Methodos Series 17

Jakub Bijak

Towards Bayesian Model-Based Demography

Agency, Complexity and Uncertainty in Migration Studies

With contributions by
Philip A. Higham · Jason Hilton · Martin Hinsch
Sarah Nurse · Toby Prike · Oliver Reinhardt
Peter W.F. Smith · Adelinde M. Uhrmacher
Tom Warnke

Springer

Methodos Series

Methodological Prospects in the Social Sciences

Volume 17

This Book Series is devoted to examining and solving the major methodological problems social sciences are facing. Take for example the gap between empirical and theoretical research, the explanatory power of models, the relevance of multilevel analysis, the weakness of cumulative knowledge, the role of ordinary knowledge in the research process, or the place which should be reserved to "time, change and history" when explaining social facts. These problems are well known and yet they are seldom treated in depth in scientific literature because of their general nature. So that these problems may be examined and solutions found, the series prompts and fosters the setting-up of international multidisciplinary research teams, and it is work by these teams that appears in the Book Series. The series can also host books produced by a single author which follow the same objectives. Proposals for manuscripts and plans for collective books will be carefully examined.

The epistemological scope of these methodological problems is obvious and resorting to Philosophy of Science becomes a necessity. The main objective of the Series remains however the methodological solutions that can be applied to the problems in hand. Therefore the books of the Series are closely connected to the research practices.

More information about this series at http://www.springer.com/series/6279

Jakub Bijak

Towards Bayesian Model-Based Demography

Agency, Complexity and Uncertainty in Migration Studies

With contributions by
Philip A. Higham • Jason Hilton • Martin Hinsch
Sarah Nurse • Toby Prike • Oliver Reinhardt
Peter W. F. Smith • Adelinde M. Uhrmacher
Tom Warnke

Springer

Jakub Bijak
Social Statistics & Demography
University of Southampton
Southampton, UK

With contributions by

Philip A. Higham
University of Southampton
Southampton, UK

Martin Hinsch
University of Southampton
Southampton, UK

Toby Prike
University of Southampton
Southampton, UK

Peter W. F. Smith
University of Southampton
Southampton, UK

Tom Warnke
University of Rostock
Rostock, Germany

Jason Hilton
University of Southampton
Southampton, UK

Sarah Nurse
University of Southampton
Southampton, UK

Oliver Reinhardt
University of Rostock
Rostock, Germany

Adelinde M. Uhrmacher
University of Rostock
Rostock, Germany

Methodos Series
ISBN 978-3-030-83038-0 ISBN 978-3-030-83039-7 (eBook)
https://doi.org/10.1007/978-3-030-83039-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To those who had to leave their homes to find a better future elsewhere

Foreword

This book, perfectly in line with the aims of the *Methodos* Series, proposes micro-foundations for migration and other population studies through the development of model-based methods involving Bayesian statistics. This line of thought follows and completes two previous volumes of the series. First, the volume *Probability and social science*, which I published in 2012 (Courgeau, 2012), shows that Bayesian methods overcome the main difficulties that objective statistical methods may encounter in social sciences. Second, the volume *Methodological Investigations in Agent-Based Modelling*, published by Eric Silverman (2018), shows that its research programme adds a new avenue of empirical relevance to demographic research.

I would like to highlight here the history and epistemology of some themes of this book, which seem to be very promising and important for future research.

Bayesian Epistemic Probability

The notion of probability originated with Blaise Pascal's treatise of 1654 (Pascal 1654). As he was dealing with games of pure chance, i.e., assuming that the dice on which he was reasoning were not loaded, Pascal was addressing objective probability, for the chances of winning were determined by the fact that the game had not been tampered with. However, he took the reasoning further in 1670, introducing epistemic probability for unique events, such as the existence of God. In a section of the *Pensées* (Pascal 1670), he showed how an examination of chance may lead to a decision of theological nature. Even if we can criticise its premises, this reasoning seems near to the Bayesian notion of epistemic probability introduced one hundred years later by Thomas Bayes (1763), defined in terms of the knowledge that humanity can have of objects.

Let us see in more detail how these two principal concepts differ.

The objectivist approach assumes that the probability of an event exists independently of the statistician, who tries to estimate it through successive experiments. As the number of trials tends to infinity, the ratio of the cases where the event occurs to the total number of observations tends towards this probability. But the very hypothesis that this probability exists cannot be clearly demonstrated. As Bruno de Finetti said clearly: probability does not exist objectively, that is, independently of the human mind (De Finetti, 1974).

The epistemic approach, in contrast, focuses on the knowledge that we can have of a phenomenon. The epistemic statistician takes advantage of new information on this phenomenon to improve his or her opinion *a priori* on its probability, using Bayes' theorem to calculate its probability *a posteriori*. Of course, this estimate depends on the chosen probability *a priori*, but when this choice is made with appropriate care, the result will be considerably improved relative to the objective probability.

When it comes to using these two concepts in order to make a decision, the two approaches differ even more. When an objectivist provides a 95% confidence interval for an estimate, they can only say that if they were to draw a large number of samples of the same size, then the unknown estimate would lie in the confidence interval they constructed 95% of the time. Clearly, this complex definition does not fit with what might be expected of it. The Bayesians, in contrast, starting from their initial hypotheses, can clearly state that a Bayesian 95% credibility interval indicates an interval in which they were justified in thinking that there was a 95% probability of finding the unknown parameter.

One may wonder why the Bayesian approach, which seems better suited for the social sciences and demography, has taken so long to gain acceptance among researchers in these domains. The first reason is the complexity of the calculations, which computers can now undertake. The example of Pierre-Simon de Laplace (1778), who presented the complex calculations and approximations (twenty pages mainly devoted to formulae) in order to solve, with the epistemic approach, a simple problem involving comparing the birth frequencies of girls and boys, is a good explanation of this reason. A second reason is a desire for an objective demography, drawing conclusions from data alone, with a minimal role for personal judgement.

Jakub Bijak was one of the first demographers to use Bayesian models, for migration forecasting (Bijak, 2010). He showed that the Bayesian approach can offer an umbrella framework for decision-making, by providing a coherent mechanism of inference. In this book, with his colleagues, he provides us with a more complete analysis of Bayesian modelling for demography.

Agent-Based or Model-Based Demography?

Social sciences, and more particularly demography, were launched by John Graunt (1662), just eight years later than the notion of probability was conceived. In his volume on the *Bills of Mortality*, Graunt used an objective probability model to estimate the age-specific probabilities of dying, under hypotheses that were rough, but the only conceivable ones at this time (Courgeau, 2012, pp. 28–34).

Later, Leonard Euler (1760) extended Graunt's model to the reproduction of the human species, introducing fertility and mortality. He used three hypotheses in order to justify his model. The first was based on the vitality specific to humans, measured by the probability of dying at each age for the members of a given population. These probabilities were assumed to remain the same in the future. The second hypothesis was based on the principle of propagation, which depended on marriage and fertility, measured by a rough approximation of fertility in a population. Again, these probabilities were to remain constant in the future. The third and last hypothesis was that the two principles of mortality and propagation are independent of each other. From these principles, Euler could calculate all the other probabilities that population scientists would want to estimate. Again, this model was computed under the objectivist probability assumptions and led to the concept of a stable population.

Later, in the twentieth century, Samuel Preston and Ansley Coale (1982) generalised this model to other populations, leading to a broader set of models of population dynamics: stable, semi-stable, and quasi-stable populations (Bourgeois-Pichat, 1994). These models were always designed assuming the objectivist interpretation of probability.

More recently, Francesco Billari and Alexia Prskawetz (2003) introduced the agent-based approach, already in use in many other disciplines (sociology, biology, epidemiology, technology, network theory, etc.) since 1970, to demography. This approach was first based on using objectivist probabilities, but more recently Bayesian inference techniques were introduced as an alternative methodology to analyse simulation models.

For Billari and Prskawetz, agent-based models pre-suppose the rules of behaviour and enable verifying, whether these micro-based rules can explain macroscopic regularities. Hence, these models start from pre-suppositions, as hypothetical theoretical models, but there is no clear way to construct these pre-suppositions, nor to verify if they are really explaining some macroscopic regularity. The choice of a behavioural theory hampers the widespread use of agent-based rules in demography, and depending on the selected theoretical model, the results produced by the agent-based model may be very different.

A second criticism of agent-based models had been formulated by John Holland (2012, p. 48). He said that "agent-based models offer little provision for agent conglomerates that provide building blocks and behaviour at higher orders of organisation." Indeed, micro-level rules find hardly a link with aggregate-level rules, and it seems difficult to think that macro-level rules may always be modelled with a micro approach: such rules generally transcend the behaviours of the component agents.

Finally, Rosaria Conte and colleagues (2012, p. 340) wondered, "how to find out the simple local rules? How to avoid *ad hoc* and arbitrary explanations? […] One criterion has often been used, i.e., choose the conditions that are sufficient to generate a given effect. However, this leads to a great deal of alternative options, all of which are to some extent arbitrary."

In front of these criticisms, this book gives preference to a model-based approach, which had already been proposed by us in Courgeau et al. (2016). This approach is based on the mechanistic theory, whereby sustained observations of some property of a population enable inferring a *functional structure*, which rules the process of generating this property. Without the inferred functional structure, this property could not come about as it does (Franck, 2002). It permits avoidance of some of the previous criticisms of agent-based models, but I will let the reader discover how the authors of this volume have improved further opportunities for constructing and verifying a mechanistic model of migration.

Conclusion

This historical and epistemological foreword on the two main and justified approaches relied on in this book by Jakub Bijak and his colleagues, Bayesian modelling and model-based demography, leaves aside many other important points that the reader will discover: migration theory, more particularly international migration theory; simulation in demography, with the very interesting set of Routes and Rumours models; cognition and decision making; computational challenges solved; replicability and transparency in modelling; and many more.

I greatly hope that that the reader will discover the importance of these approaches, not only for demography and migration studies but also for all other social sciences.

Institut national d'études démographiques
Paris, France
Daniel Courgeau

References

Bourgeois-Pichat, J. (1994). *La dynamique des populations. Populations stables, semi stables, quasi stables*. Institut national d'études démographiques, Presses Universitaires de France.

De Finetti, B. (1974). *Theory of probability* (Vol. 2). Wiley.

de Laplace, P.-S. (1780). Mémoire sur les probabilités. *Mémoires de l'Académie Royale des Sciences de Paris, 1781*, 227–332.

Euler, L. (1760). Recherches générales sur la mortalité et la multiplication du genre humain. *Histoire de l'Académie Royale des Sciences et des Belles Lettres de Berlin, 16*, 144–164.

Graunt, J. (1662). *Natural and political observations mentioned in a following index, and made upon the bills of mortality*. Tho. Roycroft for John Martin, James Allestry, and Tho. Dicas.

Pascal, B. (1654). *Traité du triangle arithmétique, avec quelques autres traités sur le même sujet*. Guillaume Desprez.

Pascal, B. (1670). *Pensées*. Editions de Port Royal.

Preston, S. H., & Coale, A. J. (1982). Age structure/growth, attrition and accession: A new synthesis. *Population Index, 48*(2), 217–259.

About the Authors

Lead Author

Jakub Bijak is Professor of Statistical Demography at the University of Southampton. He has a background in economics and over 20 years' work experience in academia and international civil service. His research focuses on demographic uncertainty, population and migration models and forecasts, and the demography of armed conflict. He has been awarded the Allianz European Demographer Award (2015) and the Jerzy Z Holzer Medal (2007) for work on migration modelling. Leader of a European Research Council project "Bayesian agent-based population studies" (www.baps-project.eu) and a Horizon 2020 project "QuantMig: Quantifying Migration Scenarios for Better Policy" (www.quantmig.eu). His email is j.bijak@soton.ac.uk.

Contributors

Philip A. Higham is Reader in Cognitive Psychology within Psychology at the University of Southampton. His research focuses on long-term human memory and metacognition (cognition about cognition), with one strand of his work focusing on the basic decision processes underlying metacognitive judgments. His email address is p.a.higham@soton.ac.uk.

Jason Hilton is Lecturer in Social Statistics and Data Science in the Department of Social Statistics and Demography at the University of Southampton. His research focuses on probabilistic projections of demographic processes, and on the management of uncertainty in agent-based demographic models. His email address is j.d.hilton@soton.ac.uk.

Martin Hinsch is a Research Fellow in the Department of Social Statistics and Demography at the University of Southampton. He is interested in emergent structures and complex systems and has conducted research in theoretical biology, bioinformatics, machine learning, epidemiology, and swarm robotics. His email address is hinsch.martin@gmail.com.

Sarah Nurse is a Research Fellow in the Department of Social Statistics and Demography at the University of Southampton. Her main research interests include migration, with a particular focus on forced migration and asylum, evaluation of migration data quality, and the demography of conflict and violence. Her email address is s.l.nurse@soton.ac.uk.

Toby Prike is a Research Fellow in the Department of Social Statistics and Demography at the University of Southampton. He holds a PhD in Psychology from Flinders University, Australia. His research looks at non-evidence-based beliefs, probabilistic reasoning, cognitive biases, and decision-making under uncertainty. His email address is t.m.prike@soton.ac.uk.

Oliver Reinhardt is a PhD student in the Modeling and Simulation Group at the University of Rostock. He holds an MSc in Computer Science from the University of Rostock. In his research, he is concerned with domain-specific modelling languages and the methodology of agent-based simulation. His email address is oliver.reinhardt@uni-rostock.de.

Peter W. F. Smith is Professor of Social Statistics in the Department of Social Statistics and Demography at the University of Southampton, and a Fellow of the British Academy. His research includes developing statistical methods for handling non-response, longitudinal data, and applications in demography, medicine, and health sciences. His email address is p.w.smith@soton.ac.uk.

Adelinde M. Uhrmacher is Professor at the Institute for Visual and Analytic Computing of the University of Rostock and head of the Modeling and Simulation Group. She holds a PhD in Computer Science from the University of Koblenz and a Habilitation in Computer Science from the University of Ulm. Her email address is adelinde.uhrmacher@uni-rostock.de.

Tom Warnke is a Research Associate in the Modeling and Simulation Group at the University of Rostock. He holds a PhD from the University of Rostock for work on domain-specific languages for modelling and simulation. His research focuses on languages for simulation experiments, design of modelling languages, and statistical model-checking. His email address is tom.warnke@uni-rostock.de.

Acknowledgements

The project BAPS: Bayesian Agent-Based Population Studies – Transforming Simulation Models of Human Migration, the results of which are presented in this book, has received generous funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 725232) for the period 2017–2022. Needless to say, this book reflects the authors' views, and the Research Executive Agency of the European Commission is not responsible for any use that may be made of the information it contains.

The project was executed by the University of Southampton, United Kingdom, and the University of Rostock, Germany. Our thanks go to our colleagues at the Centre for Population Change, University of Southampton, led by Jane Falkingham, for providing a lot of administrative, organisational and logistical support for various project activities – with special credits to Kim Lipscombe and Teresa McGowan for their continuing help and patience to our queries, however trivial. The authors also acknowledge the use of the IRIDIS High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of the work presented in this book.

A lot of the thinking behind this project, philosophical, methodological, as well as practical, emerged over several years of exchanging and pruning scientific ideas. Special thanks (in alphabetical order), for providing continuing inspiration, advice, critique and friendship, and for collaborating on many joint papers and other initiatives, go to Daniel Courgeau, Eric Silverman and Frans Willekens. The idea for including experimental data in agent-based models was stimulated by Jonathan (Jono) Gray, following the work on the Care Life Cycle project, together with Seth Bullock, Jason Noble and other colleagues, as well as by our discussions with Mathias Czaika. Our thoughts on uncertainty in agent-based modelling were, in turn, inspired by the work of David Banks, Adrian Raftery and Hana Ševčíková, and those on the role of models in demography and social science more generally by the work of Francesco Billari, Tom Burch, Robert Franck, Alexia Fürnkranz-Prskawetz, Anna Klabunde and Jan van Bavel, to whom our most sincere thanks.

In addition, some specific ideas presented in this book emerged through interactions and discussions with colleagues across different areas of modelling as well as migration research and practice. In particular, we are grateful to the organisers and participants of the following workshops: 'Uncertainty and Complexity of Migration', held in London on 20–21 November 2018; 'Rostock Retreat on Simulation', organised in Rostock on 1–3 July 2019; 'Agent-Based Models for Exploring Public Policy Planning', held at the Lorentz Center@Snellius in Leiden on 15–19 July 2019; and 'Modelling Migration and Decisions', organised in Southampton on 21 January 2020. Particular credit, in the non-individually attributable manner, compliant with the requirements of the Chatham House rule, goes to (also in alphabetical order): Rob Axtell, David Banks, Ann Blake, Alexia Fürnkranz-Prskawetz, André Grow, Katarzyna Jaśko, Leah Johnson, Nico Keilman, Ben Klemens, Elzemiek Kortlever, Giampaolo Lanzieri, Antonietta Mira, Petra Nahmias, Adrian Raftery, Hana Ševčíková, Eric Silverman, Ann Singleton, Vadim Sokolov, Sarah Wise, Teddy Wilkin and Dominik Zenner. When working on such a multidimensional topic as migration, having a multitude of perspectives to rely on has been invaluable, and we are grateful to everyone for sharing their views.

We are also indebted to Jo Russell for very careful proofreading and copyediting of the draft manuscript, and for helping us achieve greater clarity of the sometimes complex ideas and arguments. Naturally, all the remaining errors and omissions are ours alone.

On a private note, the lead author (Jakub Bijak) wishes to thank Kasia, Jurek and Basia for their support and patience throughout the writing and editing of this book during the long lockdown days of 2020–2021. Besides, having to explain several migration processes in a way that would be accessible for a year five primary student helped me better understand some of the arguments presented in this book.

Contents

List of Boxes

List of Figures

List of Tables

Part I
Preliminaries

Chapter 1
Introduction

Jakub Bijak

Population processes, including migration, are complex and uncertain. We begin this book by providing a rationale for building Bayesian agent-based models for population phenomena, specifically in the context of migration, which is one of the most uncertain and complex demographic processes. The main objectives of the book are to pursue methodological advancement in demography and migration studies through combining agent-based modelling with empirical data, Bayesian statistical inference, appropriate computational techniques, and psychological experiments in a streamlined modelling process, with the overarching aim to contribute to furthering the model-based research agenda in demography and broader social sciences. In this introductory chapter, we also offer an overview of the structure of this book, and present various ways in which different audiences can approach the contents, depending on their background and needs.

1.1 Why Bayesian Model-Based Approaches for Studying Migration?

Migration processes are characterised by large complexity and uncertainty, being some of the most uncertain drivers of population change (NRC, 2000). At the same time, migration is one of the most politically sensitive demographic phenomena in contemporary Europe (Castles et al., 2014). In a nutshell, migration is an increasingly more powerful driver of overall population dynamics across developed countries (Bijak et al., 2007; Castles et al., 2014), is socially and politically contentious, as well as being a top-priority, high-impact policy area (e.g. European Commission, 2015, 2020; UN, 2016). The so-called Syrian asylum crisis of 2015–16, and its impact on Europe and European policy and politics are prime examples of the urgent need for sound and robust scientific advice in this domain.

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_1

Unfortunately, theoretical foundations of migration remain weak and fragmented (Arango, 2000; McAuliffe & Koser, 2017), which is also to some extent true for other areas of demography (Burch, 2018). In the case of migration, tensions and trade-offs between high-level structural forces shaping the population flows and the agency of individual migrants are explicitly recognised as defining aspects of population mobility (de Haas, 2010; Carling & Schewel, 2018). Complex interrelations between various types of migration drivers operating at different levels – from individuals, to groups, to societies and states – call for more sophisticated methods of analysis than has been the case so far (Van Hear et al., 2018).

For all these reasons, among the different areas of population studies, there is a strong need to increase our understanding of migration processes. Addressing the challenges of the future requires the ability to comprehend and explain migration much better and more deeply than ever before. Currently, there is a gap between the demand for knowledge about migration, and the state of the art in this area.

From the point of view of quantitative population studies, especially those focused of human mobility, there is an acute need to fill a crucial void in formal modelling by offering new insights into the explanation of the underlying processes. Only in that way can social science help address important societal and population challenges: how the demographic processes, such as migration, can be better understood, predicted and managed. Previous efforts in that domain were largely constrained to simple approaches, with the explanatory endeavours lagging behind (for a review of formal modelling approaches from a predictive angle, see Bijak, 2010).

This book offers to fill this methodological void by presenting an innovative process for building simulation models of social processes, illustrated by an example of asylum migration, which aims to integrate behavioural and social theory with formal methods of analysis. Its key contribution is to combine in one book, novel methods and approaches of migration modelling, embedded in a joint analytical framework, while addressing some of the well-recognised philosophical challenges of model-based approaches. In particular, our main innovations include insights into human decisions and applying the formal rigour of statistical analysis to evaluate the modelling results. This combination offers novel and unique insights into some of the most challenging areas of demography and social sciences more broadly. It also bears a promise of influencing not only academics, but also practitioners and decision makers – in the area of migration and beyond – by offering methodological advice for policy-relevant simulations, and by providing a framework for decision support on their basis.

1.2 Aims and Scope of the Book

This book presents and reflects on the process of developing a simulation model of international migration route formation, with a population of intelligent, cognitive agents, their social networks, and policy-making institutions, all interacting with one another. The overarching aim of this work is to bring new insights into the theoretical and methodological foundations of demographic and migration studies, by proposing a blueprint for an interdisciplinary modelling process. In substantive terms, we aim at answering the following general question: how to introduce theoretical micro-foundations to demographic simulation studies, in particular, those of migration flows?

To that end, the book proposes a process for developing such micro-foundations for migration and other population studies through interdisciplinary efforts centred around agent-based modelling. The design of the modelling approach advocated in this volume follows recent developments in demography, computational modelling, statistics, cognitive psychology and computer science. In addition, we also offer a practical discussion on application of the proposed model-based approach by discussing a range of programming languages and environments.

In terms of the application area, the book sets out to address one of the most uncertain, complex and highest-impact population processes – international migration – which is situated at the intersection between demography and other social sciences. To address the challenges, we build on the existing literature from across a range of disciplines, incorporating in practice some of the ideas that have been proposed in terms of furthering the philosophical, theoretical and methodological perspectives involving computational social modelling.

Throughout this book, the methodological challenges of studying migration are thus addressed by bringing together interdisciplinary expertise from demography, statistics, cognitive psychology, as well as computer and complexity science. Combining them in a common analytical framework has a potential to move beyond the current state of affairs, which is largely developing in silos delineated by disciplinary boundaries (Arango, 2000). The proposed solutions can offer broader and generic methodological suggestions for analysing migration – a contemporary topic of global significance.

In particular, we offer a template for including in computational demographic models psychologically realistic micro-foundations, with an empirical basis – an aspect that is contemporarily lacking not only in migration research, but also in population studies more broadly. At the same time, the approach advocated here enables us to acknowledge and describe the fundamental epistemological limits of migration models in a formal way. To that end, some of the broader objectives of this programme of work include: identifying the inherently uncertain aspects of migration modelling, formally describing their uncertainty, providing policy recommendations under different levels of predictability of various processes, and finally offering guidance for further data collection.

In terms of the scope, the book discusses in detail the different stages and building blocks for constructing an empirically grounded simulation model of migration, and for embedding the modelling process within a wider framework of Bayesian experimental design. We use statistical principles to devise innovative computer-based simulation experiments, and to learn about the simulated processes as well as individual agents and the way they make decisions. The identified knowledge gaps are filled with information from dedicated psychological experiments on cognitive aspects of human decision making under uncertainty. In this way, the models are

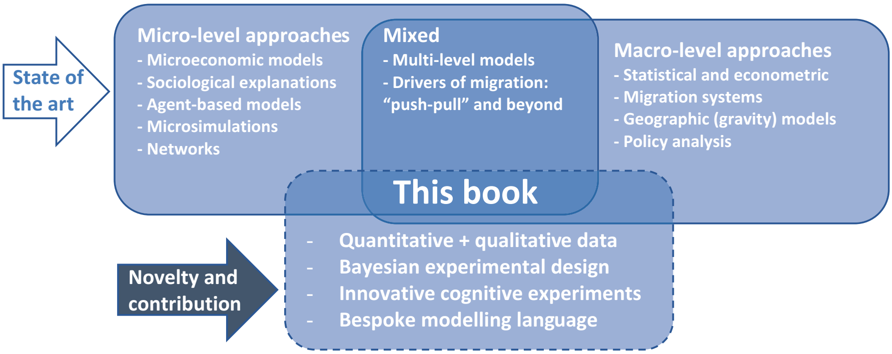


Fig. 1.1 Position of the proposed approach among formal migration modelling methods. (Source: own elaboration, based on Bijak (2010: 48))

built inductively, from the bottom up, addressing important epistemological limitations of population sciences.

The book builds upon the foundations laid out in the existing body of work, at the same time aiming to address the methodological and practical challenges identified in the recent population and migration modelling literature. Starting from a previous review of formal models of migration (Bijak, 2010), our proposed approach is specifically based on the five elements that have not been combined in modelling before. In particular, the existing micro-level approaches to migration studies, including microeconomic and sociological explanations, as well as inspirations from existing agent-based and microsimulation models, are combined here with macro-level statistical analysis of migration processes and outcomes, with the ultimate aim of informing decisions and policy analysis (see Fig. 1.1).

The novel elements included in this book additionally include combining qualitative and quantitative data in the formal modelling process (Polhill et al., 2010), learning about social mechanisms through Bayesian methods of experimental design, as well as including experimental information on human decision making and behaviour. Additionally, we develop further a dedicated programming language, ML3, to facilitate modelling migration, extending the earlier work in that area (Warnke et al., 2017). These different themes draw from the existing state of the art in migration modelling, and enhance it by adding new elements, as summarised in Fig. 1.1.

From the scientific angle, we aim to advance both the philosophical and practical aspects of modelling. This is done, first, by applying the concepts and ideas suggested in the contemporary literature to develop a model of migration routes in an iterative, multi-stage process. Second, these parallel aims are addressed by offering practical solutions for implementing and furthering the model-based research programme in demography (van Bavel & Grow, 2016; Courgeau et al., 2016; Silverman, 2018; Burch, 2018), and in social sciences more broadly (Hedström & Swedberg, 1998; Franck, 2002; Hedström, 2005).

The book draws inspiration from a wide literature. From a philosophical perspective, key ideas that underpin the theoretical discussions in this book can be found in Franck (2002), Courgeau (2012), Courgeau et al. (2016), Silverman (2018) and Burch (2018). The practical aspects of the many desired features of modelling involved, including the need for modular nature of model construction, were called for by Gray et al. (2017) and Richiardi (2017), while the need for additional, non-traditional sources of information, including qualitative and experimental data, was advocated by Polhill et al. (2010) and Conte et al. (2012), respectively.

At the same time, methods for a statistical analysis of computational experiments have also been discussed in many important reference works, for example in Santner et al. (2003). Specific applications of the existing statistical methods of analysing agent-based models can be found in Ševčíková et al. (2007), Bijak et al. (2013), Pope and Gimblett (2015) or Grazzini et al. (2017). The use of such methods – mainly Bayesian – have also been suggested elsewhere in the demographic literature, for example by Willekens et al. (2017). To that end, we propose a coherent methodology for embedding the model development process into a wider framework of Bayesian statistics and experimental design, offering a blueprint for an iterative process of construction and statistical analysis of computational models for social realms.

1.3 Structure of the Book

We have divided this book into three parts, devoted to: Preliminaries (Part I), Elements of the modelling process (Part II), and Model results, applications, and reflections (Part III). This structure enables different readers to focus on specific areas, depending on interest, without necessarily having to read the more technical details referring to individual aspects of the modelling process.

Part I lays down the foundations for the presented work. Chapter 2 focuses on the rationale and philosophical underpinnings of the Bayesian model-based approach. The discussion starts with general remarks on uncertainty and complexity in demography and migration studies. The uncertainty of migration processes is briefly reviewed, with focus on the ambiguities in the concepts, definitions and imprecise measurement; simplifications and pitfalls of the attempts at explanation; and on inherently uncertain predictions. A risk-management typology of international migration flows is revisited, focusing on asylum migration as the most uncertain and highest-impact form of mobility. In this context, we discuss the rationale for using computational models for asylum migration. To address the challenges posed by such complex and uncertain processes as migration, we seek inspiration in different philosophical foundations of demographic epistemology: inductive, deductive and abductive (inference to the best explanation). Against this background, we introduce a research programme of model-based demography, and evaluate its practical applicability to studying migration.

Part II presents five elements of the proposed modelling process – the building blocks of Bayesian model-based description and analysis of the emergence of migration routes. It begins in Chap. 3 with a high-level discussion of the process of developing agent-based models, starting from general principles, and then moving focus to the specific example of migration. We review and evaluate existing examples of agent-based migration models in the light of a discussion of the role of formal modelling in (social) sciences. Next, we discuss the different parts of migration models, including their spatial dimension, treatment of various sources of uncertainty, human decisions, social interactions and the role of information. The discussion is illustrated by presenting a prototype, theoretical model of migrant route formation and the role of information exchange, called Routes and Rumours, which is further developed in subsequent parts of the book, and used as a running example to illustrate different aspects of the model-building process. The chapter concludes by identifying the main knowledge gaps in the existing models of migration. This chapter is accompanied by Appendix A, where the architecture of the Routes and Rumours model is described in more detail.

Chapter 4 introduces the motivating example for the application of the Routes and Rumours model – asylum migration from Syria to Europe, linked to the so-called European asylum crisis of 2015–16. In this chapter, we present the process of constructing a dedicated knowledge base. The starting point is a discussion of various types of quantitative and qualitative data that can be used in formal modelling, including information on migration concepts, theories, factors, drivers and mechanisms. We also briefly present the case study of Syrian asylum migration. Subsequently, the data related to the case study are catalogued and formally assessed by using a common quality framework. We conclude by proposing a blueprint for including different data types in modelling. The chapter is supplemented by detailed meta-inventory and quality assessment of data, provided in Appendix B and available online, on the website of the research project Bayesian Agent-based Population Studies, underpinning the work presented throughout this book (www.baps-project.eu).

Chapter 5 is dedicated to presenting the general framework for analysing the results of computational models of migration. First, we offer a description of the statistical aspects of the model construction process, starting from a brief tutorial on uncertainty quantification in complex computational models. The tutorial includes Bayesian methods of uncertainty quantification; an introduction to experimental design; the theory of meta-modelling and emulators; methods for uncertainty and sensitivity analysis, as well as calibration. The general setup for designing and running computer experiments with agent-based migration models is illustrated by a running example based on the Routes and Rumours model introduced in Chap. 3. The accompanying Appendix C contains selected results of the illustrative uncertainty and sensitivity analysis presented in this chapter, as well as a brief overview of software packages for carrying out the experimental design and model analysis.

The cognitive psychological experiments are discussed in Chap. 6, following the rationale for making agent-based models more realistic and empirically grounded. Building on the psychological literature on decision making under uncertainty, the chapter starts with an overview of the design of cognitive experiments. This is followed by a presentation of three such experiments, focusing on discrete choice under uncertainty, elicitation of subjective probabilities and risk, and choice between leading migration drivers. We conclude the chapter by providing reflections on including the results of experiments in agent-based models, and the potential of using immersive interactive experiments in this context. Supplementary material included in Appendix D contains information on the study protocol and selected ethical aspects of experimental research and data collection.

Chapter 7, concluding the second part of the book, presents the computational aspects of the modelling work. We discuss the key features of domain-specific and general-purpose programming languages, by using an example of languages recently developed for demographic applications. In particular, the discussion focuses on modelling, model execution, and running simulation experiments in different languages. The key contributions of this chapter are to present a bespoke domain-specific language, aimed at combining agent-based modelling with simulation experiments, and formally describing the logical structure of models by using a concept of provenance modelling. Appendix E includes further information about the provenance description of the migration simulation models developed throughout this book, based on the Routes and Rumours template.

Part III offers a reflection on the selected outcomes of the modelling process and their potential scientific and policy implications. In particular, Chap. 8 is devoted to discussing the results of applying the model-based analytical template, combining all the building blocks listed above, and aimed at answering specific substantive research questions. We therefore follow the model development process, from the purely theoretical version to a more realistic one, called Risk and Rumours, subsequently including additional empirical and experimental data, in the version called Risk and Rumours with Reality. At the core of this chapter are the results of experiments with different models, and the analysis of their sensitivity and uncertainty. Subsequently, we reflect on the model-building process and computational implementation of the models, as well as their key limitations. The chapter concludes by exploring the remaining (residual) uncertainty in the models, and highlighting areas for future data collection. The underlying model architecture is an extension of the Routes and Rumours one, presented in Chap. 3 and Appendix A.

Subsequently, in Chap. 9, we outline the scientific and policy implications of modelling and its results. First, we discuss perspectives for furthering the model-based research agenda in social sciences, reflecting on the scientific risk-benefit trade-offs of the proposed approach. The usefulness of modelling for policy is then explored through a variety of possible uses, from scenario analysis, to foresight studies, stress testing and calibration of early warnings. To that end, we also present several migration scenarios, based on two models introduced in Chap. 8 (Risk and Rumours, and Risk and Rumours with Reality), aiming to simulate the impacts of actual policy decisions using an example of a risk-related information campaign. The chapter concludes with a discussion of the key limitations and practical recommendations for the users of the model-based approach.

The discussion in Chap. 10 focuses on the key role of transparency and replicability in modelling. Starting from a summary of the recent ‘replicability crisis’ in psychology, and lessons learned from this experience, we offer additional arguments for strengthening the formal documentation of the models constructed, including through the use of formal provenance modelling. The general implications for modelling and modellers, as well as for the users of models, are presented next.

Finally, the simulation results serve as a starting point for a broader reflection on the potential contribution of simulation-based approaches to migration research and social sciences generally. In that spirit, Chap. 11 concludes the book by summarising the theoretical, methodological and practical outcomes of the approach presented in the book in the light of recent developments in population and migration studies. We present further potential and limitations of Bayesian model-based approaches, alongside the lessons learned from implementing the modelling process proposed in the book. Key practical implications for migration policy are also summarised. As concluding thoughts, we discuss ways forward for developing statistically embedded model-based computational approaches, including an assessment of the viability of the whole model-based research programme.

1.4 Intended Audience and Different Paths Through the Book

The book is written by an interdisciplinary team with combined expertise in demography and migration studies, agent-based simulation modelling, statistical analysis and uncertainty quantification, experimental psychology and meta-cognition, as well as computer programming and simulations. We hope to demonstrate how adopting such a broad multidisciplinary approach within a common, rigorous and formal research framework opens up further exciting research possibilities in social sciences, and can help offer methodological recommendations for policy-relevant simulations. Practical applications are aided by intuitive programming advice for implementing and documenting the Bayesian model-based approach to answer real-life scientific and policy questions.

This book is primarily intended for academic and policy audiences, and aspires to stimulate new research opportunities. We hope that the presented work will be of interest to two types of academic readers. First, for demographers, sociologists, human geographers and migration scholars, it provides new methodological and philosophical insights into the possibilities offered by applying statistical rigour and empirical grounding of model-based approaches. In this way, we hope that computational demography – and demography and social sciences more generally – will benefit from engagement with new statistical, cognitive and computer science perspectives through formal, interdisciplinary modelling endeavours, which are offered throughout the whole book.

Second, for statisticians, complexity and computer scientists, as well as experimental psychologists, the book presents a case study of how the methods and approaches developed in their respective disciplines can be used elsewhere, under a common analytical umbrella. Demography can offer here a fascinating and contemporary area for the application of such research methods in a truly multi-disciplinary manner, opening up the scope for further methodological advancements. For such readers, the respective Chaps. 3, 4, 5, 6, and 7 are likely to be of interest, alongside Part III.

For non-academic readers from the areas of policy, government and civil service, working on migration, asylum, and in related domains, such as border protection, humanitarian aid, service provision, or human rights, the relevant outcomes are summarised primarily in Part III, tailored for practical applications. The focus of that part is on illustrating the possible uses of simulations by policy makers to test different scenarios concerning migration and related processes. Here, and particularly in Chap. 9, we present several ways to evaluate the efficacy of migration management measures through simulations and experimentation on a computer (*in silico*), under controlled, yet realistic conditions. More generally, such results can be of interest for policy think-tanks, government and parliamentary researchers, advisors, and independent experts as well.

Finally, the book can be used as supplementary reading for postgraduate courses, doctoral studies, and dedicated professional development training programmes, especially in the areas of formal and statistical demography, complexity science, or formal sociology. Here, we assume the prior knowledge of basic tenets of modelling and Bayesian statistics, and where relevant refer the readers to some of the key reference works and textbooks. Selected excerpts from the book, especially from Part I, can be also suited for final-year undergraduate courses in demography and complexity science, especially on methods-oriented programmes.

Chapter 2
Uncertainty and Complexity: Towards Model-Based Demography

Jakub Bijak

This chapter focuses on the broad methodological and philosophical underpinnings of the Bayesian model-based approach to studying migration. Starting from reflections on the uncertainty and complexity in demography and, in particular, migration studies, the focus moves to the shifting role of formal modelling, from merely describing, to predicting and explaining population processes. Of particular importance are the gaps in understanding asylum migration flows, which are some of the least predictable while at the same time most consequential forms of human mobility. The well-recognised theoretical void of demography as a discipline does not help, especially given the lack of empirical micro-foundations in formal modelling. Here, we analyse possible solutions to theoretical shortcomings of demography and migration studies from the point of view of the philosophy of science, looking at the inductive, deductive and abductive approaches to scientific reasoning. In that spirit, the final section introduces and extends a research programme of model-based demography.

2.1 Uncertainty and Complexity in Demography and Migration

The past, present, and especially the future size and composition of human populations are all, to some extent, uncertain. Population dynamics results from the interplay between the three main components of population change – mortality, fertility and migration – which differ with regard to their predictability. Long-term trends indicate that mortality is typically the most stable and hence the most predictable of the three demographic components. At the same time, the uncertainty of migration is the highest, and exhibits the most volatility in the short term (NRC, 2000).

Next to being uncertain, demographic processes are also complex in that they result from a range of interacting biological and social drivers and factors, acting in

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_2

non-linear ways, with human agency – and free will – exercised by the different actors involved. There are clear links between uncertainty and complexity: for mortality, the biological component is very high; contemporary fertility is a result of a mix of biological and social factors as well as individual choice; whereas migration – unlike mortality or fertility – is a process with hardly any biological input, in which human choice plays a pivotal role. This is one of the main reasons why human migration belongs to the most uncertain and volatile demographic processes, being as it is a very complex social phenomenon, with a multitude of underpinning factors and drivers.

On the whole, uncertainty in migration studies is pervasive (Bijak & Czaika, 2020). Migration is a complex demographic and social process that is not only difficult to conceptualise and to measure (King, 2002; Poulain et al., 2006), but also – even more – to explain (Arango, 2000), predict (Bijak, 2010), and control (Castles, 2004). Even at the conceptual level, migration does not have a single definition, and its conceptual challenges are further exacerbated by the very imprecise instruments, such as surveys or registers, which are used to measure it.

Historically, attempts to formalise the analysis of migration have been proposed since at least the seminal work of Ravenstein (1885). Contemporarily, a variety of alternative approaches co-exist, largely being compartmentalised along disciplinary boundaries: from neo-classical micro-economics, to sociological observations on networks and institutions (for a review, see Massey et al., 1993), or macro-level geographical studies of gravity (Cohen et al., 2008), to 'mobility transition' (Zelinsky, 1971) and unifying theories such as migration systems (Mabogunje, 1970; Kritz et al., 1992), or Massey's (2002) less-known synthesising attempt.

At the same time, the very notions of risk and uncertainty, as well as possible ways of managing them, are central to contemporary academic debates on migration (e.g. Williams & Baláž, 2011). Some theories, such as the *new economics of migration* (Stark & Bloom, 1985; Stark, 1991) even point to migration as an active strategy of risk management on the part of the decision-making unit, which in this case is a household rather than an individual. Similar arguments have been given in the context of environment-related migration, where mobility is perceived as one of the possible strategies for adapting to the changing environmental circumstances in the face of the unknown (Foresight, 2011).

Still, there is general agreement that none of the existing explanations offered for migration processes are fully satisfactory, and theoretical fragmentation is at least partially to blame (Arango, 2000). Similarly, given meagre successes of predictive migration models (Bijak et al., 2019), the contemporary consensus is that the best that can be achieved with available methods and data is a coherent, well-calibrated description of uncertainty, rather than the reduction of this uncertainty through additional knowledge (Bijak, 2010; Azose & Raftery, 2015). Due to ambiguities in migration concepts and definitions, imprecise measurement, too simplistic attempts at explanation, as well as inherently uncertain prediction, it appears that the demographic studies of migration, especially looking at macro-level or micro-level processes alone, have reached fundamental epistemological limits.

Recently, Willekens (2018) reviewed the factors behind the uncertainty of migration predictions, including the poor state of migration data and theories, additionally pointing to the existence of many motives for migration, difficulty in delineating migration versus other types of mobility, and the presence of many actors, whose interactions shape migration processes. In addition, the intricacies of the legal, political and security dimensions make international migration processes even more complex from an analytical point of view.

The existing knowledge gaps in migration research can be partially filled by explicitly and causally modelling the individuals (agents) and their decision-making processes in computer simulations (Klabunde & Willekens, 2016; Willekens, 2018). In particular, as advocated by Gray et al. (2016), the psychological aspects of human decisions can be based on data from cognitive experiments similar to those carried out in behavioural economics (Ariely, 2008). Some of the currently missing information can be also supplemented by collecting dedicated data on various facets of migration processes. Given their vast uncertainty, this could be especially important in the context of asylum migration flows, as discussed later in this chapter.

2.2 High Uncertainty and Impact: Why Model Asylum Migration?

Among the different types of migration, those related to various forms of involuntary mobility, violence-induced migration, including asylum and refugee movements, have the highest uncertainty and the highest potential impact on both the origin and destination societies (see, e.g. Bijak et al., 2019). Such flows are some of the most volatile and therefore the least predictable. They are often a rapid response to very unstable and powerful drivers, notably including armed conflict or environmental disasters, which lead people to leave their homes in a very short period (Foresight, 2011). Despite the involuntary origins, different types of forced mobility, including asylum migration, like all migration flows, also prominently feature human agency at their core: this is well known both from scholarly literature (Castles, 2004), as well as from journalistic accounts of migrant journeys (Kingsley, 2016).

As a result, and also because it is difficult to disentangle asylum migration from other types of mobility precisely, involuntary flows evade attempts at defining them in precise terms. Of course, many definitions related to specific populations of interest exist, beginning with the UN designation of a refugee, following the 1951 Convention and the 1967 Protocol, as someone who:

> "owing to well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his [*sic!*] nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country; or who, not having a nationality and being outside the country of his former habitual residence as a result of such events, is unable or, owing to such fear, is unwilling to return to it." (UNHCR, 1951/1967; Art. 1 A (2))

The UN definition is relatively narrow, being restricted to people formally recognised as refugees under international humanitarian law, even though the explicit inclusion of the notion of *fear* can help better conceptualise violence-induced migration (Kok, 2016). Broader definitions, such as those of forced displacement, range from more to less restrictive; for example, according to the World Bank:

> "forcibly displaced people [include] refugees, internally displaced persons and asylum seekers who have fled their homes to escape violence, conflict and persecution" (World Bank; http://www.worldbank.org/en/topic/forced-displacement, as of 1 September 2021).

On the other hand, the following definition of the International Association for the Study of Forced Migration (IASFM), characterises forced migrations very broadly, as:

> "Movements of refugees and internally displaced people (displaced by conflicts) as well as people displaced by natural or environmental disasters, chemical or nuclear disasters, famine, or development projects" (after Forced Migration Review; https://www.fmreview.org, as of 1 September 2021).

In several instances, pragmatic solutions are needed, so that the definition is actually determined by what can be measured, or what can be subsequently used for operational purposes by the users of the ensuing analysis. The same principle can hold for the drivers of migration and how they can be operationalised. In that spirit, Bijak et al. (2017) defined asylum-related migration as follows:

> "Asylum-related migration has therefore to jointly meet two criteria: first, it needs to be international in nature, and second, it has to be – or claimed to be – related to forced displacement, defined as forced migration due to persecution, armed conflict, violence, or violations of human rights" (Bijak et al., 2017, p.8).

This definition excludes internally displaced persons, and migrants forced to move for environment- or development-related reasons. It was also purely driven by the operational needs of the European asylum system, which was the intended user of the related analysis. For similar reasons, we use the term 'asylum migration' throughout this book, as most closely aligned with the substantive research questions that we aim to study through the lens of the model-based approach. To that end, the focus of our modelling efforts, and their possible practical applications, is on understanding the dynamics of the actual flows of people, irrespective of their legal status or specific individual circumstances.

More generally, even if a common definition could be adopted, at the higher, conceptual level, the dichotomy between forced and voluntary migration seems to some extent obsolete and not entirely valid. This is mainly attributed to the presence of a multitude of migration motives operating at the same time for a single migrant (King, 2002; Foresight, 2011; Erdal & Oeppen, 2018). The uncertainty of asylum migration is additionally exacerbated by a lack of common theoretical and explanatory framework. The aforementioned theoretical paucity of migration studies in general does not help (Arango, 2000), and the situation with respect to asylum migration is similarly problematic. Besides, in the contemporary literature there is vast disconnect between migration and refugee studies, which utilise different theoretical approaches and do not share many common insights (FitzGerald, 2015). Comprehensive theoretical treatment of different types of migration on the voluntary-forced spectrum is rare; with examples including the important work by Zolberg (1989).

One pragmatic solution can be to focus on various factors and drivers of migration, an approach systematised in the classical *push-pull* framework of Everett Lee (1966), and since extended by many authors, including Arango (2000), Carling and Collins (2018), or Van Hear et al. (2018). Specifically in the context of forced migration, Öberg (1996) mentioned the importance of 'hard factors', such as conflict, famine, persecution or disasters, pushing involuntary migrants out from their places of residence, and leading to resulting migration flows being less self-selected. A contemporary review of factors and drivers of asylum-related migration was published in the EASO (2016) report, while a range of economic aspects of asylum were reviewed by Suriyakumaran and Tamura (2016).

In addition, uncertainty of asylum migration measurement includes many idiosyncratic features, besides those common with other forms of mobility. In particular, focus on counting administrative events rather than people results in limited information being available on the context and on migration processes themselves (Singleton, 2016). As a result, on the one hand, some estimates include duplications of the records related to the same persons; while on the other hand, some of the flows are at the same time undercounted due to their clandestine nature (*idem*).

The politicisation of asylum statistics, and their uses and misuses to fit with any particular political agenda, are other important reasons for being cautious when interpreting the numbers of asylum migrants (Bakewell, 1999; Crisp, 1999). Contemporary attempts to overcome some of the measurement issues are currently undertaken through increasing use of biometric techniques, such as the EURODAC system in the European Union (Singleton, 2016), as well as through experimental work with new data, such as mobile phone records or 'digital footprints' of social media usage (Hughes et al., 2016). This results in a patchwork of sources covering different aspects of the flows under study, as illustrated in Chap. 4 on the example of Syrian migration to Europe.

Despite these very high levels of uncertainty, formal quantitative modelling of various forms of asylum-related migration remains very much needed. Its key uses are both longer-term policy design, as well as short-term operational planning, including direct humanitarian responses to crises, provision of food, water, shelter and basic aid. In this context, decisions under such high levels of uncertainty require the presence of contingency plans and flexibility, in order to improve resilience of the migration policies and operational management systems. This perspective, in turn, requires new analytical approaches, the development of which coincides with a period of self-reflection on the theoretical state of demography, or broader population studies, in the face of uncertainty (Burch, 2018). These developments are therefore very much in line with the direction of changes of the main aims of demographic enquiries over the past decades, which are briefly summarised next.

