Air traffic management (ATM) comprises a highly complex socio-technical system that keeps air traffic flowing safely and efficiently, worldwide, every minute of the year. Over the last few decades, several ambitious ATM performance improvement programmes have been undertaken. Such programmes have mostly delivered local technological solutions, whilst corresponding ATM performance improvements have fallen short of stakeholder expectations. In hindsight, this can be substantially explained from a complexity science perspective: ATM is simply too complex to address through classical approaches such as system engineering and human factors. In order to change this, complexity science has to be embraced as ATM’s ‘best friend’. The applicability of complexity science paradigms to the analysis and modelling of future operations is driven by the need to accommodate long-term air traffic growth within an already-saturated ATM infrastructure.

*Complexity Science in Air Traffic Management* is written particularly, but not exclusively, for transport researchers, though it also has a complementary appeal to practitioners, supported through the frequent references made to practical examples and operational themes such as performance, airline strategy, passenger mobility, delay propagation and free-flight safety. The book should also have significant appeal beyond the transport domain, due to its intrinsic value as an exposition of applied complexity science and applied research, drawing on examples of simulations and modelling throughout, with corresponding insights into the design of new concepts and policies, and the understanding of complex phenomena that are invisible to classical techniques.

**Andrew Cook** is a Principal Research Fellow at the University of Westminster, where he manages the Department’s ATM research. He has been a research fellow of the University for over 20 years. Leader of ten projects for EUROCONTROL, Andrew is currently engaged in several projects in SESAR’s Workpackage E (long-term and innovative research), in addition to air transport industry and EU framework research coordination. Other responsibilities include serving on the programme committee for the annual SESAR innovation conferences and on the Editorial Board of the Journal of Aerospace Operations. Andrew has published widely in ATM, across a number of domains, and frequently reviews journal submissions for Elsevier. He was editor of the first book on European ATM, published by Ashgate. Andrew holds a BA (Hons) Natural Sciences, with supplementary Distinction, and a DPhil Physiological Sciences, both from the University of Oxford.

**Damián Rivas** received his first degree in Aeronautical Engineering from the Polytechnic University of Madrid and has a PhD from Case Western Reserve University (Cleveland, Ohio). He is currently Professor in the Department of Aerospace Engineering of the Universidad de Sevilla, teaching flight mechanics and air navigation; Director of the Group of Aerospace Engineering of the University of Seville; scientific coordinator of the SESAR WP-E Complex-World Network; and a member of the programme committee of the annual SESAR Innovation Days conferences. His main research interests are aircraft trajectory optimisation, aircraft conflict detection and resolution, and uncertainty propagation in air traffic management.
This page intentionally left blank
Complexity Science in Air Traffic Management

Edited by

ANDREW COOK and DAMIÁN RIVAS
# Contents

*List of Figures*  ix

*List of Tables*  xi

*Glossary of Common Terms*  xiii

*Foreword*  xv

*Acknowledgements*  xvii

## 1 Introduction

*Marc Bourgois*

Innovative Research Manager  
EUROCONTROL Experimental Centre  
Brétigny-sur-Orge, France

## 2 Complex Network Theory

*Dr Andrew Cook*

Principal Research Fellow  
Department of Planning and Transport  
University of Westminster  
London, UK

*Dr Massimiliano Zanin*

Principal Researcher  
The INNAXIS Foundation & Research Institute  
Madrid, Spain

## 3 Complex Networks in Air Transport

*Fabrizio Lillo*

Associate Professor  
Scuola Normale Superiore  
Pisa, Italy

*Rosario N. Mantegna*

Full Professor  
Dipartimento di Fisica e Chimica  
Università degli Studi di Palermo, Italy  
and  
Center for Network Science and Department of Economics  
Central European University  
Budapest, Hungary
Salvatore Miccichè  
Associate Professor  
Dipartimento di Fisica e Chimica  
Università degli Studi di Palermo, Italy

4  Uncertainty  
Damián Rivas  
Professor  
Department of Aerospace Engineering  
Universidad de Sevilla, Spain

Rafael Vazquez  
Associate Professor  
Department of Aerospace Engineering  
Universidad de Sevilla, Spain

5  Resilience  
Dr Henk A.P. Blom  
Full Professor  
Department of Aerospace Engineering  
Delft University of Technology, The Netherlands and  
Principal Scientist  
National Aerospace Laboratory NLR  
Amsterdam, The Netherlands

Soufiane Bouarfa  
Researcher  
Department of Aerospace Engineering  
Delft University of Technology, The Netherlands

6  Emergent Behaviour  
Dr Henk A.P. Blom  
Full Professor  
Department of Aerospace Engineering  
Delft University of Technology, The Netherlands and  
Principal Scientist  
National Aerospace Laboratory NLR  
Amsterdam, The Netherlands

Dr Mariken H.C. Everdij  
Senior Scientist  
National Aerospace Laboratory NLR  
Amsterdam, The Netherlands
Contents

Soufiane Bouarfa
Researcher
Department of Aerospace Engineering
Delft University of Technology, The Netherlands

7 Data Science
Dr Massimiliano Zanin
Principal Researcher
The INNAXIS Foundation & Research Institute
Madrid, Spain

Dr Andrew Cook
Principal Research Fellow
Department of Planning and Transport
University of Westminster
London, UK

Seddik Belkoura
Researcher
The INNAXIS Foundation & Research Institute
Madrid, Spain

8 Conclusions and a Look Ahead
David Pérez
Director
The INNAXIS Foundation & Research Institute
Madrid, Spain

References
Index
This page intentionally left blank
List of Figures

3.1 Log–log plot of the daily degree complementary cumulative distribution function of the airport networks for 6 May to 2 June 2010 (AIRAC 1005) 28
3.2 Log–log plot of the daily strength complementary cumulative distribution function of the airport networks for 6 May to 2 June 2010 (AIRAC 1005) 29
3.3 Log–log plot of the daily degree complementary cumulative distribution function of the networks of navigation points for 6 May to 2 June 2010 (AIRAC 1005) 30
3.4 Log–log plot of the daily strength complementary cumulative distribution function of the networks of navigation points for 6 May to 2 June 2010 (AIRAC 1005) 31
5.1 Resilience triangle, adapted from Tierney and Bruneau (2007), with disruption moment $t_d$, moment of full performance impact $t_f$ and moment of full recovery $t_r$ 65
5.2 Examples of network resilience curves, showing network performance percentage as a function of network damage percentage 66
5.3 Overview of the organisation of airline operational control, including the communication with air traffic control 76
6.1 Complementary simulation tools are required to evaluate weak emergent behaviour along the flank and at the top of the safety pyramid of Heinrich (1931) 90
6.2 Agent-based model 92
6.3 Agents in the trajectory-based operation and tactical conflict resolution layers in A3 98
6.4 Example of eight encountering aircraft trajectories realised under the A3 design 99
6.5 Estimated event probability per aircraft per flight hour for random traffic under A3 model control at traffic demand of 3x high en-route traffic demand in 2005 100
7.1 Metric classifications 108
7.2 The dynamics of causality and correlation 109
7.3 Reconstructing multi-layer networks of delay propagation 113
7.4 Topological properties of an ATM multi-layer network 114
List of Figures

7.5 Cumulative probability distributions for airports, airlines and aircraft types 121
7.6 Topological features as a function of the number of nodes (airports) 122
7.7 Fraport’s airport operational database key interfaces 126
# List of Tables

2.1 Examples of network properties across multiple domains  
2.2 Introducing network scales  
2.3 Selected complexity metrics  
5.1 Resilience capacities in relation to robustness and dependability  
5.2 Interdependency types in critical infrastructure systems  
5.3 Dimensions and their implications for resilience analysis of the air transportation system  
5.4 Ability in modelling and analysis of types of consequences due to disruptions  
5.5 Classification of activities by airline controllers  
5.6 The four airline operations control policies and their relation to the resilience capacities  
5.7 Simulation results for policies P1–P4 in the agent-based airline operations control model  
6.1 Top five ranking sub-models in capturing hazards and non-nominal effects  
6.2 Mathematical methods integrated with agent-based modelling and simulation  
7.1 Key techniques used in data mining  
7.2 Metrics for relationships between sets of data  
7.3 Granger causality axioms  
7.4 Delayed flights (actual)  
7.5 Delayed flights (expected)  
7.6 Delayed flights (differences)  
7.7 Delayed flights (sign of differences)  
7.8 Ratios of predicted and actual flights at navigation points  
7.9 The European network for two different sizes  
7.10 Big data domains in aviation
This page intentionally left blank
Glossary of Common Terms

4D Four-dimensional
ABMS Agent-based modelling and simulation
AOC Airline operations control
ASAS Airborne separation assistance system
ATC Air traffic control
ATFM Air traffic flow management
ATM Air traffic management
CNT Complex network theory
EUROCONTROL European Organisation for the Safety of Air Navigation
FAA Federal Aviation Administration (United States)
IATA International Air Transport Association
KPI Key performance indicator
SES Single European Sky
SESAR Single European Sky ATM research (programme)
TBO Trajectory-based operations
This page intentionally left blank
Foreword

The origins of this book can be traced back to the early activities of a group of researchers interested in applying complexity science to the field of air transport in general and air traffic management (ATM) in particular. This group was subsequently supported as a research network through the Single European Sky ATM research (SESAR) programme, as a result of being awarded funding by the SESAR Joint Undertaking. The European Organisation for the Safety of Air Navigation (EUROCONTROL), as the primary contact point for the work, has provided strong support to this network. We further describe the provenance of the book and make-up of the group in our opening chapter.

ATM comprises a socio-technical system that is too complex to be researched and improved through the use of classical approaches such as systems engineering and human factors alone. The application of complexity science paradigms to the analysis and modelling of future operations is driven by the need to accommodate long-term air traffic growth within an already saturated ATM infrastructure. The validity of this view continues to gain further ground. After collaborating on various works extolling the virtues, and identifying the challenges, of applying complexity science to ATM, we decided that there was a need for a collective work on this topic, drawing together its key themes and trying to be somewhat more definitive with some of its loose ends. We also wanted to draw on the early work in the field and offer the reader some examples of applications, putting together a book that was sufficient to whet the appetite of the reader and provide further reading (we ended up with over 350 references).

It has been an interesting process. There were many blurred lines that we have tried to resolve, and many definitions that we were forced to sharpen as we shaped the text. It is typically these boundaries in science, and the accompanying debates, which are the most interesting topics to read or write about. We have tried to keep the tone reasonably informal, but instructive, and the use of acronyms light: the last thing the industry needs is any additional acronyms. Our success in the latter is reflected in the very small glossary; our success in the former, the readers will judge for themselves.

Andrew Cook (London)
and Damián Rivas (Seville)
Editors
This page intentionally left blank
Acknowledgements

Our thanks are extended to the SESAR Joint Undertaking for funding the research network, supporting its activities and providing valuable opportunities to present its work to a large audience of stakeholders.

We would like to thank Marc Bourgois of the EUROCONTROL Experimental Centre, Paris, for his excellent support during the production of this book. For over ten years, Marc has been managing innovative research activities in ATM at EUROCONTROL. He is always ready with insightful comments and suggestions on numerous research projects and initiatives, many of which have been cited within the book, and directly regarding the book itself, not least in our introductory chapter. He has also shown great patience regarding some of the delays in getting the text ready, having faith that it was in a good cause as we attended to some big questions, and the dots and commas, as best we could.

We would also like to thank Dr Victoria Williams for going beyond the call of duty in her exemplary copy-editing, drawing on her knowledge of ATM and editorial skills alike to improve the text and raise numerous helpful questions, and to Helen Varley for her meticulous final proof-reading.

Thanks also are extended to Fraport AG (Frankfurt Airport Services Worldwide) for reviewing part of Chapter 7 and permitting the use of Figure 7.7.

The production of this book was co-financed by EUROCONTROL, acting on behalf of the SESAR Joint Undertaking and the European Union. The SESAR ComplexWorld Consortium are owners of the Foreground. The opinions expressed herein are entirely those of the Consortium and do not necessarily reflect those of the SESAR Joint Undertaking, European Union or EUROCONTROL, none of whom are liable for the use of any of the content.

In the following chapters all figures and tables not separately credited are the work of the contributors to this book.
This page intentionally left blank
Chapter 1
Introduction
Marc Bourgois

This book seeks to pull together research at the interface between complexity science and air traffic management (ATM) to provide a useful resource to facilitate and inspire future work in this area. The idea for this book originated with ComplexWorld, a network of excellence funded by the Single European Sky (SES) air transport policy initiative of the European Union, in collaboration with EUROCONTROL, the European Organisation for the Safety of Air Navigation.

This chapter introduces the field of complexity science as it relates to ATM, outlines the origins of the ComplexWorld network and the diverse expertise of the participants who contributed to the book, and sets out the topics that are addressed.

1.1 Complexity in Air Traffic Management

ATM combines human actors with technical subsystems and operational procedures in a complex system characterised by the myriad concurrent interactions between its many components. Complex systems cannot be fully understood by the traditional approach of studying their individual components and subsequently summing their characteristics. This is where complex systems science is required: complex systems science provides formal modelling paradigms and analysis techniques that help in uncovering fundamental properties and suggesting design principles to sustain high performance for large non-linear systems.

It is important to differentiate complex systems science from the term ‘traffic complexity’ used in traditional ATM research. Traffic complexity, often abbreviated to ‘complexity’, is the study of factors and metrics that quantify the relationship between traffic patterns, air traffic control (ATC) sectors and associated controller workload. Traffic complexity is not the subject of this book, which instead focuses on the application of complex systems analysis approaches developed in a range of disciplines to the study of the socio-technical air transport system.

Multi-disciplinary research in complex systems and ATM is in its early days, but publications have recently increased in number and promising lines of research can now be identified. The most straightforward studies concern the statistical properties of air transport networks, essentially in search of characteristic power-law distributions (the relevance of this is clarified in Chapter 2). The first review in this field clearly shows this to be a popular application of complex network theory (CNT), but it also shows that the values found for even the most basic
network metrics may differ widely, due to differing modelling choices and possibly dataset quality or completeness issues (Zanin and Lillo, 2013). Rather than leaving the modelling entirely to complex network theorists, a multi-disciplinary approach informing the models with operational and engineering expertise from the air traffic community might lead to more robust results.

More sophisticated modelling projects have uncovered interesting properties describing the coexistence of traditional airlines operating hub-and-spoke networks with low-cost airlines operating point-to-point networks. Results from the Single European Sky ATM Research (SESAR) project ELSA, for example, show that capacity can be optimally exploited with a balance of the two airline populations; airspace exclusively populated by either traditional or low-cost airlines would provide for less optimal trajectories, with more delay or rerouting (Gurtner et al., 2015). On the same subject, multi-layered network models can determine which of those two business models contributes more to individual properties of the overall network (Cardillo et al., 2013a). For example, low-cost airlines are shown to be the main contributor to path redundancy, which increases the flight options passengers have between a given origin and destination airport.

Another popular topic of study with complex systems scientists is the phase transition, indicating a sudden regime change, for example, between free-flow and congested states of a transport network. In road transport, this has led to the Nagel-Schreckenberg model, which identifies when jamming transitions emerge by modelling the microscopic behaviour of numerous car drivers on multi-lane roads. The jamming transitions are replicated with just a few simple driving rules and moderate statistical variability. These seminal models were implemented on efficient cellular automata simulators, which have also been exploited in air traffic, for example, to model ground congestion of the taxiway system at Tokyo International airport (Mori, 2012). The Nagel-Schreckenberg model is not directly portable to airborne traffic, as planes that slowed down too much would simply fall out of the sky; nonetheless an analogous problem called ‘traffic bunching’ exists in en-route and terminal air traffic sectors. A reduction in unsafe traffic bunching is considered to be one of the potential benefits of a major conceptual change to ATM under development: four-dimensional (4D) trajectory planning. Improved understanding of traffic bunching is, therefore, a potentially useful line of research (Chalon et. al., 2007).

The core chapters of this book discuss many more examples of relevant research results and opportunities, often in considerably more detail. The references at the end of the book provide numerous further pointers into the relevant literature for the interested reader. The extent of the references list shows that something substantial is clearly growing in this field of research.

1.2 The ComplexWorld Network

Over the last ten years, the European Union has launched a number of initiatives to increase the competitiveness of the European air transport sector and to prepare the sector for sustained growth in the face of increasing external pressures. One of
these initiatives is the SES, for which four high-level goals were identified: coping with a threefold increase in traffic; improving safety performance by a factor of ten; reducing the cost of ATM by a factor of two; and reducing the environmental impact of an individual flight by 10 per cent. Much of the burden of these goals lies with a small but pivotal branch of the air transport sector: ATM.

The organisation of ATM within Europe is unfortunately highly fragmented. Each member state has its own air navigation service provider (operating its ATC centres and airport control towers), its own safety regulator and military authorities claiming priority access or exclusive use of airspace. EUROCONTROL performs centralised network management and the European Union has its own safety regulator. Also significant are the many weather information providers, the aircraft and equipment manufacturers and the airlines operating in European airspace.

To achieve the high-level goals of the SES within such a complex institutional environment, a public–private partnership, the SESAR Joint Undertaking (SJU), was established to identify, coordinate and advance the applied research and development needed for the modernisation of European ATM. As part of the research programme, ComplexWorld was established as a network of excellence to explore the potential of complex systems science to increase the understanding of the performance of the European ATM system.

Until recently, complex systems science had hardly tackled the problems of ATM; key experts on complex systems science were to be found outside of the ATM domain, most prominently in statistical physics. The original membership of the ComplexWorld network was carefully constructed to bring together experienced researchers in complex systems science with the established ATM research community.

The authors of this book are core members of the ComplexWorld network. The academic side of the ATM research community is represented by the department of aerospace engineering from the University of Seville. Two members of this group contribute to the book: Damián Rivas and Rafael Vazquez. Dr Rivas is professor, teaching flight mechanics and air navigation, and is also the scientific coordinator of the ComplexWorld network. Dr Vazquez is associate professor and teaches orbital mechanics and navigation systems. Their ATM research focus is on conflict detection and resolution, trajectory prediction and optimisation from a control theoretic perspective, and uncertainty propagation.

In addition to academic institutions, the European air traffic research community includes several national research centres; these research centres are represented in the network by AT-ONE, a joint venture of the institute of flight guidance of the German Aerospace Centre (DLR) and the air transport safety institute of the Dutch National Aerospace Laboratory (NLR). Their contributions to this book stem from Henk Blom and his co-authors Dr Mariken Everdij and PhD candidate Soufiane Bouarfa. Dr Blom and Dr Everdij have been pioneers in using complexity science approaches to research the socio-technical ATM system. Dr Blom also holds a chair in Air Traffic Management Safety at Delft University of Technology and teaches agent-based safety risk analysis.
From the statistical physics-based complexity science community, a prominent group of researchers from the University of Palermo joined the network. This group is represented here by Rosario Mantegna, Salvatore Miccichè and Fabrizio Lillo. Dr Mantegna, who holds a combined professorship at the Central European University in economics and at the University of Palermo in physics, was among the first to analyse and model social and economic systems with the concepts and tools of statistical physics as far back as 1990 (Mantegna and Stanley, 2000). Dr Miccichè is an associate professor at the faculty of medicine and researches on the characterisation of long-range correlations in bioinformatics and econophysics. Dr Lillo recently moved to the Scuola Normale Superiore di Pisa where he is a professor in quantitative finance; he is also a member of the external faculty of the Santa Fe Institute.

Additional complexity science expertise is provided by Innaxis, a Spanish foundation specialising in complexity research with extensive project experience in the energy and transport sectors. Three staff members from the Innaxis Research Institute contribute to this book: David Perez, its director, Massimiliano Zanin and Seddik Belkoura. Before joining Innaxis David Perez had a long career in the aeronautical industry, having worked for Iberia Airlines and Boeing ATM in both engineering and business development roles. He is ideally placed to identify the main contributions of complex systems research so far and to sketch an outline of the challenges ahead for complex systems scientists dedicated to the air transport domain. Dr Massimiliano Zanin is a prolific researcher in data science. Seddik Belkoura is currently pursuing his PhD on complex systems and data science applied to ATM, under the supervision of Dr Zanin.

Drawing on 20 years’ experience in research on the operational aspects of ATM, Andrew Cook provides a most useful bridge as a transport economist with much experience in the wider air transport domain. Dr Cook is a Principal Researcher in the Department of Planning and Transport at the University of Westminster and has made seminal contributions to the issues of delays and their costs from the perspective of airlines, air navigation service providers and, crucially, passengers. Andrew Cook also took on the editing responsibilities for this book, jointly with Damián Rivas.

Over the last three years these authors and several of their colleagues have joined forces in organising thematic workshops, annual conferences and in guiding and tutoring a cohort of PhD students. They have been collaborating on joint research and compiling a state-of-the-art review of complex systems science approaches, results and open questions relevant to ATM. All of it is documented in an open, collaborative authoring platform at www.complexworld.eu. This book is based not only on the research work of the authors themselves but also on the larger body of reviewed complex systems research.

1.3 Topics Covered in this Book

This book can be seen as a compilation of selected topics in complex systems science applied to air transport and ATM. The selected topics are those for which traditional ATM research has not provided satisfactory solutions and for which
complex systems science techniques hold particular promise. This section provides an overview on which concepts and examples are covered by each topic and how these topics relate to each other.

The first topic is uncertainty, in Chapter 4. At the micro scale, Damián Rivas and Rafael Vazquez consider individual flights. One aspect of the uncertainty of individual flights is captured by delays: either in observed delay statistics (often separately for the different phases of flight) or by delay propagation models. Delay cost is crucially important; authoritative studies on delays and their costs are referenced across several chapters. Another aspect of the uncertainty of an individual flight is its trajectory uncertainty, that is, the uncertainty in knowing its position at any given time. Solutions are generally based on probabilistic frameworks that model the causes of trajectory uncertainty, which include the initial conditions, aircraft performance, navigational errors and wind conditions. A classification for the most relevant research contributions is outlined.

The mesoscale corresponds to flights close enough to interact, and to weather systems. This is the realm of tactical ATC and flow management. In addition to trajectory uncertainties, causes of traffic uncertainty include human actions (by controllers and pilots), management procedures and weather uncertainty. Whereas early studies on ground holding for flow management were reliant on deterministic trajectories, recent work exploits Markov models, queuing models, stochastic dynamic programming, etc. An alternative to probabilistic models is to calculate worst-case envelopes. This approach is occasionally followed in conflict resolution strategies or in robust optimisation techniques for adverse weather avoidance. A last example at the mesoscale concerns safety assessments based on rare-event Monte Carlo simulations, which is further explored in Chapter 6, which discusses emergent behaviour. Finally, the macroscale comprises the entire air transport network and although there is a growing body of complex systems studies at this level, few have actually dealt adequately with the issue of uncertainty.

A second topic is resilience. Chapter 5 investigates the added value of complexity science for studying the impact of various disruptions. In general, a socio-technical system may absorb a disruption and adapt to it, or may first change its performance and subsequently restore and adapt. These three aspects of resilience, that is, its absorptive, restorative and adaptive capacities, allow us to relate resilience to well-known systems engineering concepts such as dependability and robustness. A further distinction is that resilience aims to address socio-technical systems, rather than technical systems.

Resilience metrics are reviewed from a wide range of scientific disciplines in which resilience has been studied, largely independently. These include ecosystems, critical infrastructures and even psychology. Though the reviewed metrics provide initial guidance, the complexity induced by the combination of social and technical components in the air transport system and the difficulty of measuring the depth and time-evolution of the impact of disturbances leave the issue of deriving a resilience metric as an open challenge.

As complexity stems from the interaction between a large number of subsystems, modelling techniques from complexity science which focus on subsystems
interaction are better suited to capture the resilience phenomenon. Agent-based models, network flow-based models and the lesser known stochastic reachability analysis and viability analysis are obvious candidate modelling techniques. In the final section of this chapter, disturbances and how they can be dealt with by human operators are illustrated by several examples of disruption management policies in a typical airline operations centre. The authors show that an agent-based model building on state-of-the-art coordination theory leads to better solutions than the heuristics currently deployed at airline operations centres.

A final topic is emergent behaviour, which is explored in Chapter 6. In complexity science a property or behaviour of a system is called emergent if it is not a property or behaviour of the individual, constituting elements of the system, but results from the interactions between its constituting elements. The further study of this broad spectrum of emergence has occupied philosophers and scholars from complexity science. This has led to a constructive definition of three categories of emergence, based on whether it can be predicted through mathematical analysis, only through simulation, or cannot be predicted even through simulation. For each of these emergence categories illustrative ATM examples are given. Emergent behaviours that can be predicted are of use in the design of a future socio-technical ATM system. A natural fit for predicting emergent behaviour is again provided by agent-based modelling and simulation (ABMS), which is a well-established complexity science paradigm. However, the use of ABMS for the purpose of identifying the capacity and safety properties of a novel socio-technical ATM design poses challenges. Pilots and controllers play a key role in safely handling incidents and avoiding their escalation. Therefore accidents are very rare events, occurring once in every 10 million or even in every billion runs, posing challenges to ABMS. This can only be managed using advanced mathematical techniques, such as compositional modelling and importance sampling of stochastic hybrid systems. An extensive model analysis of a novel self-separation concept for autonomous aircraft, based on previous work from NASA, shows that surprising levels of safety can emerge under extreme high-density operations.

The complexity science approach to these three topics suggests design principles for a high-performance socio-technical ATM system. For the discussion on these topics to be accessible to the reader unfamiliar with complexity science, several modelling paradigms and analysis techniques need to be introduced. Application of these paradigms and techniques to ATM leads to insights in their own right, in addition to supporting the more applied discussions on the topics already introduced. One modelling paradigm central to complex socio-technical systems analysis is ABMS, which is introduced and illustrated in Chapter 6. Another modelling paradigm, CNT, is explored in Chapter 2. This is a technique which not only produces compact representations but also characterises the structure of a complex system across domains as varied as transport, utilities and biology. Andrew Cook and Massimiliano Zanin introduce basic metrics that measure properties of networks, their substructures and their dynamics (further metrics focusing on output, which equates to performance, are
explored in Chapter 7). Of importance in CNT is the fact that specific metrics tend to be meaningful at specific scales; the introduction of the three typical scales (micro, meso and macro) shows its usefulness in Chapter 4. Much publicised network studies of the Internet, the world wide web and social network applications led to the understanding that real-world networks are often characterised by the small-world and scale-free properties. This makes them crucially different from the theoretical networks, such as random graphs and regular graphs, which were the focus of study in the early days of complex systems science. These real-world properties, present in the hub-and-spoke networks of airlines, for example, make man-made networks vulnerable in specific ways, a topic which is further explored in Chapter 5.

Chapter 3, on complex networks in air transport, directly applies the concepts of CNT to the air transport network. As with all modelling languages, the modeller has a wide array of choices in constructing a network representation: will the vertices of the network be airports, air traffic sectors or navigation points? Will the edges represent flights, passengers or delays? Fabrizio Lillo, Salvatore Miccichè and Rosario Mantegna review many of the alternative choices that have been made in macroscale air transport network studies in the research literature. Numerous examples of power-law distributions which reflect the small-world properties of these networks are highlighted. Other studies have gone down to the mesoscale, focusing on short-term conflict alerts or community detection techniques to identify operationally significant clusters of air traffic sectors or navigation points. Recently multi-layered network representations addressing multi-modal transport or airline competition issues have pushed the state-of-the-art in complex network modelling. Studies of the dynamics of the airport network (that is, the network with airports as vertices) provide insights into passenger dynamics, propagation of delays, resilience and the spreading of epidemics. Although the latter issue might be considered peripheral to the core of air transport research, it is a rare but famous example of the value of air transport network models in other scientific domains.

The final technique from the toolbox of complex systems analysis presented in this book is data science, addressed in Chapter 7. Data science is all about mining data so that meaningful information can be extracted about a system. Whereas complex networks or multi-agent systems create (maximally) simplified models, data science lets the data speak directly. Andrew Cook and Massimiliano Zanin, with support from Seddik Belkoura, introduce and illustrate the concepts of synchronization likelihood and Granger causality which allow deductions about relationships in data which go beyond the correlation tests known from classical statistics. Such complexity metrics help to explain the emergence of properties at the macroscale of complex systems and are complementary to classical univariate and derived metrics. No discussion on data is complete without considering data quality and the issue of big data. This chapter provides telling examples from what is probably the most used European air traffic dataset and explores the growing use of big data and the associated challenges.
This brings us to the conclusion of this book. In Chapter 8, David Perez highlights the contributions to date and the areas open for contribution by complex systems science in each of the selected topics. Additionally, he stresses the importance of collaborative tools and initiatives to support a multi-disciplinary research network for the application of complex systems science in ATM, highlights data science as a novel area deserving its own separate network initiatives and signals the importance of continued support from the European institutions.
In this chapter we set out to introduce complexity science and complex network theory (CNT). We resist the urge to plummet immediately into ATM examples. We hope that by keeping the remit of this chapter open, and drawing on a range of examples beyond air transport, it will be of interest to the reader in setting the scene, and instructive by analogy. The following chapters focus the scope for us.

Complexity science is the multidisciplinary study of complex systems. Complex systems are those that display collective behaviour – behaviour that cannot be predicted through analyses or modelling of the individual components, but emerges instead from the interactions between such components. Many of the roots of complexity science can be traced back to statistical physics, non-linear dynamics and information theory (Anderson, 1972), although a consensual definition of complexity science still does not exist. Nevertheless, complex systems invariably involve networks. Generic topological properties are present in different complex systems, suggesting that some general principles govern the creation, growth and evolution of such networks.

Central to any exploration of a complex system is CNT, to which this chapter is dedicated. It is so fundamental to the other themes of the book that it is one of the acronyms we have allowed ourselves to use: CNT. All complex systems have interconnected components, such that CNT plays a central role in complexity science (Newman, 2003; Boccaletti et al., 2006). Our discussion here is aimed at the reader who is not familiar with the theory and our objective is to cover material necessary to support the concepts presented in the following chapters, with a little extra besides. We begin with a discussion of the origins of CNT, before moving on to explore how it is used to characterise networks. This centres on different ways of describing the structural properties of networks, that is, their topologies. We also discuss the importance of scales and their associated metrics.

We define a metric as any quantitative measure, particularly one which usefully expresses some output of a system (usually performance), part of the system, or (an) agent(s) within it, usually over an aggregate scale and often as a ratio (e.g., per flight). Returning briefly to the ATM and Single European Sky (SES) context, introduced in Chapter 1, we note that here, a performance indicator (a specialised type of metric) is used for the purpose of ‘performance monitoring, benchmarking and reviewing’, whereas a key performance indicator (KPI) is for performance target setting (EUROCONTROL, 2012), and thus an even more specialised metric. As stated in Cook et al., 2013a, KPIs need to be chosen that are intelligible
(preferably to the point of being simple), sensitive (in that they accurately reflect
the aspect of performance being measured) and consistent (we cannot refine them
from one period to another without losing comparability). Trade-offs between these
desirable properties often have to be made, but even this is just the beginning of
developing the big picture. Building on these definitions, we define different classes
of complex networks before taking a look at some applications.

2.1 Representing the World with Networks

At the centre of this story we have networks and graph theory. Graph theory first
appeared in 1735 (Biggs et al., 1986) when the Swiss physicist and mathematician,
Leonhard Euler, solved the Seven Bridges of Königsberg problem: whether a path
existed across seven bridges and two islands in the city, crossing each bridge only
once. (Although there is no such possible route we can nevertheless recommend
a visit to the city, now called Kaliningrad.) In the context of graph theory, graphs
are defined as mathematical structures that show the relationships between ele-
ments. These elements are referred to as ‘nodes’ (or ‘vertices’), and the lines
between them ‘edges’ (or ‘links’). As illustrated in Table 2.1, various entities may
‘flow’, or be transported, through the network. These entities are usually of some
intrinsic value to the network, and may thus generically be considered as assets.
(We discuss negative examples later, whereby unwanted phenomena flow through
the network.)

When we are concerned about the direction of flows from one node to another,
we have a ‘directed’ graph. When we are not concerned with such distinctions,
including, for example, when we do not distinguish between the nodes, we have
an ‘undirected’ graph. These graphs can also be weighted, that is, associated with
real or integer values. For example, in an airport network each node is an airport
and a link directed from one node to another can be weighted by the number of
flights or passengers in a given time window. Indeed, such is the power of CNT
that one can assign almost any kind of nodal representation, across multiple
domains. This allows the creation of elegant representations of the governing rela-
tionships, avoiding unnecessary details about the nature of the elements them-
selves. In other words, by means of a network it is possible to represent in an
abstract form the underlying structure of a system, independently of the nature of
the system itself. We return to such characterisations in Section 2.2.

Table 2.1 shows selected examples of key network characteristics across several
domains. The flows identified are all driven by some form of energy. This is typi-
cally counted in monetary terms within the transport context, although even here
also it could be expressed as a fuel burn energy, inter alia. These flows may be
disrupted by breakage or loss of capacity, and work against metaphorical and/or
literal forms of resistance (friction).

Real-world networks are often co-dependent, such as laying water pipelines
under roads, water distribution networks being powered by electrical pumps and
inter-modal transport exchanges. More rarely, a vital edge in one network (such as
Complex Network Theory

A main road) could be the disruption event for an edge in another network (e.g., prohibiting safe species dispersal). Unlike other (biological) transport networks, the network formed by fungi is not part of the organism – rather, it is the organism.

A number of these networks also share common functional themes. Capacity is expressed through various metrics, such as pipe diameters, cable bandwidths or seating configurations for vehicles. Telecommunications terminologies for hub-and-spoke networks such as (packet) scheduling, service denials, backbones, routing protocols (with distance restrictions), traffic delivery rates, traffic forecasts, and (node) diversions have obvious analogues with air transport. We often talk of ‘downstream’ propagation effects, where the terminology is literal, in the context of water distribution, and metaphorical in other contexts.

There is an implicit trade-off that pervades transport systems, which is particularly closely echoed in telecommunications: hub-and-spoke networks are

<table>
<thead>
<tr>
<th>Network</th>
<th>Node</th>
<th>Edge</th>
<th>Flow</th>
<th>Disruption</th>
<th>Flow cost</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generic</strong></td>
<td>Collection</td>
<td>Transport</td>
<td>Asset</td>
<td>Loss of capacity</td>
<td>Energy</td>
</tr>
<tr>
<td><strong>Transport</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Air – flight-centric</td>
<td>Airport</td>
<td>Flight</td>
<td>Aircraft</td>
<td>Mechanical failure</td>
<td>Monetary</td>
</tr>
<tr>
<td>Air – pax-centric</td>
<td>Airport</td>
<td>Flight(s)</td>
<td>Passengers</td>
<td>Missed connection</td>
<td>Monetary</td>
</tr>
<tr>
<td>Urban (road)</td>
<td>Junction</td>
<td>Road segment</td>
<td>Vehicles</td>
<td>Bridge collapse</td>
<td>Monetary</td>
</tr>
<tr>
<td>Rail</td>
<td>Station</td>
<td>Track segment</td>
<td>Trains</td>
<td>Signal failure</td>
<td>Monetary</td>
</tr>
<tr>
<td>Goods</td>
<td>Warehouse</td>
<td>Road segment</td>
<td>Goods</td>
<td>Traffic congestion</td>
<td>Monetary</td>
</tr>
<tr>
<td><strong>Services/utilities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>Plant, reservoir (Sub)</td>
<td>Pipe Cables</td>
<td>Water</td>
<td>Pipe breakage</td>
<td>Energy</td>
</tr>
<tr>
<td>Electricity</td>
<td>Station</td>
<td>Cables</td>
<td>Electrons</td>
<td>Cable breakage</td>
<td>Energy</td>
</tr>
<tr>
<td>Telecoms</td>
<td>Hub, router</td>
<td>Wire/fibre</td>
<td>Data packets: electrons/photons</td>
<td>Cable breakage</td>
<td>Energy</td>
</tr>
<tr>
<td><strong>Biology/ecology</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mammalian brain</td>
<td>Distinct grey matter regions</td>
<td>White matter fibre bundles</td>
<td>Electrical impulses; Neurotransmitters</td>
<td>Breakage (e.g. disease)</td>
<td>Energy</td>
</tr>
<tr>
<td>Fungal ecology</td>
<td>Branch point, fusion, tip</td>
<td>Cord (e.g. packed with hyphae)</td>
<td>Aqueous nutrients</td>
<td>Breakage (e.g. grazing)</td>
<td>Energy</td>
</tr>
<tr>
<td>Animal ecology</td>
<td>Habitat patch</td>
<td>Landscape segment</td>
<td>Species dispersal</td>
<td>Road segment</td>
<td>Energy</td>
</tr>
</tbody>
</table>
especially efficient from an economic and design perspective but they are also particularly susceptible to short-term system failure. Failures in air transport, for example, typically manifest themselves over a period of hours and often with a timescale of one operational day. Failure types in other networks may well occur over relatively large timescales, such as years or decades. Corresponding network models consider normal ageing, which results in subsequent disruption, such as the natural wear and tear on a road, bridge, pipe or cable.

Rerouting is another common theme across many types of network. Sometimes this is (practically) instantaneous, for example, in the water distribution and telecommunications contexts. In the latter, data are insensitive to the routing (unlike passengers), as long as they are distributed within corresponding time and integrity constraints. Where possible during disruption, data packets are rather deviated over the remaining intact routing topology – rather than incurring ‘packet drops’ (somewhat analogous to flight cancellations). In the urban context, the user of the network (such as a car driver) may be able not only to autonomously adapt their route but also may even change their mode and/or choice of destination (for example, going to a different town to make purchases). While changes of route are possible in air transport, changing mode or destination is much less common.

Earlier, we touched upon the issue of disruption. While we described the movement of assets through a network, there are certain circumstances whereby negative phenomena propagate through networks and cause disruption. An example is the propagation of delay through an air transport network, where one flight causes the delay of many others, due to multiple dependencies such as connecting passengers. (Compare the failure of one bridge in a network of bridges, for example: this would not be expected to cause other bridge failures.) Such effects may be considered as the propagation of the disruption of normal functionalities, rather than the movement of some negative phenomenon. Even in the case of a tumour, a physical entity in its own right, it is usually the compromise of normal functioning tissue or neurons which manifests itself as disease, rather than the presence of the tumour tissue itself. Network topology and functionality are critical to model failure cascades, and air transport networks in particular demonstrate several special vulnerabilities. We return to this theme in Section 2.4, and in more depth in Chapter 5, which is dedicated to discussing resilience and its associated metrics.

