access data from other software via application programming interfaces (APIs) functions, thus resulting in a very flexible solution.

Such a database was built from DDR2 data, except for unit rate values that were downloaded from Eurocontrol's CRCO website. Due to the heterogeneity of formats in the input files, each file type was parsed through a different script written in the Python programming language that extracted relevant data and inserted it into the database. The database is currently made of 26 tables, plus four views (permanent stored queries that behave like virtual tables, i.e. can be queried like other tables). All tables and views are presented here as grouped into Airspace-related, Network-related, Traffic-related and post-processed categories for description purpose only.

4.4.1 Database tables and data sources

The general structure of the bespoke database and original data sources are illustrated in the current section. A detailed documentation of each table and data fields is provided in Appendix A; furthermore, some database usage examples are presented in Appendix B.

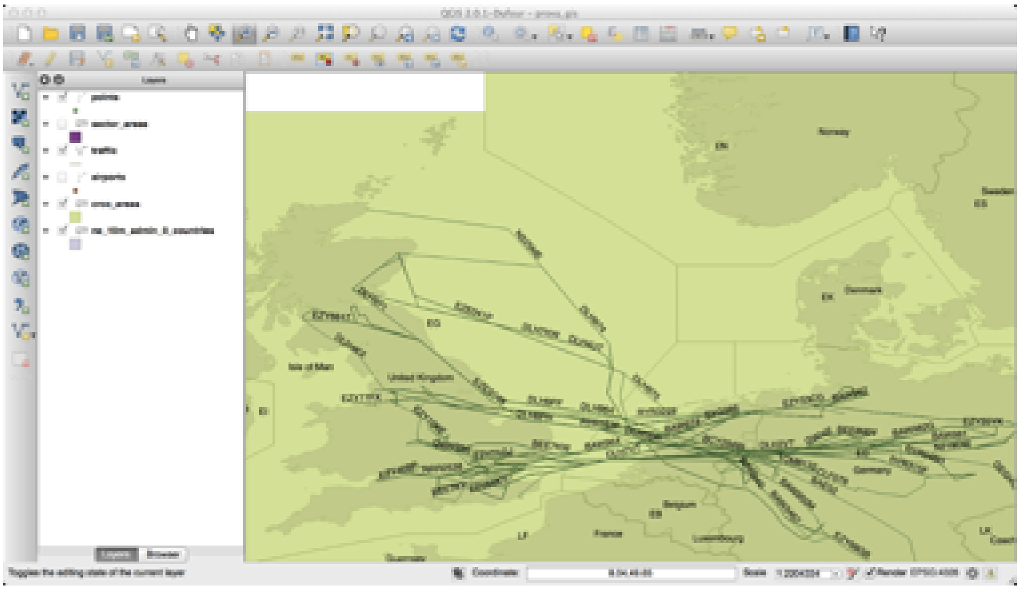


Figure 4.12: Viewing flight data in QGIS 2

Airspace-related tables

This set includes all tables that are related to sector structure, characteristics, configurations and openings (Table 4.11).

DB table name	**Data source file**	**Origin**	**Origin**
sector_areas	Airspace.spc and Sectors.gsl	NEST	Name, type (ES,CS, FIR etc.) and coordinates of airspace blocks
crco_areas	view on sector_areas table	sector_areas table	National airspaces belonging to CRCO area
sector_structure	Airspace.spc and Airblock.gar	NEST	Describes sector hierarchical relationships e.g. ES sector X is part of CS sector Y
sector_capacity	Capacity.ncap	NEST	Nominal hourly sector capacities
sector_regulations	RegPlan.nreg	NEST	Tactical variations on sector capacity
acc_configurations	Configuration.cfg	NEST	Describes different possible sector configurations
acc_configurations_openings	OpeningScheme.cos	NEST	Activation time of configurations
military_openings	AIRAC .mot files	NEST	Opening times of military sectors
unit_rates	ur.txt files	Eurocontrol website	Monthly adjusted CRCO unit rates per country

Table 4.11: Airspace-related tables

Network-related tables

This set includes data on airports and navpoints such as geographical coordinates, regulations and existing segments (Table 4.12).

DB table name	Data source file	Origin	Description
points	.nnpt and .narp files	NEST	Navpoint and airport names and coordinates
navpoints	View on points table	points table	Navpoint names and coordinates
airports	View on points table	points table	Airport names and coordinates
airport_capacity	.ncap	NEST	Nominal hourly airport capacities
airport_regulations	RegPlan.nreg	NEST	Tactical variations of airport capacities
segments	.ase	NEST	Existing links between couples of points

Table 4.12: Network-related tables

Traffic-related tables

This set comprises of all tables related to traffic, such as flight times, routes, and aircraft type (Table 4.13).

DB table name	Data source file	Origin	Description
traffic	.exp2 M1 and .so6 M1 files	NEST	General flight data i.e. departure and destination airports and times, route coordinates
traffic_det	SO6 M1	NEST	Segment-by-segment flight data
aircraft	Helicopter.txt AircraftWeight.txt	NEST	Aircraft name, type and weight

Table 4.13: Traffic-related tables

Clustered data tables

These tables are dedicated to processed data, more specifically, to the data obtained from the clustering operations (Sec. 4.14) or other elaborations on data from other tables. The column "Data origin" in Table 4.14 therefore indicates the table(s) from which the clustered data was sourced, rather than a specific DDR2 file.

DB table name	Data origin (table names)	Description
aircraft_clus	traffic, aircraft	Aircraft name and corresponding aircraft cluster name
flights	traffic, aircraft_clus, routes	Aircraft cluster and route cluster originally chosen by each flight.
routes	flights, aircraft_clus	route name and associated origin and destination airport.
routes_det	flights, intersections, aircraft_clus, routes	Route cluster/sectors intersections and corresponding average entry time, flight time, entry flight level with each aircraft cluster.
routes_navpoints	flights, traffic_det, aircraft_clus, routes	Route cluster/waypoint intersections and corresponding average arrival time and flight level with each aircraft cluster.
tv_rules_capacities	traffic_volumes, tv_flows, tv_flows_det, sector_capacity, intersections_actual	Adjusted hourly capacity values for Traffic Volumes with inconsistent original figures (as explained in Sec. 4.2.1).

Table 4.14: Clustered data tables

Chapter 5

Computational experiments

This chapter illustrates the solving approaches attempted to tackle the cPLP model presented in Chapter 3. The data instances used for the tests presented here are the same used also for the project SATURN. Since the objective of the project was to test the effectiveness of different pricing models on data instances obtained from actual traffic, these data instances are of a much larger size than those commonly presented in literature for testing bilevel models.

While the data preparation was a challenging task by itself, as described in Chapter 4, cPLP resolution was also not straightforward. The reformulation carried out to linearise the model presented in Chapter 3 (equations 3.40-3.53) introduces four constraints for each combination of viable route and departure time per flight (plus one for the integrality of the $x^{(m)}_{f,r}$ variables). The variables are in a high number themselves and therefore the size of the problem grows at a very fast rate. This is a well known issue with bilevel problems; in fact specific literature generally presents experiments run on data instances with at most one hundred users, which is at least an order of magnitude smaller than the testing instances of air traffic prepared for SATURN (this is further discussed in Sec.5.2).

While exact resolution of entire days of traffic on the whole network (around 30 000 flights, that produce a problem with approximately 700 000 integer and binary variables and almost 3 millions constraints) is clearly out of reach for exact resolution via standard optimisation techniques like branch and bound, tests have been successfully run on regional-size instances, specifically up to 5 000 flights on a single state.

A precise analysis of the computational complexity of the cPLP problem falls beyond the purpose of this work. However, all the pricing schemes illustrated in section 3.2.7 make the cPLP a hard to solve combinatorial optimisation problem, even when tested on a reduced scale, for example at regional level. The only exception is provided by the cPLP-one scheme applied to a single state, where only one rate has to be set for the whole problem. This is consistent with the Network Pricing Problem, that is proved to be $\mathcal{NP}$-Hard for all variants that require more than one price to be set in the whole problem. On the contrary, the one rate variants are solvable in polynomial time. One relevant example is illustrated in Castelli et al. (2013), where the bilevel Network Pricing Problem is applied to find the optimal unit rate value that maximises revenues for one ANSP. In this work the authors prove that the single rate case NPP is equivalent to the single-tolled arc NPP presented in Labbé et al. (1998), which was proved to have polynomial complexity, and adapt the efficient solving algorithm to the ANSP revenue maximisation case.

For all the cPLP variants different from regional cPLP-one, no efficient solving procedure exists even for the one-state case and the problem is computationally hard to solve, to the point that the branch and bound algorithm was not even able to find any integer solution on regional instances with one hour of traffic (approximately 600 flights) after two hours of execution.

To tackle these complex variants, some improvements were done on the model formulation and some simple heuristics were implemented to provide good starting solutions to the branch and bound algorithm. These allowed the solver to apply better cuts to the problem and find further solutions, but was still unable to solve to optimality the majority of the data instances within a reasonable time. This is also described in detail in Sec.5.2.

However, the simple case proved to be useful for gaining further insights on the problem. These are described in the the remainder of the chapter, organised as follows: Section 5.1 is dedicated to the calibration of the weights of the objective function components (values of K_1, K_2 and K_3 parameters in eq. 3.8). Section 5.3 analyses in detail the results obtained from applying the cPLP variants with one and two modulated rates on an 8 hour traffic instance. Heavy demand-capacity imbalances occur in this traffic sample, which specifically includes the traffic flying from, to or through French airspace between 6:00 and 14:00 (UTC time) on the reference day, September 12, 2014. Finally, Section 5.4 briefly illustrates the results obtained by a colleague within the project SATURN by solving a variant of the cPLP via a software for multiobjective optimisation that implements resolution via a Genetic Algorithm heuristic. This method does not solve the problem to optimality, but can provide many solutions that are equivalently good, from an optimisation point of view, and therefore represents a valuable tool for pointing out recurring patterns or tradeoffs among the solutions. This solving approach was applied to larger data instances than the ones used for the tests presented here, and the results obtained for the entire September 12 day on the whole European network are presented. Most observations drawn from these results confirm what pointed out for the regional case. Additionally, this approach allows to assess the effectiveness of the cPLP policy at European level, which is not possible when only one state is considered.

With the exception of Section 5.4, all tests presented in this chapter were run on a Intel i7 dual core laptop computer with 8 GB of RAM, running the Xpress solver version 7.8 64 bit (optimiser version 27.01.02). The algorithms were coded in the Mosel language, proprietary to the optimiser, version 3.8 64 bit.

5.1 Objective parameters setting

Several parameters are present in the cPLP model illustrated in Chapter 3, whose values have to be set prior to resolution. The most important ones are the peak threshold, that is, the percentage of capacity occupancy above which a sector is considered to be in a *traffic peak*, and the relative weights of the three objective function components (total shift, capacity violations and revenue neutrality violation, see eq. 3.8). These are used to weight quantities that are not directly comparable with one another, namely minutes, flights and money. It is therefore necessary to scale these quantities in order to represent a realistic tradeoff among them; at the same time, the weights should be chosen so that they do not cause numerical issues in the computational experiments.

Performing a rigorous parameter calibration falls far beyond the exploratory purpose of the present work; nevertheless, some tests were carried out through a more intuitive, although simplistic, approach, with the objective of exploring whether a stable (and realistic) configuration for the problem parameters exist and possibly gaining useful insights on further aspects of the model.

To this purpose, four medium scale traffic samples with known demand-capacity imbalances are used as testing instances. These samples represent eight hours of traffic in French airspace, specifically all flights departing, arriving or flying through France between 6:00 and 14:00 UTC from Monday, September 8, 2014 until Friday, September 12, 2014 (excluding Sep. 11 due to issues with incomplete data). The main characteristics of these traffic samples are summarised

Instance	Number of flights	Peak threshold = 0.5		Peak threshold = 0.7	
		% of peak sectors	n. of peak sectors	% of peak sectors	n. of peak sectors
12-Sep-14	4695	14	172	9	110
10-Sep-14	4476	9	110	6	74
09-Sep-14	4502	8	101	6	76
08-Sep-14	4663	9	110	6	74

Table 5.1: Characteristics of the testing instances

in Table 5.1. A simple unit rate modulation scheme is applied to this sample, where a congestion charge is added to the historical unit rate, that is, the *cPLP-one* scheme (see Sec. 3.2.7) for flying through peak sector-hour pairs (i.e., where and when demand-capacity imbalance might occur). Two values for the peak threshold were tested, namely 0.5 and 0.7, meaning that sectors with 50% (respectively, 70%) and above of occupied capacity in historic data are assigned the peak rate.

As previously mentioned, this pricing scheme is a particularly easy to solve case of cPLP; in fact, having only one price to calculate, the problem is solved in a few seconds even for data instances with several thousands of flights. The results shown in this section are obtained from a modified version of the H3 heuristic (illustrated in the following, Sec. 5.2.2), where all integer values of the peak rate between 66 and 166 (i.e., rounded value of the historical unit rate, 65.92€ with an increment up to 100€) were tested to obtain the corresponding value of the objective function.

The following Section 5.1.1 illustrate the variation in optimal solutions obtained from variating the revenue neutrality weight K_1 in 3.8; similarly, Section 5.1.2 shows how the choice of the sector violation weight affects the optimal solution obtained from the model. The airport capacity violation weight K_3 has been set to a value of one to better study the impact of route charges modulation on en route traffic redistribution with an even condition of airport-operations-related issues.

5.1.1 Shift and revenue neutrality violation (K_1 parameter)

A variation of a few Euros in a unit rate causes ANSP revenue variation in the order of the hundred thousands over a few hours. Therefore the revenue neutrality violation value (ϵ in the cPLP model, see eq. 3.9) should be scaled to a order of magnitude that allows the model to find different rate values that improve the baseline solution (where all rates are set as equal to the historical value, and therefore the revenue neutrality is equal to zero by definition). To do so, the violation ϵ is scaled by a weight K_1 in the Central Planner's objective function (eq. 3.8).

Different values of the K_1 parameter were tested, specifically negative powers of ten ranging from $1.00E^{-07}$ to $1.00E^{-02}$. The sector violation weight (K_2) was set to 60 and the airport capacity violation parameter (K_3) to 1. The results are illustrated in Table 5.2 for the September 12, 2014 data instance, for peak threshold set to 50% and 70% of nominal capacity (results tables for the other days are provided in Appendix C Sec. C.1.1). Additionally, Figures 5.1, 5.2, 5.3 and 5.4 illustrate the results in terms of variation of objective function value associated to the optimal solution found for each tested value of K_1 in the four analysed traffic samples.

The results show that values of $K_1 \geq 1.00E^{-2}$ make the revenue neutrality component in the objective function predominant on the other two and do not allow any price modulation; this can be observed since for this value of K_1 and above the optimal cPLP solution is equal to the baseline one for all instances. In one case (September 10) this happens already with

12-Sep-14						
Baseline solution						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Unit rate (€)
1.00E-07	19468.2	6374	218	6	8.18E+07	65.92
1.00E-06	19541.8	6374	218	6	8.18E+07	65.92
1.00E-05	19460	6374	218	6	0	65.92
1.00E-04	19460	6374	218	6	0	65.92
1.00E-03	19460	6374	218	6	0	65.92
1.00E-02	19460	6374	218	6	0	65.92
cPLP-one solution, peak threshold = 50%						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Peak rate (€)
1.00E-07	17309.9	8956	139	5	8.93E+07	95.92
1.00E-06	17390.3	8956	139	5	8.93E+07	95.92
1.00E-05	17309.3	8956	139	5	825894	95.92
1.00E-04	17383.6	8956	139	5	825894	95.92
1.00E-03	18126.9	8956	139	5	825894	95.92
1.00E-02	19460	6374	218	6	0	65.92
cPLP-one solution, peak threshold = 70%						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Peak rate (€)
1.00E-07	16968.5	8314	144	6	8.45E+07	92.92
1.00E-06	17044.5	8314	144	6	8.45E+07	92.92
1.00E-05	16965.1	8314	144	6	512413	92.92
1.00E-04	17011.2	8314	144	6	512413	92.92
1.00E-03	17472.4	8314	144	6	512413	92.92
1.00E-02	19460	6374	218	6	0	65.92

Table 5.2: Solution variation for different values of the K_1 parameter; results for France data instances for September 12, 2014

$K_1 = 1.00E^{-3}$. At the other side of the tested range, with $K_1 = 1.00E^{-6}$ and below, quite predictably, numerical instability problems appear in all tested instances, even in the baseline solutions (where the revenue neutrality violation term, that should always be equal to zero, is not zero). With values of K_1 between $K_1 = 1.00E^{-5}$ and $K_1 = 1.00E^{-4}$ on the contrary the solutions appear to be both flexible (i.e., optimal solutions have price modulations) and stable (i.e., no evident numerical issues arise).

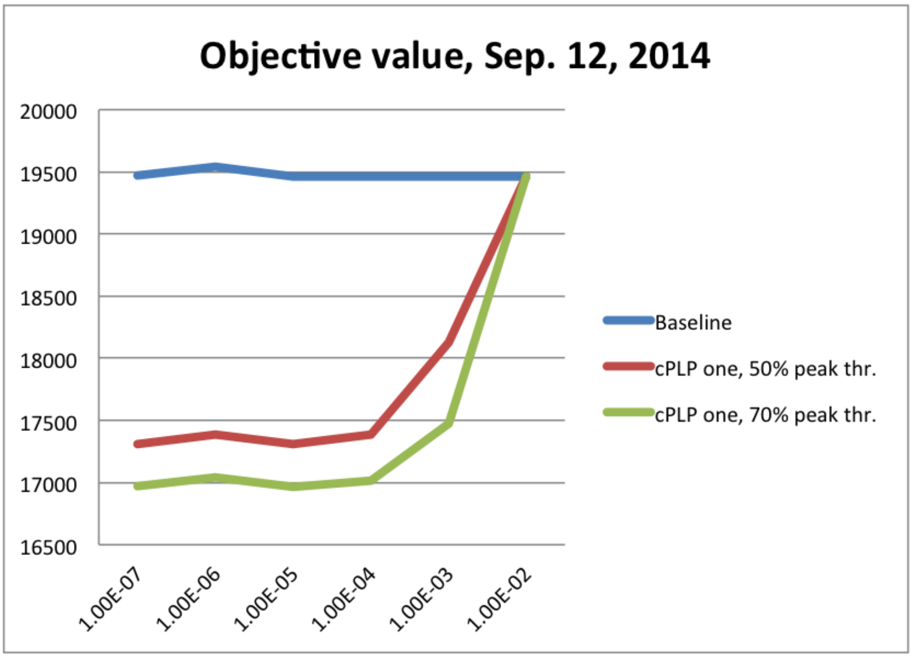


Figure 5.1: Objective value variations for different values of the K_1 parameter with the September 12 data instance

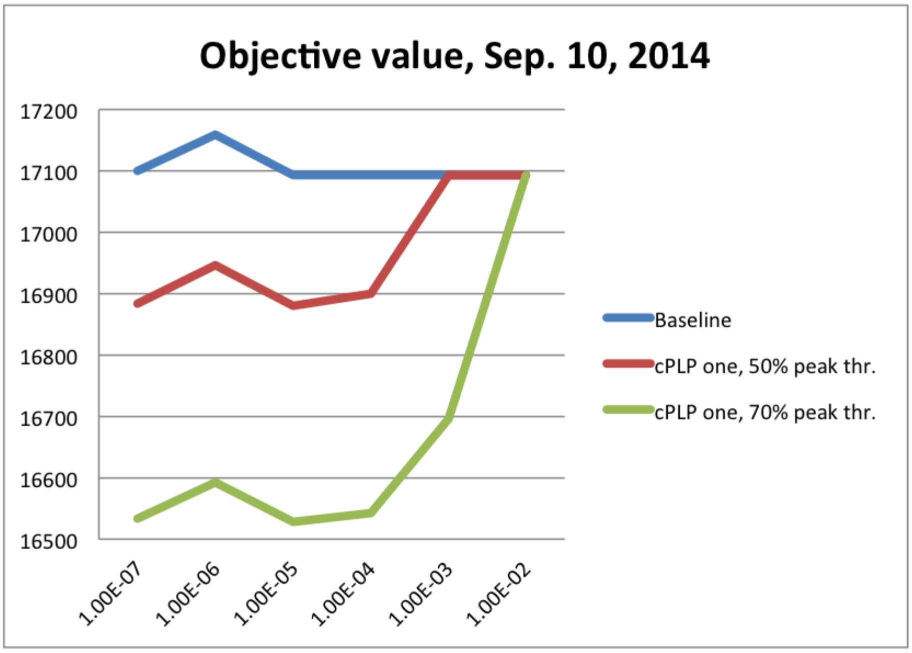


Figure 5.2: Objective value variations for different values of the K_1 parameter with the September 10 data instance

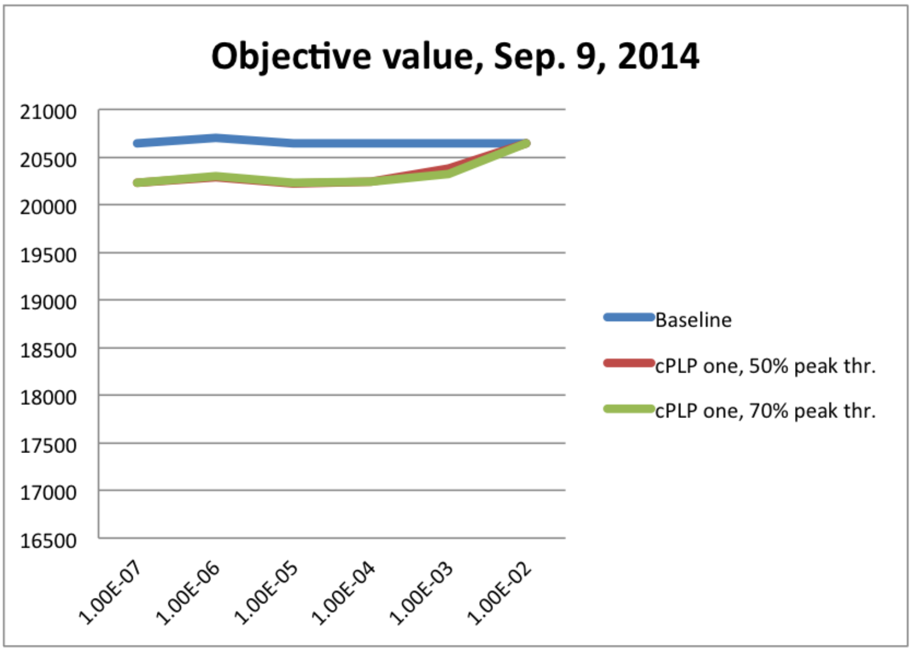


Figure 5.3: Objective value variations for different values of the K_1 parameter with the September 9 data instance

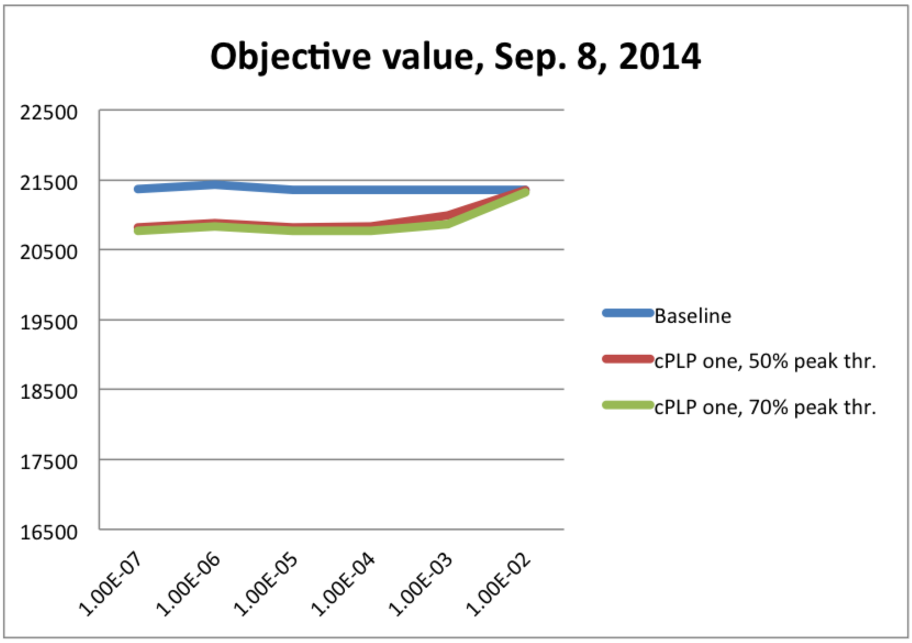


Figure 5.4: Objective value variations for different values of the K_1 parameter with the September 8 data instance

5.1.2 Shift and capacity violations (K_2 parameter)

This section investigates the issue of shift-to-capacity ratio, that is, how a variation in the K_2 parameter in the cPLP model formulation affects the solution and whether the solution stabilises from a certain value of K_2 on.

In this experiment, the revenue neutrality component in the model objective (K_1 parameter in eq. 3.8) is discarded by assigning a very low value to its weight ($K_1 = 10^{-5}$). The four days traffic sample is solved with values of the K_2 parameter ranging from 1 to 100 and corresponding optimal solution values are compared. Note that this variation affects sectors only. For airport capacity (K_3 parameter) the violations have a parameter value equal to 1.

Figures 5.5, 5.6 and 5.7 illustrate the variation obtained with a peak threshold of 50% of the nominal sector capacity in the three components of the objective function: peak rate, global shift and number of sector capacity violations. Figures 5.8, 5.9 and 5.10 show the same results obtained by setting a peak threshold of 70%. The tabular version of these results is provided in Appendix C, Sec. C.1

Results from the tests with the 50% peak threshold clearly show that the solution stabilises with values of $K_2 \geq 60$. This trend does not show in the experiments with peak threshold set to 70%. This is probably due to the fact that the higher threshold reduces the number of sector-hours in peak and therefore values higher than 100 are necessary for the K_2 parameter in order to reach a stable tradeoff with the shift component. It is also worth observing that in all experiments the optimal value of the peak rate, despite being set an upper bound of 165.92€ (100€ more than the historical value of the French unit rate, 65.92€), takes much lower values, namely between 69 and 96 in the 50% peak threshold case (see Figure 5.5) and between 66 and 103€ in the 70% case (see Figure 5.8).