2.3 Shifting Paradigm: Description, Prediction, Explanation

To trace the changes in demographic thinking about the notion of uncertainty, we need to go back to the very inception of the discipline in the seventeenth century, notionally marked by the publication of John Graunt's *Bills of Mortality* in 1662. From the outset, demography had an uneasy relationship with uncertainty and, by extension, with probability theory and statistics (Courgeau, 2012). Following a few early examples of probabilistic studies of the features of populations, the nineteenth century and the increased reliance on population censuses brought about the dominance of descriptive, and largely deterministic approaches. In that period, the questions of variation and uncertainty were largely swept under the carpet (*idem*).

Similarly, the proliferation of survey methods and data in the second half of the twentieth century offered some simple explanations of demographic phenomena in terms of statistical relationships, which still remained largely descriptive, and were missing the mechanisms actually driving population change (Courgeau et al., 2016; Burch, 2018). Only recently, especially since the 1970s and 1980s, has statistical demography begun to flourish, including a range of methods and models that apply the Bayesian paradigm, and put uncertainty at the centre of population enquiries, in such areas as prediction, small area estimation, or complex and highly-structured problems (Bijak & Bryant, 2016).

Population predictions, with their inherent uncertainty, are contemporarily seen as one of the bestselling products of population sciences (Xie, 2000). In assessing their analytical potential, Keyfitz (1972, 1981) put a reasonable horizon of population predictions at one generation ahead at most, which is already quite long, especially in comparison with other socio-economic phenomena. Within that period, the newly-born generations have not yet entered the main reproductive ages. The cohort-component mechanism of population renewal additionally ensures the relatively high levels of predictability at the population level (Lutz, 2012; Willekens, 2018): most people who will be present in a given population one generation ahead are already there.

What can reduce the predictability of population, especially in the short term, is migration, the predictive horizon of which is much shorter (Bijak & Wiśniowski, 2010), unless it is described and modelled at a very high level of generality, with very low-frequency data (Azose & Raftery, 2015). The migration uncertainty is also age-selective, affecting the more mobile age groups, such as people in the early stages of their labour market activity, more than others. This uncertainty is further amplified from generation to generation, through secondary impacts of migration on fertility and mortality rates, and through changes in the composition of populations in both origin and destination countries (for an example related to Europe, see Bijak et al., 2007).

The unpredictability of migration compounds two types of uncertainty: *epistemic*, related to imperfect knowledge, and *aleatory*, inherent to any future events, especially for complex social systems (for a detailed discussion, see Bijak & Czaika, 2020). Some migration flows are more uncertain than others, and require different analytical tools and different assumptions on their statistical properties, such as stationarity. For some processes, or over longer horizons, coherent scenarios seem to be the only reliable way of scanning the possible future pathways (see Nico Keilman's contribution to Willekens, 1990: 42–44; echoed by Bijak, 2010). Ideally, such scenarios should be equipped with solid micro-level foundations and connect different levels of analysis, from micro (individuals), to macro (populations).

Another way to describe the uncertainty of migration flows is offered by the risk management framework, with uncertainty or volatility of a specific migration type juxtaposed against its possible societal impact (Bijak et al., 2019). Under this framework, return migration of nationals is typically less volatile – and has smaller political or societal impact – than for example labour immigration of non-nationals. Seen through the lens of risk management, the violence-induced migration, including large flows of asylum seekers, refugees and displaced persons, is typically one of the most uncertain forms of mobility, also characterised by the highest societal impact (for a conceptual overview aimed at improving forecasts, see also Kok, 2016). For such highly unpredictable types of migration, early warning models may offer some predictive insights over very short horizons (Napierała et al., 2021).

Besides, despite the advances in statistical modelling, formal description and interpretation of uncertain demographic phenomena, one key epistemological gap in contemporary demography remains: the lack of explanation of the related processes, which can be especially well seen in the studies of migration. Particularly missing are solid theoretical foundations underlying the macro-level processes (see for example Burch, 2003, 2018). Numerous micro-level studies based on surveys exist, but they do not deal with the *behaviour* of individuals, only with its observable and measurable outcomes. Even the prevailing event-history and multi-level statistical studies do not offer *causal explanations* of the mechanisms driving demographic change (Courgeau et al., 2016).

In mainstream population sciences, the discussion of micro-foundations of macro-level processes has been so far very limited. Even though the importance of explicit modelling of micro-level behaviour of individuals has been acknowledged in a few pioneering studies, such as the landmark volume by Billari and Prskawetz (2003) and its intellectual descendants and follow-ups (Billari et al., 2006; van Bavel & Grow, 2016; Silverman, 2018), the associated demographic agent-based models are still in their infancy, and their theory-building and thus explanatory potential has not yet been fully accomplished, as documented in Chap. 3 on the example of migration modelling.

At the same time, various types of computational simulation models have been gaining prominence in population studies since the beginning of the twenty-first century (Axtell et al., 2002; Billari & Prskawetz, 2003; Zaidi et al., 2009; Bélanger & Sabourin, 2017), and research on the applications of computational modelling approaches to population problems is currently gaining momentum (van Bavel & Grow, 2016; Silverman, 2018). This is because computer-based simulations, such as agent-based or microsimulation models, offer population scientists many new and exciting research possibilities. At the same time, demography remains a strongly empirical area of social sciences, with many policy implications (Morgan & Lynch, 2001), for which computational models can offer attractive analytical tools.

So far, the empirical slant has constituted one of the key strengths of demography as a discipline of social sciences; however, there is increasing concern about the lack of theories explaining the population phenomena of interest (Burch, 2003, 2018). This problem is particularly acute in the case of the micro-foundations of demography being largely disconnected from the macro-level population processes (Billari, 2015). The quest for micro-foundations, ensuring links across different levels of the problem, thus becomes one of the key theoretical and methodological challenges of contemporary demography and population sciences.

2.4 Towards Micro-foundations in Migration Modelling

In order to be realistic and robust, migration (or, more broadly, population) theories and scenarios need to be grounded in solid micro-foundations. Still, in the uncertain and messy social reality, especially for processes as complex as migration, the modelling of micro-foundations of human behaviour has its natural limits. In economics, Frydman and Goldberg (2007) argued that such micro-foundations may merely involve a qualitative description of tendencies, rather than any quantitative predictions. Besides, even in the best-designed theoretical framework, there is always some residual, irreducible aleatory uncertainty. Assessing and managing this uncertainty is crucial in all social areas, but especially so in the studies of migration, given its volatility, impact and political salience (Disney et al., 2015).

In other disciplines, such as in economics, the acknowledgement of the role of micro-foundations has been present at least since the *Lucas critique* of macroeconomic models, whereby conscious actions of economic agents invalidate predictions made at the macro (population) level (Lucas, 1976). The related methodological debate has flourished for over at least four decades (Weintraub, 1977; Frydman & Goldberg, 2007). The response of economic modelling to the Lucas critique largely involved building large theoretical models, such as those belonging to the Dynamic Stochastic General Equilibrium (DSGE) class, which would span different levels of analysis, micro – individuals – as well as macro – populations (see e.g. Frydman & Goldberg, 2007 for a broad theoretical discussion, and Barker & Bijak, 2020 for a specific migration-related overview).

Existing migration studies offer just a few overarching approaches with a potential to combine the micro and macro-level perspectives: from multi-level models, that belong to the state of the art in statistical demography (Courgeau, 2007), to conceptual frameworks that potentially encompass micro-level as well as macro-level migration factors. The key examples of the latter include the push and pull migration factors (Lee, 1966), with recent modifications, such as the push-pull-plus framework (Van Hear et al., 2018), and the value-expectancy model of De Jong and Fawcett (1981). In the approach that we propose in this book, however, the link between the different levels of analysis is of statistical and computational nature, rather than being analytical or conceptual. In particular, in our approach, bridging the gap between the different levels of analysis involves building micro-level simulation models of migration behaviour, which can then be calibrated to some aspects of macro-level data.

One alternative approach for combining different levels of analysis involves building microsimulation models, whereby simulated individuals are subject to transitions between different *states* according to empirically derived rates, which are typically data-driven (Zaidi et al., 2009; Bélanger & Sabourin, 2017). Such models can be limited by the availability of detailed data, and often follow simple assumptions on the underlying mechanisms, for example Markovian 'lack of memory' (Courgeau et al., 2016). In contrast, agent-based models, based on interacting individual agents, allow for explicit inclusion of feedback effects and modelling the bidirectional impact of macro-level environment on individual behaviour and vice versa through the 'reverse causality' mechanisms (Lorenz, 2009). Still, it is recognised that many of the existing agent-based attempts are too often based on unverifiable assumptions and axioms (Conte et al., 2012).

Agent-based models focus on representing the behaviour of simulated individuals – agents – in artificial computer simulations, through applying micro-level behavioural rules to study the resulting patters emerging at the macro level. Such models, while not predictive *per se*, can be used for a variety of objectives. Epstein (2008) identified sixteen aims of modelling, from explanation, to guiding data collection, studying the range of possible outcomes, and engagement with the public. The perspective of generating explanatory mechanisms for migration through simulations and model-building, and enabling experimentation in controlled conditions *in silico*, are both very appealing to demographers (Billari & Prskawetz, 2003), and potentially also to the users of their models, including policy makers. We explore many of these aspects throughout this book.

Given the state of the art of demographic modelling, important methodological advances can be therefore achieved by building agent-based simulation models of international migration, combined in a common framework with the recent cutting-edge developments across a range of disciplines, including demography, statistics and experimental design, computer science, and cognitive psychology, the latter shedding light on the specific aspects of human decision making. This approach can enhance the traditional demographic modelling of population-level dynamics by including realistic and cognitively plausible micro-foundations.

There are several important examples of work which look at applications of agent-based modelling to social science, beginning with the seminal work of Schelling (1971, 1978). More recently, a specialised field of social simulation has emerged (Epstein & Axtell, 1996; Gilbert & Tierna, 2000), as has the analytical sociology research programme (Hedström & Swedberg, 1998; Hedström, 2005). Recently, the topic was explored, and the field thoroughly reviewed by Silverman (2018). As mentioned above, the pioneering demographic book advocating the use of agent-based models (Billari & Prskawetz, 2003) was followed by subsequent extensions and updates (e.g. Billari et al., 2006; van Bavel & Grow, 2016). In parallel, microsimulation models have been developed and extensively applied (for an overview, see e.g. Zaidi et al., 2009; Bélanger & Sabourin, 2017). In migration research, several examples of constructing agent-based models exist, such as Kniveton et al. (2011) or Klabunde et al. (2017), with a more detailed survey of such models offered in Chap. 3.

In general, agent-based models have complex and non-linear structures, which prohibit a direct analysis of their outcome uncertainty. Promising methods which could enable indirect analysis include Gaussian process (GP) emulators or meta-models – statistical models of the underlying computational models (Kennedy & O'Hagan, 2001; Oakley & O'Hagan, 2002), or the Bayesian melding approach (Poole & Raftery, 2000), implemented in agent-based transportation simulations (Ševčíková et al., 2007). In demography, prototype GP emulators have been tested on agent-based models of marriage and fertility (Bijak et al., 2013; Hilton & Bijak, 2016). A general framework for their implementation is that of (Bayesian) statistical experimental design (Chaloner & Verdinelli, 1995), with other approaches that can be used for estimating agent-based models including, for example, Approximate Bayesian Computations (Grazzini et al., 2017). A detailed discussion, review and assessment of such methods follows in Chap. 5.

Before embarking on the modelling work, it is worth ensuring that the outcomes – models – have realistic potential for increasing our knowledge and understanding of demographic processes. The discussion about relationship between modelling and the main tenets of the scientific method remains open. To that end, we discuss the epistemological foundations of model-based approaches next, with focus on the question of the origins of knowledge in formal modelling.

2.5 Philosophical Foundations: Inductive, Deductive and Abductive Approaches

There are several different ways of carrying out scientific inference and generating new knowledge. The deductive reasoning has been developed through millennia, from classical syllogisms, whereby the conclusions are already logically entailed in the premises, to the hypothetico-deductive scientific method of Karl Popper (1935/1959), whereby hypotheses can be falsified by non-conforming data. The deductive approaches strongly rely on hypotheses, which are dismissed by the proponents of the inductive approaches due to their arbitrary nature (Courgeau et al., 2016).

The classical inductive reasoning, in turn, which underpins the philosophical foundations of the modern scientific method, dates back to Francis Bacon (1620). It relies on inducing the formal *principles* governing the processes or phenomena of interest (Courgeau et al., 2016), at several different levels of explanation. These principles, in turn, help identify the key *functions* of the processes or phenomena, which are required for these processes or phenomena to occur, and to take such form as they have. The identified functions then guide the observation of the empirical properties, so that in effect, the observed variables describing these properties can illuminate the *functional structures* of the processes or phenomena as well as the *functional mechanisms* that underpin them.[1]

When it comes to hypotheses, the main problem seems to be not so much their existence, but their haphazard and often not properly justified provenance. To help address this criticism, a third, less-known way of making scientific inference has been proposed: *abduction*, also referred to as 'inference to the best explanation'. The idea dates back to the work of Charles S. Peirce (1878/2014), an American philosopher of science working in the second half of the nineteenth century and the early twentieth century. His new, pragmatic way of making a philosophical argument can be defined as "inference from the body of data to an explaining hypothesis" (Burks, 1946: 301).

Seen in that way, abduction appears as a first phase in the process of scientific discovery, with setting up a novel hypothesis (Burks, 1946), whereas deduction allows subsequently for deriving testable consequences, while modern induction allows their testing, for example through statistical inference. As an alternative classification, Lipton (1991) labelled abduction as a separate form of inductive reasoning, offering 'vertical inference' (*idem*: 69) from observable data to unobservable explanations (theory), allowing for the process of discovery. The consequences of the latter can subsequently follow deductively (*idem*). Thanks to the construction and properties of abductive reasoning, this perspective has found significant following within the social simulation literature, to the point of equating the methods with the underpinning epistemology. To that end, Lorenz (2009: 144) explicitly stated that "simulation model is an abductive process".

Some interpretations of abductive reasoning stress the pivotal role it plays in the sequential nature of the scientific method, as the stage where new scientific ideas come from in a process of creativity. At the core of the abductive process is surprise: observing a surprising result leads to inferring the hypothesis that could have led to its emergence. In this way, the (prior) beliefs, confronted by a surprise, lead to doubt and enable further, creative inference (Burks, 1946; Nubiola, 2005), which in itself has some conceptual parallels with the mechanism of Bayesian statistical knowledge updating.

There is a philosophical debate as to whether the emergence of model properties as such is of ontological or epistemological nature. In other words, whether modelling can generate new facts, or rather help uncover the patterns through improved knowledge about the mechanisms and processes (Frank et al., 2009). The latter interpretation is less restrictive and more pragmatic (*idem*), and thus seems better suited for social applications. As an example, in demography, a link between discovery (surprise) and inference (explanation) was recently established and

[1] The notion of classical induction is different from the concept of induction as developed for example by John Stuart Mill, where observables are generalised into conclusions, by eliminating those that do not aid the understanding of the processes under study, for example in the process of experimenting (Jacobs 1991). The two types of induction should not be confused. On this point, I am indebted to Robert Franck and Daniel Courgeau for detailed philosophical explanations.

formalised by Billari (2015), who argued that the act of discovery typically occurs at the population (macro) level, but explanation additionally needs to include individual (micro)-level foundations.

Abduction, as 'inference to the best explanation', is also a very pragmatic way of carrying out the inferential reasoning (Lipton, 1991/2004). What is meant by the 'best explanation' can have different interpretations, though. First, it can be the best of the candidate explanations of the probable or approximate truth. Second, it can be subject to an additional condition that the selected hypothesis is satisfactory or 'good enough'. Third, it can be such an explanation, which is 'closer to the truth' than the alternatives (Douven, 2017).

The limitations of all these definitions are chiefly linked to a precise definition of the criterion for optimality in the first case, satisfactory quality criteria in the second, as well as relative quality and the space of candidate explanations in the third. One important consideration here is the parsimony of explanation – the Ockham's razor principle would suggest preferring simple explanations to more complex ones, as long as they remain satisfactory. Another open question is which of these three alternative definitions, if any, are actually used in human reasoning (Douven, 2017)?

In any case, a lack of a single and unambiguous answer points out to lack of strict identifiability of abductive solutions to particular inferential problems: under different considerations, many candidate explanations can be admissible, or even optimal. This ambiguity is the price that needs to be paid for creativity and discovery. As pointed out by Lorenz (2009), abductive reasoning bears the risk of an *abductive fallacy*: given that abductive explanations are sufficient, but not necessary, the choice of a particular methodology or a specific model can be incorrect.

These considerations have been elaborated in detail in the philosophy of science literature. In his comprehensive treatment of the approach, Lipton (1991/2004) reiterated the pragmatic nature of inference to the best explanation, and made a distinction between two types of reasoning: 'likeliest', being the most probable, and 'loveliest', offering the most understanding. The former interpretation has clear links with the probabilistic reasoning (Nubiola, 2005), and in particular, with Bayes's theorem (Lipton, 2004; Douven, 2017). This is why abduction and Bayesian inference can be even seen to be 'broadly compatible' (Lipton, 2004: 120), as long as the elements of the statistical model (priors and likelihoods) are chosen based on how well they can be thought to explain the phenomena and processes under study. In relation to the discussion of psychological realism of the models of human reasoning and decision making (e.g. Tversky & Kahneman, 1974, 1992), formal Bayesian reasoning can offer rationality constraints for the heuristics used for updating beliefs (Lipton, 2004).

There are important implications of these philosophical discussions both for modelling, as well as for practical and policy applications. To that end, Brenner and Werker (2009) argued that simulation models built by following the abductive principles at least partially have a potential to reduce the error and uncertainty in the outcome. In particular, looking at the modelled structures of the policy or practical problem can help safeguard against at least some of the unintended and undesirable consequences (*idem*), especially when they can be identified through departures from rationality.

In that respect, to help models achieve their full potential, the different philosophical perspectives need to be ideally combined. As deduction on its own relies on assumptions, induction implies uncertainty, and abduction does not produce uniquely identifiable results, the three perspectives should be employed jointly, although even then, uncertainty cannot be expected to disappear (Lipton, 2004; Brenner & Werker, 2009). These considerations are reflected in the nascent research programme for model-based demography, the main tenets of which we discuss in turn.

2.6 Model-Based Demography as a Research Programme

The methodology we propose throughout the book is inspired by the principles of the *model-based research programme* for demography, recently outlined by Courgeau et al. (2016), who were inspired by Franck (2002). In parallel, similar propositions have been developed by other prominent authors, such as Burch (2018), in a tradition dating back to Keyfitz (1971). Among the different approaches to demographic modelling, Courgeau et al. (2016) suggested that the model-building process should follow the classical inductive principles from the bottom up. In this way, the process should start by observing the key population properties generated by the process under study (migration), followed by inferring the *functional structures* of these processes in their particular context, identifying the relevant variables, and finally conceptual and computational modelling. The results of the modelling should allow for identifying gaps in current knowledge and provide guidance on further data collection. By so doing, the process can be iterated as needed, as argued by Courgeau et al. (2016), ideally following the broad principles of classical inductive reasoning.

It is worth stressing that the proposed model-based programme is not the same as an approach that relies purely on agent-based modelling. First, the model-based approaches can involve different types of models: agent-based ones are an obvious possibility, but microsimulations or formal mathematical models can also be used, alongside the statistical models used to unravel the properties of analytical or computational models they are meant to analyse. Second, as argued in Chap. 3, agent-based models alone, especially those applied to social processes such as migration, necessarily have to make many arbitrary and *ad hoc* assumptions, unless they can be augmented with additional information from other sources – observations, experiments, and so on – as proposed in the full model-based approach advocated here. From that point of view, the model-based approach includes a (computational or analytical) model at its core, but goes beyond that – and the *process* of arriving at the final form of the model is also much more involved than the programming of a model alone.

The existing agent-based attempts at describing migration, reviewed and evaluated in more detail in Chap. 3, offer a good starting point for the model-building process. In particular, Klabunde et al. (2015) looked at the staged nature of the decision process, following the Theory of Planned Behaviour (Ajzen, 1985), whereby behaviour results from intentions, formed on the basis of beliefs, norms and attitudes, and moderated by actual behavioural control. None of the existing approaches, however, explicitly represent key cognitive aspects of decision-making mechanisms, nor do they include a comprehensive uncertainty assessment at the different levels of analysis. Our proposed model-based approach offers insights into bottom-up modelling based on a range of information sources, addressing some of the key epistemological limitations of simulations, especially of human decisions.

There are many other building blocks that can facilitate modelling: importantly, despite high uncertainty, migration is characterised by stable regularities in terms of its spatial structures (Rogers et al., 2010) and age profiles (Rogers & Castro, 1981). The latter is an outcome of links with life course and other demographic processes, such as family formation or childbearing (Courgeau, 1985; Kulu & Milevski, 2007). The role of migrant networks in the perpetuation of migration processes is also well recognised (Kritz et al., 1992; Lazega & Snijders, 2016). For such elements – networks and *linked lives* – agent-based models are a natural tool of scientific enquiry (Noble et al., 2012). Following the general philosophy of Ben-Akiva et al. (2012), it is also worthwhile distinguishing the *process* of migration decision making at the individual level, and the *context* at the group and societal levels, integrated within a common multi-level analytical model. A joint modelling of different levels of analysis was also suggested in the *Manifesto of computational social science* by Conte et al. (2012). In the same work, Conte et al. (2012) suggested that computational social science modelling should be more open to non-traditional sources of data, beyond surveys and registers, and in particular embrace tailor-made experimentation under controlled conditions.

Many of these different elements are used in the application of the model-based approach presented throughout this book. The empirical experiments focus on different aspects of human decision-making processes, such as choices between different options (Ben-Akiva et al., 2012), the role of uncertainty – especially the subjective probabilities and possible biases – as well as attitudes to risk (Gray et al., 2017), which are discussed in more detail in Chap. 6. In this way, the purpose of a scientific enquiry becomes as much about the model and the related analysis, as it is about the process of the iterative improvement of the analytical tools and an increase in their sophistication. In philosophical terms, the proposed approach also addresses the methodological suggestions made by Conte et al. (2012) that different types of empirical data are used throughout the model construction process, not merely for final validation, which is understood here as ensuring alignment between the model and some aspects of the observed reality.

Nevertheless, one important challenge of designing and implementing such a modelling process remains: how to combine simulations with other analytical methods, including statistics, as well as experiments, with a strong empirical base (Frank et al., 2009)? To that end, Courgeau et al. (2016) stressed the role of appropriate experimental design and related statistical methods to bring the different methodological threads together, and to align model-based enquiries closer with the classical inductive scientific research programme, dating back to Francis Bacon (1620;

after: *idem*). The broad tenets of this approach are followed throughout this book, and its individual components are presented in Part II.

In the model-based programme, as proposed by Courgeau et al. (2016), the objective of modelling is to *infer* the functional structures that generate the observed social properties. Here, the empirical observables are necessary, but not sufficient elements in the process of scientific discovery, given that for any set of observables, there can be a range of non-implausible models generating matching outcomes (*idem*). At the same time, as noted by Brenner and Werker (2009), the modelling process needs to explicitly recognise that the errors in inference are inevitable, but modellers should aim to reduce them as much as possible.

In what can be seen as a practical solution for implementing a version of the model-based programme, Brenner and Werker (2009:3.6) advocated four steps of the modelling process:

(1) Setting up the model based on all available empirical knowledge, starting from a simple variant, and allowing for free parameters, wherever data are not available (abduction);
(2) Running the model and calibrating it against the empirical data for some chosen outputs, excluding the implausible ranges of the parameter space (induction, in the modern sense);
(3) On that basis, classifying observations into classes, enabling alignment of theoretical explanations implied by the model structure with empirical observations (another abduction);
(4) Use of the calibrated model for scenario and policy analysis (which *per se* is a deductive exercise, notwithstanding the abductive interpretation given by Brenner & Werker, 2009).

In this way, the key elements of the model-based programme become explicitly embedded in a wider framework for model-based policy advice, which makes full use of three different types of reasoning – inductive, abductive and deductive – at three different stages of the process. Additionally, the process can implicitly involve two important checks – *verification* of consistency of the computer code with the conceptual model, and *validation* of the modelling results against the observed social phenomena (see David, 2009 for a broad discussion).

As a compromise between the ideal, fully inductive model-based programme advocated by Courgeau et al. (2016) and the above guidance by Brenner and Werker (2009), we propose a pragmatic variant of the model-based approach, which is summarised in Fig. 2.1. The modelling process starts by defining the specific research question or policy challenge that needs explaining – the model needs to be specific to the research aims and domain (Gilbert & Ahrweiler, 2009, see also Chap. 3). These choices subsequently guide the collection of information on the properties of the constituent parts of the problem. The model construction then ideally follows the classical inductive principles, where the functional structure of the problem, the contributing factors, mechanisms and the conceptual model are inferred. If a fully inductive approach is not feasible, the abductive reasoning to provide the 'best explanation' of the processes of interest can offer a pragmatic alternative.

Subsequently, the model, once built, is internally verified, implemented and executed, and the results are then validated by aligning them with observations. This step can be seen as a continuation of the inductive process of discovery. The nature of the contributing functions, structures and mechanisms is unravelled, by identifying those elements of the modelled processes without which those processes would

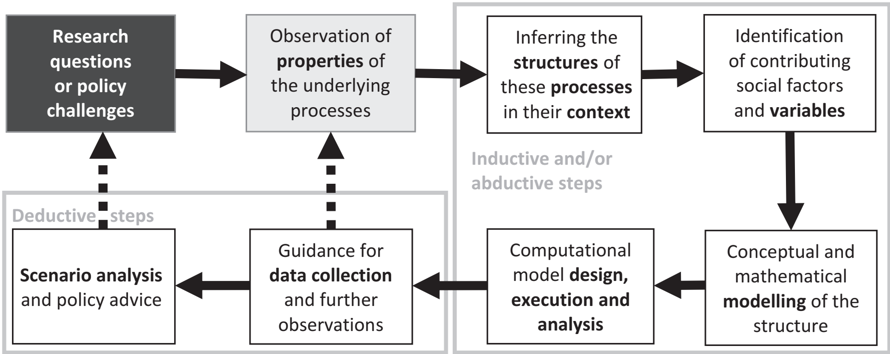


Fig. 2.1 Basic elements of the model-based research programme. (Source: own elaboration based on Courgeau et al., 2016: 43, and Brenner and Werker, 2009)

not occur, or would manifest themselves in a different form. At this stage, the model can also help identify (deduce) the areas for further data collection, which would lead to subsequent model refinements. At the same time, also in a deductive manner, the model generates derived scenarios, which can serve as input to policy advice. These scenarios can give grounds to new or amended research or policy questions, at which point the process can be repeated (Fig. 2.1).

Models obtained by applying the above principles can therefore both enable scenario analysis and help predict structural features and outcomes of various policy scenarios. The model outcomes, in an obvious way, depend on empirical inputs, with Brenner and Werker (2009) having highlighted some important pragmatic trade-offs, for example between validity of results and availability of resources, including research time and empirical data. These pragmatic concerns point to the need for initiating the modelling process by defining the research problem, then building a simple model, as a first-order approximation of the reality to guide intuition and further data collection, followed by creating a full descriptive and empirically grounded version of the model.

At a more general level, modelling can be located on a continuum from general (nomological) approaches (Hempel, 1962), aimed at uncovering idealised laws, theories and regularities, to specific, unique and descriptive (ideographic) ones (Gilbert & Ahrweiler, 2009). The blueprint for modelling proposed in this book aims to help scan at least a segment of this conceptual spectrum for analysing the research problem at hand.

In epistemological terms, the guiding principles of the abductive reasoning can be seen as a pragmatic approximation of a fully inductive process of scientific enquiry, which is difficult whenever our knowledge about the functions, structures and mechanisms is limited, incomplete, poor quality, or even completely missing. In the context of social phenomena, such as migration, these limitations are paramount. This is why the approach adopted throughout the book sees the classical induction as the ideal philosophy to underpin model-based enquiries, and the abductive reasoning as a possible real-life placeholder for some specific aspects. In this way, we aim to offer a pragmatic way of instantiating the model-based research programme in such situations, where applying the fully inductive approach for every element of the modelling endeavour is not feasible. We discuss the elements of the proposed methodology in more detail in Part II.

Part II
Elements of the Modelling Process

Chapter 3
Principles and State of the Art of Agent-Based Migration Modelling

Martin Hinsch and Jakub Bijak

Migration as an individual behaviour as well as a macro-level phenomenon happens as part of hugely complex social systems. Understanding migration and its consequences therefore necessitates adopting a careful analytical approach using appropriate tools, such as agent-based models. Still, any model can only be specific to the question it attempts to answer. This chapter provides a general discussion of the key tenets related to modelling complex systems, followed by a review of the current state of the art in the simulation modelling of migration. The subsequent focus of the discussion on the key principles for modelling migration processes, and the context in which they occur, allows for identifying the main knowledge gaps in the existing approaches and for providing practical advice for modellers. In this chapter, we also introduce a model of migration route formation, which is subsequently used as a running example throughout this book.

3.1 The Role of Models in Studying Complex Systems

Before focusing specifically on modelling human migration, it might be helpful to briefly discuss the role that models can play in analysing complex social phenomena in general. In a wider sense, models can have various purposes (Edmonds et al., 2019; Epstein, 2008); however, here we are specifically interested in the application of models to the study of complex systems. Such systems, that is, systems of many components with non-linear interactions, are notoriously difficult to analyse. Even under best experimental conditions, emergent effects can make it nearly impossible to deduce causal relationships between the behaviour and interactions of the components and the global behaviour of the system (Johnson, 2010). This issue is greatly exacerbated in those systems that are not amenable to experimentation under controlled conditions because they can neither be easily replicated nor manipulated, such as for instance large-scale weather, a species' evolutionary history, or most

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_3

medium- to large-scale social systems. In these cases, modelling can be an extremely useful – and sometimes the only – way to understand the system in question.

3.1.1 What Can a Model Do?

As argued in Chap. 2, whether a model is constructed by following inductive or abductive principles or indeed a mixture of both, and whether it is a computer simulation or a mathematical model, at its heart, it ends up being a deduction engine. It is a tool to – rigorously and automatically – infer the consequences of a set of assumptions, thereby augmenting the limited capacity of human reasoning (Godfrey-Smith, 2009; Johnson, 2010). At the most general level, we can distinguish two epistemologically distinct ways in which such a tool can be used in the context of studying complex systems: proof of causality and extrapolation.

Proof of Causality. Understanding causality in complex systems can be challenging since the links between micro- and macro-behaviour or between assumptions and dynamics tend to be opaque. A model can be used in this situation to infer specific chains of causality. By modelling a set of micro-processes or assumptions we can demonstrate – rigorously, assuming no technical mistakes have been made – which behaviour they produce.

The ability of agent-based models to link the micro- and macro-level processes and phenomena can be used to directly validate or disprove the logical consistency of a pre-existing hypothesis of the form *'(macro-level) phenomenon X is caused by (micro-level) mechanism Y'*. Alternatively, by iterating over several different (micro-level) mechanisms, the (minimum) set of assumptions required to produce a specific behaviour can be discovered (see Grimm et al., 2005; Strevens, 2016; Weisberg, 2007). It is important to note, however, that any such proof of causality can only demonstrate logical consistency of a hypothesis. Empirical research is required to prove the occurrence of the mechanism in question in a given real-world situation.

In a classical example, the famous Schelling (1971) separation model demonstrates that the observed segregation between population groups in many cities can be caused by relatively minor preferences at the individual level. Similarly, the series of 'SugarScape' models by Epstein and Axtell (1996) show that a number of population-level economic phenomena can be the result of basic interactions between very simple agents.

Extrapolation. For many complex systems, we are interested in their behaviour under conditions that are not directly empirically accessible, such as future behaviour or the reaction to specific changes in circumstances. Assuming that we already have a good understanding of a system, we can use a model to replicate the mechanisms responsible for the aspects of the system we are interested in, and use it to extrapolate the system's behaviour.

Different types of complex models of the physics of the Earth's atmosphere, for example, can be used to predict changes in local weather over the range of days on one hand, as well as the development of the global climate in reaction to human influence on the other.

3.1.2 Not 'the Model of', but 'a Model to'

At this point it is important to note that everyday use of language tends to obscure what we really do when building a model. We tend to talk about real world systems in terms of discrete nouns, such as 'the weather', 'this population', or 'international migration'. This has two effects: first, it implies that these are things or objects rather than observable properties of dynamic, complex processes. Second, it suggests that these phenomena are easy to define with clear borders. This leads to a – surprisingly widespread – 'naive theory of modelling' where we have a 'thing' (or an 'object' of modelling) that we can build a canonical, 'best' 'model of', in the same way we can draw an image of an object.

In reality, however, for both types of inference described above, how we build our model is strictly defined by the problem we use it to solve: either by the set of assumptions and behaviours we attempt to link, or by the specific set of observables we want to extrapolate. That means that for a given empirical 'object' (such as 'the weather'), we might build substantially different models depending on what aspect of that 'object' we are actually interested in. In short, which model we build is determined by the question we ask (Edmonds et al., 2019).

As an illustration, let us assume that we want to model a specific stretch of river. Things we might possibly be interested in could be – just to pick a few arbitrary examples – the likelihood of flooding in adjacent areas, sustainable levels of fishing or the decay rate of industrial chemicals. We could attempt to build a generic river model that could be used in all three cases, but that would entail vastly more effort than necessary for each of the single cases. To understand flooding risk, for example, population dynamics of the various animal species in the river are irrelevant. Not only that, building unnecessary complexity into the model is in fact actively harmful as it introduces more sources of error (Romanowska, 2015). It is therefore prudent to keep the model as simple as possible. Thus, even though we will in all three cases build a model 'of the river', the overlap between the models will be limited.

3.1.3 Complications

The main foundational task in modelling therefore consists in defining and delineating the system. First, the system needs to be defined horizontally – that is, which part of the world do we consider peripheral and which parts should be part of the model? Second, it needs also to be specified vertically – which details do we consider important? This can be quite challenging as there is fundamentally no straightforward way to determine which processes are relevant for the model output (Barth et al., 2012; Poile & Safayeni, 2016).

Defining the system can become less of a challenge, as long as we are working in the context of a proof-of-causality modelling effort, since finding which assumptions produce a specific kind of behaviour is precisely the aim of this type of modelling. However, as soon as we intend to use our model to extrapolate system behaviour, trying to include all processes that might affect the dynamics we are interested in, while leaving out those that only unnecessarily complicate the model, becomes a difficult task. As a further complication, we are in practice constrained by various additional factors, such as availability of data, complexity of implementation, and computational and analytical tractability of the simulation (Silverman, 2018). Even with a clear-cut question in mind, designing a suitable model is therefore still as much an art as a science.

3.2 Complex Social Phenomena and Agent-Based Models

Almost all social phenomena – including migration – involve at least two levels of aggregation. At the macroscopic level of the social aggregate – such as a city, social group, region, country or population – we can observe conspicuous patterns or regularities: large numbers of people travel on similar routes, a population separates into distinct political factions, or neighbourhoods in a city are more homogeneous than expected by chance. The mechanisms producing these patterns, however, lie in the interactions between the components of these aggregates – usually individuals, but also groups, institutions, and so on, as well as between the different levels of aggregation.

In order to understand or predict the aggregate patterns we can therefore try to analyse regularities in the behaviour of the aggregate (which can be done with some success, see e.g. Ahmed et al., 2016), or we can try to derive the aggregate behaviour from the behaviour of the components. The latter is the guiding principle behind agent-based modelling/models (ABM): instead of attempting to model the dynamics of a social group as such, the behaviour of the agents making up the group and their interactions are modelled. Group-level phenomena are then expected to emerge naturally from these lower-level mechanisms.

Which modelling paradigm is best suited to a given problem depends to a large degree on the problem itself; however, a few general observations concerning the suitability of ABMs for a given problem can be made. If we want to build an explanatory model, it is immediately clear that agent-based models are a useful – or in many cases the only reasonable – approach. Even for predictive modelling, however, such models have become very popular in the last decades. The advantages and disadvantages of this method have been discussed at length elsewhere (Bryson et al., 2007; Lomnicki, 1999; Peck, 2012; Poile & Safayeni, 2016; Silverman, 2018), but to sum up the most important points: agent-based models are computationally expensive, not easy to implement (well), difficult to parameterise, and are dependent on arbitrary assumptions. On the other hand, they provide unrivalled flexibility in terms of which mechanisms and assumptions to make part of the model, and describe the system on a level that is more accessible to domain experts and non-modellers than aggregate methods. Most importantly, as soon as interactions or differences between people are assumed to be an essential part of a given system's behaviour, it is often much more straightforward to model these directly and explicitly than to attempt to find aggregate solutions.

3.2.1 Modelling Migration

Migration is a prime example of a complex social phenomenon. It is ubiquitous, as well as being one of the crucial processes driving demographic change. Migration can have substantial impacts in all countries involved in the process – origin, transit and destination – in terms of demography, economy, politics and culture. As a political topic, it has also both been important and contentious. Migration complexity and the agency of migrants are some of the important reasons behind the ineffectiveness of migration policies and the reasons why they bring about unintended consequences (Castles, 2004). In recent years, migration has also found increased relevance and focus in the context of the 'digital revolution' (see e.g. Leurs & Smets, 2018; Sánchez-Querubín & Rogers, 2018).

Given the importance and implications of migration processes, there are strong scientific as well as practical incentives for a better understanding of their complexity. However, as argued in Chap. 2, while there is substantial empirical research on migration, existing theoretical studies are sparser and still largely focused on voluntary, economically motivated migration (Arango, 2000; Massey et al., 1993), with forced and asylum migration lagging behind.

3.2.2 Uncertainty

To make things even more difficult, for most of the research questions relevant to the migration processes we are unable to exclude that differences as well as interactions between individuals are an essential part of the dynamics we are interested in. At least as a starting point, this commits us to agent-based modelling as the default architecture.

In the context of migration modelling, the agent-based methodology presents two major challenges. First, as mentioned earlier, many of the processes involved in our target system are not well defined. We therefore have to be careful to take the uncertainty resulting from this lack of definition into account. This is no easy task for a simple model, but even less so for a complicated agent-based model. Second, agent-based models tend to be computationally expensive, which reduces the range of parameter values that can be tested, and thus ultimately the level of detail of any results, including through the lens of sensitivity analysis.

Moreover, in the context of migration modelling, the situation is further complicated by the fact that empirical data on many processes are quite sparse, if they exist, or of poor quality, as further exemplified in Chap. 4. For example, there may be strong anecdotal or journalistic evidence that smugglers play an important role not only in transporting migrants across the Mediterranean, but also in helping them, for instance, along the Balkan route (Kingsley, 2016). Empirically it is, however, extremely difficult to assess the prevalence of smuggling on these routes since all parties involved – smugglers, migrants, as well as law enforcement agencies – have a vested interest in understating these numbers. As another example, it is obvious that borders and border patrols are an extremely important factor in determining how many migrants arrive in which EU country. While numbers on border apprehensions exist (as for example reported by Frontex, 2018), it is unclear how these numbers map to actual border crossings, in particular taking into account repeat attempts.

As a result, we have very little hard knowledge concerning the underlying migration processes. How likely is it for migrants to be caught at the border? How much do migrants usually know about border controls? How do they use that knowledge in deciding where to go? What do migrants do if they fail to cross a border? In the light of these – and many other – grey areas in describing migration processes in detail, any modelling endeavour has to put a strong emphasis on the different guises of the associated uncertainty. In particular, we need to test not only for numeric uncertainty resulting from the intrinsic stochasticity of the modelled processes, but also for uncertainty resulting from our lack of knowledge of the processes themselves (Poile & Safayeni, 2016). While migration uncertainty and unpredictability is well acknowledged (Bijak, 2010; Castles, 2004; Williams & Baláž, 2011), simulation models still need to incorporate it in a more formal and systematic manner.

3.3 Agent-Based Models of Migration: Introducing the Routes and Rumours Model

For a long time, theoretical migration research has been dominated by statistical or equation-based flow models in the economic tradition (Greenwood, 2005). However, the rise of agent-based modelling in the social sciences in the last decades has left its mark on migration research as well. A full review of migration-related ABM studies is outside the scope of this book (but see for example Klabunde & Willekens, 2016 or McAlpine et al., 2021). Instead, we present a number of key aspects of ABMs in general and migration models in particular, and discuss how they have been approached in the existing literature.

Throughout the book we also present a running example taken from our own modelling efforts related to a model of migrant route formation linked to information spread (Routes and Rumours), different elements of which are described in successive boxes throughout this book. We attempt to clarify the points made in the main text by applying them to our example in turn. Insofar as relevant for this chapter, the documentation of the model can be found in Appendix A.

3.3.1 Research Questions

A key dimension along which to distinguish existing modelling efforts is the purpose for which the respective models have been built. The majority of ABMs of migration are built with a concrete real-world scenario in mind, often with a specific focus on one aspect of the situation: Hailegiorgis et al. (2018) for example aimed to predict how climate change might affect emigration from rural communities (among other aspects) in Ethiopia. They used data specific to that situation (including local geography) for their model. Entwisle et al. (2016) studied the effect of different climate change scenarios on migration in north Thailand using a very detailed model that includes data on local weather patterns and agriculture. Frydenlund et al. (2018) attempted to predict where people displaced by conflict in the Democratic Republic of Congo will migrate to. Their model, among other features, includes local geographical and elevation data.

Many of these very concrete models, however, while being calibrated to a specific situation are meant to provide more general insights. Suleimenova and Groen (2020), for example, modelled the effect of policy decisions on the number of arrivals in refugee camps in South Sudan. Their study was intended to provide direct support to humanitarian efforts in the area. At the same time, it serves as a showcase for a new modelling approach that the authors have developed.

A minority of studies eschew data and specific scenarios, and instead focus on more general theoretical questions. Collins and Frydenlund (2016), for example, investigated the effect of group formation on the travel speed of refugees using a purely theoretical model without any relation to specific real-world situations. In a similar vein, Reichlová (2005) explored the consequences of including safety and social needs in a migration model. Although her study was explicitly motivated by real-world phenomena, the model itself and the question behind it are purely theoretical.

Finally, some models are built without a specific domain question in mind. In these cases, the authors often explore methodological issues or put their model forth as a framework to be used by more applied studies down the line (e.g. Groen, 2016; Lin et al., 2016; Suleimenova et al., 2017). Others simply explore the dynamics arising from a set of assumptions without further reference to real-world phenomena (e.g. Silveira et al., 2006, or Hafızoğlu & Sen, 2012).

The research question underpinning the Routes and Rumours model is defined in Box 3.1.

3.3.2 Space and Topology

Migration is an inherently spatial process. Spatial distance between countries of origin and destination has long been part of macroscopic, so-called gravity models of migration (Greenwood, 2005). Agent-based models, however, make it possible to model spatial aspects of migration much more explicitly.

Box 3.1: Routes and Rumours: Defining the Question
The starting point for the Routes and Rumours model that serves as our running example was the observation, first, that very little theoretical work has been done on the migration journey itself and second, that on that journey what little information migrants have on the local conditions often is based on hearsay from other migrants (Dekker et al., 2018; Wall et al., 2017). From there, we decided to investigate the effect of the availability and transmission of information on the emergence of migration routes. In the first instance, we did not attempt to describe a *specific real-world situation*, however, but wanted to use our model to better understand the *general mechanisms* behind the interaction between information and route formation.

Our model was therefore at this point purely theoretical. Our working hypothesis was that routes – which clearly emerge in the real world – are a result more of self-organisation than optimisation and would therefore be difficult to predict, if prediction was at all possible.

How relevant space is in a given model is determined by the phenomena that a modeller is interested in. In a situation where the net flow of migration between a small number of countries or locations is being investigated, for example, spatial relationships beyond mutual distances is often not taken into account (e.g. Heiland, 2003; Lin et al., 2016, but see e.g. Ahmed et al., 2016 for a non-agent-based model that includes geographic information). There are also some models that include a spatial component but use the relative spatial position of agents solely as a simple representation of social distance (e.g. Klabunde, 2011; Reichlová, 2005).

If actual spatial detail is required, spatial information is usually represented either by a square grid or a graph. While a grid-based approach has the advantage of being straightforward to implement and understand, it does tend to be computationally heavier. Which structure works best, however, ultimately often depends on the requirements of the model and the availability of data.

Fully theoretical models tend to use simple grid-based spatial structure (Silveira et al., 2006; Collins & Frydenlund, 2016; but see Naqvi & Rehm, 2014). Similarly, spatial models built to simulate a specific scenario but without using real-world geographical data (e.g. Sokolowski et al., 2014; Werth & Moss, 2007) will often resort to this solution for convenience. While Hailegiorgis et al. (2018) used detailed rasterised data for their model, most models employing real-world data seem to be built on much simpler graph structures representing networks of, for example, cities (Groen, 2016), districts (Hassani-Mahmooei & Parris, 2012), or even entire countries (Lin et al., 2016).

Finally, in some cases, a completely different approach is used. Naivinit et al. (2010) used a grid structure but with hexagonal instead of square cells. Similarly, although the description of their model is not very detailed, it appears that Frydenlund et al. (2018) did not implement a discretised spatial representation at all, but directly used polygonal data extracted from a geographical information system (GIS). For the Routes and Rumours model, the spatial structure of the simulated world is summarised in Box 3.2.

Box 3.2: Space in the Routes and Rumours Model
Since we intended to study the emergence of migration routes, we had to take spatial structures into account. An initial version of the model showed, however, that a naive grid-based approach was too computationally costly. We settled therefore on representing cities and transport links as vertices and edges of a graph, respectively. Such a representation is sparser than a full grid, but nevertheless reflects the main topological features of the modelled landscape, which are the spatial connections between different settlements through transport links. An example topology is shown in Fig. 3.1 below.

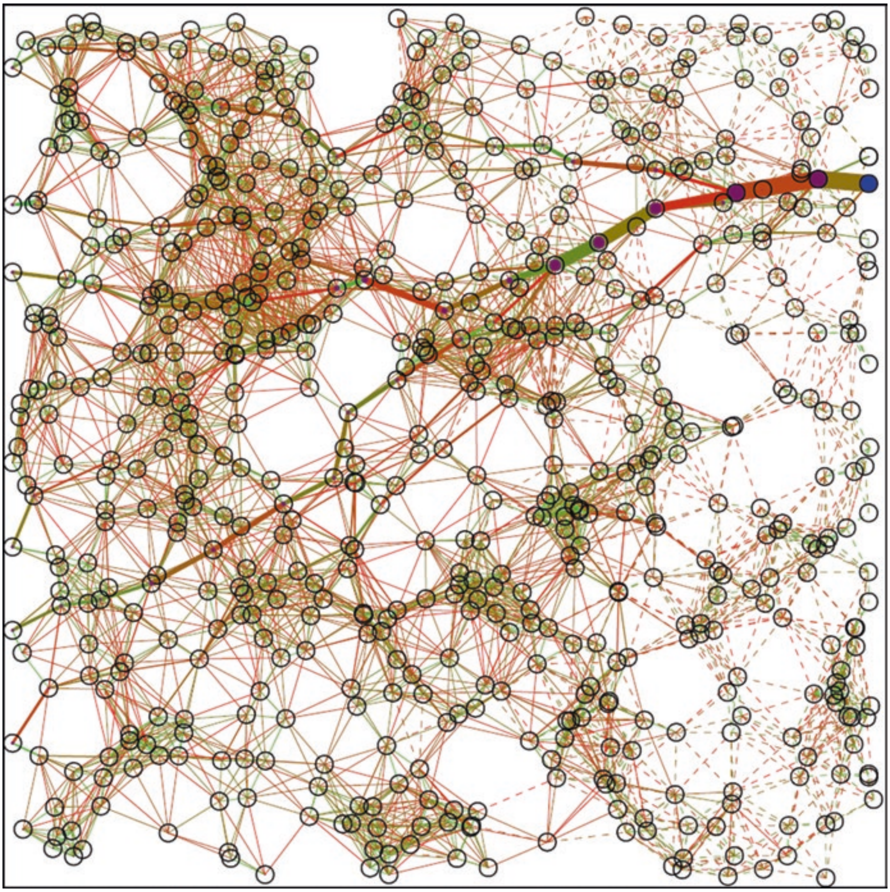


Fig. 3.1 An example topology of the world in the Routes and Rumours model: Settlements are depicted with circles, and links with lines, their thickness corresponding to traffic intensity

3.3.3 Decision-Making Mechanisms

Decision making is an essential part of most models of human migration, or indeed of most other forms of human behaviour (Klabunde & Willekens, 2016). However, which of the many different types of decisions involved a given model makes explicit varies, and is primarily a function of the question the model is used to answer.