2.2 Characterising Networks

Table 2.2 introduces the three scales that are commonly used when describing complex systems. (The mammalian brain is the most complex structure known in the universe. It is as compelling to use it as an illustrative example, as it is awe-inspiring that each reader is deploying one to read this text.) The first of these scales, the microscale, focuses on the properties of a single link or single node. At this scale, one may study how a single neuron is connected with others, the nature of these connections (chemical, electrical, etc.) and, indeed, try to model the
neuron’s internal dynamics. One step further, one can focus on the mesoscale, as an intermediate level between studying elements of a system separately and considering the system as a whole. We may define the mesoscale as any regular structure that affects the connectivity of groups of nodes in the network (Almendral et al., 2011). This may include communities, i.e., sets of nodes that are more connected among themselves than with the rest of the network, that are essential to the understanding of the structures and operations of real networks (Fortunato, 2010). Also important at the mesoscale are motifs and core-periphery structures. Motifs are patterns (more formally, recurring sub-graphs) within a network with a frequency higher than expected in random ensembles (Milo et al., 2002). Core-periphery structures arise from the combination of a densely connected inner core and a set of peripheral nodes sparsely connected with the core (Holme, 2005).

The reader may readily recognise how the concept of the mesoscale is more elusive, and less commonly studied, than the microscale or macroscale, and more difficult to assess in a quantitative way (Zanin et al., 2014). In the brain, the mesoscale is composed of groups of neurons, cooperatively processing a single stimulus. This may include groups of neurons strongly connected together, thus giving birth to a connectivity mesoscale – large groups working together during some complex tasks, i.e., a functional mesoscale. Alternatively, the mesoscale may comprise small groups of neurons collaborating to process a single stimulus, as in the case of the columns of the neocortex (Kandel et al., 2000).

Finally, the third scale is that of the system as a whole: the macroscale. Its metrics account for the overall structure of the network, addressing the movement of information through the network. For example, the number of jumps or moves needed to move from one side of the network to the other, or the importance of a node with respect to the movement of information within the network.

Table 2.3 illustrates a selection of metrics used in CNT, each of which can be set in the context of a corresponding network scale. We mentioned above that one of the powers of CNT is the ability to define nodes according to the needs of a particular analysis or representation. The interpretation of the metrics also thus depends on the context in which they are used, i.e., how the nodes are defined. The simplest example of a microscale metric is the degree. This is defined as the number of connections arriving at, or departing from, a node. While this measure reveals some information about the structure of the network (for example, it is possible to define the most central node as the one with most connections), more knowledge

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mammalian brain example</th>
<th>Air transport example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microscale</td>
<td>A single neuron</td>
<td>A single flight</td>
</tr>
<tr>
<td>Mesoscale</td>
<td>A community of neurons, cooperatively processing a single stimulus</td>
<td>The community of flights of a single airline</td>
</tr>
<tr>
<td>Macroscale</td>
<td>The brain</td>
<td>The air transport network</td>
</tr>
<tr>
<td>Metric</td>
<td>Classification</td>
<td>Basis</td>
</tr>
<tr>
<td>--------</td>
<td>----------------</td>
<td>-------</td>
</tr>
<tr>
<td>Alpha centrality</td>
<td>Microscale</td>
<td>Node</td>
</tr>
<tr>
<td>Betweenness (centrality)</td>
<td>Microscale</td>
<td>Node</td>
</tr>
<tr>
<td>Centrality</td>
<td>Microscale</td>
<td>Node</td>
</tr>
<tr>
<td>Clustering coefficient</td>
<td>Mesoscale</td>
<td>Triplets of nodes</td>
</tr>
<tr>
<td>Degree (centrality)</td>
<td>Microscale</td>
<td>Node</td>
</tr>
<tr>
<td>Degree correlation</td>
<td>Macroscale</td>
<td>Network</td>
</tr>
<tr>
<td>Eigenvector centrality</td>
<td>Microscale</td>
<td>Node</td>
</tr>
<tr>
<td>Entropy of the degree distribution</td>
<td>Macroscale</td>
<td>Network</td>
</tr>
<tr>
<td>Geodesic distance</td>
<td>Macroscale Network</td>
<td>The average number of steps needed to move between two nodes of the network.</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Global efficiency</td>
<td>Macroscale Network</td>
<td>The ease of information flow between pairs of nodes; the (generic) cost of this communication can be approximated by the distance (length) of the shortest path connecting two nodes – the normalised global efficiency is defined as the mean value of the inverse of such distances: $E_g = 1/{n(n-1)} \sum_{i,j=1} 1/d_{i,j}$.</td>
</tr>
<tr>
<td>Information content</td>
<td>Mesoscale Network</td>
<td>Assesses the presence of any mesoscale structure, by evaluating the information lost when pairs of nodes are iteratively merged together (Zanin et al., 2014).</td>
</tr>
<tr>
<td>Link density</td>
<td>Microscale Network</td>
<td>The number of links in the network, $l$, divided by the maximum number of links that could be present; for a network composed of $n$ nodes, the link density is thus $l/n(n-1)$. See Boccaletti et al., 2006; Costa et al., 2007.</td>
</tr>
<tr>
<td>Modularity</td>
<td>Mesoscale Network</td>
<td>Assesses the presence of a community structure in the network (Girvan and Newman, 2002).</td>
</tr>
<tr>
<td>Rich club coefficient</td>
<td>Microscale Node</td>
<td>For each node of degree $k$, this is the ratio $E_i / N_i$, where $E_i$ is the number of observed links and $N_i$ is the number of potential links, both numbers computed by considering only links with other nodes having degree larger than $k$ (see Section 2.4 and Chapter 3).</td>
</tr>
</tbody>
</table>

can be extracted from the aggregation of all degrees, for example, through a degree distribution, $P(k)$, which expresses the fraction of nodes in the network with degree $k$ (we return to this below).

If nodes are defined to represent some parameterisation of delay, for example, and we had a few nodes with a very high degree, this would suggest that those nodes were primarily responsible for the propagation of delay in the network. On the other hand, if all nodes had more or less the same degree, no delay multiplier node is suggested. Betweenness is a metric measuring how important a given node is for the movement of information in a network. It represents how information (e.g., delay phenomena) moves through the whole network, and is thus a macroscale metric, as is global efficiency. For the latter, distances may be defined as required (for example, Great Circle distances could be used for flights), although they are often treated as a topological distance, i.e., based on the number of links
needed to ‘move’ from one node to another. Degree, on the other hand, is a purely
local metric – taking into account only the structure of the network around the node.
Both degree and link density are microscale metrics.

The eigenvector centrality is a metric defined such that this centrality of a node
is proportional to the centralities of those to which it is connected. This measure was
first created for the analysis of social networks, where the ‘importance’ (influence)
of a person is proportional to the importance of his or her friends (for example, one
person may be very important even if he/she has only one friend, if that friend is the
US President). One method to compute the eigenvector centrality is to assign each
node an initial random (positive) amount of influence. Every node then distributes
this influence evenly among its outward neighbours, receiving likewise inward
influence. This process is iterated until each node is sending out as much influence
as it is receiving – whereby the system has reached a steady state. The amount of
influence assigned to each node under this steady state is its eigenvector centrality.
Alpha centrality is a generalisation of eigenvector centrality, whereby nodes are also
allowed to have external sources of influence. The external influences (exogenous
factors) are often assigned as a constant for every node, or as a given vector for each
node. An iterative procedure may then be used (as per the eigenvector centrality
method) to reach a steady state. In this process, a constant (alpha) determines the
trade-off between the importance of the exogenous factors against the endogenous
factors – that is, those internal to the network. When alpha is very large, only the
external influences are of consequence. When alpha is zero, only the internal con-
nectivity matters, and we effectively return to the eigenvector case. (Sometimes
alpha is defined conversely, and these definitions are then reversed.)

2.3 Classes of Complex Networks

One of the main reasons behind the success of CNT in the last decade has been its
capacity to describe characteristics of networks that are shared by a vast number
of real systems. Specifically, the backbone of connections of most real-world sys-
tems are characterised by small-world and scale-free properties. Before defining
these, however, it is necessary to go back in time, and describe the first models that
were used to study networks.

While introduced in 1735 by Euler, as discussed in Section 2.1, the person
mainly responsible for the theoretical advancements in network (then ‘graph’)
theory was Paul Erdős. Among his numerous contributions, probably the most
important was the introduction of the concept of random graphs, also known as
Erdős-Rényi graphs. Given a set of \( n \) disconnected vertices, links between all pairs
of nodes are created with a probability \( p \); therefore no coherent structure is expected
to appear, but instead, random graphs are in fact homogeneous distributions of
random connections.\(^1\)

If random graphs are characterised by a complete absence of structure, the other
extreme is represented by regular graphs, that is, networks where all nodes have
the same number of connections. Both extremes are of limited applicability in describing real-world networks, as usually natural and man-made systems present a trade-off between regularity and more complex structures. This lack of realism motivated the creation of richer models, the two most important being small-world and scale-free networks. The former appeared in 1998, when Watts and Strogatz realised that real networks were neither regular nor completely random graphs, but rather that they lay somewhere between these two extremes. Specifically, random graphs are characterised by a low mean geodesic distance and by a low clustering coefficient; on the other hand, regular graphs show high mean geodesic distances and high clustering coefficients. By analysing social and biological networks, they discovered that most of them were characterised by a low mean geodesic distance, but also by a high clustering coefficient. In order to explain how such a combination can emerge in real systems, they proposed a hybrid model: starting from a regular graph, few links are randomly deleted and replaced by random (and, therefore, not regular) long-range connections. By tuning the number of links affected by such ‘rewiring’, it was possible to create a large family of networks, all of them maintaining the high clustering of the regular initial graph, but also showing a reduced distance between nodes. The two combined properties are now widely known as the ‘small-world’ effect (Watts and Strogatz, 1998).

Finally, a fourth class of network topologies emerged when another fact about real-world networks was observed. Instead of having homogeneous nodes, that is, nodes with approximately the same number of connections, real-world networks were characterised by some highly important nodes, usually called ‘hubs’. This is especially evident in air transport networks, for example, in which a few airports connect with much of the network, for example, London, Paris, and Frankfurt, and a large number of smaller airports only connect to these hubs (Zanin and Lillo, 2013). Mathematically, these networks are called ‘scale free’, as their degree distribution $P(k)$ follows a power law (i.e., the fraction $P(k)$ of nodes in the network with $k$ connections to other nodes, follows the power law $P(k) \sim k^{-\gamma}$ for large values of $k$, where $\gamma$ is typically between 2 and 3) and thus has no characteristic scale. (We develop this further in the cases discussed in Chapter 3.) Once this characteristic was brought to light, it was necessary to create a model for describing how it can emerge in natural systems. Such a problem was solved in 1999 by Barabási and Albert, by leveraging on the concept of ‘preferential attachment’ (Barabási and Albert, 1999). The process begins with an initial (seed) network of $m$ nodes, $m$ being usually small. New nodes are then added to the network, one at a time. Each of these new nodes is connected to existing nodes with a probability that is proportional to the degree of the latter, as expressed by:

$$p_i = \frac{k_i}{\sum_{j \in G} k_j},$$

where $p_i$ is the probability of connecting with node $i$. Due to this biased attachment mechanism, highly connected nodes rapidly gain more links, thus becoming hubs
of the system: the rich get richer. We pick up this discussion in the next section, where we see how such network classifications have implications for vulnerabilities, for example, to attack.

2.4 The Internet and Initial Insights into Vulnerability

The Internet may be readily represented as a graph, with hosts, routers and servers the nodes and the physical links connecting them (typically optical fibres or copper cables) the edges. CNT researchers have thus employed a graph-based approach to model the Internet structure and simulate its traffic. Complete mapping of the Internet is difficult to achieve, however, because it is not centrally administered and changes constantly. Researchers thus usually employ coarse-grained maps that contain only the links between autonomous systems, which are sub-networks, administered separately, or use maps of the connections only at the router level. Nevertheless, such incomplete graphs have furnished important findings regarding the structure and dynamics of the Internet, and their study provides clear examples of the kinds of results produced by complex network analyses. Let us look at some of the evidence.

As a first approach, researchers focused on investigating the global structure of the Internet: specifically, on its power-law nature, following the pioneering work of Faloutsos et al. (1999). Three snapshots of the Internet at the autonomous systems level (collected between 1997 and 1998), and one instance at the router level (from 1995), were examined. Power laws were found for the degree distribution, the degree rank, and the number of pairs of nodes within $h$ edges (called ‘hops’). Although criticisms were raised that these maps were constructed in such a way as to represent a particular (i.e., not general) aspect of the network, subsequent studies (e.g., Siganos et al., 2003) confirmed the power-law nature for more snapshots of the Internet (ranging from 1997 to 2002). Furthermore, Maslov et al. (2004) identified differences in networks with the same degree distribution, such as the Internet at the autonomous systems level (in 2000) and molecular networks. A power-law degree correlation function was found in another map of the Internet, which was probably a consequence of a hierarchical structure (Ravasz and Barabási, 2003; Corominas-Murtra et al., 2013) whereby high-level controller nodes were interconnected, thus creating an inverse tree structure.

Eriksen et al. (2003) found modules that approximately matched single countries, while Rosato et al. (2008) observed that the Internet has large and highly clustered regions joined by a few links. The ‘rich club’ phenomenon was also identified in an autonomous systems network, i.e., where the highest degree nodes were well connected with each other, by Zhou and Mondragón (2004). As a final example, the modularity of the Internet was investigated using spectral analysis of autonomous systems-level networks, i.e., by detecting groups of highly connected nodes through the analysis of the eigenvectors of the adjacency matrix (Fortunato, 2010).

So what do all these findings tell us? They suggest that the Internet has a scale-free structure, in which large hubs control large groups of small nodes, organised
both according to a hierarchical and spatially constrained structure. Far from being a purely theoretical classification, understanding the class to which a real-world network belongs furnishes valuable insights with regard to optimisation and vulnerability. For example, it is well established that scale-free networks are highly resilient to random failures, as most of the nodes have a low number of connections, and therefore their removal would not dramatically affect the whole network dynamics. Nevertheless, this does come at a high price, since they are particularly vulnerable to targeted attacks (Albert et al., 2000). Specifically, a smart attack can target an important hub, disrupting the behaviour of the dependent nodes, thus affecting the functioning of the Internet in a particular geographical region. Notably, small-world and random networks behave in opposite ways. Consider a random network. In this case, an attacker could not identify a highly important node, and therefore these types of network are resilient to targeted attacks. On the other hand, since no nodes are really secondary, such an attack may be expected to cause significant damage. Deploying CNT towards an improved understanding and characterisation of such network topologies, and the identification of, for example, the underlying network class and most central nodes, often affords powerful insights into identifying critical elements of such systems.

Indeed, as may be expected, the initial analyses of the topology of the Internet paved the way to the study of its vulnerability to random failure or attacks. The vulnerability of complex networks, including the Internet at the autonomous systems level, was studied by Albert et al. (2000) with regard to random node removal (simulating node failure) or hub removal (simulating a network under attack). Results showed the scale-free networks (Barabási–Albert) model to be extremely efficient against random failures, unlike the Erdös–Rényi model. However, hub-based attacks rapidly break scale-free networks into small, isolated groups. In Holme et al. (2002), attacks were performed on complex networks (including the Internet at the autonomous systems level) removing nodes with high degree or high betweenness centrality, both calculated only at the beginning of the attack (i.e., in the original network) or at each step of attack. Similar strategies were employed for edge removal. At each node or edge removal, the average inverse geodesic length and the size of the largest connected sub-graph were obtained. It was shown that the recalculated betweenness was the most harmful to the Internet, considering edge-based attacks. For node-based attacks, the different strategies were equally harmful. Another interesting observation was that these important nodes did not necessarily have high degrees. We discuss resilience and vulnerability further in Chapter 3 and Chapter 5, in the air transport context.

2.5 Air Transport Performance and Cost Assessment

In Table 2.1 we illustrate that disruption in transport networks is typically measured in terms of costs, so we should not conclude this chapter without some reference to these costs in the air transport context. From the airline perspective, such costs
Dr Andrew Cook and Dr Massimiliano Zanin are typically measured in terms of the cost of delay. They are mostly incurred at the arrival and post-arrival phases of flight (see Section 4.2.1 for a discussion of uncertainty by phase of flight). Such effects (and associated costs) often propagate through the rest of the network as so-called ‘reactionary’ effects – as flights wait for connecting passengers and crews, for example, causing cascade effects through the network. It is difficult to overstate the importance of delay costs. In fact, in 1999, the European Commission launched the SES initiative (a paradigm shift in the design and function of European airspace – see Chapter 1), specifically in response to increasing delays.

There are important trade-offs between strategic costs (e.g., adding buffer to airline schedules) and tactical delay costs (e.g., delays on the day of operation). The cost of delay can thus be used as a basis for strategic decision-making, often in the context of airline scheduling, on which there is a wide literature. A very good introduction with extensive references is to be found in Wu, 2010. For an analysis of the strategic–tactical cost trade-offs in the context of calculating how much buffer to add to a schedule, see Cook et al., 2010. Dunbar et al. (2012) also study the airline scheduling problem, introducing a new approach to calculating and minimising the cost of propagated delay in a framework that integrates aircraft routing and crew pairing.

Focusing on tactical delay costs in Europe, in 2013, the single largest delay cause was indeed reactionary delay (45 per cent), the second largest being airline causes (31 per cent), and with air traffic flow management (ATFM) delays comprising around 11 per cent (airline-reported delays, CODA, 2014). Even though ATFM delay only comprises just over one-tenth of these delays, it is estimated at a cost of approximately EUR 660 million in 2013 (measured in euros at 2009 values (EUROCONTROL, 2014)). The derivation of these costs of delay to European airlines, taking into account passenger, crew, fuel, maintenance and fleet (strategic) costs are to be found in Cook and Tanner, 2011 (with an update to this work pending as we go to press).

In 2013, 36.1 per cent of flights were delayed by five minutes or more on departure, which was an increase of 0.6 percentage points in comparison to 2012, despite the lower traffic of 2013 (CODA, 2014). En-route ATFM delays decreased for the third consecutive year in 2013 (to 0.53 minutes per flight) – this was the lowest level recorded (EUROCONTROL, 2014). While industry reporting has historically been very much focused on reporting flight-centric metrics, it has become increasingly apparent that the average delays of (delayed) flights and passengers are not the same and that we are lacking passenger-centric metrics. Trade-offs between these metrics also need to be better understood, as they are observed to move in opposite directions for certain types of flight prioritisation. With growing political emphasis in Europe on service delivery to the passenger, and passenger mobility, how are we to measure the effectiveness of passenger-driven performance initiatives in air transport if we do not have the corresponding set of passenger-oriented metrics and understand the associated trade-offs in the context of delay propagation?
In some of the first work in this area, using large datasets for passenger bookings and flight operations from a major US airline, Bratu and Barnhart (2004) showed that passenger-centric metrics are superior to flight-based metrics for assessing passenger delays, primarily because the latter do not take account of replanned itineraries of passengers disrupted due to flight-leg cancellations and missed connections. For August 2000, the average passenger delay (across all passengers) was estimated as 25.6 minutes, i.e., 1.7 times greater than the average flight leg delay of 15.4 minutes. Based on a model using 2005 US data for flights between the 35 busiest airports, Sherry et al. (2008) concurred that “flight delay data is a poor proxy for measuring passenger trip delays”. For passengers (on single-segment routes) and flights, delayed alike by more than 15 minutes, the ratio of the separate delay metrics was estimated at 1.6. Furthermore, heavily skewed distributions of passenger trip delay demonstrated that a small proportion of passengers experienced heavy delays, which was not apparent from flight-based performance metrics (Wang, 2007; Calderón-Meza et al., 2008). Using US historical flight segment data from 2000 to 2006 to build a passenger flow simulation model to predict passenger trip times, Wang (2007) cites flight delay, load factors, cancellation (time), airline cooperation policy and flight times as the most significant factors affecting total passenger trip delay in the system. Analysing US flight data for 2007 between 309 airports to estimate passenger-centric delay metrics showed (Calderón-Meza et al., 2008) that the average trip delay for passengers over all flights was 24 minutes, while for passengers on flights delayed by at least 15 minutes, the average delay was 56 minutes.

The SESAR project POEM (Cook et al., 2013b) built the first full European network simulation model with explicit passenger itineraries and delay cost estimations (based on September 2010 data). New flight and passenger prioritisation scenarios were used to explore the trade-offs between flight-centric and passenger-centric metrics, with ratios between the two agreeing well with the US values reported above. When the network operated under flight prioritisation and aircraft wait rules that minimised airline costs, win–win outcomes were observed, saving on average a sizeable EUR 40 per flight. This project also set out to better characterise the propagation of delay through the network using complexity science methods. We pick up on this in Section 3.3, where we revisit CNT communities.

2.6 Conclusions

We have introduced the main facets of CNT: how networks arise in natural and man-made systems, how they can be characterised, and insights into how an understanding of network structure could be used to improve operational performance and resilience. To pave the way further for themes picked up in the following chapters, we conclude by presenting two extensions of CNT, which have attracted a lot of attention in recent years: temporal and multi-layer networks.
Temporal networks, are composed of edges that are not continuously active. As an example, let us consider the air transport network. In a static representation of the topology, if two airports are connected by a link, it means that a passenger can always move between them. Yet, in reality, edges are active for non-negligible periods of time: that is, flights operate only at discrete times. Clearly, the temporal structure of edge activations strongly affects the resulting dynamics, as passengers cannot connect between different flights in an arbitrary way, which also has modelling implications (Cook et al., 2013b; Zanin et al., 2009). The interested reader may refer to two useful reviews (Holme and Saramäki, 2012 and 2013).

Besides this temporal aspect, it should also be noted that the traditional CNT approach has mostly been limited to the representation of node–node interactions by means of a (generally, real) number: quantifying the weight of the corresponding graph’s connection (or link). Nevertheless, considering all links as instances of a single object can be an important constraint. It may occasionally result in not fully capturing the details present in some real-life problems, even leading to incorrect descriptions of the corresponding phenomena. A paradigmatic example of intrinsically multi-relational systems can be found in the air transport network. The traditional study of this infrastructure is based on representing it as a single-layer network, where nodes represent airports and links are direct flights between them. Yet it is clear that a more accurate mapping can be realised if airlines as a whole are considered, as passengers cannot easily connect between two flights operated by different airlines, at least by airlines belonging to different alliances (Cardillo et al., 2013b).

The realisation of the important limitation imposed by considering a single type of link has resulted in the development of a novel framework for the study of multi-layer networks, that is, graphs where several different layers of connections are taken into account (Boccaletti et al., 2014). Multi-layer networks explicitly incorporate multiple channels of connectivity and constitute the natural environment for describing systems interconnected through different categories of connections: each channel (relationship, activity, category) is represented by a layer, and the same node may have different kinds of interactions (different sets of neighbours in each layer). This allows pairs of airports to be connected by means of various types of link, representing different airlines, different aircraft types, different forms of delay propagation, and so forth. We explore multi-layer networks further in Chapter 7. Looking ahead further still for future applications of CNT: this is a discussion we save for the end of the book, in Chapter 8. We next focus specifically on complex networks in air transport, in Chapter 3.

Note

1 Many theoretical results were demonstrated through random graphs, as, for example, the expected size of the largest component (groups of connected nodes), or the critical value of $p$ for which the graph was connected (Erdős and Rényi, 1959). A comprehensive review of results obtained through random graph analysis can be found in Bollobás, 2001.
There has recently been a tremendous growth in the study and application of complex networks, as described in Chapter 2. Several studies have applied methods from complex network theory (CNT) to the air traffic system (for a review see Zanin and Lillo, 2013). As we detail in the next section, the air traffic system can naturally be seen as the superposition of different complex networks, where nodes can be airports, sectors, navigation points, etc.

Complex networks could improve both our knowledge of the air transport system and our ability to manage and control it. From an operational point of view, the methods and concepts from CNT have many potential applications. For instance, the description of the topological and metric structure of the network may help in understanding the business strategies adopted by different airlines, assessing passengers’ mobility in the presence of direct and indirect connections, or investigating the response to changes in passenger demand and to external economic forces, such as deregulation. Another aspect of interest is the dynamics of the network. This includes the propagation of delay in the airport network, as introduced in Section 2.5 and developed below, and the propagation of disturbances (e.g., capacity violation) in the sector network (Ben Amor et al., 2006). Another important example is the role of air transport in epidemic propagation due to the dynamic of passengers in the airport network spreading infectious disease (see Section 3.4.3 for more details).

The future presents many challenges for the air transport system and CNT is likely to play an increasingly significant role in tackling them. Air transport is rapidly growing worldwide. The current system will reach its capacity limits in a few years due to rising traffic demand and new business challenges. For this reason, large investment programmes like the Single European Sky ATM Research (SESAR) programme in Europe and the Next Generation Air Transportation System (NextGen) in the US have been launched. Policymakers have stressed the importance of fostering the resilience of the system and its capacity to regain mobility levels after an external shock. The future will require an increasing degree of integration between different transport modes. This problem finds a natural description in the multi-layer representation of complex networks (Boccaletti et al., 2014). These issues are not only relevant for air transport but also have important implications for society.

This review is organised as follows. In Section 3.1 we describe qualitatively the types of networks that can be used to describe the air traffic system. In Section 3.2 we review some statistical descriptors of the topological properties of these
ATM networks. Section 3.3 describes the application of the novel approach of community detection to air traffic networks and its use for the design of the airspace. In Section 3.4 we review some literature on the dynamics of air traffic networks, while in Section 3.5 we consider the importance of complex networks in the analysis of the resilience and vulnerability of the air traffic system. Finally in Section 3.6 we draw some conclusions and present some suggestions for future research.

3.1 Different Types of Air Traffic Networks

Many complex systems can be represented by one or more networks. For instance, the Internet can be represented as a set of nodes (the web pages) connected by links (the hyperlinks), but it can also be represented by considering the routers as nodes and their physical connections as links.

The air transport system can also be represented by several different networks, where the set of nodes might differ. It is important to decide which network is under investigation. Since the mobility of passengers is one of the most important aspects, networks describing it and neglecting other more traffic-oriented details (such as navigation points) have been prioritised in previous research. If mobility is the focus, the most important network is that of airports. In this network, nodes represent airports and a link between two nodes is created whenever a direct flight exists between the two airports associated with those nodes. One can see the airport network as the projection of a bipartite network (a network whose nodes are of two different types and no link exists between nodes of the same type), whose first set of nodes is composed of airports, the second set of flights, and a link exists between a flight and an airport if that flight operates to or from that airport. By projection, additional sources of information, like scheduling, types of flights, or airlines, are disregarded.

The network of airports is a directed graph, where two directed links can exist between two nodes A and B, one describing the flights from A to B and one from B to A. The projected network also has a natural weighting scheme, given by the number of flights between the two airports. This is only one of the possible weighting schemes and it neglects the difference between a route covered by large aircraft (and therefore with many passengers) and one covered by small ones. To account for this, one could associate other weights to links using, for example, the frequency of connections or the number of transported persons. From a weighted directed graph one can construct other graphs by considering the weights. For example, taking the difference or the sum between the weights from A to B and from B to A, one can construct a directed network where only one directed link exists between two nodes. Most airport networks are very close to symmetric, i.e., roughly the same number of flights (or passengers) exists from A to B and from B to A. By neglecting the weights one can obtain an unweighted network, where only topology matters, while by neglecting directionality one can obtain a simple or undirected binary
Complex Networks in Air Transport

25

graph. All these alternatives have been investigated in the literature. In any form, the flight network is probably the most investigated network in air traffic studies.

If mobility of passengers (or their delay) is the object under investigation, the representation of the airport network lacks important information. Passengers very often have to use more than one flight to get to their destination. If one considers only a static representation of the network, there is no way of knowing the real or even potential dynamics of passengers, that is, whether one would need to wait 2 or 10 hours in an airport before taking the next connecting flight. Some solutions that have been proposed to investigate the problem of indirect connectivity of passengers are further discussed below.

Airport networks can naturally be broken down into many sub-networks. For example, flight networks can be decomposed by separately considering one sub-network for each airline or for each alliance of airlines. This might highlight the difference in strategy, for example, between major national airlines and low-cost carriers. Analysis of the networks corresponding to a single airline has been performed (Li and Cai, 2004). Cardillo et al. (2013a) investigated interdependencies between the sub-networks corresponding to different airlines, or different alliances of airlines. In this case, the natural representation is the one of a multiplex or a multi-layer network. This is a graph composed of several layers, i.e., where the same nodes (airports) can be connected by different types of links (airlines). This framework could also be useful in the study of the relationships between air and other transport modes (Xu et al., 2011).

Another example of passenger-oriented networks relates to reactionary delays and their effect on passengers. Sometimes a flight cannot take off on time because of a delay to another flight. This occurs, for instance, because of the late arrival of the aircraft or crew (Pyrgiotis et al., 2013). Different networks may be created to study this phenomenon. For example, nodes may represent crews, with a link between them when they share the same aircraft. Alternatively nodes may also represent airports, connected whenever the same aircraft has to serve them in a sequential fashion. The identification of the central elements of these graphs may help in identifying which are the critical points for the dynamics of the system and would allow the creation of better mitigation strategies. A similar approach has been taken in Cook et al. (2013b) in the SESAR project POEM, touched upon in Chapter 2. Here, a Granger causality directed network has been constructed to investigate the propagation of delays. The Granger causality test is discussed in Chapter 7; for now, we note that it is capable of discriminating the causal relationship between two time series (Granger, 1969). In this network a directed link is set between airport A and airport B whenever the Granger causality test indicates that the delays occurring in airport A induce delays in airport B. By studying the communities present in such networks one is able to understand whether delays occurring in a certain airport remain confined within a relatively small set of other airports or whether they affect the system at a global level.

Apart from the airport network, other ATM networks oriented towards the study of traffic control can be defined. Airspace is a complex system, which is
hierarchically partitioned, mainly for capacity and ease of management reasons. The basic elements of this hierarchy are the sectors, which are the smallest operational pieces of airspace. Each is controlled by a pair of controllers and used by the network manager to determine whether a new flight plan can be accepted without violating the maximum traffic load. Recently, Gurtner et al. (2014) introduced the network of sectors. Each sector is a node and two nodes are connected if at least one flight goes directly from one node to the other in the considered time interval. The number of flights between two consecutive sectors can be seen as a natural weight. The network of sectors is naturally constrained by the three-dimensional structure of the real space. Recently the use of community detection on the network of sectors has been proposed as a method to improve airspace design (Gurtner et al., 2014).

Traffic networks can also be studied by considering that aircraft do not travel along the straight line (geodesic) connecting the departure and destination airport. They must follow fixed airways, defined as the union of consecutive segments between pairs of navigation points (‘navpoints’). While such constraints are imposed to improve safety (as it is easier to control ordered flows of aircraft) and capacity (the workload of controllers is reduced, and thus they can control a higher density of aircraft), bottlenecks may appear in some central zones of the airspace, or where several busy airways converge. In the network of navigation points, each node is a navigation point and two nodes are connected if at least one flight goes directly from one node to the other in the considered time interval. The natural weight is the number of flights.

Cai et al. (2012) investigated the Chinese navigation point networks. More recently, as detailed below, Gurtner et al. (2014) have studied the European navigation point network and suggested the use of community detection on this network to enhance airspace design.

Finally, we mention a recent approach to construct networks of air traffic safety events (Lillo et al., 2009). When two aircraft are too close, an automatic alarm, termed the short-term conflict alert (STCA), is activated and the air traffic controller is supposed to give instructions to the two pilots to avoid a collision. One important question is whether these alerts are isolated events or whether the aircraft initially involved are likely to be involved in other STCAs with other aircraft in the near future and so on, creating a cascade of events. This possibility signals the fact that the controller suggests a local solution without forecasting unintended consequences of his or her instructions. By using a dataset of automatically recorded STCA, Lillo et al. (2009) mapped this problem in a network of STCAs. This in turn can be mapped in a network of aircraft, where two nodes (aircraft) are connected if they were involved together in a STCA. These networks show topological regularities and may shed light on the aircraft conflict resolution dynamics.

3.2 Topologies of Air Traffic Networks and Multiplex Networks

We have seen that air traffic networks can have various characteristics and structures: networks of airports, networks of navigation points, networks of airspace sectors, bipartite networks of airports and flights, airplanes and crews, etc. Most
studies have been performed on networks where the nodes are airports (Zanin and Lillo, 2013). However networks of sectors and networks of navigation points have also been investigated.

The functional form of the degree distribution $P(k)$ (introduced in Chapter 2) is the indicator used to perform a basic investigation of the network topology. The information obtained is then typically used to define the ‘configuration model’, i.e., a basic model providing the same degree distribution as the empirical data with random connections between nodes. As with most complex systems, the air traffic network presents a high degree of heterogeneity of nodes (airports, sectors, navigation points, etc.). The configuration model is used as a null model to detect structural deviations that may not be observed if these intrinsically heterogeneous nodes were paired randomly while replicating the degree distribution. Analogous considerations are applicable when studying the strength distribution, measured by weighting the link between two airports by, for example, the number of flights or the number of passengers travelling between them.

For the network of airports, the degree distribution and the strength distribution exhibit truncated power-law behaviour (Zanin and Lillo, 2013). In other words, these air traffic networks belong to the class of ‘scale free’ complex networks. One elementary mechanism for their formation is that of preferential attachment, the tendency of new nodes to establish links with nodes that already have a high degree. The functional description of the degree distribution is therefore important in understanding both the static and the dynamic aspects of network growth.

Chapter 2 introduces various metrics used in network analysis. For airport networks, the most investigated metrics are the degree distribution, the node betweenness and its distribution, the mean length of shortest paths, and the local clustering coefficient (Li and Cai, 2004; Barrat et al., 2004; Guimerà et al., 2005; Xu and Harriss, 2008; Bagler, 2008; Han et al., 2009; Cardillo et al., 2013a; Gurtner et al., 2014). In some cases the rich club coefficient, discussed in Chapter 2, has also been investigated (Colizza et al., 2006b).

The power-law exponent characterising the degree distributions for large values of the degree has been obtained in many systems for different time intervals. The process used to estimate the power-law exponent is not standard and therefore comparison of different results and different sets and for different time intervals is not always straightforward. Even given this caveat, the power-law exponent has been detected to be between 1.0 and 3.0 when airport networks with more than 200 nodes were considered (Zanin and Lillo, 2013). The local clustering coefficient of large airport networks is usually above 0.2 and typically around 0.4 (Cardillo et al., 2013a). The mean length of shortest paths between any two nodes is typically larger than 2 and smaller than 3 for regional areas as large as one continent. The specific value depends on the airport network analysed. For the entire world airport network this value is 4.4 or 4.37 depending on the year of analysis (2000 and 2002 respectively). For the European airport network the value is close to 2.8 (Cardillo et al., 2013a).

The parameters characterising the topology of the airport networks depend slightly on the time interval used to build the specific airport network. For example,
the SESAR project ELSA has computed this exponent for each day from 8 April 2010 to 29 June 2011 (Aeronautical Information, Regulation and Control (AIRAC) cycles 1004 to 1106) using the European data collected by EUROCONTROL. The value of the power-law exponent observed for airport networks obtained each day ranged from 1.7 to 1.9 for the degree distribution and between 1.65 and 1.85 for the strength distribution.

In Figure 3.1 and Figure 3.2 we show the daily degree complementary cumulative distribution function (CCDF), i.e., the probability that the degree is larger than a given value (Figure 3.1) and the daily strength CCDF (Figure 3.2) for the airport network from 6 May to 2 June 2010 (AIRAC 1005). For each value of the x axis, different points refer to different days. There are daily dynamics observed for the two distributions but the overall behaviour is rather stable over time. Both plots use a log-log scale. The leptokurtic character of the degree and strength distributions is evident across the entire 28-day reporting cycle, meaning the probability of large degree is much larger than in a Gaussian distribution. The curves shown in the two figures are compatible with a truncated power law.

**Figure 3.1**  Log–log plot of the daily degree complementary cumulative distribution function of the airport networks for 6 May to 2 June 2010 (AIRAC 1005)
Another informative indicator for the airport networks is the normalised node betweenness. The normalised node betweenness of airports is also characterised by a probability density function with a power-law tail (observed in more than one decade). Therefore a high level of heterogeneity is observed with respect to the normalised node betweenness of the different airports. The normalised node betweenness of each airport is often analysed in parallel with its degree. This is typically done by considering the scatter plot of the node betweenness versus the node degree. Empirical analyses show that large airports tend to have large betweenness, supporting the conclusion that large airports behave like hubs.

In the European air traffic networks there exist a few exceptions to this general trend. One prominent example is Liège airport (EBLG), which sometimes has high values of betweenness and an intermediate value of the degree. More generally, empirical analyses show that a large portion of European medium-sized airports have a relatively high value of normalised node betweenness.

One possible explanation for this empirical observation is that the concept of hub should be associated with airlines rather than airports. For large traditional airlines, it is the airline that organises its own flight network around one or more airports that are the centre of operations and maintenance. Considering sub-networks associated with specific airlines, it can be concluded that the shape of the node betweenness versus node degree scatter plot is specific to each airline.

Figure 3.2 Log–log plot of the daily strength complementary cumulative distribution function of the airport networks for 6 May to 2 June 2010 (AIRAC 1005)
The air traffic networks obtained from flight trajectories and navigation points provide information on the ATM infrastructure underlying flight planning and flight management, reflecting the geographical and procedural constraints induced by the current organisation of ATM. Recently there have been empirical investigations of the topology and structure of these networks in Chinese (Cai et al., 2012) and European (Gurtner et al., 2014) airspace. A network of navigation points is obtained by considering the trajectories (planned or performed) of all flights flying in a given time window and assigning a link between two nodes when a flight trajectory is passing through two adjacent navigation points. With this definition the network for large regions is primarily planar, that is, a network that can be embedded on a planar surface without observing the crossing of links. The cumulative degree distribution of the air traffic networks obtained from navigation points is close to exponential. In the case of the cumulative strength distribution the tail significantly deviates from an exponential decay, although it is hard to classify it as a power-law decay. In Figure 3.3 and Figure 3.4 we show the daily degree CCDF (Figure 3.3) and the daily strength CCDF (Figure 3.4) for the network of navigation points for 6 May to 2 June 2010 (AIRAC 1005). For each value of the x axis different points refer to different days. Both plots use a log-log scale. For the present distributions, the leptokurtic character of the degree and strength distributions is

Figure 3.3  Log–log plot of the daily degree complementary cumulative distribution function of the networks of navigation points for 6 May to 2 June 2010 (AIRAC 1005)
much less evident across the 28 days, especially for the strength distribution. In fact, while the curves shown in Figure 3.3 are compatible with a truncated power law, the same cannot be said for the curves in Figure 3.4.

We have already discussed how the airspace system and the ATM are characterised by relationships and procedures that might be efficiently described as multiplex. With few exceptions, airport networks have been empirically analysed unconditionally with respect to the specific airline responsible for each connection. In parallel to these unconditional analyses of the projected network of airports, an analysis of a multiplex of the projected network of airports can be performed with specific links associated with each airline. Cardillo et al. (2013a) have studied how the aggregation of 37 layers associated with 37 large, distinct airlines compares with the unconditional airport network. Such an analysis of a multiplex is more informative than the unconditional one. The study indicated that the relatively high value of the local clustering coefficient (see Table 2.3) is due to triangles among airports that are primarily associated with more than just a single airline. Moreover, behaviour differs between airport networks obtained by aggregating layers of major airlines and airport networks obtained by aggregating low-cost airlines both in the cumulative degree distribution and especially in the mean length of shortest paths.