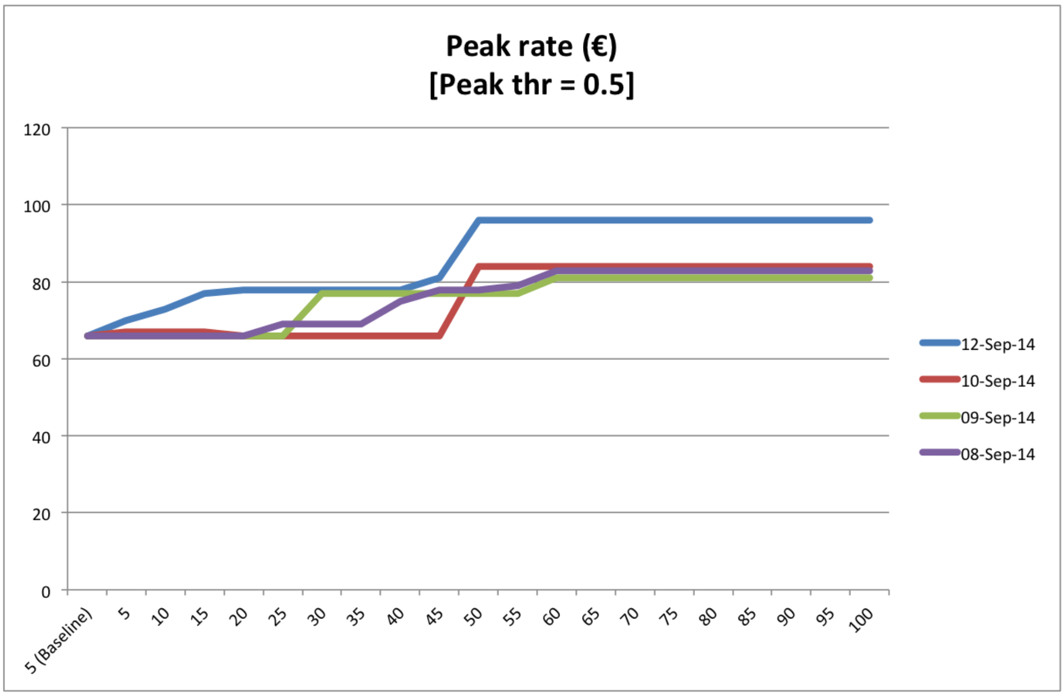


Figure 5.5: Comparison of solution variation in peak rates for different values of the K_2 parameter over four days of traffic on France with peak threshold set to 50% of nominal capacity

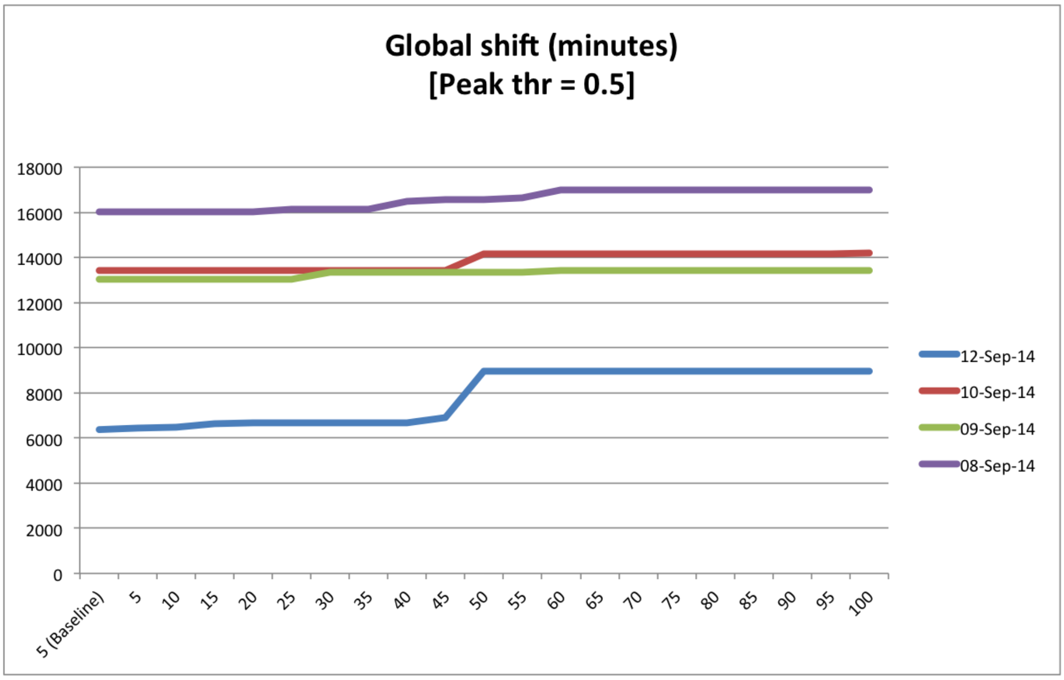


Figure 5.6: Comparison of solution variation in total shift for different values of the K_2 parameter over four days of traffic on France with peak threshold set to 50% of nominal capacity

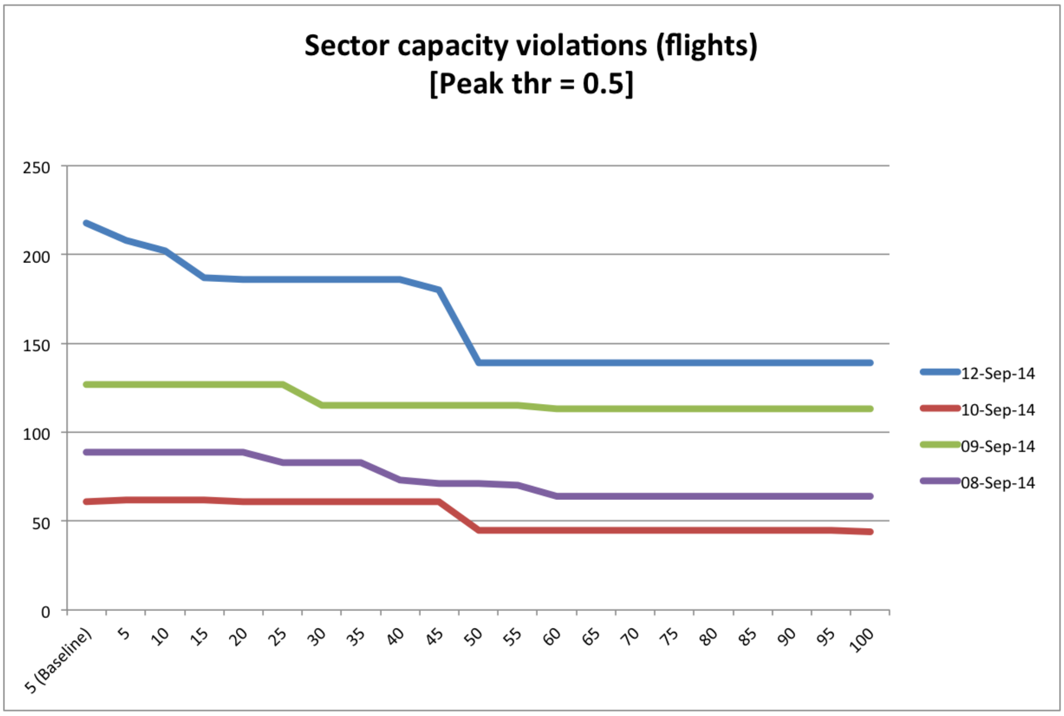


Figure 5.7: Comparison of solution variation in violated sector capacities for different values of the K_2 parameter over four days of traffic on France with peak threshold set to 50% of nominal capacity

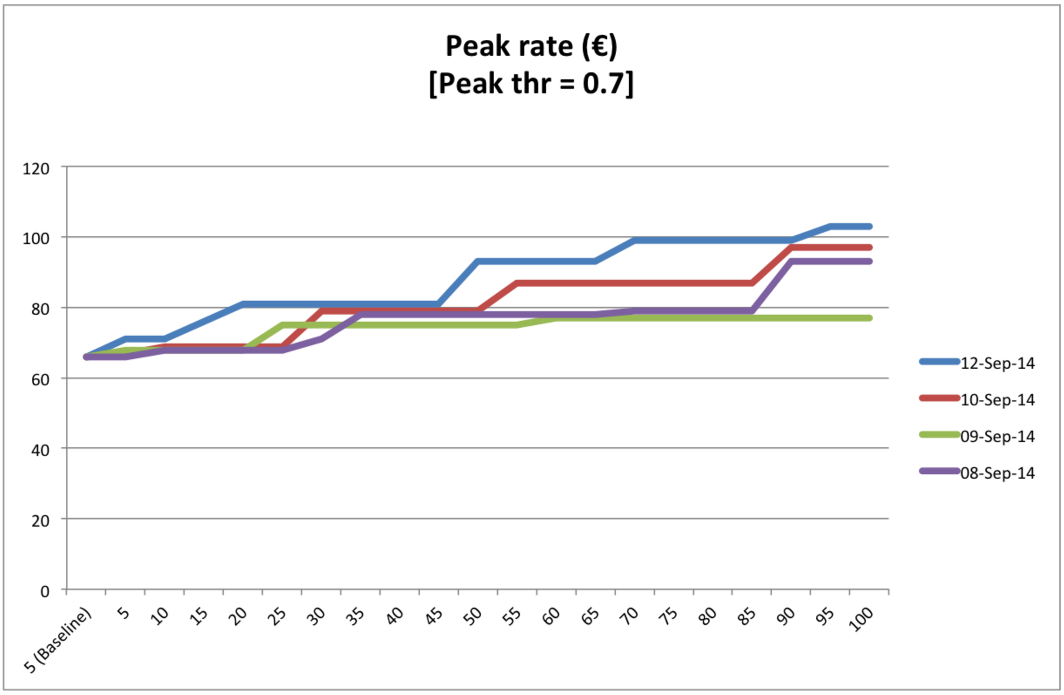


Figure 5.8: Comparison of solution variation in peak rates for different values of the K_2 parameter over four days of traffic on France with peak threshold set to 70% of nominal capacity

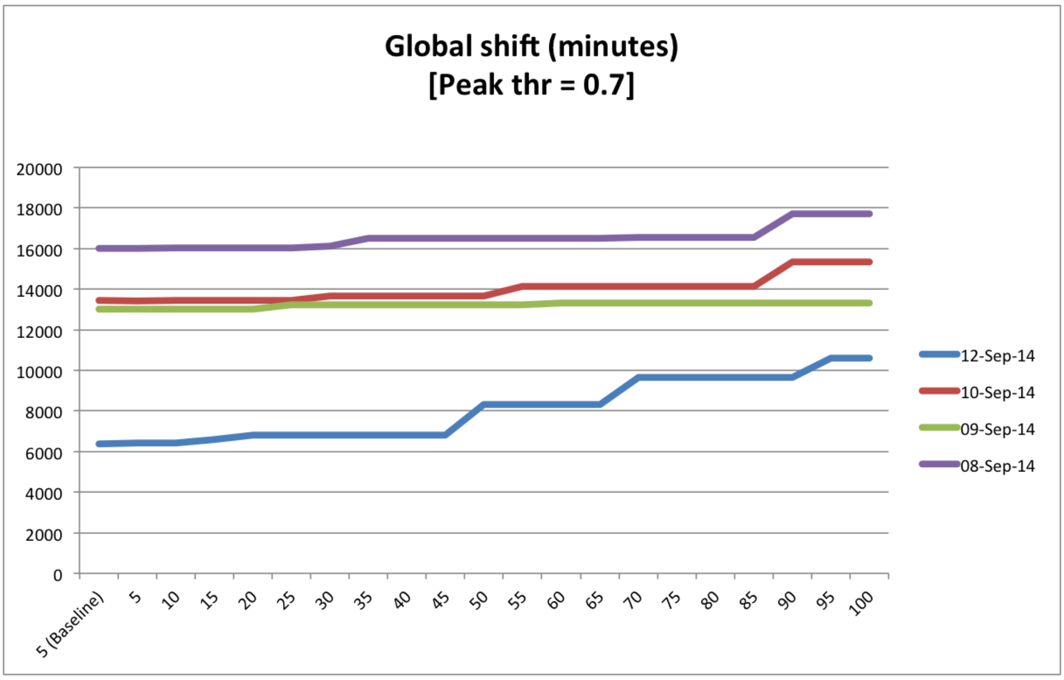


Figure 5.9: Comparison of solution variation in total shift for different values of the K_2 parameter over four days of traffic on France with peak threshold set to 70% of nominal capacity

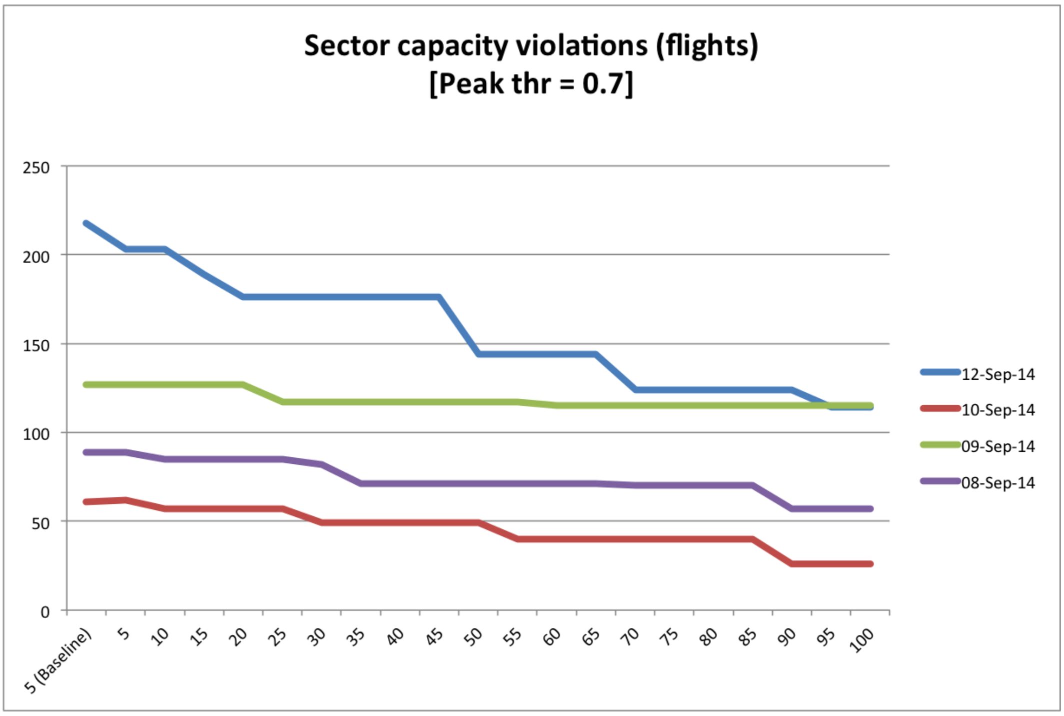


Figure 5.10: Comparison of solution variation in violated sector capacities for different values of the K_2 parameter over four days of traffic on France with peak threshold set to 70% of nominal capacity

5.2 cPLP model resolution

As previously mentioned in Chapter 3, bilevel problems are, in general, difficult to solve, due to the high number of variables and constraints resulting from the linearisation process. Literature on Network Design Problem and Network Pricing Problem, two well studied bilevel problems that share several characteristics with the cPLP presented here, although lacking the temporal dimension (i.e., followers' variables only consider alternative routes instead of routes and time shift), typically report tests run on networks with no more than a hundred users (see for example Violin 2014, Heilporn 2008). Literature on Air Traffic Flow Management models, on the other hand, that share with the cPLP the same definition for the airspace users' variables (i.e., alternative routes and departure times are considered) but lack the bilevel structure, usually presents result obtained from regional size tests (i.e., one state) with a few thousands flight (for example, data instances used in Bertsimas et al. (2011) are randomly generated instances of around six thousands flights).

The European airspace network comprises 40 national airspaces (countries belonging to the CRCO system plus Estonia and Ukraine for geographic reasons) with an average of 600 airspace sectors active every hour out of the average 1200 activated during the day. One day of traffic typically sums up to around 30 000 flights, including national, international and intercontinental traffic. Solving the ANSP-based cPLP (the variant with a peak and an off-peak rate per country) through a commercial solver software generates a problem with around 200 000 constraints and 35 000 integer and binary variables for one hour of traffic; the whole-day September 12 data instance produces a problem with 2 847 848 constraints and 702 701 binary and integer variables.

Therefore, testing the bilevel cPLP on instances obtained from real traffic data posed several challenges: the problem size is remarkable even for relatively small data samples; at the same time testing traffic samples with a few tens of flights hardly allows to draw realistic considerations on the effectiveness of the proposed pricing policy. In addition, the data itself posed a further challenge: many routing alternatives have very similar costs. This is well illustrated by Figure 5.11, that shows the shape of the ANSP-based cPLP objective function for fixed values of the off peak rate and increasing values of the peak rate (i.e., the cPLP-one scheme). The values are obtained from regional traffic samples of 4 000 - 5 000 flights for four days (the same used in Sec. 5.1, with parameters set to $K_1 = 1.00E^{-5}$, $K_2 = 60$ and $K_3 = 1$). At macroscopic level, these functions clearly have a convex shape and have a minimum, but they are not smooth and present a lot of local minima. Furthermore, all curves present a wide region of very similar solutions: September 12, that was the fourth most trafficked day of 2014 (and for which differentiated costs profiles per flights are available, see Sec. 4.2.2, which was not the case for the other days) exhibits a steeper curve, but the other, less trafficked days show a wide region where the objective function is almost flat.

This multitude of solutions with very close objective function values heavily affects the efficiency of the branch and bound algorithm implemented by the solver, which is not able to efficiently prune the solution tree with any of the tested pricing strategy, with the exception of the simplest case of a single rate on for one state. Fine tuning some parameters of the optimiser, such as tree exploration strategy or adding more of the default cuts did not bring any major change. Preliminary tests showed that the plain implementation of the cPLP model as formulated in Chapter 3 is not able to find a first integer solution on data instances of one hour of traffic even letting the execution run for as long as six hours.

Several improvements were attempted to speed up the resolution. Section 5.2.1 illustrates some bounds applied to variables and simple improvements to the model formulation. Section 5.2.2 describes some simple heuristics that were implemented to find good initial integer solu-

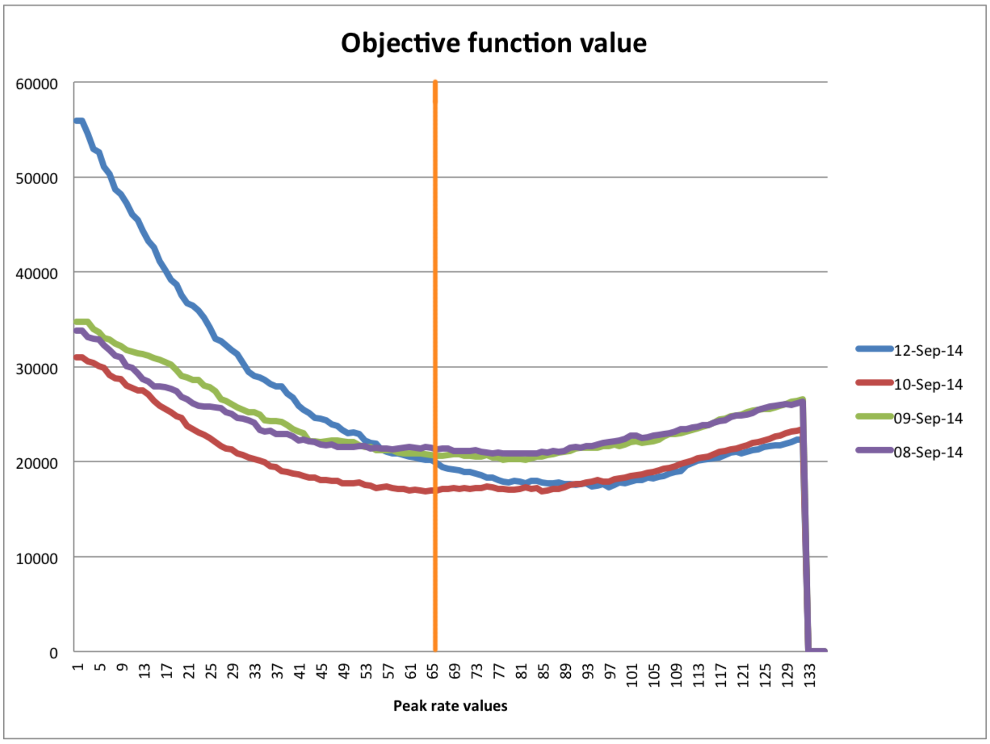


Figure 5.11: Objective function values for different values of the peak rate; the orange line represents the value of the historical unit rate (cPLP one rate modulation instances)

tions to speed up the branch and bound resolution. A further improvement that was attempted was the implementation of valid inequalities known as *Strengthened Shortest Path Inequalities* (SSPI), introduced in Heilporn (2008) for the Network Pricing Problem. These inequalities provide very good cuts for the NPP, but proved to be rather ineffective for the cPLP.

5.2.1 Problem bounds

In order to speed up solving time, some bounds and simple reformulations were added to the problem to reduce the size of the feasible region. The values applied to the bounds can be considered realistic with regard to the application.

Bounds for the capacity violations variables The model has a constraint stating that each flight can only choose one route and one departure time (eq.3.21); in the relaxed version of the problem, this constraint implies that the sum of the values of all x variables related to the same flight will always be equal to one. Therefore, if the number of flights that might enter in a sector or airport during one hour is inferior to the declared capacity for that resource, the corresponding violation variable ($\alpha_s^{(h)}$ or $\alpha_a^{(h)}$) is bounded to zero, otherwise to the difference between the capacity and potential entries.

Operational costs coefficients Airspace users aim at minimising the costs of flights, expressed in the cPLP as the sum of operational costs ($OC_{f,r}^{(m)}$) and route charges ($RC_{f,r}^{(m)}$). The operational costs term does not contain any variable and can therefore be precalculated; for each flight, the minimum $OC_{f,r}^{(m)}$ term, that is, the fixed cost of the cheapest route and departure time option for that flight, can be subtracted from either side of the minimum cost constraint (eq. 3.36). This leads to an equivalent, yet tighter formulation for the constraint.

Tight *Big-M* values and bounds on rate variables Literature suggests that in general, using Big-M values in constraints leads to loose formulations that are not handled well by solvers. Big-M terms should therefore be set to the smallest possible value that allows the constraints to hold. Finding tight values for Big-M terms can therefore help in speeding up resolution due to an improved formulation of the problem (see for example Dewez et al. 2008, for the Network Pricing Problem case). In pricing problems where upper bounds to costs or charges exist, this is a straightforward choice. In the cPLP, Big-M values appear in the linearisation constraints that bind the $y_{f,r}^{(m)}$ variables to the value of the route charges or to zero (eq. 3.38 and 3.39). In the model, the only existing bound to the charges is loosely implied by the revenue neutrality constraint (eq. 3.9). Since preliminary tests on small instances also showed that this constraint alone was not enough to guarantee that the model would provide realistic solutions, further bounds were added: first to the charges,by imposing that the modulated rates should not exceed the historical unit rate by more than 100€; then, to the overall route charges for each flight, should not exceed two times the amount that would be charged using the historical unit rates. This latter value is an upper bound to the route charges per flight, and is therefore assigned to the Big-M terms in the constraints ($N_{f,r}^{(m)}$).

5.2.2 Heuristics

Additionally to bounds on variables, some simple heuristics were also implemented, to provide the solver with feasible integer solutions and good cutoff bound values to speed up the resolution. Running these heuristics on branch and bound nodes was initially attempted, but the execution resulted infeasible in the majority of cases. It was therefore preferred to use these heuristics only for providing starting solutions to the branch and bound.

H0 (baseline solution) A first integer solution is obtained by setting:

$$P_s^{(h)} = Ur_n \quad \forall s \in S_n, n \in N, h \in H$$

that is, the rate variables equal to the unit rates of the corresponding ANSP, and solving for the $x_{f,r}^{(m)}$ variables only. This is the solution obtained with no price modulation, where the problem is reduced to only assigning the users to their minimum cost route and departure time. Since this reduces the bilinear terms of the cPLP (product of variables $P_s^{(h)} \cdot x_{f,r}^{(m)}$) to the product of constants and binary variables, the complexity of the problem is reduced and resolution time is very fast (less than three minutes for a data instance of one day of traffic and less than one second for one hour of traffic).

This solution is expected to be systematically worse than those obtained from the other heuristics. However, it provides a baseline for evaluating results obtained from price modulation since it allows to compare traffic distribution with and without modulation, keeping all other problem assumptions and conditions equal. Therefore in the following tests it is always provided as a reference.