Traditionally, modelling studies on migration were primarily invested in understanding under which conditions people decide to migrate and where they will go (Massey et al., 1993). Consequently, the two types of decisions most often included in migration models – agent-based or not – are first, whether to leave and migrate in the first place, and second, which destination to choose when migrating.

In a common type of model, the main focus lies on the conditions in the area or country of origin. In this case, migration is just one of several ways in which individuals can react to changes in local conditions, and the fate of migrants is usually not tracked beyond the decision to leave unless return migration is included (e.g. Entwisle et al., 2016). Examples of such models include Naivinit et al. (2010), Smajgl and Bohensky (2013) and Hailegiorgis et al. (2018).

Unless they are focused on a pair of countries or locations (such as the USA and Mexico, e.g. Klabunde, 2011 and Simon et al., 2016; or East and West Germany, Heiland, 2003), models that simulate the entire migration process usually include the decision to leave as well as a decision where to go. For models of internal migration this is often implemented as a detailed, spatially explicit choice of location (e.g. Frydenlund et al., 2018; Hébert et al., 2018; or Groen et al., 2020). In models of international migration, the decision is usually presented as a choice between different possible countries of destination (e.g. Reichlová, 2005 or Lin et al., 2016).

In addition, a few studies extend the scope of the analysis beyond the simple decisions to leave and where to go. As mentioned before, some models let migrants decide whether to return to their country of origin (e.g. Klabunde, 2014; Simon, 2019). Others include the option to attempt to reach the destination using illegal means (Simon et al., 2016). Finally, there are a few rare modelling studies that focus on entirely different aspects of migration, and consequently model different decisions, such as whether to join a group while travelling (Collins & Frydenlund, 2016). The way decisions are implemented also varies a lot between different studies. In some cases, the decision model is based on an established paradigm such as utility maximisation (e.g. Heiland, 2003; Klabunde, 2011; Silveira et al., 2006). In others, the model is specifically intended as a test case to study the effects of decision making, such as the inclusion of social norms in an economic model (Werth & Moss, 2007), using the theory of motivation (Reichlová, 2005) or the Theory of Planned Behaviour (Klabunde et al., 2015; Smith et al., 2010). Often, however, there does not seem to be a clear justification for the behaviour rules built into the model.

Even in models specifically aimed at prediction within a given real-world scenario, empirical validation of decision rules does not seem to be very common. If it happens, it is usually limited to calibrating the model with regression data linking migration decisions to individuals' circumstances (e.g. Entwisle et al., 2016; Klabunde, 2014; Smith, 2014). Direct validation of decision processes using, for example, survey-based information (Simon et al., 2016), is rare. For further reading on decision making in migration models we recommend the review by Klabunde and Willekens (2016).

In our case, the way the decisions about the subsequent stages of the journey are being made in the Routes and Rumours model is summarised in Box 3.3.

Box 3.3: Decisions in the Routes and Rumours Model
Since we were primarily interested in the journey itself, we assumed in our running example that individuals have already made the decision to leave their home country, but are not yet at a point where the decision as to which destination country to travel to matters. Instead, we focused on the decisions that determine the route a migrant travels, that is which city to head for next and how to get there.

In principle, agents attempt to reach their destination as quickly as possible. However, in our model the shortest path is not necessarily optimal. The quality of a route is affected by *friction*, an aggregate measure of distance and ease of travel but also the risk a specific leg of the journey entails, as well as the general quality (a stand in for e.g. availability of resources and shelter or permissiveness of local law enforcement) of waypoints. For most components of that decision, we did not have any data to draw on, so we resorted to a simple *ad hoc* model of decision making. For the effect of risk, however, we were able to incorporate data from a psychological survey (see Chap. 6).

3.3.4 Social Interactions and Information Exchange

By definition, macroscopic models have difficulty in capturing the interactions between individuals. This turns out to be a methodological issue once it becomes clear that network effects play an important role in determining the dynamics of international migration (Gurak & Caces, 1992; Massey et al., 1993). To a certain degree, and in some cases, these network effects and other interactions between individuals can be approximated at a macroscopic level (e.g. Ahmed et al., 2016; Massey et al., 1993). However, modelling interactions between individuals is substantially more straightforward in agent-based models, even though there are examples of such models of migration that either do not include any interactions between individuals at all, or only indirect interactions via some global state (e.g. Hébert et al., 2018; Heiland, 2003; Lin et al., 2016).

The simplest forms of interaction take place in movement models where proximity (Frydenlund et al., 2018) or group membership (Collins & Frydenlund, 2016) affect an agent's trajectory. If more complicated interactions are taken into account, then most often this takes the form of social networks that affect an individual's willingness and/or ability to migrate. In the simplest form, this is done by using space as a proxy for social distance (see Sect. 3.3.2) and defining an individual's 'social network' as all individuals within a specific distance in that space (e.g. Reichlová, 2005; Silveira et al., 2006). More elaborate models explicitly set up links between individuals and/or households (Simon, 2019; Smith et al., 2010; Werth & Moss, 2007), which in some cases are assumed to change over time (e.g. Klabunde, 2011; Barbosa et al., 2013).

The effects that networks are assumed to have on individuals vary and in many cases more than one effect is built into models. Most commonly, networks directly affect individuals' migration decisions either by providing social utility (e.g. Reichlová, 2005; Silveira et al., 2006; Simon, 2019) or social norms (Smith et al., 2010; Barbosa et al., 2013). Another common function is the transmission of information on the risk or benefits of migration (Barbosa et al., 2013; Klabunde, 2011; Simon et al., 2018). Direct economic benefits of networks are only taken into account in a few cases (Klabunde, 2011; Simon, 2019; Werth & Moss, 2007).

Apart from social networks, a few other types of interaction occur in agent-based models of migration. In some studies, agents make their migration decisions without any direct influence from others but interact with them in other ways, such as economically (Naivinit et al., 2010; Naqvi & Rehm, 2014) or by learning (Hailegiorgis et al., 2018), which affects their economic status and thus the likelihood of migrating.

Information and exchange of that information between migrants are the main processes we assumed to be relevant for the emergence of migration routes, and consequently had to be a core part of our model. The information dynamics within the model, as well as the mechanism for the update of agents' beliefs, are summarised in Box 3.4.

3.4 A Note on Model Implementation

A significant hurdle to the broader adoption of agent-based modelling – in particular, in the social sciences – is the specialist skill required to build these kinds of models. There are ways to lower that hurdle, such as specialised software packages (Railsback et al., 2006) or domain-specific languages (discussed in Chap. 7), however all of these come at the cost of reduced flexibility and at times very low efficiency (Reinhardt et al., 2019).

In order to leverage the full potential of agent-based modelling it is therefore often still helpful to implement these models from scratch in a general-purpose language. There is a vast array of languages and methods from which to choose. Traditionally, these fall on a spectrum marked by a trade-off between speed and convenience. At one end, we have fast, yet difficult and unwieldy 'systems-programming' style languages such as C, C++, Fortran or Rust, and at the other much simpler and more convenient, but slow languages such as Python or R. Unfortunately, the fast end of this spectrum tends to be only accessible to experienced programmers, and even then involves trading off convenience and productivity for speed.

Box 3.4: Information Dynamics and Beliefs Update in the Routes and Rumours Model

Agents in our model start out knowing very little about the area they are travelling through, but accumulate knowledge either by exploring locally or by exchanging information with agents they meet or are in contact with. This information is not only necessarily incomplete most of the time, but may also not be accurate. Through exchange it is even possible that incorrect information spreads in the population.

For each property of the environment – say, risk associated with a transport link – an agent has an estimate as well as a confidence value. Collecting information improves the estimate and increases the confidence. During information exchange with other agents, however, confidence can even decrease if both agents have very different opinions.

Our model of information exchange therefore had to fulfil a number of conditions: (a) knowledge can be wrong and/or incomplete, (b) knowledge can be exchanged between individuals, yet, crucially the exchange does not depend on objective, but only on subjective reliability of the information, and (c) agents therefore need an estimate of how certain they are that their information is correct.

Since existing models of belief dynamics do not fulfil all of these criteria, we designed a new (sub-) model of information exchange.

Formally, we used a mass action approach to model the interaction between the certainty $t \in (0, 1)$ and doubt $d = 1 - t$ components of two agents' beliefs. During interactions we assumed that these components interact independently in a way that agents can be convinced (doubt transforming to certainty through the interaction with certainty), converted (certainty of one belief is changed to certainty of a different belief through the interaction with certainty) or confused (certainty is changed to doubt by interacting with certainty if the beliefs differ sufficiently).

For two agents A and B we calculated difference in belief as

$$\delta_v = \frac{|v_A - v_B|}{v_A + v_B}.$$

The new value for doubt is then:

$$d_A' = d_A d_B + (1 - c_i) d_A t_B + c_u \delta_v t_A t_B,$$

and the new value estimate:

$$v_A' = \frac{t_A d_B v_A + c_i d_A t_B v_B + t_A t_B (1 - c_u \delta_v)\left((1 - c_e) v_A + c_e v_B\right)}{(1 - d_A')},$$

where c_i, c_e and c_u are parameters determining the amount of convincing, conversion and confusion.

Julia, a new language developed by a group from MIT (Bezanson et al., 2014), has recently started to challenge this trade-off. It has been designed with a focus on technical computing and the express goal of combining the accessibility of a dynamically typed scripting language like Python or R with the efficiency of a statically typed language like C++ or Rust. A combination of different techniques is used to achieve this goal. In order to keep the language easily accessible, it employs a straightforward syntax (borrowing heavily from MatLab) and dynamic typing with optional type annotations. Runtime efficiency is accomplished by combining strong type inference with just-in-time compilation based on the LLVM platform (Lattner & Adve, 2004). Following a few relatively straightforward guidelines, it is therefore possible to write code in Julia that is nearly as fast as C, C++ or Fortran while being substantially simpler and more readable.

Beyond simplicity and efficiency, however, Julia offers additional benefits. Similar to languages such as R or Python, it comes with interactive execution environments, such as a REPL (read-eval-print loop) and a notebook interface that can greatly speed up prototyping. It also has a powerful macro system built in that has, for example, been used to enable near-mathematical notation for differential equations and computer algebra. Some specific notes related to the Julia implementation are summarised in Box 3.5.

Box 3.5: Specific Notes on Implementation of the Routes and Rumours Model in Julia

We implemented the Routes and Rumours model in Julia from the outset. Beyond the noted combination of simplicity and efficiency, there were a few additional areas where development of the model benefitted substantially from the choice of language:

- Defining and inputting model parameters tends to be cumbersome and error-prone in static languages. Usually the addition of a parameter requires several changes at different places in the code. Using Julia's meta-programming facilities, it was straightforward to have all uses of a model parameter (definition, description, default values, input and output) generated from a single point of definition.
- Similarly, collection and output of data from the model often leads to either inefficient or scattered and fragile code. Using macros, we implemented a simple declarative interface that allows for the definition of data output in one place and mostly separate from the model code.
- As a minor benefit, we were able to use the same language to interactively analyse and graph the data generated by the simulations as for the simulation itself.
- As discussed in Chap. 7, we used Julia's macro system to implement an abstraction of event-based scheduling that is nearly as convenient as a dedicated external domain-specific language.
- Adding dynamically loadable, yet efficient, scenario modules to the model turned out to be close to trivial (see Chap. 8).

3.5 Knowledge Gaps in Existing Migration Models

As we can see, ABMs have become firmly established as a method available for migration modelling. Their application ranges from purely theoretical models to efforts to predict aspects of migration calibrated to a specific real-world situation. A variety of different topics have been tackled such as the effects of climate change on migration via agriculture, the spread of migration experiences through social networks, the formation of groups by travelling migrants, or how the local threat of violence affects numbers of arrivals in refugee camps. Methodologically, these models vary considerably as well, including for example GIS-based spatial representation, decision models based on the theory of planned behaviour, or a spatially explicit ecological model that predicts agricultural yields.

On the other hand, some notable counter-examples notwithstanding, many models in this field still tend to be simple, not at all or poorly calibrated, narrow in focus and littered with *ad hoc* assumptions. In many cases, this is despite best efforts on the part of the authors. Not only is agent-based modelling in general a very 'data hungry' method, but in addition – as further discussed in Chap. 4 and in Sect. 3.2 in this chapter – migration is a phenomenon that is inherently difficult to access empirically.

While macroscopic data on e.g. number of arrivals, countries of origin or demographic composition are sometimes reasonably accessible, microscopic data, in particular on individual decision making, can be nearly impossible to obtain (Klabunde & Willekens, 2016). Consequently, decision making – arguably the most important part of a model concerned with an aspect of human behaviour – is in most models at best calibrated with regression data (but see Simon et al., 2016 for a notable exception) and often neither calibrated, nor in other ways justified (e.g. Hébert et al., 2018).

Unfortunately, even calibration or validation against easier to obtain macroscopic data is not a given. Even some predictive studies restrict themselves to the most basic forms of validation, for example by simply showing model outcomes next to real data (e.g. Groen et al., 2020; Lin et al., 2016; Suleimenova & Groen, 2020). For a purely theoretical model, a lack of empirical reference is not necessarily a cause for concern. But if it is the express goal of a study to be applicable to a concrete real-world situation, then a certain effort towards understanding the amount as well as the causes of uncertainty in the model results should be expected. As some authors, who go to great lengths to include the available data and to calibrate the model against it, demonstrate, high-quality modelling efforts do exist (e.g. Naivinit et al., 2010; Simon et al., 2018; Hailegiorgis et al., 2018).

Another point to note is the relative paucity of theoretical studies attempting to find general mechanisms – as opposed to generating predictions of a specific situation – in the tradition of Schelling (1971) or Epstein and Axtell (1996). Of the existing examples, some stand in the tradition of abstract modelling approaches employed in physics, so that it is difficult to assess the generality of their results (Hafızoğlu & Sen, 2012; Silveira et al., 2006). All these issues additionally reinforce the need for the model-based research programme, advocated in Chap. 2, going beyond the state of the art in agent-based modelling, and including other approaches and sources of empirical information. As argued before, such efforts should be ideally guided by the principles of classical inductive reasoning.

Generally, however, we can see that formal modelling can open up new areas for migration studies. Many questions remain untouched, providing promising areas for future research. On the whole, as argued above, the primary focus of any modelling exercise should not be aimed at a precise description, explanation or prediction of migration processes, which is an impossible task, but at identifying gaps in data and knowledge. Furthermore, for any given migration system, there is no canonical model. As argued before, the models need to be built for specific purposes, and with particular research questions in mind. Of course, many such questions still have direct practical, policy or scientific relevance. Examples of such questions may include:

- What is the uncertainty of migration across a range of time horizons? What can be a reasonable horizon for attempts at predicting migration, under a reasonable description of uncertainty?
- How are the observed flows of migration likely to be formed, who might be migrating, and who would stay behind? What is the role of historical trends, migrant networks, or other drivers?
- What drives the emergence of migration routes, policies and political impacts of migration? Are migration policies only exogenous variables, or are they endogenous, driven by migration flows?
- More generally, does migration lead to feedback effects, for example through the impacts on societies, policies or markets, and how is it mediated by the level of integration of migrants?
- What are the root causes of migration, and how does migration interact with other aspects of social life? To what extent are various actors (migrants, institutions, intermediaries…) involved?
- How are migration decisions formed and put into action? Do cognitive components dominate, or are emotions highly involved as well? Does it vary between different migration types?

The specific questions, which can be driven by policy or scientific needs, will determine the model architecture and data requirements. Next, we discuss a way of assessing the data requirements of the model through formal analysis.

Chapter 4
Building a Knowledge Base for the Model

Sarah Nurse and Jakub Bijak

In this chapter, after summarising the key conceptual challenges related to the measurement of asylum migration, we briefly outline the history of recent migration flows from Syria to Europe. This case study is intended to guide the development of a model of migration route formation, used throughout this book as an illustration of the proposed model-based research process. Subsequently, for the case study, we offer an overview of the available data types, making a distinction between the sources related to the migration processes, as well as to the context within which migration occurs. We then propose a framework for assessing different aspects of data, based on a review of similar approaches suggested in the literature, and this framework is subsequently applied to a selection of available data sources. The chapter concludes with specific recommendations for using the different forms of data in formal modelling, including in the uncertainty assessment.

4.1 Key Conceptual Challenges of Measuring Asylum Migration and Its Drivers

Motivated by the high uncertainty and complexity of asylum-related migration, discussed in Chap. 2, we aim to illustrate the features of the model-based research process advocated in this book with a model of migration route formation. We have focused on the events that took place in Europe in 2015–16 during the so-called 'asylum crisis', linked mainly to the outcomes of the war in Syria. To remain true to the empirical roots of demography as a social science discipline, a computational model of asylum migration needs to be grounded in the observed social reality (Courgeau et al., 2016).

Given the nature of the challenge, the data requirements for complex migration models are necessarily multi-dimensional, and are not limited to migration processes themselves, additionally including a range of the underpinning features and

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_4

drivers. At the same time, problems with data on asylum migration are manifold and well documented (see Chap. 2). The aim of the work presented in this chapter is to collate as much information as possible on the chosen case study for use in the modelling exercise, and to assess its quality and reliability in a formal way, allowing for an explicit description of data uncertainty. In this way it can be still possible to use all available relevant information while taking into account the relative quality when deciding on the level of importance with which the data should be treated, and the uncertainty that needs to be reflected in the model.

In this context, it was particularly important to choose a migration case study with a large enough number of migrants, and with a broad range of available information and sources of data on different aspects of the flows. This is especially pertinent in order to allow investigation of the different theoretical and methodological dimensions of the migration processes by formally modelling their properties and the underlying migrant behaviour. Consequently, knowledge about the different aspects of data collection and quality of information, and a methodology for reflecting this knowledge in the model, become very important elements of the modelling endeavour in their own right.

In this chapter, we present an assessment of data related to the recent asylum migration from Syria to Europe in 2011–19. As mentioned above, we chose the case study not only due to its humanitarian and policy importance, and the high impact this migration had both on Syria and on the European societies, but also taking into account data availability. This chapter is accompanied by Appendix B, which lists the key sources of data on Syrian migration and its drivers. The listing includes details on the data types, content and availability, as well as a multidimensional assessment of their usefulness for migration models, following the framework introduced in this chapter.

Even though one of the central themes of the computational modelling endeavours is to reflect the complexity of migration, the theoretical context of our understanding of population flows has traditionally been relatively basic. As mentioned in Chap. 2, within a vast majority of the existing frameworks, decisions are based on structural differentials, such as employment rates, resulting in observed overall migration flows (for reviews, see e.g. Massey et al., 1993; Bijak, 2010). In his classical work, Lee (1966) aimed to explain the migration process as a weighing up of factors or 'drivers' which influence decisions to migrate, while Zelinsky (1971) described different features of a 'mobility transition', which could be directly observed. Most of the traditional theories do not reflect the complexity of migration (Arango, 2000), and typically fail to link the macro- and micro-level features of the migration processes, which is a key gap that needs addressing through modelling.

More recently, there have been attempts to move the conceptual discussion forward and to bridge some of these gaps. A contemporary 'push-pull plus' model (Van Hear et al., 2018) adds complexity to the original theory of Lee (1966), but fails to provide a framework that can be operationalised in an applied empirical context. The 'capability' framework of Carling and Schewel (2018) stresses the importance of individual aspirations and ability to migrate, but again fails to map the concepts clearly onto the empirical reality. In general, the disconnection between the theoretical discussions and their operationalisation – largely limited to survey-based questions on migration intentions – is a standard fixture of much of the conceptual work on migration.

In the context of displacement or forced migration, including asylum-related flows, the conceptual challenges only get amplified. As noted by Suriyakumaran and Tamura (2016), and Bijak et al. (2017), operationalisation of the conceptually complex theories of asylum migration is typically reduced to identifying a selection of available drivers to include in explanatory models. The presence of underlying structural factors or 'pre-conditions' for migration is itself not a sufficient driver of migration; very often, migration occurs following accumulation of adverse circumstances, and some trigger events, either experienced or learnt about through social networks or media. For that reason, the monitoring of the underlying drivers, such as the conflict intensity, becomes of paramount importance (Bohra-Mishra & Massey, 2011). On the other hand, the measurement of drivers comes with its own set of challenges and limitations, which also need to be formally acknowledged.

Another crucial concept to consider when modelling migration processes is how different elements of the conceptual framework interact, and what that implies for measurement. An example could be the measurement of the difficulty of different routes for migrants undertaking a journey. In this case, it is important whether a prospective route includes crossing national borders, whether those borders are patrolled, whether there is a smuggling network already operating, and whether individuals have access to the information and resources necessary to navigate all the barriers that can exist for migrants. As an overall summary measure or perception for decision making, this can be thought of as a route's friction (see Box 3.3; for a general discussion related to migration, see Stillwell et al., 2016). Friction can include either formal barriers, such as national borders and visa restrictions, or informal barriers, such as geographic distance or physical terrain. These challenges require adopting a flexible and imaginative approach to using data, for example by building synthetic indicators based on several sources, or using model-based reconciliation of data (Willekens, 1994).

4.2 Case Study: Syrian Asylum Migration to Europe 2011–19

In this section, we look at recent Syrian migration to Europe (2011–19) through the lens of the available data sources, and propose a unified framework to assess the different aspects in which the data may be useful for modelling. From a historical perspective, recent large-scale Syrian migration has a distinct start, following the widespread protests in 2011 and the outbreak of the civil war. After more than a year of unrest, in June 2012 the UN declared the Syrian Arab Republic to be in a state of civil war, which continues at the time of writing, more than nine years later. Whereas previous levels of Syrian emigration remained relatively low, the nature of the conflict, involving multiple armed groups, government forces and external nations, has resulted in an estimated 6.7 million people fleeing Syria since 2011 and a further 6.1 million internally displaced by the end of 2019, according to the UNHCR (2021, see also Fig. 4.1). The humanitarian crisis caused by the Syrian conflict, which had its dramatic peak in 2015–16, has continued throughout the whole decade.

Initial scoping of the modelling work suggests the availability of a wide range of different types of data that have been collected on the recent Syrian migration into Europe. In particular, the key UNHCR datasets show the number of Syrians who were displaced each year, as measured by the number of registered asylum seekers, refugees and other 'persons of concern', and the main destinations of asylum seekers and refugees who have either registered with the UNHCR or applied for asylum. The information is broken down by basic characteristics, including age and sex and location of registration, distinguishing people located within refugee camps and outside.

As shown in Fig. 4.1, neighbouring countries in the region (chiefly Turkey, Lebanon and Jordan, as well as Iraq and Egypt) feature heavily as countries of asylum, together with a number of European destinations, in particular, Germany and Sweden. The scale of the flows, as well as the level of international interest and media coverage, means that the development of migrant routes and strategies have often been observed and recorded as they occur. In many cases, the situation of the Syrian asylum seekers and refugees is also very precarious. By the UNHCR's account, by the end of 2017, nearly 460,000 people still lived in camps, mostly in the region, in need of more 'durable solutions', such as safe repatriation or resettlement. (This number has started to decline, and nearly halved by mid-2019). A further five million were dispersed across the communities in the 'urban, peri-urban and rural areas' of the host countries (UNHCR, 2021). The demographic structure of the Syrian refugee population generates challenges in the destination countries with respect to education provision and labour market participation, with about 53% people of working age (18–59 years), 2% seniors over 60 years, and 45% children and young adults under 18 (UNHCR, 2021).

When it comes to asylum migration journeys to Europe, visible routes and corridors of Syrian migration emerged, in recent years concentrating on the Eastern Mediterranean sea crossing between Turkey and Greece, as well as the secondary land crossings in the Western Balkans, and the Central Mediterranean sea route between Libya and Italy (Frontex, 2018). By the end of 2017, Syrian asylum migrants were still the most numerous group – over 20,000 people – among those apprehended on the external borders of the EU (of whom nearly 14,000 were on the Eastern Mediterranean sea crossing route). However, these numbers were considerably down from the 2015 peak of nearly 600 thousand apprehensions in total, and nearly 500,000 in the Eastern Mediterranean (*idem*, pp. 44–46). These numbers can be supplemented by other sad statistics: the estimated numbers of fatalities, especially referring to people who have drowned while attempting to cross the Mediterranean. The IOM minimum estimates cite over 19,800 drownings in the

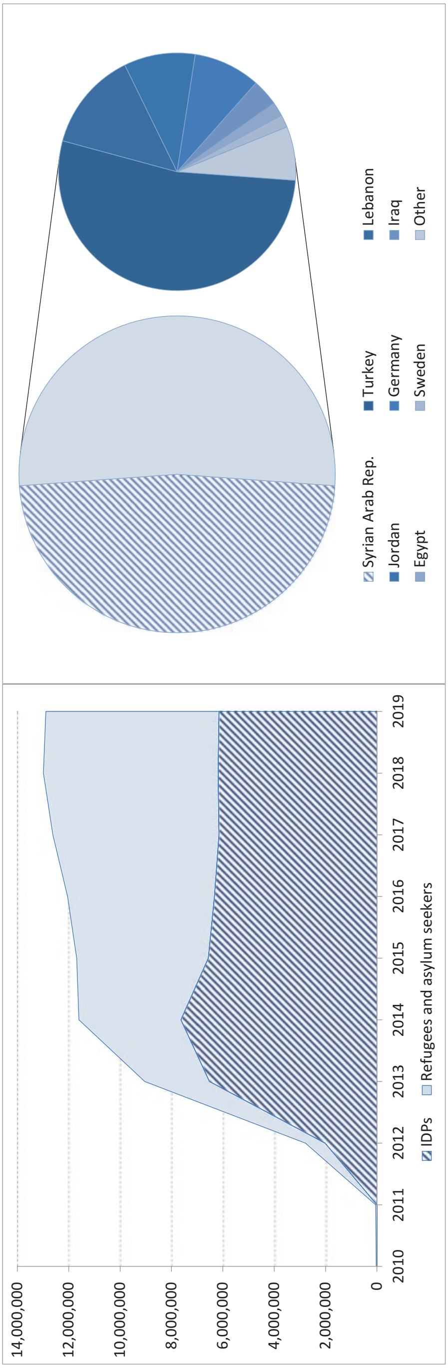


Fig. 4.1 Number of Syrian asylum seekers, refugees, and internally displaced persons (IDPs), 2011–19, and the distribution by country in 2019. (Source: UNHCR, 2021)

period 2014–19, of which 16,300 were in the Central Mediterranean. In about 850 cases, the victims were people who came from the Middle East, a majority presumed to be Syrian (IOM, 2021). In the same period, the relative risk of drowning increased to the current rate of around 1.6%, substantially higher (2.4%) for the Central Mediterranean route (*idem*).

As concerns the destinations themselves, the asylum policies and recognition rates (the proportion of asylum applicants who receive positive decisions granting them refugee status or other form of humanitarian protection) clearly differ across the destination countries, and also play a role in shaping the asylum data. Still, in the case of Syrian asylum seekers, these differences across the European Union are not large. According to the Eurostat data,[1] between 2011 and 2019, over 95% decisions to the applications of Syrian nationals were positive, and these rates were more or less stable across the EU, with the exception of Hungary (with only 36% positive decisions, and a relatively very low number of decisions made). It is worth noting here that administrative data on registrations and decisions have obvious limitations related to the timeliness of registration of new arrivals and processing of the applications, sometimes leading to backlogs, which may take months or even years to clear. Moreover, the EU statistics refer to asylum applications *lodged*, which refers to the final step in the multi-stage asylum application process, consisting of a formal acknowledgement by the relevant authorities that the application is under consideration (European Commission, 2016).

At the same time, besides the official statistics from the registration of Syrian refugees and asylum seekers by national and international authorities, specific operational needs and research objectives have led to the emergence of many other data sources. In this way, in addition to the key official statistics, such as those of the UNHCR, there exist many disparate information sets, which deal with some very specific aspects of Syrian migration flows and their drivers. These sources extend beyond the fact of registration, providing much deeper insights into some aspects of migration processes and their context. Still, the trade-offs of using such sources typically include their narrower coverage and lack of representativeness of the whole refugee and asylum seeker populations. Hence, there is a need for a unified methodology for assessing the different quality aspects of different data sources, which we propose and illustrate in the remainder of this chapter. In addition, we present a more complete survey of these sources in more detail in Appendix B, current as of May 2021, together with an assessment of their suitability for modelling.

[1] All statistics quoted in this paragraph come from the 'Asylum and managed migration' (migr) domain, table 'First instance decisions on applications by citizenship, age and sex' (migr_asydcfsta), extracted on 1 February 2021.

4.3 Data Overview: Process and Context

4.3.1 Key Dimensions of Migration Data

In the proposed approach to data collection and use in modelling, we suggest following a two-stage process of data assessment for modelling. The first stage is to identify all available data relevant to the different elements involved in the decision making and migration flows being modelled. The second stage is then to introduce an assessment of uncertainty so that it can be formally taken into account and incorporated into the model.

Depending on the purpose and the intended use in different parts of the model, the data sources can be classified by type; broadly, these can be viewed as providing either *process-related* or *contextual* information. The distinction here is made between data relating specifically to the migration processes, including the characteristics of migrants themselves, their journey and decisions on the one hand, and contextual information, which covers the wider situation at the origin, destination and transit countries, on the other. Relevant data on context can include, for example, macro-economic conditions, the policy environment, and the conflict situation in the country of origin or destination.

In addition, in order to allow the data to be easily accessed and appropriately utilised in the model, the sources can be further classified depending on the level of aggregation (macro or micro), as well as paradigm under which they were collected (quantitative or qualitative). These categories, alongside a description of source type (for example, registers, surveys, censuses, administrative or operational data, journalistic accounts, or legal texts) are the key components of meta-information related to individual data sources, and are useful for comparing similar sources during the quality assessment.

The conceptual mapping of the different stages of the migration process and their respective contexts onto a selection of key data sources is presented in Fig. 4.2, with context influencing the different stages of the process, and the process itself being simplified into the origin, journey and destination stages. For each of these stages, several types of sources of information may be typically available, although certain types (surveys, interviews, 'new data' such as information on mobile phone locations or communication exchange, social media networks, or similar) are likely to be more associated with some aspects than with others. From this perspective, it is also worth noting that while the process-related information can be available both at the macro level (populations, flows, events), or at the micro level (individual migrants), the contextual data typically refer to the macro scale.

Hence, to follow the template for the model-building process sketched in Chap. 2, the first step in assessing the availability of data for any migration-related modelling endeavour is to identify the critical aspects of the model, without which the processes could not be properly described, and which can be usefully covered by the existing data sources, with a varying degree of accuracy. Next, we present examples of such process- and context-related aspects.

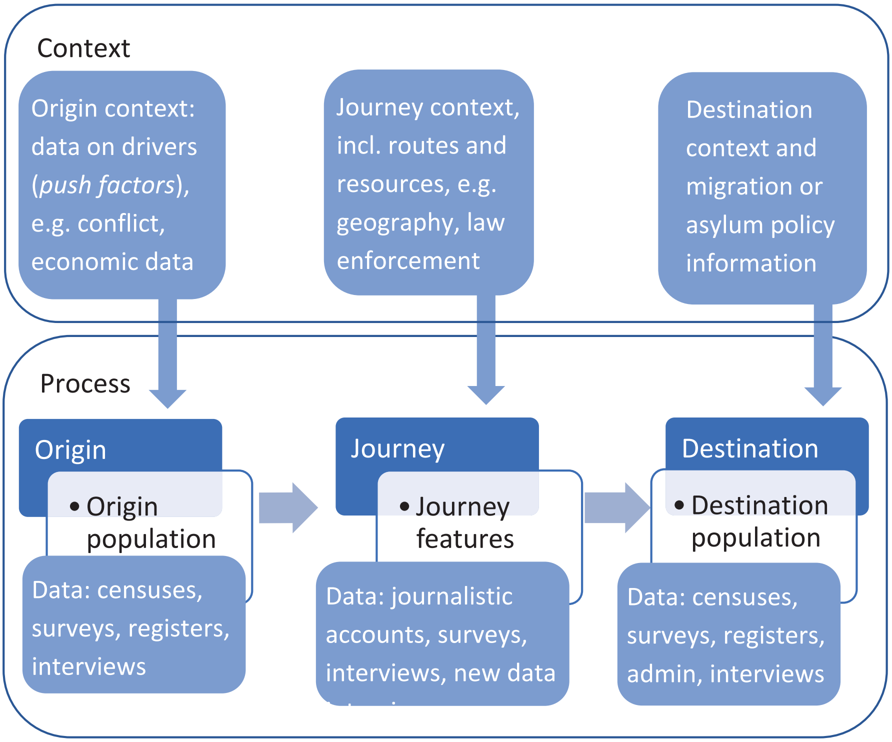


Fig. 4.2 Conceptual relationships between the process and context of migrant journeys and the corresponding data sources. (Source: own elaboration)

4.3.2 Process-Related Data

Among the process-related data, describing the various features of migration flows and migrants, be it for individual actors involved in migration (micro level) or for the whole populations (macro level), the main types of the information can be particularly useful for modelling are listed below.

Origin Populations. Information on the origin country population, such as data from a census or health surveys can be used for benchmarking. Data on age and sex distributions as well as other social and economic characteristics can be helpful in identifying specific subpopulations of interest, as well as in allowing for heterogeneity in the populations of migrants and stayers.

Destination Populations. A wide range of data on migrant characteristics, economic situation (employment, benefits), access to and use of information, intentions, health and wellbeing at the destination countries can be used for reconstructing various elements of migrant journeys, and assessing the situation of migrants at the destination. Note that with respect to migration processes, these data are typically retrospective, and can include a range of sources, from censuses and surveys, through administrative records, to qualitative interviews.

Registrations. Administrative and operational information from destination countries and international or humanitarian organisations, which register the arrival of migrants, can provide particularly timely data on numbers and characteristics as well as the timing of arrivals. These data also have clearly specified definitions due to their explicit collection purposes.

Journey. Any information available about the specific features of the journey itself also forms part of the process-related information. This could include data about durations of the different segments of the trip, or distinct features of the process of moving, which can be gauged for example from retrospective accounts or surveys, including qualitative interviews or journalistic accounts. Similarly, information on intermediaries, smugglers, and so on, as long as it is available and even remotely reliable, can be a part of the picture of the migrant journeys.

Information Flows. Availability of information on routes and contextual elements can also impact on migrants' decisions during the migration process. Even though the information itself can be contextual, its availability and trustworthiness are related to the migration process. Insights into the information availability (and its flipside: the uncertainty faced by migrants before, during and after their journeys) can be obtained from surveys, but there is an underutilised potential to use alternative sources ('new data'). The use of such data for analysis requires having appropriate legal and ethical safeguards and protocols in place, in order to ensure that the privacy of the subjects of data collection is stringently protected.

4.3.3 Contextual Data

Formal modelling offers a possibility of incorporating a wide range of different types of contextual data, shaping the migration decisions through the environment in which the migration processes take place. The list below is by no means exhaustive, and it concentrates on the four main aspects of the context – related to the origin, destination, policies, and routes.

Origin Context. Information on the situation in the countries and regions of origin can include such factors as conflict intensity, the presence of specific events or incidents, as well as reports from observers and media, and identify the key drivers related to the decision to migrate (corresponding to push factors in Lee's 1966 theoretical framework).

Destination Context. At the other end of the journey, information on destination countries, such as macro-economic data, attitudes and asylum acceptance rates, provides contextual information on the relative attractiveness of various destinations (corresponding to pull factors).

Policies and Institutions. Specifically related to the destination context, but also extending beyond it, information on various aspect of migration policy and law enforcement, including visa, asylum and settlement policies in destination and transit countries, as well as their changes in response to migration, additionally helps paint a more complete picture of the dynamic legal context of migrant decisions and of their possible interactions with those of other actors (border agents, policy makers, and so on).

Route Features. Contextual data on, for example, geographic terrain, networks, borders, barriers, transport routes and law enforcement can be used to assess different and variable levels of friction of distance, which can have long- and short-term impact on migration decisions and on actual flows (corresponding to intervening obstacles in Lee's framework). Here, information on the level of resources that are required for the journey, including availability of humanitarian aid, or intricacies of the smuggling market, as well as information on migrant access to resources, can provide additional insights into the migration routes and trajectories. Resources typically deplete over time and journey, which again impacts on decisions by determining the route, destination choice, and so on. This aspect can form a part of the set of route features mentioned above, or feature as a separate category, depending on the importance of the resource aspect for the analysis and modelling.

The multidimensionality of migration results in a patchwork of sources of information covering different aspects of the flows and the context in which they are taking place, often involving different populations and varying accuracy of measurement, which can be combined with the help of formal modelling (Willekens, 1994). At the same time, it implies the need for greater rigour and transparency, and a careful consideration of the data quality and their usefulness for a particular purpose, such as modelling.

Different process and context data are characterised by varying degrees of uncertainty, stemming from different features of the data collection processes, varying sample sizes, as well as a range of other quality characteristics. The quality of data itself is a multidimensional concept, which requires adequate formal analysis through a lens of a common assessment framework adopted for a range of different data sources that are to be used in the modelling exercise. We discuss methodological and practical considerations related to the design of such an assessment framework next, illustrated by an application to the case of recent Syrian migration to Europe.

4.4 Quality Assessment Framework for Migration Data

No perfect data exist, let alone concerning migration processes. The measurement of asylum migration requires particular care, going beyond the otherwise challenging measurement of other forms of human mobility (see e.g. Willekens, 1994). As mentioned in Chap. 2, the most widespread ways to measure asylum migration processes involve administrative data on events, which include very limited information about the context (Singleton, 2016). Other, well-known issues with the statistics involve duplicated records of the same people, for whom multiple events have been recorded, as well as the presence of undercount due to the clandestine nature of many asylum-related flows (Vogel & Kovacheva, 2008). The use of asylum statistics for political purposes adds another layer of complexity, and necessitates extra care when interpreting the data (Bakewell, 1999).

More generally, official migration statistics, as with all types of data, are social and political constructs, which strongly reflect the policy and research priorities prevalent at the time (for an example, see Bijak & Koryś, 2009). For this reason, the purpose and mechanisms of data collection also need to be taken into account in the assessment, as different types of information may carry various inherent biases. Given the potential dangers of relying on any single data source, which may be biased, when describing migration flows through modelling, multiple sources ideally need to be used concurrently, and be subject to formal quality assessment, as set out below.

4.4.1 Existing Frameworks

Assessing the quality of sources can allow us to make use of a greater range of information that may otherwise be discarded. Trustworthiness and transparency of data are particularly important for a politically sensitive topic of migration against the backdrop of armed conflict at the origin, and political controversies at the destination. Official legal texts, especially more recent ones, include references to data quality – European Regulation 862/2007 on migration and asylum statistics refers to and includes provisions for quality control and for assessing the "quality, comparability and completeness" of data (Art. 9).[2] Similarly, Regulation 763/2008 on population and housing censuses explicitly lists several quality criteria to be applied to the assessment of census data: relevance, accuracy, timeliness, accessibility, clarity, comparability, and coherence (Art. 6).[3]

Existing studies indicate several important aspects in assessing the quality of data from different sources. A key recent review of survey data specifically targeting asylum migrants, compiled by Isernia et al. (2018), provides a broad overview, as well as listing some specific elements to be considered in the data analysis. Surveys selected for this review highlight definitional issues with identifying the appropriate target population. Aspiring to clarity in definitional issues is an enduring theme in migration studies, asylum migration included (Bijak et al., 2017).

There are also several examples of existing academic studies in related areas, which aim at assessing the quality of sources of information. Specifically in the

[2] Regulation (EC) No 862/2007 of the European Parliament and of the Council of 11 July 2007 on Community statistics on migration and international protection, OJ L 199, 31.7.2007, p. 23–29, with subsequent amendments.

[3] Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses, OJ L 218, 13.8.2008, p. 14–20.

context of irregular migration, Vogel and Kovacheva (2008) proposed a four-point assessment scale for various available estimates, broadly following the 'traffic lights' convention (green, amber, red), but with the red category split into two subgroups, depending on whether the estimates were of any use or not. Recently, the traffic lights approach was used by Bijak et al. (2017) for asylum migration, and was based on six main assessment criteria: (1) Frequency of measurement; (2) Fit with the definitions; (3) Coverage in terms of time and space; (4) Accuracy, uncertainty and the presence of any biases; (5) Timeliness of data release; and (6) Evidence of quality assurance processes. In addition, similar assessments were carried out in the broader demographic studies of the consequences of armed conflict (GAO, 2006; Tabeau, 2009; Bijak & Lubman, 2016), including additional suggestions for how to address the various challenges of measurement.

4.4.2 *Proposed Dimensions of Data Assessment: Example of Syrian Asylum Migration*

The aim and nature of the modelling process imply that, while clarity of definitions is important, it is also possible to encompass a wider range of information sources and to assign different relative importance to these sources in the model. Our proposal for a quality assessment framework and uncertainty measures for different types of data is therefore multidimensional, as set out below. In particular, we propose six generic criteria for data assessment:

1. Purpose for data collection and its relevance for modelling
2. Timeliness and frequency of data collection and publication
3. Trustworthiness and absence of biases
4. Sufficient levels of disaggregation
5. Target population and definitions including the population of interest (in our case study, Syrian asylum migrants)
6. Transparency of the data collection methods

The need to identify the target population precisely is common for all types of data on migrants, but there are additional quality criteria specific to registers and survey-based sources. Thus, for register-based information an additional criterion relates to its completeness, while for surveys, their design, sampling strategy, sample sizes, and response rates are all aspects that need to be clearly set out in order to be assessed for rigour and good practice in data collection (Isernia et al., 2018).

In our framework, all criteria are evaluated according to a five-point scale, based on the traffic lights approach (green, amber, red), but also including half-way categories (green-amber and amber-red). The specific classification descriptors for assigning a particular source to a given class across all the criteria are listed in Table 4.1. Finally, for each source, a summary rating is obtained by averaging over the existing classes. This meta-information on data quality can be subsequently used in modelling either by adjusting the raw data, for example when these are known to be biased, or by reflecting the data uncertainty, when there are reasons to believe that they are broadly correct, yet imprecise.

Table 4.1 Proposed framework for formal assessment of the data sources for modelling the recent Syrian asylum migration to Europe

Criteria	Green	Amber	Red
Purpose: Is the purpose for data collection relevant to and appropriate for the aim of modelling?	Yes: aim is to estimate and/or understand migration from Syria	May be different purpose but still relevant	No: data collection for different purpose, impacting usefulness
Timeliness: Are the data published at sufficiently frequent intervals?	Yes: repeated measures published regularly	May be repeated measures but with long gaps and/or publication delays	No: one-off collection or long delay in publication
Trustworthiness: Is the source free from obvious biases or stated political aims?	Yes: evidence of impartiality	Unclear or unstated	No: clear evidence of bias
Disaggregation: Is there sufficient geographic and country of origin detail?	Yes: country of origin and destination fully disaggregated	Partial disaggregation e.g. for some variables of interest	No: not possible to identify sufficient detail
Target population and definitions: Are they Syrian migrants from specified time period?	Yes	May be a dataset including Syrian migrants	May be dataset of migrants but incorrect time period or nationality
Transparency: Is there a clearly stated purpose, design and methodology?	Yes, thorough	Yes, partial	No
Completeness:(1) Is there evidence of rigorous processes to capture and report the entire population?	Yes: stated aim and explicit strategies to achieve this	May not be sufficiently addressed but without evidence of gaps	No: evidence of gaps in dataset
Sample design:(2) Is there an appropriate sampling strategy and attempt to achieve sufficient sample size and response rate?	Yes, thoroughly described	Yes, partial	No or unclear

(1) Criterion specific to population registers
(2) Criterion specific to survey data and qualitative sources

The result of applying the seven quality criteria to 28 data sources identified as potentially relevant to modelling Syrian migration is summarised in Table 4.2 and presented in detail in Appendix B. The listing in the Appendix additionally

Table 4.2 Summary information on selected data sources related to Syrian migration into Europe

Focus and type	Process data		Context data
	Destination population	**Routes and journey**	
Macro-level sources			
- Quantitative	Mainly registrations, operational data and large survey data **Green**/**Amber** (**10**)	Data from surveys and registrations, as well as operational data **Amber** (**7**)	Official statistics of the receiving (**Green**) and sending (**Amber**/**Red**) countries (**2**)
- Qualitative			Policy, legal and other secondary information **Green**/**Amber** (**1**)
Micro-level sources			
- Quantitative	Large-scale and random surveys **Green**/**Amber** (**3**)	Targeted surveys **Amber** (**1**)	
- Qualitative	Surveys and in-depth interviews. **Amber** (**1**)	Surveys and in-depth interviews. **Amber** (**3**)	

Note: Figures in brackets (**0**) indicate the number of sources reviewed in each category. Their details are listed in Appendix B

includes 20 supplementary, general-level sources of information on migration processes, drivers or features, some aspects of which may also be useful for modelling, but which are unlikely to be at the core of the modelling exercise, and therefore have not been assessed following the same framework. For the latter group of sources, only generic information about source type and the purpose of collection is provided, alongside a basic description and access information.

On the whole, a majority of the data sources on Syrian asylum migration can be potentially useful in the modelling, at least to some degree. Most of the available data rely on registrations, operational data and surveys, and can be directly used to construct, parameterise or benchmark computational models of migration. The key proviso here is to know the limitations of the data and to be able to reflect them formally in the models. Caution needs to be taken when using some specific data sources, such as information from sending countries (in this case, Syria), due to a potential accumulation of several problems with their accuracy and trustworthiness, as detailed in Appendix B, but even for these, some high-level information can prove useful. Some suggestions as to the possible ways in which various data can be included in the models follow.

4.5 The Uses of Data in Simulation Modelling

One important consideration when choosing data to aid modelling is that the information used needs to be subsidiary to the research or policy questions that will be answered through models. For example, consider the questions about the journey (*process*), such as whether migrants choose the route with the shortest geographic distance, or is it mitigated by resources, networks and access to information? Exploring possible answers to this question would require gathering different sources of data, for example around general concepts such as 'friction' or 'resources', and would allow the modeller to go far beyond standard geographic measures of distance or economic measures of capital, respectively.

The arguments presented above lead to three main recommendations regarding the use of data in the practice of formal modelling.

First, there are no perfect data, so the expectations related to using them need to be realistic. There may be important trade-offs between different sources in terms of various evaluation criteria. For this reason, any data assessment has to be multidimensional, as different purposes may imply focus on different desired features of the data.

Second, any source of uncertainty, ambiguity or other imperfection in the data has to be formally reflected and propagated into the model. A natural language for expressing this uncertainty is one of probabilities, such as in the Bayesian statistical framework.

Third, the context of data collection has to be always borne in mind. Migration statistics – being to a large extent social and political constructs – are especially prone to becoming 'statistical artefacts' (see e.g. Bijak & Koryś, 2009), being distorted, and sometimes misinterpreted. With that in mind, the use of particular data needs to be ideally driven by the specific research and policy requirements rather than mere convenience.

One key extension of the formal evaluation of various data sources is to investigate the importance of the different pieces of knowledge, and to address the challenge of coherently incorporating the data on both micro- and macro-level processes, as well as the contextual information, together with their uncertainty assessment, in a migration model. If that could be successfully achieved, the results of the modelling can additionally help identify the future directions of data collection, strengthening the evidence base behind asylum migration and helping shape more realistic policy responses.

A natural formal language for describing the data quality or, in other words, the different dimensions of the uncertainty of the data sources, is provided by probability distributions, which can be easily included in a fully probabilistic (Bayesian) model for analysis. In the probabilistic description, two key aspects of data quality come to the fore: *bias* – by how much the source is over- or under-estimating the real process – which can be modelled by using the location parameters of the relevant distributions (such as mean, median and so on), and *variance* – how accurate the source is – which can be described by scale parameters (such as variance, standard deviation, precision, etc.). As in the statistical analysis of prediction errors, there may be important trade-offs between these two aspects: for example, with sample surveys, increasing the sample size is bound to decrease the variance, but if the sampling frame is mis-specified, this can come at the expense of an increasing bias – the estimates will be more precise, but in the wrong place.