The study of multiplex is just beginning. It will increase and may allow detection of local structures of the multiplex, such as the ones showing enhanced

![Figure 3.4](image_url)

**Figure 3.4** Log–log plot of the daily strength complementary cumulative distribution function of the networks of navigation points for 6 May to 2 June 2010 (AIRAC 1005)
co-occurrence of links (in the case of competition) or absence of co-occurrence (in the case of alliances and synergies).

3.3 Communities in Air Traffic Networks and Airspace Design

The configuration model is the simplest random model explaining the heterogeneous interconnections of the nodes of a complex network. Empirical complex networks usually deviate from the corresponding configuration model and display regions of nodes that are more interconnected (and correspondingly regions that are less interconnected) than those expected by the configuration model. In this way one can identify sets of nodes that might be characterised by similar attributes or conditions. These sets of nodes are called in network science ‘communities’. The origin of the name is due to the fact that sets of nodes with similar attributes or conditions were originally investigated in social networks (Scott, 2000). Communities (or clusters or sets of nodes) are an important aspect of understanding and modelling the architecture of a real network. The literature on social networks and the literature on network science have proposed many methods for detecting communities in complex networks (Fortunato, 2010).

Recent studies have identified communities in different air traffic networks. Community detection in air traffic networks (of airports, sectors and navigation points) might provide information about the coherence between the flights performed and the current airspace design and can be used to help design the structure of airspace from the bottom up, that is, starting from the empirically observed state of the system (Gurtner, 2014).

Airspace is hierarchically partitioned for reasons mainly related to ATM and control. At the highest level, European airspace is partitioned into multinational areas, termed ‘functional airspace blocks’. These are not yet fully implemented, but their activation is planned in the near future. Each country also has its own national airspace, which is typically partitioned into ‘area control centres’. Each of these is itself partitioned into sectors, which are the smallest unit of control, being under the direct supervision of air traffic controllers. Inside the sectors we find the navigation points constituting the grid through which the flights move. The boundaries of these multiple partitions are defined taking into account political and strategic factors and the traffic characteristics.

The empirical results obtained by performing community detection on air traffic networks of airports, sectors and navigation points provide lists of nodes of the different air traffic networks obtained from the planned and/or executed flights. These sets of airports, sectors and navigation points can be compared with standard partitioning comparison indices such as, for example, the adjusted Rand index (Hubert and Arabie, 1985) with the planned hierarchical organisation of the system.

The communities observed overlap significantly with the planned and future (in the case of the functional airspace blocks) hierarchical organisation of the
airspace. However, significant deviations between the present partitioning and the partitioning obtained by community detection analysis were also observed, suggesting that some organisational decisions, most probably arising for political and strategic reasons, are not compatible with indications arising from traffic data.

The results obtained from community detection of air traffic networks underline the difference between the bottom-up approach, with an unsupervised partition based on the traffic, and the top-down approach of the planning of existing partitions. Community detection is based on heuristic approaches that might provide different outcomes depending on the different approaches and fitness measures used (Fortunato, 2010). In the analysis of air traffic networks it has been shown that different community detection algorithms are able to capture different features of airspace organisation (Gurtner, 2014). They are able to suggest alternative ways of redesigning some structures of the airspace. For example, area control centres might be more densely connected inside and have fewer interfaces (links) with the exterior, which might create added value from an operational point of view. Another potential application is the use of the empirical detection of communities in air traffic networks to detect optimal boundaries between sectors, air traffic control (ATC) centres, national airspaces, and functional airspace blocks that require intensive coordination.

3.4 Air Traffic Network Dynamics

The structure of the air traffic network is crucial for the dynamics of variables of social, economic and technological importance. Examples range from passenger mobility to the economic impact of aircraft delays or to the optimal functioning of the ATM system. For studies of these dynamics, the air traffic network considered is that of airports.

3.4.1 Passenger Dynamics

The structure of the air traffic network affects the ability of a passenger to reach his or her destination in the shortest possible time and with fewest changes. However, purely topological metrics can be poor indicators for assessing passengers’ needs. For instance, a shortest path (in terms of number of flights) detected in a network is not necessarily usable; the network is not usually complete, so some possible connections in the shortest path are actually missing. Indeed, Malighetti, et al. (2008) show that roughly two-thirds of the fastest indirect connections (i.e., airports connected through one or more other airports) are not operated by the airlines alliance system. That clearly poses new challenges for the definition of new network metrics more representative of the specific socio-technical complex system we are considering. For example, Malighetti et al. (2008) introduce the essential betweenness as the number of unavoidable minimal paths passing through an airport. A high value of the essential betweenness indicates that the airport is a bottleneck for the traffic in the system.
Another typical issue investigated regards the travel time from one airport to another and the related necessity of a network metric that combines information about the minimum number of connections between two airports, the minimum time to travel between them and the minimum waiting time. Such issues are dealt with in the context of a hub-and-spoke network structure. For each pair of airports the shortest travel time between them is calculated for a passenger who wants to leave at a specific time and to arrive at the destination on the same day. Scheduled flight time data for a specific day are used, allowing at least one hour for a flight connection. The optimal path between the two airports is identified as being the one with a minimum travel time with a given starting time on the day. The introduction of these new metrics might therefore be useful not only for characterising the system but also for passengers who would like to exploit the optimal travel options between two airports, given their personal requirements.

The requirement for more effective network metrics therefore depends on the viewpoint from which we consider the system. The average delays of (delayed) flights and passengers are not the same. The air transport industry has until recently paid poor attention to passenger-centric metrics, its reporting being essentially flight-centric. Trade-offs between these metrics need to be better understood, as they are observed to behave in opposite ways under certain types of flight prioritisation criteria. With growing political emphasis in Europe on service delivery to the passenger, and passenger mobility, there is a need for passenger-oriented network metrics that augment the usual (purely topological) network metrics with passenger-related information. Cook et al. (2013b) investigated several flight prioritisation strategies. One promising strategy assigned departure times of aircraft at an airport based on cost minimisation (for example, passenger rebooking costs). This markedly improved a number of passenger-related delay metrics, as well as airline costs, the latter determined by reductions in passenger hard costs to the airline. The authors also investigated the impact of arrival and reactionary flight delays on passenger delays depending on the business model of airlines and the airport size. One might think that large airports would be associated with more reactionary and arrival delay, leading to larger passenger delays. Instead it was found that for some of the smaller airports, arrival delay was doubled (or even tripled) into reactionary delay. One main factor is the reduced delay recovery potential for the airline at such airports, for example through limited crew and aircraft resources. Other factors include flexible or expedited turnarounds, and whether a given airport has sufficient connectivity and capacity to re-accommodate disrupted passengers.

3.4.2 Propagation of Delays

The propagation of delays plays a crucial role in the propagation of other phenomena such as those connected to passenger dynamics. The propagation of delays is in turn affected by the structure of the air traffic network. In particular, we are interested in air traffic congestion and flight delays, where the propagation of delays can give an indication of the spread of a malfunction across the system.
Lacasa et al. (2009) consider a simple random (Erdős–Rényi) network topology. A critical transition distinguishing a free-diffusing aircraft regime from a congested one has been observed. Specifically, the percentage $P$ of aircraft that are not stuck in a node’s queue remains constant for low values of the aircraft density. If this density exceeds a threshold value, the percentage $P$ rapidly decreases. This is a typical case of phase transition that has also been observed in other transport systems (Nagel and Schreckenberg, 1992). Such behaviour is resilient to changes in the network topology. The simulations that the authors have performed, starting from the real air transport (scale-free) network, for which a power-law degree distribution is observed, are qualitatively similar to those obtained with the random (Erdős-Rényi) topology, for which a binomial degree distribution is observed.

In Section 2.5 we introduce research in the US on delay propagation modelling, in the context of costs and contrasting performance metrics. Corresponding work in Europe is less common, largely due to the lower availability of necessary (passenger) data. For reviews of the literature, the reader may consider the papers by Cook et al. (2012) and Campanelli et al. (2014). Where connectivities are modelled, such models mostly use hypothecated passenger and/or crew connectivities, although two exceptions in the European context are mentioned below.

Wonga and Tsai (2012) examine flight delay propagation using Cox regression analysis (a proportional hazards model), and find that the key factors contributing to departure delays include turnaround buffer time, aircraft type, cargo and mail handling, technical and aircraft equipment, passenger and baggage handling, and weather; for arrival delays, the key factors include block buffer time and weather. In Europe, the strong relationship between departure delay and arrival is well established, indicating relatively little change during the gate-to-gate phase (EUROCONTROL, 2014).

An analytical queuing and network decomposition model – approximate network delays (AND) – has been used to study delay propagation for a network comprising the 34 busiest airports in the US and 19 of the busiest airports in Europe (Pyrgiotis et al., 2013). The model treats airports as a set of interconnected individual queuing systems. Due to its analytical queuing engine, it does not require multiple runs (as simulations do) to estimate its performance metrics and can evaluate the impacts of scenarios and policy alternatives. The use of explicit passenger data is planned for this model. Actual passenger transfer numbers have also been used in numerical simulations of a major US hub, where it was demonstrated that when a balancing objective function was applied, each metric studied – terminal transit times of passengers, aircraft taxi times and gate conflict durations – outperformed observed values (Kim et al., 2013).

Covering 305 US airports in 2010, an agent-based model reproduced observed delay propagation patterns (Fleurquin et al., 2013). Estimated passenger and crew connectivities were identified as the most relevant factors driving delay propagation. The probability of such connections was modelled as proportional to flight connectivity levels at each airport, with such connectivity being the most parameter affecting the propagation of delays. The authors study how congested
airports form connected clusters in the network. Further developments of this work within the SESAR project TREE, extending the research into the European context, are reported in Campanelli et al., 2014, where further promising agreement with real flight performance data is reported, and the intention to incorporate real passenger connectivity data is indicated.

We mention in Chapter 2 that the SESAR project POEM built the first full European network simulation model with explicit passenger itineraries and delay cost estimations (Cook et al., 2013b). Compared to current operations, when this network operated under rules to minimise airline costs, delay propagation was found to be contained within smaller airport communities, but these communities were more susceptible to such propagation, illustrating an important trade-off. The nature of a network may, therefore, be changed not only by attacks (or failures), as with Internet hubs (see Chapter 2) but also by operational rules. Further work on community detection is planned by this team to explore these effects in more detail, and illustrations are presented regarding the importance of multiple layers for CNT metric determination, in Chapter 7.

3.4.3 Spreading of Epidemics

Air passenger transport constitutes one of the most important vectors for long-range spreading of an epidemic. This was one of the first examples of dynamic effects on the air transport network to be considered and perhaps one of the most successful studies. Colizza et al. (2006a) used real data on passenger mobility to study the role of the large-scale properties of the airline transport network in determining the global diffusion pattern of emerging diseases. To distinguish the role of the network structure in the spatio-temporal pattern of the epidemic process, the authors introduced a characterisation of the epidemic pattern by using the normalised entropy, \( H \). If the epidemic is homogeneously affecting all nodes the entropy attains its maximum value \( H=1 \). The value \( H=0 \) corresponds to one initial infected airport. To investigate the effect of the network structure, the authors consider two model networks. The first model network (called HOMN) is a homogeneous Erdős–Rényi random graph with the same number of vertices \( V \) as the real one, and is obtained as follows: for each pair of vertices \((j, l)\), an edge is drawn independently, with probability, valid for all links, proportional to \( <k>V \), where \( <k> \) is the average degree of the real network and \( V \) is the number of airports in the network. In this way, they obtain a typical instance of a random graph with a Poisson degree distribution, peaked around the average value \( <k> \) and decreasing faster than exponentially at large degree values. This is in strong contrast to the empirically detected degree distribution of the real airport network. For the second model (called HETN) they retain instead the exact topology (i.e., degree distribution) of the real network. The results seem to indicate that, when considering the airport network, the fat tails of the degree distribution in the HETN network determines to a large extent the overall properties of the empirically observed epidemic pattern.
Recently, this approach has been used to assess the international spreading risk of the 2014 West African Ebola outbreak (Gomes et al., 2014). Part of the model simulates the number of passengers travelling daily worldwide on each airline connection in the world. They conclude that ‘the risk of international spread of the Ebola virus is still moderate for most of the countries’, but ‘if the outbreak is not contained, the probability of international spread is going to increase consistently’. They used their model to simulate the effects of possible containment measures, such as the suspension of flights by some airlines, travel restrictions, ability to detect Ebola cases during international flights, etc.

3.5 Resilience and Vulnerability

We first meet resilience and vulnerability in Chapter 2. Resilience is the property by which a system is able to recover its normal functioning status after the occurrence of internal or external disturbances. In this respect, resilience can be seen as a measure of the efficiency of the system. Conversely, vulnerability is the inability to withstand the effects of disturbances (EUROCONTROL, 2009a). The concept of resilience in the context of ATM is considered thoroughly in Chapter 5. Here we are mainly interested in discussing some aspects of resilience and vulnerability that are connected with the network representation of the air transport system.

Resilience is a crucial topic in network theory. Scale-free networks (networks showing a power-law degree distribution) show an extreme vulnerability to attacks; the removal of a few key nodes that play a central role in maintaining a network’s connectivity can cause widespread significant disruption, both globally (Albert et al., 2000) and locally (Crucitti et al., 2003). These networks do, however, display a high degree of error tolerance. This result contrasts with what is found for random networks, such as the random graph model of Erdős–Rényi and the small-world model of Watts and Strogatz; both lead to a fairly homogeneous network, in which each node has approximately the same number of links and the degree distribution fast decays to zero for large degree values. These types of networks show low vulnerability to (targeted) attacks, due to the fact that, by construction, all links play approximately the same role in the network topology.

Resilience in the context of ATM systems has been mainly studied in the context of safety events (Gluchshenko and Foerster, 2013; Woltjer et al., 2013) and hardware malfunction (White et al., 2012). Here we are interested in the relationships between the topology of the air traffic network and its resilience or vulnerability. Resilience in the European network from a passenger perspective has been investigated by Cook et al. (2013b), who explored the robustness of different prioritisation rules under disruption and investigated how the network loses functionality as a consequence of major disruption at one of its nodes. The top ten airport vulnerabilities were all hubs, mostly serving a high number of direct destinations relatively poorly covered by near-neighbour airports (Cook et al., 2013a).
Chi and Cai (2004) analyse how the main topological properties of the US air transport network are changed by random failures and targeted attacks. The first issue is investigated by deactivating airports at random and thus simulating random disturbances like emergency situations or critical weather conditions. The topological properties, including average degree, clustering coefficient, diameter and efficiency, are slightly affected when several of the least connected airports are randomly removed. Such properties, however, change drastically with the removal of several of the most connected ones. The degree distribution and the weight distribution under errors behave similarly to those of the original network. The second issue is investigated by deleting the most connected nodes, as in an intentional terrorist attack. Under attacks, the degree distribution changes from a two-segment power law to a monotonic one. The under-attack weight distribution still displays a power-law tail, although with a slightly smaller exponent.

A similar approach was used by Hossain et al. (2013), who also investigated random disruptions of flight paths, for example, when airways become unavailable due to bad weather. They found that the investigated network is resilient to random failures of airports and random disruptions of flight paths. The network remains connected and incurs minimal increase in travel times and reduced ‘reachability’ when most of its nodes (airports) are removed or its edges (airways) become randomly unavailable. In the case of a targeted failure (a targeted isolated airport shutdown) the network is more sensitive to node failure by a descending order of degree as well as ‘betweenness centrality’.

A different perspective involves assessing resilience and vulnerability by considering real stress events. A typical case is that of the Eyjafjallajökull volcano eruption, which occurred in 2010, and the associated ash cloud. This topic is addressed in Chapter 5. Wilkinson et al. (2012) showed that the effect on air traffic was disproportionately severe because the network shows a truncated, scale-free distribution and has a spatial degree distribution that is uniform with distance from the centre of the network, resulting in a network that is vulnerable to spatial hazards. These distributions result from a combination of the desirability of a location, space limitations and the distance users are prepared to travel overland to an airport. It is debatable whether these factors might be changed to mitigate these disruptions and still fulfil the airlines’ and passengers’ requirements.

3.6 Conclusions

The application of CNT to air traffic management has seen significant growth in recent years. This is partly because air traffic can be seen as the superposition of different networks, including the networks of airports, sectors and navigation points. Moreover each of these networks can be seen as a multiplex – for example, by associating each layer with a different airline. The study of the topology of these networks is important for several reasons related to understanding, monitoring, controlling, and optimising the air traffic system. The topological properties of air
traffic networks are useful: (i) for studying how the air traffic has changed in recent years; (ii) for identifying the more central and vulnerable nodes of air traffic networks; and (iii) for identifying communities of nodes that can be naturally associated with units of ATC (e.g., sectors or airspaces). The topology affects the dynamics that can take place on a network, such as the mobility of passengers and goods, the propagation of disturbances (e.g., delays, congestion, and so on), and the propagation of epidemics. Despite this large range of applications of CNT to ATM, we believe that the interaction between these two fields is only beginning and new and interesting contributions will be published in forthcoming years.

In the near future, tremendous air traffic growth is expected. As the Single European Sky (SES) and SESAR programmes testify, the current system seems unable to support such an increase and new tools and methods must be developed. In addition, interoperability and connectivity between different transport modes is on the agenda of several international institutions. This creates new, exciting challenges, and problems that need to be tackled using many different approaches. We believe that CNT could provide a significant contribution to this challenge.
4.1 Uncertainty in Context

Uncertainty is loosely defined as the condition of having only partial or limited knowledge about the existing state or a future outcome. All socio-technical complex systems exhibit some uncertainty (McDaniel and Driebe, 2005), and ATM is not an exception.

Flight passengers are faced with unexpected delays and cancellations due to bad weather or traffic congestion. Connecting passengers can miss flights due to a domino effect (the accumulation of delay from several flights). Pilots face uncertain weather and wind conditions, with inexact forecasts, and have to rely on onboard equipment to fly in hazardous conditions. Air traffic controllers deal with uncertainty daily, by setting separation buffers or by sending aircraft into holding patterns.

In the past, uncertainty in ATM has been either ignored or barely mentioned, except in some specific applications such as navigation systems or the separation buffers between aircraft. Recently, researchers have begun to recognise the importance of uncertainty in the air traffic system.

Accomplishing the Single European Sky (SES) high-level goals of improving capacity and safety (see Section 1.2) requires a paradigm shift in operations through state-of-the-art, innovative technology and research. A promising approach that can improve current prediction and optimisation mechanisms towards meeting these goals is to model, analyse, and manage the uncertainty present in ATM. New ideas and concepts are emerging, such as uncertain four dimensional (4D) trajectories, uncertain ATM networks, and stochastic conflict resolution.

However, uncertainty is a multi-layered concept and can be approached from many sides. ATM researchers working on the subject lack a common language or framework that would facilitate not only dissemination of their work but also collaboration towards common objectives. Thus, the purpose of this chapter is to define with clarity the concept of uncertainty, differentiating between the various sources that affect the ATM system, and to provide a framework to study uncertainty in the context of ATM. This framework must be as general as possible, to cover all aspects of ATM and to encompass the many possible approaches to studying uncertain systems.
4.1.1 The Definition of Uncertainty

There are many definitions of uncertainty. For instance, from the point of view of an individual, any statement that he or she does not know with certitude to be either true or false, is uncertain (Lindley, 2006). In this sense, uncertainty is different for each person.

In the context of socio-technical systems and, in particular, in ATM, uncertainty is defined as the condition of being partially or totally unknown or in doubt, which can refer to the truth of certain statements (for instance, whether a flight is delayed or not) but more typically refers to the precision of certain quantitative values (such as the geographical coordinates that describe the position of a flying aircraft at a given instant or an estimated time of arrival, quantities which are never known exactly).

To analyse uncertainty, the first step is to appropriately represent it. Perhaps the most common representation of uncertainty uses probability, but this is not the only representation and is not always the best one (Halpern, 2003).

When speaking of uncertainty in a system, it is important to distinguish between objective and subjective uncertainty. In objective uncertainty, also known as fundamental uncertainty, the system is intrinsically non-deterministic, that is, it does not evolve in a deterministic way (for example, quantum systems or radioactive decay). It is inherently impossible to perfectly know its state. In ATM, some examples of objective uncertainty arise in atmospheric dynamics. For instance, in a baroclinic situation, arbitrary and unpredictable distortions might trigger cyclonic conditions, but the onset of that instability in space and time is not deterministic. Characteristics of the cyclone itself, for instance, the precise rain rate at any location within, or the central pressure minimum at a given time, cannot be predicted with precision. While baroclinic instability can be simulated by numerical weather prediction models, when experiments are repeated with slightly different (but physically equivalent) initial conditions, solutions remain close until baroclinic instability develops and then solutions diverge without bound. The differences in initial conditions might be as small as a difference of a 100th of a degree in temperature at an arbitrary location, while all other initial conditions might be equal. Thus, the forecast accuracy of numerical weather prediction models is limited and depends on the weather situation, the region, and the variable considered. As a rule of thumb, one can consider a typical limit of seven days due to cyclogenesis. Higher spatial resolution models and the integration of more observational data can improve forecasts. Probabilistic nowcast systems are beginning to be developed to explicitly account for forecast uncertainty (Megenhardt et al., 2004).

In the case of subjective uncertainty (also known as ‘hidden determinism’), the system itself is deterministic, but still its state is not perfectly known. There are several possible reasons. The inputs that affect the dynamics of the system may be partially unknown or imperfectly measured (for example, due to sensor noise), or the mathematical laws of evolution of the system may not be exactly determined (for example, when human factors play a significant role in the system). This
discussion centres around subjective uncertainty (called simply ‘uncertainty’ in this chapter), as it is common to all socio-technical complex systems to some degree due to the presence of numerous agents, complex interactions between them, human factors, imperfectly known dynamics, and measurement errors, among other factors. Objective uncertainty will manifest itself mainly through meteorological phenomena.

Uncertainties may also be classified according to their perceived potential to be reduced. Uncertainties are characterised as epistemic if there is a possibility of reducing them by gathering more data or by refining models, and as aleatory if the possibility of reducing them is not foreseen (Kiureghian and Ditlevsen, 2009).

4.1.2 Sources of Uncertainty

The sources of uncertainty that appear in ATM are very disparate, which explains the sometimes radically different approaches taken when researching them. In this chapter, they are classified into the following broad types.

• **Data uncertainty** This type of uncertainty appears when there are some data which are known, but in an inexact way, i.e., there exists some level of uncertainty in their values. Imperfect models, which produce uncertain outputs, also fall in this category. Some typical examples include the aircraft position given by a global positioning system (GPS; which includes a bounded error of the order of several metres), the aircraft take-off weight (which is seldom known perfectly), and simplified aircraft performance models (such as EUROCONTROL’s Base of Aircraft Data (BADA)).

• **Operational uncertainty** The decisions taken by individuals (pilots, air traffic controllers, or airline and airport managers) have a very significant influence on the ATM system, but are difficult to predict (even if they are based on perfectly known rules). Thus they introduce a degree of uncertainty, which can affect, for example, taxi times or departure or arrival times. Human factors play a key role in modelling this type of uncertainty.

• **Equipment uncertainty** This type of uncertainty refers to problems in equipment, ranging from aircraft or vehicle malfunction and breakdown to total failure. As a result of such problems a system can suddenly deviate from its normal mode of operation, behave erratically, or even stop working. This will in turn affect other related or neighbouring systems, increasing the level of uncertainty locally or even globally. For instance, loss of communication between aircraft and air traffic controllers due to faulty radio equipment may produce delays in other aircraft due to augmented security measures, or may, in a worst-case scenario, compromise safety.

• **Weather uncertainty** This type of uncertainty includes uncertain wind velocity, fog, temperature, snowfall and thunderstorm regions and time intervals. Meteorological conditions are accurately predictable only for a
short time horizon; therefore, flight plans have to be based on estimations, which are often far from reality. In particular, adverse weather can introduce high levels of uncertainty, not only in a particular trajectory but also in the air traffic as a whole. Strategies to deal with adverse weather may include rerouting or cancelling flights. While weather uncertainty is a combination of objective uncertainty and data uncertainty, uncertainty about meteorological conditions has a very large impact on the ATM system and thus deserves to be classified separately.

This classification is very similar to the one presented by Heidt and Gluchshenko (2012). Other authors consider different classifications of sources of uncertainty; see, for instance, Driebe and McDaniel, 2005.

For a given ATM scenario, a careful and detailed characterisation of the different sources of uncertainty that affect the situation is essential. This characterisation has different components to:

- identify and classify all the relevant sources;
- determine their statistical properties;
- identify the dependence structure between different sources;
- characterise and quantify their effects (for instance, whether they affect the whole scenario or only parts of it).

A key question is how to simplify modelling by excluding negligible sources of uncertainty; these can be found by performing a sensitivity analysis, either analytically or numerically. For instance, when studying the uncertainty of the airborne stage of the aircraft trajectory (see Section 4.2.2), a standardised sensitivity study on the full range of trajectory prediction tools would allow comparison of the sensitivity results obtained, and could, therefore, be used to identify which uncertainty sources can be removed from the analysis.

4.1.3 Scales in ATM

The analysis of how a given problem is affected by the various sources of uncertainty is linked to the relevant scale or scales of the problem. These scales are not unrelated, since uncertainty in smaller scales may propagate to larger scales, and vice versa.

There are three clearly differentiated scales depending on the level of detail and aggregation. These scales are defined in Section 2.2, for general complex systems. In the context of this chapter, it is useful to link the scales to concrete elements of the ATM system to set the stage for the discussion in the next sections. The scales and the ATM element they represent are as follows:

- **Microscale** This represents a single flight. All the uncertainties that affect a flight make the flight itself and all quantities related to it uncertain. This
is called ‘flight uncertainty’. At this smallest scale one must analyse all the uncertainty sources that affect the flight at its different stages: strategic, pre-departure, gate-to-gate (this includes ground and airborne stages), and post-arrival. Of particular interest is the analysis of the uncertainties that affect the airborne stage of the aircraft trajectory, which requires consideration of the dynamics of the aircraft and the changing environment through which the aircraft moves (the atmosphere). Flight uncertainty is, in turn, a source of uncertainty for other scales in which the flight is a basic element, for example, traffic or network problems.

- **Mesoscale** This represents the air traffic. This is an intermediate scale that focuses on a given area containing many individual aircraft that interact following a given set of rules; for instance, a terminal manoeuvring area or a sector. This scale still considers individual vehicles, but describes their activities and interactions based on aggregate relationships. The mesoscopic scale allows, for example, the assessment of the potential impact of routing rules and conflict resolution strategies on aircraft delay, or the study of propagation patterns and performance degradation. The analysis of flow management problems is also framed within this scale, because, even though they affect the air transport network, their solution requires a finer level of detail than that of the network scale.

- **Macroscale** This represents the air transport network. Air transport can be considered at the level of a regional, national, or trans-national network, or even the global ATM system. This scale integrates the state of the various ATM elements and allows focus on the network properties, giving a high-level view of the system. It is important to study how uncertainty in flights and air traffic (microscopic and mesoscopic scales) propagates to affect the macro-scale, even though some authors consider the aircraft trajectory perfectly known (without any uncertainty) at this scale. Other relevant problems include the study of how operational and weather uncertainties affect the air transport network.

These scales are separately studied in Sections 4.2, 4.3, and 4.4, respectively, since the treatment of uncertainty is rather different for each of them. Within each of these scales, one can find two types of problems related to the time horizon under analysis.

- **Estimation of the present state** Over a short-term time horizon, the main concern is to enforce safety. Data uncertainty is the principle uncertainty source. Information sharing and filtering techniques can be used to reduce uncertainty.

- **Prediction of the future state** For medium and long-term time horizons, the main concern now is efficient planning. This must consider flight plans, weather forecasts and predicted traffic. Uncertainty propagation and
‘domino effects’ have to be minimised while optimising system performance. All uncertainty source types affect this problem; for long-term planning, data uncertainty is always present, while for the medium term, weather uncertainty can have a considerable impact, causing unexpected planning changes or cancellations.

4.2 Flight Uncertainty

The individual flight lies at the smallest scale of the problem. Study of uncertainties at this microscale is important because the flight is the core element of ATM.

Flight uncertainty encompasses all the uncertainties present at the different stages of the flight. The term ‘stage’ has been used to avoid confusion with ‘phase’ (which is conventionally reserved for describing the elements of a flight’s operation from gate to gate and is not wholly appropriate for its strategic and post-arrival stages). The stages of a (typical) flight are the following.

- **Strategic** This covers the timeframe months before the flight up to two hours before the off-block time. This includes the filing of flight plans but not the air traffic flow management (ATFM) slot allocation process.
- **Pre-departure** This includes slot allocation, commences two hours before and continues up to the aircraft off-block time.
- **Gate-to-gate** This includes the ground stage (taxi-in and taxi-out) and the airborne stage.
- **Post-arrival** This commences once the aircraft is on-block.

Next, two key aspects of flight uncertainty are discussed. The first is flight delay, which is a manifestation of the temporal uncertainty of the flight. Its study is of the utmost importance as it has effects on both traffic and network scales. The second is the uncertainty associated with the aircraft airborne stage; in this chapter this is called ‘trajectory uncertainty’ (this denomination is common in the literature when studying uncertainty at the airborne stage, although, in fact, the ground movement is part of the gate-to-gate trajectory).

4.2.1 Flight Delay

The analysis and prediction of flight delay are topics of great interest with an abundant literature. In this section, we focus primarily on delay causes in the context of the phases of flight. The specific topic of delay propagation through networks is discussed in Section 3.4.2) and the cost of delay is introduced in Chapter 2.

During the strategic planning stage, uncertainties might be large. Airport slots, and the planned time for take-offs or landings, can change from season to season; IATA’s Schedules Conference is held twice a year. Uncertainty might arise if, for instance, too many flights are scheduled to depart or arrive during a time window.
It is not unusual to have up to ten flights scheduled to depart or arrive at the same time, to be resolved through operational practice at the airport and air traffic control (ATC). When airlines file flight plans far in advance, many flight details are uncertain. For example, airlines rely on the Integrated Initial Flight Plan Processing System (IFPS) to complete route information when final processing occurs 20 hours prior to the flight. In addition, predictions of sector congestion are also less reliable, and weather data are only known with reasonable reliability a very short time period before the flight. Airlines ‘direct filing’ flight plans to IFPS (five days before the flight onwards) have the opportunity to take advantage of more dynamic information as it becomes available. Even so, potential slot delay is still an unknown.

In 2008, which represents a peak in European air traffic volume (with 10.19 million flights, compared with 9.45 million flights during 2013) the standard deviation of delay for all causes during the pre-departure stage was 17 minutes (16 minutes in 2013) (EUROCONTROL, 2009b; EUROCONTROL, 2014). Around 20 per cent of the flights in 2008 were regulated by ATFM slots, with the calculated take-off time determined via receipt of a slot allocation message. According to EU Regulation No 255/2010 (in force since September 2011), member states of the European Union must ensure that airports adhere to ATFM departure slots and where the adherence is 80 per cent or less during a year, the Air Traffic Services’ units at the airport concerned must detail the actions taken to ensure future adherence. A 15-minute take-off slot tolerance is available for ATC departure sequencing purposes (EUROCONTROL, 2010a), which can be −5 minutes to +10 minutes in relation to the calculated take-off time. In 2008, 18.5 per cent of regulated flights took off outside this 15-minute slot window (this number diminished to 13 per cent in 2013). Although not all ATFM slots result in a delay, slot allocation and the management of slots contribute to uncertainty. For example, an allotted take-off time can be revised or an initial slot allocation message can be issued late, close to the off-block time. Uncertainty relating to other aspects of a flight can result in inefficient slot use – for instance, rather than delay a regulated flight in order to rectify a ground-based problem, an airline might retain the slot in the hope that the problem can be overcome and the slot used, although this may lead to the slot being wasted when it is too late to reallocate the ATM resources efficiently.

Another cause of delay during the pre-departure stage is airport ground handling, which is the planning, scheduling, and control of all aircraft turnarounds at an airport. Turnaround operations consist of several parallel and sequentially running ground-handling activities (such as disembarking and embarking passengers, unloading and loading luggage, maintenance checks, fuelling, cleaning and catering), which partly depend on each other or on the successful completion of other running processes. Uncertainty in turnaround operations is mostly characterised by variations from the predicted off-block time. Field data show that turnaround performance is highly dependent on the operational flight parameters, for example, the number of passengers, aircraft type or flight distance. Furthermore, external factors may contribute uncertainty to turnaround performance; Oreschko et al.
Damián Rivas and Rafael Vázquez (2012) have shown that the most influential factors are arrival delay, airport category, the skill level of ground staff, and weather.

Independent ground service providers handle many of the turnaround activities. The interests of the service providers may not be in line with the best interests of the airline (or passengers). This makes scheduling the activities in airport ground handling a challenging task. Efficient procedures to reduce this source of delay have been developed; for example, Mao et al. (2009) developed an algorithm for the scheduling of airport ground-handling services, including disruptions, using a heterogeneous multi-agent scheduling frame and online scheduling to cope with uncertainty, considering cooperative and non-cooperative settings.

Next, the gate-to-gate stage covers the ground phases (taxi-in and taxi-out) and the airborne phases. In 2008, the standard deviation of delay during this stage was found to be 19 minutes (typically due to sequencing, weather and other local disruptions), reducing to 17 minutes in 2013. As in the pre-departure stage, the arrival time delay is based on the airline’s last-filed estimated off-block time and is thus an underestimate of the full variability of the arrival time, as airlines may refuse to manage delay. Uncertainty once the aircraft is airborne is studied separately in Section 4.2.2. Focusing on the ground stage, the actual taxi time can differ greatly from the planned taxi time, impacting on the take-off time. Standard taxi times can be modified at short notice in response to changing airport conditions, and congestion, or the lack of it, can affect taxi times dynamically.

The post-arrival stage commences once the aircraft is on-block and it is a particularly pertinent stage for delay propagation. During this phase, the mechanics of passenger connectivities (requiring airline decision-making on whether to wait for delayed inbound passengers), crew connectivities (again with decisions on whether to hold flights for inbound delayed crews, or to implement crew swaps, where possible) and aircraft rotations and swaps come to the fore. Added to these uncertainties are the dynamics of the availability of the handling agent (and ramp team), plus the potential loss of ATFM slots due to delayed departures. This phase, therefore, plays an important role in determining ‘knock-on’ or reactionary delay (the modelling of which is discussed in Section 3.4.1). This is especially important at hubs, where passenger (and crew) connectivities may both impose particular operational constraints and offer corresponding solutions (e.g., greater passenger reaccommodation potential, Cook et al., 2013b).

Predicting flight delay at hubs is thus of particular interest. Andersson et al. (2000) use queuing models to capture the taxi-out and taxi-in processes and an integer-programming model to represent airline decision-making (attempting to capture the dynamics of the aircraft turnaround process). Balakrishna et al. (2010) introduce more complex models, developing a prediction method for taxi-out delay at Tampa International Airport using nonparametric reinforcement learning, set in the probabilistic framework of stochastic dynamic programming. When predicting delay, the time horizon of the prediction will change the accuracy of the estimates. Ohfeldt et al. (2011) found that, compared to the in-flight (shorter horizon) uncertainty, which is typically assumed to be a normal distribution, the uncertainty for
4.2.2 Trajectory Uncertainty

The most significant part of the flight is the airborne trajectory, which is affected by different sources of uncertainty. All these sources make the airborne trajectory itself uncertain.

Trajectory uncertainty analysis requires the study of the relevant sources of uncertainty, their mechanism of propagation (mainly, the dynamics of the aircraft and associated models), and also the mechanisms of uncertainty mitigation that the aircraft might have (typically, feedback loops that use navigational aids and sensors). The most important sources of uncertainty affecting the trajectory are usually the following: uncertainty in the initial conditions, uncertainty in the aircraft performance models, wind uncertainty, navigational errors, flight management system errors and operational uncertainty.

When dealing with trajectory uncertainty, the most appropriate framework is that of probability and statistics (Feller, 1968). Sources of uncertainty such as the initial mass are usually modelled by random variables, which are described by distribution functions. The uncertainty in each random variable is given by its covariance. The aircraft position at a given time can be described in a similar way; a very useful construction is that of a region of confidence, whose centre lies at the expected value of the position and includes around it a region of space where the aircraft could be located with a certain degree of probability (the volume of a region of confidence depends fundamentally on the degree of uncertainty, which is measured by the covariance matrix). However, to describe an uncertain trajectory, one needs to include a more advanced concept of probability: a stochastic process (Grimmet, 2001), in which a quantity (e.g., the aircraft position) is described by a random variable that changes with time. This allows the consideration of not only the uncertainty at a given time but also how uncertainty propagates into the future. Well-known models using differential equations describe the trajectory of an aircraft. However, if the initial conditions of these equations are uncertain or/and the differential equations contain uncertain terms, the equations become stochastic differential equations (Oksendal, 2003), the solution being an uncertain trajectory (with uncertainty changing with time).

One of the first applications of this probabilistic framework in ATM was in the field of navigation systems, which estimate the present position of the aircraft using sensor information. Since sensor measurements are corrupted by noise, the resulting navigational solution is uncertain (Grewal et al., 2000). To visualise this uncertainty, the position of an aircraft can be represented as an uncertainty ellipsoid, centred around the computed position, whose size depends on the uncertainty; the more uncertain the navigational solution, the larger the ellipsoid. The real position of the aircraft can lie anywhere within this ellipsoid, which will change size with
time as the navigational precision changes. When the uncertainty region moves through space, one actually obtains a *tube* of uncertainty; with very high probability an aircraft will be contained in this tube along its real trajectory.

In general, the methods used to study trajectory uncertainty and uncertainty propagation can be classified into two main groups.

- **Monte Carlo methods** (see Hastings, 1970.) These are computational methods that rely on repeated random sampling to compute their results. That is, one randomly selects a value for all the uncertain elements that affect the trajectory, and then computes it in a deterministic way. This is repeated as many times as necessary until the resulting trajectories are a representative sample of the uncertain trajectory itself. The main advantage of these methods is that they can be used with all types of uncertainty and do not require any complicated computation beyond solving the equations of flight mechanics many times. However, this is a very expensive method in computational terms (suffering from ‘the curse of dimensionality’; if there are many sources of uncertainty one needs many sample points and therefore it is not practicable) and requires random sampling of the sources of uncertainty, which is not always easy. Monte Carlo variants, such as the sequential Monte Carlo (Liu and Chen, 1998), use more advanced ideas (carefully selecting the samples of the sources) to improve the convergence and reduce the required number of sample points.