H1 (integer solution from root node relaxation) In the solution of the relaxation of the cPLP (i.e., the problem solved without the integrality constraints on the $x_{f,r}^{(m)}$, $\alpha_s^{(h)}$ and $\alpha_a^{(h)}$) variables) typically less than 10% of the $x_{f,r}^{(m)}$ variables take a fractional value. Starting from this observation, the following simple heuristic was developed, to obtain a feasible integer solution from the almost-integer values from the relaxation:

Step 1. LR: Solve the relaxed cPLP to obtain the corresponding $x^{(m)}_{f,r,LR}$ and $P^{(h)}_{s,LR}$ values;

Step 2. MI1: Solve the mixed integer cPLP after setting the rates equal to the $P^{(h)}_{s,LR}$ values:

$$P^{(h)}_s = P^{(h)}_{s,LR} \quad \forall s \in S, h \in H$$

and obtain the corresponding $x^{(m)}_{IP1,f,r}$ values;

Step 3. MI2: To further improve the M1 solution, solve again the mixed integer cPLP, this time setting only the $x^{(m)}_{f,r}$ variables that had the same (integer) value in both LR and MI1 solutions:

$$x^{(m)}_{f,r} = x^{(m)}_{IP1,f,r} \iff x^{(m)}_{IP1,f,r} = x^{(m)}_{f,r,LR} \qquad \forall f \in F, r \in R_f, m \in M_f$$

and obtain the corresponding $x^{(m)}_{f,r,MI2}$ and $P^{(h)}_{s,MI2}$ values. Depending on the size of the problem, this phase could take a long time to solve. It should therefore be given a time limit, beyond which the execution terminates and the best solution so far (or the MI1 solution, in case none was found in MI2) is accepted.

Note that $\alpha^{(h)}_s$ and $\alpha^{(h)}_a$ variables are not considered by the algorithm since their values represent flights that violate capacity limits. Their values therefore depend on the values of $x^{(m)}_{f,r}$.

H2-A (decomposition) This heuristic extends the H0 heuristic by iteratively solving the problem with one set of variables at the time, alternating between the route assignment variables $x^{(m)}_{f,r}$ and the rates $P^{(h)}_s$. As already explained for H0, this drastically reduces the complexity of the problem and therefore resolution is very fast. The drawback of this algorithm is its tendency to get stuck in local minima; as a consequence, the solutions provided are generally not of very high quality.

The algorithm steps are the following:

Step 1. Initialisation Solve for:

$$P^{(h)}_s = Ur_n \quad \forall s \in S_n, n \in N, h \in H$$

and obtain the corresponding $x^{(m)}_{f,r,B}$ assignment.

Step 2. Iteration A Solve for:

$$x^{(m)}_{f,r} = x^{(m)}_{f,r,B} \quad \forall f \in F, r \in R_f, m \in M_f$$

and obtain the corresponding $P^{(h)}_{s,A}$ rates.

Step 3. Iteration B solve for:

$$P^{(h)}_s = P^{(h)}_{s,A} \quad \forall s \in S, h \in H$$

and obtain a new $x^{(m)}_{f,r,B}$ assignment.

Step 4. Termination Iterate Steps 2. and 3. until the stopping criterion is met. Return the best solution obtained.

H2-B (rolling horizon) A *rolling horizon* approach was also attempted, where, starting from the *baseline solution* (H0 heuristic), only a subset of the $x_{f,r}^{(m)}$ variables is solved at each iteration together with the rates $P_s^{(h)}$, trying to improve the objective function. A *time-based* approach, that is, choosing the subset of flights to solve based on flight departure time (e.g., all flights departing between 8:00 and 8:30) did not prove to be effective, meaning that in the preliminary tests rates almost never deviated from the initialisation values (the historical unit rates).

On the contrary, a *geography-based* approach, that is, solving one state at the time, was able to improve the baseline solution. Like the decomposition heuristic H2, this one also has the tendency to stop in local minima.

The algorithm steps are the following:

Step 1. Initialisation Solve for $P_s^{(h)} = Ur_n$ to obtain the baseline solution; Sort ANSPs in decreasing order for entering flights;

Step 2. Iterations For each ANSP in the ranking obtained in Step 1, starting from the one with the highest number of entering flights, do the following:

1. Solve for the peak rate;
2. Solve for the off peak rate;

Step 3. Termination When it is not possible to further improve the solution by iterating over ANSPs, terminate the algorithm and accept the best solution obtained so far.

H2 (H2-A and H2-B combined) An improved heuristic was obtained by combining the decomposition heuristic H2-A and the geography-based H2-B by solving them in a iterative sequence, changing from one to the other whenever the one that is being executed gets blocked in a local minimum. This approach proved successful, meaning that better solutions were obtained than those from H2-A and H2-B alone.

The algorithm steps are the following:

Step 1. Initialisation Solve for $P_s^{(h)} = Ur_n$ to obtain the baseline solution;

Step 2. H2-A Improve the baseline solution through heuristic H2-A until a local minimum is found (i.e., iterations are not able to improve the solution any further);

Step 3. H2-B Run heuristic H2-B on the solution obtained from Step 2, optimising one per country at a time until a new minimum is found;

Step 4. Termination If the stopping criterion is met, terminate the algorithm and accept the best solution obtained so far; otherwise iterate again from Step 2.

H3 (solving procedure for regional cPLP-one) This simple algorithm provides very quickly a solution that is close to optimality for a single rate. It was implemented for solving the cPLP-one scheme on a single state, but provides a feasible integer solution also for the other cPLP pricing schemes. Clearly, it is not expected to provide good quality solutions for non-regional instances.

The algorithm steps are the following:

Step 1. Initialisation Set $incr = 1$, $maxIncr = 100$, $maxGrow = 5$ and $bestObj = \infty$,

Step 2. Iterations Solve for:

$$P_s^{(h)} = \begin{cases} U_n & \text{if } h \text{ is off-peak time for sector } s \\ U_n + incr & \text{if } h \text{ is peak time for sector } s \end{cases} \quad \forall n \in N, s \in Sn_n,\, h \in H, \delta \geq 1$$

and obtain the objective function value for the iteration, $iterObj$. The objective value of the previous iteration is stored in $olditerObj$.
If $iterObj \leq bestObj$, update the value of $bestObj$ and set $growCnt$ to zero:

$$bestObj = iterObj$$
$$growCnt = 0$$

If $iterObj > olditerObj$, increment $growCnt$ of 1:

$$growCnt = growCnt + 1$$

Finally, at the end of the iteration, increment $incr$ of 1:

$$incr = incr + 1$$

Step 3. Termination If $incr > maxIncr$ or $growCnt > maxGrow$ or the problem is infeasible, terminate the algorithm and accept the best solution obtained so far; otherwise iterate again from Step 2.

The algorithm terminates whenever the set bound for the rate is reached or when the value of the objective function has increased for five consecutive iterations or in case the problem becomes infeasible.

The solution obtained from this heuristic, provided to the solver for reoptimisation, allows to obtain the optimal solution values for the one-stare cPLP-one instances in a few branch and bound iterations. Tested instances with one hour of traffic were normally solved in around two minutes (one for H3, one for branch and bound); instances with eight hours of traffic (such as the one presented in Sec. 5.3) were normally solved in five to eight minutes. The solutions proved to be of good quality for regional cPLP-two as well.

Table 5.3 and 5.4 provide results from a test of cPLP-two (i.e., two rates per ANSP) run on 24 data files to compare the heuristics described. For each algorithm, the objective function value of the best solution found is presented (Table 5.3), together with run time (Table 5.4). Objective values are meant for comparison within the same instance: the lower the value, the better. Different instances have completely different solutions so no comparison should be made among objective values of different instances. The baseline solution (obtained from H0) and the best solution found by the optimiser, initialised with the solutions provided by the heuristics and with a time limit of 10 000 seconds for the execution, are also provided for comparison. Instead of running time, the optimality gap value is shown in this latter case. The data instances are regional-scale files corresponding to one hour of traffic on France from September 12 and 10, 2014, with approximately 500 to 700 flights each. Due to the reduced size of the traffic sample, capacities were scaled by a factor of 0.5 (i.e., they were halved) to simulate demand-capacity imbalance. The other parameters were set to the following values: $K_1 \geq 1.00E^{-4}$, $K_2 = 60$ and $K_3 = 1$.

Results clearly show that even these small size traffic samples are hard to solve, with very high values of gap for all instances representing heavily trafficked hours, even after almost three

		Objective value				
Instance name	Number of flights	H0	H1	H3	H5	B&B
12 Sep 2014, LF, 7:00	587	3438.93	2698.15	2835.96	2566.18	2153.38
12 Sep 2014, LF, 8:00	704	10994.20	8373.95	6785.85	5833.72	6715.59
12 Sep 2014, LF, 9:00	667	12546.30	8582.25	6573.42	5832.32	5686.01
12 Sep 2014, LF, 10:00	624	4422.00	2942.27	2069.43	2112.38	1561.44
12 Sep 2014, LF, 11:00	634	4328.00	3678.63	2720.66	2801.41	2716.67
12 Sep 2014, LF, 12:00	553	2287.00	2227.01	2062.77	2197.10	1997.56
12 Sep 2014, LF, 13:00	621	4918.43	4618.38	2996.98	2660.28	2567.44
12 Sep 2014, LF, 14:00	570	2923.29	2923.29	1952.1	2126.14	1932.12
12 Sep 2014, LF, 15:00	598	5670.75	5134.10	2437.51	3564.75	2494.77
12 Sep 2014, LF, 16:00	633	5860.81	5534.10	3421.80	3410.82	3258.91
12 Sep 2014, LF,17:00	517	1527.65	1496.09	1038.11	1464.65	943.54
12 Sep 2014, LF, 18:00	520	3028.63	3088.10	2248.50	2367.63	2064.77
10 Sep 2014, LF, 7:00	580	4087.31	3186.21	2980.48	3083.1	2980.48
10 Sep 2014, LF, 8:00	646	4875.2	4180.28	3438.64	3411.2	3150.03
12 Sep 2014, LF, 9:00	639	6721.33	4761.12	2869.85	3397.3	3291.2
10 Sep 2014, LF, 10:00	584	3625.67	3823.78	2810.78	2513.17	2378.44
10 Sep 2014, LF, 11:00	603	3320.56	3285.16	2263.55	2110.19	2083.6
10 Sep 2014, LF, 12:00	534	2103.87	1759.03	1455.99	1584.1	1314.74
10 Sep 2014, LF, 13:00	583	2264.98	2126.51	1495.44	1784.07	1495.44
10 Sep 2014, LF, 14:00	559	2511.51	1910.56	1745.27	2319.12	1582.76
10 Sep 2014, LF, 15:00	545	2749.17	2126.49	1386.09	2275.07	1381.98
10 Sep 2014, LF, 16:00	531	2041.26	1671.05	1391.22	1691.09	1381.04
10 Sep 2014, LF,17:00	487	2243.02	1857.23	1270.71	1727.09	1194.38
10 Sep 2014, LF, 18:00	474	2316.49	1715.56	1550.77	1524.82	1582.76

Table 5.3: Comparison of heuristics results - objective function values

Instance name	Solving time (seconds) Number of flights	H0	H1	H3	H5	B&B Gap at 10 000 seconds
12 Sep 2014, LF, 7:00	587	0.73	3.60	32.28	63.21	1%
12 Sep 2014, LF, 8:00	704	1.02	8.44	70.14	82.53	52%
12 Sep 2014, LF, 9:00	667	0.98	6.89	339.83	82.74	57%
12 Sep 2014, LF, 10:00	624	1.21	31.72	273.89	87.01	35%
12 Sep 2014, LF, 11:00	634	1.13	26.19	99.79	90.52	61%
12 Sep 2014, LF, 12:00	553	1.04	13.29	150.99	72.94	1%
12 Sep 2014, LF, 13:00	621	0.92	9.70	295.13	76.21	48%
12 Sep 2014, LF, 14:00	570	0.63	8.09	176.60	49.37	46%
12 Sep 2014, LF, 15:00	598	0.79	5.07	96.96	67.45	35%
12 Sep 2014, LF, 16:00	633	0.74	3.72	85.66	50.52	48%
12 Sep 2014, LF,17:00	517	0.81	4.38	149.84	49.06	8%
12 Sep 2014, LF, 18:00	520	0.57	3.17	140.25	51.30	15%
10 Sep 2014, LF, 7:00	580	0.88	11.06	176.50	90.60	16%
10 Sep 2014, LF, 8:00	646	1.25	12.41	292.61	62.13	18%
12 Sep 2014, LF, 9:00	639	1.04	10.60	62.13	84.39	46%
10 Sep 2014, LF, 10:00	584	0.94	10.26	255.24	82.12	35%
10 Sep 2014, LF, 11:00	603	1.04	15.41	73.93	65.14	21%
10 Sep 2014, LF, 12:00	534	0.87	14.55	155.94	44.04	1%
10 Sep 2014, LF, 13:00	583	0.85	14.67	199.17	70.59	1%
10 Sep 2014, LF, 14:00	559	0.66	4.95	217.06	57.73	4%
10 Sep 2014, LF, 15:00	545	0.65	8.08	123.97	59.06	1%
10 Sep 2014, LF, 16:00	531	0.62	7.00	228.81	53.85	1%
10 Sep 2014, LF,17:00	487	0.75	6.02	133.08	52.00	1%
10 Sep 2014, LF, 18:00	474	0.68	4.49	197.06	50.78	1%

Table 5.4: Comparison of heuristics results - solving time and B&B gap

hours of execution. It also appears that none of the tested heuristics is able to provide a solution close to optimality. In general the best solution is always provided by either H2 or H3 and both have reasonable execution time, considering the size of the problem. In some cases however, when provided to the branch and bound resolution, the heuristic solutions were reoptimised by fixing the values of the integer variables and the resulting solution value was much worse than the original heuristic one (see for example the case of 12 Sep 2014, LF, 8:00). Poor results from H1, which is based on the solution of the linear relaxation of the problem, confirm that the relaxation of the problem is, indeed, bad and therefore suggest that further improvements should be attempted to the formulation.

5.3 Resolution of a regional instance

The present section compares cPLP results obtained from the 8-hour regional instance from September 12, 2014 (see Table 5.1), comprising all traffic departed between 6:00 and 14:00 (UTC time) that crossed, departed from or arrived into French airspace (icao code: LF), which amounts to 4695 flights in total. This traffic sample represents the peak hours of the state that had the highest level of congestion registered in the day, in terms of airspace sectors whose capacity was breached and is therefore a valid testing instance for assessing the effectiveness of the proposed cPLP policy.

The airspace configuration in use at each hour comprises 90 to 100 active airspace sectors including terminal manoeuvring areas at major airports (see Figure 5.12 and Table 5.8).

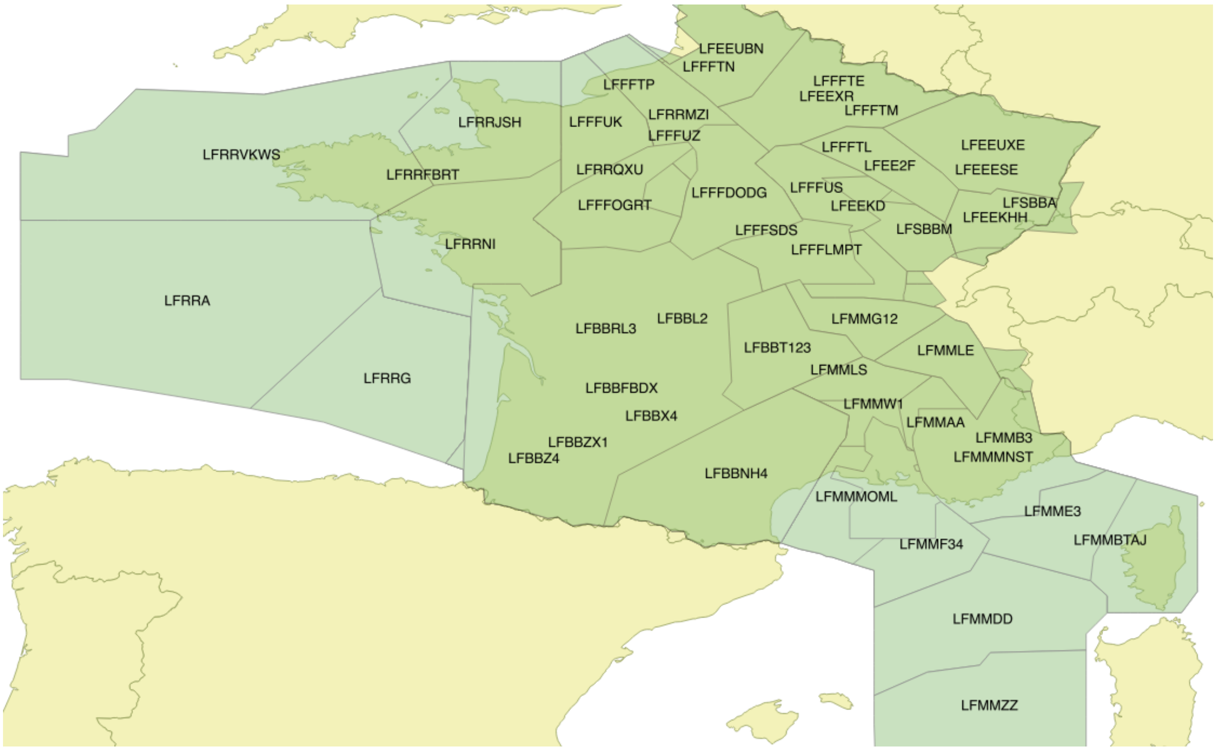


Figure 5.12: Active sectors in French airspace at 10:00 UTC on Sep.12, 2014

The traffic distribution, according to historical data on actually flown routes for this traffic sample, presents the highest traffic peaks between 9:00 and 12:00.

Approximately 90% of the active sectors in this sample have a declared nominal capacity (the remainder are considered as *unconstrained*). Note that although the results presented in the following only concern French airspace, capacity constraints on all active European sectors are considered in the problem resolution.

The peak threshold here is set to 50% of nominal capacity, and the remaining parameters of the models are set as follows: $K_1 = 10^{-5}$ for revenue neutrality violation, $K_2 = 60$ for sector capacity violation and $K_3 = 1$ for airport capacity violation.

Concerning departure time options, in the presented test each flight has a possible departure time shift of up to 30 minutes before and 30 minutes after the requested departure time for that specific flight by the AU. The shift is divided in 10-minute long slots, meaning that a flight with requested departure time set at 6:00 will be assigned one of the following departure times by the model: [5:30; 5:40; 5:50; 6:00; 6:10; 6:20; 6:30]. It is assumed that a 10-minute granularity is precise enough to describe a process that is to be applied in the strategic phase of flight planning.

The problem size with this configuration has 575 216 rows and 134 661 integer and binary variables. The following pricing schemes were tested, against historical data and the baseline solution:

cPLP with one modulated charge (cPLP-one) The modulation is applied to the peak rate only, while the off peak rate is kept equal to the historical unit rate value (65.92 €). As already mentioned, having only one rate to solve for the entire problem makes this a particularly easy to solve case of cPLP, since adapting the efficient solving procedure from Castelli et al. (2013) to this case is straightforward.

cPLP with two modulated charges (cPLP-two) The modulation is applied to both peak and off peak rate. Even if with two tariff variables only, this variant of the problem is highly complex to solve due to its combinatorial nature. The heuristics presented in Sec. 5.2 were not able to provide good enough solutions to allow the solver to speed up the branch and bound resolution and solve this instance of the problem within reasonable time limits (more than twenty hours are necessary for the solver to terminate the branch and bound execution).

At the current state, even resolution by enumeration appears to be more efficient (for the two rates case): this has also been done for comparison purpose by testing iteratively all integer combinations of peak and off peak values ranging between 1€ and the set limit of 165€ (100 € higher than the historical unit rate value, 65.92€). With this configuration, and assuming a peak rate no lower than the off peak rate, approximately 11 000 combinations have to be tested, so the procedure took several hours (approximately 10, but this time could be easily reduced by parallelising the procedure). The best solution obtained through this method is close enough to optimality (3.43% in this test) and was provided to the solver to initialise the branch and bound algorithm. The test carried out in this way allowed to solve the instance within the set limit of 1% gap.

A summary of the solution values obtained from these pricing schemes are described in Table 5.6.

Concerning route and departure time assignment to flights, the solutions obtained are evaluated according to the performance indicators described in the following. Obtained indicators values are summarised in Table 5.6, while a comparison of the different solutions in in terms of rerouted and shifted flights between each couple of pricing schemes is provided in Table 5.7.

En-route charges: sum of the route charges imposed on the flights, used to cover the costs of ANS provision; calculated according to eq. 3.54 and 3.56 for cPLP-one, to eq. 3.54 and 3.55 for cPLP-two; measured in € per flight;

Pricing scheme	Baseline	cPLP-one	cPLP-two
Peak rate (€)	65.92	95.95	95.78
Off peak rate (€)	65.92	65.92	88.15
Shift (minutes)	6374	8956	6847
N. of sector cap. viol.	218	139	143
N. of airport cap. viol.	6	5	5
Revenue neutrality viol. (€)	0	825894	904066
Solving time	6 seconds	320 seconds	11 hours

Table 5.5: Comparison of the obtained solutions

Indicator (avg. per flight)	Historical data	Baseline	cPLP-one	cPLP-two
En-route charges (€)	1563.56	1390.64	1613.88	1615.75
Operational costs (€)	7496.94	7252.65	7258.84	7255.18
Departure shift (minutes /10)	0.00	1.53	2.11	1.64
Arrival shift (minutes /10)	0.00	2.08	2.64	2.20
Horizontal flight efficiency	0.12	0.11	0.11	0.11
Temporal flight efficiency	3.92	0.89	0.94	0.92

Table 5.6: Performance indicators values obtained from the different solutions

Operational costs: sum of the costs for operating the aircraft for the assigned route duration plus the ground shift component; calculated according to eq.3.19, measured in € per flight;

Departure shift: absolute difference between the requested and assigned departure time. Measured in minutes per flight;

Arrival shift absolute difference between the arrival time obtained by departing at requested arrival time using the shortest route and the assigned arrival time. Measured in minutes per flight;

Horizontal flight efficiency: difference between the origin-destination en-route distance of assigned routes, and the great circle distance between the origin and destination, expressed as a percentage of the great circle distance;

Temporal flight efficiency: difference between the duration of the shortest route and the duration of the assigned route; measured in minutes per flight.

Results show that the proposed modulation of en-route charges does not impact flight costs in a significative way. Compared to historical data, the cPLP pricing schemes have lower average operational costs and lower average route charges. Compared to the baseline solution, however, despite practically equivalent operational costs, the modulated schemes have higher en route charges. This was quite predictable in the cPLP-one case (since charges are greater or equal to the historical unit rate applied in the baseline), less so for cPLP-two. The remaining indicators show that the baseline solution dominates the modulated charging schemes where temporal efficiency is concerned, even if the cPLP-two values appear to be very close. Values of average horizontal flight efficiency are equivalent in baseline, cPLP-one and cPLP-two (in fact, as shown in Table 5.7, the route assignment in the three solutions differs for no more than a couple hundred flights between one and another) and slightly improved in comparison with historical data.

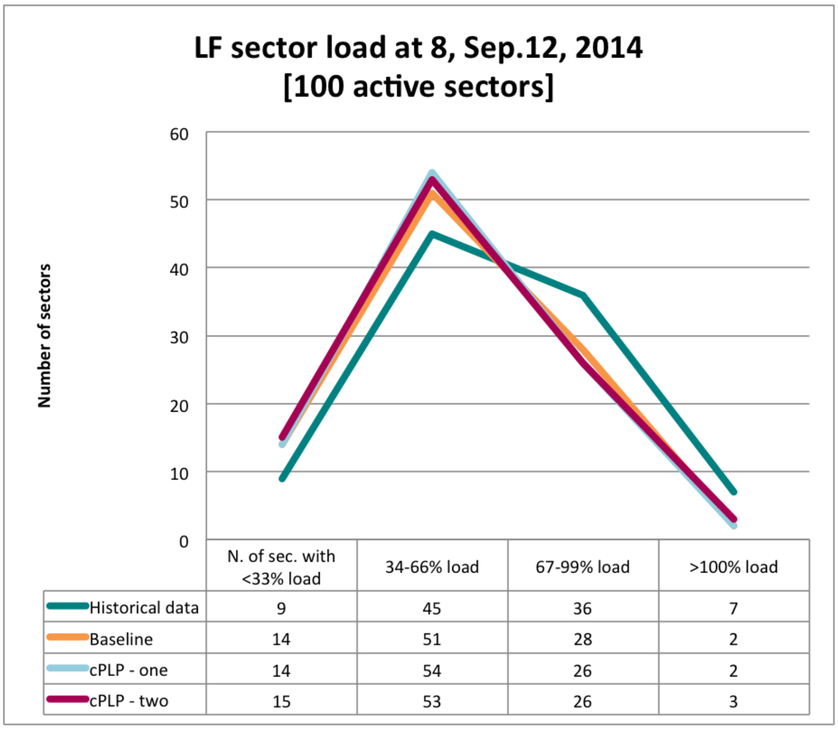
Difference in number of flights with rerouting				
	Historical	Baseline	cPLP-one	cPLP-two
Historical	0	2698	2695	2701
Baseline	-	0	201	184
cPLP-one	-	-	0	74
cPLP-two	-	-	-	0
Difference in number of flights with shift				
	Historical	Baseline	cPLP-one	cPLP-two
Historical	0	466	632	495
Baseline	-	0	249	89
cPLP-one	-	-	0	195
cPLP-two	-	-	-	0
Difference in number of flights with shift and rerouting				
	Historical	Baseline	cPLP-one	cPLP-two
Historical	0	279	364	300
Baseline	-	0	57	40
cPLP-one	-	-	0	28
cPLP-two	-	-	-	0

Table 5.7: Comparison of rerouted and shifted flights among the obtained solutions

The obtained distribution of traffic load over airspace sectors is illustrated in the charts of Figures 5.13-5.16 for the four most trafficked hours. The complete tabular version of these results, for all the hours of the tested traffic sample can be found in Appendix C Sec. C.2. Here the hourly load, defined as the ratio between the number of flights entering a sector during a certain hour and its declared nominal capacity is used as a metric for evaluating the obtained solution.