Of the eight quality assessment criteria listed in Table 4.1, the first two (purpose and timeliness) are of a general nature, and – depending on the aim of the modelling endeavours – can be decisive in terms of whether or not a given source can be used at all. The remaining ones can be broadly seen either as contributing to the bias of a source (definitions of the target populations, trustworthiness of data collection, and

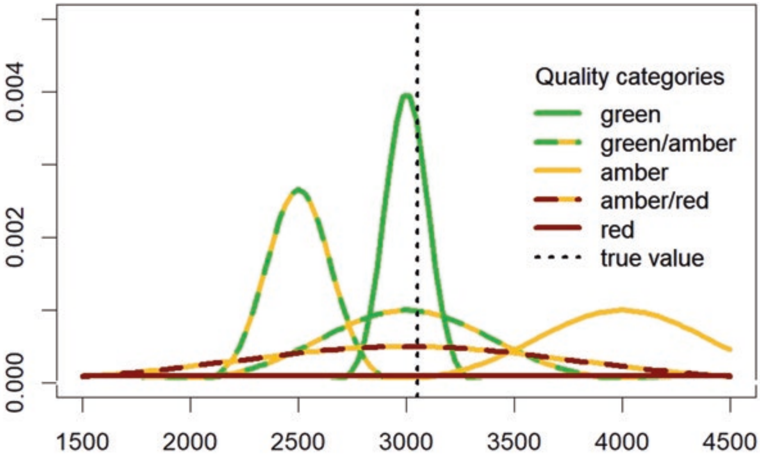


Fig. 4.3 Representing data quality aspects through probability distributions: stylised examples. (Source: own elaboration)

completeness of coverage), or to its variance (level of disaggregation, sample design, and transparency of data collection mechanisms). The interplay between these factors can offer important guidance as to what probabilistic form a given distribution needs to take, and with what parameters.

Figure 4.3 illustrates some stylised possibilities of how data falling into different quality classes can map onto the reality, depicted by the vertical black line. Hence, we would expect a source classified as 'green' to have minimal or negligible bias and relatively small variance. The 'green/amber' sources could either exhibit some bias, the extent of which can be at least approximately assessed, or maybe a somewhat larger variance – although both of these issues together would typically signify the 'amber' quality level and a need for additional care when handling the data. Needless to say, sources falling purely into the 'red' quality category should not be used in the analysis at all, while the data in the 'amber/red' category should only be used with utmost caution, given that they can point to general tendencies, but not much beyond that.

As discussed in Chap. 2, the data can enter into the modelling process at different stages. First, as summarised in Fig. 2.1, modelling starts with observation of the properties of the processes being modelled. What follows, in the inductive step of model construction, is the inclusion of information about the features and structures of the process, as well as the information on the contributing factors and drivers. Hence, at the steps following the principles of the classical inductive approach, all relevant context data need to be included, as well as micro-level data on the building blocks of the process itself. Subsequently, so that the model is validated against the reality, macro-level data on the process can be used for benchmarking. In other words, micro-level process data, as well as context data become model inputs, whereas macro-level process data are used to calibrate model outputs.

A natural way to include the uncertainty assessment of the different types of data sources is then, for the inputs, to feed the data into the model in a probabilistic form (as probability distributions), and, for the outputs, to include in the model an additional error term that is intended to capture the difference between the processes being modelled and their empirical measurements (see Chap. 5). Box 4.1 presents an illustration related to a set of possible data sources, which may serve to augment the Routes and Rumours model introduced in Chap. 3 and to develop it further, together with their key characteristics and overall assessment. More details for these sources are offered in Appendix B.

Box 4.1: Datasets Potentially Useful for Augmenting the Routes and Rumours Model

As described in Chap. 3, temporal detail and spatial information are important for this model in order to understand more about the emergence of migration routes. We focused on the Central Mediterranean route, utilising data on those intercepted leaving Libya or Tunisia, losing their lives during the sea crossing, or being registered upon arrival in Italy. One exception was the retrospective Flight 2.0 survey, carried out in Germany, which looked into the use of information by migrants during their journey. All the data included below are quantitative, reported at the macro-level (although Flight 2.0 recorded micro-level survey data), and relate to the migration process. The available data are listed in Table 4.3 below; for this model monthly totals were used. In addition, OpenStreetMap (see source S02 in Appendix B) data provides real world geographic detail. For a general quality assessment of data sources, see Appendix B, where the more detailed notes for each dataset provide additional relevant information and give some brief explanation of the reasoning behind particular quality ratings.

Table 4.3 Selection of data sources which can inform the Routes and Rumours model, with their key features and quality assessment

Reference in Appendix B	Source	Content focus	Source and time detail	Quality rating	Bias & variance
11	IOM Missing Migrants: Flows	Destination population: Interceptions by Libyan /Tunisian coastguards	Operational & admin, monthly data	**Amber**	Medium undercount & variance
12	IOM Missing Migrants: Deaths	Number of recorded deaths during Central Med crossings	Operational & journalistic, daily data	**Amber**	Medium undercount & variance
13	IOM Displacement Tracker	Destination population: Daily arrivals registered in Italy	Operational, daily data	**Green/ amber**	Small undercount & variance
24	Flight 2.0 / Flucht 2.0	Data on information use and levels of trust *en route* to Germany	One-off survey	**Amber**	Unknown bias, large variance

Source: see Appendix B for details related to individual sources

Of course, there are also other methods for dealing with missing, incomplete or fragmented data, coming from statistics, machine learning and other emerging areas of broader 'data science'. The review of such methods remains beyond the scope of this book, but it suffices to name a few, such as various approaches to imputation, which have been covered extensively e.g. in Kim and Shao (2014), or data matching, which in machine learning is also referred to as data fusion, also covered by a broad literature (e.g. Bishop et al., 1975/2007; D'Orazio et al., 2006; Herzog et al., 2007). A comprehensive recent review of the field was provided by Little and Rubin (2020). In the migration context, some of these methods, such as micro-level matching, are not very feasible, unless individual-level microdata are available with enough personal detail to enable the matching. For ethical reasons, this should not be possible outside of very secure environments under strictly controlled conditions; therefore this may not be the right option for most applied migration research questions. Better, and more realistic options include reconciliation of macro-level data through statistical modelling, such as in the Integrated Modelling of European Migration work (Raymer et al., 2013), producing estimates of migration flows within Europe with a description of uncertainty. Such estimates can then be subject to a quality assessment as well, and be included in the models following the general principles outlined above.

4.6 Towards Better Migration Data: A General Reflection[4]

As discussed before, the various types of contemporary migration data, as well as other associated information on the related factors and drivers, are still far from achieving their potential. The data are typically available only after a time delay, which poses problems for applications requiring timeliness, such as rapid response in the case of asylum migration. Data on migrants, as opposed to counts of migration events, are still relatively scarce, and particularly lacking are longitudinal studies involving migrant populations. The existing data are not harmonised, nor are they exactly 'interoperable' – ready to be used for different purposes or aims, with tensions between particular policy objectives and the information the data can provide.

No matter what practical solutions are adopted for the use of migration data in modelling, several important caveats need to be made when it comes to the interpretation of the meaning of the data. As argued above, the data themselves are

[4] Part of the discussion is inspired by a debate panel on migration modelling, held at the workshop on the uncertainty and complexity of migration, in London on 20–21 November 2018. The discussion, conducted under the Chatham House rule (no individual attribution), covered two main topics: migration knowledge gaps and ways to fill them, and making simulation models useful for policy. We are grateful to (in alphabetical order) Ann Blake, Nico Keilman, Giampaolo Lanzieri, Petra Nahmias, Ann Singleton, Teddy Wilkin and Dominik Zenner for sharing their views.

social constructs and the product of their times, and as such, are not politically neutral. These features put the onus on the modellers and users, who need to be aware of the social and political baggage associated with the data. Besides the need to be conscious of the context of the data collection, there can be a trap associated with bringing in too much of the analysts' and modellers' own life experience to modelling. This, in turn, requires particular attention in the context of modelling of migration processes that are global in nature, or consider different cultural contexts than the modellers' own.

Similar reservations hold from the modelling point of view, especially when dealing with agent-based models attempting to represent human behaviour. Such models often imply making very strong value judgements and assumptions, for example with respect to the objective functions of individual agents, or the constraints under which they operate. The values that are reflected in the models need to be made explicit, also to acknowledge the role of the research stakeholders, for the sake of transparency and to ensure public trust in the data. It has to be clear who defines the research problem underlying the modelling, and what their motivations were.

Another aspect of trust relates to the new forms of data, such as digital traces from social media or mobile phones, where their analytical potential needs to be counterbalanced by strong ethical precautions related to ensuring privacy. This is especially crucial in the context of individual-level data linking, where many different sources of data taken together can reveal more about individuals than is justified by the research needs, or than should be ethically admissible. This also constitutes a very important challenge for traditional data providers and custodians, such as national and international statistical offices and other parts of the system of official statistics, whose future mission can include acting as legal, ethical and methodological safeguards of the highest professional standards with respect to migration data collection, processing, storage and dissemination.

Another important point is that the modelling process, especially if employed in an iterative manner, as argued in Chap. 2 and throughout this book, can act as an important pathway towards discovering further gaps in the existing knowledge and data. This is a more readily attainable aim than a precise description or explanation of migration processes, not to mention their prediction. Additionally, this is the place for a continuous dialogue between the modellers and stakeholders, as long as the underpinning ideas and concepts are well defined, simple, clear and transparent, and the expectations as to what the data and models can and cannot deliver are realistic.

To achieve these aims, open communication about the strengths and limitations of data and models is crucial, which is one of the key arguments behind an explicit treatment of different aspects of data quality, as discussed above. These features can help both the data producers and users better navigate the different guises of the uncertainty and complexity of migration processes, by setting the minimum quality standards – or even requirements – that should be expected from the data and models alike. A prerequisite for that is a high level of statistical and scientific literacy, not only of the users and producers of data and models, but also ideally among the general public. To that end, while the focus of this chapter is on the limitations of various sources of data, and what aspects of information they are able to provide, the next one looks specifically at the ways in which the formal model analysis can help shed light on information gaps in the model, and also utilise empirical information at different stages of the modelling process.

Chapter 5
Uncertainty Quantification, Model Calibration and Sensitivity

Jakub Bijak and Jason Hilton

Better understanding of the behaviour of agent-based models, aimed at embedding them in the broader, model-based line of scientific enquiry, requires a comprehensive framework for analysing their results. Seeing models as tools for experimenting *in silico*, this chapter discusses the basic tenets and techniques of uncertainty quantification and experimental design, both of which can help shed light on the workings of complex systems embedded in computational models. In particular, we look at: relationships between model inputs and outputs, various types of experimental design, methods of analysis of simulation results, assessment of model uncertainty and sensitivity, which helps identify the parts of the model that matter in the experiments, as well as statistical tools for calibrating models to the available data. We focus on the role of emulators, or meta-models – high-level statistical models approximating the behaviour of the agent-based models under study – and in particular, on Gaussian processes (GPs). The theoretical discussion is illustrated by applications to the Routes and Rumours model of migrant route formation introduced before.

5.1 Bayesian Uncertainty Quantification: Key Principles

Computational simulation models can be conceptualised as tools for carrying out "opaque thought experiments" (Di Paolo et al., 2000), where the links between model specification, inputs and outputs are not obvious. Many different sources of uncertainty contribute to this opaqueness, some of which are related to the uncertain state of the world – the reality being modelled – and our imperfect knowledge about it, while others relate to the different elements of the models. In the context of computational modelling, Kennedy and O'Hagan (2001) proposed a taxonomy of sources of error and uncertainty, the key elements of which encompass: model inadequacy – discrepancy between the model and the reality it represents; uncertainty in observations (including measurement errors); uncertainty related to the unknown model parameters; pre-specified parametric variability, explicitly included in the

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_5

model via probability distributions; errors in the computer code; and residual variability, left after accounting for every other source.

The tools of probability and statistics, and in particular Bayesian statistics, offer a natural way of describing these different sources of uncertainty, by expressing every modelled quantity as a random variable with a probability distribution. The mechanism of Bayesian inference, by which the prior quantities (distributions) are combined with the likelihood of the data to yield posterior quantities, helps bring together the different sources of knowledge – data and *prior knowledge*, the latter for example elicited from experts in a given domain.

There is a long history of mutual relationships between Bayesian statistics and social sciences, including demography, dating back to the seminal work of Thomas Bayes and Pierre-Simon de Laplace in the late eighteenth century (Courgeau, 2012, see also Foreword to this book). A thorough introduction to Bayesian statistics is beyond the scope of this book, but more specific details on Bayesian inference and applications in social sciences can be found in some of the excellent textbooks and reference works (Lynch, 2007; Gelman et al., 2013; Bryant & Zhang, 2018), while the use of Bayesian methods in demography was reviewed in Bijak and Bryant (2016).

The Bayesian approach is especially well-suited for carrying out a comprehensive analysis of uncertainty in complex computational models, as it can cover various sources and forms of error in a coherent way, from the estimation of the models, to prediction, and ultimately to offering tools for supporting decision making under uncertainty. In this way, Bayesian inference offers an explicit, coherent description of uncertainty at various levels of analysis (parameters, models, decisions), allows the expert judgement to play an important role, especially given deficiencies of data (which are commonplace in such areas as migration), and can potentially offer more realistic assessment of uncertainty than traditional methods (Bijak, 2010).

Uncertainty quantification (UQ) as a research area looking into uncertainty and inference in large, and possibly analytically intractable, computational models, spanning statistics, applied mathematics and computing, has seen rapid development since the early twenty-first century (O'Hagan, 2013; Smith, 2013; Ghanem et al., 2019). The two key aspects of UQ include propagating the uncertainty through the model and learning about model parameters from the data (calibration), with the ultimate aim of quantifying and ideally reducing the uncertainty of model predictions (*idem*). The rapid development of UQ as a separate area of research, with distinct methodology, has been primarily motivated by the increase in the number and importance of studies involving large-scale computational models, mainly in physical and engineering applications, from astronomy, to weather and climate, biology, hydrology, aeronautics, geology and nuclear fusion (Smith, 2013), although with social science applications lagging behind. A recent overview of UQ was offered by Smith (2013), and a selection of specific topics were given detailed treatment in the living reference collection of Ghanem et al. (2019). For the reasons mentioned before, Bayesian methods, with their coherent probabilistic language for describing all unknowns, offer natural tools for UQ applications.

The main principles of UQ include a comprehensive description of different sources of uncertainty (error) in computational models of the complex systems under study, and inference about the properties of these systems on that basis. To do that, it relies on specific methods from other areas of statistics, mathematics and computing, which are tailored to the UQ problems. These methods, to a large extent, rely on the use of *meta-models* (or *emulators*, sometimes also referred to as surrogate models) to approximate the dynamics of the complex computational models, and facilitate other uses. Specific methods that have an important place in UQ include uncertainty analysis, which looks at how uncertainty is propagated through the model, and sensitivity analysis, which aims to assess which elements of the model and, in particular, which parameters matter for the model outputs (Oakley & O'Hagan, 2002). Besides, for models with predictive ambitions, methods for calibrating them to the observed data become of crucial importance (Kennedy & O'Hagan, 2001). We discuss these different groups of methods in more detail in the remainder of this chapter, starting from a general introduction to the area of statistical experimental design, which is underpinning the construction and calibration of meta-models, and therefore provides foundations for many of the UQ tools and their applications.

5.2 Preliminaries of Statistical Experimental Design

The use of tools of statistical experimental design in the analysis of the results of agent-based models starts from the premise that agent-based models, no matter how opaque, are indeed experiments. By running the model at different parameter values and with different settings – that is, experimenting by repeated execution of the model *in silico* (Epstein & Axtell, 1996) – we learn about the behaviour of the model, and hopefully the underlying system, more than would be possible otherwise. This is especially important given the sometimes very complex, non-transparent and analytically intractable nature of many computational simulations.

Throughout this chapter, we will define an experiment as a process of measuring a "stochastic response corresponding to a set of … input variables" (Santner et al., 2003, p. 2). A computer experiment is a special case, based on a mathematical theory, implemented by using numerical methods with appropriate computer hardware and software (*idem*). Potential advantages of computer experiments include their built-in features, such as replicability, relatively high speed and low cost, as well as their ability to analyse large-scale complex systems. Whereas the quality standards of natural experiments are primarily linked to the questions of randomisation (as in randomised control trials), blocking of similar objects to ensure homogeneity, and replication of experimental conditions, computer experiments typically rely on deterministic or stochastic simulations, and require transparency and thorough documentation as minimum quality standards (*idem*).

Computer experiments also differ from traditional, largely natural experiments thanks to their wider applicability, also to social and policy questions, with different ethical implications than experiments requiring direct human participation. In some social contexts, other experiments would not be possible or ethical. For example, analysing optimal ways of evacuating people facing immediate danger (such as fire or flood), very important for tailoring operational response, cannot involve live experiments in actual dangerous conditions. In such cases, computer experiments can provide invaluable insights into the underlying processes, possibly coupled with ethically sound natural experiments carried out in safe conditions, for example on the ways large groups of people navigate unknown landscapes.

To make the most of the computer experiments, their appropriate planning and design becomes of key importance. To maximise our information gains from experimentation, which typically comes at a considerable computational cost (as measured in computing time), we need to know at which parameter values and with which settings the models need to be run. The modern statistical theory and practice of experimental design dates back to the agricultural work of Sir Ronald Fisher (1926), with the methodological foundations fully laid out, for example, in the much-cited works of Fisher (1935/1958) and Cox (1958/1992). Since then, the design of experiments has been the subject of many refinements and extensions, with applications specifically relevant for analysing computer models discussed in Santner et al. (2003) and Fang et al. (2006), among others.

The key objectives of the statistical design of experiments are to help understand the relationship between the inputs and the outcome (response), and to maximise information gain from the experiments – or to minimise the error – under computational constraints, such as time and cost of conducting the experiments. The additional objectives may include aiding the analytical aims listed before, such as the uncertainty or sensitivity analysis, or model-based prediction.

As for the terminology, throughout this chapter we use the following definitions, based on the established literature conventions. Most of these definitions follow the conventions presented in the Managing Uncertainty in Complex Models online compendium (MUCM, 2021).

Model (simulator) "A representation of some real-world system, usually implemented as a computer program" (MUCM, 2021), which is transforming inputs into outputs;

Factor (input) "A controllable variable of interest" (Fang et al., 2006, p. 4), which can include model parameters or other characteristics of model specification.

Response (output) A variable representing "specific properties of the real system" (Fang et al., 2006, p. 4), which are of interest to the analyst. The output is a result of an individual run (implementation) of a model for a given set of inputs.

Calibration The analytical process of "adjusting the inputs so as to make the simulator predict as closely as possible the actual observation points" (MUCM, 2021);

Calibration parameter "An input which has … a single best value" with respect to the match between the model output and the data (reality), and can be therefore used for calibration (MUCM, 2021);

Model discrepancy (inadequacy) The residual difference between the observed reality and the output calibrated at the best inputs (calibration parameters);

Meta-model (emulator, surrogate) A statistical or mathematical model of the underlying complex computational model. In this chapter, we will mainly look at statistical emulators.

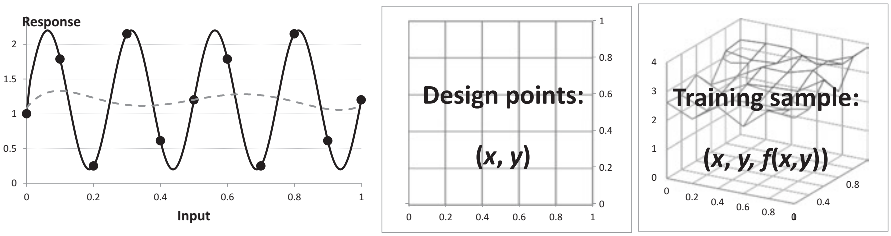


Fig. 5.1 Concepts of the model discrepancy (left), design (middle) and training sample (right). For the discrepancy example, the real process (solid line) is $f(x) = 1.2 \sin(8\pi x)$, and the model (dashed line) is a polynomial of order 6, fitted by using ordinary least squares. The calibration parameters are then the coefficients of the polynomial, and the model discrepancy is the difference between the values of the two functions. (Source: own elaboration)

Design "A choice of the set of points in the space of simulator inputs at which the simulator is run" (MUCM, 2021), and which then serve as the basis for model analysis;

Training sample Data comprising inputs from the design space, as well as the related outputs, which are used to build and calibrate an emulator for subsequent use in the analysis.

The diagrams in Fig. 5.1 illustrate the concepts of model discrepancy, design and training sample.

There are different types of design spaces, which are briefly presented here following their standard description in the selected reference works (Cox, 1958/1992; Santner et al., 2003; Fang et al., 2006). To start with, a **factorial design** is based on combinations of design points at different levels of various inputs, which in practice means being a subset of a hyper-grid in the full parameter space, conventionally with equidistant spacing between the grid points for continuous variables. As a special case, the **full factorial design** includes *all* combinations of *all* possible levels of *all* inputs, whereas a **fractional factorial design** can be any subset of the full design. Due to practical considerations, and the 'combinatorial explosion' of the number of possible design points with the increasing number of parameters, limiting the analysis to a fractional factorial design, for the sake of efficiency, is a pragmatic necessity.

There are many ways in which fractional factorial designs can be constructed. One option involves random design, with design points randomly selected from the full hyper-grid, e.g. by using simple random sampling, or – more efficiently – stratified sampling, with the hyper-grid divided into several strata in order to ensure good coverage of different parts of the parameter space. An extension of the stratified design is the Latin Hypercube design – a multidimensional generalisation of a two-dimensional idea of a Latin Square, where only one item can be sampled from each row and each column, similarly to a Sudoku puzzle. In the multidimensional case, only one item can be sampled for each level in every dimension; that is, for every input (*idem*).

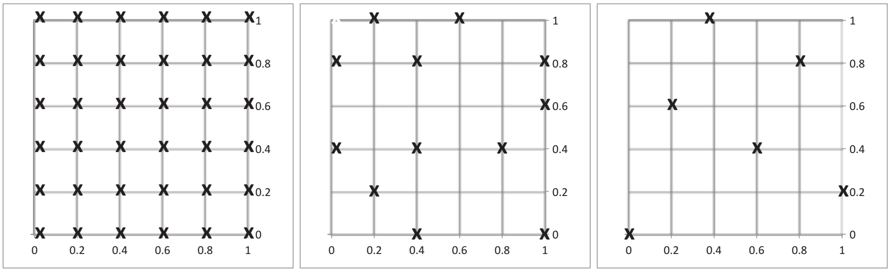


Fig. 5.2 Examples of a full factorial (left), fractional factorial (middle), and a space-filling Latin Hypercube design (right). (Source: own elaboration)

More formally, with a discrete Latin Hypercube design we ideally want to cover the whole range of the distribution of each of the K input variables, X_i. For each i, let this input range be divided into N equal parts (bins), from which N elements satisfying the Latin Hypercube rule can be sampled. This can be done in $[N^K (N-1)^K \dots 1^K]/[N (N-1) \dots 1] = (N!)^{K-1}$ different ways. Among those, some designs can be space filling, with points spread out more evenly in the multidimensional space, while some others are non-space filling, leaving large 'gaps' without sampling points, which is undesirable. In practice, the available algorithms try ensuring that the design is as much space filling as possible, for example by maximising the minimum distances between the design points, or minimising correlations between factors (Ranjan & Spencer, 2014). Examples of a full factorial, fractional factorial, and a space-filling Latin Hypercube design spaces for a 6 × 6 grid are shown in Fig. 5.2.

Generally, Latin Hypercube samples have desirable statistical properties, and are considered more efficient than both random and stratified sampling (see the examples given by McKay et al., 1979). One alternative approach involves model-based design, which requires a model for the results that we expect to observe based on any design – for example an emulator – as well as an optimality criterion, such as minimising the variance, maximising the information content, or optimising a certain decision based on the design, in the presence of some loss (cost) function. The optimal model-based design is then an outcome of optimising the criterion over the design space, and a typical example involves design that will minimise the variance of an emulator built for a given model.

If the parameter space is high-dimensional, it is advisable to reduce the dimensionality first, to limit the analysis to those parameters that matter the most for a given output. This can be achieved by carrying out pre-screening, or sequential design, based on sparse fractional factorial principles, which date back to the work of Davies and Hay (1950). Among the different methods that have been proposed for that purpose, Definitive Screening Design (Jones & Nachtsheim, 2011, 2013) is relatively parsimonious, and yet allows for identifying the impact of the main effects of the parameters in question, as well as their second-order interactions.

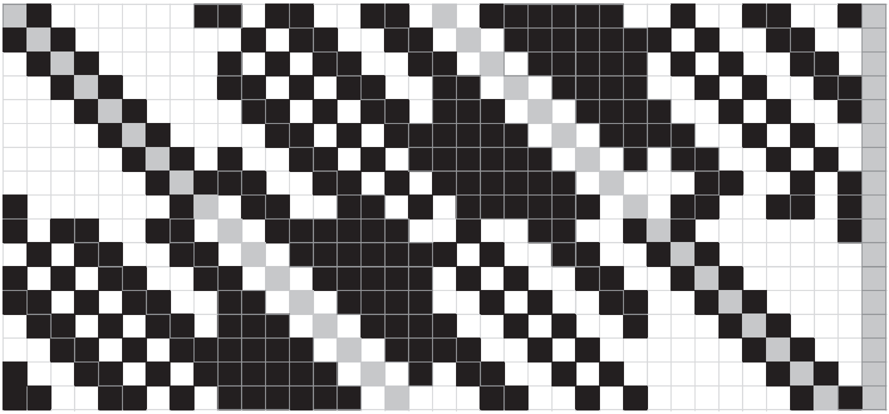


Fig. 5.3 Visualisation of a transposed Definite Screening Design matrix D′ for 17 parameters. Black squares correspond to high parameter values (+1), white to low ones (−1), and grey to middle ones (0). (Source: own elaboration)

The Definitive Screening Design approach is based on so-called conference matrices $\mathbf{C}_{m \times m}$, such that $1/(m–1)$ $\mathbf{C'C} = \mathbf{I}_{m \times m}$, where m is either the number of parameters (if m is even), or the number of parameters plus 1 (if m is odd). The elements of matrix **C** can take three values: +1 for the 'high' values of the respective parameters, 0 for the 'middle' values, and −1 for the 'low' values, where the specifics are set by the analyst after looking at the possible range of each parameter. The design matrix **D** is then obtained by stacking the matrices **C**, –**C** and a vector of middle values, **0**, so that $\mathbf{D'} = [\mathbf{C'}, –\mathbf{C'}, \mathbf{0'}]'$ (Jones & Nachtsheim, 2011, 2013). The rows of matrix **D′** represent parameters (if m is odd, the last row can be omitted), and the columns represent the design points, at which the pre-screening experiments are to be run: $2m + 1$ if m is even, and $2m + 3$ if m is odd. An example of a design matrix **D′** for $m = 17$ parameters, implying 37 design points, is illustrated in Fig. 5.3.

Once the model is run, either a descriptive exploration of the output, or a formal sensitivity analysis (see Sect. 5.4) can indicate which parameters can be dropped without much information loss. In Box 5.1, we present an illustration of the proposed approach for the Routes and Rumours migration model, which was introduced in Chap. 3, with some detailed results reported in Appendix C.

Other methods that can be used for pre-screening of the model parameter space include Automatic Relevance Determination (ARD), and Sparse Bayesian Learning (SBL), dating back to the work of MacKay (1992), which both use Bayesian inference to reduce the dimensionality of the parameter space by 'pruning' the less relevant dimensions (for an overview, see e.g. Wipf & Nagarajan, 2008). From the statistical side, these methods link with Bayesian model selection (Hoeting et al., 1999) and the Occam's razor principle, which favours simpler models (in this case, models with fewer parameters) over more complex ones. From the machine

Box 5.1: Designing Experiments on the Routes and Rumours Model
This running example illustrates the process of experimental design and analysis for the model of migrant routes and information exchange introduced in Chap. 3. In this case, the pre-screening was run on $m = 17$ parameters: six related to information exchange and establishing or retaining contacts between the agents; four related to the way in which the agents explore their environment, with focus on the speed and efficiency; four describing the quality of the routes, resources and the environment; and three related to the resource economy: resources and costs.

The Definitive Screening Design was applied to the initial 17 parameters, with 37 design points as shown in Fig. 5.3, with the low, medium and high values corresponding to ¼, ½ and ¾ of the respective parameter ranges. At these points, four model outputs were generated: *mean_freq_plan*, related to agent behaviour, describing the proportion of time the agents were following their route plan; *stdd_link_c*, describing route concentration, measuring the standard deviation of the number of visits over all links; *corr_opt_links*, linked to route optimality, operationalised as the correlation of the number of passages over links with the optimal scenario; and *prop_stdd*, measuring replicability, here approximated by the standard deviation of traffic *between* replicate runs (see also Bijak et al., 2020). For the first three outputs, 10 samples were taken at each point, to allow for the cross-replication error in the computer code, while the fourth one already summarised cross-replicate information.

The results of the model were analysed by using Gaussian process emulators fitted in the GEM-SA package and used for conducting a preliminary sensitivity analysis (Kennedy & Petropoulos, 2016, see also Sects. 5.3 and 5.4). Across the four outputs, five parameters related to information exchange proved to be of primary importance: the probabilities of losing a contact (*p_drop_contact*), communicating with local agents (*p_info_mingle*), communicating with contacts (*p_info_contacts*), and exchanging information through communication (*p_transfer_info*), as well as the information noise (*error*). The sensitivity analysis indicated that these five parameters were jointly responsible for explaining between 30% and 83% of the variation of the four outputs, and almost universally included the top three most influential parameters for each output. For further experiments, two parameters related to exploration were also manually included, to make sure that the role of this part of the model was not overlooked. These were the speed of learning about the environment (*speed_expl*), and probability of finding routes and connecting links during the local exploration (*p_find*). Detailed results in terms of shares of variances attributed to individual inputs are reported in Appendix C.

The results proved largely robust to changes in the random seed, especially when a separate variance term for the error in computer code (the 'nugget' variance) was included, and also when comparing them with the outcome of a standard ANOVA procedure. For the further steps of the analysis, a Latin Hypercube sample design was generated in GEM-SA, with $N = 65$ design points, and six replicates of the model run at each point, so with 390 samples in total. This sample was used to build and test emulators and carry out uncertainty and sensitivity analysis, as discussed in the next section.

learning side, these approaches also have common features with support vector machines (Tipping, 2001). As the ARD and SBL methods are quite involved, we do not discuss them here in more detail, but a fuller treatment of some of the related approaches can be found, for example, in Neal (1996).

5.3 Analysis of Experiments: Response Surfaces and Meta-Modelling

There are several ways in which the results of complex computational experiments can be analysed. The two main types of analysis, linking to different research objectives, include *explanation* of the behaviour of the systems being modelled, as well as the *prediction* of this behaviour outside of the set of observed data points. In this chapter, broadly following the framework of Kennedy and O'Hagan (2001), we look specifically at four types of explanations:

- Response of the model output to changes in inputs, both descriptive and model-based.
- Sensitivity analysis, aimed at identifying the inputs which influence the changes in output.
- Uncertainty analysis, describing the output uncertainty induced by the uncertain inputs.
- Calibration, aimed at identifying a combination of inputs, for which the model fits the observed data best, by optimising a set of calibration parameters (see Sect. 5.2).

Notably, Kleijnen (1995) argued that these types of analysis (or equivalent ones) also serve an internal modelling purpose, which is model *validation*, here understood as ensuring "a satisfactory range of accuracy consistent with the intended application of the model" (Sargent, 2013: 12). This is an additional model quality requirement beyond a pure code *verification*, which is aimed at ensuring that "the computer program of the computerized model and its implementation are correct" (*idem*). In other words, carrying out different types of explanatory analysis, ideally together, helps validate the model internally – in terms of inputs and outputs – as well as externally, in relation to the data. Different aspects of model validation are reviewed in a comprehensive paper by Sargent (2013).

At the same time, throughout this book we interpret *prediction* as a type of analysis involving both *interpolation* between the observed sample points, as well as *extrapolation* beyond the domain delimited by the training sample. Extrapolation comes with obvious caveats related to going beyond the range of training data, especially in a multidimensional input space. Predictions can also serve the purpose of model validation, both out-of-sample, by assessing model errors on new data points, outside of the training sample, as well as in-sample (cross-validation), on the same

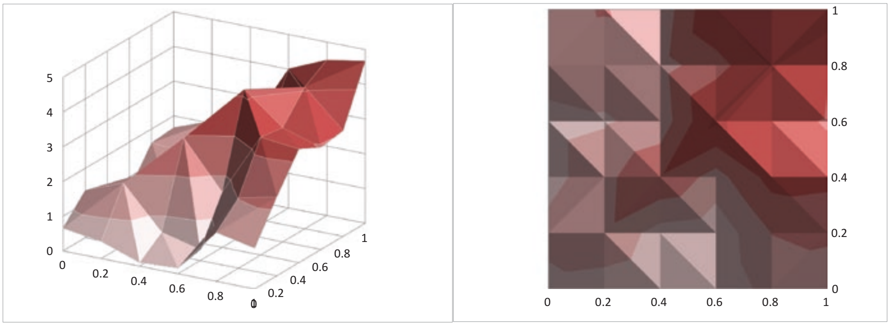


Fig. 5.4 Examples of piecewise-linear response surfaces: a 3D graph (left) and contour plot (right). (Source: own elaboration)

data points, by using such well-known statistical techniques as leave-one-out, jackknife, or bootstrap.

In all these cases, mainly because of computational constraints – chiefly the time it takes the complex computer models to run – it is much easier to carry out the explanatory and predictive analysis based on the surrogate meta-models. To that end, we begin the overview of the methods of analysis by discussing response surfaces and other meta-models in this section, before moving to the uncertainty and sensitivity analysis in Sect. 5.4, and calibration in Sect. 5.5.

The first step in analysing the relationships between model inputs and outputs is a simple, usually graphical description of a *response surface*, which shows how model output (response) varies with changes in the input parameters (for a stylised example, see Fig. 5.4). This is useful mainly as a first approximation of the underlying relationships, although even at this stage the description can be formalised, for example by using a regression meta-model, either parametric or non-parametric. Such a simple meta-model can be estimated from the data and allows the inclusion of some – although not all – measures of error and uncertainty of estimation, mainly those related to the random term and parameter estimates. The typical choices for regression-based approximations of the response surfaces include models including just the main (first-order) effects for the individual parameters, as well as those additionally involving quadratic effects, and possible interaction terms (Kleijnen, 1995). Other options include local regression models and spline-based non-parametric approaches.

The uses of emulators based on Gaussian processes date back to approaches that later became known as Kriging, named after South African geostatistician, Danie G Krige, who developed them in early 1950s[1] (Cressie, 1990). The more recent developments, specifically tailored for the meta-analysis of complex computational models, are largely rooted in the methodology proposed in the seminal papers of

[1] It is worth noting that, according to Cressie (1990), similar methods were independently proposed already in the 1940s by Herman Wold, Andrey Nikolaevich Kolmogorov and Norbert Wiener.

Kennedy and O'Hagan (2001) and Oakley and O'Hagan (2002), presenting the construction and estimation of Bayesian GP emulators.

The basic description of the GP emulation approach, presented here after Kennedy and O'Hagan (2001, 431–434), is as follows. Let the (multidimensional) model inputs $\mathbf{x}$ from the input (parameter) space $\mathbf{X}$, $\mathbf{x} \in \mathbf{X}$, be mapped onto a one-dimensional output $y \in \mathbf{Y}$, by the means of a function f, such that $y = f(\mathbf{x})$. The function f follows a GP distribution, if "for every $n = 1, 2, 3, \ldots$, the joint distribution of $f(\mathbf{x_1}), \ldots, f(\mathbf{x_n})$ is multivariate normal for all $\mathbf{x_1}, \ldots, \mathbf{x_n} \in \mathbf{X}$" (*idem*: 432). This distribution has a mean m, typically operationalised as a linear regression function of inputs or their transformations $\mathbf{h}(\cdot)$, such that $m(\mathbf{x}) = \mathbf{h}(\mathbf{x})' \boldsymbol{\beta}$, with some regression hyperparameters $\boldsymbol{\beta}$. The GP covariance function includes a common variance term across all inputs, σ^2, as well as a non-negative definite correlation matrix between inputs, $\mathbf{c}(\cdot,\cdot)$. The GP model can be therefore formally written as:

$$f(\cdot) \mid \beta, \sigma^2, R \sim MVN\left(m(\cdot); \sigma^2 \mathbf{c}(\cdot,\cdot)\right) \tag{5.1}$$

The correlation matrix $\mathbf{c}(\cdot,\cdot)$ can be parameterised, for example, based on the distances between the input points, with a common choice of $\mathbf{c}(\mathbf{x_1}, \mathbf{x_2}) = \mathbf{c}(\mathbf{x_1} - \mathbf{x_2}) = \exp(-(\mathbf{x_1} - \mathbf{x_2})' \mathbf{R} (\mathbf{x_1} - \mathbf{x_2}))$, with a *roughness matrix* $\mathbf{R} = \mathrm{diag}(r_1, \ldots, r_n)$, indicating the strength of response of the emulator to particular inputs. To reflect the uncertainty of the computer code, the matrix $\mathbf{c}(\cdot,\cdot)$ can additionally include a separate variance term, called a *nugget*. Kennedy and O'Hagan (2001) discuss in more detail different options of model parameterisation, choices of priors for model parameters, as well as the derivation of the joint posterior, which then serves to calibrate the model given the data. We come back to some of these properties in Sect. 5.5, devoted to model calibration.

In addition to the basic approach presented above, many extensions and generalisations have been developed as well. One such extension concerns GP meta-models with heteroskedastic covariance matrices, allowing emulator variance to differ across the parameter space. This is especially important in the presence of phase transitions in the model domain, whereby model behaviour can be different, depending on the parameter combinations. This property can be modelled for example by fitting two GPs at the same time: one for the mean, and one for the (log) variance of the output of interest. Examples of such models can be found in Kersting et al. (2007) and Hilton (2017), while the underpinning design principles are discussed in more detail in Tack et al. (2002).

Another extension concerns multidimensional outputs, where we need to look at several output variables at the same time, but cannot assume independence between them. Among the ideas that were proposed to tackle that, there are natural generalisations, such as the use of multivariate emulators, notably multivariate GPs (e.g. Fricker et al., 2013). Alternative approaches include dimensionality reduction of the output, for example through carrying out the Principal Components Analysis (PCA), producing orthogonal transformations of the initial output, or Independent Component Analysis (ICA), producing statistically independent transformations

(Boukouvalas & Cornford, 2008). One of their generalisations involves methods like Gaussian Process Latent Variable Models, which use GPs to flexibly map the latent space of orthogonal output factors onto the space of observed data (*idem*).

Given that GP emulators offer a very convenient way of describing complex models and their various features, including response surfaces, uncertainty and sensitivity, they have recently become a default approach for carrying out a meta-analysis of complex computational models. Still, the advances in machine learning and increase of computational power have led to the development of meta-modelling methods based on such algorithms, as classification and regression trees (CART), random forests, neural networks, or support vector machines (for a review, see Angione et al., 2020). Such methods can perform more efficiently than GPs in computational terms and accuracy of estimation (*idem*), although at the price of losing analytical transparency, which is an important advantage of GP emulators. In other words, there appear to be some trade-offs between different meta-models in terms of their computational and statistical efficiency on the one hand, and interpretability and transparency on the other. The choice of a meta-model for analysis in a given application needs therefore to correspond to specific research needs and constraints. Box 5.2 below continues with the example of a migration route model introduced in Chap. 4, where a GP emulator is fitted to the model inputs and outputs, with further details offered in Appendix C.

Box 5.2: Gaussian Process Emulator Construction for the Routes and Rumours Model

The design space with seven parameters of interest, described in Box 5.1 was used to train and fit a set of four GP emulators, one for each output. The emulation was done twice, assuming that the parameters are either uniformly or normally distributed. The emulators for all four output variables (*mean_freq_plan*, *stdd_link_c*, *corr_opt_links* and *prop_stdd*) additionally included code uncertainty, described by the 'nugget' variance term. The fitting was done in GEM-SA (Kennedy & Petropoulos, 2016). In terms of the quality of fit, the root mean square standardised errors (RMSSE) were found to be in the range between 1.59 for *mean_freq_plan* and 1.95 for *stdd_link_c*, based on a leave-20%-out cross-validation exercise, which, compared with the ideal value of 1, indicated a reasonable fit quality. Figure 5.5 shows an example analysis of a response surface and its error for one selected output, *mean_freq_plan*, and two inputs, *p_transfer_info* and *p_info_contacts*, based on the fitted emulator. Similar figures for the other outputs are included in Appendix C. For this piece of analysis, all the input and output variables have been standardised.

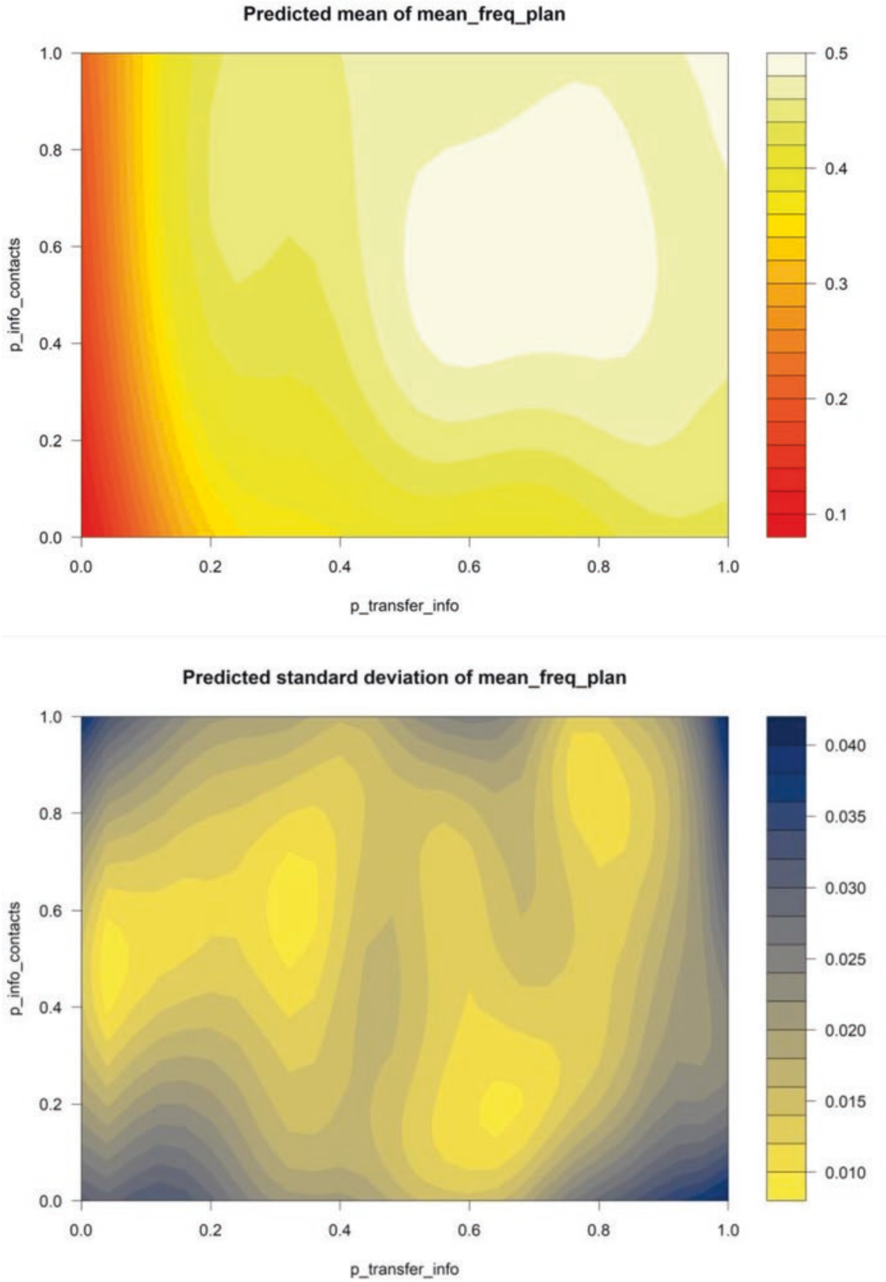


Fig. 5.5 Estimated response surface of the proportion of time the agents follow a plan vs two input parameters, probabilities of information transfer and of communication with contacts: mean proportion (top) and its standard deviation (bottom). (Source: own elaboration)

5.4 Uncertainty and Sensitivity Analysis

Once fitted, emulators can serve a range of analytical purposes. The most immediate ones consider the impact of various model inputs on the output (response). Questions concerning the uncertainty of the output and its susceptibility to the changes in inputs are common. To address these questions, *uncertainty analysis* looks at how much error gets propagated from the model inputs into the output, and *sensitivity analysis* deals with how changes in individual inputs and their different combinations affect the response variable.

Of the two types of analysis, uncertainty analysis is more straightforward, especially when it is based on a fitted emulator such as a GP (5.1), or another meta-model. Here, establishing the output uncertainty typically requires simulating from the assumed distributions for the inputs and from posterior distributions of the emulator parameters, which then get propagated into the output, allowing a Monte Carlo-type assessment of the resulting uncertainty. For simpler models, it may be also possible to derive the output uncertainty distributions analytically.

On the other hand, the sensitivity analysis involves several options, which need to be considered by the analyst to ascertain the relative influence of input variables. Specifically for agent-based models, ten Broeke et al. (2016) discussed three lines of enquiry, to which sensitivity analysis can contribute. These include insights into mechanisms generating the emergent properties of models, robustness of these insights, and quantification of the output uncertainty depending on the model inputs (ten Broeke et al., 2016: 2.1).

Sensitivity analysis can also come in many guises. Depending on the subset of the parameter space under study, one can distinguish *local* and *global* sensitivity analysis. Intuitively, the local sensitivity analysis looks at the changes of the response surfaces in the neighbourhoods of specific points in the input space, while the global analysis examines the reactions of the output across the whole space (as long as an appropriate, ideally space-filling design is selected). Furthermore, sensitivity analysis can be either *descriptive* or *variance-based*, and either *model-free* or *model-based*, the latter involving approaches based on regression and other meta-models, such as GP emulators.

The descriptive approaches to evaluating output sensitivity typically involve graphical methods: the visual assessment (‘eyeballing’) of response surface plots (such as in Fig. 5.4), correlations and scatterplots can provide first insights into the responsiveness of the output to changes in individual inputs. In addition, some of the simple descriptive methods can be also model-based, for example those using standardised regression coefficients (Saltelli et al., 2000, 2008). This approach relies on estimating a linear regression model of an output variable y based on all standardised inputs, $z_{ij} = (x_{ij} - x_i)/\sigma_i$, where x_i and σ_i are the mean and standard deviation of the ith input calculated for all design points j. Having estimated a regression model on the whole design space $\mathbf{Z} = \{(z_{ij}, y_j)\}$, we can subsequently compare the absolute values of the estimated coefficients to infer about the relative influence of their corresponding inputs on the model output.

Variance-based approaches, in turn, aim at assessing how much of the output variance is due to the variation in individual inputs and their combinations. Here again, both model-free and model-based approaches exist, which differ in terms of whether the variance decomposition is analysed directly, based on model inputs and outputs, or whether it is based on some meta-model that is fitted to the data first. As observed by Ginot et al. (2006), one of the simplest, although seldom used methods here is the analysis of variance (ANOVA), coupled with the factorial design. Here, as in the classical ANOVA approach, the overall sum of squared differences between individual outputs and their mean value can be decomposed into the sums of squares related to all individual effects (inputs), plus a residual sum of squares (Ginot et al., 2006). This approach offers a quick approximation of the relative importance of the various inputs.