- **Non-Monte Carlo methods** To avoid the computational problem of Monte Carlo methods, other techniques have been proposed to study uncertainty propagation in dynamical systems. Halder and Bhattacharya (2011) classify those methods into two categories:
  - *parametric* in which one evolves in time the statistical moments, such as mean and covariance, but the probability distribution remains unknown. An example is the generalised polynomial chaos method (Prabhakar et al., 2010).
  - *non-parametric* in which the probability density function itself is evolved. An example is the method presented in Halder and Bhattacharya 2011, which finds the probability density function by solving a stochastic Liouville equation.

These methods help to analyse and describe uncertain trajectories and uncertainty propagation along the trajectory. However, there are also methods that deal with the problem of uncertainty mitigation, i.e., reduction in the uncertainty of the trajectory. This is usually only possible if new information is obtained, for example by sensors or communication with other agents; this new information carries its own uncertainty, and the question arises of how to update the uncertain trajectory to incorporate the new information, in such a way that the uncertainty in the obtained trajectory is reduced as much as possible. The algorithms that deal with this issue are called ‘filtering’ algorithms, and the most widely known is the
Kalman filter (Anderson and Moore, 1979). Although the Kalman filter was initially developed to solve a navigation system problem (how to fuse information from inertial-type sensors with information from other sensors to obtain the most accurate possible navigational solution), it can be applied in any situation in which one has an estimate of the uncertainty of the trajectory and new data (also with a degree of uncertainty) is obtained.

While stochastic differential equations have been used for decades to assess the accuracy of navigation systems, the application of the method to trajectories is rather recent and the literature dealing directly with trajectory uncertainty and uncertainty propagation is scarce. Polynomial chaos is used to study both uncertainty propagation (Prabhakar et al., 2010) and trajectory estimation (Dutta and Bhattacharya, 2010) for hypersonic flight dynamics with uncertain initial data. Vazquez and Rivas (2013) have developed methods to study the impact of initial mass uncertainty on fuel consumption during the cruise flight phase. Zheng and Zhao (2011) developed a statistical model of wind uncertainties and applied it to stochastic trajectory prediction in the case of straight, level aircraft flight trajectories. Weather uncertainty was considered by Matthews et al. (2009), Yen et al. (2003), and Nilim et al. (2003). Navigational errors and flight management system errors were studied by Grewal et al. (2000) and Kim et al. (2009), respectively. Crisostomi et al. (2009) combine Monte Carlo methods with deterministic worst-case methods to analyse trajectory prediction for ATC. Some of these methods can also be useful in trajectory optimisation problems; for instance, Fisher and Bhattacharya (2010) use a polynomial chaos approach to solve optimal trajectory generation problems with probabilistic uncertainty in system parameters, and Li et al. (2014) apply the same method to optimise trajectories to minimise time in climb for supersonic aircraft, while considering the uncertainties in the aerodynamic data.

4.3 Traffic Uncertainty

The flight and network scales take a microscopic and a macroscopic view of the ATM system, respectively. While the points of view of the flight (microscopic view) and the air transport network (macroscopic view) are of great importance, there is another key, intermediate scale (the mesoscopic scale), which appears at the level of the terminal manoeuvring area or an air traffic sector. In this scale, microscopic uncertain objects (flights and their trajectories) interact with management procedures and separation rules, and with the uncertain atmospheric conditions. This generates a dynamic, rapidly changing environment where ATC has to take decisions (such as rerouting or holding) trying to fulfil the conflicting objectives of improving performance (increasing capacity, following user-preferred trajectories, minimising delays) and guaranteeing safety (avoiding incidents and especially accidents). Problems such as conflict detection and resolution and the analysis of flow management are framed within this mesoscopic scale.
When studying a scenario at the mesoscopic scale, one finds many sources of uncertainty, which affect the scenario:

- uncertain trajectories, which need to be predicted by air traffic controllers with a large degree of unknown data (making the predictions even more uncertain);
- operational uncertainty, which is partly due to human factors, but also due to the interaction between sometimes ambiguous or conflicting management rules;
- discrepancies between the weather forecast and the real weather, which are particularly significant in the case of adverse weather.

The two main types of scenarios that appear at this scale are the terminal manoeuvring area and the air traffic sector:

- The terminal manoeuvring area is often the most complex scenario, as aircraft converging to the runway (incoming flights) interact with each other and with outgoing flights, though there are many sensors and flight aids which help to decrease the uncertainty to a more manageable level.
  The main concern is safety, followed by delays and capacity.
- An air traffic sector typically presents a lesser level of complexity, as traffic in general is not convergent, but the level of uncertainty is higher due to decreased sensor coverage and the lack of accuracy in the weather forecast. Safety has to be enforced while at the same time maximising capacity and minimising deviations from user-preferred trajectories.

When studying air traffic under uncertain conditions, one important and very difficult problem is flow management. This was first addressed by Odoni (1987), who considered that congestion could arise at any point in the trajectory: at the airport of origin or destination or en route (at a waypoint or sector). It was understood from the beginning that the flow management problem is stochastic in nature, as it involves, among other things, a prediction (forecast) of weather; however, to simplify the problem most authors have treated aircraft trajectories deterministically. The first approach to solving the problem was to hold aircraft on the ground to reduce en-route delay. Vranas et al. (1994) considered a multi-airport, ground-holding programme, taking into account the propagation of delay, and solved the problem using a heuristic approach. The inclusion of both ground-holding and en-route decisions was first addressed by Helme (1992), but deterministic models are still used.

Alonso et al. (2000) were the first to include uncertain factors in both the airport and airspace, and Mukherjee and Hansen (2009) made use of more realistic weather scenarios where decisions are dynamically taken and reviewed as a function of updated weather forecasts. Another approach (Nilim et al., 2004) models the problem as a Markov decision process with the weather evolving as a Markov chain.
Chang et al. (2010) consider the use of ground delay, cancellation, cruise speed modification, air holding and diversion as recourse actions when determining how aircraft are sent towards a sector under the uncertainty of weather; the problem is solved using rolling horizon methods. Kim et al. (2009) derived service time distributions for different flight phases and assessed traffic flow efficiency using queuing network models. Clarke et al. (2009) developed a methodology to study the capacity of a volume of airspace in the presence of weather uncertainty (stochastic capacity) and formulated a stochastic dynamic programming algorithm for traffic flow management; they presented a stochastic routing algorithm that provides guidance for routing aircraft in the presence of the uncertainties of adverse weather. Knorr and Walter (2011) analyse the impact of trajectory uncertainty on sector complexity and workload, considering the following sources of uncertainty: trajectory data quality, model quality, and operational procedures. Sengupta et al. (2014) considered the problem of ATFM under airspace capacity uncertainty arising from weather or environmental effects, and used stochastic programming to develop risk-hedge decisions resulting in the least delay at a specified level of acceptable variance.

Another important problem in any air traffic scenario is the existence of aircraft conflicts and loss of separation, which requires the use of conflict detection and resolution algorithms (Valenzuela and Rivas, 2011; Ruiz et al., 2013). These algorithms perform three different processes:

• predicting that a conflict is going to occur in the future;
• communicating the detected conflict to a human operator;
• assisting in the resolution of the conflict.

These three processes can be affected in different ways by uncertainties:

• in the estimation of the present state of the aircraft (observable states, sensor errors);
• in the prediction of their future states (trajectory prediction);
• in the communication to the human operator (system malfunction, operator misinterpretation);
• in the resolution process (stochastic resolution, negotiation process).

Kuchar and Yang (2000) identified two fundamental methods to handle uncertainty in trajectory prediction in conflict detection and resolution.

• **Worst-case method** It is assumed that an aircraft will perform a manoeuvre selected from a set of possible operations. If any one of these manoeuvres could cause a conflict, then a conflict is predicted. This method is conservative in the sense that conflict alerts are triggered whenever there is any possibility of a conflict within the definition of the worst-case trajectory model. Durand and Alliot (1997) used this approach to consider the
effect of ground and vertical speeds uncertainties, while Tomlin et al. (1998) used it to determine the minimal unsafe operating region for each aircraft.

- **Probabilistic method** The uncertainties are modelled to describe potential variations in the future trajectory of the aircraft. This is usually done in one of two ways. The first approach consists in adding a position error to a nominal trajectory, from which the conflict probability can be derived (Isaacson and Erzberger, 1997; Vela et al., 2009). The second approach (Prandini et al., 2000) is to develop a complete set of possible future trajectories, each weighted by a probability of it occurring; the trajectories are then projected into the future to determine the probability of conflict.

In conflict resolution, uncertainties can be found if stochastic processes are used to obtain the resolution manoeuvres, as is done by Durand and Alliot (1997) who use genetic algorithms. Another example can be found in distributed algorithms, which usually require a negotiation process where the aircraft agree a coordinated solution. This negotiation process is commonly affected by uncertainties, such as delays in the communications, lack of knowledge of the algorithms used by the other aircraft, etc. However, these uncertainties can be mitigated by replacing the negotiating process with mathematical criteria that the resolution manoeuvres must meet, such as those proposed by Narkawicz and Muñoz (2011). An example of a conflict solver that includes uncertainty can be found in the Complete Air Traffic Simulator (CATS), which is a general-purpose simulator for en-route traffic that incorporates uncertainty in ground speed and in climbing and descending rates, as described in Alliot et al. (1997).

Another outstanding question for both traffic flow management and conflict resolution is how much traffic can be safely accommodated. To perform safety risk assessments, Monte Carlo simulations can be used. Blom et al. (2006) explain the key issues in these types of simulations: first, one has to develop an appropriate simulation model and a sound way to speed up the simulation. One also has to validate the simulation model versus the real operation, and the simulation-supported approach has to be embedded within the safety risk assessment of the total operation. Given that safety-critical events are rare occurrences, advanced techniques have to be used. Blom and Bakker (2011) performed an accident risk assessment for a self-separation concept of operations, using agent-based modelling and rare-event Monte Carlo simulations. Their agent-based modelling was based on a stochastic hybrid model, dynamically coloured Petri nets (Everdij et al., 2006a) (see Chapter 6). In Andrews et al. (2006) another approach is presented; they performed a fault analysis of a highly automated separation assurance system. The growth and decay of risk during a failure of the system was evaluated using fault tree methods that integrate risk over time.

One of the most interesting challenges at the traffic scale is to increase the robustness of air traffic under highly uncertain conditions (such as adverse weather
or a congested air transport network). Adverse weather represents a challenging limiting factor of air traffic capacity for ATM. In practice, it requires traffic managers to reroute all flights affected. Similarly, a congested air transport network (a manifestation of the network scale affecting the traffic scale) is another limiting factor. Heidt and Gluchshenko (2012) address the problem of uncertainty and robustness in the context of ATM planning. In this work, different time horizons are considered (strategic, pre-tactical and tactical planning), taking into account the impact level of uncertainty, because a source of uncertainty that plays a crucial role for one planning horizon can be irrelevant for others. Hauf et al. (2013) address the need to model the impact of adverse weather on ATM performance and design guidance algorithms to navigate through or around thunderstorms. Because the problems of conflict resolution and weather avoidance can arise simultaneously, some authors develop integrated algorithms to solve them. For instance, Lauderdale and Erzberger (2014) propose a unified solution in which three algorithms are applied sequentially to solve the problems of separation conflicts, arrival sequencing, and weather-cell avoidance in en-route airspace; safety buffers are included to reduce the effect of aircraft performance uncertainties and efficiently handle complex weather environments.

4.4 Network Uncertainty

To analyse uncertainty at the macroscale of the air transport network (be it at a regional, national, trans-national or global level), it is best to abstract and integrate the various complex and heterogeneous ATM elements in a way that allows the assessment of uncertainty and other properties of interest without needing to include too much detail (which would be impractical or even impossible if dealing with the whole ATM system). A promising framework that allows such an abstraction is complex network theory (CNT), which is described in Chapter 2 and whose specific application to ATM is detailed in Chapter 3.

CNT is particularly well suited to study the air transport network, especially when subject to uncertainty. For instance, even under normal operating conditions, the delay due to trajectory uncertainty of several flights can accumulate, and, if above a certain threshold, cause congestion, which will produce a sudden change in the network topology, as some nodes may become unreachable. In the same fashion, weather uncertainty can produce sudden, almost-instantaneous changes in the network as the ATM system tries to accommodate such occurrences. These changes may result in unexpected configurations due to complex interactions between many agents and the rules of operation. Thus, CNT allows the study of the impact of uncertainty on the overall performance of the system: capacity, predictability, stability, robustness, and even resilience.

Chapter 3 presents numerous examples of how CNT can be applied to diverse ATM problems. However, there seems to be a lack of studies in which the uncertainty of the system is explicitly taken into account. In CNT, uncertainty has been
typically understood as the presence of errors in measurements. For instance, when analysing the protein–protein interaction networks, nodes represent the protein of a cell and a link between two of them is created when a correlation in their expression level exists. When measuring such correlations, there is always an error: thus, non-existent correlation may be detected, and existing correlation may be neglected (von Mering et al., 2002). To solve this problem, several algorithms have been developed, which try to assess the ‘validity’ of links in the network (Clauset et al., 2008; Zhou et al., 2009; Guimerà and Sales-Pardo, 2009). One interesting idea is the concept of uncertain links (that is, links that may or may not exist); however, it has not yet been exhaustively studied. In Ahnert et al. (2007), the concept of an uncertain complex network has been realised by using weighted networks. Specifically, the authors suppose that the weight associated with each link (between 0 and 1) is equivalent to the probability of that link existing. Then, complex weighted networks can be represented as an ensemble of random networks, whose analysis can then be performed by using standard probabilistic metrics. In a similar way, Zanin (2011) proposed the interpretation of link weights as a measure of the uncertainty of such links, i.e., of the probability of their occurrence. By transforming the adjacency matrix of the network, it was possible to associate a probability of existence to each path inside the graph.

Other examples of the use of CNT to study ATM problems are included in Lillo et al. (2011), who have investigated how airport network topology is related to the statistical properties of flight delays and how it affects the propagation of flight delays from one airport to another; and Sanchez et al. (2011), who studied the link between specific prioritisation rules applied to flights in cases of capacity shortfalls at nodes and the network behaviour and stability, both locally and at network-wide scales.

A different approach is to consider that the structure of the network is deterministic, but that, from time to time, a link (or a node) suffers a failure or an attack. A failure is random, but an attack requires the participation of an external agent, which tries to weaken the network by suppressing the most vulnerable (or the most important) element. It has been shown that the topology of the network defines its vulnerability (Albert et al., 2000; Holme et al., 2002). Following this line of research, Cardillo et al. (2013b) analysed the resilience of the European air transport network against random flight failures. A multiplex network formalism (with the set of flights of each airline as an interdependent network) was used to solve the passenger’s rescheduling problem. A comparison with the ‘singleplex’ approach shows that the multiplexity strongly affects the robustness of the European air network.

The sparsity of results in this area of research is due to the concept of uncertain complex networks being not yet mature. Thus, there is a need to develop complex network models that incorporate a degree of randomness that can properly capture the uncertain dynamics of the ATM network. While weighted complex networks constitute a promising starting point to developing such a theory, the main problem is to develop ways to incorporate microscale and mesoscale uncertainties in the
analysis of the structure and behaviour of the air transport network. Another interesting and potentially useful concept is that of evolving networks. For instance, it would be of great interest to develop dynamic network models that describe how the air transport network and, in particular, the connection between nodes, changes when subjected to adverse weather conditions. Agent-based modelling or stochastic networks stand as promising tools to model the modifications under such conditions. These models might help to identify the ‘best’ rules (in some optimal sense) that should be followed to minimise the impact of such changes on the network topology.

4.5 Conclusions

This chapter has provided an overview of how uncertainty affects the ATM system at different scales. Given that uncertainty is ubiquitous in ATM, if the objectives of increasing air capacity and safety are to be met, there is a clear need to reduce or at least manage uncertainty. Another of the greatest challenges faced by the future ATM system is the integration of new airspace users and the increase in delegating capacity and safety critical traffic management functions to automated systems. The accommodation of these new airspace users, which will have to coexist with conventional users, a widely reorganised airspace and the increased level of automation, will require a paradigm shift with regard to ATM. It will be necessary to develop essential trajectory management functions to efficiently manage heterogeneous traffic with increasing use of autonomous ATM systems, in uncertain conditions. The development of such systems has to focus on how to deal with sources of uncertainty and their propagation through the system.

To address these challenges, uncertainty must be directly taken into account. While new concepts are emerging, such as uncertain 4D trajectories, uncertain air traffic networks, or stochastic conflict resolution, and there are pioneering studies considering at least some degree of uncertainty in the system, there is still a lack of a common language or framework to further collaboration between researchers to deal with the challenges posed by uncertainty. The main objective of this book has been to provide such a framework and review past results.

Thus, there are still many open problems in ATM that require for their solution the use of uncertainty concepts and tools. A small sample of these problems is given according to the different scales of ATM that are introduced in Section 4.1.3.

At the flight scale, one of the main challenges is to define and compute the trajectory uncertainty with precision. For that, one must analyse the propagation of the different sources of uncertainty along the trajectory, which requires finding the right stochastic models to allow the description of uncertain trajectories with sufficient accuracy while being sufficiently tractable. The first element to consider is what kind of aircraft motion model is to be used for trajectory prediction (six degrees of freedom versus three degrees of freedom). This has a considerable impact on the inputs required for computing a trajectory and, therefore, on what
sources of uncertainties have to be considered. Although each approach has its own characteristics, there are some inputs (sources of uncertainty) that are common to all of them: the initial conditions, the aircraft performance model, the weather model and the Earth model. All trajectory predictors, regardless of the implementation used for their development, require these four inputs, whose influence on trajectory uncertainty is unclear and must be determined. Another important outcome of this analysis is to determine whether the different uncertainty sources amplify or decay along the trajectory. Next, a validation process for such stochastic prediction tools is required, to prove that the simulated data is appropriate for modelling parts of the real ATM process. This can be achieved by comparison either with real flight data or with other established prediction tools. When it comes to 4D trajectories as envisaged by SESAR, a new challenge arises: the lack of real input data to compare with. Apart from a very few flight trials that performed the initial 4D tracking operation, there is no flight data from real aircraft, doing real 4D trajectory tracking. It is not representative to compare 4D trajectory prediction with real flight data without an active 4D tracking algorithm.

For analysis at the air traffic scale, at present, uncertainty is mostly handled with deterministic algorithms (traffic flow management, conflict detection, conflict resolution, etc.) where the level of uncertainty is just guessed and used as a buffer (by setting a minimum separation distance). If capacity is to be maximised and user-preferred trajectories implemented, while still enforcing safety, it is necessary to use stochastic algorithms that include uncertainty models in their formulation. These algorithms should take into account all different sources of uncertainty present in the problem and compute a solution. The solution would be stochastic in nature, so that, for instance, the probability of having a capacity as large as possible is maximised (rather than directly maximising a value of capacity).

While it is evident that uncertainty has a significant impact on safety assurance levels, there have been few studies directly relating the effect of uncertainty (and its propagation) on levels of safety. If safety is to be greatly increased, it is necessary to improve the understanding of this relationship. In particular, it would be of great interest to investigate what types of uncertainty are most significant for the safety of air traffic, that is, to perform a sensitivity analysis with respect to uncertainties. To carry out this analysis one could think of safety key performance indicators (KPIs) as functions of the degree of uncertainty; this functional dependence could be obtained through a simulation study (using, for instance, rare-event Monte Carlo techniques) so that one would be able to quantify the variation of the safety KPIs when the levels of uncertainty change. Such a study would allow one to determine which uncertainties are most detrimental to safety and are worth mitigating (if at all possible) to improve safety levels.

At the network scale, one of the main difficulties faced by researchers beginning to study network uncertainty is that the field is relatively new, at least compared with the flight and traffic scales. While some pioneering work was presented in Section 4.4, there is a need for further work in the area. This represents an opportunity for new researchers to shape this novel field. More specifically, some
particular challenges are the study of how trajectory uncertainty affects the network, how the propagation of delays affects the network capacity and eventually leads to the congestion of the system (a phase change), and also how to modify operational procedures to minimise the probability of this happening.

Finally, a source of uncertainty that affects all scales is the modelling of weather. Given that deterministic weather forecasts are not accurate, probabilistic models are necessary. One current trend is to use ensemble prediction forecasts, which attempt to characterise and quantify the inherent prediction uncertainty based on ensemble modelling; see, for instance, Hacker et al., 2003. Ensemble forecasting is a prediction technique that aims to generate a representative sample of the possible future states of the atmosphere. An ensemble forecast is a collection of typically 10 to 50 weather forecasts which may be obtained in different ways based on time-lagged, multi-model, and/or multi-initial conditions approaches (Arribas et al., 2005; Lu et al., 2007). Different national meteorological offices provide ensemble prediction forecasts, which can be used to model weather uncertainty, be it at the flight, traffic, or network scale.
This page intentionally left blank
Chapter 5
Resilience

Dr Henk A.P. Blom and Soufiane Bouarfa

Thanks to influential work by Hollnagel and other researchers (Hollnagel et al., 2006, 2009), the value of resilience in air transport has been well recognised in behavioural sciences. The objective of this chapter is to show that air transport can benefit significantly from the study of resilience from the perspective of complexity science. This allows knowledge from behavioural sciences to be combined with the systematic modelling and analysis approach of complexity science.

Civil air transport is an example of a large and complex socio-technical system. It comprises interactions between different types of entities, including technical systems, operational stakeholders, regulators, and consumers (DeLaurentis and Ayyalasomayajula, 2009). Technology plays a central role, as does the social context within which the various parties operate. This socio-technical system copes with many differing internal and external disruptions that test its resilience. These events may interact with each other, potentially creating a cascade of other events that may span over different spatial and temporal scales. In current air transport, disruptions are managed by operators at airlines, airports, and air traffic control (ATC) centres, and may impact on the overall performance of the socio-technical system, for example, some flights are rerouted, some aircraft or crew are exchanged, and some passengers are rebooked. Managing disruptions involves trade-offs which are created by the complexities inherent to both the processes managed and the finite resources of operational systems (Hollnagel, 2009). For instance, in the case of congested airspace, air traffic controllers might ask airlines to reroute their flights. In such a situation, improving safety comes at the cost of economy. Potentially, there are conflicting goals leading to dilemmas and bottlenecks that must be resolved. Nevertheless most problems are adequately solved, and most of these events pass without substantial inconvenience to passengers.

In some cases, however, the resilience of the air transport system falls short, resulting in significant flight delays. A typical example is bad weather, which may jeopardise the normal operation of an airport or a sector and induces ‘ripple’ effects (propagation) throughout the network. Another example is that of a malfunctioning aircraft being stuck with its passengers at a distant airport, as a result of which all passengers are delayed many hours.

In addition to regular cases with limited consequences, there are rare cases with very severe consequences. These severe consequences can be divided into two categories: network-wide consequences dramatically affecting the performance of the system; and catastrophic accidents involving one or more aircraft. The former happens in the case of external events to which the air transport network is
vulnerable (see Section 3.5), such as the outbreak of a viral disease causing passengers and airlines to change their travel behaviour (for example, SARS in 2003 and Ebola in 2014) or volcanic ash impacting air travel in a large area (for example, the Icelandic volcano in 2010). Cases of the latter include fatal runway incursions (for example, the Linate runway collision in 2001), fatal mid-air collisions (for example, the Überlingen mid-air event in 2002), and the loss of control of an aircraft flying through a hazardous weather system (for example, the Air France crash in the Atlantic Ocean in 2009). Some external events belong to both categories, for example, the 9/11 terrorist action in 2001 led to fatal accidents and caused the shutdown of air travel in a large area.

The examples above show a wide variety of significant events with major consequences. However, thanks to the resilience of the air transport system, many significant events have negligible consequences. To increase the resilience of the air transport system, there is a need to identify, understand, and model the interdependencies of the air transport system and analyse its response to the large variety of possible disturbances. This chapter aims to show that a complexity science perspective can be a valuable asset in meeting this need. In particular, the chapter aims to answer the following questions: What is resilience and how is it measured? Why use complexity science to model and analyse resilience? Which complexity science approaches can be used? The chapter also demonstrates the benefits of applying complexity science and behavioural science to an airline problem. This specific application concerns airline operations control (AOC), whose core functionality is to provide resilience to the large variety of disruptions that happen on the day of operation.

This chapter is organised as follows. Section 5.1 addresses resilience capacities. Section 5.2 examines various resilience metrics from the literature. Section 5.3 introduces complexity science approaches for studying resilience. Section 5.4 provides a convincing resilience application of using complexity science in air transport. Section 5.5 provides conclusions.

5.1 Resilience Capacities

Resilience comes from the Latin word ‘resilio’, meaning ‘to jump back’, and is increasingly used in various disciplines to denote the ability to absorb strain and bounce back from unfavourable events. The term was initially used in the field of mechanics as ‘the ability of a metal to absorb energy when elastically deformed and then to release it upon unloading’ (Hoffman, 1948). Holling (1973) extended the resilience concept to ecological systems as the ‘persistence of systems and of their ability to absorb change and disturbance and still maintain the same relationships between populations or state variables’. Since then, various other extensions of resilience have been introduced in other domains, such as economics, organisational science and safety science.

Recently, Francis and Bekera (2014) conducted a systematic review of the resilience developments across multiple domains, and identified the following three
resilience capacities: (i) absorptive capacity, (ii) adaptive capacity, and (iii) restorative capacity. Absorptive capacity is the degree to which a system can absorb the impacts of system disruptions and minimise consequences with little effort (Vugrin et al., 2010). The practice of incorporating adequate buffer capacity in anticipation of increased stress on the system is an absorptive endowment. It is considered to be a proactive measure to absorb potential shocks. Adaptive capacity is the ability of a system to adjust to undesirable situations by undergoing some internal changes. Adaptive capacity is distinguished from absorptive capacity in that an adaptive system can change its response. A system’s adaptive capacity includes the ability to forecast adverse events, recognise threats, and reorganise after the occurrence of an adverse event. Finally, restorative capacity is the ability to recover or bounce back from disruptive events and return to normal or improved operations.

Table 5.1 shows what the three resilience capacities mean for resilience-related concepts like robustness and dependability. Robustness is defined as the ability of elements, systems, and other units of analysis to withstand a given level of stress or demand without suffering degradation or loss of function (MCEER, 2006). This definition is consistent with the absorptive capacity described by Francis and Bekera (2014). Hence, a socio-technical system that has absorptive capacity only is robust. ‘System dependability’ is the collective term used in systems engineering to describe a system’s availability performance and its influencing factors: reliability, performance, maintainability performance and maintenance support performance (IEC, 1990). Thus, a dependable system has both absorptive and restorative capacities. In comparison to dependability, resilience is an endowed or enriched property of a system that is capable of effectively combating (absorbing, adapting to, and rapidly recovering from) potentially disruptive events.

Robustness and dependability are system properties that are well addressed throughout systems engineering. For air transport, this means that the key challenges in analysing resilience are twofold: 1) to address the absorptive and restorative capacities for a complex socio-technical system rather than for a complex technical system, and 2) to improve the adaptive capacities of both the absorption of disturbances and the recovery from a system performance degradation due to disruptions.

Placing emphasis on improving the adaptive capacity to absorb disturbances concurs with Hollnagel et al. (2009), who define a resilient system as one with ‘the

<table>
<thead>
<tr>
<th>Related system properties</th>
<th>Resilience capacities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absorptive</td>
</tr>
<tr>
<td>Robustness</td>
<td>+</td>
</tr>
<tr>
<td>Dependability</td>
<td>+</td>
</tr>
<tr>
<td>Resilience</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 5.1 Resilience capacities in relation to robustness and dependability
intrinsic ability to adjust its functioning prior to, during, or following changes and disturbances, and thereby sustain required operations under both expected and unexpected conditions’. In the safety domain, Hollnagel (2014) explains that this definition of resilience engineering reveals a need to study ‘what may go right’, rather than the traditional approach of studying only ‘what may go wrong’. The traditional and novel approaches are referred to as Safety-I and Safety-II respectively.

5.2 Resilience Metrics

This section examines resilience metrics from the literature, covering different domains including ecosystems, critical infrastructure systems, networks, organisations, information systems, psychology, and transport systems.

5.2.1 Ecosystems

For ecosystems, Gunderson et al. (2002) distinguished between two resilience measures: engineering resilience and ecological resilience. Engineering resilience describes the ability to return to the steady state following a perturbation (Pimm, 1984; Varian, 1992; Tilman, 1996; Scheffer, 2009), that is, it implies only one stable state and global equilibrium. Ecological resilience relates to the amount of disturbance that a system can absorb before it changes state (Holling, 1996; Gunderson et al., 2002; Scheffer, 2009), that is, it emphasises conditions far from any stable steady state, where instabilities can ‘flip’ a system into another regime of behaviour (Gunderson et al., 2002). So, ecological resilience is measured by the magnitude of disturbance that can be absorbed before the system redefines its structure by changing the variables and processes that control behaviour (Gunderson et al., 2002). For engineering resilience, the only possible measures for resilience are near-equilibrium ones, such as a characteristic return time to a global equilibrium following a disruption, or the time difference between the moments of disruption and of full recovery.

5.2.2 Critical Infrastructure Systems

For earthquake engineering, Tierney and Bruneau (2007) suggested measuring resilience by considering the functionality of an infrastructure system after a disaster has occurred, and also the time it takes for a system to return to its pre-disaster level. Their suggestion was based on the observation that resilient systems reduce the probabilities of failure, the consequences of failure, and the time for recovery. This concept is illustrated by the ‘resilience triangle’ in Figure 5.1, which represents the performance degradation due to damage caused by earthquake disruption(s), as well as the pattern of restoration and recovery over time.

The higher the resilience of a system, the smaller the size (depth and duration) of the triangle. Bruneau et al. (2003) expressed resilience as follows:
Renschler et al. (2010) proposed a framework to measure earthquake resilience at the community scale, integrating several dimensions such as population, environment, physical infrastructure, and economic development into one resilience index.

Li and Lence (2007) defined resilience $R_e(t_f,t_r)$ as the conditional probability that given full performance impact at time $t_f$, the system is fully recovered at time $t_r$, i.e.

$$R_e(t_f, t_r) = P\left\{ F(t_r) \geq F_0 \right\} \left\{ F(t_f) < F_0 \right\}$$

where $F(t_f)$ and $F(t_r)$ are the performance levels at $t_f$ and $t_r$, respectively, and $F_0$ is the original stable system performance level (100% level in Figure 5.1). Attoh-Okine et al. (2009) extended the conditional probability approach of Li and Lence (2007) with a ‘belief’ function to capture incomplete data in urban infrastructure systems.

Francis and Bekera (2014) have proposed quantifying resilience $R_e$ as follows:

$$R_e = S_p \frac{F(t_r)}{F_0} \frac{F(t_f)}{F_0}$$

where $F_0$ is the original stable system performance level (100% level in Figure 5.1); $F(t_f)$ is the post-disruption performance level (at point B in Figure 5.1); $F(t_r)$ is the performance at a new stable level after recovery efforts have been exhausted (at point D in Figure 5.1); and $S_p$ is the speed recovery factor (slope of BD).
Ayyub (2014) proposed expressing the resilience $R_t$ metric as follows:

$$R_t = \frac{t_d + \alpha (t_f - t_d) + \beta (t_r - t_f)}{t_r}$$

where $\alpha$ and $\beta$ are the ratios of mean performance levels during periods $(t_d, t_f)$ and $(t_f, t_r)$ respectively versus the pre-disruption performance level.

Musman and Agbolosu-Amison (2014) proposed capturing resilience in terms of mission risk. According to their definition, resilience can be computed as being: (1) a utility-based performance metric that indicates how well the system responds in the face of one or more incidents (where incidents are assumed to have occurred); (2) a probability that some events might occur to bring the system to some specified unacceptable level of performance; or (3) a risk estimate that combines the probability of incidents with a utility-based measure of the performance changes that result when incidents occur.

5.2.3 Networks

In network theory, Najjar and Gaudiot (1990) proposed measures for network resilience $R_N(p)$ and relative network resilience $R_{NR}(p)$, where $R_N(p)$ is defined as the number of node failures a network can sustain while remaining connected with a probability $(1-p)$, and $R_{NR}(p)$ is defined as the ratio of network resilience $R_N(p)$ to the number $N$ of nodes in the network.

Garbin and Shortle (2007) generalised this to a network resilience metric in the form of actual network performance (or the percentage of the normal network performance) as a function of the network damage (see Figure 5.2). Examples of parameters that characterise networks are demand, topology, capacity, and routing. Garbin and Shortle (2007) also proposed using the area under the curve in Figure 5.2 as an index for the resilience of a network.

![Figure 5.2 Examples of network resilience curves, showing network performance percentage as a function of network damage percentage](image-url)

Adapted from Garbin and Shortle (2007).
Rosenkrantz et al. (2009) proposed metrics to quantify the resilience of service-oriented networks under node and edge failures. The metrics are based on the topological structure of the network and the manner in which services are distributed over the network. They made a distinction between network edge resilience and network node resilience. A network is said to be \( k \)-edge failure resilient if no matter which subset of \( k \) or fewer edges fails, each resulting sub-network is self-sufficient. A network is said to be \( k \)-node failure resilient if no matter which subset of \( k \) or fewer nodes fails, each resulting sub-network is self-sufficient. They presented algorithms to determine the maximum number of node and edge failures that can be tolerated by a given service-oriented network, and to optimally allocate services over a given network so that the resulting service-oriented network can tolerate single node or edge failures.

Henry and Ramirez-Marquez (2012) expressed resilience as the ratio of recovery to loss suffered by the system. This means that if the recovery is equal to the loss, then the system is fully resilient, and if there is no recovery, then no resilience is exhibited. They acknowledged that quantifying resilience requires identification of a quantifiable and time-dependent system-level delivery function, also called a ‘figure of merit’ (such as delay, connectivity, flow, etc.). In systems where multiple figures of merit are considered, an event could be disruptive with respect to one figure of merit but not disruptive with respect to another figure of merit. Therefore for a holistic analysis of system resilience, the system must be analysed with respect to all figures of merit that are relevant and important (Henry and Ramirez-Marquez, 2012).

5.2.4 Organisations and Information Systems

Dalziell and McManus (2004) suggested measuring resilience by assessing the total impact on key performance indicators (KPIs) between the time of disruption and the recovery time, where these indicators are real-valued measures at a certain moment in time for the corresponding key performance areas. The variation of a specific indicator is measured and plotted against time from the start of the disruption \( t_d \) until full recovery \( t_r \). The resilience then represents a weighted sum of the areas under the indicator curves.

Zobel and Khansa (2012) introduced a general approach for characterising cyber infrastructure resilience in the face of multiple malicious cyber attacks. Their proposed technique accounts for the amount of loss incurred by an information system in the face of multiple cyber attacks, and it captures the strength and timing of these attacks.

5.2.5 Psychology

In psychology, various psychometric scales have been developed to assess the resilience of individuals. For instance, Wagnild and Young (1993) developed a Likert scale to identify the degree of individual resilience, considering it a positive personality characteristic that enhances individual adaptation. The scale consists
of 25 items, each rated with a seven-point agreement scale. Smith et al. (2008) proposed a ‘brief resilience scale’ to assess the ability to bounce back or recover from stress.

Other Likert scales include the Baruth protective factors inventory, the Connor-Davidson scale, and the resilience scale for adults (see Ahern et al., 2006, for a detailed review).

5.2.6 Transport Systems

Chen and Miller-Hooks (2012) defined a resilience indicator that considers the ability of the freight transport network to cope with the negative consequences of disruptions. The indicator explicitly accounts for the network topology, operational attributes, and the impact of potential recovery activities. Such activities might be undertaken in the immediate aftermath of the disruption to meet target operational service levels while adhering to a fixed budget.

Omer et al. (2013) identified three resilience metrics to measure the impact of disruptions on the performance of a road-based transport system. The three identified metrics were the travel time resilience, environmental resilience, and cost resilience. The resilience values were measured by introducing hypothetical disruptions to a network model of a regional transport network.

Gluchshenko and Foerster (2013) proposed a qualitative measure for resilience in air transport based on recovery time. They introduced three degrees of resilience, namely: (i) high resilience, when the time of deviation is considerably longer than recovery time; (ii) medium resilience, when the time of deviation and recovery time are approximately equal; and (iii) low resilience, when the time of deviation is considerably shorter than the recovery time.

Hughes and Healy (2014) proposed a qualitative framework to measure the resilience of road and rail transport systems, through dedicated measurement categories for technical and organisational dimensions. The framework involves an initial determination of the context of the resilience assessment, followed by a detailed assessment of resilience measures, which combine to generate a resilience score ranging from 4 (very high resilience) to 1 (low resilience).

Janic (2015) provides an alternative resilience indicator for air transport networks analogous to the indicator proposed by Chen and Miller-Hooks (2012) for intermodal freight transport. The indicator considers the network’s inherent properties and the set of actions for mitigating costs and maintaining the required safety level. Because mitigating actions include delaying, rerouting and/or cancelling flights, Janic (2015) defines this indicator as the ratio of the on-time and delayed flights achieved to the total number of scheduled flights during a specific time period. Janic (2015) also proposed the measurement of the resilience of an air transport network by estimating the sum of the weighted resilience of each individual airport.

Resilience can also be expressed in terms of mission risk (Musman and Agbolosu-Amison, 2014). In air transport, one well-studied mission risk metric is the reach probability for an aircraft trajectory (Prandini and Hu, 2006, 2008; Blom et al.,
2007, 2009). Let $P_{\text{Reach}}^{i,j}(h, d)$ be the probability that the difference in the 3-dimensional position $\{S^i_t - S^j_t\}$ of aircraft pair $(i,j)$ hits or enters a disk $D(h,d)$ of height $h$ and diameter $d$, over a finite time interval $[0,T]$, i.e.

$$P_{\text{Reach}}^{i,j}(h, d) = \text{Prob} \left\{ \exists t \in [0,T] \text{ such that } \{S^i_t - S^j_t\} \in D(h, d) \right\}$$

Then the reach probability $P_{\text{Reach}}^{i,j}(h, d)$ for aircraft $i$ is obtained by a summation over these $P_{\text{Reach}}^{i,j}(h, d)$’s for all $j \neq i$, i.e.

$$P_{\text{Reach}}^{i,j}(h, d) = \sum_{j \neq i} P_{\text{Reach}}^{i,j}(h, d)$$

In Section 6.3 this reach probability is evaluated for an air traffic application with $h = 0$ and $d$ ranging from 0.1 nm to 6 nm. Hence $P_{\text{Reach}}^{i}(h, d)$ is here a metric for the probability that the mission fails to deliver a horizontal miss distance of $d$ or higher between aircraft $i$ and all other aircraft. Similarly, the complement $1 - P_{\text{Reach}}^{i,j}(h, d)$ is the probability that the mission successfully maintains horizontal separation of $d$ or higher between aircraft $i$ and all other aircraft.