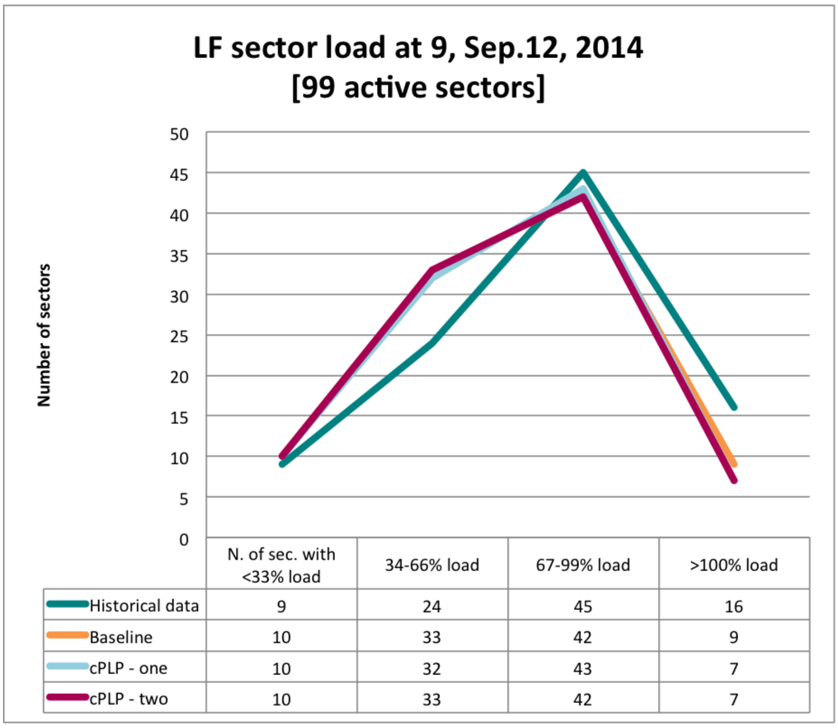
The effects of traffic redistribution, in terms of number of sectors that are in peak in the historical data and baseline solution, compared to the figures obtained after running cPLP-one and cPLP-two on the traffic sample, are illustrated in Table 5.8. Finally, Figures 5.17-5.20 further illustrate the geographical distribution of traffic load over French upper airspace obtained from the different solutions (Fig. 5.18 for the baseline, 5.19 for cPLP-one, 5.20 for cPLP-two) and historical data (Fig. 5.17).

Both price modulation schemes prove to be effective in improving traffic load over airspace sectors. Obtained sector loads figures in the cPLP route assignments are systematically lower than historical data, and improved compared to the baseline, although not by astonishing values. The number of sectors with more than 100% load (i.e., sectors with violated capacity) also decreases in all hours with cPLP, as the traffic moves across less utilised sectors. This is particularly evident in Figure 5.16, which represents sector traffic load between 11:00 and 12:00. Here the number of sectors with more than 100% load is halved in both cPLP solutions compared to the baseline, while the number of sectors with 33%-66% traffic load increases. What is more common however having traffic shifted from the capacity breached sectors to the heavily loaded ones (66%-99% load). This is probably due to geographic reasons, since high demand-capacity imbalances tend to involve areas wider than a single sector, especially within the same country. The entity of this improvement however is maybe better appreciated when considering the number of individual capacity breaches (number of sector capacity violations in Table 5.6), which amounts to 218 in the baseline solution and is reduced to 139 and 143 in cPLP-one and cPLP-two solutions respectively.



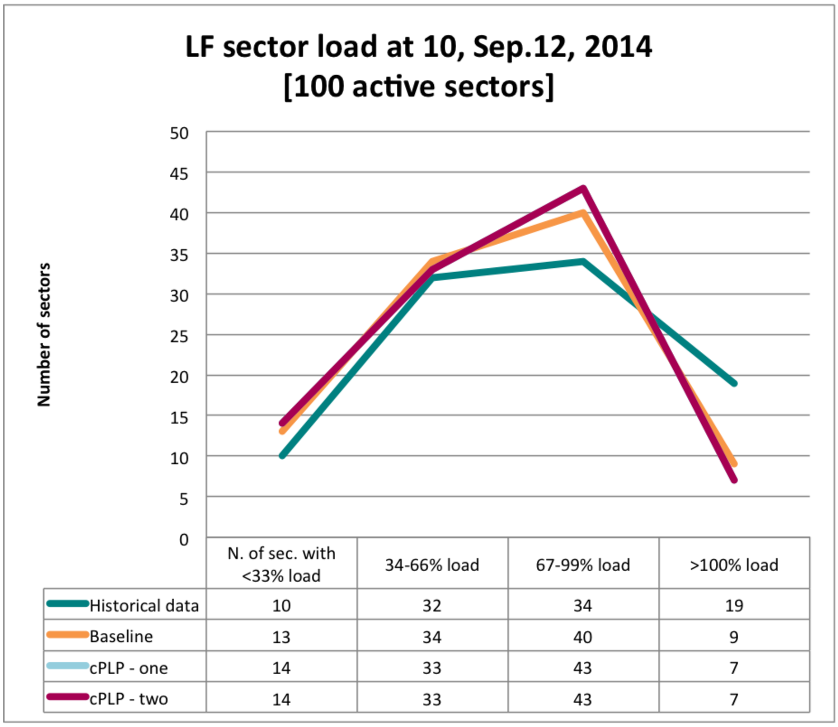
	N. of sec. with <33% load	34-66% load	67-99% load	>100% load
Historical data	9	45	36	7
Baseline	14	51	28	2
cPLP - one	14	54	26	2
cPLP - two	15	53	26	3

Figure 5.13: Sectors load on French airspace obtained from the different solutions between 8:00 and 9:00 UTC



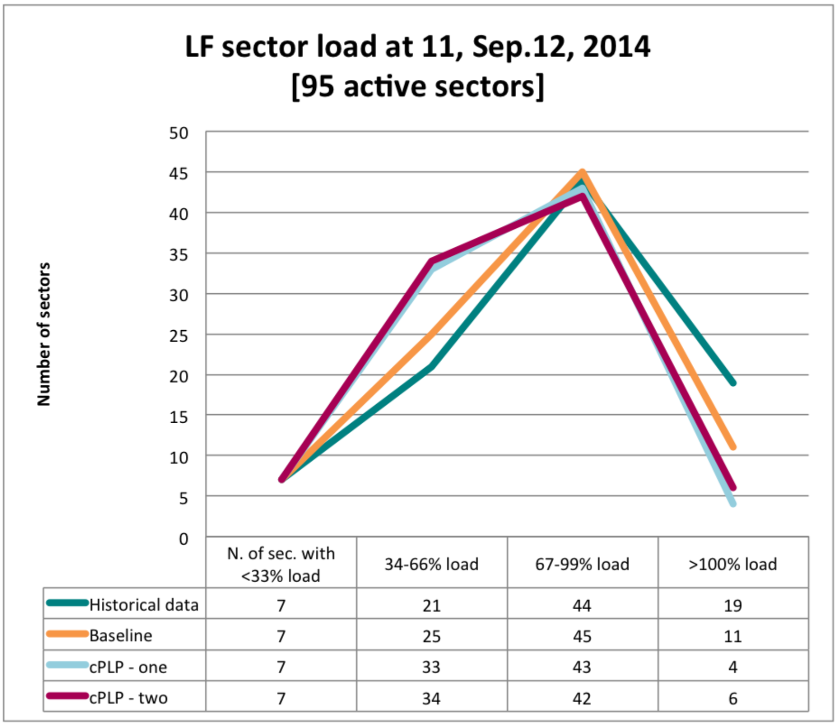
	N. of sec. with <33% load	34-66% load	67-99% load	>100% load
Historical data	9	24	45	16
Baseline	10	33	42	9
cPLP - one	10	32	43	7
cPLP - two	10	33	42	7

Figure 5.14: Sectors load on French airspace obtained from the different solutions between 9:00 and 10:00 UTC



	N. of sec. with <33% load	34-66% load	67-99% load	>100% load
Historical data	10	32	34	19
Baseline	13	34	40	9
cPLP - one	14	33	43	7
cPLP - two	14	33	43	7

Figure 5.15: Sectors load on French airspace obtained from the different solutions between 10:00 and 11:00 UTC



	N. of sec. with <33% load	34-66% load	67-99% load	>100% load
Historical data	7	21	44	19
Baseline	7	25	45	11
cPLP - one	7	33	43	4
cPLP - two	7	34	42	6

Figure 5.16: Sectors load on French airspace obtained from the different solutions between 11:00 and 12:00 UTC

Time	N. of active sectors	N. of peak sectors (>50% load)			
		Historical data	Baseline	cPLP - one	cPLP - two
6:00-6:59	90	55	4	6	4
7:00-7:59	95	61	39	31	37
8:00-8:59	100	71	57	53	55
9:00-9:59	99	81	73	72	73
10:00-10:59	100	73	73	72	73
11:00-11:59	95	79	74	71	70
12:00-12:59	95	70	54	52	50
13:00-13:59	97	68	59	55	57
14:00-14:59	94	73	34	35	34
15:00-15:59	94	66	0	1	0

Table 5.8: Number of peak sectors in historical data, baseline solution and after applying cPLP

A final word for comparing the two tested modulation schemes between them: it is hard to assess which one provides a better solution, at least from the test on this traffic sample: cPLP-one often shows a better traffic distribution (in terms of lower values of violated sector capacities) and revenue neutrality value; cPLP-two, in turn has a total shift more than 20% lower than cPLP-one (see Table 5.6), but obtained by raising not only the Peak rate, but the off peak too (meaning, some excess traffic is re-routed to neighbouring countries).

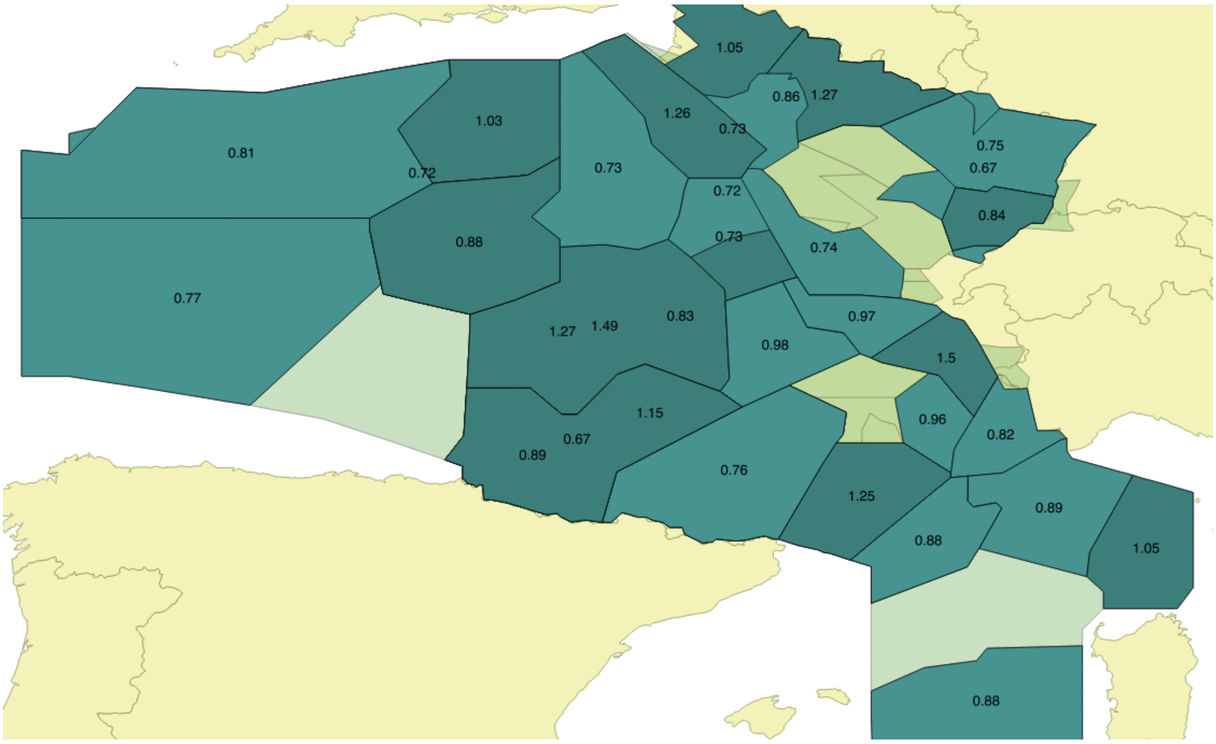


Figure 5.17: Sectors with heavy traffic load (lighter colour: 67% to 99%, darker colour: above 100%) in historical data at 10:00 UTC, Sep.12, 2014

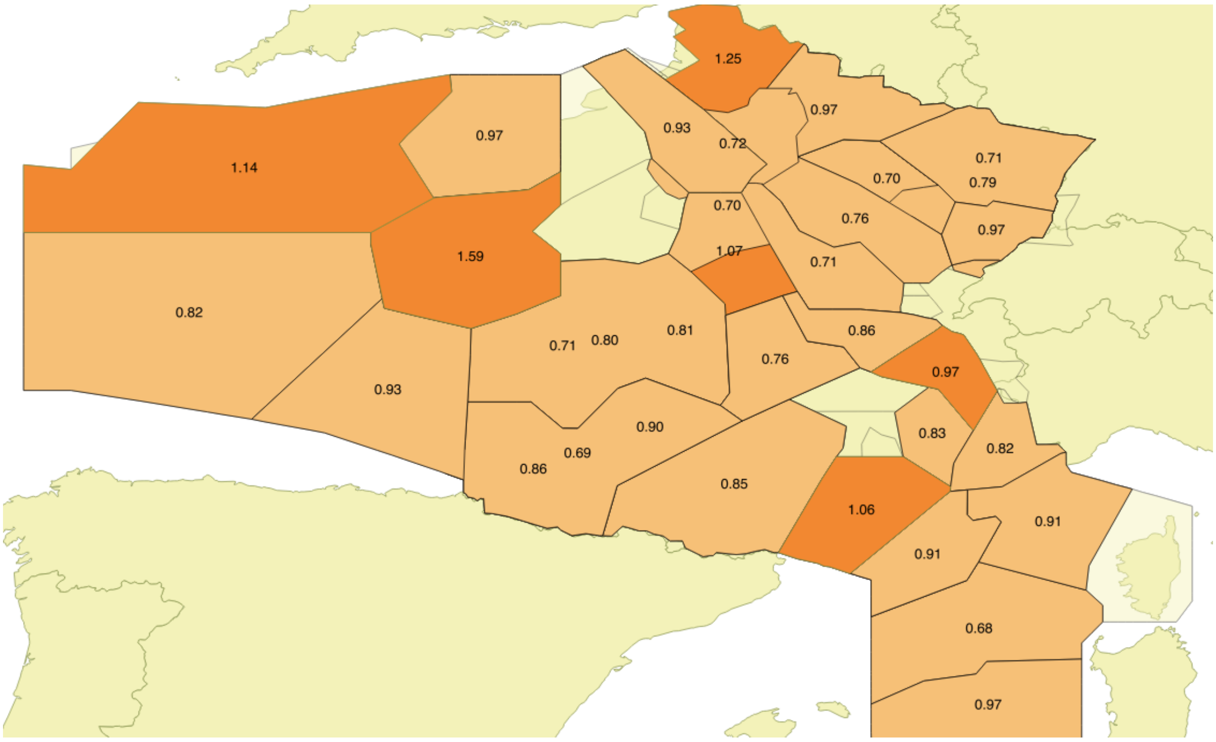


Figure 5.18: Sectors with heavy traffic load (lighter colour: 67% to 99%, darker colour: above 100%) in the baseline solution at 10:00 UTC, Sep.12, 2014

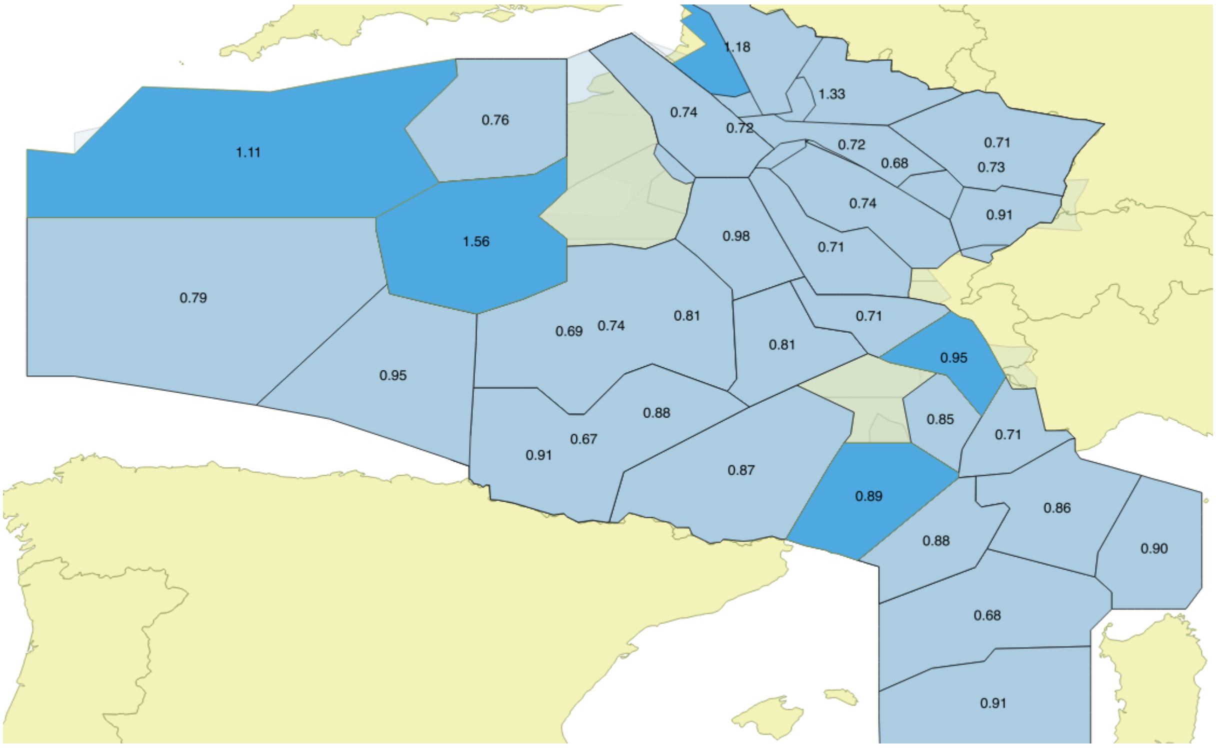


Figure 5.19: Sectors with heavy traffic load (lighter colour: 67% to 99%, darker colour: above 100%) in the cPLP one solution at 10:00 UTC, Sep.12, 2014

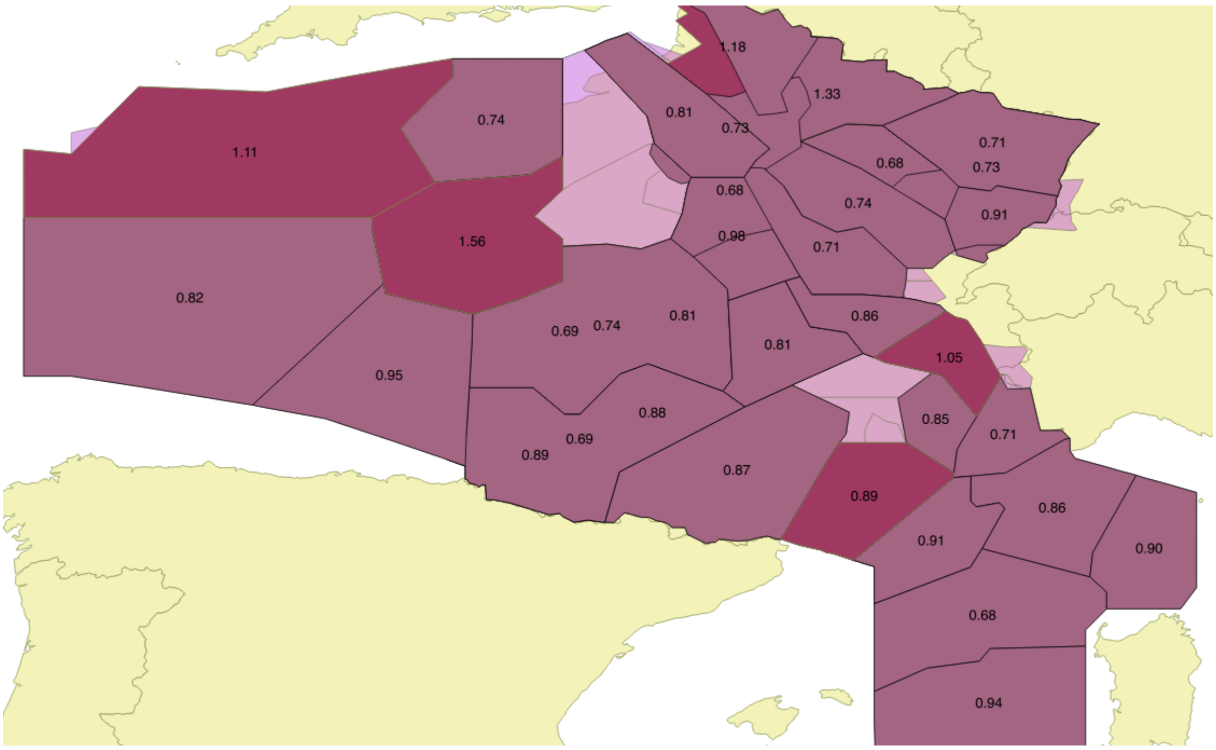


Figure 5.20: Sectors with heavy traffic load (lighter colour: 67% to 99%, darker colour: above 100%) in the cPLP-two solution at 10:00 UTC, Sep.12, 2014

5.4 Further results from the project SATURN

The current section illustrates some results obtained by a colleague within the project SATURN (as illustrated in deliverables D.4.2 and D.6.5 SATURN 2015*a*,*b*) that tackled cPLP resolution through a different approach from the one illustrated in the present work, specifically an implementation in the multiobjective optimisation software modeFrontier. This software performs optimisation through Genetic Algorithms (GA) and was able to solve much larger data instances, including the entire September 12, 2014 data instance (approximately 29 000 flights) on the whole European network.

In general, Genetic Algorithms belong to a class of solution methods known as evolutionary metaheuristic. This is a class of stochastic algorithms whose steps are loosely inspired by the principles of evolution and natural selection. In a GA groups of candidate solutions are generated, selected, variated and evolved throughout the iterations according to the following outline:

Genetic Algorithm

1. **Initialisation** Generate a random population of n candidate solutions for the problem (individuals or chromosomes);
2. **Fitness** Evaluate the quality of each individual according to a defined fitness function;
3. **Selection** Select two individuals (parents) from the population according to their fitness (the better fitness, the bigger chance to be selected);
4. **Crossover** With a crossover probability cross over the parents to form new individuals (offspring). If no crossover is performed, the offspring are copies of the parents;
5. **Mutation** With a mutation probability mutate offspring at each locus (position in the solution);
6. **Accepting** Place new offspring in a new population;
7. **Test** If the end condition is satisfied, stop, and return the best solution in current population; otherwise repeat from step 2;

The stochastic element is fundamental in GAs to avoid the creation of a number of almost identical sub-optimal individuals. Being heuristic methods, there is no guarantee that the final solution will be the optimal one, but they are able to explore the solution search space and eventually converge to a solution in a reasonable amount of time even for very large problems.

To tackle the cPLP, in its two-charges modulation variant, the model was implemented in an in-house extension of the modeFrontier software called the MOGASI algorithm (see Costanzo et al. 2014). Specifically, the implemented variant of cPLP is a multiobjective optimisation problem, that is, the three components of the cPLP model are treated as independent objective functions; the solving procedure delineates the Pareto front with the optimal solutions obtained from the different possible combinations of the objectives (i.e., referring to the formulation used so far, equivalent to trying all possible combinations of the K_1, K_2 and K_3 parameters). The solving procedure applies the genetic algorithm to the peak and off peak rates: for a candidate assignment of rates, the corresponding optimal routing choices for the for the flights are calculated, and these latter are used to compute the corresponding objective value and therefore the *fitness* of that particular set of tariffs.

The pricing variables are defined as:

$$Pp_n = Ur_n + \delta_{p,n} \forall n \in N \tag{5.1}$$

$$Pp_n = Ur_n + \delta_{o,n} \forall n \in N \tag{5.2}$$

where $\delta_{p,n}$ and $\delta_{o,n}$ are respectively peak and off-peak variations on the historical unit rate.

An additional constraint is set on the revenue neutrality violation term ϵ_n stating that the total percentage variation over all ANSPs should not exceed 20%:

$$\sum_{n \in N} |\epsilon_n| \leq 0.2 \tag{5.3}$$

Finally, the maximum allowed average capacity breach is also bounded not to exceed 25% of the total capacitated sectors and airports:

$$\frac{\sum_{s \in S, h \in H} \frac{\alpha_s^{(h)}}{Qs_s^{(h)}} + \sum_{a \in A, h \in H} \left(\frac{\alpha_{a,dep}^{(h)}}{Q_{a,dep}^{(h)}} + \frac{\alpha_{a,arr}^{(h)}}{Q_{a,arr}^{(h)}} + \frac{\alpha_{a,gl}^{(h)}}{Q_{a,gl}^{(h)}} \right)}{\sum_{s \in S, h \in H} \alpha_s^{(h)} + \sum_{a \in A, h \in H} (\alpha_{a,dep}^{(h)} + \alpha_{a,arr}^{(h)} + \alpha_{a,gl}^{(h)})} \leq 0.25 \tag{5.4}$$

The solving program was run on the daily instance file for September 12, 2014. This file comprises 29 539 flights and capacity constraints for all sectors and airports. It lacks however traffic volumes capacities (that were implemented in a later phase of SATURN and account for more than 50% of actual capacity constraints), and therefore the resulting solutions have a much lower shift and number of capacity breaches in proportion to the results illustrated in Section 5.3 for the French airspace alone.