The state-of-the-art approaches, however, are typically based on the decomposition of variance and on so-called Sobol' indices. Both in model-free and model-based approaches, the template for the analysis is the same. Formally, let overall output variance in a model with K inputs be denoted by $V = \mathrm{Var}[f(\mathbf{x})]$. Let us then define the *sensitivity variances* for individual inputs i and all their multi-way combinations, denoted by V_i, V_{ij}, …, $V_{12\ldots K}$. These sensitivity variances measure by how much the overall variance V would reduce if we observed particular sets of inputs, x_i, $\{x_i, x_j\}$ … $\{x_1, x_2 \ldots x_K\}$, respectively. Formally, the sensitivity variances can be defined as $V_S = V - E\{\mathrm{Var}[f(\mathbf{x})|\mathbf{x}_\mathbf{S} = \mathbf{x}^*_\mathbf{S}]\}$, where $\mathbf{S}$ denotes any non-empty set of individual inputs and their combinations. The overall variance V can then be additively decomposed into terms corresponding to the inputs and their respective combinations (e.g. Saltelli et al., 2000: 381):

$$V = \Sigma_i \; V_i + \Sigma_{i<j} \; V_{ij} + \ldots + V_{12\ldots K} \tag{5.2}$$

Based on (5.2), the sensitivity indices (or Sobol' indices) S can be calculated, which are defined as shares of individual sensitivity variances in the total V, $S_i = V_i/V$, $S_{ij} = V_{ij}/V$, …, $S_{12\ldots K} = V_{12\ldots K}/V$ (e.g. Sobol', 2001; Saltelli et al., 2008). These indices, adding up to one, have clear interpretations in terms of variance shares that can be attributed to each input and each combination of inputs.

The model-based variant of the variance-based approach is based on some meta-model fitted to the experimental data; such a meta-model can involve, for example, a Bayesian version of the GP, which was given a fully probabilistic treatment by Oakley and O'Hagan (2004). Another special case of the sensitivity analysis is *decision-based*: it looks at the effect of varying the inputs on the decision based on the output, rather than the output as such. Again, this can involve model-based approaches, which can be embedded within the Bayesian decision analysis, coupling the estimates with loss functions related to specific outputs (*idem*).

In addition to the methods for global sensitivity analysis, local methods may include evaluating partial derivatives of the output function $f(\cdot)$ – or its emulator – in the interesting areas of the parameter space (Oakley & O'Hagan, 2004). In practice, this is often done by the means of a 'one-factor-at-a-time' method, where one of the model inputs is varied, while others are kept fixed (ten Broeke et al., 2016). This approach can help identify the type and shape of one-way relationships (*idem*). In terms of a comprehensive treatment of the various aspects of sensitivity analysis, a detailed overview and discussion can be found in Saltelli et al. (2008), while a fully probabilistic treatment, involving Bayesian GP emulators, can be found in Oakley and O'Hagan (2004). In the context of agent-based models, ten Broeke et al. (2016) have provided additional discussion and interpretations, while applications to demographic simulations can be found for example in Bijak et al. (2013) and Silverman et al. (2013).

To illustrate some of the key concepts, the example of the model of migration routes is continued in Box 5.3 (with further details in Appendix C). This example summarises results of the uncertainty and global variance-based sensitivity analysis, based on the fitted GP emulators.

Box 5.3: Uncertainty and Sensitivity of the Routes and Rumours Model

In terms of the uncertainty of the emulators presented in Box 5.2, the fitted variance of the GPs for standardised outputs, representing the uncertainty induced by the input variables and the intrinsic randomness (nugget) of the stochastic model code, ranged from 1.14 for *mean_freq_plan*, to 1.50 for *stdd_link_c*, to 1.65 *corr_opt_links*. The nugget terms were respectively equal 0.009, 0.020 and 0.019. For the cross-replicate output variable, *prop_stdd*, the variances were visibly higher, with 4.15 overall and 0.23 attributed to the code error.

As for the sensitivity analysis, for all four outputs the parameters related to information exchange proved most relevant, especially the probability of exchanging information through communication, as well as the information error – a finding that was largely independent of the priors assumed for the parameters (Fig. 5.6). In neither case did parameters related to exploration matter much.

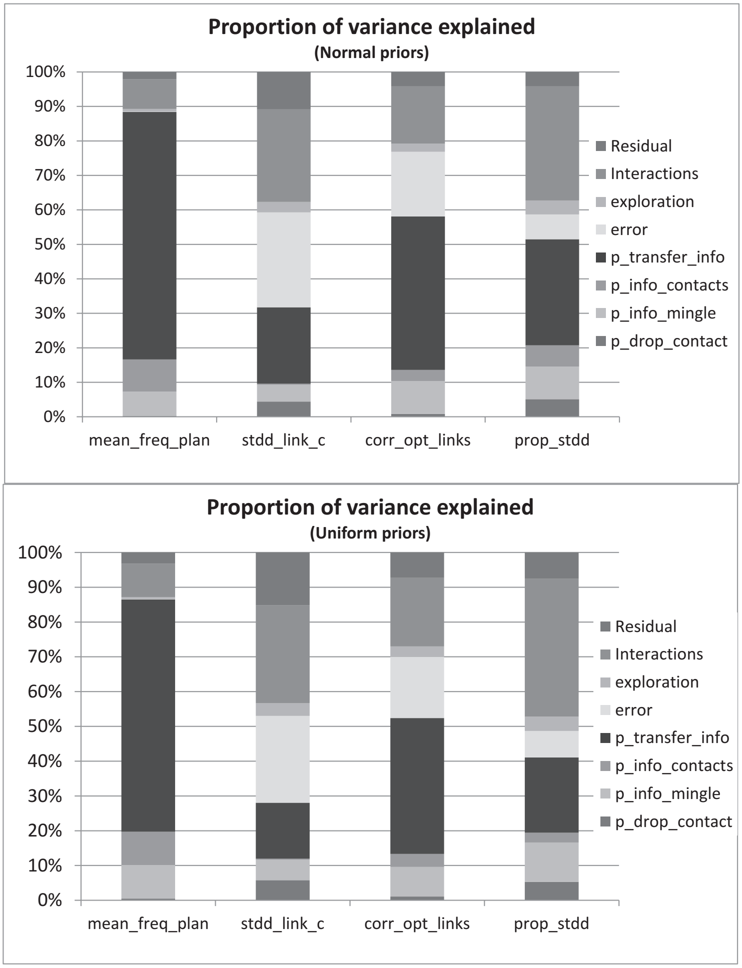


Fig. 5.6 Variance-based sensitivity analysis: variance proportions associated with individual variables and their interactions, under different priors. (Source: own elaboration)

5.5 Bayesian Methods for Model Calibration

Emulators, such as the GPs introduced in Sect. 5.3, can serve as tools for calibrating the underlying complex models. There are many ways in which this objective can be achieved. Given that the emulators can be built and fitted by using Bayesian methods, a natural option for calibration is to utilise full Bayesian inference about the distributions of inputs and outputs based on data (Kennedy & O'Hagan 2001; Oakley & O'Hagan, 2002; MUCM, 2021). Specifically in the context of agent-based models, various statistical methods and aspects of model analysis are also reviewed in Banks and Norton (2014) and Heard et al. (2015).

The fully Bayesian approach proposed by Kennedy and O'Hagan (2001) focuses on learning about the calibration parameters **θ** of the model or, for complex models, its emulator, based on data. Such parameters are given prior assumptions, which are subsequently updated based on observed data to yield calibrated posterior distributions. However, as mentioned in Sect. 5.3, even at the calibrated values of the input parameters, model discrepancy – a difference between the model outcomes and observations – remains, and needs to be formally acknowledged too. Hence, the general version of the calibration model for the underlying computational model (or meta-model) f based on the training sample **x** and the corresponding observed data $z(\mathbf{x})$, has the following form (Kennedy & O'Hagan, 2001: 435; notation after Hilton, 2017):

$$z(\mathbf{x}) = \rho f(\mathbf{x}, \theta) + \delta(\mathbf{x}) + \varepsilon(\mathbf{x}). \tag{5.3}$$

In this model, $\delta(\mathbf{x})$ represents the discrepancy term, $\varepsilon(\mathbf{x})$ is the residual observation error, and ρ is the scaling constant. GPs are the conventional choices of priors both for $f(\mathbf{x}, \boldsymbol{\theta})$ and $\delta(\mathbf{x})$. For the latter term, the informative priors for the relevant parameters typically need to be elicited from domain experts in a subjective Bayesian fashion, to avoid problems with the non-identifiability of both GPs (*idem*).

The calibrated model (5.3) can be subsequently used for prediction, and also for carrying out additional uncertainty and sensitivity checks, as described before. Existing applications to agent-based models of demographic or other social processes are scarce, with the notable exception of the analysis of a demographic micro-simulation model of population dynamics in the United Kingdom, presented by Hilton (2017), and, more recently, an analysis of ecological demographic models, as well as epidemiological 'compartment' models discussed by Hooten et al. (2021).

Emulator-based and other more involved statistical approaches are especially applicable wherever the models are too complex and their parameter spaces have too many dimensions to be treated, for example, by using simple Monte Carlo algorithms. In such cases, besides GPs or other similar emulators, several other approaches can be used as alternative or complementary to the fully Bayesian inference. We briefly discuss these next. Detailed explanations of these methods are beyond the scope of this chapter, but can be explored further in the references (see also Hooten et al., 2020 for a high-level overview, with a slightly different emphasis).

- **Approximate Bayesian Computation (ABC).** This method relies on sampling from the prior distributions for the parameters of a complex model, comparing the resulting model outputs with actual data, and rejecting those samples for which the difference between the outputs and the data exceeds a pre-defined threshold. As the method does not involve evaluating the likelihood function, it can be computationally less costly than alternative approaches, although it can very quickly become inefficient in many-dimensional parameter spaces. The theory underpinning this approach dates to Tavaré et al. (1997), with more recent overviews offered in Marin et al. (2012) and Sisson et al. (2018). Applications to calibrating agent-based models in the ecological context were discussed by van der Vaart et al. (2015).
- **Bayes linear methods, and history matching.** In this approach, the emulator is specified in terms of the two first moments (mean and covariance function) of the

output function, and a simplified (linear) Bayesian updating is used to derive the expected posterior moments given the model inputs and outputs from the training sample, under the squared error loss (Vernon et al., 2010). Once built, the emulator is fitted to the observed empirical data by comparing them with the model outputs by using measures of *implausibility*, in an iterative process known as history *matching* (*idem*). For many practical applications, especially those involving highly-dimensional parameter spaces, the history matching approach is computationally more efficient than the fully Bayesian approach of Kennedy and O'Hagan (2001), although at the expense of providing an approximate solution (for more detailed arguments, see e.g. the discussion of Vernon et al., 2010, or Hilton, 2017). Examples of applying these methods to agent-based approaches include a model of HIV epidemics by Andrianakis et al. (2015), as well as models of a demographic simulation and fertility developments in response to labour market changes (the so-called Easterlin effect) by Hilton (2017).

- **Bayesian melding.** This approach 'melds' two types of prior distributions for the model output variable: 'pre-model', set for individual model inputs and parameters and propagated into the output, and 'post-model', set directly at the level of the output. The two resulting prior distributions for the output are weighted (linearly or logarithmically) by being assigned weights a and ($1–a$), respectively, and the posterior distribution is calculated based on such a weighted prior. The underpinning theory was proposed by Raftery et al. (1995) and Poole and Raftery (2000). In a recent extension, Yang and Gua (2019) proposed treating the pooling parameter a as another hyper-parameter of the model, which is also subject to estimation through the means of Bayesian inference. An example of an application of Bayesian melding to an agent-based modelling of transportation can be found in Ševčíková et al. (2007).
- **Polynomial chaos.** This method, originally stemming from applied mathematics (see O'Hagan, 2013), uses polynomial approximations to model the mapping between model inputs and outputs. In other words, the output is modelled as a function of inputs by using a series of polynomials with individual and mixed terms, up to a specified degree. The method was explained in more detail from the point of view of uncertainty quantification in O'Hagan (2013), where it was also compared with GP-based emulators. The conclusion of the comparison was that, albeit computationally promising, polynomial chaos does not (yet) account for all different sources of uncertainty, which calls for closer communication between the applied mathematics and statistics/uncertainty quantification communities. A relevant example, using polynomial chaos in an agent-based model of a fire evacuation, was offered by Xie et al. (2014).
- **Recursive Bayesian approach.** This method, designed by Hooten et al. (2019, 2020), aims to make full use of the natural Bayesian mechanism for sequential updating in the context of time series or similar processes, whereby the posterior distributions of the parameters of interest are updated one observation at a time. The approach relies on a recursive partition of the posterior for the whole series into a sequence of sub-series of different lengths (Hooten et al. 2020), which can be computed iteratively. The computational details and the choice of appropriate sampling algorithms were discussed in more detail in Hooten et al. (2019).

We conclude this chapter by providing an example of calibrating the migration route formation model, which is presented in Box 5.4.

Box 5.4: Calibration of the Routes and Rumours Model

In order to demonstrate the use of calibration techniques, a set of representative values from the previous set of experimental samples was treated as 'observed data' against which to calibrate. Principal components were taken from a normalised matrix of samples of the output variables *mean_freq_plan, corr_opt_links,* and *stdd_link_c* to transform to a set of orthogonal coordinates. The variable *prop_stdd* was not used because it refers to summaries of repeated simulations; these cannot even theoretically be observed, as they would correspond outcomes from many different possible histories. Following Higdon (2008), separate GP emulators were then fitted to the mean of the principal component scores at each design point, with the variation over repetitions added as a variance term that is allowed to vary over the design. The DiceKriging R package was used to fit all emulators (Roustant et al., 2012), and k-fold cross validation indicated that the emulators captured the variation in the simulator reasonably well. A simplified but multivariate version of the model discussed in Sect. 5.3 was employed for the purposes of calibration, with ρ set to 1 and with the discrepancy and observation error terms assumed to independently and identically (normally) distributed. Posterior distributions for the unknown calibration parameters $\boldsymbol{\theta}$ were obtained from this model using the `stan` Bayesian modelling package (Stan Development Team, 2021). Non-informative Beta(1,1) priors were used for the calibration parameters.

Figure 5.7 shows the resultant calibrated posterior distributions. As the sensitivity analysis showed, *p_transfer_info* has the greatest effect on simulator outputs, and therefore we gain more information about this parameter during the calibration process, while the posteriors indicate that a wide range of values of other parameters could replicate the observed values, given our uncertainty about the simulator and about reality, and taking into account the stochasticity of the simulator itself. Still, the wide uncertainty in the posterior distributions for the most parameter values is not surprising: it reflects the high uncertainty of the process itself. In a general case, such high residual errors remaining after calibration could illuminate the areas where the uncertainty might be either irreducible (aleatory), or at least difficult to reduce given the available set of calibration data that was used for that purpose.

Figure 5.8 shows that the resulting calibrated predicted emulator outputs are close to the target values (red dotted lines). This means that running the simulator on samples from the calibrated posterior of the input parameters is expected to produce a multivariate distribution of output values centred on our observed values.

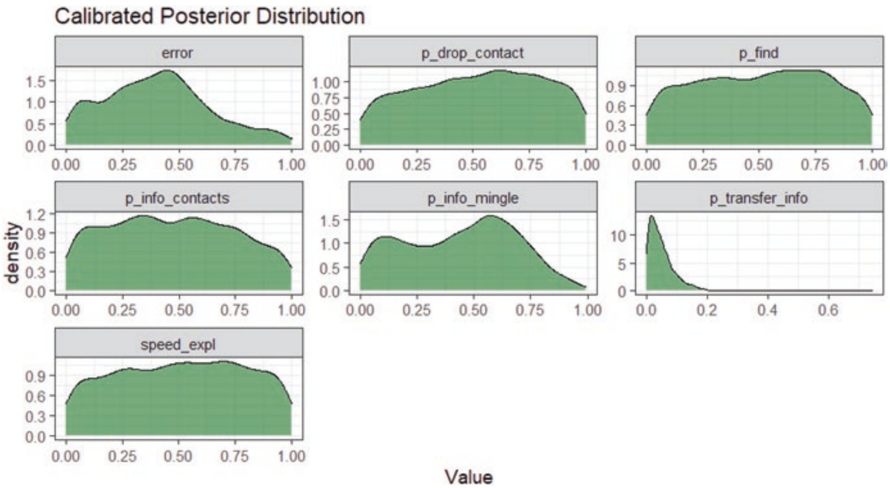


Fig. 5.7 Calibrated posterior distributions for Routes and Rumours model parameters

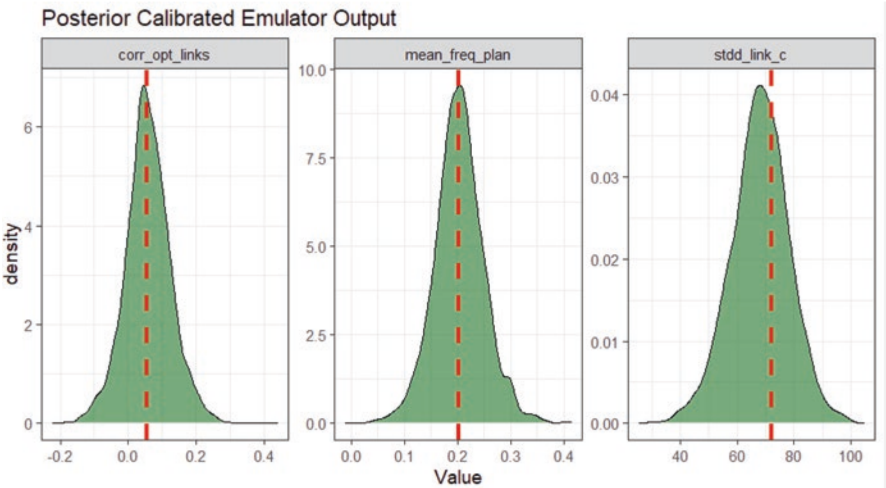


Fig. 5.8 Posterior calibrated emulator output distributions

Chapter 6
The Boundaries of Cognition and Decision Making

Toby Prike, Philip A. Higham, and Jakub Bijak

This chapter outlines the role that individual-level empirical evidence gathered from psychological experiments and surveys can play in informing agent-based models, and the model-based approach more broadly. To begin with, we provide an overview of the way that this empirical evidence can be used to inform agent-based models. Additionally, we provide three detailed exemplars that outline the development and implementation of experiments conducted to inform an agent-based model of asylum migration, as well as how such data can be used. There is also an extended discussion of important considerations and potential limitations when conducting laboratory or online experiments and surveys, followed by a brief introduction to exciting new developments in experimental methodology, such as gamification and virtual reality, that have the potential to address some of these limitations and open the door to promising and potentially very fruitful new avenues of research.

6.1 The Role of Individual-Level Empirical Evidence in Agent-Based Models

Agents are the key feature that distinguish agent-based models from other forms of micro-simulation. Specifically, within agent-based models, agents can interact with one another in dynamic and non-deterministic ways, allowing macro-level patterns and properties to emerge from the micro-level characteristics and interactions within the model. This key feature of agent-based models means that insights into individual behaviour from psychology and behavioural economics, such as behaviours, personalities, judgements, and decisions, are even more crucial than for other modelling efforts. Within this chapter, we provide an outline as to why it is important to incorporate insights from the study of human behaviour within agent-based models, and give examples of the processes that can be used to do this. As in other chapters within this book, agent-based models of migration are used as an exemplar,

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_6

however, the information and processes described are applicable to a wide swathe of agent-based models.

Traditionally, many modelling efforts, including agent-based models of demographic processes, have relied on normative models of behaviour, such as expected utility theory, and have assumed that agents behave rationally. However, descriptive models of behaviour, commonly used within psychology and behavioural economics, provide an alternative approach with a focus on behaviour, judgements, and decisions observed using experimental and observational methods. There are many important trade-offs to consider when deciding which approaches to use for an agent-based model and which level of specificity or detail to use. For example, normative models may be more likely to be tractable and already formalised, which gives some key advantages (Jager, 2017). In contrast, many social scientific theories based on observations from areas such as psychology, sociology, and political science may provide much more detailed and nuanced descriptions of how people behave, but are also more likely to be specified using verbal language that is not easily formalised. Therefore, to convert these social science theories from verbal descriptions of empirical results into a form that can be formalised within an agent-based model requires the modeller to make assumptions (Sawyer, 2004). For example, there may be a clear empirical relationship between two variables but the specific causal mechanism that underlies this relationship may not be well established or formalised (Jager, 2017). Similarly, there may be additional variables within an agent-based model that were not incorporated in the initial theory or included in the empirical data. In situations such as these, it often falls to the individual modeller(s) to make assumptions about how to formalise the theory, provide formalised causal mechanisms, and extend the theory to incorporate any additional variables and their potential interactions and impacts.

When it comes to agent-based models of migration, the extent to which empirical insights from the social sciences are used to add complexity and depth to the agents varies greatly (e.g., see Klabunde & Willekens, 2016 for a review of decision making in agent-based models of migration). Additionally, because migration is a complex process that has wide-ranging impacts, there are many options and areas in which additional psychological realism can be added to agent-based models. For example, the personality of the agent is likely to play a role and may be incorporated through giving each agent a propensity for risk taking. Previous research has shown that increased tolerance to risk is associated with a greater propensity to migrate (Akgüç et al., 2016; Dustmann et al., 2017; Gibson & McKenzie, 2011; Jaeger et al., 2010; Williams & Baláž, 2014), and therefore incorporating this psychological aspect within an agent-based model may allow for unique insights to be drawn (e.g., how different levels of heterogeneity in risk tolerance influence the patterns formed, or whether risk tolerance matters more in some migration contexts than others). Additionally, the influence of social networks on migration has been well established (Haug, 2008) so this is also a key area where there may be benefits to adding realism to an agent-based model (Klabunde & Willekens, 2016; Gray et al., 2017). A review of existing models and empirical studies of decision making in the context of migration is offered by Czaika et al. (2021).

When it is believed that an agent-based model can be improved through incorporating additional realism or descriptive insights, designing and implementing an experiment or survey can be a very useful way to gain data, information, and insights. However, there are several different approaches that can be used to derive insights from the social sciences and other empirical literature to inform agent-based models before taking the step of engaging in primary data collection. The first, and most straightforward approach, is to examine the existing literature to see which insights can be gleaned and how people have previously attempted to address the same or similar issues (e.g., if the modeller wants to incorporate emotion or personality into an agent-based model, there are existing formalisms that may be appropriate for use in such instances; Bourgais et al., 2020).

Even if there are no agent-based or other models that have previously addressed the specific research issues or concerns in terms of formalising and incorporating the same descriptive aspect, there may still be pre-existing data that can be used to answer any specific questions that may arise or additional realism that could be incorporated. However, in this situation the modeller will still have to take the additional difficult steps of extracting the information from the existing data or theory (likely a verbal theory) and formalising it for inclusion within an agent-based model. Finally, if it emerges that there are neither pre-existing implementations within a model nor an existing formalism, and there are no verbal theories or relevant data that can be used to build formalisms for inclusion, then it may be time to engage in dedicated primary data collection, and design an experiment and/or survey of the modeller's own design (see also Gray et al., 2017).

When designing a survey or experiment, it is important to keep in mind the specific goal of the data collection. For example, in terms of agent-based modelling, the goal may be to use the data to inform parameters within the model, or it may be to compare and contrast several different decision rules to decide which has the strongest empirical grounding to include within the model. In the following sections, we outline several experiments that were conducted to better inform agent-based models of asylum migration. The descriptions we provide serve as exemplars, and include an outline of the development of key questions for each experiment, a brief overview of how each experiment was implemented and the methodologies used for the experiments, and finally a discussion of how the data collected in each experiment can be used to inform an agent-based model of migration.

6.2 Prospect Theory and Discrete Choice

The first set of psychological experiments conducted to better inform agent-based models of migration focused on discrete choice within a migration context. Traditionally, most agent-based models of migration have used expected utility and/or made other assumptions of rationality when building their models (see also the description of neoclassical theories of migration, summarised in Massey et al., 1993). That is, they make assumptions that agents within the models will behave in the way that they ‘should’ behave based on normative models of optimal behaviour. However, research within psychology and behavioural economics has called many of these assumptions into question. The most famous example of this is prospect theory, developed by Kahneman and Tversky (1979) and subsequently updated to become cumulative prospect theory (Tversky & Kahneman, 1992). Based on empirical data, prospect theory proposes that people deviate from the optimal or rational approaches because of biases in the way that they translate information from the objective real-world situation to their subjective internal representations of the world. This has clear implications for how people subsequently make judgements and decisions. Some of the specific empirical findings related to judgement and decision making that are incorporated within prospect theory include loss aversion, overweighting/underweighting of probabilities, differential responses to risk (risk seeking for losses and risk aversion for gains), and framing effects.

Prospect theory was also a useful first area in which to conduct experiments to inform agent-based models of migration because, unlike many other theories of judgement and decision making based on empirical findings, it is already formalised and can therefore be implemented more easily within models. Indeed, in previous work, de Castro et al. (2016) applied prospect theory to agent-based models of financial markets, contrasting these models with agent-based models in which agents behaved according to expected utility theory. De Castro et al. (2016) found that simulations in which agent behaviour was based on prospect theory were a better match to real historical market data than when agent behaviour was based on expected utility theory. Although the bulk of research on prospect theory has focused on financial contexts (for reviews see Barberis, 2013; Wakker, 2010), there is also growing experimental evidence that prospect theory is applicable to other contexts. For example, support for the theory has been found when outcomes of risky decisions are measured in time (Abdellaoui & Kemel, 2014) or related to health such as the number of lives saved (Kemel & Paraschiv, 2018), life years (Attema et al., 2013), and quality of life (Attema et al., 2016).

Czaika (2014) applied prospect theory to migration patterns at a macro-level, finding that the patterns of intra-European migration into Germany were consistent with several aspects of prospect theory, such as reference dependence, loss aversion, and diminished sensitivity. However, because this analysis did not collect micro-level data from individual migrants, it is necessary to assume that the macro-level patterns observed occur (at least partially) due to individual migrants behaving in a way that is consistent with prospect theory. This is a very strong assumption, which risks falling into the trap of the ecological fallacy. At the same time, however, there are also a variety of studies that have examined risk preferences of both economic migrants (Akgüç et al., 2016; Jaeger et al., 2010) and migrants seeking asylum (Ceriani & Verme, 2018; Mironova et al., 2019), and can therefore provide data about some individual level behaviour, judgments and decisions to inform agent-based models of migration. Bocquého et al. (2018) extended this line of research further, using the parametric method of Tanaka et al. (2010) to elicit utility functions from asylum seekers in Luxembourg, finding that the data supported prospect theory over expected utility theory. However, these previous studies examining risk and the application of prospect theory to migration still used standard financial tasks, rather than collecting data within a migration context specifically.

Based on the broad base of existing empirical support, we decided to apply prospect theory to our agent-based models of migration and therefore designed a dedicated experiment to elicit prospect theory parameters within a migration context. There are a variety of potential approaches that can be used to elicit prospect theory parameters (potential issues due to divergent experimental approaches are discussed in Sect. 6.4). To avoid making *a priori* assumptions about the shape of the utility function, we chose to use a non-parametric methodology adapted from Abdellaoui et al. (2016; methodology presented in Table 6.1). Participants made a series of choices between two gambles within a financial and a migration context. For each choice, both gambles presented a potential gain or loss in monthly income (50% chance of gaining and 50% chance of losing income; see Fig. 6.1 for an example trial). Using this methodology, we elicited six points of the utility function for gains and six points for losses. We then analysed the elicited utility functions for financial and migration decisions to test for loss aversion, whether there was evidence of concavity for gains and/or convexity for losses, and whether there were differences between the migration and financial contexts (see Appendix D for more details on the preregistration of the hypotheses, sample sizes, and ethical issues).

There are many ways that the results from these experiments can be used to inform agent-based models of migration. The first and perhaps simplest way is to add loss aversion to the model. Because the data collected were within the context of relative changes in gains and losses for potential destination countries, these results can be used within the model to create a distribution of population level loss aversion, from which each agent is assigned an individual level of loss aversion (to allow for variation across agents). Therefore, rather than making assumptions about the extent of loss aversion present within a migration context, instead, each agent within the model would weight potential losses more heavily than potential gains, following the empirical findings from the experiment in a migration context. Similarly, after fitting a function to the elicited points for gains and losses, it is possible to again use this information to inform the shape of the utility functions that are given to agents within the model. That is, the data can be used to inform the extent to which agents place less weight on potential gains and losses as they get further from the reference point (usually implemented as either the current status quo or the currently expected outcome). For example, the empirical data inform us whether people consider a gain of $200 in income to be twice as good as a gain of $100, or only one and a half times as good when they are making a decision.

An additional advantage of including the financial context within the same experiment is that it allows for direct comparisons between that context and a migration context. Therefore, because there is a wide body of existing research on decision making within financial contexts, if the results are similar across conditions then that may provide some supporting evidence that this body of research can be relied on when applied to migration contexts. Conversely, if the results reveal that

Table 6.1 Procedure for eliciting utility functions

Step	Elicitation equation	Value elicited	Prespecified values
1	$G_{(p)}L \sim x_0$	L	All stakes: $x_0 = 0$, $p = 0.5$ Small stakes: $G = 250$, $l = 50$, $g = 50$ Medium stakes: $G = 500$, $l = 100$, $g = 100$ Large stakes: $G = 1000$, $l = 200$, $g = 200$
2	$x_1^+ \sim G_{(p)}x_0$	x_1^+	
3	$x_1^- \sim L_{(p)}x_0$	x_1^-	
4	$x_{1(p)}^+\mathcal{L} \sim x_{0(p)}l$	$\mathcal{L}$	
5	$x_{2(p)}^+\mathcal{L} \sim x_{1(p)}^+l$	x_2^+	
6	$x_{3(p)}^+\mathcal{L} \sim x_{2(p)}^+l$	x_3^+	
7	$x_{4(p)}^+\mathcal{L} \sim x_{3(p)}^+l$	x_4^+	
8	$x_{5(p)}^+\mathcal{L} \sim x_{4(p)}^+l$	x_5^+	
9	$x_{6(p)}^+\mathcal{L} \sim x_{5(p)}^+l$	x_6^+	
10	$\mathcal{G}_{(p)}\, x_1^- \sim g_{(p)}x_0$	$\mathcal{G}$	
11	$\mathcal{G}_{(p)}\, x_2^- \sim g_{(p)}x_1^-$	x_2^-	
12	$\mathcal{G}_{(p)}\, x_3^- \sim g_{(p)}x_2^-$	x_3^-	
13	$\mathcal{G}_{(p)}\, x_4^- \sim g_{(p)}x_3^-$	x_4^-	
14	$\mathcal{G}_{(p)}\, x_5^- \sim g_{(p)}x_4^-$	x_5^-	
15	$\mathcal{G}_{(p)}\, x_6^- \sim g_{(p)}x_5^-$	x_6^-	

Notes: elicitation procedure taken from Abdellaoui et al. (2016) with some prespecified values altered. The step column shows the order in which values are elicited from participants. The elicitation equation shows the structure used for each elicitation. The value elicited column shows the value that is being elicited at that step. Elicited values were initially set so that both gambles had equivalent utility. The prespecified values column shows the values within the elicitation equations that are prespecified rather than being elicited. The size of the prespecified values were chosen to be approximately equidistant in terms of utility rather than in terms of raw values. Therefore, there is a larger gap between the medium and large stakes than between the medium and small stakes to account for diminishing sensitivity for values further from the reference point. x_0 = reference point, x_1^+ through x_6^+ = the six points of the utility function elicited for gains, x_1^- through x_6^- = the six points of the utility function elicited for losses, p = probability of outcomes, G = a prespecified (large) gain, L = an elicited loss equivalent to G in terms of utility, l = a prespecified loss, L = an elicited loss, g = a prespecified (small) gain, $\mathcal{G}$ = an elicited gain. The tilde (~) denotes approximate equivalence or indifference between the two alternative options

there are differences between the contexts, then it highlights that modellers should show caution when applying financial insights to other contexts. The presence of differences between contexts would highlight the need to collect additional data within the specific context of interest, rather than relying on assumptions, formalisations, or parameter estimates developed in a different context.

Fig. 6.1 An example of the second gain elicitation (x_2^+) within a migration context and with medium stakes. As shown in panel A, x_2^+ is initially set so that both gambles have equivalent utility. The value of x_2^+ is then adjusted in panels B to F depending on the choices made, eliciting the value of x_2^+ that leads to indifference between the two gambles. (Source: own elaboration in Qualtrics)

6.3 Eliciting Subjective Probabilities

The key questions for the second set of psychological experiments emerged from the initial agent-based models presented in Chap. 3 and analysed in Chap. 5. These models highlighted the important role that information sharing and communication between agents can play in influencing the formation and reinforcement of migration routes. Because these aspects played a key role in influencing the results produced by the models, (as indicated by the preliminary sensitivity analysis of the influence of the individual model inputs on a range of outputs, see Chap. 5), it became clear that we needed to gather more information about the processes involved to ensure the model was empirically grounded.

To achieve these aims, we designed a psychological experiment with these specific questions in mind so that the data could be used to inform parameters for the model. Prior to implementing the experiment, we reviewed the relevant literature across domains such as psychology, marketing, and communications to examine what empirical data existed as well as which factors had previously been shown to be relevant. Throughout this process, we kept the specific case study of asylum seeker migration in mind, giving direction and focus to the search and review of the literature. This process led us to focus in on two key factors that were directly relevant to the agent-based model and had also previously been examined within the empirical literature: the source of the information and how people interpret verbal descriptors of likelihood or probability.

Regarding the source of the information, we chose to focus on three specific aspects of source that existing research had shown to be particularly influential: expertise, trust, and social connectedness. Research into the role of source expertise had shown that people are generally more willing to change their views and update their beliefs when the source presenting the information has relevant expertise (Chaiken & Maheswaran, 1994; Hovland & Weiss, 1951; Maddux & Rogers, 1980; Petty et al., 1981; Pilditch et al., 2020; Pornpitakpan, 2004; Tobin & Raymundo, 2009). Trust in a source has also been shown to be a key factor in the interpretation of information and updating of beliefs, with people more strongly influenced by sources in which they place a higher degree of trust (Hahn et al., 2009; Harris et al., 2016; McGinnies & Ward, 1980; Pilditch et al., 2020; Pornpitakpan, 2004). Finally, social connectedness has been found to be an important source characteristic, with people more strongly influenced by sources with whom they have greater social connectedness. For example, people are more influenced by sources that are members of the same racial or religious group and/or sources with whom they have an existing friendship or have worked with collaboratively (Clark & Maass, 1988; Feldman, 1984; Sechrist & Milford-Szafran, 2011; Sechrist & Young, 2011; Suhay, 2015).

The other key aspect was the role of verbal descriptions of likelihood and how people interpret and convert these verbal descriptors into a numerical representation (Budescu et al., 2014; Mauboussin & Mauboussin, 2018; Wintle et al., 2019). This was of particular relevance for the agent-based model of migration because it directly addresses the challenge of converting information from a more fuzzy, verbal description into a numerical response that is easily formalised and can be included within a model. Examining verbal descriptions of likelihood allowed us to address questions such as 'when someone says that it is likely to be safe to make a migration journey, how should that be numerically quantified' which is a key step for formalising these processes within the agent-based model.

Having established the areas of focus through an iterative process of generating questions via the agent-based model and reviewing existing literature, it was then possible to design an experiment that provides empirical results to inform the model, and also has the potential to contribute to the scientific literature more broadly by addressing gaps within the literature. We were able to do this by selecting sources that were relevant for asylum seeker migration and also varied on the key source characteristics of expertise, trust, and social connectedness. These choices were also informed by previous research conducted in the Flight 2.0/Flucht 2.0 research project on the media sources used by asylum seekers before, during, and after their journeys from their country of origin to Germany (Emmer et al., 2016; see also Chap. 4 and Appendix B). The specific sources that were chosen for inclusion in the experiment were: a news article, a family member, an official organisation, someone with relevant personal experience, and the travel organiser (i.e., the person organising the boat trip). Additionally, we randomised the verbal likelihood that was communicated by each source to be one of the following: *very likely, likely, unlikely, or very unlikely* (one verbal likelihood presented per source). For example, a participant may read that a family member says a migration boat journey across the sea is likely to be safe, that an official organisation says the trip is unlikely to be safe, that someone with relevant personal experience says it is very unlikely to be safe, and so on (see Fig. 6.2 for an example).

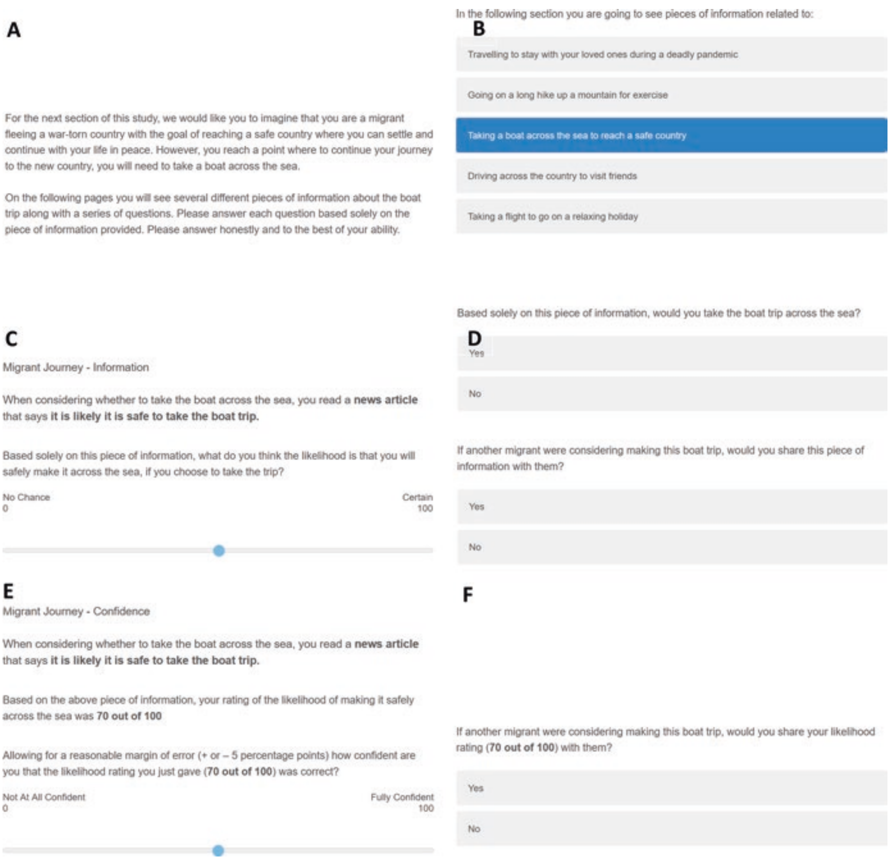


Fig. 6.2 Vignette for the migration context (panel A), followed by the screening question to ensure participants paid attention (panel B) and an example of the elicitation exercise, in which participants answer questions based on information from a news article (panels C to F). (Source: own elaboration in Qualtrics)

After seeing each piece of information, participants judged the likelihood of travelling safely (0–100) and made a binary decision to travel (yes/no). Additionally, they indicated how confident they were in their likelihood judgement, and whether they would share the information and their likelihood judgement with another traveller. Participants also made overall judgements of the likelihood of travelling safely and hypothetical travel decisions based on all the pieces of information, and indicated their confidence in their overall likelihood judgement, and whether they would share their overall likelihood judgement. At the end of the experiment, participants indicated how much they trusted the five sources in general, as well as whether they had ever seriously considered or made plans to migrate to a new country, and whether they had previously migrated to a new country (again, see Appendix D for details on the preregistration, sample sizes, and ethical issues).

Conducting this experiment provided a rich array of data that can be used to inform an agent-based model of asylum seeker migration. For example, it becomes relatively straightforward to assign numerical judgements about safety to information that agents receive within an agent-based model because data has been collected on how people (experiment participants) interpret phrases such as ‘the boat journey across the sea is likely to be safe’. It is also possible to see whether these interpretations vary depending on the source of the information, such as whether ‘likely to be safe’ should be interpreted differently by an agent within the model depending on whether the information comes from a family member or an official organisation. Additionally, because we collected overall ratings it is possible to examine how people combine and integrate information from multiple sources to form overall judgements. This information can be used within an agent-based model to assign relative weights to different information sources, such as weighting an official organisation as 50% more influential than a news article, a family member as 30% less influential than someone with relevant personal experience, and so on.

To more explicitly illustrate this, the data collected in this experiment were used to inform the model presented in Chap. 8. Specifically, because for each piece of information participants received they provided both a numerical likelihood of safety rating and a binary yes/no decision regarding whether they would travel, it was possible to calculate the decision threshold at which people become willing to travel, as well as how changes in the likelihood of safety ratings influence the probability that someone will decide to travel. We could then use these results to inform parameters within the model that specify how changes in an agent’s internal representation of the safety of travelling translate into changes in the probability of them making specific travel decisions.

6.4 Conjoint Analysis of Migration Drivers

In the third round of experiments, *conjoint analysis* is used to elicit the relative weightings of a variety of migration drivers. Specifically, the focus is on characteristics of potential destination countries and analysing which of these characteristics have the strongest influence on people's choices between destinations. The impetus for this experimental focus again came from some key questions within both the model and the migration literature more broadly. In relation to the model, this line of experimental inquiry arose because the model uses a graphical representation of space that the agents attempt to migrate across towards several potential end cities (end points), with numerous paths and cities present along the way.

In the initial implementations of the Routes and Rumours model, there was no differentiation between the available end points. That is, the agents within the model simply wanted to reach any of the available end cities/points and did not have any preference for some specific end cities over others. This modelling implementation choice was made to get the model operational and to provide results regarding the importance of communication between agents and agent exploration of the paths/cities. However, to enhance the realism of the agent-based model and make it more directly applicable to the real-world scenarios that we would like to model, it became clear that it was important for the end cities to vary in their characteristics and the extent to which agents desire to reach them. Therefore, it was important to gather empirical data about the characteristics of potential end destinations for migration as well as how people weight the different characteristics of these destinations and make trade-offs when choosing to migrate.

Previous research has examined the various factors that influence the desirability of migration destination countries (Carling & Collins, 2018). Recently, a taxonomy of migration drivers has been developed, made up of nine dimensions of drivers and 24 individual driving factors that fit within these nine dimensions (Czaika & Reinprecht, 2020). The nine dimensions identified were: demographic, economic, environmental, human development, individual, politico-institutional, security, socio-cultural, and supra-national. The breadth of areas covered by these dimensions helps to emphasise the large array of characteristics that may influence the choices migrants make about the destination countries of interest.

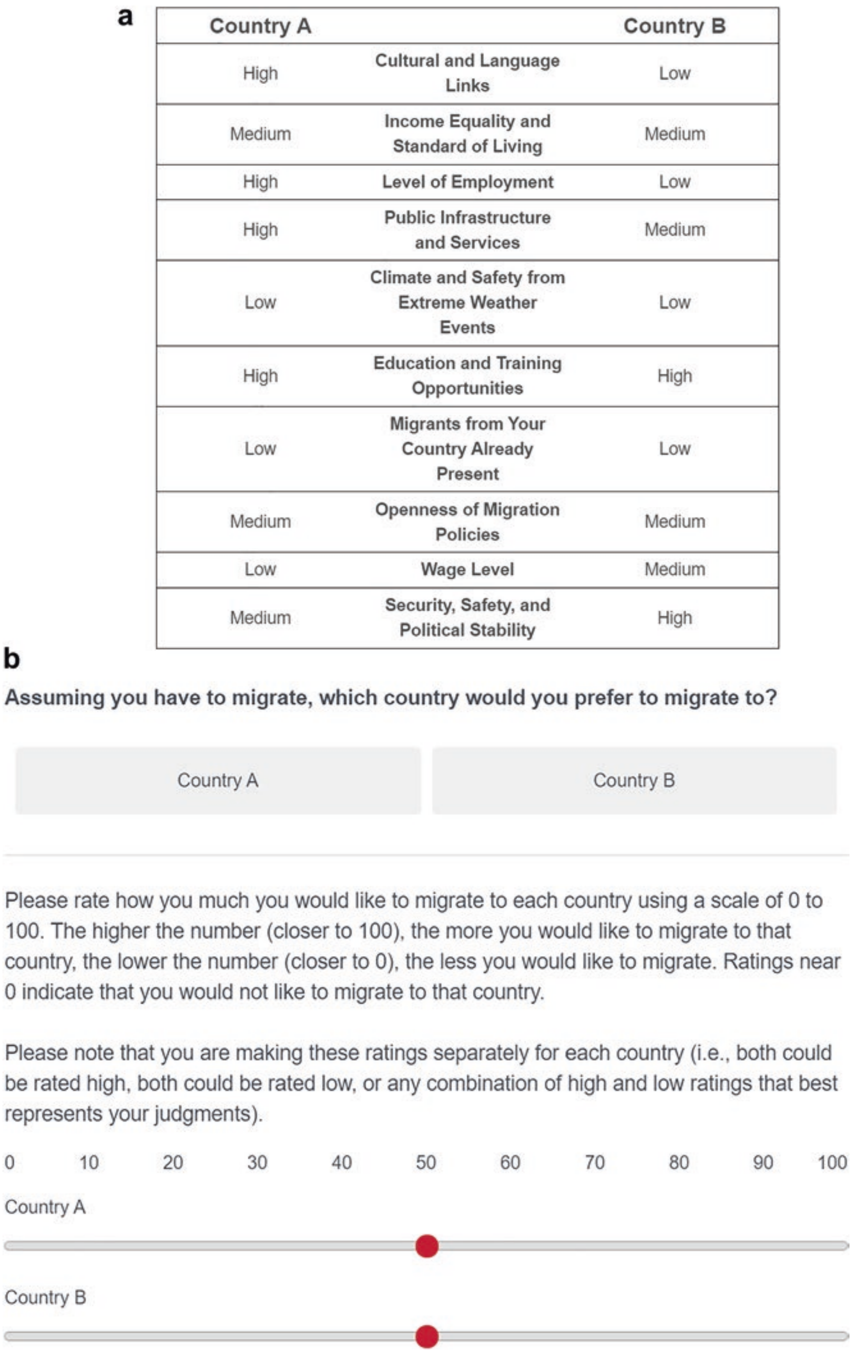
Research using an experimental approach has also previously been used to examine the importance of a variety of migration drivers, in Baláž et al. (2016) and Baláž and Williams (2018). Both these studies examined how participants searched for information related to wages, living costs, climate, crime rate, life satisfaction, health, freedom and security, and similarity of language (Baláž et al., 2016), as well as the unemployment rate, attitudes towards immigrants, and whether a permit is needed to work in the country (Baláž & Williams, 2018). Additionally, in both studies participants were asked about their previous experience with migration so that results could be compared between migrants and non-migrants. The results of these studies showed that, consistent with many existing neo-classical approaches to migration studies (Borjas, 1989; Harris & Todaro, 1970; Sjaastad, 1962; Todaro, 1969), participants were most likely to request information on economic factors and also weighted these factors the most strongly in their decisions. Specifically, wages and cost of living were the most requested pieces of information and had the highest decision weights. However, they also found that participants with previous migration experience placed more emphasis on non-economic factors, being more likely to request information about life satisfaction and to give more weight to life satisfaction when making their decisions. This suggests that non-economic factors can also play an important role in migration, and that experience of migration may make people more likely to consider and place emphasis on these non-economic factors.

Building on the questions derived from the agent-based model and this previous literature, we decided to conduct an experiment informing the conjoint analysis of the weightings of a variety of migration drivers. Specifically, the approach taken was to examine the existing literature to identify the key characteristics of destination countries that are present and may be relevant for the destination countries within our model. Therefore, we examined the migration drivers included in the previous experimental work (Baláž et al., 2016; Baláž & Williams, 2018) as well as the taxonomy of driver dimensions and individual driver factors (Czaika & Reinprecht, 2020) along with a broader literature review to come up with a long-form list of migration drivers that could potentially be included. Then, through discussions with colleagues and experts within the area of migration studies,[1] we reduced the list down to focus in on the key drivers of interest, while also ensuring the specific drivers chosen provide at least partial coverage across the full breadth of the driver dimensions identified by Czaika and Reinprecht (2020). Specifically, the country-level migration drivers chosen for inclusion were: average wage level, employment level, number of migrants from the country of origin already present, cultural and linguistic links with the country of origin, climate and safety from extreme weather events, openness of migration policies, personal safety and political stability, education and training opportunities, income equality and standard of living, and public infrastructure and services (e.g., health).