5.2.7 Usability in Air Transport

From this literature review of resilience metrics one may conclude that there are multiple approaches to measuring resilience. The key question is which of these resilience metrics from various domains are most appropriate for air transportation? To make some progress we address this for the possible types of consequences identified in the introduction:

i. negligible consequences;
ii. catastrophic accidents involving one or more aircraft;
iii. significant local performance consequences;
iv. network-wide performance consequences.

For consequences of types (iii) and (iv), it is tempting to use the triangle in Figure 5.1 as a measure of the lack of resilience of the system considered in response to the disruption(s). Then engineering resilience is very effective in measuring the duration (A–D) of the resilience triangle in Figure 5.1. Typically, this duration is a measure of the extra time needed to complete a (safe) recovery from the disturbance. However, the real difficulty is how to measure the depth (A–B) of the immediate post-disruption performance degradation in the resilience triangle. The resilience metrics developed in various domains form an illustration of the difficulty in measuring this depth. As suggested by Dalziell and McManus (2004), one approach would be to measure this depth in terms of a weighted sum of multi-dimensional KPIs that are commonly used by the air transport community.

Consequences of types (i) and (ii) are not well captured by the resilience triangle interpretation. Consequence (i) means that there is no triangle at all. Consequence (ii) simply implies that there may be loss of aircraft hull(s) and passenger
lives, rather than recovery. The measure needed for type (i) consequences is one of ecological resilience, which characterises the amount of disruption that can be handled in such a way that the consequences are negligible. This leads to a shortlist of two remaining metrics: the psychological metrics (e.g., Likert scales) and the mission risk metric (e.g., reach probability). Because resilience metrics for individual humans only are insufficient for the complex socio-technical air transport system, the mission risk metric seems to be the best candidate. One advantage of the mission risk metric is that its complement forms a metric for mission success.

None of these metrics measures the individual contribution of the adaptive capacity separately from measuring the contributions of the absorptive and restorative capacities. To capture the effect of adaptive capacities, two measurements are required: one for the full complex system and another one for the complex system in which the adaptive capacities have been nullified.

Collecting real resilience data for air transport is challenging. To do so, one has to wait for particular, potentially rare, disruptions to happen. Considering resilience in the design of a novel operational concept presents more challenges, requiring the use of appropriate complexity science modelling and analysis approaches.

5.3 Complexity Science Perspective

5.3.1 Complex System Interdependencies

To improve the resilience of the air transport system, it is critical to identify, understand, and model system interdependencies (Ouyang, 2014). The performance of air transport operations, particularly under disruptive events, is dependent upon a set of highly interdependent subsystems including airlines, airports, and ATC centres. These subsystems are often connected at multiple levels through a wide variety of mechanisms, such that an interdependency exists between the states of any given pair of subsystems or components. Rinaldi et al. (2001) defined an interdependency as a bidirectional relationship between two infrastructures through which the state of each infrastructure influences or is correlated to the state of the other. As a simple example, airlines and airports are interdependent. An airport closure (for example, due to weather, limited capacity, or ATC strike) might cause airlines to cancel or divert their flights. At the same time, decisions made at an AOC centre influence and depend on airport processes (for example, gate change, passenger luggage). In normal air transport operations, some interdependencies are invisible, but under disruptive scenarios they emerge and become obvious. An illustration of this is the 2010 Eyjafjallajökull volcano eruption in Iceland which caused the closure of airspace of many European countries, and millions of passengers to be stranded at airports around the world.

Rinaldi (2004) identified four primary classes of interdependencies in critical infrastructure systems; these are presented in Table 5.2. An infrastructure system is defined by the US president’s commission on critical infrastructure protection
Resilience

(1997) as a network of independent, mostly privately owned, man-made systems and processes that function collaboratively and synergistically to produce and distribute a continuous flow of essential goods and services. Such a system is considered to be critical when its incapacity or destruction would have a debilitating impact on defence and economic security.

Modelling interdependencies in air transport is a complex, multidimensional, multidisciplinary problem. Table 5.3 lists some of the dimensions associated with system interdependencies that complicate resilience analysis. To model such interdependencies, there is a need for the systematic application, validation, and integration of modelling approaches. This view aligns with a common view in the literature that for the analysis of the resilience of complex critical infrastructure systems, multiple modelling and simulation approaches need to be integrated into a unifying framework that accounts for various dimensions (Ouyang, 2014). Each approach is appropriate for a certain number of resilience applications, depending on the components being modelled. Overall, the unifying framework can be used to assess the effectiveness of various resilience improvement strategies, supporting both strategic and tactical decision-making.

### 5.3.2 Complexity Science Approaches for Studying Resilience

Ouyang (2014) provided a comprehensive review of various complexity science modelling approaches and grouped them into several broad types: agent-based approaches, network-based approaches, empirical approaches, systems’ dynamics-based approaches, economic theory-based approaches, and other approaches such as hierarchical holographic modelling, the high-level, architecture-based method, Petri-nets, dynamic control systems theory, and Bayesian networks. These approaches have subsequently been systematically assessed against several resilience improvement strategies for critical infrastructure systems and the types of interdependencies they cover (Ouyang, 2014). Overall, agent-based methods and network flow-based methods appear to have the widest and most proven

<table>
<thead>
<tr>
<th><strong>Interdependency type</strong></th>
<th><strong>Definition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical interdependency</td>
<td>When the state of two systems is each dependent on the material output(s) of the other.</td>
</tr>
<tr>
<td>Cyber interdependency</td>
<td>When the state of a system depends on information transmitted through the information infrastructure.</td>
</tr>
<tr>
<td>Geographic interdependency</td>
<td>When the state of a system can change due to a local environmental event.</td>
</tr>
<tr>
<td>Logical interdependency</td>
<td>When the state of each of two systems is dependent on the state of the other via a mechanism other than one of the three listed above.</td>
</tr>
</tbody>
</table>
applicability, since they cover more resilience improvement strategies corresponding to the three resilience capacities than are covered by other approaches. Meanwhile, viability theory and stochastic reachability analysis (Bujorianu, 2012; Martin et al., 2011) are particularly adept at allowing researchers to model and analyse the various forms of uncertainty in air transport (see Chapter 4), and can be applied in both agent-based and network-based models. These four complementary modelling and analysis approaches are discussed in subsequent sections.

5.3.3 Agent-Based Modelling and Simulation

Agent-based modelling and simulation (ABMS) is increasingly recognised as a powerful approach to model complex socio-technical systems and to capture their emergent behaviour (Chan et al., 2010; Holland, 1998). This is because it can

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Implications for resilience analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple stakeholders</td>
<td>Stakeholders have different motivations and problems that drive the modelling requirements.</td>
</tr>
<tr>
<td>Multiple spatial scales</td>
<td>Scope of scenarios ranges from airports to the whole European airspace or to the global scale. Scale affects the resolution and quantity of interdependency data required for models.</td>
</tr>
<tr>
<td>Multiple time scales</td>
<td>Different events have varying timescales of relevance. The dynamics of the impacts vary from minutes (e.g., normal activities by the operators), to days (e.g., bad weather), up to years or even decades (e.g., catastrophic accidents).</td>
</tr>
<tr>
<td>Multiple key performance areas</td>
<td>Multiple competing key performance areas exist in air transport, e.g., safety, capacity, economy, environment. Resilience analysis should be performed with respect to the full spectrum of these key performance areas.</td>
</tr>
<tr>
<td>Cascading and higher order effects</td>
<td>Disruptions at one airport can propagate to other airports, creating second and higher order disruptions.</td>
</tr>
<tr>
<td>Socio-technical perspective</td>
<td>The air transport system is a socio-technical system. Behavioural responses can influence the efficiency and safety of operations (e.g., situation awareness of operators, or passenger response to an infectious disease).</td>
</tr>
<tr>
<td>Disruption management plans</td>
<td>Recovery procedures influence the state of a system during a disruption and may affect coordination among various stakeholders, e.g., disruption management by AOC.</td>
</tr>
<tr>
<td>Regulations</td>
<td>Regulations influence operational behaviour as well as the response to and recovery from disruptions, e.g., cancelling a flight due to curfew at a destination airport.</td>
</tr>
<tr>
<td>Growing demand</td>
<td>Constant growth in the number of flights, aircraft and airports. Rapid change of the market (from a small number of national airlines to the recent appearance of many companies with new business models).</td>
</tr>
</tbody>
</table>
Resilience 73

represent important phenomena resulting from the characteristics and behaviours of individual agents and their interactions (Railsback and Grimm, 2011). Burmeister et al. (1997) discuss the benefits of using this approach in domains that are functionally or geographically composed of autonomous subsystems, where the subsystems exist in a dynamic environment, and the subsystems have to interact flexibly. According to Burmeister et al. (1997), ABMS can be used to structure and appropriately combine the information into a comprehensible form. For a complex socio-technical system, the approach provides the tools for analysing, modelling, and designing the whole system in terms of agents, each with its own set of local tasks, capability and interactions with the other agents. Agents can be described at a high level of abstraction, yet the resulting composition is very efficient. Burmeister et al. (1997) conclude that ABMS reduces the complexity in systems design by making available abstraction levels that lend themselves to a more natural way of modelling in the problem domain. In the same vein, Jennings (2000) outlines that the requirements of ABMS are highly compatible with those of complex systems’ development. He shows that an ABMS approach is particularly well suited to complex systems because: (a) they provide an effective way of partitioning the problem space of a complex system; (b) they provide a natural means to model complex systems through abstraction; and (c) they capture the interactions and dependencies. In Chapter 6, ABMS is further explained regarding its ability to identify emergent behaviour in complex socio-technical designs.

5.3.4 Network-based Methods

As explained in Chapter 2, network theory is used to investigate the structure and topology of networks, and has applications in many disciplines including computer science, economics, sociology and operations research. Network-based methods are particularly useful for analysing the complex structure of large-scale systems. For instance, centrality measures can quantify the relative importance of network nodes and links (Newman, 2004). Dependency analysis between the nodes can calculate higher-order and cascading effects. Ouyang (2014) has classified network-based methods into two main categories: topology-based methods and flow-based methods. The former models a network based on its topology; the latter takes into account the service or flow made and delivered by the system. According to Ouyang (2014), network flow-based methods relate to all three resilience capacities, in contrast to topology-based methods, which relate only to the absorptive capacity. Both types of method are relevant to air transport. Examples of topology-based methods include those of Guimerà et al. (2005), who analysed the topology of the worldwide air transport network; Chi and Cai (2004) who analysed how the topological properties of the US airport network are affected when a few airports are no longer operational (for example, due to failures or attacks); and Li and Cai (2004), who studied the airport network of China. A complementary example of results obtainable by network flow-based approaches is the analysis of delay in the US airspace system (Meyn et al., 2004) using the airspace concept evaluation system (ACES) simulator.
5.3.5 **Stochastic Reachability Analysis**

The primary aim of stochastic reachability analysis is to evaluate the probability that a system can reach a target state starting from a given initial state. This is especially of interest in air transport where the system should be kept outside an unsafe region of the state space, and the control input can be chosen so as to avoid this unsafe region. Modern applications of stochastic reachability analysis have become increasingly complex. This complexity is due to the rich interactions, complicated dynamics, randomness of environment, uncertainty of measurements and tolerance to faults (Bujorianu, 2012). Examples of illustrative applications in air transport include the work of Prandini and Hu (2006, 2008), who use stochastic reachability analysis to study aircraft conflict detection, and of Blom et al. (2007, 2009), who use stochastic reachability analysis to study collision risk in ATM (see sections 6.2–6.3).

5.3.6 **Viability Theory**

Viability theory (Aubin, 1991) was originally developed to study dynamical systems which collapse or badly deteriorate if they leave a given subset of the state space. Therefore the objective is to keep the system in the part of the state space where it can survive, that is, where it is viable. In follow-up research by Aubin et al. (2002), viability theory has been extended to hybrid dynamical systems. Recently, Martin et al. (2011) have explained that viability theory provides a natural mathematical framework for the modelling and analysis of resilience in complex systems. In general, viability theory can be applied in a wide range of domains, including cognitive sciences, finance, economics and the sociological sciences. An example application in air transport is obstacle avoidance, which also appears in numerous application fields. Other examples include using viability algorithms to compute wind optimal routes to reach an airport in minimal time, and computing safety envelopes of an aircraft in different phases of flight (Aubin et al., 2011).

5.3.7 **Use in Air Transport**

The use of these methods in resilience modelling and analysis in air transport may depend on the specific application. Below and in Table 5.4 we make this more precise for the four types of consequences addressed earlier, i.e. (i) negligible consequences; (ii) catastrophic accidents involving one or more aircraft; (iii) significant local performance consequences; and (iv) network-wide performance consequences.

For types (i), (ii) and (iii) consequences, pilots and controllers may play a key role in reacting appropriately to various events. In such cases ABMS seems the most appropriate approach. For type (ii) consequences, it is explained in Section 6.3 that agent-based modelling and analysis has to be combined with mathematical methods from the stochastic reachability domain; without these mathematical methods the Monte Carlo simulation of an agent-based model may
be impracticable. In contrast with traditional safety risk analysis, an ABMS approach can cover both Hollnagel’s (2014) Safety-I (i.e. ‘what can go wrong’) and Safety-II (i.e. ‘what can go right’). This dual capability of ABMS is illustrated in Section 6.4 for an advanced airborne self-separation concept of operations.

For consequences of types (i), (iii) and (iv), the network flow-based methods seem to be the most logical fit as long as human involvement does not play a key role. Otherwise, ABMS might be the better choice. The airspace concept evaluation system used by Meyn et al. (2004) is a network flow-based method that uses an agent-based architecture, illustrating that, in practice, the network and agent-based methods tend to be integrated. If an agent-based or a network flow-based model has been developed in a proper mathematical setting, then this model can also be applied to viability and reachability analyses for the specific application considered.

5.4 Airline Disruption Case

The aim of this section is to illustrate the use of ABMS to study the effect of four coordination strategies on the response of an airline to a disruption influencing many human and technical agents. The specific resilience metric used is engineering resilience, which originates from the ecosystem domain (Section 5.2.1).

5.4.1 Airline Disruption Management

An airline’s flight schedule is subject to many disruptions, including deteriorating weather, passenger delays and aircraft- or crew-related problems. Each such disruption may be detrimental to the delivery of the daily fleet schedule of an airline and to the smooth and timely movement of passengers from their origins to their destinations. Operators at the AOC centre take corrective actions to recover from disruptions. This can only be done through interaction with teams external to the airline, for example, at ATC centres and airports. Within AOC centres, many operators with different roles interact and coordinate towards achieving a common goal, namely to manage disruption so that their airline operations adhere to the strategic plan (schedule) as closely as possible.
Bruce (2011a, 2011b) has systematically studied decision-making processes in six AOC centres, seeking advice from an expert panel of airline operations management staff and operators to ensure broader views (for example, in terms of gender, age, years of experience in the airline industry, years of experience in the AOC domain, and previous occupation). Figure 5.3 gives a high-level overview of a

![Diagram of the organisation of airline operations control including the communication with air traffic control](image)

**Figure 5.3** Overview of the organisation of airline operations control including the communication with air traffic control  
*Source:* Bouarfa et al. (2014), with permission of the American Institute of Aeronautics and Astronautics, Inc.
5.4.2 Four Disruption Management Policies

To understand the impact of various policies on the performance of airline disruption management, four different policy types have been defined. Three of them (P1–P3) are based on established practices, the fourth (P4) is based on recent coordination theory.

Policies P1–P3 are based on Bruce 2011a and 2011b. As shown in Table 5.5, these three policies capture three different decision-making styles and coordination strategies used by airline controllers. Under P1, airline controllers identify straightforward considerations such as aircraft patterns and availability, crew commitments and maintenance limitations. Under P2, airline controllers have a greater comprehension of the problem. They take into account the more complex consequences of the problem. Under P3, airline controllers demonstrate thinking beyond the immediacy of the problem and examine creative ways to manage the disruption.

The fourth policy (P4) is based on the theory of Klein et al. (2005) regarding the coordination of joint activity by multiple actors. This theory identifies three process types that are required for effective coordination, namely: (A) the criteria for joint activity processes; (B) the satisfying requirements for joint activity; and (C) the choreography of joint activity. The criteria for joint activity (A) are that participants in the joint activity agree to support the coordination process and to prevent its breakdown. If these criteria are satisfied, the parties have to fulfil certain

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Policy 1 Elementary</th>
<th>Policy 2 Core</th>
<th>Policy 3 Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintenance Information</td>
<td>Accept information source and content and act on information given about a maintenance situation.</td>
<td>Challenge/query information about a maintenance situation.</td>
<td>Seek alternative information and recheck source and reliability.</td>
</tr>
<tr>
<td>Crewing</td>
<td>Await crew from inbound aircraft.</td>
<td>Challenge crew limits/Seek extensions to crew duty time.</td>
<td>Seek alternative crew (e.g., from nearby base or other aircraft).</td>
</tr>
<tr>
<td>Curfews</td>
<td>Curfews are not taken into account.</td>
<td>Identify curfews and work within them.</td>
<td>Seek curfew dispensation.</td>
</tr>
<tr>
<td>Aircraft</td>
<td>Seek first available aircraft.</td>
<td>Request high-speed cruise.</td>
<td>Combine flights to free up aircraft.</td>
</tr>
</tbody>
</table>

Adapted from Bruce (2011a).
requirements (B) such as making their actions predictable, sustaining common ground, and being directable. The choreography (C) for achieving these requirements is a series of phases that are guided by various signals and coordination devices, in support of effective coordination. To apply this novel policy to AOC, Bouarfa et al. (2014, 2015) have studied and defined (informal) coordination rules to which AOC agents should adhere.

Table 5.6 shows which of the three resilience capacities apply to each of the four AOC policies. Only policies P3 and P4 have adaptive capacities. Under P3, individual controllers examine creative ways to manage or avoid disruptions, thereby adjusting their strategies. Under P4, controllers develop adaptive capacity at the team level through coordination with each other (Klein et al., 2005). Policy P1 lacks not only adaptive capacity but also restorative capacity. The explanation is that under P1, controllers act on information given about a certain situation without challenging it. For instance, if information is coming about an instrument indication problem from the pilot, the controller would turn the aircraft back to the airport with its passengers. However, such a decision is not always needed, as the problem could be a loose wire when the instrument was changed and could be fixed at the next airport. Therefore, the rapidity of return to normal operations and ability to adjust are lacking in P1. Finally, P2 is different from P1 in that it does have restorative capacity, as controllers take into account the more complex consequences of the problem, challenge and request additional information, thus expediting the recovery process.

### 5.4.3 Airline Disruption Scenario

To assess the impact of the four policies (P1–P4) we consider a challenging scenario that is well evaluated in the literature (Bruce, 2011a, 2011b). The scenario concerns a mechanical problem with an aircraft on the ground at Charles de Gaulle (CDG) airport, which is scheduled for a long-haul flight to a fictitious airport in the Pacific (PC). The flight was progressively delayed at CDG for three hours due to mechanical unserviceability, to the extent that the operating crew members were eventually unable to complete the flight within their legal duty time.
This scenario was also considered by a panel of airline operations management experts. They developed several alternatives, and subsequently identified the best solution, which was to reroute the flight from CDG to PC to include a stopover in Mumbai (BOM). In parallel, a replacement flight crew was flown in as passengers of a scheduled flight from PC to BOM, to replace the delayed crew on the flight part from BOM to PC. The question is how the outcomes of ABMS of an AOC centre compare to this best solution found by the expert panel.

5.4.4 Agent-based Simulation Results

For each of the four disruption management policies P1–P4, an agent-based model has been developed (Bouarfa et al., 2015). The variations in these policies lead to differences in terms of the sequence of agent involvement, the information being exchanged, and the sequence of activities.

Table 5.7 presents some of the agent-based simulation results obtained for the four policies. See Bouarfa et al., 2015, for more complete agent-based simulation results, such as various costs. P3 concurs with the best solution identified by the expert panel. However the outcomes of P1 and P2 are significantly worse, and the outcome of P4 outperforms even the expert panel result. To understand these differences, the agent-based simulation results have carefully been analysed.

Under policies P1 and P2, AOC operators make decisions at the core or elementary level and with limited coordination, as a result of which the disruption is not efficiently managed. The aircraft mechanical problem was eventually fixed, however the flight was cancelled. As a result, the 420 passengers were accommodated in hotels (i.e. greatly inconvenienced). This unfavourable outcome can be explained

<table>
<thead>
<tr>
<th>AOC policy</th>
<th>Flight</th>
<th>Aircraft mechanical problem</th>
<th>Crew problem</th>
<th>Passengers problem (all at a cost to the airline)</th>
<th>Minimum disruption management time</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>Cancelled</td>
<td>Fixed</td>
<td>Not resolved</td>
<td>Pax. accommodated in hotel</td>
<td>28 min</td>
</tr>
<tr>
<td>P2</td>
<td>Cancelled</td>
<td>Fixed</td>
<td>Not resolved</td>
<td>Pax. accommodated in hotel</td>
<td>28 min</td>
</tr>
<tr>
<td>P3</td>
<td>Delayed</td>
<td>Fixed</td>
<td>Resolved</td>
<td>Pax. delayed due to fixing aircraft and due to flying via BOM</td>
<td>28 min</td>
</tr>
<tr>
<td>P4</td>
<td>Delayed</td>
<td>Fixed</td>
<td>Resolved</td>
<td>Pax. delayed due to fixing aircraft</td>
<td>19 min</td>
</tr>
</tbody>
</table>

See Bouarfa et al. (2015) for full results.
by the possible actions at level 1 and 2 by the crew controller, i.e. ‘await crew from inbound aircraft’ and ‘seek extensions to crew duty time’. Crew controllers at this level mainly consider crew sign-on time and duty time limitations and try to work within these constraints. In this scenario, none of the actions solves the crew problem.

Under policy P3, airline operations controllers consider complex crewing alternatives such as transferring crew from another airport. Therefore, under P3 the decision was made to divert the flight to BOM and position another crew from PC into BOM. Here, both the delayed crew and the replacement crew were able to operate in one tour of crew duty time. In comparison to policies P1 and P2, policy P3 has a much better outcome from both the airline and passenger perspective. The time required to manage the disruption for policy P3 is equal to that for P1 and P2.

Under policy P4, airline agents make decisions at the elementary level, like P1–P2, but under a healthy coordination regime. Therefore the aircraft, crew, and passenger problems were resolved with minimum disruption. The main difference between P4 and the other policies – P1–P3 – is that airline agents now act according to coordination rules (Bouarfa et al., 2015) that account for all joint activity phases (criteria, requirements, and choreography). Thus, for instance, when the crew controller cannot find a crew, he signals his understanding about the situation and the difficulties he is facing. Likewise, the airline operations supervisor signals his understanding back to the crew controller just to be sure of the crew situation or to give the crew controller a chance to challenge his assumptions. Such a process of communicating, testing, updating, tailoring, and repairing mutual understandings is aimed at building common ground prior to starting the choreography phase (Klein et al., 2005). By updating the crew controller on changes outside his view, and coordinating by agreement, precedent and salience, he managed together with the crew controller to solve the crew problem before moving to the next coordination phase. In the scenario considered, P4 was therefore able to identify a possibility that had not been identified by any of the other three policies, nor by the expert panel: the flight crew that had landed the aircraft at CDG had received sufficient rest to fly the delayed aircraft directly to PC instead of enjoying their scheduled day off in Paris. Passengers had a minimum delay compared to the previous policies (P1–P3), as they only had to wait for the aircraft to be fixed. Another relevant difference between P4 and the other policies P1–P3 is the shorter time needed to manage the disruption.

5.5 Conclusions

Thanks to scholars from behavioural sciences, it has become clear that for the future development of air transport, resilience to various types of possible disruptions should be studied. The possible consequences of such disruptions may be negligible, or there may be significant consequences such as catastrophic accidents, significant local consequences, or very severe network-wide consequences. This chapter has conducted a systematic study of what complexity science has to offer to the study of resilience in future air transport for the various types of consequence.
A socio-technical system is said to be resilient when it has adaptive capacities in addition to absorptive and restorative capacities. A system that has absorptive capacity only is called robust. A system that has absorptive and restorative capacities is called dependable. Robustness and dependability are properties of systems that are well addressed through systems engineering. For air transport, this means that the key resilience challenges are not only to address a complex socio-technical system rather than a complex technical system but also to improve the adaptive capacities. These adaptive capacities concern both the phase of disruption absorption and the phase of recovering from a system performance degradation due to disruption.

Placing emphasis on improving the adaptive capacity in absorbing disruptions concurs with the resilience engineering definition of Hollnagel et al. (2009) for use in ATM research, which states that a resilient system is one which has ‘the intrinsic ability to adjust its functioning prior to, during, or following changes and disturbances, and thereby sustain required operations under both expected and unexpected conditions’. In the safety domain, Hollnagel (2014) explains that this resilience engineering view reveals a need to study ‘what may go right’, rather than the traditional approach of studying only ‘what may go wrong’. The traditional and novel approaches are called ‘Safety-I’ and ‘Safety-II’ respectively.

In the literature, qualitative and quantitative resilience metrics have been developed in various domains. The qualitative measures are of two types: ecological resilience and engineering resilience. Ecological resilience is a measure for the amount of disruption that the socio-technical air transport system can absorb before it leads to significant changes in its key performance areas. Engineering resilience is a measure of the duration between the moment of significant reduction in its KPIs and the moment of recovery.

Most resilience metrics relate to engineering resilience, that is, they address recovery rather than the avoidance of significant consequences. Exceptions are the psychological metrics (e.g., Likert scales) for individual human performance (Ahern et al., 2006), and mission risk, such as reach probability for conflict and collision risk in ATM (Prandini and Hu, 2008; Blom et al., 2009).

No resilience metric in the literature is able to capture the effect of the adaptive capacities of a socio-technical system separately from the capture of the effects of the absorptive and restorative capacities. As has been shown in Section 5.5, an effective way to address this problem is to develop a proper model of the socio-technical system considered, and subsequently perform two measurements: one for the full model, and the other for a version of the model in which the adaptive capacities are nullified.

Complexity science provides powerful modelling and analysis methods, the most important of which are ABMS, network flow-based methods, stochastic reachability, and viability theory. When human operators play a key role in the specific resilience aspect to be studied, then ABMS is the logical choice. When the resilience issue to be studied is concerned with propagation of disruption effects through a network, then a network flow-based method is the preferred choice. When both aspects play a role, then a network flow-based approach that uses agent-based architecture might be used. Once a proper agent-based or network flow-based
model has been developed, this may be used as a basis for stochastic reachability analysis or viability theory. These complexity science approaches allow a model of the socio-technical air transport system to be used to assess the possible effects upon KPIs by increasing the size of disruptions and by varying disruption management strategies in each of the three capacities. The practical working of this approach is demonstrated in Section 5.5 by quantifying the impact of adopting changes in coordination policies by AOC, for example, by making them more or less adaptive.

In conclusion, this chapter has shown that a complexity science approach to resilience in air transport has significant potential to both strengthen and broaden the resilience engineering approach of Hollnagel et al. (2006, 2009). It has also been demonstrated that a complexity science-based approach to resilience yields practical results for the air transport system. This great potential brings with it several valuable directions for research, such as:

- to further develop and apply mission risk metrics that capture the effect of absorptive and adaptive capacities of the air transport system to both separation related and non separation-related disruptions;
- to further develop metrics relevant to the recovery and adaptation of the air transport system from performance degradation due to disruptions;
- to further the development and application of ABMS for the evaluation of both positive and negative impacts of potential resilience improvements in the future designs in ATM and air transport operations;
- to further the development of network flow-based modelling and its integration with ABMS for the evaluation of recovery from network wide performance degradation in the air transport system;
- to further develop the application of reachability and viability theories to the air transport system, by taking advantage of the above-mentioned network-flow and agent-based model developments.

Note

1 In systems engineering, reliability is the ability of a system or component to perform its required functions under stated conditions for a specified period of time. One should note that this systems engineering definition of reliability is more restricted than what is meant when we refer to a ‘reliable’ airline. Such an airline is indeed reliable in the sense of the systems engineering definition. However, an airline also needs to be adaptive in response to unexpected adverse conditions, in order to perform in a competitive market. This entails getting passengers (and their bags) to their destinations (reasonably on time) and, indeed, having a reputation for doing so. Successful airlines thus have an adaptive capacity, rendering them resilient.
In complexity science a property or behaviour of a system is called ‘emergent’ if it is not a property or behaviour of the constituting elements of the system, but results from the interactions between them. In the socio-technical air transport system these interactions are between human operators, technical systems and procedures at airlines, airports and air traffic centres. Safety is a clear example of an emergent property of the air transport system: it is not a property of any one constituent element, but results from the interactions between the constituting elements. Like other emergent properties and behaviours, safety results from decades of evolutionary development. For example, each commercial air accident is well investigated, with the findings often leading to improvements to elements of the air transport system. The result has been a steady improvement of safety despite increasing air traffic (Roelen and Blom, 2013). One may therefore conclude that current aviation works thanks to the implicit use of emergent behaviour.

To accommodate expected growth in commercial air traffic, significant changes to the air transport system are in development both in the US and in Europe. In addition to accommodating much higher traffic demands, the challenge is to improve performance across other key performance areas, including safety and efficiency (SESAR, 2007a). It is, therefore, essential to study and understand the effects of changes to individual elements and interactions on emergent properties and behaviours (Bar-Yam, 2003; Holmes, 2004; SESAR, 2007b). A key issue for future air traffic design is the identification of unknown emergent risks (EUROCONTROL, 2010b). Unanticipated hazards could arise as a result of new concepts, tools, or procedures. Similarly, positive emergent behaviours may also be currently unknown. Woods et al. (2010) explain that while new paradigms (e.g., airborne self-separation) could give rise to new vulnerabilities, they could also remove existing ones. More generally, new emergent behaviour that is not well understood often leads to poor performance. Once new emergent behaviour is better understood, it may be possible to generate emergent behaviour that yields performance improvements. The challenge, which this chapter addresses, is how to identify and study emergent behaviour during the early design stage of future ATM.

This chapter is organised as follows. Section 6.1 studies emergent properties and behaviour in air transport. This starts with a review of the various emergence perspectives developed in the literature, followed by illustrative examples of emergence, and concludes with a study of what this means for the identification and
analysis of emergence in future air transport designs. Section 6.2 provides a systematic identification of methods to identify and analyse emergence in air transport and Section 6.3 presents an application of these methods to an advanced design. Section 6.4 provides concluding remarks.

6.1 Emergent Behaviour in Air Transport

6.1.1 Perspectives on Emergent Behaviour

Philosophers have long been interested in the concept of emergence, and especially in trying to establish a common definition for this vague yet very useful concept. Aristotle (350 BC) referred to it, saying ‘the whole is something over and above its parts, and not just the sum of them all’. Mill (1872) used the example of water to illustrate the same idea, noting that the physical properties of water cannot be predicted from knowledge of the properties of hydrogen and oxygen atoms. The term ‘emergent’ was coined by Lewes (1875), who argued that certain phenomena produce ‘qualitative novelty’ or material changes that cannot be expressed in simple quantitative terms. Quoting Lewes: ‘The emergent is unlike its components insofar as these are incommensurable, and it cannot be reduced to their sum or their difference’ (Lewes, 1875).

A fundamental aspect of emergence is that it refers to a property of a system or structure on a higher level of organisation. Bedau (1997, 2008) calls this a ‘macro-property’. A property is called ‘emergent’ if it is not a property of any one element of the system: it is a macro-property that cannot be a micro-property. For example, water can have the property of transparency, but an individual water molecule cannot. The difficulty in understanding this is that an emergent property of a system also arises out of the properties of its elements and the relationships between them, though often there is no simple explanation. Transparency of water cannot simply be explained by studying individual water molecules, yet it emerges from the properties of these molecules. Much of the literature on emergence has focused on understanding the relations between micro-processes and macro-properties. The study of emergence found renewed interest in the late twentieth century with the growth of scientific interest in complexity and the development of new, non-linear mathematical tools (Corning, 2002). This came at the expense of ‘reductionism’, which implies the ability to understand all phenomena completely in terms of the processes from which they are composed.

A more contemporary interpretation of the reductionist view is ‘supervenience’ (Horgan, 1993; Kim, 1991). A macro property supervenes on a micro property if there cannot be a macro-difference without a micro-difference. For example, there can be no difference in the temperature of a gas without some difference in the behaviour of its molecules. Although the idea of supervenience seems straightforward, ‘problems’ with the concept have triggered debate (Kim, 1991).
Another important term in the literature on emergence is ‘downward causation’. Several interpretations of the term are in circulation. A ‘strong’ definition is proposed by Sperry (1964), who states that downward causation means that the macro properties have the power to control micro processes. Other authors propose ‘weaker’ definitions; Campbell (1974) states that in downward causation the micro processes are restrained by and act in conformity to macro properties. For this weaker definition, many more examples can be found in nature and culture (Heylighen, 1995).

Bedau (2008) defines ‘nominal emergence’ as a ‘macro property’ that cannot be a ‘micro property’. He argues that there are three main types of nominal emergence: resultant, weak and strong:

- **Resultant emergence** is nominal emergence where the macro properties are derivable from the micro processes without simulation. Hence resultant emergence can be fully explained in terms of the micro processes.

- **Weak emergence** is the notion of a macro property that is derivable from the micro processes, but only by simulation, that is, by watching how it unfolds in time. The micro-level interactions are interwoven in such a complicated network that the macro properties have no simple explanation. Weak emergence may involve downward causation in the sense that the micro properties cannot be explained by micro–micro interactions, but only by objective macro properties that unify an otherwise heterogeneous collection of micro instances.

- **Strong emergence** is nominal emergence in which the emergent properties are supervenient properties with irreducible macro causal powers. These macro causal powers have effects both at the macro-to-macro level and at the macro-to-micro level; the latter effects refer to downward causation. The ‘irreducibility’ part explains that the macro properties are autonomous (cannot be derived) from the micro processes. The macro properties do depend on their micro processes, though this is not derivable through simulation.

Bedau’s example of strong emergence is ‘consciousness’, which cannot be derived, even in principle, from the physical properties of human beings, including their genes, neurological connections and DNA. In the absence of other examples, Bedau (2008) argues that although strong emergence has had a prominent place in philosophical discussions, its scientific credentials are poor, whereas weak emergence is consistent with materialism and is scientifically useful. Bedau also proposes splitting the concept of weak emergence into three types: a) weak emergence that in principle is derivable without simulation but in practice must be simulated; b) weak emergence that in principle is underivable except by finite feasible simulation; c) weak emergence that is underivable except by simulation, but the requisite simulation is unfeasible or infinite. Bedau applies a deterministic view on the types of simulations considered, which, for example, precludes Monte Carlo simulation.
Chalmers (2002) includes a notion of ‘unexpectedness’ or ‘surprise’ in the definition of emergence, which leads to some alternative definitions for strong and weak emergence. Other authors also refer to the notion of surprise, like Sanz (2004), who defines emergence as ‘just systemic behaviour that is difficult to predict in advance’. Bedau (2008) explains that he left the notion of surprise absent on purpose, due to it being rather subjective. Instead, Bedau claims that with his definition of weak emergence in terms of simulation he is presenting objectivist approaches to emergence, though he notes that his classification is not exhaustive.

Zarboutis and Wright (2006) explain that emergence is closely related to ‘adaptation’; in complexity science, each micro component interacts with its neighbouring ones, and adapts its internal organisation to meet individual criteria. They do this by using local information and usually while being unaware of the properties that emerge at the macro level. The process of micro-to-macro emergence simultaneously undergoes a form of macro-to-micro control that aims to assure that the macro properties are meaningful. The complex system ideally finds a form of ‘optimal adaptation’, which balances emergence and hierarchical control. The notion of adaptation seems to find a balance between micro-to-macro and macro-to-micro causation by adding various feedback loops.

Bar-Yam (2004) develops a mathematical theory behind Bedau’s strong emergence. Partly based on this, Fromm (2005) places the philosophical emergence views of Bedau and Chalmers within the context of complex multi-agent systems. This leads to a splitting of Bedau’s weak emergence into two types, and to clear scientific credentials for strong emergence. Fromm’s type I emergence corresponds to the ‘resultant emergence’ of Bedau (2008). Fromm’s type II is Bedau’s weak emergence with single feedback from the macro property to the micro processes. Fromm’s type III is Bedau’s weak emergence with multiple macro properties as well as multiple feedbacks between macro–macro and macro–micro. Type III emergence typically applies when a multi-agent system incorporates intelligent agents (for example, those who can learn). Fromm’s type IV emergence is Bedau’s strong emergence, though with the novel insight that macro properties may emerge in a complex multi-agent system in a hierarchically hidden, multi-level way. This makes it feasible that Bedau’s strong emergence conditions are satisfied without violating any law of physics. Fromm’s examples of strong emergence are life and culture. Life is a strong emergent property based on genes, genetic code and nuclei/amino acids. Culture is a strong emergent property based on memes, language and writing systems.

6.1.2 Emergent Behaviour in Air Transport

Air transport is a complex system in which several emergent properties, phenomena and behaviours appear (Everdij et al., 2011). This subsection gives some illustrative examples, and tries to classify them in terms of Bedau’s ‘resultant emergence’, ‘weak emergence’ and ‘strong emergence’.
Resultant emergence describes macro properties that cannot be micro properties, but that can be derived from all the micro properties, without using simulation. The many examples from air transport include:

- the function of a technical system on-board an aircraft or on the ground. The function of a technical system is a result of the fixed roles of the components, but it cannot be fulfilled by these components in isolation;
- thermodynamic properties such as pressure, volume, temperature, which play a role in airworthiness issues;
- for some air transport operations, such as procedural management of oceanic air traffic, the risk of an aircraft collision can be predicted without simulations, using information about the parallel route structure, the traffic flows, and the statistics of large deviations by aircraft from their agreed route (Reich, 1966).

One example of strong emergence is the safety culture in an airline, or in an air traffic control (ATC) centre. It is an evolutionary product of routine aspects of everyday practice and rules, of management and organisational structures, and of national–cultural behaviours (Ek et al., 2007; Gordon et al., 2007). Even through simulation, causal relationships in this safety culture have not been revealed (Sharanskykh and Stroeve, 2011).

Weak emergence is abundant in air transport as the complex critical infrastructure involves each of the following key interdependencies identified by Rinaldi (2004).

- **Physical interdependency** Two systems are physically interdependent if the state of each depends upon the material output(s) of the other. This applies to the many aircraft in a sector and their interdependency with ATC in that sector.
- **Cyber interdependency** A system has a cyber-interdependency if its state depends on information transmitted through the information infrastructure. This applies, for example, to the information exchange between entities on the ground and in the air.
- **Geographic interdependency** Systems are geographically interdependent if a local environmental event can create state changes in all of them, such as when a major weather front enforces distant aircraft and air traffic sectors to make significant changes to their plans.
- **Logical interdependency** Two systems are logically interdependent if the state of each depends upon the state of the other via some mechanism that is not a physical, cyber, or geographic connection. This applies to the explicit involvement of humans (e.g., pilots, controllers) in the critical decision-making processes.