The solving program ran for approximately 36 hours, during which approximately 100 000 solutions, of which 30 000 feasible, were generated. Among all the feasible ones, the scatter chart in Figure 5.21 displays a collection of selected Pareto solutions (green squares) in terms of two of the three objectives, specifically the capacity violations (y axis) and the total shift (x axis). The red square represents the baseline solution computed with the historical unit rates. The GA is able to identify, for both objectives, a number of equally good alternatives that dominate the baseline solution. Furthermore, the plot confirms the same negative correlation pointed out in Section 5.1 between the objectives: shifting flights can significantly reduce capacity violations and solutions alternative to the baseline always represent a tradeoff between these two components.

Figure 5.22 graphically compares the objectives values for the four Pareto solutions selected in Figure 5.21 in terms of total shift, average capacity violation, revenue neutrality violation and number of violated capacities. The blue line is later referred to as S1, the green line as S2, the light blue as S3 and the purple line S4. None of these solutions have optimal values in all objectives. Instead, they all represent equally valid trade-offs between requirements, achieved by redistributing air traffic in different ways. The red line (S0) represents the baseline solution. It exhibits a medium Total Shift while the revenue neutrality is equal to zero by definition. The solution represented by the light blue line (S3) can be considered the best tradeoff among objectives, under several aspects: the revenue neutrality violation is no worse than the other non-baseline solutions and both total shift and capacity violations indicators are better than or nearly equal to the baseline values.

Figure 5.23 shows peak and off-peak rates for the four selected solutions only in the case at least one of them deviates by more than four € from the September 2014 unit rate. For some states, such as Belgium and Luxembourg (EB), Finland (EF), and Slovenia (LJ), both peak and off-peak rates appear to be higher than historical rates in all solutions. In other cases, such as for Spain (LE), the opposite occurs, with all other intermediate alternatives being possible: for example The Netherlands (EH), where off peak rates are always lower, and peak rates are always higher, than the historical tariff.

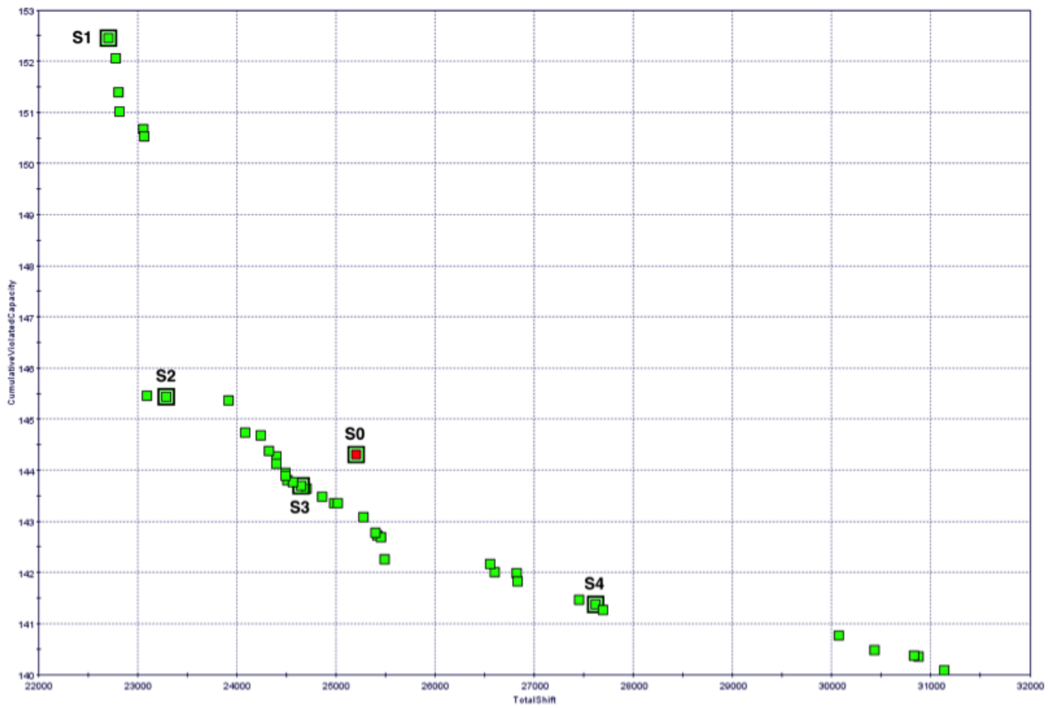


Figure 5.21: Pareto front for Total Global Shift vs. Capacity Violation (Source: SATURN Deliverable 6.5)

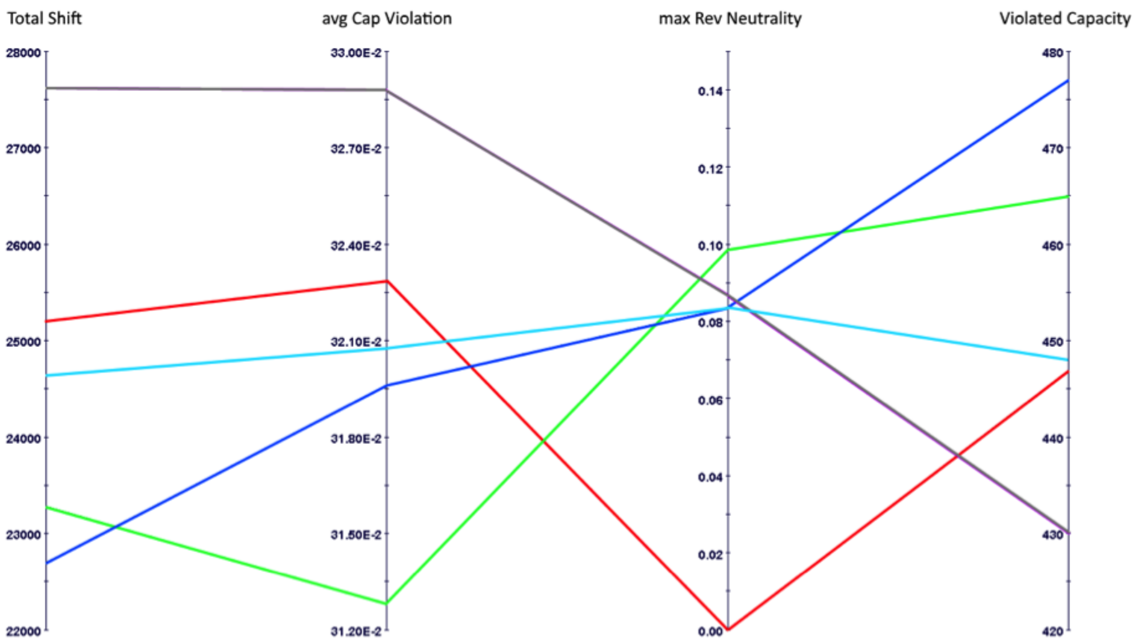


Figure 5.22: Tradeoffs among the baseline solution (red line) and five Pareto-solutions (Source: SATURN Deliverable 6.5)

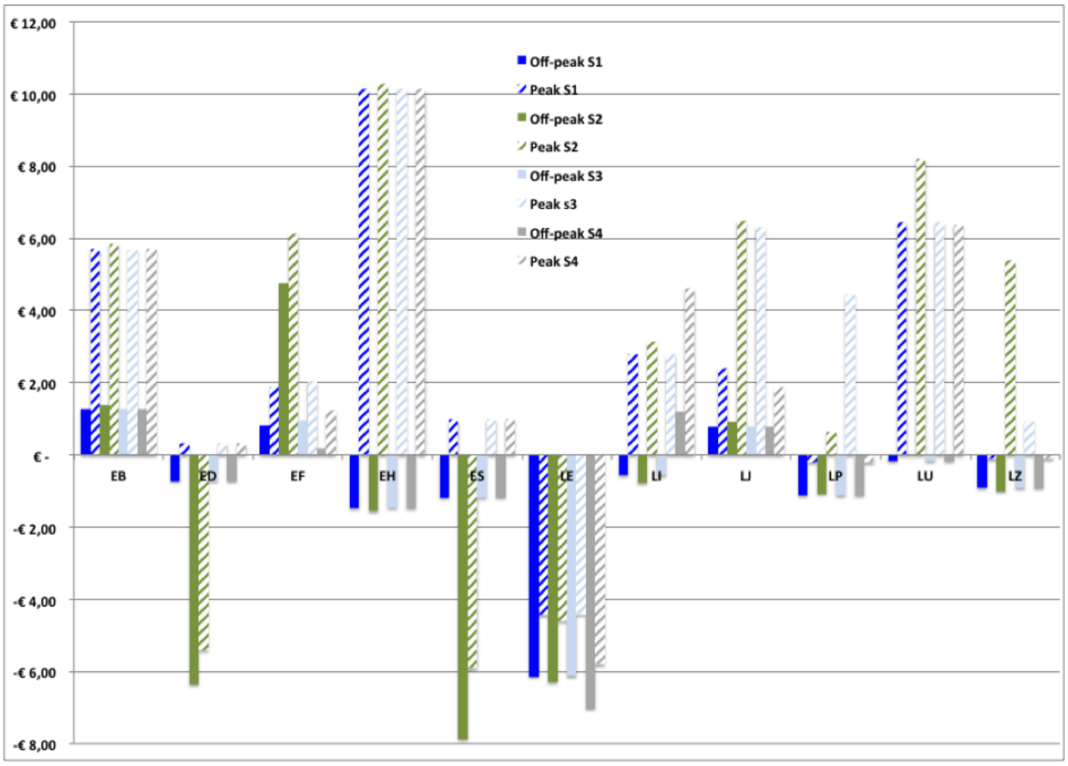


Figure 5.23: Peak and off-peak rates for selected ANSPs (Source: SATURN Deliverable 6.5)

N. of flights with different routes					
	S0	S1	S2	S3	S4
S0	0	858	550	258	541
S1	-	0	525	642	1236
S2	-	-	0	320	936
S3	-	-	-	0	644
S4	-	-	-	-	0
N. of flights with different shift					
	S0	S1	S2	S3	S4
S0	0	1000	685	308	643
S1	-	0	629	744	1420
S2	-	-	0	412	1114
S3	-	-	-	0	740
S4	-	-	-	-	0

Table 5.9: Number of flights that use different routes and number of flights that have different total shifts per solutions pair (Source: SATURN Deliverable 6.5)

Table 5.9 compares the different solutions in terms of number of flights that fly different routes and number of flights that have different shifts. It shows that on average less than 5% of flights was shifted or re-routed. This figure would likely increase if traffic volume capacities were also considered in the resolution, but the current proportion is not incoherent with the proportions obtained from the 8 hour regional example presented in Section 5.3 (Table 5.7), which considers peak hours on one of the regions with the highest demand-capacity imbalances in Europe.

The results from this test suggest that it is possible to improve the baseline solution at European level through a very limited spatial and/or temporal redistribution of traffic. This however confirms the previously stated observation that such improvements come at a price: in areas with an high demand-capacity imbalance there might not be viable alternative routes through airspace with spare capacity, therefore reducing capacity breaches inevitably increases the total amount of shift and vice versa. What could be a suitable and realistic trade-off between the two is an issue left open for discussion.

Conclusion

This thesis investigated the possibility of mitigating demand-capacity imbalances in the European ATM system by providing signals to airspace users through the modulation of en-route charges in the strategic level of flight planning, months in advance of the day of operations. The aim of this approach is to redistribute traffic in excess by inducing rerouting or shifts in the schedule. Previous research on this topic is extremely scarce, with only a handful of works available in literature. Therefore, most research carried out within the SESAR WP-E project SATURN (within which this thesis was developed) has been exploratory, spreading across different disciplines and facing several issues that do not commonly affect more mature fields of research, where established methodologies and practices exist. This exploratory character is a prominent trait also of this thesis, which borrows principles and techniques from sectors as different as economics, computer science, data engineering and operations research. It represents a seminal work and, as such, it clearly has some limitations. For example, the optimisation model presented, which represents the application of a centralised peak load pricing policy at European level by a central authority and the reaction of airspace users that route their flights accordingly, is fully deterministic and presents some simplistic assumptions. In this sense, the model could be improved by taking some degree of uncertainty in consideration. This, together with improvements to the proposed formulation, suggests compelling directions for future research. Additionally, the interaction of ANSPs among themselves, that is, investigating peak load pricing in a *decentralised* approach, would also represent an interesting development of this work.

In any case, data availability currently constitutes a major limitation to research on European air traffic. The sources available have very restrictive licensing agreements and the data itself requires a significative effort in cleaning, combination and harmonisation before becoming suitable for testing mathematical models. Data preparation and management represented a major challenge for the work presented here and for the SATURN project, and having finally been able to test these models on data instances obtained from real traffic represents a significant achievement that currently has no predecessor in scientific literature on European ATFM models.

To the best of our knowledge, in this sense, this work represents an interesting contribution also in regard to bilevel problems applications. The tests presented here were run on traffic samples that are at least ten times larger than the data instances commonly presented in bilevel literature. Data files with a hundred users are normally considered hard to solve; the regional example presented here has almost 4 700 flights and it has been solved to near optimality for a non trivial case with two rates for one ANSP.

A final word goes to the quality of the results obtained, in the project, but also illustrated in this thesis: they are realistic with regard to the application, and suggest that the proposed methodology works. They confirm that a mild modulation of air navigation charges could indeed help in redistributing air traffic in Europe in a more sustainable way for future growth.

Appendices

Appendix A

Database for air traffic data

The present appendix illustrates the structure of the bespoke database (DB) developed for storing and analysing air traffic data. The main rationale for developing the database was to overcome the limitation of the NEST software (as explained in Sec. 4.1.2 and 4.4) and to provide an easier access to its air traffic data sets. The DB was mainly used as a support tool for performing data analysis and for building the necessary data instance files to test the models illustrated in Chapter 3.

A.1 Database structure

The database has two user-defined schemas: the `public` schema, where the original data is stored and the `saturn_units` schema where data obtained from elaborations or data analysis from the public schema is stored.

A.1.1 The `public` schema

The `public` schema contains the following tables, sourced as illustrated in Table 4.11:

Airspace-related tables

Table name	**Data source file**	**Origin**
`acc_configurations`	Configuration.cfg	NEST
`acc_configurations_openings`	OpeningScheme.cos	NEST
`crco_areas`	view on sector_areas table	
`military_openings`	AIRAC .mot files	NEST
`sector_areas`	Airspace.spc and Sectors.gsl	NEST
`sector_capacity`	Capacity.ncap, Activation.nact	NEST
`sector_regulations`	RegPlan.nreg	NEST
`sector_structure`	Airspace.spc and Airblock.gar	NEST
`traffic_volumes`	TrafficVolume.ntfv	NEST
`tv_flows`	flow.nflw	NEST
`tv_flows_det`	flow.nflw	NEST
`unit_rates`	ur.txt files	EUROCONTROL website

Table A.1: Data sources for airspace-related tables

The tables listed above are described in detail in the following.

`acc_configurations` **table** This table contains data about existing sector opening configurations for Area Control Centres (ACCs).

Field Name in DB	Data Type	Info
acc	string	ICAO designator for this ACC
conf	string	ACC configuration name
sector	string	Sectors belonging to this ACC configuration
airac	smallint	Reference Aeronautical Information Regulation And Control (AIRAC) cycle

Table A.2: `acc_configurations` table

`acc_configurations_openings` **table** This table contains information on opening times and dates for Area Control Centres configurations.

Field Name in DB	Data Type	Info
acc	string	ICAO designator for this ACC
date	date (DD:MM:YY)	Day of validity for this configuration setting
start_time	time (HH:MM:SS)	start time for this configuration setting
end_time	time (HH:MM:SS)	end time for this configuration setting
conf	string	name of the ACC configuration
airac	smallint	Reference Aeronautical Information Regulation And Control (AIRAC) cycle

Table A.3: `acc_configurations_openings` table

`military_openings` **table** This table contains information on opening times and dates for military airspace sectors.

Field Name in DB	Data Type	Info
icao	string	ICAO designator for this sector
start_date	date (DD:MM:YY)	start date for this configuration setting
end_date	date (DD:MM:YY)	end date for this configuration setting
start_time	time (HH:MM:SS)	start time for this configuration setting
end_time	time (HH:MM:SS)	end time for this configuration setting
top_fl	integer	upper vertical boundary for contained airspace (if applicable)
bottom_fl	integer	lower vertical boundary for contained airspace (if applicable)
type	string	EX (=opening time from Entry to Exit) or E (= opening time at Entry, no restriction at Exit)
airac	smallint	Reference Aeronautical Information Regulation And Control (AIRAC) cycle

Table A.4: `military_openings` table

`sector_areas` **table** This table contains identification and geographic data for airspace sectors.

Field Name in DB	Data Type	Info
icao	string	ICAO designator for this airspace
long_name	string	Extended name for this airspace, if applicable
type	string	type of airspace (AB, ES, CS etc.)
airac	smallint	Reference Aeronautical Information Regulation And Control (AIRAC) cycle
geog	MultiPolygon	Area of this airspace

Table A.5: `sector_areas` table

The `crco_areas` view on the `sector_areas` table contains only national airspaces belonging to the Central Route Charges Office (CRCO) area, that is, countries that delegate EUROCONTROL for collecting air navigation charges.

`sector_capacity` table This table contains information on hourly capacities for air sectors and traffic volumes (TVs), that is, the maximum number of entries allowed per hour in the corresponding airspace. If the capacity belongs to a sector, the `tv` field contains the code "AS" ("airspace"). If the capacity is associated to a traffic volume, the field `icao` contains the name of the air sector associated to that traffic volume.

Field Name in DB	Data Type	Info
icao	string	ICAO designator for this sector
date	date (DD:MM:YY)	day of validity for this capacity
start_time	time (HH:MM:SS)	start time of validity for this capacity
end_time	time (HH:MM:SS)	end time of validity for this capacity
capacity	integer	capacity of the sector
tv	string	reference traffic volume
active	Boolean	constraint activation for this AIRAC

Table A.6: `sector_capacity` table

`sector_regulations` table This table contains information on Air Traffic Control (ATC) regulations (i.e., tactical adjustments) applied to sectors' capacities on the specified date and time.

Field Name in DB	Data Type	Info
reg_id	string	ID of the regulation
icao	string	ICAO designator for this sector
tv	string	Traffic volume for this regulation (if present)
start_date	date (DD:MM:YY)	start day of validity for this capacity
end_date	date (DD:MM:YY)	end day of validity for this capacity
start_time	time (HH:MM:SS)	start time of validity for this capacity
end_time	time (HH:MM:SS)	end time of validity for this capacity
sww	integer	Slot window width in minutes
ssw	integer	Slot slice width in minutes
reason	char[1]	Reason for the regulation (see Notes for details)
capacity	integer	capacity of the sector

Table A.7: `sector_regulations` table

The `reason` field contains information on why each regulation was applied. The following single letter codes are used:

Code	Corresponding reason
C	ATC CAPACITY
G	AERODROME CAPACITY
W	WEATHER
T	EQUIPMENT (ATC)
E	EQUIPMENT (NOT ATC)
I	INDUSTRIAL ACTION
M	MILITARY ACTIVITY
R	ATC ROUTING
S	ATC STAFFING
V	ENVIRONMENTAL ISSUES
D	DE-ICING
P	SPECIAL EVENT
O	OTHER REASON
U	UNKNOWN REASON

Table A.8: Codes used for indicating motivations for ATC capacity regulations

`sector_structure` table This table contains information on the hierarchical structure of airspace sectors, specifying which airspace is part of which sector. Both `airspace_icao` and `sector_icao` fields refer to the airspace portions defined in the `sector_areas` table.

Field Name in DB	Data Type	Info
airspace_icao	string	ICAO designator for container airspace
sector_icao	string	ICAO designator for contained airspace
top_fl	integer	top vertical boundary for contained airspace (if applicable)
bottom_fl	integer	bottom vertical boundary for contained airspace (if applicable)
airac	smallint	Reference Aeronautical Information Regulation And Control (AIRAC) cycle

Table A.9: `sector_structure` table

The hierarchical relations between types of airspace portions (referring to the codes used for the `type` field in the `sector_areas` table are the following:

Type of airspace	Code	Component	Code
Elementary sector	ES	Airblock	AB
Collapsed sector	CS	Elementary sector	ES
Sector cluster	CLUS	Elementary sector	ES
ATC Unit Airspace	AUA	Sector cluster Elementary sector	CLUS ES
ATC Unit Airspace Group	AUAG	ATC Unit Airspace	AUA
Military collapsed sector	CRSA	Elementary collapsed sector	ERSA
Area Control Centre	ACC	Sector cluster Elementary sector	CLUS ES
Flight Information Region	FIR	Area Control Centre	ACC
National airspace	NAS	Flight Information Region	FIR
Large region (typically a continent)	AREA	National airspace	NAS

Table A.10: Airspace sectors hierarchy

`traffic_volumes` table This table contains the definitions of Traffic Volumes (TV) for the considered Aeronautical Information Regulation And Control (AIRAC) cycle.

Field Name in DB	Data Type	Info
tv_name	string	Designator for this Traffic Volume (TV)
ref_location	string	ICAO designator for this airspace/airport/navpoint
ref_location_type	string	N = navpoint; A = arp; AS = airport set; S = sector; ACC = acc
min_fl	smallint	Lower altitude for this TV. Empty if ref_location_type = A
max_fl	smallint	Upper altitude for this TV. Empty if ref_location_type = A
role	string[1]	If ref_location_type= A, role can be: D = departure, A = arrival, G = global; empty otherwise
airac	smallint	Reference Aeronautical Information Regulation And Control (AIRAC) cycle

Table A.11: `traffic_volumes` table

tv_flows table This table specifies which Traffic Volume Flows (TVF) rules are included or excluded from each Traffic Volume (TV) definition for the considered Aeronautical Information Regulation And Control (AIRAC) cycle.

Field Name in DB	Data Type	Info
tv	string	Designator for this Traffic Volume (TV)
flow	string	Designator for this Traffic Volume Flow (TVF)
included_excluded	string	Defines whether TVF should be included or excluded from the TV; IN = included ; EX = excluded
airac	smallint	Reference Aeronautical Information Regulation And Control cycle

Table A.12: tv_flows table

tv_flows_det table This table contains the definitions for Traffic Volume Flows (TVF) rules for the considered Aeronautical Information Regulation And Control (AIRAC) cycle.

Field Name in DB	Data Type	Info
flow	string	Designator for this Traffic Volume Flow (TVF) rule
from_to	string	Type of flow: FROM or TO
ref_location	string	ICAO designator for this airspace/airport/navpoint
ref_location_type	string	N = navpoint; A = arp; AS = airport set; S = sector; ACC = acc
airac	smallint	Reference Aeronautical Information Regulation And Control cycle

Table A.13: tv_flows_det table

unit_rates table This table contains the monthly-adjusted unit rates for air navigation charges for each Air Navigation Service Provider (ANSP) belonging to the CRCO area.

Field Name in DB	Data Type	Info
icao	string	ICAO designator for this country
ur	money	unit rate for this country (in Euro)
month	integer	month of validity for this unit rate
year	integer	year of validity for this unit rate

Table A.14: unit_rates table

Network-related tables

Table name	Data source file	Origin
airports	View on points table	
airport_capacity	.ncap, .nact	NEST
airport_regulations	RegPlan.nreg	NEST
navpoints	View on points table	
points	.nnpt and .narp files	NEST
segments	.ase	SAAM dataset

Table A.15: Data sources for network-related tables

The tables listed above are described in detail in the following.

airport_capacity table This table contains information on hourly capacities for airports, that is, the maximum number of departures, arrivals or total movements (as specified by the type field) allowed per hour in the specified aerodrome.

Field Name in DB	Data Type	Info
icao	string[4]	ICAO designator for this airport
tv	string	traffic volume for this capacity
date	date	year of validity for this capacity
start_time	time (HH:MM:SS)	start time for this capacity
end_time	time (HH:MM:SS)	end time for this capacity
capacity	integer	capacity of this airport
role	char[1]	role of constraint: A = Arrival, D= Departure, G = Global
active	boolean	constraint activation for this AIRAC

Table A.16: airport_capacity table

`airport_regulations` table This table contains information on Air Traffic Control (ATC) regulations (i.e., tactical adjustments) applied to airports' capacities on the specified date and time. The `reason` field contains information on why each regulation was applied. The letter codes are the same used for sector regulations (see A.8).

Field Name in DB	**Data Type**	**Info**
reg_id	string	ID of the regulation
icao	string[4]	ICAO designator for this airport
tv	string	Traffic volume for this regulation (if present)
start_date	date (DD:MM:YY)	start day of validity for this capacity
end_date	date (DD:MM:YY)	end day of validity for this capacity
start_time	time (HH:MM:SS)	start time of validity for this capacity
end_time	time (HH:MM:SS)	end time of validity for this capacity
sww	integer	Slot window width in minutes
ssw	integer	Slot slice width in minutes
reason	char[1]	Reason for the regulation (see Notes for details)
capacity	integer	capacity of the sector
role	char[1]	role of constraint: A = Arrival, D= Departure, G = Global

Table A.17: `airport_regulations` table

`points` table This table contains names and geographic coordinates for airports and navigation points. The field `type` specifies to which of either category each entry belongs. Two views are defined on this table, `airports` and `navpoints` to collect only entities belonging to either category.

Field Name in DB	**Data Type**	**Info**
icao	string [5]	ICAO designator for this navpoint/airport
long_name	string	Extended name for this nav./airport, if applicable
type	string	type of point (nav= navpoint, apt= airport)
geog	Point	Coordinates for this navpoint/airport
iata	string	IATA name of the airport if type="apt", empty otherwise
slot_width	smallint	airport departure slot width in minutes if type="apt", empty otherwise

Table A.18: `points` table

`segments` table This table contains data on network segments, that is, couples of navigation points used for defining routes and average flying time between them.