Having identified the key drivers for inclusion, the approach used to examine this specific question was an experiment using a conjoint analysis design (Hainmueller et al., 2014, 2015). In a conjoint analysis experiment, participants are presented with a series of trials, each of which presents alternatives that contain information on a number of key attributes (in this case, migration drivers). This approach allows researchers to gain information about the causal role of a number of attributes within a single experiment, rather than conducting multiple experiments or one excessively long experiment that examines the role of each individual attribute one at a time (Hainmueller et al., 2014). Additionally, because all of the attributes are presented together on each trial, it is possible to establish the weightings of each attribute relative to the other presented attributes. That is, a conjoint analysis design allows the analyst to establish not only whether wages have an effect, but how strong that effect is relative to other drivers such as employment level or education and training opportunities. An example of the implementation of the conjoint analysis experiment is presented in Fig. 6.3.

Another benefit of the conjoint analysis approach is that because weightings are revealed at least somewhat implicitly (rather than in designs that explicitly ask participants about the weightings or importance they place on specific attributes),

[1] With special thanks to Mathias Czaika.



a

Country A		Country B
High	Cultural and Language Links	Low
Medium	Income Equality and Standard of Living	Medium
High	Level of Employment	Low
High	Public Infrastructure and Services	Medium
Low	Climate and Safety from Extreme Weather Events	Low
High	Education and Training Opportunities	High
Low	Migrants from Your Country Already Present	Low
Medium	Openness of Migration Policies	Medium
Low	Wage Level	Medium
Medium	Security, Safety, and Political Stability	High

b

Assuming you have to migrate, which country would you prefer to migrate to?

Country A

Country B

Please rate how you much you would like to migrate to each country using a scale of 0 to 100. The higher the number (closer to 100), the more you would like to migrate to that country, the lower the number (closer to 0), the less you would like to migrate. Ratings near 0 indicate that you would not like to migrate to that country.

Please note that you are making these ratings separately for each country (i.e., both could be rated high, both could be rated low, or any combination of high and low ratings that best represents your judgments).

Fig. 6.3 Example of a single trial in the conjoint analysis experiment (panel A) and the questions participants answer for each trial (panel B). (Source: own elaboration in Qualtrics) and because multiple attributes are presented at the same time, participants may be less influenced by social desirability because they can use any of the attributes present to justify their decision. This is supported by a study by Hainmueller et al. (2015) who found that a paired conjoint analysis design did best at matching the relative weightings of attributes for decisions on applications for citizenship in Switzerland when these weightings were compared to a real-world benchmark (the actual results of referendums on citizenship applications). For these reasons, within the present study we also ask participants to explicitly state how much they weight each variable, allowing for greater understanding of how well people's stated and revealed preferences align with each other. This comparison between implicit and explicit weightings is also expected to reveal the extent to which people are aware of, and able or willing to communicate the relative value they place on the country attributes that motivate them to choose one destination country over another.

The results from this conjoint analysis experiment can be used to inform the agent-based model by collecting empirical data on the relative weightings of various migration drivers. Additionally, because the experimental data are collected at an individual level, it is also possible to observe to what extent these weightings are heterogenous between individuals (e.g., whether some individuals place more emphasis on safety while others care more about economic opportunities). These relative weightings can then be combined with real-world data on actual migration destination countries or cities to calculate 'desirability' scores for potential migration destinations within the model, either at an aggregate level or, if considerable heterogeneity is present, by calculating individual desirability scores for each agent to properly reflect the differences in relative weightings found in the empirical data. The model can then be rerun with migration destinations that vary in terms of desirability to examine what effects this has on aspects such as agent behaviour, route formation, and total number of agents arriving at each destination.

6.5 Design, Implementation, and Limitations of Psychological Experiments for Agent-Based Models

When designing and implementing psychological experiments, there are several key aspects that must be considered to ensure that valid and reliable conclusions can be drawn from the experiment. Although both reviewing the existing empirical literature and experimental methods have great potential to contribute to the design and implementation of agent-based models, there are also some serious limitations with these approaches. No single experiment or set of experiments is ever perfect, and there are often trade-offs that must be made between various competing interests when designing and implementing a study. In the following section, we discuss several key aspects of designing and implementing psychological experiments using examples from Sects. 6.2, 6.3, and 6.4. The aspects covered include confounding variables, measurement accuracy, participant samples, and external validity of experimental paradigms. In addition to guidance on how these aspects can be addressed we also discuss the limitations of the experimental approaches used (and many psychological experiments more broadly) and suggest ways to overcome these limitations.

When designing a psychological experiment it is important to consider the potential for confounds to influence the outcome (Kovera, 2010). Confounding occurs when there are multiple aspects that vary across experimental conditions, meaning that it is not possible to infer whether the changes seen are due to the intended experimental manipulation, or occur because of another aspect that differs between the conditions. For example, in the experiment discussed in Sect. 6.3, we were interested in the influence of information source on the judgements and decisions that were made. Therefore, we included information from sources such as a news article, an official organisation, and a family member. However, we ensured that the actual information provided to participants was kept consistent regardless of the source (e.g., 'the migrant sea route is unlikely to be safe') rather than varying the information across the source formats, such as by presenting a full news article when the source was a news article or a short piece of dialogue when the source was a family member. To examine the role of source, it was crucial that the actual information provided was kept consistent because otherwise it would be impossible to tell whether differences found were due to changes in the source or because of another characteristic such as the length or format of the information provided. However, the drawback in choosing to keep the information presented identical across sources is that the stimuli used are less representative of their real-world counterparts (i.e., the news articles used in the study are less similar to real-world news articles), highlighting that gaining additional experimental control to limit potential confounds can come at the cost of decreasing external validity.

Another key issue to consider is the importance of measurement (for a detailed review see Flake & Fried, 2020). Although a full discussion and evaluation is beyond the scope of the current chapter, some aspects of measurement related issues are made particularly clear through the experiment described in Sect. 6.2. Within this study, we wanted to elicit parameters related to prospect theory. However, previous research by Bauermeister et al. (2018) found that, relevant for prospect theory, the estimates of risk attitudes and probability weightings for the same participants depended on the specific elicitation methodology used. Specifically, Bauermeister et al. compared the methodology from Tanaka et al. (2010) and Wakker and Deneffe (1996), and found that the elicited estimates for participants were more risk averse when the former approach was used, whereas they were more biased in their probability weightings when the latter method was applied (with greater underweighting of high probabilities and overweighting of low probabilities). This raises serious concerns around the robustness of findings, because it suggests that the estimates of prospect theory parameters gathered may be conditional on the experimental methodology used and therefore these estimates are incredibly difficult to generalise and apply to an agent-based model. We attempted to address these issues by using the non-parametric methodology of Abdellaoui et al. (2016), since it requires fewer assumptions than many other elicitation methods. However, the findings of Bauermeister et al. (2018) still highlight the extent to which the results of studies can be highly conditional on the specific methodology and context in which the study takes place, and therefore may be difficult to generalise.

Issues with the typical samples used within psychology and other social sciences have been well documented for many years now (Henrich et al., 2010). Specifically, it has long been pointed out that the populations used for social science research are much more Western, Educated, Industrialised, Rich, and Democratic (WEIRD) than the actual human population of the Earth (Henrich et al., 2010; Rad et al., 2018). This bias means that much of the data within the social sciences literature that can be used to inform agent-based models may not be applicable whenever the social process or system being modelled is not itself comprised solely of WEIRD agents. Even though this issue has been known about for quite some time, there has not yet been much of a shift within the literature to address it. Arnett (2008) found that between 2003 and 2007, 96% of the participants of experiments reported in top psychology journals were from WEIRD samples.

More recently, Rad et al. (2018) found that 95% of the participants of the experiments published in *Psychological Science* between 2014 and 2017 were from WEIRD samples, suggesting that even though a decade had passed, there had been little change in the extent to which non-WEIRD populations are underrepresented within the psychological literature. Despite their being relatively little research conducted with non-WEIRD samples, that research has produced considerable evidence that there are cultural differences across many areas of human psychology and behaviour, such as visual perception, morality, mating preferences, reasoning, biases, and economic preferences (for reviews see Apicella et al., 2020; Henrich et al., 2010). Of particular relevance for the experiments discussed in the previous sections, Falk et al. (2018) found that economic preferences vary considerably between countries and Rieger et al. (2017) found that, although descriptively, the results from nearly all of the 53 countries they surveyed were consistent with prospect theory, the estimates for the parameters of cumulative prospect theory differed considerably between countries. Therefore, if there is a desire to use results from the broader literature or from a specific study to inform an agent-based model, then it is important for researchers to ensure that the participants included within their studies are representative of the population(s) of interest, rather than continuing to sample almost entirely from WEIRD populations and countries.

The issue of the extent to which findings from experimental contexts can be generalised to the real-world has also received considerable attention across a wide range of fields (Highhouse, 2007; Mintz et al., 2006; Polit & Beck, 2010; Simons et al., 2017). As highlighted by Highhouse (2007), many critiques of experimental methodology place an unnecessarily large emphasis on surface-level ecological validity. That is, the extent to which the materials and experimental setting appear similar to the real-world equivalent (e.g., how much the news articles used as materials within a study look like real-world news articles). However, provided the methodology used allows for proper understanding of "the process by which a result comes about" (Highhouse, 2007, p. 555), then even if the experiment differs considerably from the real world, the information gained is still helpful for developing theoretical understanding that can then be tested and applied more broadly. In the context of asylum migration, additional insights can be gained from some related areas, for example on evacuations during terrorist attacks or natural disasters (Lovreglio et al., 2016), where agent-based models are successfully used to predict and manage the actual human behaviour (e.g. Christensen & Sasaki, 2008; Cimellaro et al., 2019; see also an example of Xie et al., 2014 in Chapter 5). Conceptually, one common factor in such circumstances could be the notion of *fear* (Kok, 2016).

Nonetheless, migration is an area in which the limitations of lab or online-based experimental methods and the difficulty of truly capturing and understanding the real-world phenomena of interest becomes clear. Deciding to migrate introduces considerable disruption and upheaval to an individual or family's life, along with potential excitement at new opportunities and discoveries that might await them. How then can a simple experiment or survey conducted in a lab or online via a web browser possibly come close to capturing the real-world stakes or the magnitude of the decisions that are faced by people when they confront these situations in the real world? This problem is likely even more pronounced for migrants seeking asylum, who are likely to be making decisions under considerable stress and where the decisions that they make could have actual life or death consequences. Given the large body of evidence showing that emotion can strongly influence a wide range of human behaviours, judgments, and decisions (Lerner et al., 2015; Schwarz, 2000), it becomes clear that it is incredibly difficult to generalise and apply findings from laboratory and online experimental settings in which the degree of emotional arousal, emotional engagement, and the stakes at play are so greatly reduced from the real-world situations and phenomena of interest.

For the purpose of the modelling work presented in this book, we focus therefore on incorporating the empirical information elicited on the subjective measures (probabilities) related to risky journeys and the related confidence assessment (Sect. 6.3). The process is summarised in Box 6.1.

Box 6.1: Incorporating Psychological Experiment Results Within an Agent-Based Model

Incorporating the results of psychological experiments with an agent-based model may not be a straightforward task, because the specific method of implementation will vary greatly depending on the setup and structure of the model. Therefore, this brief example is designed to outline how results from the experiment in Sect. 6.3 have been incorporated into an agent-based model of migration (see Chap. 8 for more details on the updated version of the model).

In the updated version of the original Routes and Rumours model introduced in Chap. 3, called 'Risk and Rumours' (see Chap. 8), agents make safety ratings for the links between cities within the simulation, and these ratings subsequently effect the probability that they will travel along a link. Within the updated Risk and Rumours model, agent beliefs about risk are represented as an estimate *v_risk*, with a certainty measure *t_risk*, bounded between 0 and 1.

(continued)

Box 6.1 (continued)

Within the model, agents form these beliefs based on their experiences travelling through the world as well as by exchanging information with other agents. There is also a scaling parameter for risk, *risk_scale* which is greater than 1. Based on the above, for risk-related decisions, an agent's safety estimate for a given link (s) is derived as:

$$s = t\_risk * \left(1 - v\_risk\right)^{risk\_scale} * 100$$

The logit of the probability to leave for a given link (p) is then calculated as:

$$p = I + S * s$$

The results of the experiment in Sect. 6.3 are incorporated within the model through the values of the intercept I and slope S. These variables take agent-specific values drawn from a bivariate normal distribution, the parameters for which come from the results of a logistic regression conducted on the data collected in the experiment. In this way, the information gained from the psychological experiment about how safety judgments influence people's willingness to travel is combined with the beliefs that agents within the model have formed, thereby influencing the probability that agents will make the decision to travel along a particular link on their route.

6.6 Immersive Decision Making in the Experimental Context

The development of more immersive and engaging experimental setups can provide an exciting avenue to address several of the concerns outlined in the previous section. Increasing immersion within experimental studies is particularly helpful for addressing concerns related to realism and emotional engagement of participants. One potentially beneficial approach that can be used to increase emotional engagement, and thereby at least partially close the emotional gap between the experimental and the real-world, is through 'gamification'. Research has shown that people are motivated by games and that playing games can satisfy several psychological needs such as needs for competence, autonomy, and relatedness (Przybylski et al., 2010; Ryan et al., 2006).

Additionally, Sailer et al. (2017) showed that a variety of aspects of game design can be used to increase feelings of competence, meaningfulness, and social connectedness, feelings that many researchers are likely to want to elicit in participants to increase immersion and emotional engagement while they are completing an experiment. Using gamification to increase participant engagement and motivation does not even require the inclusion of complex or intensive game design elements.

Lieberoth (2014) found that when participants were asked to engage in a discussion of environmental issues, simply framing the task as a game through giving participants a game board, cards with discussion items, and pawns increased task engagement and self-reported intrinsic motivation, even though there were no actual game mechanics.

To improve the immersion and emotional engagement of participants in experimental studies of migration, we plan to use gamification aspects in future experiments. Specifically, we aim to design a choose-your-own adventure style of game to explore judgements and decision making within asylum migration context. Inspiration for this approach came from interactive choose-your-own adventure style projects that were developed by the BBC (2015) and Channel 4 (2015) to educate the public about the experiences of asylum seekers on their way to Europe.[2] We plan to use the agent-based models of migration that have been developed to help generate an experimental setup, and then combine this with aspects of gamification to develop an experiment that can be 'played' by participants. For example, by mapping out the experiences, choices, and obstacles that agents within the agent-based models encounter as well as the information that they possess, it is possible to generate sequences of events and choices that occur, and then design a choose-your-own adventure style game in which real-world participants must go through the same sequences of events and choices that the agents within the model face. This allows for the collection of data from real-world participants that can be directly used to calibrate and inform the setup of the agents within the agent-based model, while simultaneously also having the advantage of being more immersive, engaging, and motivating for the participants completing the experiment.

Improvements in technology also allow for the development of even more advanced and immersive experiments in the future, using approaches such as video game modifications (Elson & Quandt, 2016), and virtual reality (Arellana et al., 2020; Farooq et al., 2018; Kozlov & Johansen, 2010; Mol, 2019; Moussaïd et al., 2016; Rossetti & Hurtubia, 2020). Elton and Quandt (2016) highlighted that by using modifications to video games, it is possible for researchers to have control over many aspects of a video game, allowing them to design experiments by operationalising and manipulating variables and creating stimulus materials so that participants in experimental and control groups can play through an experiment in an immersive and engaging virtual environment. At the same time, observational studies based on information from online games allow for studying many aspects of social reality and social dynamics, which may be relevant for agent-based models, such as networks and their structures, collaboration and competition, or inequalities (e.g. Tsvetkova et al., 2018).

The increased availability and decreased costs of virtual reality headsets have also allowed for researchers to test the effectiveness of presenting study materials and experiments within virtual reality. Virtual reality has already been used to

[2] For the interactive versions of these online tools, see https://www.bbc.co.uk/news/world-middle-east-32057601 and http://twobillionmiles.com/ (as of 1 January 2021).

examine phenomena such as pedestrian behaviour and traffic management (Arellana et al., 2020; Farooq et al., 2018; Rossetti & Hurtubia, 2020), behaviour during emergency evacuations (Arellana et al., 2020; Moussaïd et al., 2016), and the bystander effect (Kozlov & Johansen, 2010). It has also been applied to a wide range of areas within economics and psychology (for a review see Mol, 2019). In the context of agent-based simulation models, hybrid approaches, with human-computer interactions, have also been the subject of experiments (Collins et al., 2020). These new technological developments allow for the simulation and manipulation of experimental environments in ways that are simply not possible using standard experimental methods, or would be unethical and dangerous to study in the real world. They allow researchers to take several steps towards closing the gap between the laboratory and the real world, and open the door to many exciting new research avenues.

Chapter 7
Agent-Based Modelling and Simulation with Domain-Specific Languages

Oliver Reinhardt, Tom Warnke, and Adelinde M. Uhrmacher

Conducting simulation studies within a model-based framework is a complex process, in which many different concerns must be considered. Central tasks include the specification of the simulation model, the execution of simulation runs, the conduction of systematic simulation experiments, and the management and documentation of the model's context. In this chapter, we look into how these concerns can be separated and handled by applying domain-specific languages (DSLs), that is, languages that are tailored to specific tasks in a specific application domain. We demonstrate and discuss the features of the approach by using the modelling language ML3, the experiment specification language SESSL, and PROV, a graph-based standard to describe the provenance information underlying the multi-stage process of model development.

7.1 Introduction

In sociological or demographic research, such as the study of migration, simulation studies are often initiated by some unusual phenomenon observed in the macro-level data. Its explanation is then sought at the micro-level, by probing hypotheses about decisions, actions, and interactions of individuals (Coleman, 1986; Billari, 2015). In this way, theories about decisions and behaviour of individuals, as well as data that are used as input, for calibration, or validation, contribute to the model generation process at the micro- and macro-level respectively. Many agent-based demographic simulation models follow this pattern, e.g., for fertility prediction (Diaz et al., 2011), partnership formation (Billari et al., 2007; Bijak et al., 2013), marriage markets (Zinn, 2012) as well as migration (Klabunde & Willekens, 2016; Klabunde et al., 2017). Whereas typically, data used for calibration and validation focuses on the macro-level, additional data that enter the model-generating process at micro-level add both to the credibility of the simulation model (see Chaps. 4 and 6) and to the complexity of the simulation study.

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_7

An effective computational support of such simulation studies needs to consider various concerns. These include specifying the simulation model in a succinct, clear, and unambiguous way, its efficient execution, executing simulation experiments flexibly and in a replicable manner (see Chap. 10), and making the overall process of conducting a simulation study, including the various sources and the interplay of model refinement and of simulation experiment execution, explicit. Given the range of concerns, domain-specific languages (DSLs) seem particularly apt to play a central role within supporting simulation studies, as they are aimed at describing specific concerns within a specific domain (Fowler, 2010). In DSLs, abstractions and notations of the language are tailored to the specific concerns in the application domain, so as to allow the stakeholders to specify their particular concerns concisely, and others in an interdisciplinary team to understand these concerns more easily. The combination of different DSLs within a simulation study naturally caters for the separation of different concerns required for handling the art and science of conducting simulation studies effectively and efficiently (Zeigler & Sarjoughian, 2017).

In this chapter, we explore how different DSLs can contribute to (a) agent-based modelling (and present implications for the efficient execution of these models) based on the modelling language ML3, (b) specifying simulation experiments based on the simulation experiment specification language SESSL, and finally, (c) to relating the activities, theories, data, simulation experiment specifications, and simulation models by exploiting the provenance standard PROV. We also discuss a salient feature of DSLs, that is, that they constrain the possibilities of the users in order to gain more computational support, and the implication for use and reuse of the language and model.

7.2 Domain-Specific Languages for Modelling

DSLs for modelling are aimed at closing the gap between model documentation and model implementation, with the ultimate goal to conflate both in an executable documentation. Two desirable properties of a DSL for modelling are practical expressiveness, describing the ease of specifying a model in the language as well as how clearly more complex mechanisms can be expressed, and succinctness. Whereas the number of the used lines of code can serve as an indication for the latter, the former is difficult to measure. Practical expressiveness must not be confused with formal expressiveness, which measures *how many* models can theoretically be expressed in the language, or, in other words, the genericity of the language (Felleisen, 1991).

7.2.1 Requirements

A necessary prerequisite for achieving practical expressiveness is to identify central requirements of the application domain before developing or selecting the DSL. These key requirements related to agent-based models, specifically in the migration context, are listed below.

Objects of Interest. In migration modelling, the central objects of interest are the individual migrants and their behaviour. With an agent-based approach, migrants are put in the focus and represented as agents. In contrast to population-based modelling approaches, such an agent-based approach allows modelling of the heterogeneity among migrants. Each migrant agent has individual attribute values and an individual position in the social network of agents. As a consequence, agent-based approaches allow modelling of how the situation and knowledge of an individual migrant influences his or her behaviour. In addition to the migrant as the central entity, other types of actors can be modelled as agents in the system, for example government agencies or smugglers. Although these might correspond to higher-level entities, depicting them as agents facilitates modelling of the interaction between different key players in migration research.

Dynamic Networks. Agent-based migration models need to include the effects of agents' social ties on their decisions and vice versa. Therefore, both the local attributes of an agent and its network links to other agents should be explicitly represented in the modelling language. It is also crucial to allow for several independent networks between agents. This becomes particularly important when combining different agent types as suggested above, for example to distinguish contact networks among migrants from contacts between migrants and smugglers. Note that encoding changes in the networks can be challenging, both in the syntax of the DSL as well as in the simulator implementation.

Compositionality. Agent-based simulation models can become complex quickly due to many interconnected agents acting in parallel. All agents can act in ways that change their own state, the state of their neighbours, or network links. A DSL can address this complexity by supporting compositional modelling. As stated by Henzinger et al. (2011, p. 12), "[a] compositional language allows the modular description of a system by combining submodels that describe parts of the system". An agent-based model as described above can be decomposed into parts on several levels. First, different types of agents can be distinguished. Second, different types of behaviour of a single type of agent can be described independently. Both improve the readability of the model, as different parts of the model can be understood individually.

Decisions. A central goal of this simulation study is to deepen our understanding of migrants' decision processes (see Chaps. 3 and 6). Modelling these decisions in detail, and the migrants' knowledge on which they are based, is therefore inevitable.

The DSL must therefore be powerful and flexible enough to express them. In addition, the language must not be limited to a single model of decision making, to enable an implementation and comparison of different decision models.

Formal Semantics. Simulation models are often implemented in an *ad hoc* fashion. If a model is instead specified with a DSL and that DSL has a formal definition, it becomes possible to interpret the model or parts of it based on formal semantics. The semantics of a DSL for modelling maps a given model to a mathematical structure of some class, often a stochastic process. For example, many modelling approaches in computational biology are based on Continuous-Time Markov Chains (De Nicola et al., 2013). In addition to helping the interpretation of a model, establishing the connection between the DSL and the underlying stochastic process also informs the design of the simulation algorithm and, for example, allows reasoning over optimisations. Thus, DSLs for agent-based modelling of migration benefit from having a formal definition.

Continuous Time. In agent-based modelling, there are roughly two ways to consider the passing of time. The first approach is the so-called 'fixed-increment time advance,' where all agents have the opportunity to act on equidistant time points. Although that approach is the dominant one, it can cause problems that threaten the validity of the simulation results (Law, 2006, 72 *ff*). First, the precise timing of events is lost, which prohibits the analysis of the precise duration between events (Willekens, 2009). Second, events must be ordered for execution at a time point, which can introduce errors in the simulation. The alternative approach is called 'next-event time advance' and allows agents to act at any point on a continuous time scale. This approach is very rarely used in agent-based modelling, but can solve the problems above. Therefore, a DSL for agent-based modelling of migration should allow agents to act in continuous time.

7.2.2 The Modelling Language for Linked Lives (ML3)

Based on the above requirements we selected the Modelling Language for Linked Lives (ML3). ML3 is an external domain-specific modelling language for agent-based demographic models. In this context, *external* means that it is a new language independent of any other, as opposed to an *internal* DSL that is embedded in a host language and makes use of host language features. ML3 was designed to model life courses of interconnected individuals in continuous time, specifically with the modelling of migration decisions in mind (Warnke et al., 2017). That makes ML3 a natural candidate for application in this project. In the following Box 7.1, we give a short description of ML3, with examples taken from a version of the Routes and Rumours model introduced in Chap. 3, available at https://github.com/oreindt/routes-rumours-ml3, and relate it to the requirements formulated above.

Box 7.1: Description of the Routes and Rumours Model in ML3

Agents: The primary entities of ML3 models are agents. They represent all acting entities of the modelled system, including individual persons, but also higher-level actors, such as families, households, NGOs or governments. An agent's properties and behaviour are determined by their type. Any ML3 model begins with a definition of the existing agent types. The following defines an agent type `Migrant`, to represent the migrants in the Routes and Rumours model:

```
Migrant(
  capital : real,
  in_transit : bool,
  steps : int
)
```

Agents of the type `Migrant` have three attributes: their capital, which is a real number (defined by the type real after the colon), for example an amount in euro; and a Boolean attribute, that denotes if they are currently moving, or staying at one location; and the number of locations visited so far.

Agents can be created freely during the simulation. To remove them, they may be declared 'dead'. Dead agents do still exist, but no longer act on their own. They may, however, still influence the behaviour of agents who remain connected to them.

Links: Relationships between entities are modelled by links. Links, denoted by `<->`, are bidirectional connections between agents of either the same type (e.g., migrants forming a social network), or two different types (e.g., migrants residing at a location that is also modelled as an agent). They can represent one-to-one (`<->` e.g., two agents in a partnership), one-to-many (`<->` e.g., many migrants may be at any one location, but any migrant is only at one location), or many-to-many relations (`<->` e.g., every migrant can have multiple other migrant contacts, and may be contacted by multiple other migrants). The following defines the link between migrants and their current location in the Routes and Rumours model:

```
location:Location[1]<->[n]Migrant:migrants
```

This syntax can be read in two directions, mirroring the bidirectionality of links: from left to right, it says that any one `[1]` agent of the type `Location` may be linked to multiple `[n]` agents of the type `Migrant`, who are referred to as the location's `migrants`. From right to left, any `Migrant` agent is linked to one `Location`, which is called its `location`. ML3 always preserves the consistency of bidirectional links. When one direction is changed, the other is changed automatically. For example, when a new location is set for a migrant, it is automatically removed from the old location's migrants, and added to the new location's migrants.

(continued)

Box 7.1 (continued)

Function and procedures: The ability to define custom functions and procedures adds expressive power to ML3, allowing complex operations, and aiding readability and understandability by allowing for adding a layer of abstraction where necessary. Unlike many general-purpose programming languages, ML3 distinguishes functions, encapsulating calculations that return a result value, and procedures, containing operations that change the model state. Both are bound to a specific agent type, making them related to methods in object-oriented languages. A library of predefined functions and procedures aids with common operations. The following function calculates the cost of travel from the migrant's current location to a potential destination (given as a function parameter):

```
Migrant.move_cost(?destination : Location) : real :=
    costs_move * ego.location.link
    .filter(?destination in alter.endpoints).only().friction
```

The value of this function is calculated from the base cost of movement (the model parameter `costs_move`), scaled by the friction of the connection between the two locations, which is gained by filtering all outgoing ones using the predefined function filter, and then unwrapping the only element from the set of results using `only()`. The keyword `ego` refers to the agent the function is applied to. Procedures are defined similarly, with -> -ing the:=.

Rules: Agents' behaviour is defined by rules. Every rule is associated with one agent type, so that different types of agents behave differently. Besides the agent type, any rule has three parts: a guard condition, that defines *who* acts, i.e., what state and environment an agent of that type must be in, to show this behaviour; a rate expression, that defines *when* they act; and the effect, that defines *what* they do. With this three-part formulation, ML3 rules are closely related to stochastic guarded commands (Henzinger et al. 2011). The following (slightly shortened) excerpt from the Routes and Rumours shows the rule that models how migrants begin their move from one location to the next:

```
1 Migrant
2 | !ego.in_transit                          // guard
3 @ ego.move_rate()                          // rate
4 -> ego.in_transit := true                  // effect
5    ego.destination := ego.decide_destination()
```

The rule applies to all living agents of the type `Migrant` (line 1). Like in a function or procedure, `ego` refers to one specific agent to which the rule is applied. According to the `guard` denoted by | (line 2) the rule applies to all

(continued)

Box 7.1 (continued)

migrants who are currently not in transit between locations. The `rate` following `@` (line 3) is given by a call to the function `move_rate`, where a rate is calculated depending on the agent's knowledge of potential destinations. The value of the rate expression is interpreted as the rate parameter of an exponential distribution that governs the waiting time until the effect is executed. Rules with certain non-exponential waiting times may be defined with special keywords (see Reinhardt et al., 2021). The effect is defined in lines 4 and 5, following `->`. The migrant decides on a destination and is now in transit to it.

In general, the guard and rate may be arbitrary expressions, and may make use of the agent's attributes, links (and attributes and links of linked agents as well), and function calls. The effect may be an arbitrary sequence of imperative commands, including assignments, conditions, loops, and procedure calls. The possibility of using arbitrary expressions and statements in the rules is included to give ML3 ample expressiveness to define complex behaviour and decision processes. The use of functions and procedures allows for encapsulating parts of these processes to keep rules concise, and therefore readable and maintainable.

For each type of agent, multiple rules can be defined to model different parts of their behaviour, and the behaviour of different types of agents is defined in separate rules. The complete model can therefore be composed from multiple sub-models covering different processes, each consisting of one or more rules. Formally, a set of ML3 rules defines a Generalised Semi-Markov Process (GSMP), or a Continuous-time Markov Chain (CTMC) if all of the rules use the default exponential rates. The resulting stochastic process was defined precisely in Reinhardt et al. (2021).

7.2.3 Discussion

Any domain-specific modelling language suggests (or even enforces), by the metaphors it applies and the functionality it offers, a certain style of model. Apart from the notion of linked agents, which is central for agent-based models, for ML3, the notion of behaviour modelled as a set of concurrent processes in continuous time is also of key importance. This is in stark contrast to commonly applied ABM frameworks such as NetLogo (Wilensky, 1999), Repast (North et al., 2013), or Mesa (Masad & Kazil, 2015), which are designed for modelling in a stepwise, discrete-time approach. If in a simulation model events shall occur in continuous time, these events need to be scheduled manually (Warnke et al., 2016). In this regard, and with its firm grounding in stochastic processes, ML3 is more closely related to stochastic process algebras, which have also been applied to agent-based systems before (Bortolussi et al., 2015). Most importantly, this approach results in a complete separation of the model itself, and its execution. ML3's rules describe these processes declaratively, without including code to execute them (which we describe in the next section of this chapter). This makes the model more succinct, accessible and maintainable.

The result of applying ML3 to the Routes and Rumours model was twofold (Reinhardt et al., 2019). On the one hand, the central concepts of ML3 were well suited to the model, especially in separating the different kinds of behaviour into multiple concurrent processes for movement, information exchange, exploration and path planning. Compared to the earlier, step-wise version of the model (Hinsch & Bijak, 2019), this got rid of some arbitrary assumptions necessitated by the fixed time step, e.g., that movement to another location would always take one unit of time. In the continuous-time version, time of travel can depend on the distance and friction between the locations without restrictions.

On the other hand, it became apparent that some aspects of the model were difficult to express in ML3. In particular, ML3 knows only one kind of data structure: the set. This hindered modelling the migrants' knowledge about the world and the exchange of knowledge between migrant agents. These processes could be expressed, but only in a cumbersome way that, in addition, was highly inefficient for execution. The reason for this lack of expressive power is rooted in ML3's design as an external DSL, with a completely new syntax and semantics independent of any existing language. The inclusion of all the capabilities that general purpose languages have in regards to data structures would be possible, but would be unreasonable due to the necessary effort.

While the application of ML3 in this form was deemed impractical for the simulation model, insights from its application very much shaped the continued model development. The model was redesigned in terms of continuous processes, using the macro system of a general-purpose language (in this case, Julia) to achieve syntax similar to ML3's rules, as this excerpt, equivalent to the rule shown above, demonstrates:

```
1 @processes sim agent::Agent begin
2 ...
3 @poisson(move_rate(agent, sim.par))
4   ~ ! agent.in_transit
5   => start_move!(agent, sim.model.world, sim.par)
```

Line 1 is equivalent to line 1 in the ML3 rule (Box 7.1), with the difference that in ML3 the connection to an agent type is declared individually for every rule, while this version does it for a whole set of processes. Lines 3 to 5 contain the same three elements (guard, rate, effect) as ML3 rules, but with the order of the first two switched. The effect was put in a single function `start_move`, which contains code equivalent to that in the effect of the ML3 rule. This Julia version is, however, not completely able to separate the simulation logic from the model itself, but requires instructions in the effect, to trigger the rescheduling of events described in the next section.

In terms of language design, this endeavour showed the potential of redesigning ML3 as an internal DSL. ML3's syntax for expressions and effects already closely resembles object-oriented languages. Embedding it in an object-oriented host-language would allow the use of a similar syntax and other host-language features, such as complex data structures, type systems as well as tooling, for generating and debugging models.

7.3 Model Execution

When a simulation model is specified, it must be executed to produce results. If the model is implemented in a general-purpose language, this usually just means executing the model code. However, if specified in a DSL such as ML3, the model specification does not contain code for the execution, which is handled by a separate piece of software: the simulator. Given a model and an initial model state, i.e., a certain population of agents, the simulator must sample a trajectory of future states. For models with exponentially distributed waiting times, such as ML3, algorithms to generate such trajectories are well established, many of them derived from Gillespie's Stochastic Simulation Algorithm (SSA) (Gillespie, 1977). In the following, we describe a variation of the SSA for ML3. A more detailed and technical description can be found in Reinhardt and Uhrmacher (2017). The implementation in Java, the ML3 simulator, is available at https://git.informatik.uni-rostock.de/mosi/ml3.

7.3.1 Execution of ML3 Models

We begin the simulation with an initial population of agents, our state s, which is assumed at some point in time t (see Fig. 7.1a). As described in Sect. 7.2, each ML3 agent has a certain type, and for each type of agent there are a number of stochastic rules that describe their behavior. Each pair of a living agent a and a rule r matching the agent's type, where the rule's guard condition is fulfilled, yields a possible state transition (or *event*), given by the rule's effect applied to the agent. It is associated with a stochastic waiting time T until its occurrence, determined by an exponential distribution whose parameter is given by the rule's rate applied to the agent $\lambda_r(a, s)$. To advance the simulation we have to determine the event with the smallest waiting time Δt, execute its effect to get a new state s' and advance the time to the time of that event $t' = t + \Delta t$.

As per the semantics of the language, the waiting time T is exponentially distributed:

$$P\left(T \le \Delta t\right) = 1 - e^{-\lambda_r(a,s)\cdot\Delta t}. \tag{7.1}$$

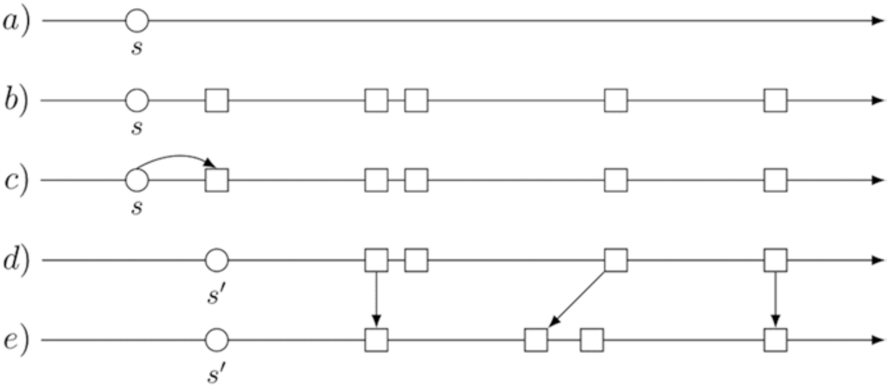


Fig. 7.1 Scheduling and rescheduling of events. We begin in state s at some time t depicted as the position on the horizontal time line (**a**). Events (squares) are scheduled (**b**). The earliest event is selected and executed (**c**), resulting in a new state s' at the time of that event (**d**). Then, affected events must be rescheduled (**e**)

This distribution can be efficiently sampled using inverse transform sampling (Devroye, 1986), i.e. by sampling a random number u from the uniform distribution on the unit interval and applying the distribution function's inverse:

$$\Delta t = -\frac{1}{\lambda_r\left(a,s\right)} \cdot \ln u \tag{7.2}$$

Using this method, we can sample a waiting time for every possible event (Fig. 7.1b). We can then select the first event, and execute it (Fig. 7.1c). In practice, the selection of the first event is implemented using a priority queue (also called the event queue), a data structure that stores pairs of objects (here: events) and priorities (here: times), and allows retrieval of the object with the highest priority very efficiently.

After the execution of this event, the system is in a new state s' at a new time t'. Further, we still have sampled execution times for all events, except the one that was executed (Fig. 7.1d). Unfortunately, in this changed state, these times might no longer be correct. Some events might no longer be possible at all (e.g., the event was the arrival of a migrant at their destination, so other events of this agent no longer apply). For others, the waiting time distribution might have changed. And some events might not have been possible in the old state, but are in the new (e.g., if a new migrant entered the system, new events will be added). In the worst case, the new state will require the re-sampling of all waiting times. In a typical agent-based model, however, the behaviour of any one agent will not directly affect the behaviour of many other agents. Their sampled times will still therefore be valid. Only those events that are affected will need to be re-sampled (Fig. 7.1e). In the ML3 simulator this is achieved using a dependency structure, which links events to attribute and link values of agents. When the waiting time is sampled, all used attributes and links are stored as dependencies of that event. After an event is executed, the events dependent on the changed attributes and links can then be retrieved. A detailed and more technical description of this dependency structure can be found in Reinhardt and Uhrmacher (2017).

In Box 7.2 below, Algorithm 1 shows the algorithm described above in pseudo-code, and algorithm 2 shows the sampling of a waiting time for a single event.

Box 7.2: Examples of Pseudo-Code for Simulating and Scheduling Events

Algorithm 1 Simulation Algorithm.
s: the current state, given as a set of agents
t: the current time
m: the model, given as a set of rules
Q: the event queue
D: the dependency structure

```
// schedule all potential events in the event queue
for each a ∈ s, r ∈ m:
    if !dead(a, s) and typeof(a) = typeof(r):
        schedule(r, a)

while t < t_end:
    // select the next event from the queue
    (r, a, Δt) := pop(Q)

    // advance simulation time
    t := t + Δt

    // execute the event
    s := effect(r)(a, s)

    // reschedule the executed event
    schedule(r, a)

    // reschedule all affected events
    for each (r, a) ∈ affected(D):
        schedule(r, a)
```

Algorithm 2 Schedule.
(r, a): the event to schedule
s: the current state, given as a set of agents
t: the current time
m: the model, given as a set of rules
Q: the event queue
D: the dependency structure

```
if !dead(a, s) and guard(r)(a, s):
    u ~ Uniform(0,1)
    Δt := − 1/rate(r)(a,s) · ln u
    push(Q,r,a,Δt)
else:
    remove(Q, r, a)
update(*D*)
```

7.3.2 Discussion

The simulation algorithm described above is abstract in the sense that it is independent of the concrete model. The model itself is only a parameter for the simulation algorithms – in the pseudo-code in Algorithm 1 in Box 7.2 it is called *m*. As a result, the simulator, i.e., the implementation of the simulation algorithm, is model-independent. All the execution logic can hence be reused for multiple models. This not only facilitates model development, it also makes it economical to put more effort into the simulator, as this effort benefits many models.

On the one hand, this effort can be put into quality assurance, resulting in better tested, more reliable software. A simulator that has been tested with many different models will generally be more trustworthy than an *ad hoc* implementation for a single model (Himmelspach & Uhrmacher, 2009). On the other hand, this effort can be put into advanced simulation techniques. One of these techniques we have already covered: using continuous time. The simulation logic for a discrete-time model is often just a simple loop, where the events of a single time step are processed in order, and time is advanced to the next step. The simulation algorithm described above is considerably more complex than that. But with the simulator being reusable, the additional effort is well invested. Separation of the modelling and the simulation concerns serves as an enabler for continuous-time simulation. Similarly, more efficient simulation algorithms, e.g., parallel or distributed simulators (Fujimoto, 2000), simulators that exploit code generation (Köster et al., 2020), or approximate the execution of discrete events (Gillespie, 2001) developed for the language, will benefit all simulation models defined in this language.

The latter leads us back to an important relationship between the expressiveness of the language and the feasibility and efficiency of its execution. The more expressive the modelling language, and the more freedom it gives to the modeller, the harder it is to execute models, and especially to do so efficiently. The approximation technique of Tau-leaping (Gillespie, 2001), for example, cannot simply be applied to ML3, as it requires the model state and state changes to be expressed as a vector, and state updates to be vector additions. ML3 states – networks of agents – cannot be easily represented that way. Ideally, every feature of the language is necessary for the model, so that implementing the model is possible, but execution is not unnecessarily inefficient. DSLs, being tailored to a specific class of models, may achieve this.

7.4 Domain-Specific Languages for Simulation Experiments

With the increasing availability of data and computational resources, simulation models become ever more complex. As a consequence, gaining insights into the macro- and micro-level behaviour of an agent-based model requires increasingly complex simulation experiments. Simulation experimentation benefits from using DSLs in several ways.

- They allow specifying experiments in a readable and succinct manner, which is an advantage over using general-purpose programming or scripting languages to implement experiments.
- They facilitate composing experiments from reusable building blocks, which makes applying sophisticated experimental methods to simulation models easier.
- They help to increase the trustworthiness of simulation results by making experiment packages available that allow other researchers to reproduce their results.

In this section, we illustrate these benefits by showing how SESSL, a DSL for simulation experiments, is applied for simulation experiments with ML3 and give a short overview of other current developments regarding DSLs for simulation experiments.

7.4.1 Basics

The fundamental idea behind using a DSL for specifying experiments is to provide a syntax that captures typical aspects of simulation experiment descriptions. Using this provided syntax, a simulation experiment can be described succinctly. This way, a DSL for experiment specification 'abstracts over' individual simulation experiments, by creating a general framework covering different specific cases. The commonalities of the experiments become then part of the DSL, and the actual experiment descriptions expressed in the DSL focus on the specifics of the individual experiments.

One experiment specification DSL is the 'Simulation Experiment Specification on a Scala Layer' (SESSL), an internal DSL that is embedded in the object-functional programming language Scala (Ewald & Uhrmacher, 2014). SESSL uses a more refined approach to abstracting over simulation experiments. Between the language core and the individual experiments, SESSL employs simulation-system-specific *bindings* that abstract over experiments with a specific simulation system. Whereas the language core contains general experiment aspects such as writing observed simulation output to files, the bindings package experiment aspects are tailored to a specific simulation approach, such as specifying which simulation outputs to observe. This way, SESSL can cater to the differences between, for example, conducting experiments with population-based and agent-based simulation models: whereas population-based models allow a direct observation of macro-level outputs, agent-based models might require aggregating over agents and agent attributes. Another difference is the specification of the initial model state, which, for an ML3 model, might include specifying how to construct a random network of links between agents.

To illustrate how experimentation with SESSL works, we now consider an example experiment specified with SESSL's binding for ML3 (Reinhardt et al., 2018). The following listing shows an excerpt of an experiment specification for the Routes and Rumours model. Such an SESSL experiment specification is usually saved in a Scala file and can be run as a Scala script.

```
1 execute {
2   new Experiment with Observation {
3     model              = "routes.ml3"
4     replications       = 10
5     stopTime           = 100
6     set("p_find_links" <~ 0.5)
7     observeAt(stopTime)
8
9     initializeWith(JSON("init50.json"))
10    val migrants = observe("migrants" ~ agentCount(agentType = "Migrant"))
11    // additional lines elided
12  }
13 }
```

In an SESSL experiment, a number of options are available. For example, in the listing above, the model file, the number of replications, and the stop time of each simulation run are set in lines 3–5. Line 6 is an example of setting the value of a model input parameter, and line 7 specifies that model outputs are recorded when a simulation run terminates. These are examples of settings that are part of virtually all experiments and, therefore, belong to the SESSL core. The lines 9 and 10, in contrast, refer to settings that are ML3-specific and packaged in the SESSL binding for ML3. Line 9 specifies a JSON file that is used to create an initial population for each simulation run. An ML3-specific observable, which counts the number of `Migrant` agents, is configured in line 10.

Which options are available in an experiment depends on the binding used, but also the creation of the experiment as in line 2. Here, the experiment is configured to include observation options (**with** `Observation`). With such 'mix-ins,' SESSL allows a high degree of flexibility. Some mix-ins are packaged in the SESSL core and provide generic features; others belong to bindings and contain simulation-system-specific features. For example, the `Observation` mix-in above is part of the binding for ML3, and provides commands to record observations from ML3 simulation runs, such as `agentCount`.

This example shows how recurring aspects of simulation experiments can be efficiently expressed. Through bindings and mix-ins, SESSL allows for packaging code and making it available for reuse across experiments. As a result, the actual experiment specification focuses on the specifics of the experiment with little syntactical overhead.

7.4.2 Complex Experiments

The specification of more complex experiments in SESSL exploits the abstraction over different simulation systems. Many experimental methods can be integrated with the generalisation of simulation experiments in the SESSL core. As a result, those methods can be applied to any experiment for any simulation system. Examples of experimental methods that are realised this way are algorithms to create designs of experiments, which work with the inputs of an experiment (e.g., set in the experiment shown above), or algorithms that process the outputs.

We demonstrate this by fitting a regression meta-model to the Routes and Rumours model, based on a central composite design (see Reinhardt et al., 2018 for background). Based on the experiment specification shown above, three changes are necessary to integrate these experimental methods with the experiment. First, the mix-ins `CentralCompositeDesign` and `LinearRegression` are added to the experiment:

```
new Experiment with ... with CentralCompositeDesign with LinearRegression {
```

To the configuration options of the experiment we add the specification of the design.

```
centralComposite("p_drop_contact" <~ interval(0.0, 1.0), "p_info_mingle" <~ interval(
0.0, 1.0), ...)
```

Lastly, the linear regression is applied to the collected simulation results.

```
withExperimentResult { result =>
  val regr = fitLinearModel(result)("p_drop_contact", "p_info_mingle", ...)(migrants)
  println(regr.fittedFunction)
  println(regr.rSquared)
}
```

This is an example of the extensibility of internal DSLs such as SESSL. The `withExperimentResult` block allows injecting arbitrary user code that is invoked when the experiment (all replications of all design points) is finished. Here, we use the function `fitLinearModel` to obtain a regression meta-model `regr` for the observed result, the given factors, and the observable migrants. The fitted function and the r^2 goodness-of-fit measure are written as output.

7.4.3 Reproducibility

In addition to making specifying and executing simulation experiments easier, DSLs can also help to make experiments reproducible (for a general discussion, see Chap. 10). As experiments are typically single files, they can be easily distributed to other researchers, who can then execute the experiments and confirm their results. This way, textual DSLs and, in particular, internal DSLs facilitate packaging experiments in an executable fashion, in contrast to, for example, GUI-based experimentation tools. However, the execution of an experiment requires additional software that must be acquired and installed. SESSL solves this challenge by employing Apache Maven (https://maven.apache.org/), an industry-grade software project management tool, and its associated infrastructure. We give a short summary of the idea below.