In air transport, these interdependencies are well organised and induce various types of weakly emergent macro properties and behaviours. Because of the
involvement of human agents (e.g., pilots, controllers), the macro-properties are typically of Fromm’s type III. In some exceptional events, the interdependencies may fail. There are two categories of such exceptional events. The first is events that force the dynamics of the air transport system outside its normal mode of operation and therefore dramatically affect the performance of the system. The second category is safety-critical behaviour involving one or more aircraft that in rare cases leads to accidents. These two categories are not necessarily mutually exclusive.

Examples of the first category of exceptional events which can cause interdependencies to fail include phase transitions and percolation in a network. A phase transition describes many locally interacting elements causing a collective phase change (returning to the example of water, a physical analogue is the melting of ice, i.e., a transition from the solid to the liquid phase). Typically, there exists a critical point that marks the passage from one phase to another (Helbing, 2001). For example, in road traffic a relatively small increase of traffic demand may lead to a sudden decrease in travel velocity, which may decrease the total traffic flow (Kerner, 2004). This specific kind of phase transition is largely absent in air traffic, due to caps imposed on air traffic flows. Percolation in a network describes probabilistic, network-wide propagation between sites or subsystems across links in the network. In air transport, there are several networks where propagation may happen; for example, the spatio-temporal propagation of congestion over airspace sectors (Ben Amor and Bui, 2012; Conway, 2005) or the propagation of the queuing of passengers through the whole air transport system.

Safety and safety perception are examples of the second category of exceptional events relating to weak emergence. Safety is the complement of the macro property safety risk. Safety perception is a weak emergent macro awareness by a human agent (e.g., a pilot or controller) regarding the safety of those flights that fall under his/her responsibility. Both safety risk and safety perception emerge from and influence behaviour at various other macro levels. For example, the propagation of one or more hazards through the air transport system may create a condition under which the application of established procedures by crew or ATC unintentionally causes the situation to deteriorate. This may occur, for example, when there are differences in situation awareness between different agents in the system, and these differences cannot be recognised by any of the agents (De Santis et al., 2013). This kind of emergent behaviour can have fatal outcomes. Examples include the Linate runway collision in 2001, the Ueberlingen mid-air collision in 2002 and the Air France crash in the Atlantic Ocean in 2009 due to loss of control of the aircraft. Fortunately, most of the time, the exceptional safety-critical behaviour is resolved prior to evolving into an accident.

One example of such exceptional safety-critical behaviour that is identified by simulation is described by Stroeve et al. (2013). A concept of operations is considered in which frequent active runway crossings take place on a departure runway in good visibility conditions. To limit the potential risks related to such operations, the concept included a runway incursion alerting system to warn the air traffic
controller in situations in which a departing and a crossing aircraft simultaneously make use or start making use of the runway. According to early safety assessments using traditional approaches, such as fault trees and event trees, the alerting system would significantly reduce risk. However, in Monte Carlo simulations of a dynamic risk model of the actors, systems and interactions, the risk-decreasing contribution of the alerting system for the air traffic controller appeared small. In the simulation, in most situations in which the alerting system enables the air traffic controller to warn the pilot, the pilot of one of the involved aircraft has already identified and started to resolve the conflict. If in time-critical situations the conflict were not detected by the pilots, it would often not be resolved via the alerting system either, for example, because of a late alert, a delay in the communication line between controller and pilots, or a late or inappropriate reaction of the controller or pilot.

The described effect was discovered only after developing and simulating an agent-based risk model that covered the totality of interactions of components, including their variability in performance over time. The complexity of air transport operations involves a combinatorial explosion of the many events that may occur in a dynamic way and the many uncertainties involved, such that certain aspects of safety risk can only be studied through simulation. The human mind is simply not able to grasp the many combinations of events occurring later or earlier than average, or resolutions of situations that are implemented differently, even when supported by graphical tools such as tree-based schemes or analytical equations. Monte Carlo simulations made it possible to identify how the operation evolves through time in a dynamic way, addressing the combinatorial explosion and allowing specific behaviour to emerge.

6.1.3 Identifying Emergence in Future Air Traffic Design

The challenge posed by SESAR and NextGen is to make significant changes in air transport. SESAR (2006) has formulated strategic objectives to increase both capacity and economy by a factor of two, and safety by a factor of ten. To achieve these challenging objectives, changes in infrastructure are foreseen. For example, cyber interdependency will be strengthened by system-wide information management, physical interdependency will be strengthened by trajectory-based operations (TBO), and logical interdependency will be strengthened by changing the roles of pilots and controllers. These changes may affect emergent behaviours.

Following Bedau (2008), simulation is required to capture as yet unknown weak emergence. There are three established types of simulation tool available:

- **Human-in-the-loop simulation (HITL)** This works well for the identification of weak emergent behaviour that happens under normal conditions, for example, to identify that a pilot or controller tends to use a technical system or procedure in a different way than that intended by the developers.
• **Network flow-based simulation** This works well for identifying how specific propagation patterns in the air transport network change as a result of a new design, for example, to identify the impact of the design change on the traffic flows in the case of a significant disturbance, such as bad weather (Gong et al., 2012).

• **Agent-based modelling and simulation (ABMS)** This works well when there are many interacting agents, particularly if these agents have intelligence, e.g., pilots and controllers. Shah et al. (2005) explain that this approach can identify emergent behaviour in air transport in which human agents play a key role. Monechi et al. (2013) use an agent-based modelling approach to identify phase transitions when the traffic flow in a sector is not capped.

The problem is that these three simulation tools cannot capture emergent behaviours of exceptional safety critical events, nor can they be used to analyse the safety risk of a novel design. The gap between these established simulation approaches and what is required is depicted in Figure 6.1. At the base of the safety pyramid there are the high-frequency controller and pilot actions, which may happen in the order of 10 to 100 events per flight hour. These events are well analysed by human-in-the-loop simulation and by ABMS. However, these methods leave unexplored the weak emergent behaviour that happens along the flank and at the top of the safety pyramid. Halfway up the flank, there are rarer incidents happening in the order of 1 per 10,000 flight hours. Just below the top there are accidents, which

![Figure 6.1](image)

**Figure 6.1** Complementary simulation tools are required to evaluate weak emergent behaviour along the flank and at the top of the safety pyramid of Heinrich (1931)

*Source: Blom (2013).*
happen in the order of 1 per 10 million flight hours. At the top you have mid-air collisions, which may happen in the order of 1 per billion flight hours. The ratio between the event frequencies at the top versus those at the base is in the order of $10^{10}$. This gap is bridged by agent-based safety risk analysis, which is explained in the next section.

6.2 Agent-based Emergent Behaviour Analysis

6.2.1 Agent-based Modelling and Simulation of Complex Socio-technical Systems

ABMS is increasingly recognised as a powerful approach to understanding complex socio-technical systems exhibiting emergent behaviour (Holland, 1998). This method has the ability to identify and analyse known and unknown emergent behaviours (Chan et al., 2010). It can represent important phenomena resulting from the characteristics and behaviours of individual agents and their interactions (Railsback and Grimm, 2011). Burmeister et al. (1997) discuss the benefits of using an agent-based approach in domains that are functionally or geographically composed of autonomous subsystems, where the subsystems exist and interact in a dynamic environment. According to Burmeister et al. (1997), agent-based modelling can be used to structure and appropriately combine the information into a comprehensible form. For a large, complex system such as a traffic system, agent-based modelling provides the tools to analyse, model, and design the whole system in terms of its subsystems, each with its own set of local tasks and capability. The integration of the subsystems can then be achieved by modelling the interactions between the subsystems. So, agent-based modelling provides abstraction levels that make it simpler and more natural to deal with the scale and complexity of problems in these systems. Agent components can be described at a high level of abstraction, yet they support a systematic compositional modelling approach. Burmeister et al. (1997) conclude that agent-based modelling reduces the complexity in systems design by making available abstraction levels that lend themselves to a more natural way of modelling. In the same vein, Jennings (2000) argues that ABMS and complex system development requirements are highly compatible. He shows that agent-based modelling techniques are particularly well suited to complex systems because: a) they provide an effective way of partitioning the problem space of a complex system; b) they provide a natural means of modelling complex systems through abstraction; and c) they capture the interactions and dependencies within the system.

Because ABMS is applied in many different domains, there are multiple definitions of ‘agent’. Among the various definitions are:

- an agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors (Russel and Norvig, 2009);
• an agent is an autonomous system situated within a part of an environment, which senses that environment and acts on it, over time, in pursuit of its own agenda and so as to effect what it senses in the future (Franklin and Graesser, 1997);

• an agent is a computer system that is situated in some environment and that is capable of autonomous action in this environment in order to meet its delegated objectives (Wooldridge, 2009).

In air transport, where different actors, hardware, and software are interacting elements of a complex socio-technical system, we consider agents as entities that are able to perceive and act upon (part of) their environment (see Figure 6.2). These agents may be humans, systems, organisations, and any other entity that pursues a certain goal. For instance, an air traffic controller can be viewed as an agent perceiving his/her environment (displays, alerting systems, runway availability, etc.) and acting upon this environment (for example, through communicating with other agents like pilots or other controllers, or turning off runway stop-bars remotely). The environment of agent \( i \) is understood as everything surrounding agent \( i \).

6.2.2 Agent-based Modelling of Hazards

Traditional safety studies assume well defined cause–effect links that propagate the effects of events contributing to the safety risk (e.g., sequential or epidemiological safety models). Recent views indicate that such models may not adequately represent the complexity of modern socio-technical systems (Hollnagel and Woods, 2005; Hollnagel et al., 2006). Instead, ABMS forms a logical choice for the safety-risk analysis from a socio-technical perspective (Blom et al., 2001a, 2001b, 2006; Stroeve et al., 2009). By distinguishing a number of agents and their interactions,
the overall process can be considered to be emerging from the individual agent processes. This not only provides a transparent way of structuring the model, which supports the analysis both conceptually and computationally, but also makes the model easier to maintain, resulting in local model refinements instead of global changes. Stroeve et al. (2013) made a systematic comparison of the agent-based approach against a sequence-based approach for a runway crossing operation. This study revealed many advantages of the former approach, including considerable differences in the risk results obtained. The main disadvantage is that agent-based safety risk analysis requires computational modelling expertise that differs from the expertise of traditional safety analysts.

In safety risk analysis the multi-agent model is coded in a language which includes the capability to generate random numbers. This allows a computer to run a large number (N) of different simulations with the agent-based model of the operation considered. This is called ‘Monte Carlo simulation’. By counting the number C of crashes over all of these N runs, one gets an estimated crash probability of C/N per run. One of the clear advantages is that Monte Carlo simulation provides a much more powerful approach in handling the combinatorially many possible event sequences. In a classical risk assessment approach, the analyst needs to identify the possible event sequences prior to systematic quantification of risk. With Monte Carlo simulation, this is not required. Instead, the agent-based model is Monte Carlo simulated to assess the probabilities with which particular event sequences and outcomes occur. Monte Carlo simulation has another advantage: for the catastrophic outcomes of a simulation run, one can look back and see how the trajectories evolved prior to the crash. By doing so for a sufficient number of Monte Carlo simulated crashes, one can gain insight into exactly what happens along the slope of the safety pyramid.

All kinds of hazards and non-normal events play important roles. Recently a systematic study has been completed regarding the agent-based modelling of the various kinds of such hazards (Bosse et al., 2013), and the importance of various sub-models has been assessed in terms of the percentage of hazards that can be captured (Blom et al., 2013). Table 6.1 presents the five most important sub-models, revealing a remarkable aspect of agent-based modelling in air transport.

<table>
<thead>
<tr>
<th>Top 5 sub-models</th>
<th>% of hazards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-agent situation awareness differences</td>
<td>41.4</td>
</tr>
<tr>
<td>Technical systems modes (configurations, failures)</td>
<td>19.9</td>
</tr>
<tr>
<td>Basic human errors (slips, lapses, mistakes)</td>
<td>18.0</td>
</tr>
<tr>
<td>Human information processing</td>
<td>14.3</td>
</tr>
<tr>
<td>Dynamic variability (e.g. aerodynamics)</td>
<td>8.6</td>
</tr>
</tbody>
</table>
The dynamic variability sub-model captures, for example, the dynamic movement of aircraft in the form of a set of differential equations. It captures 8.6 per cent of hazards. This sub-model is often used in aviation simulation studies. The human information processing model (Wickens and Holland, 2000) captures 14.3 per cent. This model is also often used in aviation, for example, for the simulation of human performance (Corker et al., 2008; Foyle and Hooey, 2008). Basic human error models for slips, lapses and mistakes (Reason, 1990) are widely used in classical risk analysis across all safety-critical domains, including the nuclear and chemical industries. Modelling these human slips, lapses and mistakes captures 18 per cent of hazards. Technical systems modes, including system configurations and system failures, are also widely used in classical safety risk analysis. This sub-model captures 19.9 per cent of hazards.

The highest-ranking sub-model is multi-agent situation awareness (MA-SA) differences, at 41.4 per cent, which is more than twice the percentage of the second ranked sub-model. The MA-SA sub-model (Stroeve et al., 2003) is an extension of the situation awareness model of Endsley (1995). The extension allows the possibility that agents in the ATM system may exacerbate differences in situation awareness while having no means to recognise that these differences exist. This is comparable to what happens in the game of Chinese whispers. In contrast to Chinese whispers, it is not only human agents that contribute to this propagation but also technical systems agents. Fortunately these multi-agent situation awareness differences do not often enter into the current ATM system unnoticed. If they do, this can lead to high-risk propagation of these differences to other agents.

A simple example of multi-agent situation awareness difference propagation is ‘Level bust’. A pilot of aircraft A receives an instruction from his/her air traffic controller to climb to an altitude level of 31,000 ft. Assume that the pilot mishears the instruction as 32,000 ft and enters this into the flight management system. The flight management system will level-off aircraft A at 32,000 ft instead of the 31,000 ft that is expected by the air traffic controller. The air traffic controller may have previously instructed aircraft B to fly at 32,000 ft near the intended level-off point of aircraft A, and is now forced to instruct aircraft B to deviate to, say, 33,000 ft, to avoid a potential collision. In this example the situation awareness difference sneaks in during the communication between the pilot of aircraft A and the air traffic controller. This then propagates to a situation awareness difference between the controller and the flight management system of aircraft A. In the current ATM system the propagation of these differences is only noticed when aircraft A does not level off at 31,000 ft. At such a late stage, there is little time left to avoid a potential collision.

Although ‘Level Bust’ is well known in ATM, the idea of capturing this phenomenon through multi-agent situation awareness difference propagation modelling is not. This kind of multi-agent situation awareness difference propagation appears to apply in a significant percentage of commercial aviation accidents. Hence, this approach could enable the prediction of high-risk situations in a future air transport design.
6.2.3 Integration of Mathematical Methods

Because the timescales of events at the top and bottom of the safety pyramid are very different, a straightforward Monte Carlo simulation of an agent-based model might take a lifetime. The solution is to integrate agent-based modelling with dedicated mathematical tools. This approach has become popular in financial mathematics and in particle physics, but its application for safety risk analysis is an innovative development. Table 6.2 provides an overview of mathematical tools that have been usefully integrated with agent-based models.

One of the key mathematical tools is the formalism of stochastically and dynamically coloured Petri nets (SDCPNs) (Everdij and Blom, 2005, 2010). This mathematical formalism specifies a model which assures a one-to-one connection between the agent-based model and certain stochastic process properties. The formalism supports a one-to-one relation with the evolution equations of Fokker-Planck-Kolmogorov (Bect, 2010; Krystul et al., 2007) and with the theory of probabilistic reachability analysis for stochastic hybrid systems (Blom et al., 2007, 2009; Bujorianu, 2012; Prandini and Hu, 2006).

The SDCPN formalism supports specifying and composing interacting infrastructure networks hierarchically and allows their stochastic analysis and Monte Carlo simulation. It has been proven by Everdij and Blom (2005, 2010), that a process generated using the SDCPN formalism (e.g., through Monte Carlo simulation) is mathematically equivalent to a generalised stochastic hybrid process (Bujorianu, 2012). Therefore, processes generated using stochastically and dynamically coloured Petri nets inherit the stochastic analysis powers of generalised stochastic hybrid processes, one of which is the strong Markov property (Krystul et al., 2007). These powerful relationships provide a sound mathematical basis for accelerated convergence of Monte Carlo simulations.

In Table 6.2, importance sampling, conditional Monte Carlo simulation, and the interacting particle system approach are all examples of accelerated Monte Carlo techniques. Importance sampling is a widely used technique for rare event simulations (Asmussen and Glyn, 2007; Bucklew, 2004). Conditional Monte Carlo simulation consists of decomposing catastrophic risk simulations into a

<table>
<thead>
<tr>
<th>Table 6.2</th>
<th>Mathematical methods integrated with agent-based modelling and simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stochastically and dynamically coloured Petri nets</td>
<td></td>
</tr>
<tr>
<td>Fokker–Planck–Kolmogorov evolution</td>
<td></td>
</tr>
<tr>
<td>Importance sampling</td>
<td></td>
</tr>
<tr>
<td>Conditional Monte Carlo simulation</td>
<td></td>
</tr>
<tr>
<td>Interacting particle system</td>
<td></td>
</tr>
<tr>
<td>Sensitivity/elasticity analysis</td>
<td></td>
</tr>
<tr>
<td>Uncertainty quantification</td>
<td></td>
</tr>
</tbody>
</table>
sequence of conditional Monte Carlo simulation problems and combining the results of these conditional simulations into the assessed risk value. The integrated particle system approach supports probability estimation of a rare event by introducing a sequence of intermediate event conditions that always precede the event of interest. The rare event probability is determined as the product of the conditional probabilities of reaching these intermediate events. The conditional probabilities are estimated by simulating in parallel several copies of the process, that is, each copy is considered as a particle following the trajectory generated by the process dynamics (Blom et al., 2007). Cérou et al. (2006) have proven that under certain conditions this approach yields unbiased risk probabilities, which distinguishes the integrated particle system approach from the popular Restart method (Villén-Altamirano and Villén-Altamirano, 2002). The main condition that is required to ensure unbiased estimation is that the simulated process must have the strong Markov property. Developing a model using stochastically and dynamically coloured Petri nets ensures that the generated process satisfies this strong Markov requirement.

Uncertainty is inherent to safety risk analysis. It can be analysed and quantified using mathematical tools including sensitivity/elasticity (log sensitivity) analysis. To assess the bias and uncertainty in the modelled risk level, Everdij et al. (2006b) developed the following stepwise approach:

1. **Identify potential differences between the simulation model and reality.** This concerns differences between: i) the values assumed in the simulation model and the real parameter values; ii) differences between the model structure and the structure in reality; iii) differences due to hazards that are not modelled; and iv) differences between the operational concept modelled and real operations.

2. **Assess the size/probability of the differences.** For each parameter, a bias factor and a corresponding uncertainty interval are assessed. For other types of differences, a value is assessed for the probability that the difference applies in the case considered.

3. **Assess the elasticity (log-sensitivity) of assessed risk level for changes in parameter values.** Additional Monte Carlo simulations are conducted with the model to assess the elasticities (log-sensitivities) of the assessed safety risk to changes in its parameter values.

4. **Assess the effect of each potential parameter value difference on the risk outcome.** The bias and uncertainty intervals of each parameter are combined with the risk elasticities.

5. **Assess the effect of the non-parameter differences.** For the non-parameter types of differences, a conditional risk bias given the difference exists is assessed and this is combined with the probability that the difference exists (step 2).

6. **Determine the joint effect of all differences.** The joint effect of all differences on the bias and uncertainty interval of the risk is determined.
The integration of these mathematical techniques within ABMS has been shown to be highly effective in ATM safety risk analysis (Stroeve et al., 2013), which has recently been recognised by the FAA and EUROCONTROL as a valuable approach in evaluating advanced ATM developments (EUROCONTROL/FAA, 2014; Fota et al., 2014).

6.3 Case Study: Airborne Self-separation

6.3.1 Advanced Autonomous Aircraft (A3) Operation

This section illustrates the application of agent-based safety risk analysis to an advanced airborne self-separation design. The idea of airborne self-separation (or free flight) is that pilots manage separation using an airborne separation assistance system (ASAS). Although free flight has been researched since its ‘invention’ (RTCA, 1995), a dispute has continued between two schools of researchers. One school believes that the emergent behaviour and properties of free flight are such that it can safely accommodate high traffic demand. The other school disagrees. Positive pilot findings (Ruigrok and Hoekstra, 2007) during human-in-the-loop simulations of free flight have not resolved this dispute. There is a need for a systematic understanding of rare emergent behaviours and the safety macro properties of free flight. To this end, Blom and Bakker (2011, 2012, 2013, 2014) have conducted an agent-based safety risk analysis of an advanced airborne self-separation design; this forms the background of the case study presented in this section.

The airborne self-separation design considered is the advanced autonomous aircraft (A3) design (Cuevas et al., 2010; Gelnarová and Casek, 2009). This A3 design is largely based on an advanced free flight design of NASA (2004), which was later published in a conference paper (Wing and Cotton, 2011). The related NASA design has been shown to work well in pilot-in-the-loop simulations (Consiglio et al., 2010). The A3 design has a four-layered architecture:

- strategic flow control layer;
- TBO layer;
- tactical conflict resolution layer;
- collision avoidance layer.

The strategic flow control layer assumes a centrally organised control of air traffic flows such that local traffic demands will stay below certain limits that can safely and efficiently be accommodated by the next layers. In the TBO layer, each aircraft determines a four-dimensional (4D) trajectory plan that aims to be conflict-free from the 4D trajectory plans of other aircraft. This 4D trajectory plan is broadcast to all other aircraft. If there are conflicts between broadcast 4D trajectory plans, the aircraft involved will be triggered to iterate until the overall situation is conflict-free. The tactical conflict resolution layer aims to resolve any short-term conflicts through tactical course changes. This is needed if an aircraft
deviates too much from its 4D trajectory plan, or if the iteration to conflict-free trajectory plans has not been completed in time, for example. The collision avoidance layer works as a last resort if a serious conflict is unavoidable, but a collision can still be avoided through a rapid climb or descent.

Each of these layers involves many interacting micro processes with the involvement of all flights concerned. This leads to macro behaviour at each of these four levels, that is, behaviour that cannot be observed at the micro process level. There is flow control-induced macro behaviour at a strategic level, TBO-induced macro behaviour at the medium-term planning level, tactical conflict resolution macro behaviour at the short term level, and airborne collision avoidance macro behaviour at the last resort level. At each level humans play a key role in the decision-making, and there is feedback from the macro behaviour at one level to the micro-behaviour at another level. This makes the A3 concept design a nice example of Fromm’s type III emergent behaviour.

Of these four layers, the first (strategic flow control) and the last (collision avoidance) exist, but the two layers in the middle differ a lot from current ATM. For this reason it was decided to perform an agent-based safety risk analysis for these two middle layers. Figure 6.3 gives a high-level overview of the agent-based

![Figure 6.3 Agents in the trajectory-based operation and tactical conflict resolution layers in A3](source: Blom (2013).)
model developed, which shows the relevant agents in these two layers, as well as their main interactions. These interactions include deterministic and stochastic relationships, as is appropriate for the agents considered. The agent-based model has been developed in a hierarchical way. For each aircraft there is an agent representing the aircraft itself, an agent representing the guidance, navigation and control (GNC), an agent for the ASAS, an agent for the pilot flying (PF), and an agent for the pilot not flying (PNF). Agents representing global communication, navigation and surveillance (CNS) and the environment interact with all of the aircraft. From a multi-agent perspective the two middle layers are not recognisable as separate architectural layers.

Following the development, implementation and verification of this agent-based model, Monte Carlo simulations have been conducted. Figure 6.4 shows a top view of an example outcome of a single Monte Carlo simulation run for a scenario of eight conflicting aircraft. Opposing aircraft in this scenario start at 135 nautical miles (nm) apart; the straight lines joining each aircraft’s origin and its

Figure 6.4  Example of eight encountering aircraft trajectories realised under the A3 design

Note: The aircraft started at the locations of the diamonds and stopped shortly before reaching the opposite diamond. The circle in the middle has a diameter of 10 nm (18.5 km).
Source: Blom (2013).
destination intersect at the centre. Without a functioning ASAS this would be unsafe.

### 6.3.2 Rare-event Monte Carlo Simulations

The key rare-event Monte Carlo simulation results of Blom and Bakker (2011, 2012, 2013, 2014) are obtained for random traffic scenarios with traffic demand multiple times greater than that in a busy en-route sector on a busy day in 2005. To accomplish this through Monte Carlo simulation of a limited number of aircraft, use has been made of a periodic boundary condition (Rapaport, 2004). The key outcome of these rare-event Monte Carlo simulations is given in Figure 6.5. The curve shows the occurrence probabilities per flight hour of miss distance values occurring in the range from 6 nm down to 100 m. The curve starts at a miss distance value of 6 nm at a probability value of 1, and then goes down when the miss distance gets smaller. The curve in Figure 6.5 consists of four parts.

![Figure 6.5 Estimated event probability per aircraft per flight hour for random traffic under A3 model control at traffic demand of 3x high en-route traffic demand in 2005](image)

Note: The bracket I shows the frequency of a miss distance below 3.33 nm in UK en-route traffic (NATS, 2011).

• The first part is a horizontal line from 6 nm to 5.3 nm at a probability level of 1.
• The second part of the curve covers the steep slope between 5.3 nm and 4.3 nm.
• The third part of the curve is an almost horizontal line from 4.3 nm to ~1 nm.
• The fourth part of the curve bends down between miss distances of ~1 nm and 100 m.

During the first and third parts of this curve the event probability per flight hour hardly decreases with the miss distance. In the second and fourth parts of the curve the event probability is clearly going down. The reduction during the fourth part is known as the ‘Providence’ factor in mid-air collision risk. The fixed level of the third part reflects the limited dependability of the ASAS supporting systems. By far the largest reduction happens during the second part of the curve; this corresponds to the tactical conflict resolution layer. At 5.3 nm the second part of the curve starts bending down, then goes through a rapid fall around 5.0 nm, and continues to fall until the dependability of the ASAS supporting systems is reflected in the third part of the curve.

Figure 6.5 also shows a reference point in the form of a bracket I, representing the frequency of current events in controlled UK airspace for which the miss distance between aircraft is below 3.33 nm (66 per cent of the applicable minimum separation criteria) (NATS, 2011). The position of this reference point above the curve shows that for a traffic density triple the highest density observed in 2005, the simulated A3 design performs much better than current operations.

The curve in Figure 6.5 does not show a distinct contribution from the TBO layer. This raises the question whether the tactical layer can produce the same results without a functioning TBO layer. To verify this, the rare-event Monte Carlo simulation has been repeated for a situation in which the TBO layer is not properly working. The sharp second part of the curve in Figure 6.5 appeared to completely disappear, and instead a slowly decreasing second part of the curve remained (Blom and Bakker, 2013). This shows that an effective TBO layer is a prerequisite for the tactical conflict resolution layer to deliver the very sharp second part in the curve of Figure 6.5.

It is standard practice that the distance between centre lines of 4D trajectory plans does not become smaller than the minimum horizontal separation criterion (which is 5 nm) plus an extra buffer of several nautical miles to allow for typical navigational deviations from the centre line. However, the curve in Figure 6.5 has been obtained for situations in which the distance between centre lines of 4D plans is equal to the minimum horizontal separation criterion of 5 nm, i.e., there is no extra buffer. This means that for the A3 design, the extra buffer for typical navigational deviations is not required.

While the accuracy of wind forecasts has improved in recent years, occasional large errors can significantly affect the performance of trajectory prediction tools.
Blom and Bakker (2012, 2014) have shown the impact upon the risk curve in Figure 6.5 of systematic wind field prediction errors of up to 30 ms$^{-1}$ (60 kt); this slightly reduces the sharp downfall between 5 nm and 4 nm and moves the curve 1 nm closer to the frequency bracket I for current operations. However, this is much better than the 3 nm reported by Consiglio et al. (2009) for the TBO layer alone, meaning that the combination of the TBO layer and the tactical conflict resolution layer handles wind field prediction errors of 30 ms$^{-1}$ – far better than the TBO layer can do alone.

Blom and Bakker (2013, 2014) have also found that no phase transitions occur if traffic flow is steadily increased from three times to six times the highest en-route traffic demands in 2005.

Finally, Blom and Bakker (2012, 2014) have provided a comparison with the future target level of safety values. This shows that the airborne self-separation TBO concept has the potential to deliver SESAR’s very high future safety targets. The tactical layer appeared to work unexpectedly well in managing uncertainties that are not resolved in time by the TBO layer. This is a very positive emergent behaviour that goes beyond the expectations of the A3 designers.

### 6.4 Conclusions

In complexity science a property or behaviour of a system is called ‘emergent’ if it is not a property or behaviour of the constituting elements of the system, but rather results from the interactions between them. This chapter has studied commercial air transport from the perspective of emergence. In order to accommodate the expected growth in commercial air transport, both in the US and in Europe, significant changes to its socio-technical system are being developed, including possible changes in the roles of pilots and controllers. For such complex socio-technical designs, it is essential to study and understand the effects of the novel interactions between the many individual elements in order to be able to identify and address their positive and negative emergent properties and behaviours.

In Section 6.1, challenges have been identified regarding the identification and analysis of emergent behaviours of a novel air transport design. This has been accomplished in three steps. First, an outline has been given of different perspectives of emergence in the literature. This started by explaining that an emergent property is a macro-property that cannot be a micro-property. Next, it was outlined that in the literature different types of emergence have been identified, ranging from resultant emergence, which can be predicted through analysis, through weak emergence, which can be predicted through simulation, to strong emergence, which cannot even be predicted through simulation. Commercial air transport displays various examples of each of these three types of emergence, such as resultant emergence from complex technical systems in an aircraft, safety culture
as a strong emergence example, and many socio-technical system examples of weak emergence, the most demanding of which are rare emergent behaviours that happen along the flank and at the top of the air transport safety pyramid presented.

This chapter has identified several methods of analysing weak emergence in a future air traffic design, including the safety risk property. Since human operators (e.g., pilots and controllers) continue to play a key role in such designs, ABMS has been identified as the key approach. The chapter has also determined that ABMS is in need of two complementary approaches: 1) a systematic agent-based modelling of hazards and disturbances; and 2) a combination of ABMS with rare-event Monte Carlo simulation methods. Various sub-models play a significant role here, in particular one that addresses differences in situation awareness between different agents. All sub-models are to be integrated and rare-event Monte Carlo simulations are to be used to analyse macro properties that happen at various frequencies along the slope of Heinrich’s (1931) safety pyramid.

The agent-based safety risk analysis approach has been demonstrated for an advanced airborne self-separation socio-technical design. Since its invention under the name ‘free flight’, this kind of application has led to a dispute between two schools of researchers. One school believes that free flight brings positive emergent behaviours. The other school believes the opposite. The specific free-flight design evaluated has a four-layered architecture, including a strategic flow control layer, a TBO layer, a tactical conflict resolution layer, and a collision avoidance layer. Each of these layers involves many interacting micro processes that lead to macro behaviour at each of the four layers. It has been explained how an agent-based model was developed for the TBO and tactical conflict resolution layers and their interactions. Subsequently, this model was used to conduct rare-event Monte Carlo simulations. The results clearly show that the advanced airborne self-separation design yields various positive emergent behaviours. The findings clearly support the school of believers in free flight.

In conclusion, emergent behaviours that can be predicted are of use in the design of a future socio-technical ATM system. It has also been demonstrated that weak emergent behaviour of future designs can be identified and analysed by the application of ABMS in combination with agent-based hazard modelling and rare-event Monte Carlo simulation. This opens valuable directions for follow-up research:

- to further develop ABMS tools for application in the socio-technical air transport system;
- to further develop and incorporate agent-based hazard modelling and rare-event Monte Carlo simulation in above-mentioned ABMS tools;
- to use the above tools for the evaluation of emergent behaviours of future ATM designs along the flank and at the top of the air transport safety pyramid.
Note

1 Chinese whispers is a game in which the first player whispers a phrase or sentence to the next player. Each player successively whispers what that player believes he or she heard to the next. The last player announces the statement to the entire group. Errors typically accumulate in the retellings, so the statement announced by the last player differs significantly, and often amusingly, from the one uttered by the first.
Chapter 7
Data Science

Dr Massimiliano Zanin, Dr Andrew Cook
and Seddik Belkoura

In this chapter, we explore the holistic concept of data science. Data science is concerned with the extraction (sometimes called ‘mining’) of meaning from data. Kantardzic (2011) explains how much of modern science is based on first-principle models to describe systems, starting with a basic model (such as Maxwell’s equations for electromagnetism, only later empirically proven), which are then verified (or otherwise) by experimental data to estimate some of the parameters. However, in many domains, such first principles are not known, and/or the system is too complex to be formalised mathematically. Through data mining, there is currently a “paradigm shift from classical modelling and analyses based on first principles to developing models and the corresponding analyses directly from data” (Kantardzic, 2011). This often requires the use of machine-learning algorithms to identify patterns and relationships across data elements, even when those datasets are large, incomplete and noisy (Witten and Eibe, 2005).

As with complexity science, there is no universally agreed definition of data science, but the basic principle involves moving from data, which are of practically no use in their unrefined state and often in very large datasets, often via some automated process, to the stage of producing usable information that offers the users insights that were not available from the raw data alone. Some of the analogies with mining are indeed irresistible, for example, where very large volumes of raw materials need to be extracted, sifted and refined before we produce something as valuable as a cut diamond. As with complexity science, data science is also multidisciplinary, drawing on such overlapping domains as mathematics, statistics, programming, modelling, machine learning, indexing techniques, stream data processing, data warehousing and evaluation environments, scalable analytics and data visualisation. Data science also has commonalities with complexity science, such as through applied network theory and the use of tools from statistical physics.

The reader may also have encountered the terms ‘knowledge discovery’ and ‘knowledge discovery in databases’ (KDD), which usually refer to drawing on big data and data visualisation challenges. This is intimately linked with data mining, and, as the name implies, places special emphasis on the extraction of patterns that can be considered as ‘knowledge’ – we return again to meaning. Both data mining and knowledge discovery comprise a core component of data science. They embrace several classical statistical techniques and a number of the methods we identify in this chapter – not least artificial intelligence (particularly machine
learning) supported by the development of large databases (Nisbet et al., 2009). Such tasks are concerned with data acquisition, pattern recognition, feature selection, clustering and classification, and often through to model deployment.

This chapter thus deals with techniques for extracting meaningful information from complex systems – specifically from data. It is to be hoped that in future, data science will go one step further, and deliver decision-making tools for ATM. We begin by discussing why classical approaches alone are unlikely to keep pace with the needs of analysts in this domain, before moving on to the central issue of understanding and establishing cause-and-effect relationships. We conclude with a discussion of some of the day-to-day practicalities of extracting, presenting and interpreting such data in a meaningful way, and the topic of big data: increasingly in the air transport spotlight.

7.1 Shortcomings of the Classical Approach for Data Metrics and Mining

Despite there being a blurred line separating data mining from classical statistical analysis, we will sketch out part of that line. In short, data science has a major computation/data processing component, unlike classical statistics, and is much more focused on understanding system dynamics. In this context, two important facts should be taken into account. Firstly, not all classical statistical theories are included in data mining. Secondly, data mining goes far beyond classical statistics. Let us start by recalling the difference between the frequentist and Bayesian views of probability. According to the former view, probability is just the frequency of an event, when a long series of trials is executed. This implies that the event may happen or not (more technically, that the (null) hypothesis is true or not): a probability is assigned to the set of events, not to a single event. On the other hand, Bayesian statistics interprets probability as a measure of how strongly one believes in the hypothesis. In other words, it attaches a probability to each event, and not to the whole set. This difference is one of the fundamentals of data mining. Let us suppose that we have a set of events, and that we want to construct a model capable of classifying them into two families, e.g., ‘safe’ and ‘unsafe’. According to the frequentist view, a probability can only be associated with the whole set, such that, for example, half of them would be safe, half unsafe – nothing can be said about a single event. A Bayesian statistician would try to associate a probability with each event, thus preparing the ground for a predictive model. Bayesian statistics should thus not only be included in data science, but indeed considered as its foundation, thanks to the many data mining algorithms that it has inspired.

While classical statistics remains important, it is not the only type of mathematics that should be well understood by the data scientist. For example, many inferential techniques and machine-learning algorithms depend heavily on linear algebra. Key data science processes are grounded in matrix mathematics and have much less to do with classical statistics. Overall, data scientists may need
substantial breadth and depth in their knowledge of statistics to manage this discipline. Sometimes, we may even encounter conflicts between classical schools of thought and data science, as we discuss later, regarding the use of p values with big data. Table 7.1 summarises five key techniques used in data mining. The list is non-exhaustive and we have focused on common examples.

**Table 7.1 Key techniques used in data mining**

<table>
<thead>
<tr>
<th>Technique</th>
<th>In more detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support vector machines</td>
<td>Given observations that belong to two different categories, the researcher often wants to decide to which of these two classes new observations belong. Considering the data as points in n-dimensional space, the algorithm searches the (n-1) hyperplane separating the two sets of data with the largest distance between them. This separator enables a straightforward classification of new points.</td>
</tr>
<tr>
<td>K-nearest neighbours</td>
<td>This method is similar to support vector machines, based on classifying a test point given others that are already classified. Given a set of points, variables in a multidimensional space, with the corresponding class to which they belong, a new test point is compared and classified using the most frequent group in k-nearest neighbours.</td>
</tr>
<tr>
<td>Decision tree</td>
<td>This enables the classification of processes, specifically the prediction of membership of objects or cases into categories. It is performed by excluding possibilities within a logical scheme. At each branching of the tree, a characteristic is tested, so that the possible outcomes (memberships) are reduced. This technique might be useful to identify the cause or the scenarios in which a particular feature appears.</td>
</tr>
<tr>
<td>Neural networks</td>
<td>This method is used to achieve pattern and sequence recognition, data processing (signal separation, noise identification), adaptive control, or time series prediction. More generally, it allows the researcher to infer a function from observations, i.e., to assess the relationships between stimuli and the global response of the network. It is based on the connections between a given set of nodes, which can be altered to achieve a desired signal.</td>
</tr>
<tr>
<td>Bayesian networks</td>
<td>This method is used to infer causes or effects given some known variables (e.g., disease state from symptoms presented) as well as the probabilistic relationships between some of them. A Bayesian network comprises a set of statistical variables on top of the nodes of a network. These variables could be observables, hidden (latent) variables, or unknown parameters or hypotheses. The probability of a child node codifying a variable to be positive or negative is a function depending exclusively on its parents, as in a Markov process. Bayesian networks are useful for computing the probability for a series of outputs given some inputs, or the probability of having a set of inputs, hidden variables or hypotheses, after having observed a set of outcomes.</td>
</tr>
</tbody>
</table>
Classical models are not often well suited to the exploration of complex systems. Holistic, network treatments are required, rather than bottom-up, disaggregate approaches and extrapolations of averages over multiple scales. Relationships between different phenomena, including cause–effect relationships, are often overlooked in a purely theoretical analysis. Indeed, as we saw in Chapter 6, emergence in complex systems produces behaviours at the macroscale that cannot be predicted by means of the analysis of each single element at the microscale. By extending and complementing this framework, considering the system as a complex network, we may significantly advance the state of the art by embracing a new set of analytical methods and corresponding metrics.