Field Name in DB	**Data Type**	**Info**
orig	string [5]	beginning of the segment
dest	string [5]	end of the segment
avg_travel_time	time interval	avg time for this segment (avg from traffic data)
type	integer	type of segment (e.g. ,CDR)
geog	LineString	Coordinates of this segment

Table A.19: `segments` table

Traffic-related tables

Table name	Data source file	Origin
`aircraft`	Helicopter.txt AircraftWeight.txt	NEST
`intersections`	.t5 M1 files	SAAM/NEST
`traffic`	.exp2 M1 and .so6 M1 files	SAAM/NEST
`traffic_det`	SO6 M1	SAAM/NEST

Table A.20: Data sources for traffic-related tables

The tables listed above are described in detail in the following.

`aircraft` table This table contains relevant data about aircraft, that is, identifier, type and maximum takeoff weight (MTOW).

Field Name in DB	Data Type	Info
name	string	ICAO designator for aircraft
type	string	type of aircraft (heli = helicopter)
weight	float	nominal MTOW in metric tonnes

Table A.21: `aircraft` table

`intersections` table This table contains data about intersections between planned routes (M1) and airspace sectors. All sectors are considered, including national airspaces and collapsed sectors, even sectors that were not included in the actual active configurations.

Field Name in DB	Data Type	Info
uuid	string	CALLSIGN-ADEP-ADES-EOBD-EOBT
sector	string	sector name
entry_date	date (DD:MM:YY)	sector entry date
exit_date	date (DD:MM:YY)	sector exit date
entry_time	time (HH:MM:SS)	sector entry time
exit_time	time (HH:MM:SS)	sector exit time
before_sec_opening	boolean	Y if flight was in sector before it was opened
after_sec_closure	boolean	N if flight was in sector when it was closed
entry_fl	smallint	entry flight level
exit_fl	smallint	exit flight level
distance	float	total distance flown in sector
entry_start_pt	string	start point of entry segment
entry_end_pt	string	end point of entry segment
exit_start_pt	string	start point of exit segment
exit_end_pt	string	end point of exit segment
entry_distance	float	flown distance on entry segment
exit_distance	float	flown distance on exit segment

Table A.22: `intersections` table

`traffic` **table** This table contains general information about flights, like identifiers, departure and arrival dates and times and geographic representation of the route.

Field Name in DB	**Data Type**	**Info**
adep	string [4]	ICAO designator for departure airport
ades	string [4]	ICAO designator for arrival airport
aircraft	string [4]	ICAO designator for aircraft
rfl	smallint	Requested Flight Level
dep_date	date (DD:MM:YY)	Actual departure date
dep_time	time (HH:MM:SS)	Actual departure time
arr_date	date (DD:MM:YY)	Arrival date (= end_date of last flight segment in so6)
arr_time	time (HH:MM:SS)	Arrival time (= end_time of last flight segment in so6)
callsign	string	Call sign, up to 7 chars
airline	string	ICAO designator for airline
uuid	string	CALLSIGN-ADEP-ADES-EOBD-EOBT
planned_dep_date	date (DD:MM:YY)	Planned departure date
planned_dep_time	time (HH:MM:SS)	Planned departure time (should be the expected time shown at the airport)
type	string	type of flight (NB! Empty in DDR2 data sets)
most_pen_reg	string	ID of most penalizing regulation
saamid	integer	SAAM ID of the flight
atfm_delay	smallint	Assigned ATFM delay
continued	integer	SAAM ID of the following flight using the same aircraft
dep_slot	time (HH:MM:SS)	Beginning time of departure airport slot
arr_slot	time (HH:MM:SS)	Beginning time of arrival airport slot
geog	LineString	Coordinates of the route
geog3d	3D LineString	3D Coordinates of the route

Table A.23: `traffic` table

`traffic_det` **table** This table contains detailed information about initially planned (M1) routes, specifically, the sequence of air navigation points and corresponding arrival time and flight level.

Field Name in DB	**Data Type**	**Info**
uuid	string	CALLSIGN-ADEP-ADES-EOBD-EOBT
seq	integer	n-th segment of this flight
start_pt	string	name of first point of the segment
end_pt	string	name of last point of the segment
start_time	time (HH:MM:SS)	segment enter time
end_time	time (HH:MM:SS)	segment exit time
start _date	date (DD:MM:YY)	segment enter date
end_date	date (DD:MM:YY)	segment exit date
start_fl	integer	segment enter flight level
end_fl	integer	segment exit flight level
status	integer	0 = climb 1= descent 2 = cruise
length	float	length of the segment

Table A.24: `traffic_det` table

A.1.2 The saturn_units schema

In order to test the developed optimisation models, some further elaboration was necessary on the data. As illustrated in 4.14, some clustering operations were performed in order to identify distinct routes within a certain distance and to estimate average flight time and operational costs for different types of aircraft on different routes. Additionally, as illustrated in 4.2.1, some elaborations were carried out on capacity data for traffic volumes in order to obtain realistic capacity figures to use for testing the models.
Results of these elaborations were stored in a separate schema, named `saturn_units`, to keep the original DDR2 data stored in the `public` schema separate from data derived from further elaborations.

The tables in the `saturn_units` schema and their content are described in the following:

Table name	Data origin
aircraft_clus	public.aircraft
flights	public.traffic public.traffic_det
routes	public.intersections
routes_det	public.intersections
routes_navpoints	public.traffic_det public.intersections
tv_rules_capacities	public.intersections_actual public.traffic_volumes public.tv_flows public.tv_flows_det

Table A.25: tables in `saturn_units` schema

`aircraft_clus` table This table contains clustering information for aircraft. Each aircraft identifier is assigned to one of thirteen clusters according to closeness of its maximum takeoff weight to the twelve defined in Cook and Tanner (2011), plus one cluster for small aircraft with MTOW $<$ 10T that are used as references for estimating operational costs.

Field Name in DB	Data Type	Info
aircraft	string	aircraft identifier from table public.aircraft
cluster	string	cluster to which this aircraft is assigned

Table A.26: `aircraft_clus` table

Aircraft cluster names are the following: A319, A320, A321, A332, AT43, AT72, B733, B734, B735, B738, B744, B752, B763, DH8D, E190, OTHR.

`flights` table This table contains information on route and aircraft cluster associated with the original initial flight plan (M1) of each flight.

Field Name in DB	Data Type	Info
route_id	string	route unique identifier
uuid	string	flight unique identifier from public.traffic table
aircraft_clus	string	cluster to which the original aircraft of flight identified by uuid belongs

Table A.27: `flights` table

`routes` table This table contains basic information on route clusters and their geographic data representation.

Field Name in DB	Data Type	Info
route_id	string	route unique identifier
origin	string	ICAO identifier of departure airport
dest	string	ICAO identifier of arrival airport
geog	geography	coordinates of the route

Table A.28: `routes` table

`routes_det` table This table contains detailed information on sector intersections for route clusters, namely average entry flight level, entry and fly-through time for all aircraft clusters for each route cluster.

Field Name in DB	Data Type	Info
route_id	string	route unique identifier
seq	integer	sequence number of sectors within this route
sector	string	sector name
*_time	time interval	average flight time in sector for aircraft cluster *
*_entry_time	time interval	average entry time in sector after departure for aircraft cluster *
*_entry_fl	real	average entry flight in sector for aircraft cluster *

Table A.29: `routes_det` table

The asterisk (*) at the beginning of field names implies that that field name starts with the aircraft cluster name, meaning there is one such fields for each of the thirteen clusters used for aircraft type.

`routes_navpoints` **table** This table contains detailed information on navigation point sequence for route clusters, namely average entry flight level and arrival time for all aircraft clusters for each route cluster.

Field Name in DB	**Data Type**	**Info**
route_id	string	route unique identifier
navpoint	string	ICAO identifier of navigation point
*_entry_time	time interval	average arrival time from departure to this navpoint using this route for aircraft cluster *
*_entry_fl	real	average arrival flight level to this navpoint using this route for aircraft cluster *

Table A.30: `routes_navpoints` table

As for the `routes_det` table, the asterisk (*) in the field name implies the aircraft cluster name, meaning there is one such fields for each of the thirteen clusters used for aircraft type.

`tv_rules_capacities` **table** This table contains adjusted capacities for traffic volumes, calculated from actual radar intersection data (from M3 T5 files, stored in the `intersections_actual` table).

Field Name in DB	**Data type**	**Info**
tv_name	string	Name of the TV
ref_location	string	Name of the sector regulated by the TV
acc	string	Name of the ACC to which the sector belongs
hour	smallint	Hour of application of the adjusted capacity; e.g., 18 = from 18:00 to 18:59
capacity	smallint	Adjusted capacity to apply to the sector
date	date	Date on which capacity should be applied

Table A.31: `tv_rules_capacities` table

Appendix B

Database use examples

This chapter presents some examples of the capabilities of the bespoke database. Each of the following four examples presents a different query that is unlikely to be easily performed in NEST. All examples are based on DDR2 data for the set of 25159 flights that operated on Thursday 14th November 2013. Queries were run on an i7 dual core at 1.7 GHz and with 8 GB of Ram and computational times are shown.

Example 1: transit through a specified airspace

Find all flights that departed on 14th November 2013 that transited Greek airspace.

SQL code for the query:

```sql
select *
from traffic
where dep_date = '2013-11-14'
and ST_Intersects(
        geog,
        (select geog
        from sector_areas
        where icao = 'LG')
        )
```

Listing 1: SQL query for example 1: transit through a specified airspace

1118 flights are selected and this result is obtained in 34 869 ms, as shown by the tabular output in PgAdmin3 database management graphic interface (Figure B.1):

The same output format is obtained when executing the query within user-made programs. Most programming languages provide software libraries for easy handling of data retrieved from a database: Figure B.2) shows the visualisation of the output from the query in QGIS 2.

The same data can be obtained in NEST through the Statistics dialogue box (Figure B.3, left), but a considerable amount of manual selection is required to set the query up. Alternatively, the same selection can be performed via the Traffic Queries dialogue, which allows map visualisation and result export in .so6 format, but requires a .so6 file as input (Figure B.3, right).

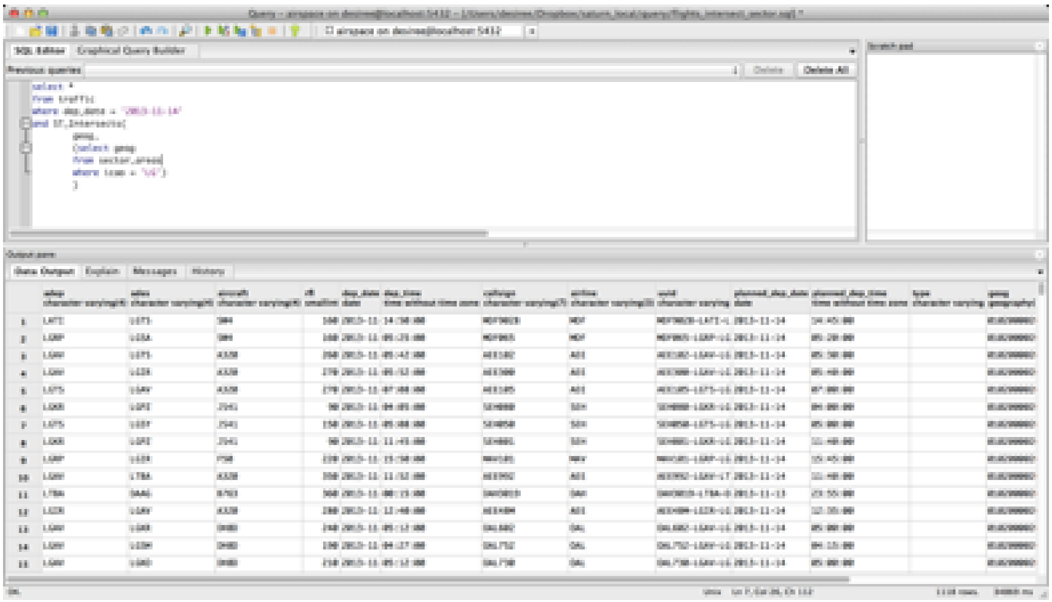


Figure B.1: Query output in tabular form for Database Example 1

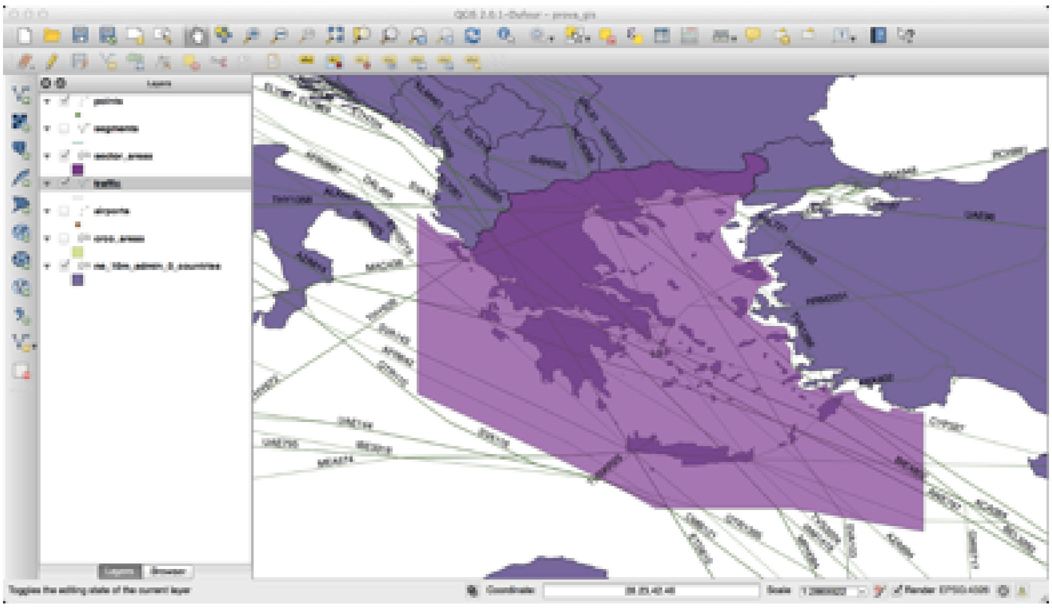


Figure B.2: Visualisation of the query output for Database Example 1

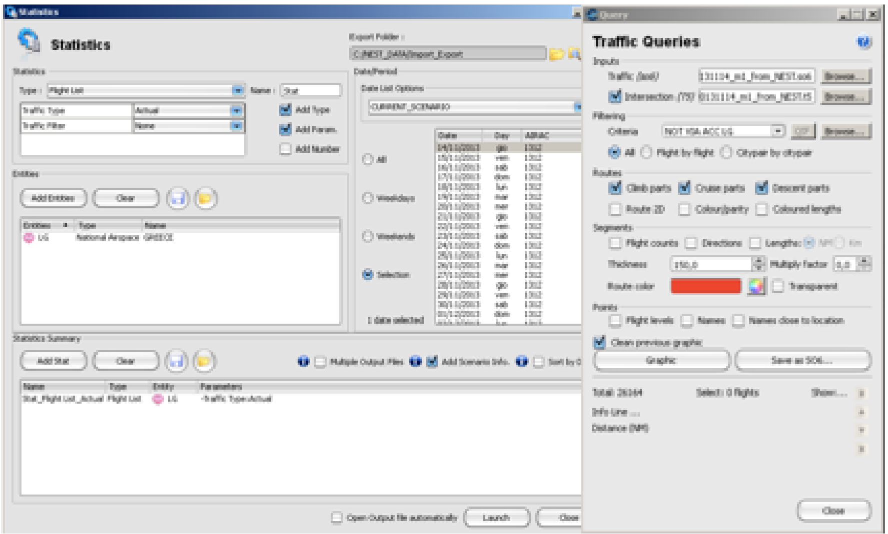


Figure B.3: Query output for Database Example 1 available in NEST Statistics dialogue box (left) and Traffic Queries dialogue box (right)

Example 2: exclusion of CDR routes

Find all flights that departed on 14th November 2013 that did not fly any CDR routes.

SQL code for the query:

```
select *
from traffic
where uuid not in (
        select distinct uuid
        from traffic_det
        where (start_pt, end_pt) in (
                select orig, dest
                from segments
                where type in (20, 21, 22, 23, 40, 41, 42, 43, 60,
                61, 62, 63, 80, 81, 82, 83, 100, 101, 102, 103,
                120, 121, 122, 123)
                )
        and start_date = '2013-11-14')
and dep_date = '2013-11-14'
```

Listing 2: SQL query for example 2: exclusion of CDR routes

The meaning of the numeric codes for segment type are explained in the DDR2 Manual, under the reference for .ASE segment files (Eurocontrol 2014*c*, Section 10.3.7):

```
0=NO, 1=NORMAL, 2=ARRIVAL, 3=DEPARTURE (permanent rte segment)
20=NO, 21=NORMAL, 22=ARRIVAL, 23=DEPARTURE (CDR Generic)
40=NO, 41=NORMAL, 42=ARRIVAL, 43=DEPARTURE (CDR 1)
60=NO, 61=NORMAL, 62=ARRIVAL, 63=DEPARTURE (CDR 2)
80=NO, 81=NORMAL, 82=ARRIVAL, 83=DEPARTURE (CDR 3)
100=NO, 101=NORMAL, 102=ARRIVAL, 103=DEPARTURE (CDR 1+2)
120=NO, 121=NORMAL, 122=ARRIVAL, 123=DEPARTURE (CDR 1+3)
```

The query can therefore be easily modified for excluding only flights using any combination of the above.

18 650 flights (out of 25 159 total for the day) are selected and this result is obtained in 29 542 ms, as shown in tabular output in PgAdmin3 (Figure B.4):

Figure B.5 shows the output visualisation in QGIS 2.

Note that there is no straightforward way to perform this query within NEST.

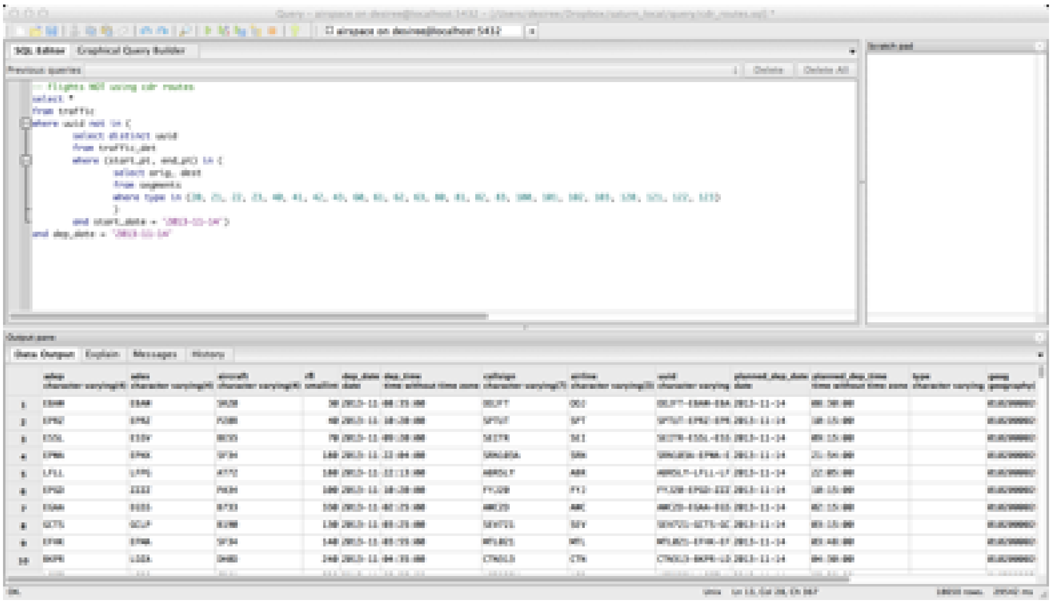


Figure B.4: Query output in tabular form for Database Example 2

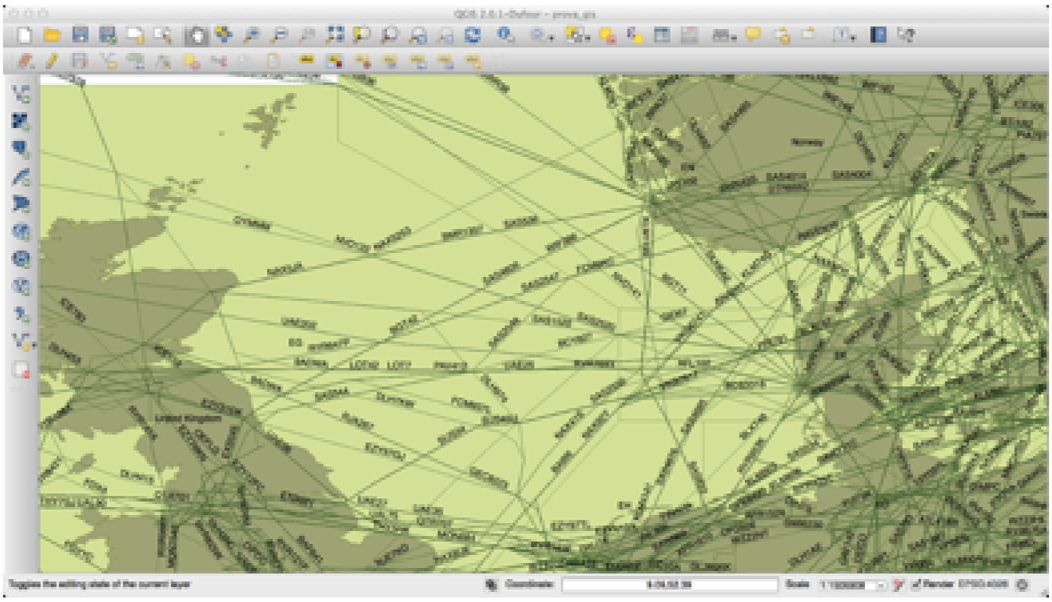


Figure B.5: Visualisation of the query output for Database Example 2

Example 3: most regulated sectors

Find the sectors with the higher number of regulations for AIRAC 1312.

AIRAC 1312 ran from 14 November to 11 December 2013. During this cycle, a total of 993 regulations concerning airspace sectors or related traffic volumes were issued. This query ranks sectors according to the number of regulations they were affected by.

SQL code for the query:

```sql
select count(*) as num, icao, tv
from sector_regulations
where start_date >= '2013-11-14' and start_date <= '2013-12-11'
group by icao,tv order by num desc
```

Listing 3: SQL query for example 3: most regulated sectors

The code above identifies sector names and corresponding traffic volumes, together with the number of regulations affecting them (identified by the `count(*)` function). The order by num desc statement sorts the results in descending order.
In the current example the date interval corresponds to beginning date and end date of AIRAC 1312. The query however can be performed on any interval, like calendar months, seasons or an entire year or reduced to a week or a single day.
With the query 185 entities are selected and this result is obtained in 2499 ms, as shown in the tabular output in PgAdmin3 (Figure B.6).

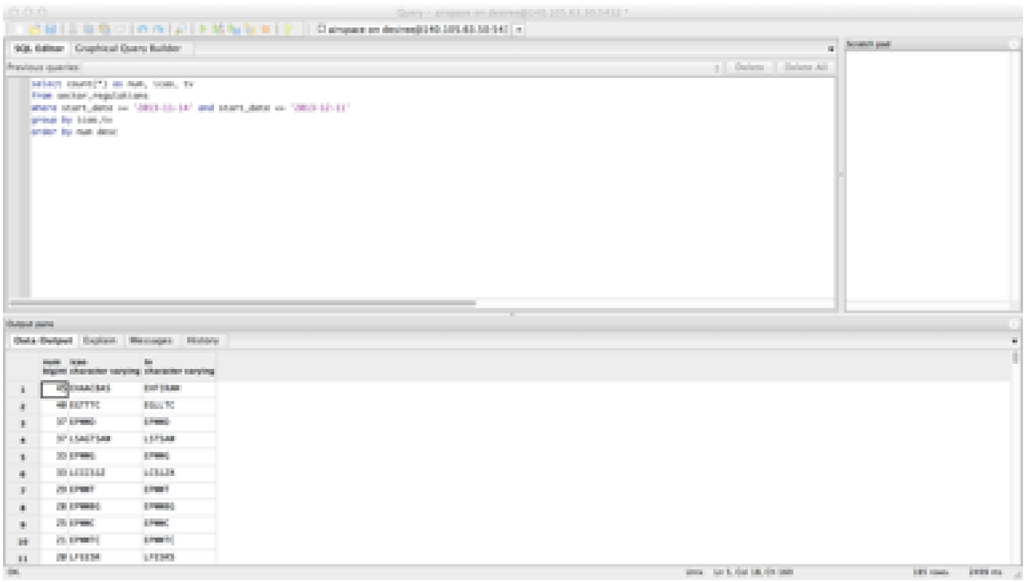


Figure B.6: Query output in tabular form for Database Example 3

The 20 most regulated sectors are visualised in QGIS 2 (Figure B.7).