Each SESSL experiment is accompanied by a Maven configuration file (called `pom.xml`) that contains details about the software artefacts needed to execute the experiment. Those software artefacts might have their own dependencies, which are automatically resolved by Maven. For example, an SESSL experiment with an ML3 model must only declare its dependency on the SESSL binding for ML3, which in turn depends on the SESSL core and the ML3 simulation package. To execute an experiment, Maven checks whether all dependencies are already installed and, if not, downloads and installs all missing software artefacts automatically. Thus, these downloads are only necessary for the first execution of the experiment. An example of packaging an experiment this way is the SESSL-ML3 quickstart package, which is available from https://git.informatik.uni-rostock.de/mosi/sessl-ml3-quickstart.

7.4.4 Related Work

Using a tailored language to specify simulation experiments was pioneered by the 'Simulation Experiment Description Markup Language' (SED-ML) (Waltemath et al., 2011). SED-ML aims at computational biology and, being based on XML, is a machine-readable rather than human-readable language. In contrast to SESSL, where experiments are executable standalone artefacts, SED-ML is an exchange format for experiments that can be written and read by tools in the computational biology domain.

In the area of agent-based simulation, some tools support simple experiments. Repast Simphony, for example, provides an interface for 'Batch Runs,' which are simple parameter sweeps (Collier & Ozik, 2013); Netlogo's BehaviorSpace module (Wilensky, 2018) enables parameter sweeps as well. Both approaches allow importing and exporting experiments as XML files. In contrast to SED-ML, however, these XML files are tool-specific and cannot be used to port an experiment from one tool to another. More complex experiments can be implemented by writing code that generates such files. For example, this approach has been used to apply

Simulated Annealing (an optimisation algorithm) to a Repast Simphony model (Ozik et al., 2014). More recently, an R package with a DSL-like interface has been published that implements complex experiments by generating XML files for NetLogo (Salecker et al., 2019).

To gain more independence from concrete tools, simulation experiments can also be represented in a more abstract form, for example in *schemas* (Wilsdorf et al., 2019). Such a schema describes a machine-readable format of the salient aspects of a simulation experiment, which can then be used to (semi-) automatically generate representations of that experiment in concrete tool formats.

7.4.5 Discussion

Using DSLs emphasises the role of simulation experiments as standalone artefacts. Experiments and their parts can be composed and reused largely independently of a concrete simulation model, as they are defined in their own DSL. The DSL implementation is then responsible for executing a given experiment specification for a given model. In other words, DSLs for simulation experiments allow separation of the concerns of developing a model on the one hand, and designing experiments for a model on the other.

One central advantage of DSLs for simulation experiments is the potential for reuse. First, it becomes possible to reuse components of simulation experiments and compose new experiments from them. This is particularly useful when applying complex experimental methods to a simulation model, as these methods can be implemented based on an experiment abstraction that represents the commonalities of all simulation systems. By mapping a concrete simulation system to this abstraction, as SESSL's bindings do, all methods become applicable. But the term 'reuse' can also refer to complete experiments. One relevant example is conducting the same experiment with two different implementations of a model or two different models of the same phenomenon. By confirming that the results from both experiment executions match, the models can be *cross-validated*.

Finally, expressing simulation experiments with DSLs also facilitates capturing the role of experiments and their relation to simulation models in the course of a simulation study, which is studied in the following section by using the concept of formal provenance modelling.

7.5 Managing the Model's Context

Understanding how the data and theories have entered the model-generating process is central for assessing a simulation model, and the simulation results that are generated based on this simulation model. This understanding also plays a pivotal role in the reuse of simulation models, as it provides valuable information as to for which applications a given model might be valid.

Documentation of agent-based models has been standardised in the ODD protocol (Overview, Design concepts, Details; see Grimm et al., 2006), which is regularly applied in many fields, including the social sciences (Grimm et al., 2020). However, ODD only includes small parts of the wider context, how a simulation model has been generated, mostly in the 'purpose' and 'input data' elements. Some more information (especially on analysis) is included by TRACE (Schmolke et al., 2010; Grimm et al., 2014), which, when applied to an agent-based model, might include an ODD documentation of the model itself. Both of these approaches rely on extensive textual descriptions, which might easily add up to 30 pages (see, e.g., Klabunde et al., 2015).

Instead of textual description, we propose a more formal approach, i.e., using PROV (Groth & Moreau, 2013), which represents a provenance standard, to describe how a simulation model has been generated (Ruscheinski & Uhrmacher, 2017). *Provenance* refers to "information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness" (Groth & Moreau, 2013).

PROV represents provenance information as a directed acyclic graph. This graph contains different types of nodes, including *entities* (shown as circles), e.g., data, theories, simulation model specifications, or simulation experiment specifications, and *activities* (shown as squares), such as calibration, validation, analysing, refining, or composing. Edges represent relationships between nodes, the most prominent ones being *used by* and *generated by*. For example, the entities simulation model and data may be used by the activity calibration, and as a result, a calibrated simulation model as well as an experiment specification be generated by this activity. DSLs do not need to be executable, and in fact PROV is not; however, it allows for storage of the information in a structured manner in a graph database and consequently, for it to be queried.

In this way, the analyst can query, for instance, which data have been used for validating or calibrating a particular model, or retrieve all validation experiments that have been executed with simulation models and upon which a particular simulation model is based. If DSLs, such as ML3, are used for specifying the simulation model, and other DSLs, such as SESSL, are used for specifying the simulation experiments, then these simulation experiments can be reused for future model versions (Peng et al., 2015) and may be re-executed automatically (Wilsdorf et al., 2020). Besides, provenance information can be stored and retrieved at different levels of detail (Ruscheinski et al., 2019). We illustrate this based on the Routes and Rumours model.

Figure 7.2 shows an example of a provenance graph, based on Box 5.1 in Chap. 5. It describes in detail how a sensitivity analysis was conducted. The provenance graph begins with the Routes and Rumours model, as defined in Chap. 3, on the very left (*M*). For the purpose of this example, we omit the process of the model creation, and the entities on which it is based. At first, as described in the second paragraph in Box 5.1, a Definitive Screening Design was applied on the 17 model parameters,

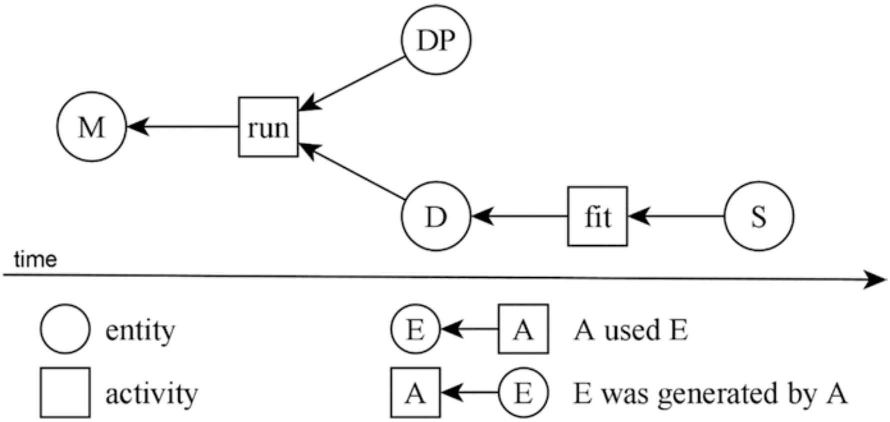


Fig. 7.2 Provenance graph for model analysis based on Box 5.1 in Chap. 5. (Source: own elaboration)

and simulation runs were performed on the 37 resulting design points. We model these two steps as a single process (*run*), which generated two entities: the design points (*DP*) produced in the design step, and the data produced by the simulation runs (*D*).

Subsequently, GP emulators were fitted to the data in the next step (*fit*), yielding the emulators and the information about sensitivity they contain (*S*) as a result. If this was conducted using a DSL such as SESSL (see Sect. 7.4), or even a general-purpose programming language, the processes (*run*) and (*fit*) would have yielded the corresponding code as additional products, which would appear as additional entities, and could be used to easily reproduce the results. However, the analysis was performed with GEM-SA, a purely GUI-based tool, so there is no script, or anything equivalent.

Figure 7.3 (see Appendix E for details) shows a broader view of the whole modelling process in less detail, including multiple iterations of models (Mi), their analysis, psychological experiments, and data assessment. The whole analysis shown in Fig. 7.2 is then folded into the process $a1$, the first step of the broader analysis of the Routes and Rumours model. The analysis shown above uses that model ($M3$) as an input, and produces sensitivity information as an output ($S1$). The process is additionally linked to the methodology proposed by Kennedy and O'Hagan (2001), denoted as ($K01$), and thereby indirectly related to the later steps of the process, in which a similar analysis is repeated on subsequent versions of the model.

To give the provenance graph meaning, appropriate information about the individual entities and activities must be provided. The type of entity or activity determines what information is necessary. That might be a textual description (e.g., ODD for models, or a verbal description of the processes as in Box 5.1), code (potentially in a domain-specific language), or the actual data and relevant meta-data for data-entities. In our case, to provide sources of this information, in Appendix E we mostly refer to the appropriate chapters and sections of this book.

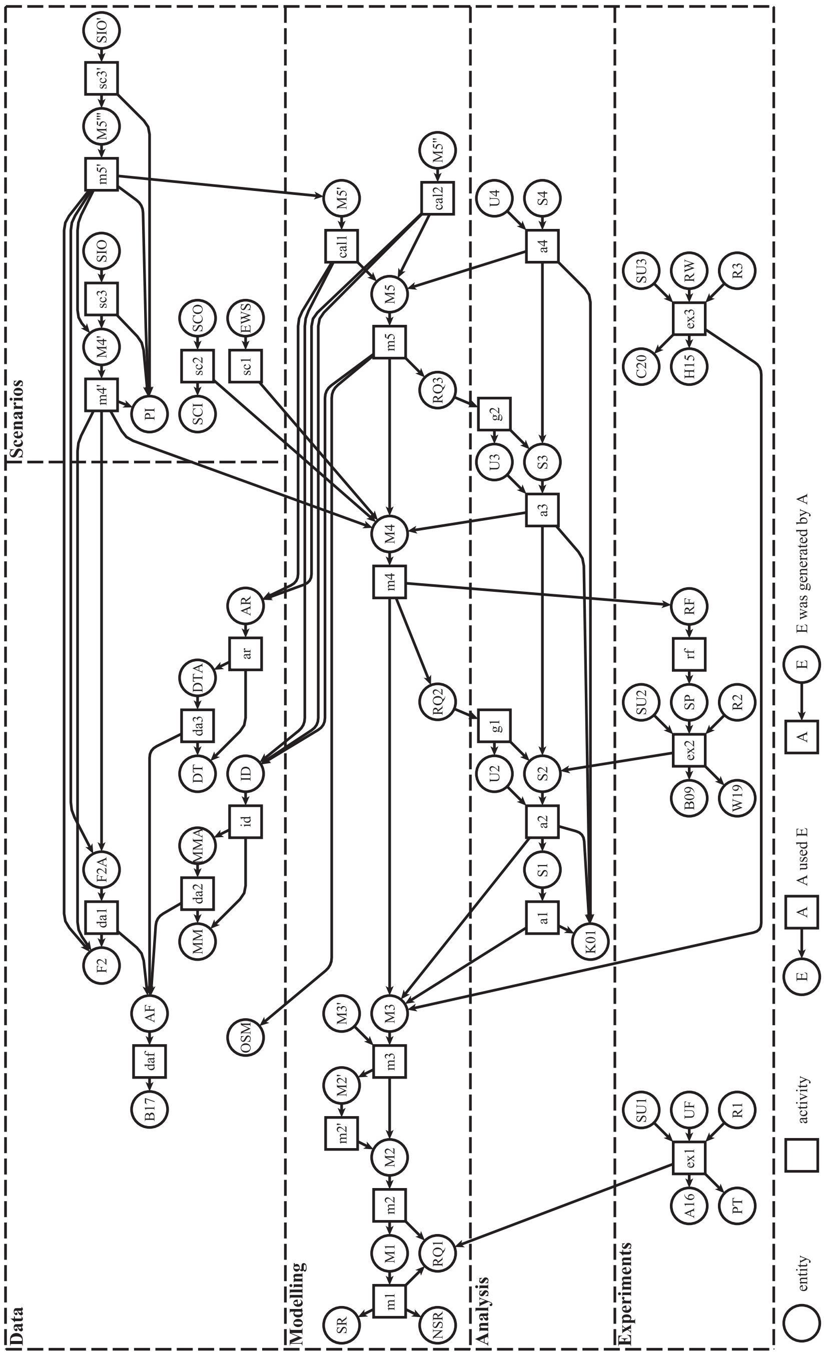


Fig. 7.3 Overview of the provenance of the model-building process – for details, see Appendix E. (Source: own elaboration)

Of course, as a natural extension, a provenance model may also span multiple simulation studies on related subjects, relating current research to previous research, for example if a model developed in one study reuses parts of a previous model (Budde et al., 2021). For this purpose, standardised provenance models included in model repositories such as CoMSES/OpenABM can be used.

7.6 Conclusion

Conducting a complex simulation study is an intricate task, in which a variety of different concerns have to be considered. We have identified some of the central ones, i.e., specifying a simulation model, executing simulation runs, conducting complex simulation experiments, and documenting the context and history of a simulation model, and demonstrated how domain-specific languages can be employed to tackle them separately. A domain-specific modelling language allows for a succinct model representation, making use of suitable metaphors. With the application of the ML3 to the Routes and Rumours model, we have demonstrated the value of such metaphors, e.g., ML3's rules to model concurrent processes. At the same time, DSLs put a limitation to the kinds of models that can be expressed. This limitation of expressive power, however, has benefits for the execution of simulation runs, in that limitations allow for more efficient simulation algorithms. A DSL that is too powerful for its purpose might hence be equally impractical. This highlights an important trade-off for selecting a suitable DSL – and for designing such a language in the first place. DSLs for simulation experiments allow the specification of such experiments in a readable and succinct way. Such executable experiment specifications may then be shared and reused, improving reproducibility of results.

Finally, PROV, a graph-based language for provenance modelling, allows the specification of a model's history and context in a way that is accessible to both human readers and computational processing. This is especially important for creating and documenting subsequent model versions as part of the iterative process advocated throughout this book, including several different elements, such as model versions, languages and formalisms used, empirical and experimental data, elements of analysis (meta-modelling and sensitivity) and their results, and so on. The creation of such model is presented in Chap. 8, and the role of individual elements in the whole model-building process, as well as its scientific and practical implications, are discussed throughout Part III of the book.

Part III
Model Results, Applications, and Reflections

Chapter 8
Towards More Realistic Models

Martin Hinsch, Jakub Bijak, and Jason Hilton

This chapter is devoted to the presentation of a more realistic version of the model, Risk and Rumours, which extends the previous, theoretical version (Routes and Rumours) by including additional empirical and experimental information following the process described in Part II of this book. We begin by offering a reflection on the integration of the five elements of the modelling process, followed by a more detailed description of the Risk and Rumours model, and how it differs from the previous version. Subsequently, we present selected results of the uncertainty and sensitivity analysis, enabling us to make further inference on the information gaps and areas for potential data collection. We also present model calibration for an empirically grounded version of the model, Risk and Rumours with Reality. In that way, we can evaluate to what extent the iterative modelling process has enabled a reduction in the uncertainty of the migrant route formation. In the final part of the chapter, we reflect on the model-building process and its implementation.

8.1 Integrating the Five Building Blocks of the Modelling Process

The move from a data-free, theoretical agent-based model to one that represents the underlying social processes and reality more closely, requires making advances in all five areas presented in Part II of this book. The model itself needs to be further developed to answer more specific research questions in a more realistic scenario, the data and experimental information need to be collected, ideally guided by the statistical analysis where possible, and the modelling language and formalism need to be chosen so that they serve the new modelling aims and purposes.

In the context of the migration model presented in this book, we have therefore set out to create a more realistic version of the simulation of the migration routes into Europe. To make the model better resemble real-life scenarios, the notion of personal risk was introduced into the modelled world – in this case, the chance of

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_8

not being able to make it safely to the destination and, in extreme cases, of perishing along the way. This was intended to align the scenario more closely with the sad reality of the deadly maritime crossings from North Africa and Turkey into Europe, especially via the Central Mediterranean route, where at least 17,400 people have perished between 2014 and January 2021 – a majority of the more than 21,300 deaths in the whole Mediterranean basin in that period[1] (Frontex, 2018; IOM, 2021, see also Chap. 4).

In particular, by extending the model and its purpose, we were interested in investigating whether our model could be used to test the claim – which was made by some parties within the EU – that an increased risk on the Mediterranean would lead to a decrease in 'pull factors' of migration and thus a decrease in the number of arrivals (for a critical discussion of this idea, see e.g. the *Death by Rescue* report by Heller and Pezzani 2016, as well as other studies, overviews and briefs, *such* as Cusumano & Pattison, 2018; Cusumano & Villa, 2019; and Gabrielsen Jumbert, 2020). This is the type of research question that does not necessarily imply predictive capabilities in a simulation model, but rather seeks to illuminate the mechanisms and trade-offs involved in the interplay between risk, information, communication, and decisions.

In our case, the starting point for the model extension was the theoretical Routes and Rumours model, presented in Chap. 3 and Appendix A. Each of the subsequent building blocks – the empirical data, statistical analysis, psychological experiments, and the discussion around the choice of an appropriate programming language – as well as the changes made to the model itself as it was further developed to serve the purpose, were then used to augment the simulated reality in the light of the knowledge that became available as the modelling process unfolded.

Of course, as discussed before, identifying the empirical basis for the model proved challenging. Of the many different data sources on asylum migration discussed in Chap. 4 and Appendix B, only a handful were directly applicable to the new version of the model, and of those, only a couple ended up being used. The potentially applicable sources concentrated mainly on the process data on registered arrivals in Europe, (uncertain) risk-related data on the deaths in the Mediterranean, and survey-based indications of the sources of information used by migrants along the way (see Box 4.1).

The statistical analysis discussed in Chap. 5 served as a way of focusing the model on the most important aspects of the route dynamics, while at the same time allowing its development in other areas. To that end, the key findings regarding the sensitivity of the model outputs to a small set of information-related variables enabled us to concentrate on the key defining features of the underlying social mechanisms driving route formation, which in this case was focused on information exchange. At the same time, as was expected given the nature of migration

[1] The relative risk of death is also far higher on the Central Mediterranean route than elsewhere: the minimum estimates suggest the risk of dying of 2.4% in 2016–19 (confirmed deaths and disappearances to attempted crossings), as compared to 0.4% on the other Mediterranean routes: Eastern and Western – a six-fold difference (IOM, 2021).

processes, the levels of uncertainty surrounding the modelled route formation and the impact of its drivers (via model parameters), remained high – and higher than in the Routes and Rumours model.

On the one hand, the results of the statistical analysis carried out on the first, theoretical version of the model (Routes and Rumours), helped therefore delineate the possible uses of the psychological experiments in enhancing the simulation. In particular, the design of the second set of experiments discussed in Chap. 6, looking at the attitudes to risk and eliciting subjective probabilities of a safe journey depending on the source of information, was directly informed by both the model design and sensitivity analysis reported above. The data from this experiment were then directly used in informing the way the agents respond to different types of information in the current model version.

On the other hand, the choice of a modelling language also influenced the model-building, albeit indirectly. Despite the model development continuing in a general-purpose programming language (Julia) rather than a domain-specific one (ML3), the new version as described in Chap. 3 includes some aspects of the model formalism and semantics, uncovered through parallel implementation in both languages (Reinhardt et al., 2019). This mainly relates to using the continuous definition of time and to modelling of events through the waiting times, as recommended in Chap. 7. At the same time, the provenance description of the model helped understand the mechanics of the modelling process itself, and offered a more systematic way in which to extend the first version of the model.

Throughout the remainder of this chapter, we present the results of following the modelling process discussed before, in the form of a more realistic and empirically grounded, yet still explanatory rather than predictive model of migration route formation. In comparison with Routes and Rumours, the focus goes beyond the role of information and choice between different options under uncertainty, and now additionally includes risk and risk avoidance, with potentially very serious consequences for the agents. We discuss the motivation for the specific elements of the construction of the resulting Risk and Rumours model, as well as a detailed description of its constituting parts next.

8.2 Risk and Rumours: Motivation and Model Description

Most of the capabilities required by our model in order to be able to test whether increased risk could lead to a reduction in arrivals were already in place in the Routes and Rumours version, except for one crucial one: the presence of risk, and the rules governing the agents' decisions in relation to risky circumstances, the addition of which was the key feature of the new version, called Risk and Rumours. Other than that, in the previous version the agents already reacted in real (simulated) time to the changes in travel conditions. Here, the continuous time paradigm offers a much more natural environment for framing the process of information flow and belief update, devoid of the artificial constraints imposed by the granularity of time steps and scheduling problems in discrete simulations (Chap. 7). Furthermore, the agents' decisions are based not only on their subjective (and possibly imperfect) knowledge, which could be exchanged with other agents, mediated by the levels of trust, or gained by exploring the environment, but also by different levels of risk and attitudes towards it.

Contrary to the previous version, and to keep the Risk and Rumours model consistent, both internally and with the reality it aims to represent, in this version of the model it is possible for agents to die, which removes them from the simulation entirely. For the sake of simplicity, we assume that the agents can only die when moving across transport links. As with the other processes in the continuous-time version of the model, death happens stochastically at a certain rate. The rate of death for a given link is calculated from a risk value associated with each link that represents the expected probability of an agent dying when crossing that link, and the expected time it takes to cross that link. The death rates can be taken from the empirical data, such as the Missing Migrants project (see Chap. 4), either applied directly as model inputs, or used to calibrate the outputs.

The agents' information on the transport links now also includes corresponding knowledge about risk, which they are able to learn about and communicate in the same way as for the links' friction and other properties of their environment (see Chap. 3). Still, this is the one aspect of the new version of the model that is of crucial importance from the point of view of examining substantive research questions, many of which – implicitly or explicitly – rely on some assumptions about the attitudes of prospective migrants towards risk, and on the decisions taken in this light.

To that end, the risk-based decision making in the current version of the model is directly informed by the empirical experiments on subjective probabilities, risk attitudes and confidence in the ensuing decisions according to the source of information, as described in Sect. 6.3. Here, we used a logistic regression of the (stated) probability of making a decision to travel against the (stated) perceived level of risk, to parameterise a bivariate normal distribution. From this distribution, we draw for each agent individual values for the slope S and intercept I of the logit-linear function mapping the probability of travel, p (as per the experimental setup), and the agent's perceived risk, s. As discussed in more detail in Box 6.1 in Sect. 6.5, the logit of the probability to travel can then be calculated as $p = I + S * s$. In this version of the model the value of p is transformed into a probability, and used as part of the cost calculation on which the agents' path planning is based. For specific details on the calculation of risk functions, including the role of risk scaling factors, see Box 6.1 in Sect. 6.5, as well as the online material referenced in Appendix A.

In terms of the topology of the new version of the model, for simulating the effect of elevated risk we implemented a 'virtual Mediterranean' by keeping the risk at very low levels (0.001) for most links in the world, but increasing it in all links overlapping a rectangular region that ran across half of the width of the simulated area (the red – darker – central area in Fig. 8.1, showing the model topology).

In order to be able to run simulation experiments based on complex pre-defined scenarios such as, for example, policy interventions or changes in the agents' environment over time, we further added a generic 'plug-in' scenario system to the

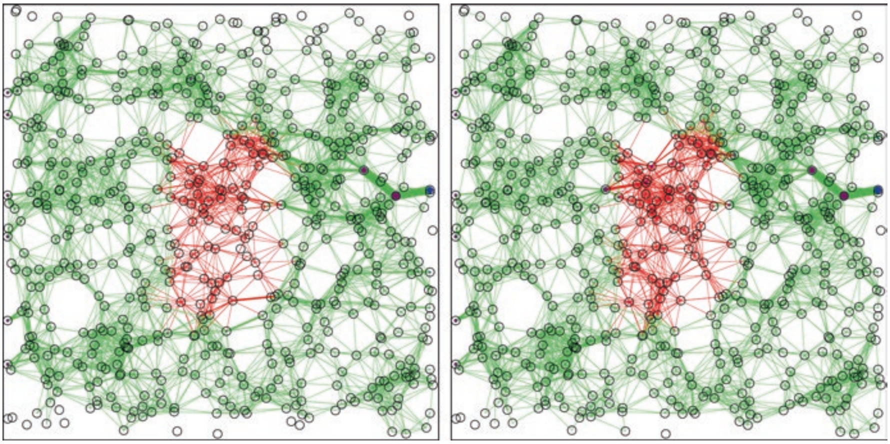


Fig. 8.1 Topology of the Risk and Rumours model: the simulated world with a link risk represented by colour (green/lighter – low, red/darker – high) and traffic intensity shown as line width. In this scenario, cautious agents (left) take traffic routes around the high-risk area, whereas agents exhibiting risky behaviour (right) take the shortest paths, crossing through the dangerous parts of the map**.** (Source: own elaboration)

model. This makes it possible to load additional code during the runtime of the simulation that, for example, changes the values of some parameters at a pre-defined time, or occasionally modifies the properties of some parts of the simulated world.

Examples of policy-relevant simulations generated by this model are described in more detail in Chap. 9. Their implementation required three such 'plug-in' scenario modules: two of them simulate simple changes in the external conditions of departures (the migrant generating process) and travel conditions, namely a change in departure rate at a given time, and change in the level of risk in the high-risk area at a given time. The third module simulates a government information campaign to make migrants aware of the high risk of crossing a dangerous area (here, our virtual Mediterranean) under varying levels of trust in official information sources informed by the Flight 2.0/Flucht 2.0 survey (see Box 4.1 in Sect. 4.5, and Appendix B for source details), as well as by the psychological experiment on eliciting subjective probabilities, reported in Chap. 6 (Sect. 6.2).

In this module, the information campaign has been implemented by introducing a simulated 'government agent' who has full knowledge concerning the high-risk area, who then interacts with a certain probability with agents present in the entry cities (see Appendix A). If an interaction takes place, the migrant agent in question exchanges information with the government agent analogous to the information exchange happening during regular agent contacts, albeit with modified trust levels.

In addition to providing insights into the topology of the modelled world, Fig. 8.1 offers some preliminary descriptive findings about the role of risk and risk attitudes, based on a single model run. In this example, the agents are on average either more or less risk-taking, which is in line with the qualitative findings of the first cognitive experiment, on eliciting the prospect curves (Sect. 6.2). These differences in attitudes to risk have a clear impact on the number of journeys undertaken by agents through the high-risk area. As expected, the more cautious agents are more likely to attempt travelling around, while in the scenario with higher risk tolerance, the intensity of travel through the high-risk area is visibly elevated. Some further substantive questions, which can be posed within the context of the Risk and Rumours setup, are examined for several policy-relevant scenarios generated by the model, presented in Chap. 9. Before that, however, an important intermediate question is: what is driving the behaviour observed in the model? As discussed in Chap. 5, the uncertainty and sensitivity analysis can offer at least some indications in that respect. We discuss this step of the analysis of the model behaviour next.

8.3 Uncertainty, Sensitivity, and Areas for Data Collection

To analyse the behaviour of the Risk and Rumours model itself, we follow the template from Chap. 5, with a few modifications. To start with, we limit the analysis to four model parameters related to information exchange, which were previously identified as key in Chap. 5 and one parameter related to the speed of exploration of the local environment (*speed_expl*), plus five additional free parameters, not identified from the data, yet crucial for the mechanism of the model. These additional parameters are related to the perceptions of risk, and the detailed list of all ten parameters used for uncertainty and sensitivity analysis is provided in Table 8.1.

Table 8.1 Parameters of the Risk and Rumours model used in the uncertainty and sensitivity analysis

Parameter	Description	Range
p_drop_contact	Probability of an agent losing a contact from their network	[0, 1]
p_info_contacts	Probability of an agent communicating with their own contacts	[0, 1]
p_transfer_info	Probability of exchanging information through communication	[0, 1]
Error	Measure of information error (0: perfect information, 1: full noise)	[0, 1][a]
speed_expl	Speed of taking up information when exploring locally	[0, 1]
risk_scale	Measure of how the chance of survival scales to the perceived safety as measured in the experimental data from Chap. 6	[4, 20]
p_notice_death	Two parameters that determine how likely it is that an agent notices another agent's death and how strongly that affects risk perception	[0, 1]
speed_risk		[0, 1]
speed_expl_risk	A parameter depicting how quickly the perceived risk is updated by local exploration of the environment	[0, 1]
path_penalty_risk	Penalty in terms of additional costs for risk associated with a given stretch of route, relative to movement and resource costs	$[0, \infty)$[b]

Notes: [a]For uncertainty and sensitivity analysis, limited to [0, 0.5] given minimal variability beyond this range. [b]For the analysis, limited to [0, 10] for practical reasons. (Source: own elaboration)

This time, our focus is on two key outputs: the number of arrivals, and the number of drownings, as the ultimate human cost of undertaking perilous migration journeys. Both of these outputs are analysed globally, but can also be looked at as time series of the relevant variables for more specific policy-related questions and for setting up coherent scenarios, as discussed further in Chap. 9.

Given the number of parameters to be studied in this version of the model, there is no need to carry out extensive pre-screening, so the analysis can focus on assessing the uncertainty of the outputs and their sensitivity to the individual model inputs, in order to unravel the dynamics of the system and interactions between its different components. As before, standard experimental design, based on Latin Hypercube Samples, is applied, with 80 design points and five replicates per point.

The main results of the sensitivity and uncertainty analysis of the Risk and Rumours model are reported in Table 8.2. For the two outputs considered – the number of arrivals and the number of deaths – three parameters related to information exchange, introduced in Chap. 5, remain of pivotal importance. The key parameter is the probability of exchanging information through direct communication (*p_transfer_info*), followed by the probability of communicating with an agent's contacts (*p_info_contacts*) and of losing contacts (*p_drop_contact*). From the newly-added parameters, depicting the relationships with risk, the most important are those related to the speed of updating the information about risk (*speed_expl_risk*), and to the mapping between the objective risk of death and its subjective assessment (*risk_scale*). The interactions between these parameters also play a role in shaping both outputs, as shown in Table 8.2.

The mean and variance levels of the expected model outputs indicate that on average, across the whole ten-dimensional parameter space, per each run with 10,000 travelling agents, the model generates nearly 7800 arrivals and 2200 deaths, although with some non-negligible variation. The resulting death rate, of around 22%, is clearly by an order of magnitude higher than would be observed even on a high-risk maritime crossing, such as Central Mediterranean. This suggests that the model needs to be properly calibrated to the empirical data on deaths in order for it to be more representative of the underlying reality of migration journeys. The estimated total variance in the code output translates into standard deviations of nearly 1150 for arrivals and over 650 for deaths, indicating considerable disparities across the whole parameter space. On the other hand, the impact of code uncertainty on the total estimated emulator variance is relatively small: the σ^2 term for the code variability 'nugget' is two orders of magnitude smaller than the overall fitted variance term of the emulator, σ^2. On the whole, the fit of the underlying GP emulator is reasonable, with the root mean squared standardised error (RMSSE) above two for both outputs, somewhat larger than the ideal levels of one, which would indicate that the emulator results are close to the model outputs.

Figure 8.2 illustrates the response surfaces with respect to the two parameters describing the relationship with risk (*risk_scale* and *speed_expl_risk*), over their space of variability defined in Table 8.1, [4, 20] × [0, 1]. The predicted values of the GP emulator, means and standard deviations, are shown for the two outputs: numbers of arrivals and deaths. For simplicity, only the results assuming Normal prior

Table 8.2 Uncertainty and sensitivity analysis for the Risk and Rumours model

Sensitivity analysis				
Input\output	Arrivals		Deaths	
Input prior:	Normal	Uniform	Normal	Uniform
p_drop_contact	**3.006**	**2.851**	**10.700**	**9.130**
p_info_contacts	**6.092**	**4.990**	**15.823**	**16.784**
p_transfer_info	**57.644**	**48.593**	**40.864**	**38.264**
error	0.145	0.176	2.330	**2.712**
speed_expl	0.718	0.564	0.533	0.597
risk_scale	**2.746**	**4.297**	**3.863**	**3.868**
p_notice_death	0.184	0.215	0.138	0.152
speed_risk	0.183	0.212	0.261	0.195
speed_expl_risk	**4.597**	**4.739**	**10.097**	**9.371**
path_penalty_risk	0.991	1.562	0.655	0.542
Interactions	**18.260**	**22.809**	**11.522**	**12.790**
Residual	5.433	8.994	3.215	5.595
Total % explained	**94.567**	**91.006**	**96.785**	**94.405**
Uncertainty analysis (Normal prior)				
Mean of expected code output	7763.92		2236.99	
Variance of expected code output	4608.59		777.78	
Mean total variance in code output	1,315,010		428,657	
Fitted sigma^2	1.3160		1.2289	
Nugget sigma^2	0.0111		0.0193	
Cross-validation (leave 20% out)				
RMSE	152.30		116.33	
RMSPE (%)	67.73%		6.05%	
RMSSE (standardised)	2.5165		2.3836	

The experiments were run on 80 Latin Hypercube Sample design points, with five repetitions per point. The values **in bold** correspond to inputs with visible (>2.5%) shares of attributed variance. (Source: own elaboration in GEM-SA. (Kennedy & Petropoulos, 2016))

distributions of inputs are shown, and the values for the remaining parameters are set at arbitrary, yet realistic values.[2] As can be seen from Fig. 8.2, both outputs show clear gradients along both risk-related parameter dimensions, with arrivals increasing and deaths decreasing with both *risk_scale* and *speed_expl_risk*, and with lower uncertainty estimated for 'middle' values of both parameters than around the edges of the respective graphs.

The results of the sensitivity analysis additionally point to the areas of further data collection, in particular with respect to **information transfers over networks** (parameters *p_transfer_info*, *p_info_contacts*, and *p_drop_contact*), **mapping of**

[2] Here, we assume *p_info_contacts* = *p_transfer_info* = 0.8, *p_drop_contact* = 0.5, *p_info_mingle* = 0.5, *error* = 0.1, *p_notice_death* = 0.8, *speed_risk* = 0.7, and *path_penalty_risk* = 5. Note that as per the outcomes of the sensitivity analysis reported in Table 8.2, only the first three of these parameters really matter.

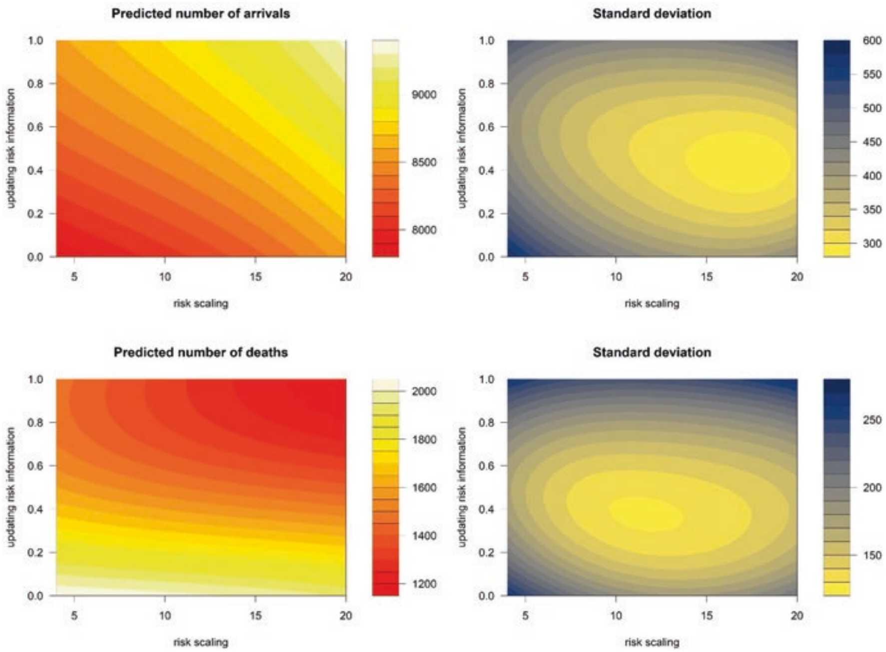


Fig. 8.2 Response surfaces of the two output variables, numbers of arrivals and deaths, for the two parameters related to risk. (Source: own elaboration in GEM-SA, Kennedy & Petropoulos, 2016)

objective and subjective risk measures (*risk_scale*), and the speed of **updating the information about risk through observation** (*speed_expl_risk*). These are the areas where the information gains in the model are likely to be the highest, and at the same time, where the existing evidence base is scarce or non-existent. Here, as discussed in Chap. 6, carrying out the more interactive and immersive cognitive experiments on decision making would bear a promise of producing results that may be less influenced by the respondent bias, which is a concern for respondents with no lived experience of migration, not to mention asylum migration. Setting up such an experiment can additionally be helped by carrying out a dedicated qualitative survey, specifically targeted at asylum seekers and refugees, the results of which would inform the experimental protocol and help manage some ethical issues related to the sensitivity of the topic.

Still, even within the confines of the current model, there is scope for further inclusion of selected data sources, discussed in Chap. 4, in order to make it even closer aligned with the reality the model aims to represent. We discuss these additions, leading to the creation of a new version of the model, called Risk and Rumours with Reality, and the process of calibrating this model to observed data by using Bayesian statistical methods, in the next section of this chapter.

8.4 Risk and Rumours with Reality: Adding Empirical Calibration

As discussed before, during the so-called ‘migration crisis’ following the Arab Spring and the Syrian civil war, attempts to cross the Mediterranean via the Central route, from Libya and Tunisia to Italy and Malta, saw a massive increase (Chap. 4). The European Union reacted to these developments by implementing a ‘deterrence’ strategy, in cooperation with North African states. This strategy relied on making it harder for humanitarian rescue missions to operate in the Mediterranean, while at the same time boosting efforts by coast guards in Libya and Tunisia to intercept asylum seekers’ boats before they could reach international waters. As mentioned before, the available data indicate that between 2015 and 2019 these policy changes could have led to a strong increase in interceptions at the African coast, and also to a greater number of fatalities, especially on the Central Mediterranean route (Frontex, 2018; IOM, 2021; see Sects. 4.2 and 8.1). The concomitant reduction in sea arrivals in Southern Europe, however, seems to indicate that their harrowing humanitarian costs notwithstanding these policy changes at least accomplished their declared goal.

It should be possible to test if this ‘deterrence hypothesis’ is true – that is, whether the effect of deterrence can indeed explain the reduction in the number of arrivals – by using an empirically calibrated model of migration that includes the effects of perceived risk on the migrants’ decisions. A full test of the hypothesis goes beyond the scope of this book; however, in the following discussion we demonstrate the first steps towards such a test, by calibrating the Risk and Rumours model against the refugee situation in the Mediterranean in the years 2016–2019, and thus creating a new version, Risk and Rumours with Reality. Setting up the modelling framework for this exercise involved four additional processes: (1) specifying the topology of the transport network, (2) extracting and assessing data on fatality and interception rates, (3) reassessing the sensitivity of the adjusted model to key parameters, and finally (4) calibrating the parameter values based on the empirical information.

To begin with, to define a geographically-plausible model topology for the network of cities and links between them in the model, we extracted the geographical locations of the most important cities in North Africa, the Levant and on the Turkish coast as well as some important landing points for refugee boats in Italy, Malta, Cyprus and Greece from OpenStreetMaps (using OpenRouteService – source S02 in Appendix B). From the same data source, we calculated travel distances between these locations to be used as a proxy for the friction parameter. The resulting map is shown in Fig. 8.3.

In terms of data for the period 2016–2019, the number of interceptions at the Tunisian and Libyan coasts as well as numbers of presumed fatalities are available from IOM (2021) (see also Chap. 4, with sources 11 and 12 listed and discussed in more detail in Appendix B). Since we do not know the number of departures, we have to infer fatality and interception rates for each year by using arrivals (*idem*) in the corresponding year. For this, we assume that every migrant will attempt

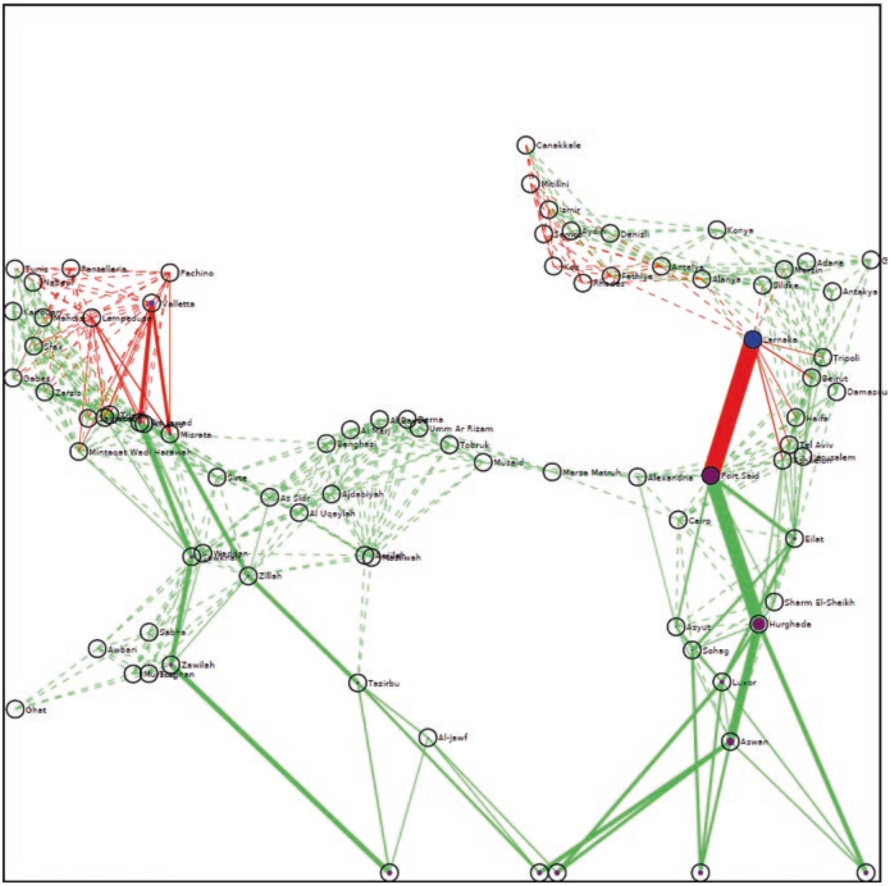


Fig. 8.3 Basic topological map of the Risk and Rumours with Reality model with example routes: green/lighter (overland) with lower risk, and red/darker (maritime) with higher risk. Line thickness corresponds to travel intensity over a particular route for a randomly-selected model run, with dashed lines denoting unused routes. (Source: own elaboration based on OpenStreetMaps)

departure until they either manage to make the crossing, or die. Intercepted migrants wait a certain amount of time and then make another attempt. Based on these assumptions we can estimate the interception probability as $p_i = N_i/(N_i + N_a + N_d)$ and probability of dying as $p_d = N_d/(N_i + N_a + N_d)$, where N_i denotes the number of interceptions, N_a – number of arrivals, and N_d – number of fatalities.

In the third step, we revisited the sensitivity and uncertainty of the revised version of the model to different parameters, with the detailed results reported in Table 8.3. In this iteration of the analysis, there is a noteworthy decrease in the share of the variance explained by individual parameters in comparison with previous model versions. There is also visibly higher impact of the parameter interactions, as well as other, residual factors that drive the model behaviour, which are not yet fully accounted for in the model, such as the changes in the intensity of migrant departures.

Table 8.3 Uncertainty and sensitivity analysis for the Risk and Rumours with Reality model

Sensitivity analysis				
Input\output	Arrivals		Deaths	
Input prior:	Normal	Uniform	Normal	Uniform
p_drop_contact	2.454	**4.413**	**14.361**	**9.539**
p_info_contacts	**7.292**	**9.118**	**4.877**	**5.550**
p_transfer_info	0.855	0.740	0.923	1.094
error	0.781	0.676	2.390	2.499
speed_expl	**2.985**	**4.134**	**7.619**	**4.844**
risk_scale	**3.135**	**4.495**	1.923	1.589
p_notice_death	0.874	0.756	0.688	0.814
speed_risk	0.668	0.578	1.319	1.564
speed_expl_risk	1.589	**2.540**	0.885	1.050
path_penalty_risk	**3.413**	**3.973**	0.575	0.682
Interactions	**34.389**	**39.076**	**64.153**	**51.182**
Residual	41.566	29.502	0.287	19.594
Total % explained	**58.434**	**70.499**	**99.713**	**80.406**
Uncertainty analysis (Normal prior)				
Mean of expected code output	9483.28		179.59	
Variance of expected code output	8311.37		2.27	
Mean total variance in code output	576,153		183.68	
Fitted sigma^2	1.6179		1.0892	
Nugget sigma^2	0.0158		0.3946	
Cross-validation (leave 20% out)				
RMSE	105.786		13.87	
RMSPE (%)	1.15%		9.06%	
RMSSE (standardised)	1.2577		2.4834	

The experiments were run on 80 Latin Hypercube Sample design points, with five repetitions per point. The values **in bold** correspond to inputs with visible (>2.5%) shares of attributed variance. (Source: own elaboration in GEM-SA, Kennedy & Petropoulos, 2016)

To increase the alignment of the model with reality further, by using the three outputs discussed above, N_i, N_a and N_d, we selected a number of parameters that had emerged as being the most important in the sensitivity analysis – such as *path_penalty_risk*, *p_info_contacts*, *p_drop_contact* and *speed_expl* – as well as the two most important parameters determining the agents' sensitivity to risk – *risk_scale* and *path_penalty_risk*. We subsequently calibrated the model using a Population Monte Carlo ABC algorithm (Beaumont et al., 2009) with the rates of change in the numbers of arrivals and interceptions between the years, as well as the fatality rates per year, as summary statistics. The rates of change were used in order to at least approximately get rid of the possible biases identified for these sources during the data assessment presented in Chap. 4 (in Table 4.3), tacitly assuming that these biases remain constant over time. A similar rationale was applied for using fatality rates. Here, the assumption was that the bias in the numerator (number of deaths) and in the denominator (attempted crossings) were of the same, or similar magnitude.

We ran the model for 2000 simulation runs spread over ten iterations, with 500 time periods for each run, corresponding to 5 years in historical time, 2015–19, with the first year treated as a burn-in period. Under this setup, however, the model turned out not to converge very well. Therefore, we additionally included the between-year changes in departure rates to the parameters to be calibrated. With this change we were able to closely approximate the development of the real numbers of arrivals and fatalities for the years 2016–19 in our model (see also Chap. 9).

In parallel, we have carried out calibration for two outputs together (arrivals and interceptions) based on the GP emulator approach, the results of which confirmed those obtained for the ABC algorithm. Specifically, we have estimated the GP emulator on a sample of 400 LHS design points, with twelve repetitions at each point, and 13 input variables, including three sets of departure rates (for 2017–19). The emulator performance and fit were found reasonable, and the results proved to be sensitive to the prior assumptions about the variance of the model discrepancy term (see also Chap. 5).

Selected results of the model calibration exercise are presented in Fig. 8.4 in terms of the posterior estimates of selected model parameters: as for the ABC estimates, we did not learn much about most of the model inputs, except for those related to departures. This outcome confirmed that our main results and qualitative conclusions were broadly stable across the two methods of calibration (ABC and GP emulators), strengthening the substantive interpretations made on their basis. To illustrate the calibration outcomes, Fig. 8.5, presents the trajectories of the model runs for the calibrated period. These two Figs. 8.4 and 8.5 – are equivalent to Figs. 5.7 and 5.8 presented in Chap. 5 for the purely theoretical model (Routes and Rumours), but this time including actual empirical data, both on inputs and outputs, and allowing for a time-varying model response.