We introduce metrics in Chapter 2 and encounter them in Chapter 3 (where we discuss their fitness for purpose) and Chapter 5 (in the specific context of resilience). Figure 7.1 shows a metric classification presented by Cook et al. (2013a). The term ‘classical’ metrics is used to denote those that are pre-defined, univariate (draw on one variable in the data) and do not use complexity science techniques. Some of these types of metric are already commonly in use, such as average aircraft delay and (even) reactionary delay. ‘Non-classical’ metrics defines both ‘complexity’ metrics (drawn from complexity science), and ‘derived’ metrics, which are in contrast to the classical metrics in that they are not (fully) pre-defined but are derived from the data iteratively and are typically multivariate – often outputs from data reduction techniques such as a factor obtained from factor analysis. We introduced several complexity metrics in Chapter 2. Figure 7.1 shows that these relationships are not wholly mutually exclusive; it could indeed be configured in subtly different ways. As with defining data science and classical statistics, the boundary between categories of metrics is not always well defined. Such boundary regions are often the most interesting in science. For example, how well do non-complexity metrics, particularly derived ones (e.g., factors from factor analysis), capture certain features of ATM system dynamics (e.g., delay propagation) compared with complex network theory (CNT) metrics? Due to the very nature of derived (data-driven) metrics, we are more likely to make unexpected findings, and to deepen our understanding through being prompted

![Figure 7.1 Metric classifications](image-url)
(forced, one may argue) to explain counterintuitive results, than we would through the use of classical metrics alone. In Section 7.3 we develop this discussion further, with a specific focus on the use of metrics in the performance assessment context.

7.2 Tackling Cause and Effect in the Complexity Context

In order to establish causal relationships between data it is necessary to turn to rather more powerful techniques than those simply describing associations. Classical statistical instruments, like, for instance, correlation analysis, are only able to assess the presence of some common (equivalent) dynamics between two or more systems. However, correlation does not imply causality. In what follows, we review why this point is important in real systems analyses, and how it can be addressed.

Figure 7.2 shows a typical situation, in which the researcher can observe two systems, $B$ and $C$, while another system is present but cannot directly be observed ($A$, in grey). The ideal situation would require $B$ and $C$ to be isolated, such that the dynamics of the latter are only ‘forced’ by the former (continuous arrow in the left panel). In this case, a correlation between $B$ and $C$ indeed points towards the presence of causality between these systems. Nevertheless, life is seldom this simple. The right panel of the same figure shows a more common situation, in which the hidden system $A$ is forcing both $B$ and $C$. In this case, the correlation we observe between the latter is just that, a correlation, and only the spurious outcome of the hidden forcing. Due to this, it is important to distinguish between three distinct forms of relationship:

- **Correlation** This is defined as a departure of two or more random variables from independence. Nevertheless, it usually means linear correlation, that is, the presence of a linear relationship between two variables.
- **Synchronisation** This occurs when two systems evolve in a coordinated manner through time. Thus, synchronisation can be seen as a kind of correlation developing through time. Furthermore, synchronisation does not

![Diagram](image.png)

**Figure 7.2** The dynamics of causality and correlation
require identical dynamics; on the contrary, the two systems may evolve in different ways, while their dynamics are nevertheless tied in some non-trivial way (a concept known as \textit{generalised synchronisation}). See Pikovsky et al., 2001, and Strogatz, 2002, for further details.

- \textbf{Causality} This is the relationship between an event (the \textit{‘cause’}) and a second event (the \textit{‘effect’}), where the second event is understood as a physical consequence of the first. (In ‘normal’ (e.g. non-subatomic) systems the effect occurs second – we ignore the strange world of quantum dynamics!)

In recent decades, numerous metrics have been proposed to detect the presence of such relationships between sets of data representing the evolution of (real) systems. In Table 7.2, the most important are shown, grouped according to the type of relationship they are able to detect: i.e., linear and non-linear. Due to its simplicity and wide range of application, we present a short description of Granger causality. Beyond describing its characteristics, our aim is to introduce the reader to the problem of detecting causality, and thus of better defining what causality is. Granger causality is held to be one of only a few tests capable of detecting the presence of causal relationships between time series. The test is an extremely powerful tool for assessing information exchange between different elements of a system, and understanding whether the dynamics of one of them is led by the other(s). It was originally developed by Nobel Prize winner Clive Granger (Granger, 1969) and although it was applied largely in the field of economics (Hoover, 2001) it has received a lot of attention in the analysis of biomedical data (Brovelli et al., 2004; Kaminski et al., 2000; Roebroeck et al., 2005). The two axioms, on which this test is based, are stated in Table 7.3.

\begin{table}[h!]
\centering
\caption{Metrics for relationships between sets of data}
\begin{tabular}{lll}
\hline
 & \textbf{Non-linear} & \textbf{Linear} \\
\hline
\text{Correlation} & Mutual Information (Li, 1990) & Pearson’s linear correlation \\
\text{Synchronisation} & Lyapunov exponent (Pecora et al., 1997) & Frank and Althoen (1994) draw a contrast with causality) \\
 & Synchronisation likelihood (Stam and Van Dijk, 2002) & \\
\text{Causality} & Transfer entropy (Staniek and Lehnertz, 2008) & Granger causality (Granger, 1969) \\
 & Symbolic analysis (Cánovas et al., 2011; Zanin et al., 2012) & \\
\hline
\end{tabular}
\end{table}
Therefore, a time series, \( q \), is considered to Granger-cause another time series, \( p \), if the inclusion of past values of the series \( q \) can improve the process of forecasting the values of the time series \( p \). In this case, the future evolution of \( p \) also depends on the past values of \( q \). Also, it should be noted that two time series presenting a high correlation, or two time series that are ‘forced’ by a third system, do not usually pass the Granger causality test: as they have similar values, one of them cannot convey useful information for the forecast of the other. In more detail, the test is calculated as follows. Firstly, we suppose that we are comparing data corresponding to two observables \( x \) and \( y \), as, for example, measures of controller workload and the appearance of safety events, with a high time resolution (e.g., observables are collected at ten-minute intervals). The first step involves the use of a standard forecasting technique on the time series \( y \), namely autoregression. In other words, we try to forecast one of the values of the time series by analysing all the values available before that moment in time. In the second step, the same process is performed, but also including data from the time series \( x \). Finally, both forecasts are compared: if the one obtained by using information about \( x \) is better than using the \( y \) forecast alone, a relationship between them exists. The statistical significance of this causality can also be assessed through an F-test (with a corresponding \( p \) value: which we discuss in Section 7.4; see Frank and Althoen, 1994, for a good introduction to the F-test).

Two important characteristics of Granger causality should be noted. First, the concept of causality, as defined by Granger, requires a time evolution: thus, it cannot be applied to static datasets or to observables describing different systems. Also, Granger causality is not symmetric: the analysis of \( y \) given \( x \) may return different values compared with an analysis of \( x \) given \( y \). This is, of course, in line with the concept of causality: if a system is controlled, or forced, by a second system, this relationship may not be mutual. Claims of causality from (multiple) bivariate time series should always be treated with caution, as true causality can only be assessed if the set of the two time series contains all possible and relevant information and action sources for the problem (Granger, 1980) – a condition that real-world experiments can only rarely satisfy (Zanin and Papo, 2013).

A final note of caution regarding the evaluation of a time series: here, the ‘stationarity’ of the data being studied is one of the most important analytical

<table>
<thead>
<tr>
<th>Axiom</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Causes must precede their effects in time.</td>
</tr>
<tr>
<td>2</td>
<td>Information relating to a cause’s past must improve the prediction of the effect above and beyond information contained in the collective past of all other measured variables (including the effect).</td>
</tr>
</tbody>
</table>
requirements, as a large number of techniques applied to such data require that, at some point, the data under study be stationary for the conclusions to hold. What does this mean? An observable is defined as stationary if its statistical properties (e.g., mean) are all constant over time. In such a case, it would not be possible to associate the mean of any sub-sample of the data with a timescale – that is, there is no trend in the data. While a physical time series cannot (usually) be proven to be stationary, the researcher may suppose the stationarity for a given timescale, make a transformation to make it stationary, or use techniques robust with respect to non-stationarity. A typical example is seasonal economic data, which need to be transformed (made stationary) before analysis.

7.3 ATM Performance Revisited: Network Layers

As introduced in Section 7.1, just as classical models are not well suited to the exploration of complex systems, we should not limit ourselves to the use of classical metrics in our endeavours to better understand ATM. As identified, most of the metrics currently in use in air transport performance assessment may be described as classical metrics, such as average delay. The metrics and techniques we use to assess any system are largely determined by the way in which we think about that system and the framework we use to formalise it. As we saw in chapters 2 and 3, a natural way of depicting the air transport system is by means of networks, where nodes represent airports and pairs of them are connected whenever some kind of relationship pertains – from the presence of a flight, up to more complex delay propagation relationships (Zanin and Lillo, 2013).

We also comment in Chapter 2 that in recent years researchers have realised that interactions between the constituent elements of complex systems seldom develop by means of a single channel. For instance, in a social network, information exchange may happen orally, electronically, or even indirectly. People interact according to different types of relationships, like friendships and co-working: each affecting the type of information transmitted. Consequently, a full network representation may require different types, or ‘layers’, of links, and thus multi-layer network representations (Boccaletti et al., 2014). If we are to represent a real system, which is expected to have a multi-layer structure, by means of complex networks, it would then be insightful to understand the magnitude of the distortions created by the use of a single-layer representation. In other words, a fundamental question should be addressed: to what extent are single-layer networks representative of the dynamics occurring at different layers? It is known that neglecting the multi-layer structure, or working with a ‘projected’ network, may alter our perception of the topology and dynamics, leading to an incorrect understanding of the properties of the system (Buldyrev et al., 2010; Vespignani, 2010).

In this section, we review some results obtained in the case of the propagation of delays in the air transport network, considering the dimension created by multiple airlines.
As depicted in Figure 7.3, two possibilities for network reconstruction are available, when starting from the raw delay time series (bottom left). In one way (top left), we could average all of the dynamics corresponding to an airport: as this yields one time series per node, the result is the creation of a single-layer network, representing the projected dynamics. From another perspective (bottom right), we could directly create a multi-layer network by considering each airline as a layer. Afterwards, such a structure can be projected into a single-layer graph, thus creating a ‘topological’ projection.

Pursuing our initial question, we should ask when these two projections are really relevant to understanding the problem being examined, that is, the propagation of delays. Towards this objective, we can calculate a set of topological metrics, both in the layers constituting the multi-layer network, and in the two projections, and then compare them. Figure 7.4 shows these results. Specifically, panels A and B respectively represent the out-degree centrality and alpha centrality (see Table 2.3) of the most central node in the dynamical projection (horizontal dashed line), in the network corresponding to the projected topology (horizontal solid line), and in all layers of the multi-layer representation (vertical bars). The two small graphs on the right respectively represent the same results for the most central node for the first and second layer. Furthermore, panels C, D and E respectively represent the (global) efficiency, clustering coefficient and entropy of the degree distribution (see Table 2.3) for the two projected networks and for the layers of the multi-layer representation.
Figure 7.4  Topological properties of an ATM multi-layer network
Adapted with permission from Zanin, 2015.
In each case, a strong disparity across the results is obtained. Specifically, the most important node in one network is seldom of major relevance in other networks. Also, topological features like (global) efficiency and the clustering coefficient take significantly different values depending on the type of projection considered. This has important consequences from an engineering point of view. For example, if one were to identify the most important node (airport) in terms of delay propagation, the use of different projections would yield different results, providing no guarantee that the most central node in the projected network is indeed relevant when the multi-layer dynamics are considered. The effects of changes in the topology of a network are presented in the next section. For the moment, the important message is that when studying the dynamics and performance of air transport networks: layers matter.

7.4 Data Integrity

7.4.1 The Classical Context

A discussion of air traffic data in the context of data science cannot be complete without reference to data cleaning. Here, we therefore take some time out of the pure complexity context, first of all, to summarise some important principles regarding data integrity in the classical context, before we develop the discussion into the complexity domain. We will take an example of the POEM project (introduced in Chapter 3), which used EUROCONTROL’s PRISME dataset detailing European flight movements. Cleaning of these data may be categorised into four major tasks.

1. Regarding aircraft identification, logic checks revealed coding inconsistencies between aircraft type, wake turbulence category and aircraft registration. These discrepancies were substantially resolved using International Civil Aviation Organization documentation and national aircraft registers.

2. Regarding airline identification, each aircraft operator needed to be categorised by its primary type of operation, i.e., full-service, regional, low-cost carrier or charter. Complications arose, however, due to unknown operators in the source data, the presence of non-commercial passenger airlines, wet leases and instances of the same aircraft flown by different operators on the same day.

3. Regarding time-related fields, all times had to be converted to local times to enable schedule checking. Daylight saving time adjustments, although synchronised throughout Europe, are manifested through many different implementations globally. A bigger problem was correcting the reported schedule times, for which independent schedule data were used for validation. It was particularly problematic identifying certain flights during significant delays and cancellations – for example, simplifying somewhat,
was flight ABC123 at 12:05 the more heavily delayed 10:00 or the almost on-time 12:00?

4 Taxi times were essentially missing from the data in most cases and had to be imputed from known distributions. In addition, numerous other data-cleaning tasks were required.

In total, 14 per cent of the flights had to be ‘demoted’ (deactivated with regard to metric inclusion, but operationally active) in the dataset, due to having one or more unreliable and unresolvable data fields. While the state of the art has moved on somewhat with newer generations of such traffic data (e.g., EUROCONTROL’s DDR2 service for flight, capacity and airspace data), the 14 per cent value may give the reader some idea of the potential size of this issue. We will return to this in a moment. It is also worth noting that special attention should be devoted to the problem of cleaning trajectory data. This problem has been tackled in several works (Zanin et al., 2011; Zanin, 2013; Zanin, 2014). Particular care should be taken regarding: radar trajectories that are not updated, that is, that just represent the planned trajectory; the presence of airspaces where radar data are not reported, thus creating ‘gaps’ in the trajectories, and incorrect time stamps, which may make an aircraft seem to fly at supersonic velocities, for example.

Consider the data in Table 7.4, which shows the number of flights from an airport by time of day, all delayed by somewhere between 15 and 20 minutes, and how much delay propagation they caused in the rest of the network. We have codified the latter as ‘low’, ‘medium’ and ‘high’, but the actual definitions used do not matter for the purposes of this illustration. We would expect a very large delay to cause more propagation in the network than a small one, but by focusing entirely on delays of a similar magnitude (15–20 minutes, say; any narrowish range would suffice) we can better examine whether there is a relationship between the time of day and the amount of propagation (at least for the main thrust of this illustration – it would be a little more complicated in practice). We might expect that delays in the morning would cause greater propagation, simply because there is more time left in the day for the delay to cause knock-on effects, before all (or, at least most) aircraft are at their intended stations by the end of the day. A common test to explore this type of relationship is the chi-square test. It works on the elegantly simple principle of comparing how many observations we have in each cell of Table 7.4 with how many we would expect, if there were no relationship between

<table>
<thead>
<tr>
<th>Scheduled time / Delay propagation</th>
<th>Morning</th>
<th>Afternoon</th>
<th>Evening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>10</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>Medium</td>
<td>10</td>
<td>60</td>
<td>30</td>
</tr>
<tr>
<td>High</td>
<td>10</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>
time of day and propagation effect. There are 60 flights in the first row (causing low propagation), 30 flights in the morning (first column), and 220 flights in total (whole table). If there were no relationship between these two variables, by simple proportions, we would expect \(60/220 \times 30 = 8.18\) flights in the top-left cell, i.e., morning flights causing low propagation. The actual count of such flights is 10 (Table 7.4). We can build up all these expected numbers, by the same simple principle, to produce Table 7.5 (note that the total number of flights here is also 220).

The chi-square test, then, basically uses the differences between the actual and expected values to determine if there is a relationship between the two variables (time of day and propagation intensity). The results, based on Table 7.6, give a \(p\) value of 0.004. Some readers will already know that this means that we would conclude that there is a ‘highly significant’ relationship between the two variables. At this point, it is hoped that those readers who do not know what the \(p\) value represents will be content with a rather simplified explanation thereof, and those who do will excuse us for presenting such a simplistic explanation. The \(p\) value can thus be (pretty loosely) explained as the probability of being wrong, if you were to decide that the two variables were dependent, given the observed data. (We can hear sharp intakes of breath from statisticians everywhere.) So, \(p = 0.004\) implies only a 0.4 per cent chance of being wrong, which we consider to be a pretty low chance, and so we conclude that the variables are indeed dependent. (We normally set thresholds of \(p < 0.05\) (‘significant’) and \(p < 0.01\) (‘highly significant’) to describe our results.) The reader interested in more about the chi-square test and \(p\) values may wish to consult further sources.¹

Now comes the rub. Imagine that there was some type of error in our data, and 15 of the 60 flights in the central cell of our table had been misclassified, and should have been in the cell above, making 35 flights in the afternoon causing ‘low’ delay

<table>
<thead>
<tr>
<th>Table 7.5</th>
<th>Delayed flights (expected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduled time / Delay propagation</td>
<td>Morning</td>
</tr>
<tr>
<td>Low</td>
<td>8.18</td>
</tr>
<tr>
<td>Medium</td>
<td>13.64</td>
</tr>
<tr>
<td>High</td>
<td>8.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 7.6</th>
<th>Delayed flights (differences)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduled time / Delay propagation</td>
<td>Morning</td>
</tr>
<tr>
<td>Low</td>
<td>+1.82</td>
</tr>
<tr>
<td>Medium</td>
<td>-3.64</td>
</tr>
<tr>
<td>High</td>
<td>+1.82</td>
</tr>
</tbody>
</table>
propagation, and 45 ‘medium’ (central cell). What effect would this have on our p value? After all, we have only reclassified just under 7 per cent (15 out of 220) of our flights. The result is that the p value soars up to 0.24, such that, statistically speaking, we would now be very sure that time of day and delay propagation propensity were not related variables! The ‘problem’ was with the distribution of the data, in that the central cell dominated the contribution to making the original p value low, as this cell produced quite a large difference between actual and expected values. What lessons can we draw from this example? Here are some of them:

- statistical results (as expressed through p values, for example) are not the last word in understanding relationships: they need to be taken in context, especially with regard to the sample size (which itself is very much a matter of context – 220 would be luxuriously large in many natural science experiments);
- a set of results may be strongly driven by just one particular classification or subset of observations, which could easily be lower than the resolution of our data (recall above that 14 per cent of the flight data records were problematic, double the 7 per cent value in our example – an unseen problem in uncleaned data);
- all of this has still only established an association, rather than causality;
- even the original p value was not related to any consistent pattern in the data.

Let us look a little further at this last point. Table 7.7 is a repetition of Table 7.6, just showing the signs of the differences. If there were some overall pattern to the data, for example, some kind of take-home message that delays in the morning resulted in (or, more strictly, were associated with) greater delay propagation, then we would have wanted all the +s to be collected in one corner of the table, and all the −s in the other. Instead, we have exactly the opposite: no pattern at all. We have also thus demonstrated that a perfectly legitimate p value might lead us to a conclusion that is correct, but of no use. In fact, after picking through the bones of this result, we can now see that the original p value told us little more than that there were a lot of aircraft in the central cell! On the other hand, we cannot draw conclusions from data inspection alone (or very rarely, at least), no matter how good our data visualisation software is. We need to know: how large is large enough?

<table>
<thead>
<tr>
<th>Scheduled time / Delay propagation</th>
<th>Morning</th>
<th>Afternoon</th>
<th>Evening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Medium</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>High</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
</tbody>
</table>
Let us next consider a network of navigation points, as introduced in Chapter 3 and, specifically, the node degree $D(i)$, i.e., the number of connections of the $i$th navigation point with other navigation points, in a certain time window. Another relevant variable might be the ratio $r(i)$ between the number of aircraft $N_a(i)$ that actually fly over the $i$th navigation point and the planned number of flights $N_p(i)$ flying over it in a certain time window: $r(i) = N_a(i)/N_p(i)$. The ratio $r$ can be considered as a local measure of the predictability of the system. The Pearson correlation coefficient between the variables $D(i)$ and $r(i)$ might be used to test whether there is a relationship between the connectivity of a certain navigation point and its predictability. In other words, one might use the Pearson correlation coefficient (see Table 7.2) to understand whether or not major deviations from planned trajectories occur at navigation points that are highly connected. Suppose we have the situation as shown in Table 7.8.

One can easily compute the Pearson correlation coefficient, returning a negative value $r = -0.37$. This would indicate that when connectivity increases, predictability decreases. However, as it is, the Pearson correlation coefficient does not take into account the fact that some navigation points are poorly used while others experience higher traffic. Therefore, it would be more appropriate to weight the navigation points with their (planned) occupancies. This can be achieved by using a weighted Pearson correlation coefficient.

If we use as weight $W_i$ the number of flights $N_p(i)$ flying over the $i$th navigation point, then the weighted Pearson coefficient returns $r_w = +0.44$. So we are moving from a negative correlation of 37 per cent to a positive correlation of 44 per cent when we take into account the role of each navigation point as measured by actual traffic load. This would indicate that when connectivity increases, predictability increases as well. This may be interpreted as follows: when there is a high traffic load, air traffic controllers try to avoid changes in the aircraft trajectories.

In fact, one navigation point in our set is quite peculiar. On the one hand, it is the only point with full predictability; however, it is also a point with very low connectivity and traffic load. When we remove this navigation point, we have $r = +0.44$ and $r_w = +0.72$. While the unweighted Pearson correlation coefficient shows

<table>
<thead>
<tr>
<th>Navigation point</th>
<th>Degree</th>
<th>$N_p$</th>
<th>$N_a$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>30</td>
<td>13</td>
<td>0.43</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>20</td>
<td>10</td>
<td>0.50</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>25</td>
<td>15</td>
<td>0.60</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>60</td>
<td>45</td>
<td>0.75</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>15</td>
<td>5</td>
<td>0.33</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>15</td>
<td>4</td>
<td>0.27</td>
</tr>
</tbody>
</table>
a dramatic change, the change in the weighted Pearson correlation coefficient is
less pronounced. This shows again that not only can a relatively small change in
the input data change our results but also that taking proper account of data weight-
ings (here, the traffic load of the navigation points) may well reduce the impact of
peculiar data.

7.4.2 The Complexity Context

We started Section 7.4.1 with a purely classical example and moved on to one
containing aspects of CNT through the inclusion of degrees of nodes (in this case,
navigation points). Let us now extend such concepts to the fuller characterisation
of complex networks. Specifically, we may ask under what conditions the value
associated with a topological feature is really relevant, and whether such values
may change as a function of the way that data are sampled and analysed. Far from
being just a theoretical problem, changes in the perceived topology can yield
important differences regarding the corresponding dynamical processes at the
nodes, and our understanding thereof. Let us consider further the case of delay
propagation. A high clustering coefficient, for example, implies a random topology,
with groups of nodes densely interconnected. Airports will therefore have similar
centralities, and resources (such as spare aircraft to deploy during disruption)
should be equally shared among them. On the other hand, low clustering coeffi-
cients can indicate scale-free topologies: in such cases, resources should be con-
centrated in the few important hubs.

Reviewing the numerous research papers on the topology of air transport (see
Zanin and Lillo, 2013), a clear feature emerges: even considering the same sub-
network (e.g., the same country), very different values for some topological metrics
are reported. For instance, the clustering coefficient of the Italian network varies
between 0.07 (Guida and Maria, 2007) and 0.418 (Zanin et al., 2008). What is the
driver of these differences? The answer lies in the shape of the degree distribution
of the nodes. Figure 7.5 depicts the cumulative probability distributions for air-
ports, airlines and aircraft types, corresponding to a four-month sample of all Euro-
pean flights. Each panel thus reports the probability of finding an airport (airline,
or aircraft type) with more than a given number of flights, as a function of this latter
parameter. Especially in the latter two cases, there is a clear linear tendency in the
lower log–log plot, indicating that both distributions may be well described by a
power law and thus that no characteristic scale can be defined (see Chapter 3). Also,
in the case of airport operations, even if the distribution is more complex than a
power law, the strong decay characteristic of a Gaussian distribution is not pres-
ent. It is important to note that much research only considers a subset of the most
important (well connected) airports and airlines, for example. Yet such scale-free
distributions imply that, independent of the applied threshold, the discarded ele-
ments (nodes) account for a share of operations larger than their degree (see Table 2.3)
may suggest. In other words, we need to be particularly careful, since applying
different thresholds may yield completely different structures.
Figure 7.5  Cumulative probability distributions for airports, airlines and aircraft types
This is further confirmed by Figure 7.6, which shows the evolution of some common topological metrics as a function of the number of airports included in the network. Airports are sorted according to the number of their flights, and a link is established between pairs thereof when at least one flight is present (i.e., independent of the airline operating the flight, or of the aircraft type). It can be seen that, after an initial transient that includes 100–200 airports, all metrics evolve according to a linear or asymptotic trajectory.

This is also exemplified in Table 7.9, which compares the metric values for two networks composed of 100 and 1,000 airports, respectively. All metrics show strong variations, especially the link density (from 0.770 to 0.081) and the mean geodesic distance (from 1.23 to $\infty$ – indicating that the second network is disconnected).

In summary, it is important for the researcher to understand the implications of choosing the size and scope of the system under analysis (both in terms of the number of instances and the number of nodes and airlines considered, for example). Applying a threshold very often introduces bias to the way in which the system is represented, and thus how the corresponding processes are understood. As with the

![Figure 7.6 Topological features as a function of the number of nodes (airports)](https://example.com/figure7.6.png)
simple, classical example illustrated of delayed flights by time of day, relatively small changes in how we examine the data can dramatically change the results. Such differences may even yield conflicting results.

### 7.5 Big Data

Finally, and as we have introduced above, no chapter on data science would be complete without a section on big data. Although this is another topic of considerable and increasing depth and breadth, we hope to at least give some insights here into the field and its application in air transport in particular.

As with several other topics we have explored, there is no unique definition of exactly what constitutes ‘big’ data. It is something of a moving target, and from the hindsight position of a decade (or probably rather earlier) from now, it might seem rather less ‘big’. However, it may generally be applied to collections of data too large to process using traditional desktop and server tools. Furthermore, one Rubicon that once crossed, will remain crossed, is the associated transition from structured data (e.g., from passenger surveys), typically queried by Structured Query Language (SQL) to far more heterogeneous data (such as those generated by social media, video analytics, mobile sensors, website clickstreams and many other types of activity and monitoring logs) and far bigger datasets – exceeding the processing capacity of conventional database systems (O’Reilly Media, 2012). Indeed, there has been a phenomenal growth of these data: every minute, Google receives 2 million search requests, over 500 new websites are created, 200 million emails are sent, Apple has around 50,000 app downloads and there are 100,000 tweets on Twitter (Domo Inc., 2015). Large volumes have presented analysts with challenges for data often not amenable to a pure SQL approach, or even to SQL at all. (Such big datasets are often called ‘NoSQL’ databases, the ‘No’ usually meaning ‘not

<table>
<thead>
<tr>
<th>Topological metric</th>
<th>Top 100 airports</th>
<th>Top 1,000 airports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering coefficient</td>
<td>0.842</td>
<td>0.507</td>
</tr>
<tr>
<td>Degree correlation</td>
<td>−0.024</td>
<td>0.0148</td>
</tr>
<tr>
<td>Efficiency</td>
<td>0.885</td>
<td>0.494</td>
</tr>
<tr>
<td>Entropy of the degree distribution</td>
<td>0.777</td>
<td>0.744</td>
</tr>
<tr>
<td>Link density</td>
<td>0.770</td>
<td>0.0813</td>
</tr>
<tr>
<td>Maximum degree</td>
<td>0.990</td>
<td>0.500</td>
</tr>
<tr>
<td>Mean geodesic distance</td>
<td>1.23</td>
<td>∞</td>
</tr>
<tr>
<td>Normalised information content</td>
<td>0.754</td>
<td>0.440</td>
</tr>
</tbody>
</table>

*Note:* Values to three significant figures. See Table 2.3 for definitions.
only’, as opposed to simply ‘not’). The supply of such data often outstrips storage capacities, let alone most (classical) analytical and integration capabilities, which present particular demands on open architectures. The availability of such data has also posed new challenges for intelligent management (such as pre-filtering), plus data mining and visualisation methods, if the user is to realise primary goals such as significantly improved business intelligence and predictive analytics, often supporting the development of new products and services, and better behavioural modelling. Mainstream analytical tools are unable to manage the volume, granularity and real-time velocity that big data present.

Many institutions have thus had to deal with the continuously growing volume of data, requiring more and more storage space and processing power. Data processing has quickened with parallel processing (DeWitt and Gray, 1992), put-data-in-memory (Herlihy and Moss, 1993) and distributed computing (Zaharia et al., 2012). This has also fostered the development of cloud computing (a specialised form of distributed computing) and hosted platforms (Mell and Grance, 2011), on the one hand, and the emergence of new array-based (e.g., SciDB; see also Brown, 2010), column store (e.g., ParAccel; see also Stonebraker et al., 2005) and NoSQL (e.g., Vertica; see also Cattell, 2011) databases, on the other hand. Meeting such challenges, cloud computing has opened up new possibilities for the aviation industry (Marks, 2014) and leveraging big data is becoming less and less constrained to those with access to more specialist resources.

To analyse big data, not only are faster algorithms necessary but two other challenges also present themselves: adapting significance testing and data visualisation. Regarding the former, and echoing central themes in Section 7.2, Granger (1998, 2003) has reminded us that p values will rarely be above the significance cut-off (alpha) level chosen (e.g., 0.05), for sufficiently large samples. In other words, a statistically significant p value will often be generated, although this might well be effectively meaningless and (primarily) due to the large sample size. In such outcomes, (practically) no information about the size or the importance of the effect tested is actually given by the p value (Cohen, 1994). Such p values should thus be complemented by supporting effect-size calculations, exploring the sensitivity of the p value to sample size (Jacob, 1990; Mingfeng et al., 2013; Orwin, 1983; Rosenthal, 1994) or the ‘many sets of data’ (MSoD) methodology (Ehrenberg, 1990; Murray and Ehrenberg, 1993) that addresses, inter alia, replicability. This involves, for example, repeating the study, introducing some variations, then assessing the generalisability of the results, their limitations and corresponding determinant factors. These themes are discussed with concise case studies and advice for researchers by Kennedy et al. (2014), including a discussion on MSoD, calling for more attention on empirically building prior knowledge and establishing patterns in wider contexts (facilitated by big data), and less on single studies (and associated p values).

Big data visualisation offers great potential (National Academies Press, 2013; Russom, 2011). This reaches far beyond simple graphics, which may result in uninformative, dense plots, due to the ‘cluttering’ effect (Amraii et al., 2014).
Instead, the visualisation of big data relies on intelligent sampling, segmentation and streaming techniques. Heat maps are a common example of an output of such a process – these often depict virtual data landscapes rather than pertaining to geographical representations, although below we cite an example of the latter. ATM is a field, in particular, where better data visualisation could be of enormous benefit to its practitioners, but the implementation of change is a slow one and tactical tools have to go through rigorous safety and certification procedures. In general, air transport and ATM are probably no further behind the development curve than most other industries, although certainly not as advanced as more pioneering fields in this respect, such as biology and biochemistry. Data streaming is the mechanism applied to maintain real-time processing and management of incoming data (e.g., from sensors) which cannot be stored. This can be associated with intelligent displays (starting to show the data before the entire file has been processed) and the predictive mining of continuous feeds (e.g., using machine learning). For further general discussion on big data, visualisation and analysis, the reader may refer to the useful publication by the National Academies Press (2013). We next turn more specifically to the air transport context, introducing several associated domains in Table 7.10.

Some airports are already using data visualisation methods (including heat maps) for geo-mapping passenger pathways through terminal areas, using

<table>
<thead>
<tr>
<th>Domain</th>
<th>Types of data managed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bookings/transactions</td>
<td>Seats sold, prices, elasticities, route demand, points of sale, ancillary services and up-selling</td>
</tr>
<tr>
<td>Loyalty programmes</td>
<td>Demographics, travel histories, preferences, (other) marketing opportunities</td>
</tr>
<tr>
<td>Airport operations</td>
<td>Facilities used, times on gate, checked vs carry-on bags, above- and below-wing services, heat maps and bottleneck predictions; mobile engagement: location-aware (GPS, wifi, beacons) and personalised services for employees and customers; passenger spend (maximising revenues) and airport access profiling; security planning, building management/maintenance and energy consumption systems</td>
</tr>
<tr>
<td>Flight operations</td>
<td>Flight plans, fuel loadings, weight/balance data, taxi times, flown trajectories, real-time aircraft health and customer tracking</td>
</tr>
<tr>
<td>Supporting information</td>
<td>Airspace availabilities and congestion monitoring, support services (e.g. meteorological), fleet and revenue tracking, social media interfaces</td>
</tr>
<tr>
<td>Forward planning</td>
<td>Future capacity and investment planning for infrastructure development (projections may draw on not only airport and regional data but also national and global events and trends)</td>
</tr>
</tbody>
</table>

Adapted from Marks, 2014.
predictive analytics to combine real-time and historical data to improve yield management at concessions. Passengers may, for example, be sent text or email messages at the airport with a special offer for a given retail outlet during the next hour, tailored to national (or even personal) spend profiles. With most major airports now offering a mobile app (e.g., for flight status data), this type of targeting is likely to grow. Drawing on historical buying patterns, significant opportunities are thus being explored for better consumer profiling and segmentation, producing commercial trend data and better targeting for loyalty programmes.

We consider more specifically as a brief case study the situation at Frankfurt Airport — a leading exponent of the use of big data. In approximate numbers, the airport serves over 150,000 passengers a day, handles up to 100 aircraft movements per hour, using 150 flight gates, 200 aircraft positions and with 4 runways at its disposal, plus 400 daily trains serving the airport (thus making it one of the highest volume intermodal hubs in Europe). Passengers have potentially large amounts of data associated with them, such as reservation, consumer tracking, geo-location and socio-demographic information. Aircraft may also generate several terabytes of data per flight from their sensors. Nevertheless, the challenge at the airport (for example) is not so much the volume of the data, but rather its complexity and interfacing across aviation, ground-handling and retail subsystems. Fraport’s airport operational database (AODB) is called ‘INFOplus’ and operates some 60 interfaces to such systems (see simplified view in Figure 7.7).

![Fraport’s airport operational database key interfaces](image)

**Figure 7.7** Fraport’s airport operational database key interfaces

Reproduced with permission.
This is modular in structure and migration-capable for specific customers, assimilating highly heterogeneous data, such as:

- meteorological;
- ATM;
- aircraft movement (time stamps, loads, engine performance);
- passenger (individual level, but anonymised);
- public transport supply.

These are connected with online dashboards, e.g. showing noise emissions data on the airport’s public sustainability homepage, and internal management systems that include punctuality forecasts. Near-real-time aircraft position data not only feed airport collaborative decision-making (A-CDM) processes but also one of the more advanced features of the airport’s use of big data: predictive analytics. In conjunction with visualisation through heat maps, inter alia, this is used to forecast passenger bottlenecks at known pinch-points at the airport (such as immigration control), several hours ahead. For this, multiple parameters from each flight (operator, aircraft type, load factor variations, country of origin, time and season, two-year historical data matching, etc.) are combined with airport sensor (and other flow) data, plus sophisticated passenger-forecasting tools. Each passenger’s pathway through the airport is simulated in an agent-based model, with an update on predicted bottlenecks made every five minutes. Comparisons are made between outcomes based on the current configuration, and those with alternative, pre-emptive changes, such as reassigning an inbound flight from a gate at Terminal B to a remote stand plus bus solution at (less-congested) Terminal A. In parallel to optimising (retail) products and services mentioned earlier, increasing passengers’ positive experiences, this type of flow management is essentially aimed at obviating negative events. Clearly, far from being a buzzword, big data are already being used to improve the passenger experience and business performance of the airport.

### 7.6 Conclusions

We have left a few questions unanswered above, which we at least try to start answering here. For example, we asked how well non-complexity metrics and methods capture certain features of ATM system dynamics (such as delay propagation) compared with those of complexity science. To summarise our view: we need both. Throughout this book we have sought, inter alia, to demonstrate what complexity science brings to the table regarding our understanding of various facets of ATM, and we project these thoughts forward in Chapter 8. We would advocate that a complementary approach is needed, whereby data science and CNT techniques complement existing (classical) methods, critically adding to our understanding of how ATM works. In Chapter 2, we proffer that key performance indicators (KPIs)
need to be chosen that are intelligible (often simple), sensitive, and consistent (we
cannot refine them from one period to another without losing comparability). How-
ever, the concomitant disadvantages are that it is then difficult to adapt them in
response to new data or methods, and that they may not afford the best real under-
standing of system dynamics. Let us take an example. When the mean, European
en-route air traffic flow management (ATFM) delay per flight (see Chapter 4)
decreased from 4.5 minutes in 1999, to 0.53 in 2013 (EUROCONTROL, 2014),
getting very close to this KPI target of 0.5 minutes, this observation in itself does
little to improve our understanding of performance. It says (almost) nothing about
the temporal and spatial dimensions, or the variability by types of operation. At
best, it tells us that probably more things got better than got worse. This is not to
say that it is not a useful KPI at a high level, but it tells us only what the overall
outcome is, not how we are doing it or, moreover, how we might better achieve it.

Such are the challenges that data science equips us better to explore, in particu-
lar, the trade-offs between KPIs and the interactions at lower levels with other
performance indicators. Through data mining and knowledge discovery we may
indeed often also be presented with unexpected findings, even wholly counterintui-
tive results. Being prompted to explain these outcomes is part of the process of
gaining deeper insights into system dynamics. As useful as these techniques are,
the outputs do not usually come with a ready-made explanation for the researcher.
The complementarity between the classical and non-classical methods may mean
that the latter furnish insights that we would previously have been without, allow-
ing us to go back and investigate further using selected classical methods. The
researcher always has choices to make, and often these are not easy. We are not yet
at the stage where we can present a big dataset to our software and the answer we
want appears several seconds or minutes later. We still (largely) design the experi-
ments, and use our judgement to determine the constraints of an analysis, the speci-
fication of nodes and edges in our network, and how and when we use p values.
Sometimes we may apply parallel methods to the same problem, whereby analo-
gies emerge in the design – for example, between the choice of factors in a factor
analytic model and the choice of nodes in a CNT model. We are also usually comp-
pelled to apply varying degrees of judgement even before any of these analyses
start, at the data cleaning stage, which itself will also typically have consequences
for the analyses.