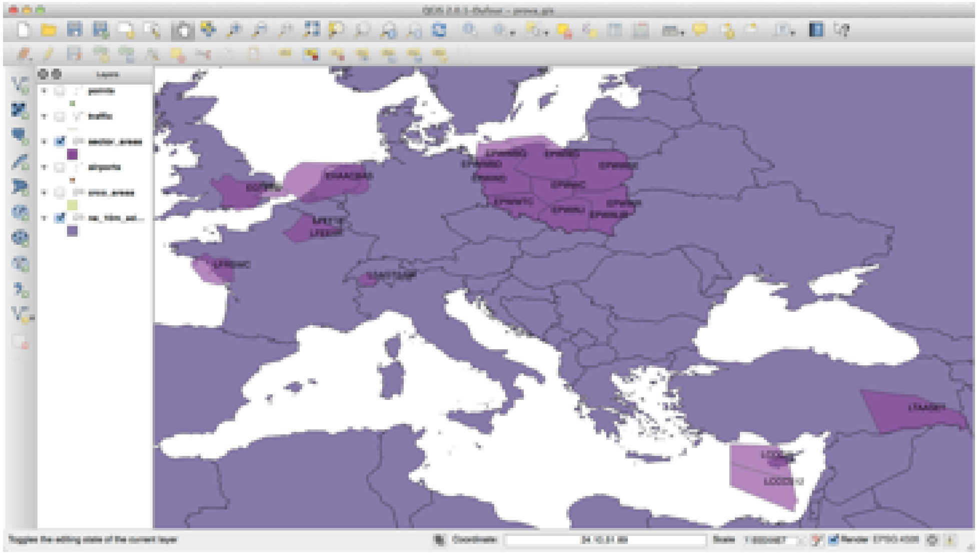


Figure B.7: Visualisation of the query output for Database Example 3

The query could be further refined by filtering on specific regulation reasons only.

Example 4: traffic with mutual exclusion characteristics

Find all origin-destination pairs of flights between the UK and Germany that departed on 14th November 2013, and flew exclusively through either Belgium or the Netherlands, together with the distance flown in each country.

On 14th November 2013, 442 flights are reported between Germany and UK (in either direction): 107 flights also cross the Belgian airspace only, 204 flights the Dutch airspace only, two flights the Danish airspace only, and the remaining 129 flights both the Belgian and Dutch airspaces. These 442 flights connect 173 different origin and destination airport pairs (O/D pairs). This query identifies the O/D pairs that have at least two flights such that one flight crosses the Belgian airspace only and the other one the Dutch airspace only. Out of 173, only six O/D pairs are selected, which comprise 44 flights.

SQL code for the query:

```sql
select uuid, adep, ades, aircraft, geog, weight as ac_weight,
-- Length of Adep - Entry-Pt segment
 ST_Length(
        ST_intersection(
                (select geog from crco_areas where icao = 'EG'),
               (select geog from traffic where uuid = tr.uuid )
        ),true) as EG_distance,
-- Length of Entry - Pt Exit-Pt segment in the Netherlands
 ST_Length(
        ST_intersection(
               (select geog from crco_areas where icao = 'EH'),
                (select geog from traffic where uuid = tr.uuid )
        ),true) as EH_distance,
-- Length of Entry - Pt Exit-Pt segment in Belgium
 ST_Length(
        ST_intersection(
                (select geog from crco_areas where icao = 'EB'),
                (select geog from traffic where uuid = tr.uuid)
        ),true) as EB_distance,
-- Length of Exit-Pt -> Ades segment
 ST_Length(
        ST_intersection(
                (select geog from crco_areas where icao = 'ED'),
                (select geog from traffic where uuid = tr.uuid )
        ),true) as ED_distance
from traffic as tr left join aircraft as ac on tr.aircraft = ac.name
where (adep,ades) in (
        --O/D couples flying through the Netherlands
        select distinct adep, ades
         from traffic
         where dep_date = '2013-11-14'
         and ((adep like 'EG%' and ades like 'ED%') or (adep like 'ED%' and ades like 'EG%'))
         -- exclude flights through Belgium
         and not ST_intersects (geog, (select geog from crco_areas where icao = 'EB'))
         -- exclude flights through Denmark
         and not ST_intersects (geog, (select geog from crco_areas where icao = 'EK'))
         intersect
         --O/D couples flying through Belgium
        select distinct adep, ades
         from traffic
         where dep_date = '2013-11-14'
         and ((adep like 'EG%' and ades like 'ED%') or (adep like 'ED%' and ades like 'EG%'))
         -- exclude flights through the Netherlands
         and not ST_intersects (geog, (select geog from crco_areas where icao = 'EH'))
         -- exclude flights through Denmark
         and not ST_intersects (geog, (select geog from crco_areas where icao = 'EK'))
 )
```

Listing 4: SQL query for example 4: traffic with mutual exclusion characteristics

The first block of the code above selects the required flight data, calculates the distances flown in each national airspace by intersecting the respective areas with each flight route and calculates the length of the intersection (select block); the second block of code (from and where clauses) filters the flights that match the given requirements, that is, crossing Belgian or Dutch airspace only (thus excluding all flights flying through Denmark and through both Belgium and the Netherlands). This filtering too is performed through an intersection between geographical data. 44 flights are selected and this result is obtained in 12 830 ms, as shown in tabular output in PgAdmin3 (Figure B.8).

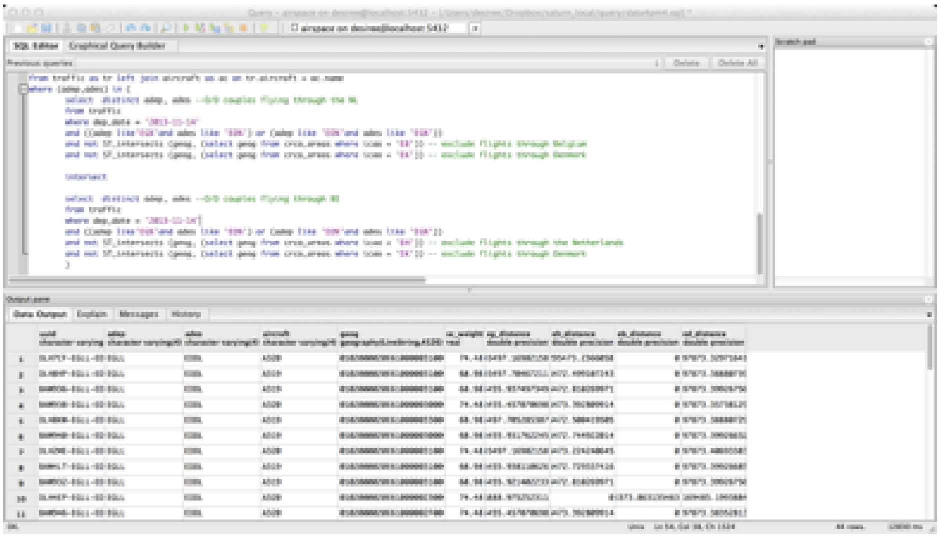


Figure B.8: Query output in tabular form for Database Example 4

Output visualisation in QGIS 2 is shown in Figure B.9.

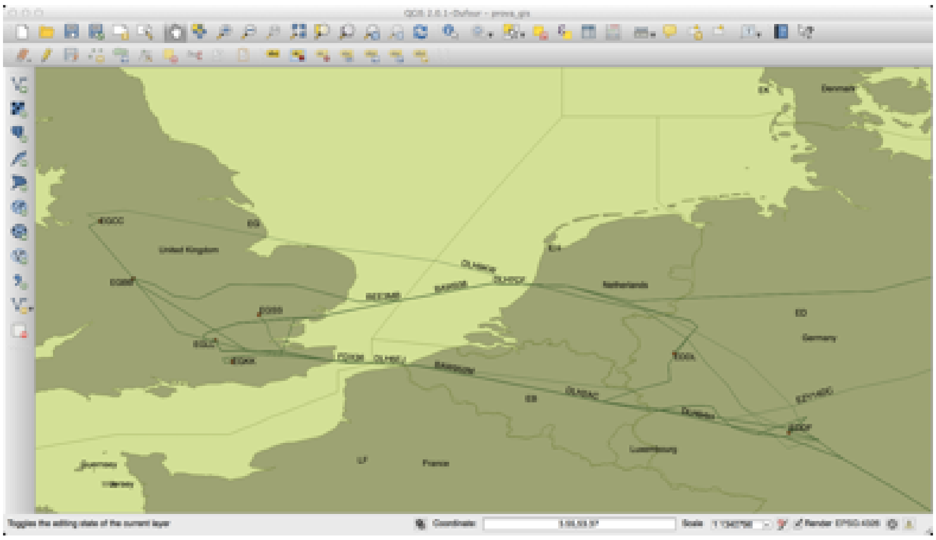


Figure B.9: Visualisation of the query output for Database Example 4

Even though this query is far more complex than those presented in the previous examples, it is however, an easier way to solve the proposed request compared with NEST.

Appendix C

Computational experiments tables

C.1 Parameters setting (Section 5.1)

C.1.1 Shift and revenue neutrality violation (Sec. 5.1.1)

10-Sep-14						
Baseline solution						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Unit rate (€)
1.00E-07	17099.7	13433	61	0	6.66E+07	65.92
1.00E-06	17159.6	13433	61	0	6.66E+07	65.92
1.00E-05	17093	13433	61	0	0	65.92
1.00E-04	17093	13433	61	0	0	65.92
1.00E-03	17093	13433	61	0	0	65.92
1.00E-02	17093	13433	61	0	0	65.92
cPLP one solution, peak threshold = 50%						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Peak rate (€)
1.00E-07	16883.8	14177	45	0	6.84E+07	83.92
1.00E-06	16945.4	14177	45	0	6.84E+07	83.92
1.00E-05	16879.3	14177	45	0	226219	83.92
1.00E-04	16899.6	14177	45	0	226219	83.92
1.00E-03	17093	13433	61	0	0	65.92
1.00E-02	17093	13433	61	0	0	65.92
cPLP one solution, peak threshold = 70%						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Peak rate (€)
1.00E-07	16532.7	14126	40	0	6.68E+07	86.92
1.00E-06	16592.8	14126	40	0	6.68E+07	86.92
1.00E-05	16527.7	14126	40	0	169235	86.92
1.00E-04	16542.9	14126	40	0	169235	86.92
1.00E-03	16695.2	14126	40	0	169235	86.92
1.00E-02	17093	13433	61	0	0	65.92

Table C.1: Solution variation for different values of the K_1 parameter; results for France data instances for September 10, 2014

09-Sep-14						
Baseline solution						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Unit rate (€)
1.00E-07	20645.8	13019	127	0	6.79E+07	65.92
1.00E-06	20706.9	13019	127	0	6.79E+07	65.92
1.00E-05	20639	13019	127	0	0	65.92
1.00E-04	20639	13019	127	0	0	65.92
1.00E-03	20639	13019	127	0	0	65.92
1.00E-02	20639	13019	127	0	0	65.92
cPLP one solution, peak threshold = 50%						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Peak rate (€)
1.00E-07	20226.9	13440	113	0	6.94E+07	80.92
1.00E-06	20289.4	13440	113	0	6.94E+07	80.92
1.00E-05	20222	13440	113	0	202367	80.92
1.00E-04	20240.2	13440	113	0	202367	80.92
1.00E-03	20379.8	13330	115	0	149810	76.92
1.00E-02	20639	13019	127	0	0	65.92
cPLP one solution, peak threshold = 70%						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Peak rate (€)
1.00E-07	20235.7	13329	115	0	6.70E+07	76.92
1.00E-06	20296	13329	115	0	6.70E+07	76.92
1.00E-05	20230	13329	115	0	103075	76.92
1.00E-04	20239.3	13329	115	0	103075	76.92
1.00E-03	20323.8	13219	117	0	84807.5	74.92
1.00E-02	20639	13019	127	0	0	65.92

Table C.2: Solution variation for different values of the K_1 parameter; results for France data instances for September 9, 2014

08-Sep-14						
Baseline solution						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Unit rate (€)
1.00E-07	21362.9	16012	89	4	6.93E+07	65.92
1.00E-06	21425.3	16012	89	4	6.93E+07	65.92
1.00E-05	21356	16012	89	4	0	65.92
1.00E-04	21356	16012	89	4	0	65.92
1.00E-03	21356	16012	89	4	0	65.92
1.00E-02	21356	16012	89	4	0	65.92
cPLP one solution, peak threshold = 50%						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Peak rate (€)
1.00E-07	20824	16973	4	64	7.03E+07	82.92
1.00E-06	20887.3	16973	4	64	7.03E+07	82.92
1.00E-05	20819.1	16973	4	64	211103	82.92
1.00E-04	20838.1	16973	4	64	211103	82.92
1.00E-03	20994.2	16496	4	73	114214	74.92
1.00E-02	21356	16012	4	89	0	65.92
cPLP one solution, peak threshold = 70%						
K_1 value	Objective value	Shift (min.)	N. of sector cap. violations	N. of airport cap. violations	Rev. neut. violation (€)	Peak rate (€)
1.00E-07	20772.8	16502	71	4	6.84E+07	77.92
1.00E-06	20834.4	16502	71	4	6.84E+07	77.92
1.00E-05	20767	16502	71	4	100515	77.92
1.00E-04	20776.1	16502	71	4	100515	77.92
1.00E-03	20866.5	16502	71	4	100515	77.92
1.00E-02	21324.6	16049	85	4	17161.1	67.92

Table C.3: Solution variation for different values of the K_1 parameter; results for France data instances for September 8, 2014

C.1.2 Shift and capacity violations (Sec. 5.1.2)

			12-Sep-14		
K_2	Peak rate	Objective value	Total shift (min)	Sector Cap. violations	Airport Cap. violations
5 (Baseline)	65.92	7478.18	6374	218	6
5	69.92	7478.16	6424	208	6
10	72.92	8500.3	6466	202	6
15	76.92	9459.42	6640	187	6
20	77.92	10389.4	6655	186	6
25	77.92	11319.4	6655	186	6
30	77.92	12249.4	6655	186	6
35	77.92	13179.4	6655	186	6
40	77.92	14109.4	6655	186	6
45	80.92	15024.5	6910	180	6
50	95.92	15919.9	8956	139	5
55	95.92	16614.9	8956	139	5
60	95.92	17309.9	8956	139	5
65	95.92	18004.9	8956	139	5
70	95.92	18699.9	8956	139	5
75	95.92	19394.9	8956	139	5
80	95.92	20089.9	8956	139	5
85	95.92	20784.9	8956	139	5
90	95.92	21479.9	8956	139	5
95	95.92	22174.9	8956	139	5
100	95.92	22869.9	8956	139	5
			10-Sep-14		
K_2	Peak rate	Objective value	Total shift (min)	Sector Cap. violations	Airport Cap. violations
5 (Baseline)	65.92	13744.7	13433	61	0
5	66.92	13731.6	13415	62	0
10	66.92	14041.6	13415	62	0
15	66.92	14351.6	13415	62	0
20	65.92	14659.7	13433	61	0
25	65.92	14964.7	13433	61	0
30	65.92	15269.7	13433	61	0
35	65.92	15574.7	13433	61	0
40	65.92	15879.7	13433	61	0
45	65.92	16184.7	13433	61	0
50	83.92	16433.8	14177	45	0
55	83.92	16658.8	14177	45	0
60	83.92	16883.8	14177	45	0
65	83.92	17108.8	14177	45	0
70	83.92	17333.8	14177	45	0
75	83.92	17558.8	14177	45	0
80	83.92	17783.8	14177	45	0
85	83.92	18008.8	14177	45	0
90	83.92	18233.8	14177	45	0
95	83.92	18458.8	14177	45	0
100	83.92	18680.8	14177	45	0

Table C.4: Solution variation for different values of the K_2 parameter; results for France data instances for September 12 and 10, 2014 with peak threshold set to 50% of nominal capacity

			09-Sep-14		
K_2	Peak rate	Objective value	Total shift (min)	Sector Cap. violations	Airport Cap. violations
5 (Baseline)	65.92	13660.8	13019	127	0
5	65.92	13660.8	13019	127	0
10	65.92	14295.8	13019	127	0
15	65.92	14930.8	13019	127	0
20	65.92	15565.8	13019	127	0
25	65.92	16200.8	13019	127	0
30	76.92	16786.9	13330	115	0
35	76.92	17361.9	13330	115	0
40	76.92	17936.9	13330	115	0
45	76.92	18511.9	13330	115	0
50	76.92	19086.9	13330	115	0
55	76.92	19661.9	13330	115	0
60	80.92	20226.9	13440	113	0
65	80.92	20791.9	13440	113	0
70	80.92	21356.9	13440	113	0
75	80.92	21921.9	13440	113	0
80	80.92	22486.9	13440	113	0
85	80.92	23051.9	13440	113	0
90	80.92	23616.9	13440	113	0
95	80.92	24181.9	13440	113	0
100	80.92	24746.5	13440	113	0
			08-Sep-14		
K_2	Peak rate	Objective value	Total shift (min)	Sector Cap. violations	Airport Cap. violations
5 (Baseline)	65.92	16467.9	16012	89	4
5	65.92	16467.9	16012	89	4
10	65.92	16912.9	16012	89	4
15	65.92	17357.9	16012	89	4
20	65.92	17802.9	16012	89	4
25	68.92	18225.9	16140	83	4
30	68.92	18640.9	16140	83	4
35	68.92	19055.9	16140	83	4
40	74.92	19427	16496	73	4
45	77.92	19788	16582	71	4
50	77.92	20143	16582	71	4
55	78.92	20493	16632	70	4
60	82.92	20824	16973	64	4
65	82.92	21144	16973	64	4
70	82.92	21464	16973	64	4
75	82.92	21784	16973	64	4
80	82.92	22104	16973	64	4
85	82.92	22424	16973	64	4
90	82.92	22744	16973	64	4
95	82.92	23064	16973	64	4
100	82.92	23384	16973	64	4

Table C.5: Solution variation for different values of the K_2 parameter; results for France data instances for September 9 and 8, 2014 with peak threshold set to 50% of nominal capacity

K_2	Peak rate	Objective value	Total shift (min)	Sector Cap. violations	Airport Cap. violations
			12-Sep-14		
5 (Baseline)	65.92	7478.18	6374	218	6
5	70.92	7447.98	6419	203	6
10	70.92	8462.98	6419	203	6
15	75.92	9461.15	6612	189	6
20	80.92	10355.2	6821	176	6
25	80.92	11235.2	6821	176	6
30	80.92	12115.2	6821	176	6
35	80.92	12995.2	6821	176	6
40	80.92	13875.2	6821	176	6
45	80.92	14755.2	6821	176	6
50	92.92	15528.5	8314	144	6
55	92.92	16248.5	8314	144	6
60	92.92	16968.5	8314	144	6
65	92.92	17688.5	8314	144	6
70	98.92	18364.5	9670	124	6
75	98.92	18984.5	9670	124	6
80	98.92	19604.5	9670	124	6
85	98.92	20224.5	9670	124	6
90	98.92	20844.5	9670	124	6
95	102.92	21457.5	10613	114	6
100	102.92	21457.5	10613	114	6
			10-Sep-14		
K_2	Peak rate	Objective value	Total shift (min)	Sector Cap. violations	Airport Cap. violations
5 (Baseline)	65.92	13744.7	13433	61	0
5	66.92	13738.5	13422	62	0
10	68.92	14032.5	13456	57	0
15	68.92	14317.5	13456	57	0
20	68.92	14602.5	13456	57	0
25	68.92	14887.5	13456	57	0
30	78.92	15144.6	13668	49	0
35	78.92	15389.6	13668	49	0
40	78.92	15634.6	13668	49	0
45	78.92	15879.6	13668	49	0
50	78.92	16124.6	13668	49	0
55	86.92	16332.7	14126	40	0
60	86.92	16532.7	14126	40	0
65	86.92	16732.7	14126	40	0
70	86.92	16932.7	14126	40	0
75	86.92	17132.7	14126	40	0
80	86.92	17332.7	14126	40	0
85	86.92	17532.7	14126	40	0
90	96.92	17689.7	15343	26	0
95	96.92	17819.7	15343	26	0
100	96.92	17949.7	15343	26	0

Table C.6: Solution variation for different values of the K_2 parameter; results for France data instances for September 12 and 10, 2014 with peak threshold set to 70% of nominal capacity

			09-Sep-14		
K_2	Peak rate	Objective value	Total shift (min)	Sector Cap. violations	Airport Cap. violations
5 (Baseline)	65.92	13660.8	13019	127	0
5	67.92	13650.6	13009	127	0
10	67.92	14285.6	13009	127	0
15	67.92	14920.6	13009	127	0
20	67.92	15555.6	13009	127	0
25	74.92	16150.7	13219	117	0
30	74.92	16735.7	13219	117	0
35	74.92	17320.7	13219	117	0
40	74.92	17905.7	13219	117	0
45	74.92	18490.7	13219	117	0
50	74.92	19075.7	13219	117	0
55	74.92	19660.7	13219	117	0
60	76.92	20235.7	13329	115	0
65	76.92	20810.7	13329	115	0
70	76.92	21385.7	13329	115	0
75	76.92	21960.7	13329	115	0
80	76.92	22535.7	13329	115	0
85	76.92	23110.7	13329	115	0
90	76.92	23685.7	13329	115	0
95	76.92	23685.7	13329	115	0
100	76.92	24260.7	13329	115	0
			08-Sep-14		
K_2	Peak rate	Objective value	Total shift (min)	Sector Cap. violations	Airport Cap. violations
5 (Baseline)	65.92	16467.9	16012	89	4
5	65.92	16467.9	16012	89	4
10	67.92	16909.8	16049	85	4
15	67.92	17334.8	16049	85	4
20	67.92	17759.8	16049	85	4
25	67.92	18184.8	16049	85	4
30	70.92	18607.8	16137	82	4
35	77.92	18997.8	16502	71	4
40	77.92	19352.8	16502	71	4
45	77.92	19707.8	16502	71	4
50	77.92	20062.8	16502	71	4
55	77.92	20417.8	16502	71	4
60	77.92	20772.8	16502	71	4
65	77.92	21127.8	16502	71	4
70	78.92	21477.8	16567	70	4
75	78.92	21827.8	16567	70	4
80	78.92	22177.8	16567	70	4
85	78.92	22527.8	16567	70	4
90	92.92	22863.9	17723	57	4
95	92.92	23148.9	17723	57	4
100	92.92	23148.9	17723	57	4

Table C.7: Solution variation for different values of the K_2 parameter; results for France data instances for September 9 and 8, 2014 with peak threshold set to 70% of nominal capacity

C.2 Exact resolution on regional instances (Sec.5.3)

Time	N. of active sectors	N. of sec. with <33% load	N. of sec. with 34-66% load	N. of sec. with 67-99% load	N. of sec. with >100% load	N. of unconstrained sectors	N. of peaks (>50% load)
	historical data						
6:00-6:59	90	19	30	21	17	3	55
7:00-7:59	95	13	45	23	7	7	61
8:00-8:59	100	9	45	36	7	3	71
9:00-9:59	99	9	24	45	16	5	81
10:00-10:59	100	10	32	34	19	5	73
11:00-11:59	95	7	21	44	19	4	79
12:00-12:59	95	11	38	32	11	3	70
13:00-13:59	97	9	39	38	3	8	68
14:00-14:59	94	5	32	41	10	6	73
15:00-15:59	94	4	30	34	10	16	66

Table C.8: Sector loads in historical data

Time	N. of active sectors	N. of sec. with <33% load	N. of sec. with 34-66% load	N. of sec. with 67-99% load	N. of sec. with >100% load	N. of unconstrained sectors	N. of peaks (>50% load)
	baseline solution						
6:00-6:59	90	71	16	0	0	3	4
7:00-7:59	95	22	53	13	0	7	39
8:00-8:59	100	14	51	28	2	5	57
9:00-9:59	99	10	33	42	9	5	73
10:00-10:59	100	13	34	40	9	4	73
11:00-11:59	95	7	25	45	11	7	74
12:00-12:59	95	12	50	25	5	3	54
13:00-13:59	97	11	44	35	3	4	59
14:00-14:59	94	26	53	11	1	3	34
15:00-15:59	94	80	1	0	0	13	0

Table C.9: Sector loads in baseline solution

Time	N. of active sectors	N. of sec. with <33% load	N. of sec. with 34-66% load	N. of sec. with 67-99% load	N. of sec. with >100% load	N. of unconstrained sectors	N. of peaks (>50% load)
	cPLP one rate						
6:00-6:59	90	67	19	1	0	3	6
7:00-7:59	95	26	55	9	0	5	31
8:00-8:59	100	14	54	26	2	4	53
9:00-9:59	99	10	32	43	7	7	72
10:00-10:59	100	14	33	43	7	3	72
11:00-11:59	95	7	33	43	4	8	71
12:00-12:59	95	12	52	24	5	2	52
13:00-13:59	97	12	50	28	1	6	55
14:00-14:59	94	22	61	8	0	3	35
15:00-15:59	94	80	1	0	0	13	1

Table C.10: Sector loads in cPLP - one rate solution

Time	cPLP two rates						
	N. of active sectors	N. of sec. with <33% load	N. of sec. with 34-66% load	N. of sec. with 67-99% load	N. of sec. with >100% load	N. of unconstrained sectors	N. of peaks (>50% load)
6:00-6:59	90	73	14	0	0	3	4
7:00-7:59	95	23	53	14	0	5	37
8:00-8:59	100	15	53	26	3	3	55
9:00-9:59	99	10	33	42	7	7	73
10:00-10:59	100	14	33	43	7	3	73
11:00-11:59	95	7	34	42	6	6	70
12:00-12:59	95	13	51	24	5	2	50
13:00-13:59	97	11	47	32	1	6	57
14:00-14:59	94	27	55	8	1	3	34
15:00-15:59	94	79	1	0	0	14	0

Table C.11: Sector loads in cPLP - two rates solution

Appendix D

Acronyms

ACC Area Control Centre

ACE ATM Cost-Effectiveness

ADEP Airport of departure

ADES Airport of destination

AFIL Air Filed Flight Plan

AIRAC Aeronautical Information Regulation and Control

AIS Aeronautical Information Service

ANS Air navigation services

ANSP Air navigation service provider

AO Airline Operator

API Application programming interface

ATC Air Traffic Control

ATCO Air Traffic Control Officer

ATFCM Air traffic flow and capacity management

ATFM Air traffic flow management

ATM Air traffic management

ATS Air traffic services

AU Airspace User

BADA Base of Aircraft Data

BP Binary Programming

CASA Computer Assisted Slot Allocation

CBL Customer Baseline Load (variant of RTP pricing method)

CBCP Credit-based Congestion Pricing (pricing method)

CBP Credit-based Pricing (pricing method)