This complementary, or meta-methodological, approach also mitigates what is
sometimes called ‘research enculturalisation’, whereby a field of research adheres
too narrowly to its own received wisdoms and culture. Data science is here to stay;
CNT is finding its feet in some domains where it has shown initial promise. We
count ATM among these. Deficiencies in our various models are unlikely to be
solved overnight, although we should not forget the cost-effectiveness of these
models, which afford insights that would be impossible to obtain through costly
experimentation with the actual systems themselves. As computing power is ever
improving (one wonders what paradigm shift will replace cloud computing), we
are able to build bigger and (hopefully) better models, using larger and larger
datasets. The powerful alliance of CNT and data science applied to complex networks is set to allow us to see the important trees in big woods.

Key challenges going forward particularly with regard to managing big data include maintaining open architectures in the context of an increasing diversity of both data formats and demands from client interfaces. Growing issues may also emerge with respect to ensuring the reliability of such data as interpretative and management systems become more automated still. Without clean, unbiased and complete data inputs, we cannot expect to get quality outputs. There may be a trend towards increasing dynamic data consumption (and metrics) as the cost of warehousing may not decrease sufficiently to offset the growing volumes of data. Current objectives look firmly towards removing a reliance on separate data stores and moving increasingly to centralised visions. Specifically regarding a key feature of ATM, despite progress in A-CDM processes, there is also significant development work ahead, as these systems often still require substantial manual intervention and have relatively poor integration of passenger data. Ultimately, the true value of big data is the extraction of meaning across multiple domains and the coordinated application thereof in a business and operational context. We are edging ever closer to realising the concept of the airport as a smart city, with improved mobility, efficiency and resilience.

Note

1 For an introduction to the p value, Frank and Althoen, 1994, is a great overall statistics textbook. Tull and Hawkins, 1993, offer a brief introduction to the chi-square test; much more detail is available in the dedicated work of Everitt, 1977. For the more adventurous, perhaps, Ziliak and McCloskey, 2011, offer a fascinating challenge to the fundamentals of modern statistics, including the use of the p value itself.
This page intentionally left blank
Chapter 8
Conclusions and a Look Ahead

David Pérez

8.1 Collaborative Research – The Way Forward

Research networks act as alliances, encouraging the development of knowledge and scientific leadership, and are essential elements of an increasingly collaborative research environment. Networks such as ComplexWorld (introduced in Chapter 1) thrive on the contributions of their members, especially those with specific knowledge background in niche subjects. This research network is an example of a successful alliance between entities committed to the development of a particular subject. Complexity science was not an obvious research discipline to promote in ATM, which has long research cycles and well established research themes; however it has proved to be a new and worthwhile discipline that adds significant value to the field.

Inspired by the great work of Harvard Business School Professor Clayton M. Christensen, the network has revolved around disruptive concepts rather than researching technologies already investigated by more conventional research avenues. The team committed to the definition and description of new, disruptive research lines and, leveraging their experience as research leaders, looked deeply into the existing knowledge on complexity science to define complexity science as a research theme.

The founding members of the consortium consolidated the input received from different channels and defined the subject matter. This was only possible with synergy between the institutions contributing to the mission. As complexity science is inherently multifaceted, a unified vision of the application of complexity research in ATM would have been neither feasible nor desirable. Instead, the team successfully compiled key areas addressing the core elements of complexity in the domain. This in-depth perspective would not have been possible without fruitful collaboration between the different entities involved and also without the promotion of the European institutions. It is through this type of collaboration that new avenues of research are being created – paving the way for additional research initiatives, which address the elements described in this book.

Although the subject of complexity in ATM was fairly nascent at the time of the network’s formation, research efforts made thus far have served as the foundation and inspiration for the themes developed in this book. In addition, it would be remiss not to mention the work and engagement from the ATM research community. From attending different events to providing constant input and new ideas,
this research community has helped to shape the ATM complexity research subject.

8.2 Taking Stock – A Critical Reprise

Complex network theory (CNT), introduced in Chapter 2, offers an excellent springboard to explore the behaviour of the air transport system. By leveraging the understanding of pervasive network structures of other domains, it enables the application of powerful tools capable of exploring the behaviour of the system in a way that traditional methods cannot achieve. Complex network analysis offers the possibility of expanding the current performance framework (that is, the collection of metrics that are monitored to evaluate the evolution of the system) to a more comprehensive and executable framework. For instance, analysing the performance of the system from a passenger perspective simultaneously requires the introduction of network concepts deploying a temporal structure and also multi-layer frameworks capable of exploring the performance of the passenger connections. Other types of complex network representations in air transport allow analysis from different perspectives and provide insights into various aspects of the system, as described in Chapter 3. Research on the topology of the airport network, navigation aids and other structural elements provides information about the supply elements of the aviation system and, therefore, valuable information on how the network is defined. Analysis of the resilience properties of the system during disturbance, regarding delay propagation, and in response to (potential) safety-critical events, plus many other properties of system behaviour, are also achievable through the right representation of the network.

One of the most important goals of the management of air transport is to ensure predictability in operations, but the concept of uncertainty, covered in Chapter 4, has been widely ignored in ATM research. Certainty, at different scales in air traffic operations, not only brings punctuality to the users of the air transport service but also enhances safety, impacts on capacity and allows implementation of cost-mitigation measures. Uncertainty is not easy to define due to its various aspects, sources and scales. In fact, the lack of a common definition among researchers is clear. A proposal for the definition and classification of uncertainty is presented in Chapter 4 and this will allow future research in the area to build on a shared foundation.

These uncertainties may lead to disruption of the socio-technical air transport system, which represents a serious issue to all stakeholders. The impact of the perturbations is propagated throughout the system. Infrastructure providers, such as airports or air navigation service providers, need to adapt to new scenarios, coordinated by the network managers. The proper management of these disruptions by various stakeholders is addressed through the concept of resilience, which is discussed in Chapter 5. The difficulty of defining, understanding and assessing how a distributed socio-technical system might manage disruptions without modifying
its performance or might revert to an original state after being altered is a key design aspect. The challenge of understanding resilience is strongly related to the interaction between different aviation elements, or its complexity. For this reason, resilience has been included in the list of disciplines in which complexity science can be applied to air transport.

Several resilience metrics and modelling techniques have been identified in Chapter 5, which begins with the definition of three resilience capacities: absorptive capacity, restorative capacity and adaptive capacity. Network-flow and agent-based modelling have proven effective when assessing the impact of the different strategies. Both techniques are capable of simulating the behaviour of the different stakeholders involved in a complex decision-making process with greater accuracy than classical methods. Agent-based modelling has advantages if the resilience problem studied involves human decision-making; network-flow modelling has advantages if the resilience problem involves network-wide issues. By integrating agent-based and network flow approaches, it is feasible to address both issues within a simulation. Stochastic reachability and viability theory are approaches that may complement agent-based and network flow modelling in particular fields such as collision risk analysis and safety critical control.

Agent-based modelling is also one of the main techniques available to capture emergent behaviour. Emergence is a central concept to complex systems theory and several kinds of emergence can be found within the ATM system, as illustrated in Chapter 6. The interaction between different stakeholders very often produces behaviours that cannot be inferred from the analysis of the individual components. Phenomena such as certain safety incidents, delay propagation or capacity constraints can only be explained at the network level due to the interaction of the different elements. Emergent behaviours that can be predicted, however, are of use in the design of a future socio-technical ATM system. Weak emergent behaviour of future designs can be identified and analysed by the application of agent-based modelling and simulation in combination with agent-based hazard modelling and rare-event Monte Carlo simulation. This opens valuable directions for follow-up research, as is highlighted in the conclusions of Chapter 6.

The extraction of meaningful information from the various datasets generated by air transport systems and operations is an area central to complex systems, and can provide methodologies and tools to enable insights into the complex behaviour of the air transport system. These concepts, addressed in Chapter 7, go beyond using data solely to calibrate or validate models (both tasks are of strong interest and value in the aerospace modelling community). Data-driven techniques could provide insights into behaviour without any preconceived model.

As is explained in Chapter 7, ‘big data’ comprises those data that exceed the processing capacity of conventional database systems (O’Reilly Media, 2012). This type of data requires special treatment due to its volume, speed, and/or incompatibility with structures of traditional database architectures. Alternative means of processing these data are needed to be able to extract value through statistical or data mining analysis, or to build new products or services. During the last decade
cost-effective solutions have become available to manage this type of data. While the large volume is the most known aspect that is addressed through big data architectures, the velocity of incoming data and the vast variety of data types are additional aspects that need to be taken into account.

Unfortunately, data often appear incomplete or redundant, and security, validation issues and data acquisition challenges are barriers that need to be addressed. All of these challenges can only be addressed with domain-specific expertise regarding aviation data. Inspired by the 2002 Nobel Prize winner, Sydney Brenner, who stated ‘how so much progress depends on the interplay of techniques, discoveries and new ideas, probably in that order of decreasing importance’ (Robertson, 1980), the ATM research community has recognised that the analysis of the complexity of ATM operations through data is only achievable if a set of data science techniques is in place. Knowledge discovery and data-mining techniques (see Chapter 7) are widely used in different fields to extract meaningful information from datasets. However, despite the inherent complexity of different air transport phenomena and the difficulties of modelling the interaction between the different elements, the existing techniques of data science developed specifically for data analysis for this domain are almost nil.

8.3 A Look Further Ahead

Much progress is expected from science and technologies developed by the SESAR exploratory research programme, which should enable a breakthrough to be made in the way air traffic is handled in Europe. After a few years of network activity, complexity science has progressed on different ATM fronts such as safety, capacity, delay propagation, passenger delay, data security, cost models and resilience. However, there are still several areas where further advances are expected. The definition of the main research lines within the field of complexity science in ATM provides a guideline for future years of research within the SESAR exploratory research programme. Furthermore, applied research work in various areas should nurture the main body of knowledge reported in this book. In particular, the following avenues are worth exploring in the context of the current aviation challenges.

8.3.1 Performance Frameworks

The four-hour door-to-door challenge, as presented by the European High Level Group on Aviation Research in the ‘FlightPath 2050’ report (European Commission, 2011), represents an ambitious vision for Europe to significantly increase passenger mobility. To achieve this type of vision, current performance frameworks require important updates from a methodological and technological point of view.

Additional research on CNT will boost the methodologies available for performance assessment, enabling the performance framework to advance from basic
statistical analysis to a scientific, data-driven performance framework. With this, the network manager can inform decision-makers on the behaviour of the system as a network, and offer advice on the impact on policy actions or technological improvements.

Finding the balance between a sophisticated performance framework that provides simple, easy to understand indicators while efficiently capturing the performance of the system will require additional research in different areas, combining complex network analysis, data science and data visualisation techniques. Future performance frameworks should include a variety of metrics that capture the different characteristics of system behaviour, from the passenger perspective to the resilience of the system, from delay propagation to the topological properties of the network of airports, aerospace sectors or navigation facilities. Additionally, performance frameworks should be capable of providing predictive analytics for different time horizons.

8.3.2 Knowledge Discovery and Data Mining

In the knowledge discovery field, data is the raw material required to start any analysis. However, the real value lies with what is found inside the data when analysed with proper tools. Data science is a multi-disciplinary paradigm that complements other complexity science methodologies, such as complex network analysis or agent-based modelling, but presents its own research challenges.

More research on a big data infrastructure targeting air transport is needed. A big data infrastructure should be capable of providing with ease, regardless of the size of the datasets, fluent access to the most important datasets that encapsulate the real behaviour of the system, including trajectory data, weather, slot allocations and safety events. Much of this information is fairly accessible, but difficult to aggregate and synthesise. Some of this information may be confidential and the appropriate paradigms for secure computation will allow researchers to derive the right insights without revealing commercially sensitive information. Important efforts have been made already in the context of system-wide information management (SWIM), but more research is needed in this context.

Knowledge discovery and data mining is a complex subject requiring a skills set of mathematical expertise combined with computer science knowledge to implement and optimise the performance of the algorithms. Operational experience is usually needed, or at least advantageous. The harmonious integration of all these concepts involves much more than data mining algorithms or big data infrastructure alone. Successful data science paradigms combine technologies from such fields to support the visualisation of business problems, for example, aviation safety, from a data perspective. There is a fundamental structure to data-analytic thinking, and basic methods that should be understood, involving principles, processes and techniques for understanding phenomena via the automated analysis of data (Provost and Fawcett, 2013), with the ultimate goal of improving decision-making. Therefore, data science will be developed as a combination of disruptive
technologies that is driving game-changing paradigms in the management of complex systems. Additionally, the development of big data predictive analytics dashboards, capable of supporting decision-making visually to optimise the performance of complex systems, will allow us to develop data-driven innovative concepts.

Large datasets are definitely needed to characterise the uncertainty of the system and incorporate uncertainty and its sources in the definition of operational concepts, instead of forcing deterministic operational concepts to deal with uncertainty. This challenge will lead to the development of adaptive systems capable, for example, of having dynamic sector capacity as a function of uncertainty (such as weather) or as a function of the topological or dynamic complexity of the traffic structures. Defining truly dynamic sector capacity should be done taking uncertainty as a core input. Other areas of interest in this context could be new runway throughput management paradigms, through mining datasets potentially relevant to runway performance, such as historical weather conditions, traffic congestion, slot allocations, air traffic control procedures, runway occupancy times, apron stands assigned to each aircraft and taxi-times.

8.3.3 Uncertainty

Uncertainty in network management is a novel area of research that could ultimately impact on the way network topology is understood. In a more integrated European aviation transport economy, analysis of the topology of the network at different scales and a deeper understanding of the importance of the different network elements should lead to different macroeconomic strategies to maximise the impact of investments and, in general, to policies driven by the optimisation of network resources.

Flight delay is the most important aspect of uncertainty at the flight level. The large economic impact of this type of uncertainty justifies further research within the area, particularly in the development of predictive analytics, despite valuable research work completed and the continued ex post reporting from EUROCONTROL. At the trajectory and traffic level, uncertainty impacts on the potential safety of air transport operations, which is usually maintained through costly safety buffers. For instance, sector capacity management lacks the necessary uncertainty management tools that would allow more efficient use of resources, i.e., management of sector capacity. A more active, integrated management of safety and uncertainty levels is critical so as not to rely solely on such safety buffers to manage potential overloads. The potential availability of large quantities of data along with state-of-the-art knowledge discovery mechanisms would enable decision support tools to be developed, capable of managing uncertainty at the tactical level. The potential impact of these techniques supports the need for making data open to researchers. The data could potentially be used to address future data uncertainty as a field of research.

Weather is a source of uncertainty that requires particular attention due to its impact at all scales. Considering this criticality, current European and North
American ATM research programmes hinge on a concrete research effort that analyses the interactions of weather with the rest of the ATM technologies, ultimately shaping how weather uncertainty concepts are introduced in ATM technologies.

8.3.4 Metrics

The search for widely accepted resilience metrics that could be incorporated into the air transport performance framework has been active for some years. The impact of disturbance on the reliability of schedules is a major area to have been studied, although resilience metrics should also tackle other key performance areas.

The first steps in making available the European datasets needed to compute resilience have been taken, but significant hurdles remain. These relate to the availability of data for research, processes to automate data cleaning and data research infrastructures to support the main modelling techniques used to provide a quantitative picture of the resilience of the air transport system (complex network modelling, data science and agent-based modelling).

Improving data-driven decision-making through automatic analysis of data is a multidisciplinary endeavour that includes statistical analysis, artificial intelligence, machine learning, computer science, data infrastructure, network theory and scalable analytics. The field indeed extends beyond the complexity science realm, and further efforts are recommended to the research community to achieve maturity in all of these disciplines in ATM.

For this, two areas were identified as particularly relevant for data-related progress beyond the state of the art in the complexity knowledge field in ATM: the study of causality and the development of metrics capable of capturing the system’s complexity behaviour. In the case of causality, some work has been completed in adapting and applying Granger’s work (see Chapter 7) in ATM, as previously achieved in biomedical and economics datasets. The initial results are promising, but due to the number of very speculative cause–effect relationships frequently mentioned, more work is needed to develop a body of knowledge and tools that can properly analyse series of datasets and identify those relationships.

Regarding the development of complexity metrics, it is very well known that despite the complexity features shown by the air transport system, most ATM performance dashboards use simplified statistical measurements that fail to capture the complexity. This results in a condensed picture of the actual performance of the system, which sometimes leads analysts to invalid conclusions. For instance, horizontal (across space) and vertical (through time) averaging of delay data (for example, average air traffic flow management (ATFM) delay per flight, per year in Europe) fail as metrics to measure delay, as the distribution of delay is far from a Normal distribution across time or space. Work has been carried out by the complexity science community on advancing these metrics, but further research is needed in this area, requiring collaboration between data science and complex network analysis.
8.3.5 Modelling Tools

Current modelling tools need to keep improving to become capable, data-driven, scientifically founded research and design tools. For instance, the use of agent-based modelling is very useful in the early identification of emergent behaviours in a new concept of operations in which different agents need to collaborate or compete to find a solution. The research community needs a strong emphasis on establishing the foundations of an agent-based environment capable of simulating a variety of scenarios and operational concepts. Most agent-oriented methodologies provide a rather vague conceptualisation of the agent-environment in the design phase and even in the execution phase, with a lack of precision regarding which entities are part of the agent-environment, and which are not, and the resources that those agents are actually using. An environment capable of encapsulating the available resources and services to be accessed by the different agents would help to achieve more realistic research. Further research is needed in this field to correctly understand emergent behaviour of future designs of the socio-technical air transport system.

8.4 Closing Thought

Several research areas explored in this book will require dedicated future effort, building their own communities and attracting talent from other disciplines. ComplexWorld has organised various conferences on complex systems to foster the multidisciplinary nature of the field and to link with the complex networks communities. One field in particular is worth mentioning: data science applied to aviation. This originally grew as a field of research within complexity science as a discipline utilised to examine the complex relationships between the different agents through analysis of datasets. However, it has flourished as a research area in its own right, requiring specific resources far removed from the original scope.

Great achievements have been accomplished within the last years in the field of complexity science as applied to air transport. While maturing research is expected to flow into the technological roadmaps in coming years, additional valuable opportunities are waiting to be addressed.


References


References


References


Delay Pattern, Proceedings of the International Conference on Research in Airport Transportation (ICRAT 2012). (Berkeley, California, USA.)


References


This page intentionally left blank
Index

accidents 6, 61, 69, 74, 83, 88, 90, 94; risk assessment 54
adaptation, emergence (comparison) 86
adaptive capacity 63–4
advanced autonomous aircraft (A3) 97–100
agent, definitions 91–2
agent-based models 9, 35, 54, 92, 127, 133, 138; airline operations control 79; delay propagation 35–6; hazards 92–4; risk 89, 97
agent-based modelling and simulation (ABMS) 9, 71–3, 79–80, 90; application 91–2; complex socio-technical systems 91–2; dual capability 75; future 138; human-in-the-loop 89; mathematical methods, integration 95–7; power 91; requirements 73; safety risk 89, 97; tools 90, 138; types 89–90
agent-based safety risk analysis 91, 93, 97, 98, 103
airborne self-separation, 9: advanced autonomous aircraft (A3) operation 97–100; rare-event Monte Carlo simulation 100–2
airborne separation assistance system (ASAS) 97, 99; supporting systems 101
aircraft: identification 115; movement 127; trajectories 99, 119; turnaround process, dynamics 48–9
airline controllers, activities (classification) 77
airline disruption: agent-based simulation results 79–80; management 75–7; policies 77–9; scenario 78–9
airline operations control (AOC): application 62; operators, corrective actions 75–6; organisation, overview 76; policies, resilience capacities (relationship) 78
airport collaborative decision-making (A-CDM) 127, 129
airport networks 23, 24, 25, 26, 27, 28, 29, 31, 32, 36; informative indicator parameters (usage) 27–8; topology (characterisation) 24, 27, 36
airport operational database (AODB) 126
airports: ground handling, pre-departure stage 47–8; nodes (number) 24; operational database key interfaces (Fraport) 126; topological features (function) 27, 28, 122
airspace: concept evaluation system (ACES) simulator 74; hierarchic partition 25–6, 32; system 25–6, 31
airspace concept evaluation system (ACES) simulator, usage 74
airspace design, communities 26, 32–3
air traffic: design, emergence (identification) 89–91; growth, accommodation 83
air traffic control (ATC) 1, 47; centres 33; communication, usage 76
air traffic flow management (ATFM) 137; delays 20; slot allocation process 46
air traffic management (ATM) 1, 127; complexity 1–2; context 9–10; infrastructure, information (provision) 30; organisation, fragmentation 3; performance, network layers 112–15; probabilistic framework, applications 49–50; scales, usage 44–6
air traffic networks: communities 32–3; community detection 7; dynamics 33–7; epidemics, spread 36–7; passenger dynamics 33–4; results 33; topologies 26–32; types 24–6
air transport: complexity science, usage 74–5; complex networks, usage 7, 23, 27, 32; elements, interactions 92; emergence, examples 87; emergent behaviour 84, 86–9; interdependencies, modelling 71; management, goals 132; network resilience, measurement 68; operations 87; performance/cost assessment 19–21; topologies 38, 120; usability, resilience metric 69–70

algorithms: community detection 33; conflict resolution 53; filtering 50–1; processes 53; viability 74

alpha centrality 14, 113–14

analytical queuing 35

Aristotle 84

arrival: delay 35; times 43

asymptotic trajectory 122

autonomous systems-level networks 18

Base of Aircraft Data (BADA) 43

Bayesian approach 106–7

belief function 65

betweenness 14–15, 19, 29, 33, 38

bias, assessment 96

bias and uncertainty, assessment 96

big data 106, 123–7, 133–4; domains 125;

information 135

brief resilience scale 68

catastrophic accidents 69

causality 109–12; detection, problem 109–10, 118; Granger causality axioms 111

cause and effect 92–3, 106, 108–12

centrality 14

Chinese whispers, game 94

chi-square test 116–17

Christensen, Clayton M. 131

civil air transport 61

classical metrics, usage 108, 112

clustering coefficient 14, 17, 27, 31, 38, 113–15, 120–3

collaborative research 131

collision avoidance layer 97

community detection 24, 26, 32, 33, 36

complementary simulation tools, usage (requirement) 90

Complete Air Traffic Simulator (CATS) 54

complexity 5–6; context 120–3; context, cause and effect (tackling) 109–12; knowledge 137; metrics 7, 137; traffic, 1; types 14–15

complexity metrics 7, 14, 108, 137

complexity science 6–7, 70–5; agent-based modelling and simulation (ABMS) 72–3; applicability in ATM 131;
stochastic reachability analysis 74;

viability theory 74

complex networks 2–4; classes 16–18;

models 56–7; usage 7, 23

complex network theory (CNT) 1, 6–7, 23;

abstraction 55; interdependencies 70–1; metrics, comparison 108–9;

review 132–4

complex socio-technical systems, 24, 25, 27, 33; agent-based modelling and simulation 91–2

ComplexWorld (network) 1–4, 131

conditional Monte Carlo simulation 95–6

conditional probability approach 65

conflict detection and resolution algorithms, usage 53–4

connectivity mesoscale 13

correlation c.f. causality 109–12

covariance matrix, usage 49

critical infrastructure systems:

interdependency types 71;

resilience metric 64–6

cyber interdependency 87

cyclone, characteristics 42

data: challenges 134; cleaning 115–16, 128; distribution, problem 118; ‘many sets of’ methodology 124; open architecture 124, 129; packet drops 12; processes 106–7; safety data 26; sets, relationships (metrics) 110;

streaming 125; traffic data 33;

uncertainty 43; visualisation 105, 118, 124–5, 127

data integrity 115–23; classical context 115–20

data metrics: classical approach,

shortcomings 106–9, 127–8;
classifications 108
data mining 105, 106–9, 124–5, 128; knowledge discovery, relationship 135–6; techniques 107
data science 7, 105–29, 135
DDR2 service (EUROCONTROL) 116
decision tree 107
degree 13–16, 113, 119–20; distribution 15, 17–18, 27–31, 36–8, 120;
entropy 14, 113–14, 122–3
degree correlation 14, 18, 122–3
delay 15–16, 33; cost of 19–21; passenger-related 34; reactionary 25; see also propagation
dependability, resilience capacities (relationship) 63
differentiated scales 44–5
directed graph 10, 14; network 24–5
disembarking passengers 47–8
disruption 12, 19–20, 37–8; consequences, modelling/analysis (ability) 75;
management policies 77–80; moment, display 65; research 131
domains: 10–11; agent-based modelling and simulation (ABMS), application 91–2; network properties, examples 11
domino effects 46
downward causation 85
dynamic risk model, Monte Carlo simulations 89
dynamic variability sub-model 94
earthquake engineering 64
ecological resilience 64
economic theory-based approaches 71
ecosystem: domain 75–6; resilience metric 64
edges see links
effect-size calculations 124
eigenvector centrality 14, 16
ELSA (SESAR project) 2, 28
emergence 84, 133; adaptation, comparison 86; examples 87; identification
89–91; interdependencies 87; resultant emergence 85; strong emergence 85; types 86; weak emergence 85
emergent behaviour 6, 83; analysis 91–7; impact 84–91; perspectives

84–6; weak emergent behaviour (evaluation), complementary simulation tools (usage requirement) 90
engineering resilience 64
entropy (degree distribution) 14; normalised 36; transfer 110
epidemics, spread (air traffic networks) 7, 36–7
equipment uncertainty 43
Erdős, Paul 16
Erdős-Rényi graphs 7, 16, 19, 35–7
Euler, Leonhard 10, 16
EUROCONTROL 1, 28; DDR2 service 116; performance 3; PRISME dataset 115
events: exceptional, 88; probabilities 100; safety events 26, 37; timescales 95
filtering algorithms 50–1
flight: delay 35, 46–9, 136; prioritisation 20; scales 51–2; stages 46–8;
uncertainty 45, 46–51
FlightPath 2050 report 134–5
flow control-induced macro behaviour 98
flow management 52; problems 45
Fokker-Planck-Kolmogorov 95
four-dimensional (4D) trajectory 41; plan 97–8
Frankfurt Airport, big data usage 126
Fraport, airport operational database key interfaces 126
free-diffusing aircraft regime, distinction 35
free flight 97
F-test, usage 111
full performance impact, moment (display) 65
full recovery, moment (display) 65
functional airspace blocks 32, 33
future state, prediction 45–6
gate-to-gate flight stage 46
Gaussian distribution 120
generalised synchronisation 110
geodesic distance, mean 15, 17, 19, 122–3
geographic interdependency 87
global efficiency 15, 113–15, 122–3
global positioning system (GPS), usage 43
Granger causality axioms 25, 111
Granger, Clive 110, 137
granularity 124
graph theory 10, 16
Great Circle distances 15–16
guidance, navigation, and control
(GNC) 99

hazards: agent-based modelling 92–4;
capture, sub-models ranking 93
heat maps 125, 127
heterogeneous data, assimilation 127
HETN (model) 36
hidden determinism 42–3
HOMN (network) 36
horizontal separation criteria 101
hub(s) 17–19, 29, 35, 37, 48, 120, 126;
hub-and-spoke network 2, 7,
11–12, 34; hub–hub connections
14; Internet 36
human-in-the-loop simulation (HITL) 89
importance sampling 95–6
incidents 51, 66, 90, 133; role of pilots 6;
see also accidents
in-flight (shorter horizon) uncertainty
48–9
INFOplus 126
information content 15, 122–3
information systems, resilience metric 67
Integrated Initial Flight Plan Processing
System (IFPS) 47
interdependency: classes 70–1; cyber
interdependency 87; geographic
interdependency 87; logical
interdependency 87; modelling
71; physical interdependency
87; types (critical infrastructure
systems) 71
Internet: mapping 18; representation
18–19; topology, analyses 19
joint activity, choreography 77–8
Kalman filter 51
key performance indicators (KPIs) 9,
127–8; multi-dimensional KPIs 69;
total impact 67
K-nearest neighbours 107
knock-on effects 116–17
knowledge discovery 105, 128; data
mining, relationship 135–6
Königsberg (Kaliningrad), Seven Bridges
of 10
level bust 94
Likert scales 70
linear trajectory 122
link density 15, 122–3
links (edges) 10, 24, 25, 27, 30, 122, 128;
attack 19
logical interdependency 87
low-cost airline 2, 31
machine learning 105–7, 125
macroscale 5, 7, 13, 45, 84–5
Markov chain 52–3
Markov models 5, 107
medium-term planning level 98
mesoscale 5, 13, 45, 52;
meteorological conditions, prediction
ability 43–4
metrics 137–8; classifications 108;
definition 9–10; see also
complexity; networks; passenger-
oriented metrics
microscale 5, 13, 44–5, 84–5
mid-air collisions 62, 91; risk 101;
Ueberlingen 88
mission risk metric 70
modularity 15
moment of full performance impact,
display 65
moment of full recovery, display 65
Monte Carlo methods 51; uncertainty
propagation classification 50
Monte Carlo simulation 93; conditional
95–6; dynamic risk model 89;
preclusion 85; rare-event 5,
100–2; run, outcome 99–100;
usage 54
motifs 13
multi-agent model 93
multi-agent situation awareness (MA-SA)
94
multi-layer networks see networks
multi-modal transport 7
multiplex 31–2
Nagel-Schreckenberg (model) 2, 35

navigation points: actual flights, ratios 119; predicted flights, ratio 119

navigation system accuracy (assessment), stochastic differential equations (usage) 51

networks: airport 23–5, 27–8, 31–6; air traffic 24–32; autonomous systems-level 18; Bayesian 107; characterisation of 12–16, 120–3; complex 23, 27, 32; damage in 66–7; decomposition of 25; directed 24–5; flow-based methods 72, 90; metrics 23, 34–6; multi-layer 2, 21–2, 23, 25, 112–15; neural 107; passenger-oriented 19–21, 25; percolation in 88; random 16, 19, 35–7; real-world 7, 10; scale-free 17–19, 35, 37, 38, 120; temporal 21–2

Next Generation Air Transportation System (NextGen) 23, 89

nodes (vertices) 10, 26, 37, 38, 112–13, 115, 120–2, 128; attack 19; connection, probability 17–18; defining 15–16; see also betweenness; degree

nominal emergence: definition 85; macro property 85; types 85

non-linear (system) 1

non-Monte Carlo methods 50

non-nominal effects, sub-models ranking 93

non-parameter differences, effect (assessment) 96

NoSQL databases 123

observable states 53

obstacle avoidance 74

off-block time 47–8; last-filed estimation 48

on-board technical system, function 87

operation, airborne self-separation concept 75

operational uncertainty 43

operator misinterpretation 53

organisation systems, resilience metric 67

out-degree centrality 113

p value 117–18, 124, 128

parameter values (changes), risk level assessment (elasticity assessment) 96

passenger: dynamics 33–4; mobility 20, 25

passenger-oriented: metrics 20; networks 19–21, 24–5

Pearson correlation coefficient, computation 119–20

performance: consequences 69; frameworks 134–5

Petri net 54; stochastically and dynamically coloured (SDCPN) 95

phase transition 2

physical interdependency 87

pilot-in-the-loop simulations 97

pilot not flying (PNF) 99

POEM (SESAR project) 21, 25, 36, 115

power law 17–18, 27–8, 30–1, 38, 120; behaviour 27; distributions 1, 7, 35, 37

predictive analytics 124–7, 135, 136

preferential attachment 87

present state, estimation 45

PRISME dataset (EUROCONTROL) 115

probabilistic models 5; methods 42, 54

projected network 24, 31; usage 112

propagation: delay 7, 12, 15, 20–1, 23, 25, 34–6, 39, 46, 48, 52, 56, 108, 112–13, 115–18, 120; models/modelling 5, 94; uncertainty 3, 45–6, 49–51, 57

protein–protein interaction 56

providence factor 101

psychology, resilience metric 67–8

psychometric scales, development 67–8

public transport supply 127

random: attacks 38; graph 7, 16–17, 36; model 32

rare-event Monte Carlo simulation 5, 100–2, 133

reach probability 69, 70

reality, simulation model (differences) 96; joint effect, determination 96; size/probability, assessment 96

real-time velocity 124

real-world networks 10, 17; co-dependency 10–11

reductionism, expense 84
<table>
<thead>
<tr>
<th>Index</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>168</td>
<td></td>
</tr>
<tr>
<td>research enculturalisation, 128</td>
<td>Single European Sky (SES) 1; context 9–10; high-level goals 3, 41</td>
</tr>
<tr>
<td>research networks, 131</td>
<td>Single European Sky ATM Research (SESAR) 2, 28, 89; exploratory research program 134</td>
</tr>
<tr>
<td>resilience 5, 37–8, 61; airline operations control policies, relationship 78; analysis, dimensions/implications 72; ATM system context 37; brief resilience scale 68; capacities 62–4; capture 66; data, collection 70; degrees 68; developments, review 62–3; engineering resilience 64; expression 68–9; indicator, defining 68; modelling techniques 133; qualitative measure 68; quantification 67; recovery to loss ratio, expression 67; robustness/dependability, relationship 63; study, complexity science approach 71–2; triangle 64, 65</td>
<td>Single European Sky ATM Research Joint Undertaking (SJU) 3</td>
</tr>
<tr>
<td>resilience metrics 5, 64–70; air transport usability 69–70; consequences 69; critical infrastructure systems 64–6; ecosystems 64; information systems 67; networks 66–7; organisation systems 67; psychology 67–8; transport systems 68–9</td>
<td>single-layer networks 112</td>
</tr>
<tr>
<td>resultant emergence 85–7</td>
<td>situation awareness 72, 88, 103; multi-agent (MA-SA) 93, 94</td>
</tr>
<tr>
<td>rich club coefficient 15, 18, 27</td>
<td>small-world networks 19; model 37; properties 7, 16, 17</td>
</tr>
<tr>
<td>risk assessment 96; spread 37</td>
<td>socio-technical system 5–6, 33–4, 61–3; approach 73; disruption 132–3; uncertainty 43</td>
</tr>
<tr>
<td>road transport 2; system performance 68</td>
<td>states: future, prediction 45–6; present, estimation 45</td>
</tr>
<tr>
<td>robustness, resilience capacities (relationship) 37, 63</td>
<td>stationarity 111–2</td>
</tr>
<tr>
<td>safety-critical events 54, 88–9, 111</td>
<td>stochastic differential equations, usage 51</td>
</tr>
<tr>
<td>safety risk 97; agent-based analysis 93, 96</td>
<td>stochastic dynamic programming, probabilistic framework 48–9</td>
</tr>
<tr>
<td>scales 7, 13, 51–2; brief resilience 68; Likert 70; psychometric, development 67–8; usage 43</td>
<td>stochastic reachability 133; analysis 74</td>
</tr>
<tr>
<td>Schedules Conference (IATA) 46–7</td>
<td>strategic flow control 98–9; layer 97</td>
</tr>
<tr>
<td>secure computation 135</td>
<td>strong emergence 85, 86</td>
</tr>
<tr>
<td>strong Markov 95, 96</td>
<td>Structured Query Language (SQL), usage 123</td>
</tr>
<tr>
<td>Structured Query Language (SQL), usage 123</td>
<td>support vector machines 107</td>
</tr>
<tr>
<td>surprise, notion 86</td>
<td>synchronisation 109–10; generalised 110; likelihood 110</td>
</tr>
<tr>
<td>system: dependability 63; malfunction 53</td>
<td>system-wide dynamics 71</td>
</tr>
<tr>
<td>system-wide information management (SWIM) 135</td>
<td>system-wide information management (SWIM) 135</td>
</tr>
<tr>
<td>tactical conflict resolution layer 97; agents, usage 98</td>
<td>targeted attacks, vulnerability 19</td>
</tr>
<tr>
<td>targeted attacks, vulnerability 19</td>
<td>taxi-out delay 48–9</td>
</tr>
<tr>
<td>taxi-out delay 48–9</td>
<td>taxi times 43; absence 116</td>
</tr>
<tr>
<td>taxi times 43; absence 116</td>
<td>terminal manoeuvring area (TMA) 52</td>
</tr>
<tr>
<td>terminal manoeuvring area (TMA) 52</td>
<td>thermodynamic properties 87</td>
</tr>
<tr>
<td>thermodynamic properties 87</td>
<td>time horizon: problems, types 45–6; short-term time horizon 45</td>
</tr>
<tr>
<td>time horizon: problems, types 45–6; short-term time horizon 45</td>
<td>time series 110–13</td>
</tr>
<tr>
<td>time series 110–13</td>
<td>topologies 17, 26–32, 35, 37; methods 73–4; projections 113</td>
</tr>
<tr>
<td>Traffic uncertainty 51–5</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td></td>
</tr>
<tr>
<td>Trajectory analysis 49, 116, 119; prediction 53; uncertainty 49–51</td>
<td></td>
</tr>
<tr>
<td>Trajectory-based operation (TBO) 89; agents, usage 98; layer 101–2</td>
<td></td>
</tr>
<tr>
<td>Transport systems: resilience metric 68–9; trade-off 11–12</td>
<td></td>
</tr>
<tr>
<td>TREE project (SESAR) 36</td>
<td></td>
</tr>
<tr>
<td>Unbiased risk probabilities 96</td>
<td></td>
</tr>
<tr>
<td>Uncertainty 7–8, 41–6, 55–7, 136–7; definition 42–3; network 55–7; sources 43–4</td>
<td></td>
</tr>
<tr>
<td>Undirected graph 10, 24</td>
<td></td>
</tr>
<tr>
<td>Unexpectedness, notion 86</td>
<td></td>
</tr>
<tr>
<td>Vertices see nodes</td>
<td></td>
</tr>
<tr>
<td>Viability theory 74</td>
<td></td>
</tr>
<tr>
<td>Vulnerability 18–19, 37–8</td>
<td></td>
</tr>
<tr>
<td>Weak emergence 85, 86; abundance 87; interdependencies 87; simulation tools 90</td>
<td></td>
</tr>
<tr>
<td>Weather uncertainty 5, 43–4, 51, 136–7; ensemble forecasting 59</td>
<td></td>
</tr>
<tr>
<td>Wind forecasts, accuracy (improvement) 101–2</td>
<td></td>
</tr>
<tr>
<td>Workload (controller) 1, 26, 111</td>
<td></td>
</tr>
<tr>
<td>Worst-case envelope 5; method 53–4</td>
<td></td>
</tr>
<tr>
<td>Worst-case trajectory model 53–4</td>
<td></td>
</tr>
</tbody>
</table>