CDR Conditional route

CFMU Central Flow Management Unit (now NMOC)

CHT Charter (carrier)

CNS Communication Navigation Surveillance

CP (1) Central planner

CP (2) Congestion Pricing (pricing method)

CPLP Centralised peak load pricing (pricing method)

CPP Critical Peak Pricing (pricing method)

CRCO Central Route Charges Office

CS Collapsed sector

CTFM Current Tactical Flight Model

CTOT Calculated takeoff Time

DB Database

DBMS Database management system

DDR2 Demand Data Repository (second phase)

DRP Demand Reduction Programs (pricing method)

EASA European Aviation Safety Agency

E/R En-route

EC European Commission

ECAC European Civil Aviation Conference

EOBT Estimated off-block time

ES Elementary sector

ETFMS Enhanced Tactical Flow Management System

ETO Estimated time over

ETOT Estimated take-off time

ETS EU's Emission trading scheme

EU European Union

EUR Euro (currency)

EXP2 Combination of Expand (exp) and Flight Info (flf) files (DDR2)

FAB Functional Airspace Block

FCFS First come first served

FDC Fully Distributed Costs (pricing method)

FIR Flight Information Region

FL Flight level

FPFS First Planned First Served

FPL Flight plan

FSC Full service carrier

FTFM Filed Tactical Flight Model

GA Genetic Algorithm

GAT General Air Traffic

GIS Geographic Information System

IATA International Air Transport Association

ICAO International Civil Aviation Organization

IER Interruptible Electricity Rate (pricing method)

IFPS Integrated Initial Flight Plan Processing System

IFPU2 IFPS Bretigny unit

IFPZ IFPS Zone

IFR Instrument flight rules

IM Infrastructure Manager

IP Integer Programming

LCC Low-Cost Carrier

LMP Locational Marginal Pricing (pricing method)

LP Linear Programming

M1 Model 1 (DDR2)

M2 Model 2 (DDR2)

M3 Model 3 (DDR2)

MIP Mixed Integer Programming

MTOW Maximum Take-Off Weight

MUAC Maastricht Upper Area Control Centre (ACC)

NDP Network Design Problem (bilevel optimisation problem)

NEST NEtwork Strategic Tool (software)

NM (1) Nautical Mile

NM (2) Network Manager

NMOC Network Manager Operations Centre (formerly CFMU)

NMP Non-Monetary Pricing (pricing method)

NOP Network Operations Plan

NPP Network Pricing Problem (bilevel optimisation problem)

NREG Regulation plan file

NSA National Supervisory Authorities

OAP Operational Air Traffic

OD Origin-destination

pgAdmin Open source software graphical user interface for PostgreSQL

PLP Peak Load Pricing (pricing method)

PostGIS Open source software support for geographic objects in PostgreSQL

PostgreSQL Open source object-relational database system

PRB Performance Review Body

QGIS Open source Geographic Information System

QoS Quality of Service

RAD Route Availability Document

REG Regional carrier

RFL Requested flight level

RP Ramsey Pricing

RTFM Regulated Tactical Flight Model

RTP Real Time Pricing (pricing method)

SA Slot Analyzer

SAAM System for traffic Assignment and Analysis at a Macroscopic level

SATURN Strategic Allocation of Traffic Under Redistribution in the Network (WP-E Project)

SES Single European Sky

SESAR Single European Sky ATM Research

SID Standard instrument departure

SJU SESAR Joint Undertaking

SO System optimum

SO6 Traffic segments file (DDR2)

SOBT Scheduled off-block time

SQL Structured Query Language

SRMC Short Run Marginal Costs (pricing method)

STAR Standard instrument arrival

STATFOR Statistics and Forecast Service (Eurocontrol)

TDP Time-dependent usage pricing (pricing method)

TIP Time-independent usage pricing (pricing method)

TMA Terminal manoeuvring area

TOU Time of usage (pricing method)

TV Traffic volume

TVF Traffic volume flow

UAC Upper area control centre

UIR Upper Information Region

UO User Optimum

UTA Upper (traffic) control area

UTC Coordinated universal time

VAT Value-added tax

VFR Visual flight rules

Bibliography

Abdulaal, M. & LeBlanc, L. J. (1979), 'Continuous equilibrium network design models', *Transportation Research Part B: Methodological* **13**(1), 19–32.

Albalate, D., Bel, G. & Fageda, X. (2009), 'Privatization and regulatory reform of toll motorways in Europe', *Governance* **22**(2), 295–318.

Anania, L. & Solomon, R. J. (1997), Flat: the minimalist price, *in* 'Internet economics', MIT press, pp. 91–118.

Andreatta, G. & Odoni, A. (2001), 'Preliminary investigation on market-based demand management strategies for European airports and air routes. Report for EUROCONTROL Experimental Centre.'.

Avlonitis, G. J. & Indounas, K. A. (2005), 'Pricing objectives and pricing methods in the services sector', *Journal of Services Marketing* **19**(1), 47–57.

Barta, V. & Kazda, P. A. (2014), Simulation of Bosnia and Herzegovina airspace sectorization and its influence on FAB CE, *in* 'Procedia from The International Conference on Air Transport INAIR 2014', pp. 13–18.

Bertsimas, D., Lulli, G. & Odoni, a. (2011), 'An Integer Optimization Approach to Large-Scale Air Traffic Flow Management', *Operations Research* **59**(1), 211–227.

Black, J. W. & Larson, R. C. (2007), 'Strategies to Overcome Network Congestion in Infrastructure Systems', *Journal of Industrial and Systems Engineering* **1**(2), 97–115.

Borenstein, S., Jaske, M. & Rosenfeld, A. (2002), Dynamic pricing, advanced metering, and demand response in electricity markets, Technical report, UC Berkeley.

Bouhtou, M., van Hoesel, S., van der Kraaij, A. F. & Lutton, J.-L. (2007), 'Tariff optimization in networks', *INFORMS journal on computing* **19**(3), 458–469.

Bracken, J. & McGill, J. (1973), 'Mathematical programs with optimization problems in the constraints', *Operations Research* (21), 37–44.

Brotcorne, L., Labbé, M., Marcotte, P. & Savard, G. (2000), 'A bilevel model and solution algorithm for a freight tariff-setting problem', *Transportation Science* **34**(3), 289–302.

Brotcorne, L., Labbé, M., Marcotte, P. & Savard, G. (2001), 'A bilevel model for toll optimization on a multicommodity transportation network', *Transportation Science* **35**(4), 345–358.

Carlson, L. & Nolan, J. (2005), 'Pricing access to rail infrastructure in Canada', *Canadian Journal of Administrative Sciences* **22**(1), 45–57.

Castelli, L., Labbé, M. & Violin, A. (2013), 'A network pricing formulation for the revenue maximization of european air navigation service providers', *Transportation Research Part C: Emerging Technologies* **33**, 214–226.

Castelli, L. & Ranieri, A. (2007), Air navigation service charges in Europe, *in* 'Proceedings of the 7th USA–Europe ATM Research & Development Seminar, Barcelona, Spain'.

Clark, D. D. (1997), Internet cost allocation and pricing, *in* 'Internet economics', MIT Press, pp. 215–252.

Cocchi, R., Shenker, S., Estrin, D. & Zhang, L. (1993), 'Pricing in computer networks: Motivation, formulation, and example', *IEEE/ACM Transportation Networks* **1**(6), 614–627.

Cook, A. J. (2007), *European air traffic management: principles, practice, and research*, Ashgate Publishing, Ltd.

Cook, A. J. & Tanner, G. (2015), European airline delay cost reference values, 2015 update, Technical report.

Costanzo, S., Castelli, L. & Turco, A. (2014), MOGASI: A multi-objective genetic algorithm for efficiently handling constraints and diversified decision variables, *in* 'Engineering Optimization 2014', CRC Press, pp. 99–104.

Dafermos, S. & Sparrow, F. T. (1971), 'Optimal resource allocation and toll patterns in user-optimised transport networks', *Journal of Transport Economics and Policy* **5**(2), 184–200.

de Groot, R. S., Wilson, M. a. & Boumans, R. M. (2002), 'A typology for the classification, description and valuation of ecosystem functions, goods and services', *Ecological Economics* **41**(3), 393–408.

de Palma, A. & Lindsey, R. (2011), 'Traffic congestion pricing methodologies and technologies', *Transportation Research Part C: Emerging Technologies* **19**(6), 1377–1399.

Delgado, L. (2015), European route choice determinants, *in* 'Proceedings of the Eleventh USA/Europe Air Traffic Management Research and Development Seminar, Lisbon, Portugal'.

Deschinkel, K., Farges, J.-L. & Delahaye, D. (2002), 'Optimizing and assigning price levels for air traffic management', *Transportation Research Part E: Logistics and Transportation Review* **38**(3), 221–237.

Dewez, S., Labbé, M. & Marcotte, P. (2008), 'New formulations and valid inequalities for a bilevel pricing problem', *Operations Research Letters* **36**(2), 141–149.

Dong, J., Mahmassani, H. S., Erdoğan, S. & Lu, C.-C. (2011), 'State-dependent pricing for real-time freeway management: Anticipatory versus reactive strategies', *Transportation Research Part C: Emerging Technologies* **19**(4), 644–657.

Duthaler, C. & Finger, M. (2011), Congestion management in the european electricity sector: Lessons from the european air transport sector, *in* 'Energy Market (EEM), 2011 8th International Conference on the European', IEEE, pp. 2–7.

Eliasson, J., Hultkrantz, L., Nerhagen, L. & Rosqvist, L. S. (2009), 'The stockholm congestion–charging trial 2006: Overview of effects', *Transportation Research Part A: Policy and Practice* **43**(3), 240–250.

EU (2005), *European Conference of Ministers of Transport: railway reform & charges for the use of infrastructure*, OECD.

Eurocontrol (2011), Conditions of application of the route charges system and conditions of payment, Central Route Charges Office, Doc. 11.60.02, May 2011., Technical report.

Eurocontrol (2014*a*), ATM Cost-Effectiveness (ACE) 2012 Benchmarking Report with 2013-2017 outlook Prepared by the Performance Review Unit (PRU) with the ACE Working Group, Technical report.

Eurocontrol (2014*b*), Customer guide to charges, Central Route Charges Office, January 2014 - Version 8., Technical report.

Eurocontrol (2014*c*), DDR2 Reference Manual v.1.0.1, Technical report.

Eurocontrol (2014*d*), European Network Operations Plan 2014-2018/19, Technical report.

Eurocontrol (2015*a*), ATM Cost-Effectiveness (ACE) 2013 Benchmarking Report with 2014-2018 outlook Prepared by the Performance Review Unit (PRU) with the ACE Working Group, Technical report.

Eurocontrol (2015*b*), EUROCONTROL Seven-Year Forecast February 2015 7-year IFR Flight Movements and Service Units Forecast : 2015-2021, Technical report.

Eurocontrol (2015*c*), Performance Review Report: An Assessment of Air Traffic Management in Europe during the Calendar Year 2014, Technical report.

Eurocontrol (2015*d*), Principles for establishing the cost-base for en route charges and the calculation of the unit rates, April 2015 update, Technical report.

Eurocontrol (2015*e*), Standard Inputs for EUROCONTROL Cost-Benefit Analyses, Technical report.

European Commission (2006), Commission Regulation (EC) No 1794/2006 of 6 December 2006, laying down a common charging scheme for air navigation services., Technical report.

European Commission (2010), Commission Regulation (EU) No 1191/2010 of 16 December 2010, amending Regulation (EC) No. 1794/2006 laying down a common charging scheme for air navigation services., Technical report.

Falkner, M., Devetsikiotis, M. & Lambadaris, I. (2000), 'An overview of pricing concepts for broadband IP networks', *Communications Surveys Tutorials, IEEE* **3**(2), 2–13.

Fan, W. & Jiang, X. (2013), 'Tradable mobility permits in roadway capacity allocation: Review and appraisal', *Transport Policy* **30**, 132–142.

Garcia, A., Hong, M. & Barrera, J. (2012), 'cap and trade for congestion control', *Dynamic Games and Applications* **2**(3), 280–293.

Gibbens, R. J. & Kelly, F. P. (1999), 'Resource pricing and the evolution of congestion control', *Automatica* **35**(12), 1969–1985.

Gulipalli, P. K., Kalmanje, S. & Kockelman, K. (2011), 'Credit-based congestion pricing: expert expectations and guidelines for application', *Journal of the Transportation Research Forum* **47**(2).

Gulipalli, P. K. & Kockelman, K. M. (2008), 'Credit-based congestion pricing: A Dallas-Fort Worth application', *Transport Policy* **15**(1), 23–32.

Gupta, A., Stahl, D. O. & Whinston, A. B. (1997), 'A stochastic equilibrium model of Internet pricing', *Journal of Economic Dynamics and Control* **21**(4-5), 697–722.

Hansen, P., Jaumard, B. & Savard, G. (1992), 'A new branch-and-bound rules for linear bilevel programming', *SIAM Journal on Scientific and Statistical Computing* **5**(13), 1194–1217.

Harrod, S. (2013), 'Auction pricing of network access for North American railways', *Transportation Research Part E: Logistics and Transportation Review* **49**(1), 176–189.

Heilporn, G. (2008), Network pricing problems: complexity, polyhedral study and solution approaches, PhD thesis, Université Libre de Bruxelles.

Hoffman, K., Berardino, F. & Hunter, G. (2013), 'Congestion pricing applications to manage high temporal demand for public services and their relevance to air space management', *Transport Policy* **28**(C), 28–41.

Hogan, W. W. (1998), Competitive electricity market design: A wholesale primer, Technical report, John F. Kennedy School of Government, Harvard University.

Hylen, B. (2001), 'Access to the rail network in some european countries', *Swedish National Road and Transport Research Institute, VTI Notat 47A-2001* .

Jeroslow, R. (1985), 'The polynomial hierarchy and a simple model for competitive analysis', *Mathematical Programming* **32**, 146–164.

Jiang, H. & Jordan, S. (1995), 'The role of price in the connection establishment process', *European Transactions on Telecommunications and Related Technologies* **6**(4), 421–430.

Joe-Wong, C., Ha, S. & Chiang, M. (2011), Time-dependent broadband pricing: Feasibility and benefits, *in* 'Distributed Computing Systems (ICDCS), 2011 31st International Conference on', IEEE, pp. 288–298.

Jordan, S. & Jiang, H. (1995), 'Connection establishment in high-speed networks', *Selected Areas in Communications, IEEE Journal on* **13**(7), 1150–1161.

Jovanović, R., Tošić, V., Čangalović, M. & Stanojević, M. (2014), 'Anticipatory modulation of air navigation charges to balance the use of airspace network capacities', *Transportation Research Part A: Policy and Practice* **61**, 84–99.

Kahn, A. E., Cramton, P., Porter, R. H., Tabors, R. D. et al. (2001), Pricing in the california power exchange electricity market: should california switch from uniform pricing to pay-as-bid pricing?, Technical report, Blue Ribbon Panel Report, California Power Exchange.

Kelly, F. P., Maulloo, A. K. & Tan, D. K. (1998), 'Rate control for communication networks: shadow prices, proportional fairness and stability', *Journal of the Operational Research society* **49**, 237–252.

Keon, N. J. & Anandalingam, G. (2003), 'Optimal pricing for multiple services in telecommunications networks offering quality-of-service guarantees', *IEEE/ACM Trans. Netw.* **11**(1), 66–80.

Kockelman, K. M. & Kalmanje, S. (2005), 'Credit-based congestion pricing: a policy proposal and the publics response', *Transportation Research Part A: Policy and Practice* **39**(7), 671–690.

Kockelman, K. M. & Lemp, J. D. (2011), 'Anticipating new-highway impacts: Opportunities for welfare analysis and credit-based congestion pricing', *Transportation Research Part A: Policy and Practice* **45**(8), 825–838.

Labbé, M., Marcotte, P. & Savard, G. (1998), 'A bilevel model of taxation and its application to optimal highway pricing', *Management science* **44**(12-part-1), 1608–1622.

Labbé, M. & Violin, A. (2013), 'Bilevel Programming and Price Setting Problems.', *4OR quarterly journal of the Belgian, French and Italian Operations Research Societies* **11**(1), 1–30.

Laurits R. Christensen Associates, I. (2009), Study of Competition in the Railroad Industry. Analysis of competition, capacity and service quality, Technical Report November, The Surface Transportation Board Washington, DC.

Leal de Matos, P. et al. (2001), 'Yield management for privatised air traffic control?', *Journal of the Operational Research Society* **52**(8), 888–895.

LeBlanc, L. J. (1975), 'An algorithm for the discrete network design problem', *Transportation Science* **9**(3), 183–199.

Letzler, R. (2010), 'Using incentive preserving rebates to increase acceptance of critical peak electricity pricing', *Center for the Study of Energy Markets working paper* **162**.

Li, F. (2007), 'Continuous locational marginal pricing (clmp)', *Power Systems, IEEE Transactions on* **22**(4), 1638–1646.

Lindsey, R. (2006), 'Do economists reach a conclusion on highway pricing? The intellectual history of an idea', *Econ Journal Watch* **3**(2), 292–379.

Litman, T. (2005), 'London congestion pricing–implications for other cities', *CESifo DICE Report* **3**(3), 17–21.

Lou, Y., Yin, Y. & Laval, J. A. (2011), 'Optimal dynamic pricing strategies for high-occupancy/toll lanes', *Transportation Research Part C: Emerging Technologies* **19**(1), 64–74.

Lufthansa (2014), Lufthansa Annual report 2014, Technical report.

Luo, Z.-Q., Pang, J.-S. & Ralph, D. (1996), *Mathematical programs with equilibrium constraints*, Cambridge University Press.

Macário, R., Teixeira, P., Sánchez-Borrà, M., López-Pita, A. & Casas Esplugas, C. (2007), Rail infrastructure pricing for intercity passenger services in europe: Possible impacts on the railways competitive framework, *in* 'International Conference Series on Competition and Ownership in Land Passenger Transport, Hamilton Island, Queensland, Australia'.

Mackie-mason, J. K. & Varian, H. R. (1995), 'Pricing Congestible Network Resources', *IEEE Journal on Selected Areas in Communications* **13**(7), 1141–1149.

Marcotte, P. & Patriksson, M. (2007), 'Traffic equilibrium', *Handbooks in Operations Research and Management Science* **14**, 623–713.

Masuda, Y. & Whang, S. (1999), 'Dynamic pricing for network service: Equilibrium and stability', *Management Science* **45**(6), 857–869.

Meeus, L., Purchala, K. & Belmans, R. (2005), 'Development of the internal electricity market in europe', *The Electricity Journal* **18**(6), 25–35.

Mohring, H. & Harwitz, M. (1962), *Highway benefits: An analytical framework*, Northwestern University Press.

Murphy, J. & Murphy, L. (1994), Bandwidth allocation by pricing in atm networks., *in* 'Proceedings of the IFIP TC6 Second International Conference on Broadband Communications II', pp. 333–351.

Nash, C. A. & Matthews, B. (2003), Rail infrastructure charges–the issue of scarcity, *in* 'Proceedings of the First Conference on Railroad Industry Structure, Competition and Investment', pp. 7–8.

Nash, J. F. et al. (1950), 'Equilibrium points in n-person games', *Proc. Nat. Acad. Sci. USA* **36**(1), 48–49.

Odlyzko, A. (1999), Paris metro pricing for the internet, *in* 'Proceedings of the 1st ACM Conference on Electronic Commerce', EC '99, ACM, New York, NY, USA, pp. 140–147.

Parris, C., , Parris, C. & Ferrari, D. (1992), A resource based pricing policy for real-time channels in a packet-switching network., Technical report, International Computer Science Institute Berkeley, CA.

Patriksson, M. (1994), *The traffic assignment problem: models and methods*, Courier Dover Publications.

Peter, B. (2003), Railway infrastructure: pricing and investment, Internal Report, Technical report, TU Berlin.

Pigou, A. C. (1920), 'The economics of welfare', *McMillan&Co., London* .

R. G. Lipsey, K. L. (1956), 'The general theory of second best', *The Review of Economic Studies* **24**(1), 11–32.

Raffarin, M. (2004), 'Congestion in european airspace: a pricing solution?', *Journal of Transport Economics and Policy* **38**(1), 109–125.

Roozbehani, M., Dahleh, M. A. & Mitter, S. K. (2012), 'Volatility of power grids under real-time pricing', *Power Systems, IEEE Transactions on* **27**(4), 1926–1940.

Rothengatter, W. (2003), 'How good is first best? marginal cost and other pricing principles for user charging in transport', *Transport policy* **10**(2), 121–130.

Ryanair (2015), Financial highlights - Annual report 2015, Technical report.

SATURN (2015*a*), D4.2-Pure Pricing Mechanisms, Technical report.

SATURN (2015*b*), D6.5-Final report (Dissemination), Technical report.

Sgouridis, S. & Kennedy, S. (2010), 'Tangible and fungible energy: Hybrid energy market and currency system for total energy management. A Masdar City case study', *Energy Policy* **38**(4), 1749–1758.

Sheffi, Y. (1985), *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall, Englewood Cliffs, NJ.

Shenker, S., Clark, D., Estrin, D. & Herzog, S. (1996), Pricing in computer networks: Reshaping the research agenda, Technical report, ACM Computer Communication Review.

Shimizu, K. & Aiyoshi, E. (1981), 'A new computational method for Stackelberg and min-max problems by use of a penalty method', *IEEE Transactions on Automatic Control* **26**(2), 460–466.

Shimizu, K., Ishizuka, Y. & Bard, J. (1997), *Nondifferentiable and Two-Level Mathematical Programming*, Kluwer Academic Publishers.

Steer Davies Gleave (2015), Policy options for the modulation of charges in the Single European Sky , Technical report.

Tian, L.-J., Yang, H. & Huang, H.-J. (2013), 'Tradable credit schemes for managing bottleneck congestion and modal split with heterogeneous users', *Transportation Research Part E: Logistics and Transportation Review* **54**, 1–13.

Van den Berg, V. A. & Verhoef, E. T. (2014), 'Congestion pricing in a road and rail network with heterogeneous values of time and schedule delay', *Transportmetrica A: Transport Science* **10**(5), 377–400.

Van Hoesel, S. (2008), 'An overview of Stackelberg pricing in networks', *European Journal of Operational Research* **189**(3), 1393–1402.

Vicente, L. N. & Calamai, P. H. (1994), 'Bilevel and multilevel programming: A bibliography review', *Journal of Global optimization* **5**(3), 291–306.

Vicente, L., Savard, G. & Júdice, J. (1994), 'Descent approaches for quadratic bilevel programming', *Journal of Optimization Theory and Applications* **81**(2), 379–399.

Violin, A. (2014), Mathematical programming approaches to pricing problems, PhD thesis, Université Libre de Bruxelles.

Von Stackelberg, H. (1952), *The Theory of Market Economy*, Oxford University Press.

Walters, A. A. (1961), 'The theory and measurement of private and social cost of highway congestion', *Econometrica: Journal of the Econometric Society* pp. 676–699.

Wang, X., Yang, H., Zhu, D. & Li, C. (2012), 'Tradable travel credits for congestion management with heterogeneous users', *Transportation Research Part E: Logistics and Transportation Review* **48**(2), 426–437.

Wardrop, J. G. (1952), Some theoretical aspects of road traffic research., *in* 'ICE Proceedings: Engineering Divisions', Vol. 1, Thomas Telford, pp. 325–362.

Wu, D., Yin, Y., Lawphongpanich, S. & Yang, H. (2012), 'Design of more equitable congestion pricing and tradable credit schemes for multimodal transportation networks', *Transportation Research Part B: Methodological* **46**(9), 1273–1287.

Yang, H. & Wang, X. (2011), 'Managing network mobility with tradable credits', *Transportation Research Part B: Methodological* **45**(3), 580–594.

Yin, Y. & Lou, Y. (2009), 'Dynamic tolling strategies for managed lanes', *Journal of Transportation Engineering* **135**(2), 45–52.

Zanin, M., Perez, D., Chatterjee, K., Kolovos, D. S., Paige, R. F., Horst, A. & Rumpe, B. (2014), 'On demand data analysis and filtering for inaccurate flight trajectories', *arXiv preprint arXiv:1409.2304* .

Zarnikau, J. W. (2010), 'Demand participation in the restructured Electric Reliability Council of Texas market', *Energy* **35**(4), 1536–1543.

Zhang, X. & Yang, H. (2004), 'The optimal cordon-based network congestion pricing problem', *Transportation Research Part B: Methodological* **38**(6), 517–537.

Ringraziamenti

Quando inizi a scrivere la pagina dei ringraziamenti di una tesi, ti rendi conto che un altro capitolo di vita si sta per chiudere. Qualcuno penserà che quando la tesi in questione è la terza, uno ci abbia anche ormai fatto l'abitudine. Non è così. C'è sempre quel momento di esitazione ("Ma ci siamo veramente?"), seguito da una sana dose di incertezza ("Domani mi sveglierò e non saprò cosa fare..") e di dubbi su quel che seguirà ("E adesso?"). Inevitabilmente si guarda al passato, si cerca di ricordarsi gli inizi, e forse allora ci si rende conto di avere qualche anno in più e qualche granello di ignoranza in meno sulle spalle. Un equo scambio, il più delle volte.

E poi ci si rende conto che non solo merito proprio. Che se anche la tesi porta il tuo nome, non esisterebbe senza le persone che ti hanno aiutato a svilupparla, che ti hanno spronato quando la pigrizia o lo sconforto ti aleggiavano attorno, che erano là ad ascoltarti quando avevi bisogno di parlare o che semplicemente ti trascinavano fuori dalla tana quando veramente non ne potevi più.

A tutte queste persone è dedicata la pagina dei ringraziamenti. Non abbiatemene a male se non la allungo ulteriormente citando i meriti di ciascuno. Tanto chi siete lo sapete già benissimo! Ce l'abbiamo fatta.. Grazie di cuore!