In the light of the results for the three successive model iterations, one important question from the point of view of the iterative modelling process is: to what extent

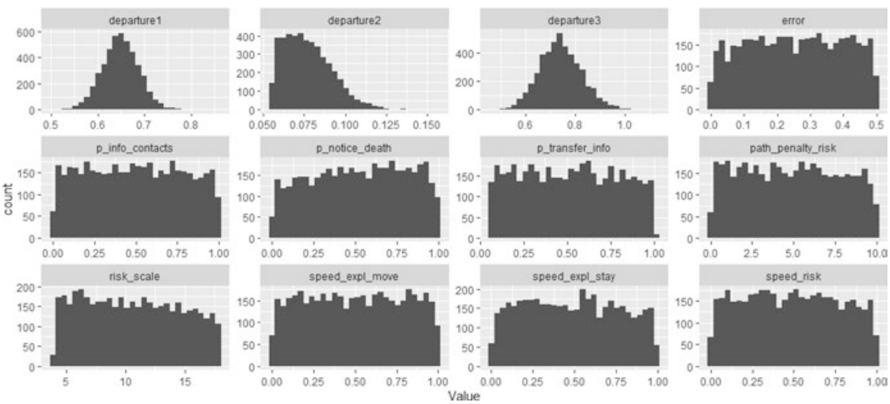


Fig. 8.4 Selected calibrated posterior distributions for the Risk and Rumours with Reality model parameters, obtained by using GP emulator. (Source: own elaboration)

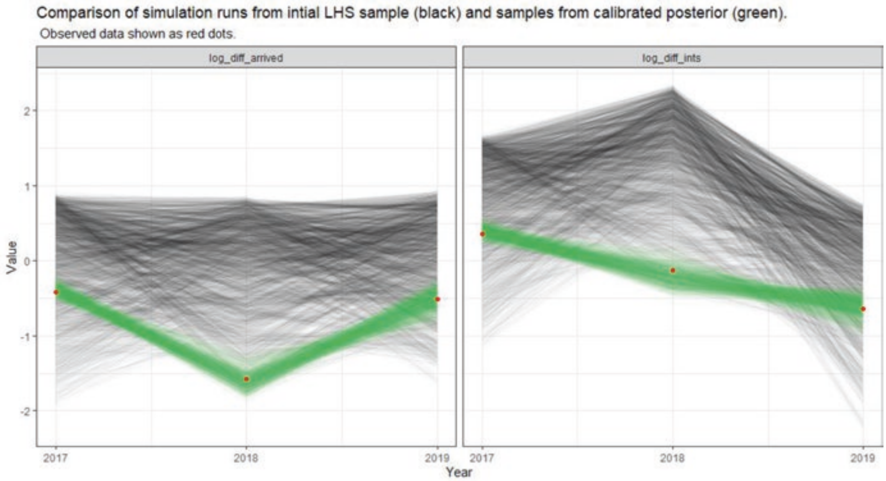


Fig. 8.5 Simulator output distributions for the not calibrated (black/darker lines), and calibrated (green/lighter lines) Risk and Rumours with Reality model. For calibrated outputs, the simulator was run at a sample of input points from their calibrated posterior distributions. (Source: own elaboration)

Table 8.4 Uncertainty analysis – comparison between the three models: Routes and Rumours, Risk and Rumours, and Risk and Rumours with Reality, for the number of arrivals, under Normal prior for inputs

Indicator\Model	Routes & Rumours	Risk & Rumours	Risk & Rumours with Reality
Mean of expected code output	9272.02	7763.92	9483.28
Variance of expected code output	46.41	4608.59	8311.37
Mean total variance in code output	**17,639**	**1,315,010**	**576,153**
Fitted sigma^2	9.4513	1.3160	1.6179
Nugget sigma^2	0.3062	0.0111	0.0158

Source: own elaboration in GEM-SA. (Kennedy & Petropoulos, 2016)

does adding more empirically relevant detail to the model, but at the expense of increased complexity, change the uncertainty of the model output? To that end, Table 8.4 compares the results of the uncertainty analysis for the number of arrivals in three versions of the model: two theoretical (Routes and Rumours and Risk and Rumours), and one more empirically grounded (Risk and Rumours with Reality). The results of the comparison are unequivocal: the key indicator of how uncertain the model results are, the mean total variance in code output (shown in bold in Table 8.4) is by nearly two orders of magnitude larger for the more sophisticated version of the theoretical model, Risk and Rumours, than for the basic one, Routes and Rumours. On the other hand, the inclusion of additional data in Risk and

Rumours with Reality, enabled reducing this uncertainty more than two-fold. Still, the *variance* of the expected code output turned out to be the largest for the empirically informed model version.

At the same time, reduction in the mean model output for the number of arrivals is not surprising, as in Risk and Rumours, *ceteris paribus*, many agents may die during their journey, especially while crossing the high-risk routes. In the Risk and Rumours with Reality version, the level of this risk is smaller by an order of magnitude (and more realistic). This leads to adjusting the mean output back to the levels seen for the Routes and Rumours version, which is also more credible in the light of the empirical data, although this time with a more realistic variance estimate. In addition, the fitted variance parameters of the GP emulator are smaller for both Risk and Rumours models, meaning that in the total variability, the uncertainty related to the emulator fit and code variability is even smaller. In the more refined versions of the model, uncertainty induced by the unknown inputs matters a lot.

Altogether, our results point to the possible further extensions of the models of migrant routes, as well as to the importance of adding both descriptive detail and empirical information into the models, but also to their intrinsic limitations. Reflections on these issues, and on other, practical aspects of the process of model construction and implementation, are discussed next.

8.5 Reflections on the Model Building and Implementation

In terms of the practical side of the construction of the model, and in particular the more complex and more empirically grounded versions (respectively, Risk and Rumours, and Risk and Rumours with Reality), the modifications that were necessary to make the model ready for more empirically oriented studies were surprisingly easy to implement. In part, this was due to the transition to an event-based paradigm which, as set out in Chap. 7, tends to lead to a more modular model architecture.

Additionally, we found that it was straightforward to implement a very general scenario system in the model. Largely this is because Julia – a general-purpose programming language used for this purpose – is a dynamic language that makes it easy to apply modifications to the existing code during the runtime. Traditionally, dynamic languages (such as Python, Ruby or Perl) have bought this advantage with substantially slower execution speed and have therefore rarely been used for time-critical modelling. Statically-compiled languages such as C++ on the other hand, while much faster, make it much harder to do these types of runtime modifications. Julia's just-in-time compilation, however, offers the possibility to combine the high speed of a static language with the flexibility provided by a dynamic language, making it therefore an excellent choice for agent-based modelling.

As concerns the combination of theoretical modelling with empirical experiments, one conclusion we can draw is that having a theoretical model first makes designing the empirical version substantially easier. Only after implementing, running, and analysing the first version of the model (see Chap. 3) were we able to determine which pieces of empirical information would be most useful in developing the model further. This also makes a strong case for using a model-based approach not only as a tool for theoretical research, but also as a method to guide and inspire empirical studies, reinforcing the case for iterative model-based enquiries, advocated throughout this book (see Courgeau et al., 2016).

In terms of the future work enabled by the modelling efforts presented in this book, the changes implemented to the model through the process we describe would also make it easy to tackle larger, empirically oriented projects that go beyond the scope of this work. In particular, with a flexible scenario system in place, we could model arbitrary changes to the system over time. For example, using detailed data on departures, arrivals and fatalities around the Mediterranean (see Chap. 4) as well as the timing of some crucial policy changes in the EU affecting death rates, we would be able to better calibrate the model parameters to empirical data. In the next step, we could then run a detailed analysis of policy scenarios (see Chap. 9) using the calibrated model to make meaningful statements on whether an increased risk does indeed lead to a reduction of arrivals.

Similar types of scenarios can involve complex pattern of changes in the border permeability, asylum policy developments, and either support or hostility directed towards refugees in different parts of Europe between 2015 and 2020. A well-calibrated model, together with an easy way to set up complex scenarios, would allow investigating the effectiveness of actual as well as potential policy measures, relative to their declared aims, as well as humanitarian criteria. An example of applying this approach in practice based on the Risk and Rumours with Reality model is presented in Chap. 9. In addition, the adversarial nature of some of the agents within the model, such as law enforcement agents and migrant smugglers, can be explicitly recognised and modelled (for a thorough, statistical treatment of the adversarial decision making processes, see Banks et al., 2015).

At a higher level, model validation remains a crucial general challenge in complex computational modelling. As laid out in Chaps. 4, 5 and 6, and demonstrated above, the additional data and 'custom-made' empirical studies, coupled with a comprehensive sensitivity and uncertainty of model outcomes, can be a very useful way of directly improving aspects of a model that are known to be underdefined. In order to be able to test the overall validity of the model, however, it ideally has to be tested and calibrated against known outcomes.

One possible way of doing that would entail focusing on a limited real-world scenario with relatively good availability of data. The assumption would then be that a good fit to the data in a particular scenario implies a good fit in other scenarios as well. For example, we could use detailed geographical data on transport topology in a small area in the Balkans, combined with data on presence of asylum seekers in camps, coupled with registration and flow data, to calibrate the model parameters. An indication of the 'empirical' quality of the model is then its ability to track historical changes in these numbers, spontaneous or in reaction to external factors. Given the level of spatial detail that would be required to design and calibrate such models, they remain beyond the scope of our work; however, even the version of the model presented throughout this book, and more broadly the iterative process of arriving at successive model versions in an inductive framework, enables making some conclusions and recommendations for practical and policy uses.

This discussion leads to a more general point: what lessons have we learned from the iterative and gradual process of model-building and its practical implementation? The proposed process, with five clearly defined building blocks, allows for a greater control over the model and its different constituent parts. Analytical (and theoretical) rigour, coherence of the assumptions and results, as well as an in-built process of discovery of the previously unknown features of the phenomena under study, can be gained as a result. Even though some elements of this approach cannot be seen as a purely inductive way of making scientific advances, the process nonetheless offers a clear gradient of continuous ascent in terms of the explanatory power of models built according to the principles proposed in this book, following Franck (2002) and Courgeau et al. (2016).

In terms of the analysis, the coherent description of phenomena at different levels of aggregations also helps illuminate their mutual relationships and trade-offs, as well as – through the sensitivity analysis – identify the influential parts of the process for further enquiries. Needless to say, for each of the five building blocks in their own right, including data analysis, cognitive experiments, model implementation and analysis, as well as language development, interesting discoveries can be made.

At the same time, it is also crucial to reflect on what the process does not allow. The proposed approach is unlikely to bring about much change in a meaningful reduction of the uncertainty of the social processes and phenomena being modelled. This is especially visible in the situations where uncertainty and volatility are very high to start with, such as for asylum migration. This point is particularly well illustrated by the uncertainty analysis presented in the previous section: introducing more realism in the model in practice meant adding more complexity, with further interacting elements and elusive features of the human behaviour thrown into the design mix. It is no surprise then that, as in our case, this striving for greater realism and empirical grounding has ultimately led to a large increase in the associated uncertainty of the model output.

In situations such as those described in this chapter, there are simply too many 'moving parts' and degrees of freedom in the model for the reduction of uncertainty to be even contemplated. Crucially, this uncertainty is very unlikely to be reduced with the available data: even when many data sources are seemingly available, as in the case of Syrian migration to Europe (Chap. 4), the empirical material that corresponds exactly to the modelling needs, and can be mapped onto the sometimes abstract concepts used in the model (e.g., trust, confidence, information), is likely to be limited. This requires the modellers to make compromises, and make sometimes arbitrary decisions, or leave the model parameters underspecified and uncertain, which increases the errors of the outputs further.

These limitations underline high levels of aleatory uncertainty in the modelling of such a volatile process as asylum migration. Even if the inductive model-building process can help reduce the epistemic uncertainty to some extent, by furthering our knowledge on different aspects of the observed phenomena, it also illuminates clearly the areas we do not know about. In other words, besides learning about the social processes and how they work, we also learn about what we do not know, and may never be able to know. Besides an obvious philosophical point, variably attributed to many great thinkers from Socrates to Albert Einstein (*passim*), that the more we know, the more we realise what we do not know, this poses a fundamental problem for possible predictive applications of agent-based models, even empirically grounded.

If simulation models of social phenomena are to be realistic, and if they are to reflect the complex nature of the processes under study, their predictive capabilities are bound to be extremely limited, maybe except for very specific and well-defined situations where exact description of the underlying mechanisms is possible. At the same time, such models allow for knowledge advances in making possible, and furthering the depth and nuance of, theoretical explanations. The process we propose in this book additionally enables the researchers to identify gaps and future research directions, so that the modelling process of a given phenomenon could continue. We discuss some ideas in terms of the possible scientific and policy impacts in the next chapter, with examples based on the current versions of the Risk and Rumours model, both theoretical, and empirically grounded.

Chapter 9
Bayesian Model-Based Approach: Impact on Science and Policy

Jakub Bijak, Martin Hinsch, Sarah Nurse, Toby Prike, and Oliver Reinhardt

In this chapter, we summarise the scientific and policy implications of the Bayesian model-based approach, starting from an evaluation of its possible advantages, limitations, and potential to influence further scientific developments, policy and practice. We focus here specifically on the role of limits of knowledge and reducible (epistemic), as well as irreducible (aleatory) uncertainty. To that end, we also reflect on the scientific risk-benefit trade-offs of applying the proposed approaches. We discuss the usefulness of proposed methods for policy, exploring a variety of uses, from scenario analysis, to foresight studies, stress testing and early warnings, as well as contingency planning, illustrated with examples generated by the Risk and Rumours models presented earlier in this book. We conclude the chapter by providing several practical recommendations for the potential users of our approach, including a blueprint for producing and assessing the impact of policy interventions in various parts of the social system being modelled.

9.1 Bayesian Model-Based Migration Studies: Evaluation and Perspectives

Following the Bayesian model-based approach in the context of modelling a route network of asylum migration has led to some specific scientific conclusions, reported in Chap. 8, but equally has left several gaps remaining and open for further enquiry. In this section, we look at the contributions in the areas of modelling, data evaluation, psychological experiments, and computing and language development, and the perspectives for enhancing them through more research in specific domains.

In substantive terms, our modelling work suggests that the migrant journey itself – which has received only sparse treatment in migration literature so far – is an important part of migration processes. We were able to show that the dynamics of the uptake and transfer of information by migrants strongly affects the emergence of migration routes. Based on this work, we can also pose specific empirical

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_9

questions concerning migration itself, but also with respect to human behaviour more generally, that will substantially improve our ability to model and understand social systems. At the same time, we can utilise different types of data (micro and macro, quantitative and qualitative, contextual and process-related) in a way that explicitly recognises their quality and describes uncertainty to be included in the models. This is especially important given the paucity of data on such complex processes as migration: here, a formal audit of data quality, as presented in Chap. 4, is a natural starting point.

Still, large gaps in available empirical knowledge of migration remain, which makes any kind of formal modelling challenging. For one, data on many processes that are known to be important are missing or sparse, especially at individual level. Even with a case study such as the recent Syrian asylum migration, there are parts of the process with little or no data, and the data that exist rarely measure specifically what the modellers may want them to. The challenge is to identify and describe the limitations of the data while also identifying how and where they may be useful in the model, and to make consistent comparisons across a wide range of data sources, with a clearly set out audit framework.

More fundamentally, however, we often do not even know which of the possible underlying processes occur in reality, and even if they do, how they affect migration. Besides, human behaviour is intrinsically hard to model, and not well understood in all the detail. Finally, the combination of a large spatially distributed system with the fact that imperfect spatial knowledge is a key part of the system dynamics leads to some technical challenges, due to the sheer size of the problem being modelled.

One key piece of new knowledge generated from the psychological experiments thus far is that migration decision making deviates from the rationality assumptions often used. We found that people exhibit loss aversion when making migration decisions (they weight losses more heavily than gains of the same magnitude), as well as that people show diminished sensitivity for gains in monthly income (i.e., they are less responsive to potential gains as they get further from their current income level). We have also found that people differentially weight information about the safety of a migration journey depending on the source of the information. Specifically, this information seems to be weighted most strongly when it comes from an official organisation, while the second most influential source of information seems to be other migrants with relevant personal experience.

When conducting cognitive experiments and adding greater psychological realism to agent-based models of migration, several important obstacles remain. One key challenge is how to simulate complex real-world environments within the confines of an online or lab-based experiment. Migration decisions have the potential to change one's life to a very large extent, be associated with considerable upheaval, and, in the case of asylum migration, occur in life-threatening circumstances. For ethical reasons, no lab-based or online experiment can come close to replicating the real-world stakes or magnitude of these decisions. This is a major challenge for both designing migration decision-making experiments and for applying existing insights from the decision-making literature to migration. Another important challenge is that migration decisions are highly context dependent and influenced by a huge number of factors. Therefore, even if it were possible to gain insight into specific aspects of migration decision making, important challenges would remain: establishing the extent to which these insights are applicable across migration decision-making contexts, and understanding and/or making reasonable assumptions about how various factors interact.

In terms of computation, the languages we developed show that the benefits of domain-specific modelling languages (e.g., separation of model and simulation, easy to implement continuous time), that are already known in other applications domains (such as cell biology), can also apply to agent-based models in the social sciences. The models gradually developed and refined in this project, and other models of social processes intended to give a better understanding of the dynamic resulting from individual behaviour, have a strong emphasis on the agents' knowledge and decision making.

However, modelling knowledge, valuation of new information, and decision making requires much more flexible and powerful modelling languages than the ones typically used in other areas. For example, we found that the modelling language needs to support complex data structures to represent knowledge. As the resulting language would share many features of general-purpose programming languages, it should be embedded into such a general-purpose language, rather than be implemented as an external domain-specific language.

In addition, our parallel implementation of the core model in two different programming languages demonstrated the value of independent validation of simulation code. To understand and evaluate a simulation model, it is not enough to know how it works; it is also necessary to know why it is designed that way. Provenance models can supplement (or partially replace) existing model documentation standards (such as the ODD or ODD+D protocols, the '+D' in the latter referring to Decisions, Müller et al., 2013; Grimm et al., 2020; see also Chap. 7), showing the history and the foundations of a simulation model. This is especially pertinent for those models, such as ours, which are to be constructed in an iterative manner, by following the inductive model-based approach.

At the same time, the key language design challenge for this kind of models seems to be finding a way to design the language in such a way that it is:

- powerful and flexible enough;
- easy to use, easy to learn and (perhaps most importantly) easy to read; and
- possible to execute efficiently.

For the provenance models, a key challenge is to identify the entities and processes that need to be included, and the relevant meta-information about them. Some of this is common to all simulation studies, independent of the modelling method or the application domain. At the same time, other aspects are application-specific (e.g., certain kinds of data are specific to demography, or to migration studies, and some information specific to these types of data is relevant). This meta-information can be gathered with the help existing documentation standards, such as ODD, which additionally underscores the need for a comprehensive data and data quality audit, as outlined in Chap. 4.

9.2 Advancing the Model-Based Agenda Across Scientific Disciplines[1]

Based on the experience with interdisciplinary model development, and building on the list of outstanding challenges identified in the previous section, we can make some tentative predictions on how model-based approaches and their components may develop in the future.

In terms of migration modelling as such, the further developments are likely to happen in a number of key areas. At this point any modelling effort is necessarily limited by the availability of empirical knowledge in the most general sense – data and other information alike. This means that models have to be either purely conceptual, exploring generic dynamics of the system without specific relation to a concrete real-world scenario, or great effort has to be invested into correctly identifying the uncertainty of model results. However, it is worth noting that statistical models, such as those from the Bayesian uncertainty quantification toolbox, can help shed light even on the behaviour of purely conceptual or theoretical models, without any empirical data, through uncertainty and sensitivity analysis.

The analysis of model results does not at present rely on a standard toolkit of approaches, but on the various methods of uncertainty quantification and emulation, such as those presented in Chap. 5, all of which offer substantial promise. The exploration of the model space can additionally involve tools of artificial intelligence, such as neural networks, especially when the more traditional methods, such as GP emulators, do not work very well, for example in the presence of tipping points or phase transitions between different model regimes. Here, more work needs to be carried out on comparing the results, applicability, and trade-offs of using different meta-models for analysis.

A large part of future progress in modelling migration – or other social systems – depends therefore on improvements in our empirical understanding of the processes under study. Methodologically, it seems promising to try to better understand how the empirical uncertainty in the data and other information leads to uncertainty in modelling results. More fundamentally, we do not have at this point a good understanding of the limits as well as the potential of modelling social phenomena in general. This is an area that will hopefully see increased activity in the future.

When it comes to data, a more tailored application of empirical information to different settings and scenarios is needed, with different uses in mind. Recognition that different data sources are more or less important or useful depends on what is being modelled, and on the research questions or policy objectives of users. Data inventories and formal quality assessments offer a starting point, informing the modellers and users what information is available, but also – perhaps even more

[1] This section includes additional contributions by participants of the workshop “Modelling Migration and Decisions”, Southampton, 21 January 2020. Many thanks go to André Grow, Katarzyna Jaśko, Elzemiek Kortlever, Eric Silverman, and Sarah Wise for providing the voices in the discussion.

importantly – which knowledge gaps remain. At the moment, there is still untapped potential with using digital trace data, for example from mobile phones or social media, to inform modelling. Of course, such data would need to come not only with proper ethical safeguards, but also with knowledge of what they actually represent, and an honest acknowledgement of their limitations.

As the data inventory grows and the quality assessment framework is applied to different settings, the criteria for comparison may be applicable more consistently. For example, it is easier to assess the relative quality of a particular type of source if a similar source has already been assessed. On the whole, the data assessment tools may also be used to identify additional gaps in available data, by helping decide which data would be appropriate for the purpose and of sufficient quality, and therefore can inform targeted future data collection. The quality assessment framework can also encourage the application of rigorous methods of data collection and processing before its publication, in line with the principles of open science.

Besides any statistical analysis, the use of empirical data in modelling can involve face validity tests of the individual model output trajectories, which would confirm the viability of individual-level assumptions. This approach would provide confirmation, rather than validation, of the model workings, and that the process of identifying data gaps and requirements could be iterative. At a more general level, having specific principles and guidelines for using different types of individual data sources in modelling endeavours would be helpful – in particular, it would directly feed into the provenance description of the formal relationships within the model, in a modular fashion. There is a need for introducing minimum reporting requirements for documentation, noting that the provenance models discussed in Chap. 7 are in fact complementary, rather than competing with narrative-based approaches, such as the ODD(+D) protocols (Müller et al., 2013; Grimm et al., 2020).

With cognitive experiments for modelling, one key area for future advancement is the development of experimental setups that reduce the gap between experiments and the real-world situations they are attempting to investigate. The more immersive and interactive experiment suggested in Chap. 6 would attempt to advance experimental work on decision making in this direction, and we expect that future work will continue to develop along these lines. Additionally, it will be crucial for future experimental work to examine the interplay of multiple factors that influence migration decisions simultaneously, rather than focusing on individual factors one at a time.

As also mentioned in Chap. 6, another key challenge is how to map the data from the experimental population to a specific population of interest, such as migrants, including asylum seekers or refugees. The external validity of the experiments, and their capacity for generalisation, is especially important given the cultural and socio-economic differences between experiment participants. One promising possibility, subject to ethical considerations, consists in 'dual track' experimentation on different populations at the same time, to try to estimate the biases involved. This could be done, for example, via social media, targeting the groups of interest, and comparing the demographic profiles with the samples collected by using traditional methods.

Furthermore, necessary psychological input on the structures of decision making to be used in the modelling process can be offered by formal description frameworks, such as the belief-desire-intention (BDI) model of Rao and Georgeff (1991), augmented by additional formal models for memory, information exchange, and so on. For migration and similar problems (mobility, relocations, evacuations…), modelling the decision processes for 'stayers' can be as important as for 'movers', and thus the information on perceived needs and expectations of both groups is key.

In addition, more detailed theoretical work and structured analysis of the already existing literature are also expected to play a key role in improving our knowledge of complex migration decision making. There is a strong need to combine and integrate existing findings from a range of application areas and scientific disciplines, in order to form a more cohesive understanding of the individual and combined impact of various factors on migration decision making (Czaika et al., 2021), and enhance our overall comprehension of the processes involved.

Finally, in computational terms, while we can demonstrate the advantages of the developed domain-specific language, it is hardly possible to create a generic tool that can be readily used by a wider modelling community within a range of large projects, like the one presented throughout this book. Preparing tools, documentation, teaching of the language, and so on, are all very long-term, community-based efforts. One approach to make the developed methods more available for a wider group of users could be to try to include them (or parts of them) into existing tools for agent-based modelling, such as NetLogo, Repast, or Mesa, for example in a form of add-ons for such tools.

As for the practicalities of modelling, one important feature of domain-specific languages is that, despite their being to some extent restricted by construction, they enable the separation of the model logic – the formal description of the model and the underlying processes – from the logic of the programming language. Internal domain-specific languages, embedded as libraries in well-known general-purpose languages, such as Julia, Java or Python, offer a sound compromise solution.

In terms of provenance, future work could lie in automating the provenance modelling in order to aid the modellers in the process. Creating a detailed provenance model, while valuable, can be a demanding and very time-consuming endeavour. To overcome that, provenance information could be, for example, already annotated in the model code, with references to the theory or data sources underling a specific model component, and a provenance model (or at least a part of it) could then be automatically constructed from those annotations.

At a more general level, there are some important implications of our approach for the art and science of modelling. First, while different models can serve different purposes (Epstein, 2008), they are very useful for expanding the imagination of modellers and users alike and for framing the conversation around the processes and systems they are trying to represent. The act of formal modelling forces the assumptions, concepts, and outcome measures to be made and operationalised explicitly, which is already an important step in the direction of fuller transparency and more robust science.

Second, no canonical modelling approaches for social processes exist, or can exist, given the complex and context-dependent nature of many aspects of the social realm. Still, having a catalogue of models, and possibly their individual submodules, can offer future modellers a very helpful toolbox for describing and explaining the mechanisms being modelled. At the same time, the modellers need to be clear about the model epistemology and limitations, and it is best when a model serves to describe one, well-defined phenomenon. In this way, models can serve as a way to formalise and embody the "theories of the middle range", a term originally coined by Merton (1949) to denote "partial explanation of phenomena … through identification of core causal mechanisms" (Hedström & Udehn, 2011), and further codified within the wider Analytical Sociology research programme (Hedström & Swedberg, 1998; Hedström, 2005; Hedström & Ylikoski, 2010).[2] In this way, the modelling gives up on the unrealistic aspiration of offering grand theories of social phenomena. This in turn enables the modellers to focus on answering the research questions at the 'right' level of analysis, which choice may well be a pragmatic and empirical one.

Third, the pragmatic considerations around how to carry out model-based migration enquiries in practice are often difficult and idiosyncratic, but this can be partially overcome by identifying examples of existing good practice and greater precision about the type of research questions such models can answer. At the same time, there is acute need for being mindful of the epistemological limitations of various modelling approaches. A related issue of how to make any modelling exercises suitable and attractive for users and policy-makers additionally requires a careful managing of expectations, to highlight the novelty and potential of the proposed modelling approaches, while making sure that what is offered remains realistic and can be actually delivered.

One important remaining research challenge, where we envisage the concentration of more work in the coming years, is how to combine the different constituting elements of the modelling process together. Here again, having agreed guidelines and examples of good practice would be helpful, both for the research community and the users. In terms of the quality of input data and other information sources, there is a need to be explicit about what various sources of information can tell us, as well as about the quality aspects – and here, explicit modelling of the model provenance can help, as argued in Chap. 7 (see, in particular, Fig. 7.3).

In future endeavours, for multi-component modelling to succeed, establishing and retaining open channels for conversation and collaboration across different scientific disciplines is crucial, despite natural constraints in terms of publication and conference 'silos'. For informed modelling of complex processes such as migration, it is imperative to involve interdisciplinary research teams, with modelling and analytical experts, and diverse, yet complementary expertise of subject matter. Open discussions around good practice, exploring different approaches to modelling and decisions, matter a lot both for the practitioners as well as theorists and methodologists, especially in such a complex and uncertain area as migration. Importantly, this also matters if models are to be used as tools of policy support and advice. We discuss the specific aspects of that challenge next.

[2] We are particularly grateful to André Grow for drawing our attention to this interpretation.

9.3 Policy Impact: Scenario Analysis, Foresight, Stress Testing, and Planning

In the context of practical implications for the users of formal models, it is a truism to say that any decisions to try to manage or influence complex processes, such as migration, are made under conditions of high uncertainty. Broadly speaking, as signalled in Chap. 2, we can distinguish two main types of uncertainty. The *epistemic uncertainty* is related to imperfect knowledge of the past, present, or future characteristics of the processes we model. The *aleatory uncertainty*, in turn, is linked to the inherent and irreducible randomness and non-determinism of the world and social realm (for a discussion in the context of migration, see Bijak & Czaika, 2020). The role of these two components changes over time, as conjectured in Fig. 9.1, with diminishing returns from current knowledge in the more distant future, which is dwarfed by the aleatory aspects, driven by ever-increasing complexity. Importantly, the influences of uncertain events and drivers accumulate over time, and there is greater scope for surprises over longer time horizons.

In the case of migration, the epistemic uncertainty is related to the conceptualisation and measurement of migration and its key drivers and their multi-dimensional environments or 'driver complexes', acting across many levels of analysis (Czaika & Reinprecht, 2020). In addition, the methods used for modelling and for assessing human decisions in the migration context also have a largely epistemic character. Conversely, systemic shocks and unpredictable events affecting migration and its drivers are typically aleatory, as are the unpredictable aspects of human behaviour, especially at the individual level (Bijak & Czaika, 2020). At a fundamental level, the future of any social or physical system remains largely open and indeterministic,

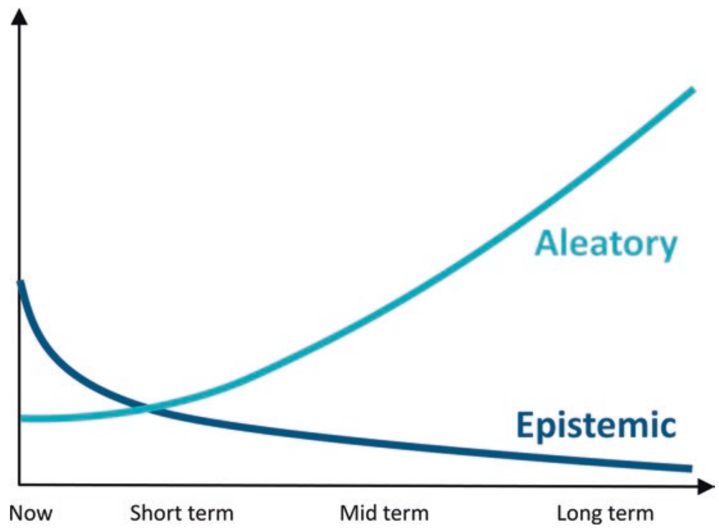


Fig. 9.1 Stylised relationship between the epistemic and aleatory uncertainty in migration modelling and prediction

with social systems additionally influenced by the irreducible uncertainty of human free will – or, in other words, agency (for a full philosophical treatment, see e.g. Popper, 1982).

In this context, an important question with practical and policy bearings is: can following the Bayesian model-based template help manage the different types of migration uncertainty across a range of time horizons? Given that different types of uncertainty dominate in different temporal perspectives, the usefulness of the proposed approach for policy and other practical applications depends on the horizon in question. An important distinction here is that while the epistemic uncertainty can be reduced, the aleatory one cannot, and needs to be managed instead. At the same time, formal modelling and probabilistic description of uncertainty can help address both these challenges.

The areas for possible reduction of the epistemic uncertainty have been highlighted throughout this book. The uncertainty in the data can be controlled, possibly by using formal quality assessment methods and combining information from different sources (Chap. 4); the features of the underpinning social mechanisms, embodied in model parameters, can be identified by formal model calibration (Chap. 5); and the knowledge on human decision making can be enhanced by carefully designed experiments (Chap. 6). Bearing in mind that there are trade-offs between the model precision and feasibility of its construction, an iterative modelling process, advocated in this book, can help identify the knowledge gaps, and thus delineate and possibly reduce epistemic uncertainty.

Given the presence of the aleatory uncertainty, in the strict predictive sense, any models of complex systems can only be valid at most in the short term, and only if uncertainty is properly acknowledged. Nevertheless, models can still be helpful for many other purposes across a range of time horizons, helping to manage policy and operational responses in the face of the aleatory uncertainty. Here, a variety of possibilities exist, from early warnings and stress testing in the short term, to long-range scenario analysis and foresight, all of which can help contingency planning (Bijak & Czaika, 2020).

9.3.1 Early Warnings and Stress Testing

Early warnings and stress testing are particularly useful for short term, operational purposes, such as humanitarian relief, border operations, or similar. What is required of formal models in such applications is a very detailed description, ideally aligned with empirical data. This description should be linked to the relevant policy or operational outcomes of interest, especially if the models are to be benchmarked to some quantitative features of the real migration system. Here, the models can be additionally augmented by using non-traditional data sources, such as digital traces from mobile phones, internet searches or social media, due to their unparalleled timeliness. In particular, formal simulation models can help calibrate early warning systems, by allowing to set the response thresholds at appropriate levels (see Napierała et al., 2021). At the same time, models can help with stress testing of the existing migration management tools and policies, by indicating with what (and how extreme) events such tools and policies can cope. One stylised example of such applications for the Risk and Rumours version of the migration route formation model is presented in Box 9.1.

Box 9.1: Model as One Element of an Early-Warning System

In the simplest example, corresponding to the operational needs of decision makers in the area of asylum migration, let us focus on the total number of arrivals at the destination, and on how this variable develops over time. There are clear short-term policy and planning needs here, related to the adequate resources for accepting and processing asylum applications, as well as providing basic amenities to asylum seekers: food, clean water, and shelter; possibly also health and psychological care, as well as education for children. All these provisions scale up with the number of new arrivals.

One example of a method for detecting changes in trends is the cumulated sum ('cusum') approach originating from statistical quality control (Page, 1954). In its simplest form, the cusum method relies on computing cumulative sums of the control variable, for example of the deviations of the observed migrant arrivals from a baseline level, and triggering a warning when a certain threshold h is reached. After a warning is triggered, the cumulative sum may then be reset to zero, to allow the system to adjust to the new levels of migration flows. Formally, if z_t is the variable being monitored, observed at time t, the cusum can be defined as $V_t = \max(0, V_{t-1} + z_t)$, where $V_0 = 0$. The use of the cusum approach to asylum migration has been discussed by Napierała et al. (2021).

Setting the threshold h at which the cusum method would trigger a warning is one of the key challenges of the approach, with visible trade-offs between false alarms (costly overreaction) and unwarranted complacency (costly lack of action). Simulation models, and even theoretical ones, such as the Risk and Rumours introduced in Chap. 8, can help shed light on the consequences of setting the thresholds at different levels. An illustration of this application is shown in Fig. 9.2, which presents a cusum chart based on the numbers of daily arrivals y_t simulated by the model. The variable under monitoring, z_t, measures a standardised number of arrivals, assuming that the average number under normal conditions is 10 persons daily, with a standard deviation of 2, so that $z_t = (y_t - 10)/2$. In real-life applications, this mean and standard deviation can, for example, correspond to the operational capacity of services that register new arrivals, and provide them with the basic necessities, such as food and shelter. To be able to respond effectively, such services need an early warning signal when the situation begins to depart from the normal conditions.

(continued)

Box 9.1 (continued)

In Fig. 9.2, a range of warnings issued at different levels of the threshold h are presented, denoted by black horizontal lines: solid for $h = 1$, dashed for $h = 2$ and dotted for $h = 4$. A warning is generated whenever the cusum line reaches a threshold. This means that for $h = 1$, the first warning, for the first wave of arrivals, is generated at time (day) $t = 90$, for $h = 2$ one day later, and for $h = 4$ three days later. For the second wave of arrivals, the warnings are generated almost synchronously: at $t = 178$ for $h = 1$ and at $t = 179$ for $h = 2$ or $h = 4$. At the same time, the threshold set at $h = 1$ generates false alarms at $t = 145$ and 146. Different thresholds have clearly varying implications for the timely operational response: while $h = 1$ leads to false alarms, and $h = 4$ may mean unnecessary delays, jeopardising the response, the threshold of $h = 2$ seems to be generating warnings about the right time. In this way, an agent-based model can be used to calibrate the threshold level of an early warning system for a given type of situation, bearing in mind the different implications of complacency on the one hand, and overreacting to the data signal on the other.

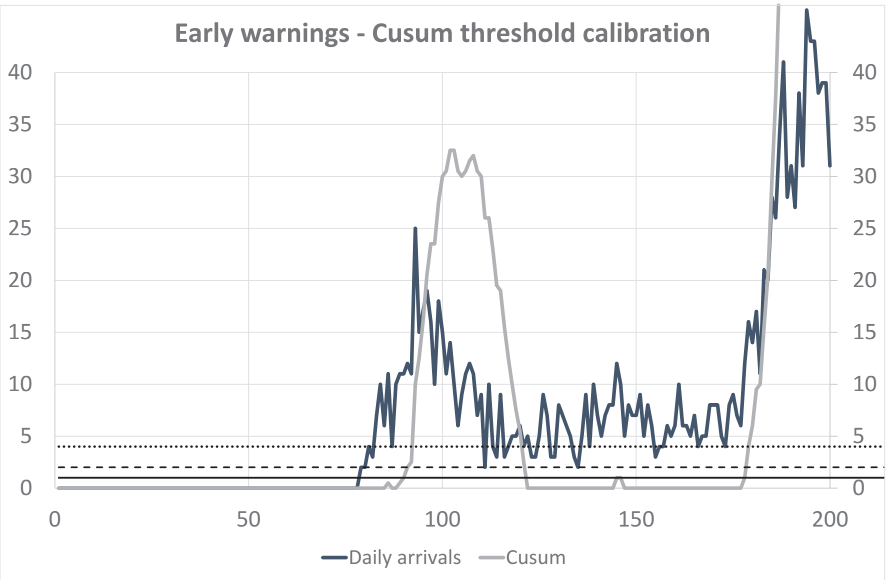


Fig. 9.2 Cusum early warnings based on the simulated numbers of daily arrivals at the destination in the migrant route model, with different reaction thresholds

9.3.2 Forecasting and Scenarios

At the other end of the temporal spectrum, foresight and scenario-based analyses, deductively obtained from the model results (see Chap. 2), are typically geared for higher-level, more strategic applications. Given the length of the time horizons, such approaches can offer mainly qualitative insights, and offer help with carrying out the stimulus-response ('what-if') analyses, as discussed later. This also means that these models can be more approximate and broad-brush than those tailored for operational applications, and can have more limited detail of the system description. An illustration of how an agent-based model can be used to generate scenarios of the emergence of various migration route topologies is offered in Box 9.2, in this case with specific focus on how migration responds to unpredictable exogenous shocks, rather than examining the reactions of flows to policy interventions, which is discussed next.

Box 9.2: Model as a Scenario-Generating Tool

To help decision makers with more strategic planning, formal scenarios – coherent model-based descriptions of the possible development of migration flows based on some assumptions on the developments of migration drivers – offer insights into the realm of possible futures, to which policy responses might be required. Ideally, to be useful, such scenarios need to be broad and imaginative enough, while at the same time remaining formal: an important advantage provided by modelling (Chap. 3). Here, scenarios based on agent-based models offer an alternative to other approaches to macro-level scenario setting with micro-foundations, such as, for example, the more analytical dynamic stochastic general equilibrium (DSGE) models used in macroeconomics (see Chap. 2; for a migration-related review and discussion, see also Barker & Bijak, 2020). One important feature of agent-based models in this context is that, being based on simulations, they do not require assumptions ensuring the analytical tractability of the problem, as is the case with DSGE or similar approaches.

As an illustration, we offer a range of scenarios generated by the theoretical version of the Risk and Rumours model presented in Chap. 8, under four sets of assumptions: the baseline one, as discussed before, for the different effects of risk on path choice among the agents ('risk-taking' versus 'cautious'), and varying levels of initial knowledge and communication ('informed' versus 'uninformed'), in each case for ten replicate runs. The scenarios illustrate the reaction of migrant arrivals to two exogenous shocks. The first is an increase in the number of the departures (and arrivals) of migrants seeking asylum from time $t = 150$, for example as a consequence of a deteriorating security situation caused by armed conflict in the countries of origin. The second shock simulates a situation where it becomes more difficult to cross a geographical barrier, such as the Mediterranean Sea, from time $t = 200$. In this

(continued)

case, the risk of the loss of life on the way increases, also due to external factors – these may be related to weather conditions, or to a smaller number of rescue efforts undertaken, for example caused by a global pandemic, a political crisis, or as a matter of political choice.

The outcomes of the various scenarios generated by the Risk and Rumours model are illustrated in Fig. 9.3. Unsurprisingly, the increased number of departures translates into an increased number of arrivals (with a time lag), and the number of fatalities reacts instantaneously to the deteriorating chances of a safe crossing. The differences for the number of arrivals obtained under different sets of assumptions are minimal, but for the number of deaths, there is a clear reduction in the fatalities under the higher levels of initial information and communication, although with considerable between-replicate variability, depicted by grey shading. This points to the information about safety of various routes as a possible area for a promising policy intervention, which is explored further in Box 9.3.

9.3.3 Assessing Policy Interventions

Contingency planning and stress-testing of migration policies and migration management systems can work across different time horizons. Such applications either require numerical input, which restricts the possible applications to shorter-term uses, or not, allowing also qualitative exploration of the space of model outcomes in the long run. In either case, the goal of the associated 'what-if' modelling exercise and the ensuing policy analysis is to assess the results of different assumptions and possible policy or operational interventions based on model results. In the migration context, possible examples may include the rerouting or changes of migration flows in response to multilateral changes of migration policies, recognition rates, information campaigns, and deploying other policy levers. Box 9.3 contains an illustrative example related to an information campaign on the safety of crossings.

As can be seen in Fig. 9.4, especially in comparison to the scenarios reported earlier in Fig. 9.3, the information campaign has barely any effect on the two model outcomes, except for minimally increasing death rates in trusting and risk-taking agents. Interestingly, the level of trust in the official information does not seem to play the role in the outcomes (Fig. 9.4). Part of the reason is that, regardless of whether the information campaign is trusted or not, it provides information about topology – possible paths and crossings – which the agents otherwise would not have access to. This effect can counterbalance any gains from the information campaign as such, especially in the situations when the agents trust the information they receive, but choose to ignore the warnings. This is an example of a mechanism possibly leading to unintended consequences of an in principle well-meaning migration policy (see Castles, 2004).

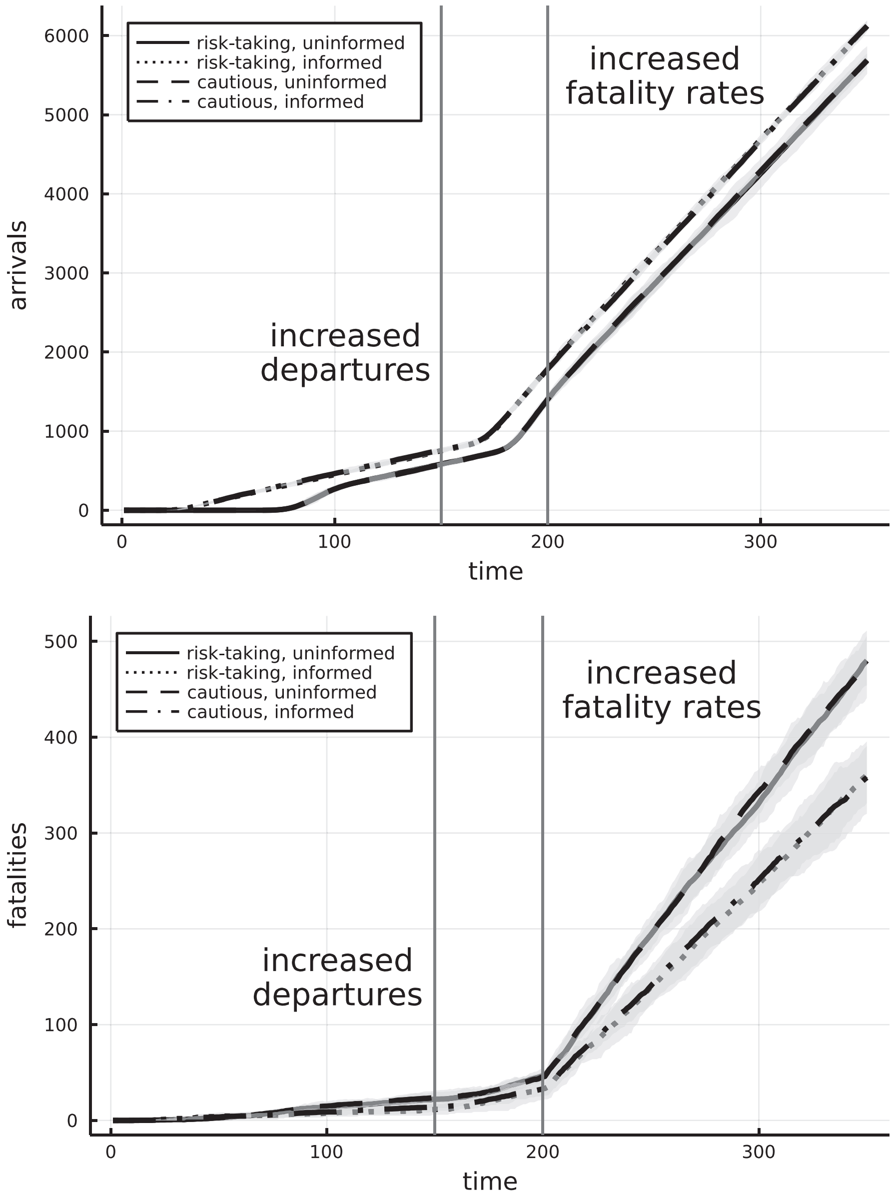


Fig. 9.3 Scenarios of the numbers of arrivals (top) and fatalities (bottom), assuming an increased volume of departures at t = 150, and deteriorating chances of safe crossing from t = 200. Results shown for the low and high effects of risk on path choice ('risk-taking' and 'cautious') and levels of initial knowledge and communication ('informed' and 'uninformed'), including between-replicate variation (grey shade)

Box 9.3: Model as a 'What-If' Tool for Assessing Interventions
Similar to scenarios driven by external shocks to the migration system, the models can serve as tools for examining 'what-if' type responses to changes to the system that can be driven by policies. As signalled in Box 9.2, a relevant example can refer to information campaigns, and to how the different ways of injecting reliable information into the system impacts the outcomes of the modelled migration flows – and of fatalities. Another question here is whether the policy tools work as envisaged by the policy makers, or if they can generate unintended consequences, and if so, what they are.

The example presented in this box is also inspired by a monitoring and evaluation study of information campaigns among prospective migrants carried out in Senegal (Dunsch et al., 2019), as well as by the findings from the Flight 2.0/Flucht 2.0 project (Emmer et al., 2016). Here, we first use the theoretical version of the Risk and Rumours model to examine the impact of a public information campaign carried out by official authorities, introduced in response to the increased number of fatalities during migrant journeys in a range of scenarios introduced in Box 9.2. The resulting trajectories of arrivals and deaths are presented in Fig. 9.4. We use the theoretical model to ascertain the possible direction and magnitude of impact of such an information campaign. The results are subsequently contrasted with those obtained for the empirically grounded model version (Risk and Rumours with Reality), shown in Box 9.4, to check whether they stay robust to additional information included in the model.

Whether the insights discussed above can be also gained from the model calibrated to the actual data series is another matter. To test it, in Box 9.4 we repeat the 'what-if' exercise introduced before, but this time for the Routes and Rumours with Reality version of the model, calibrated by using the Approximate Bayesian Computation (ABC) approach, described in Sect. 8.4.

On the whole, the results of scenarios, such as those presented in Boxes 9.3, and 9.4, can go some way towards answering substantive research and policy questions. This also holds for the questions posed in Chap. 8, as to whether increased risk – as well as information about risk – can bring about a reduction in fatalities among migrants by removing one possible 'pull factor' of migration. As can be seen from the results, this is not so simple, and due to the presence of many trade-offs and interactions between risk, peoples' attitudes, preferences, information, and trust, the effect can even be neutral, or even the opposite to what was intended. This is especially important in situations when different agents may follow different – and sometimes conflicting – objectives (see Banks et al., 2015). These findings – even if interpreted carefully – strengthen the arguments against withdrawing support for migrants crossing the perilous terrain, such as the Central Mediterranean (see Heller & Pezzani, 2016; Cusumano & Pattison, 2018; Cusumano & Villa, 2019; Gabrielsen Jumbert, 2020).

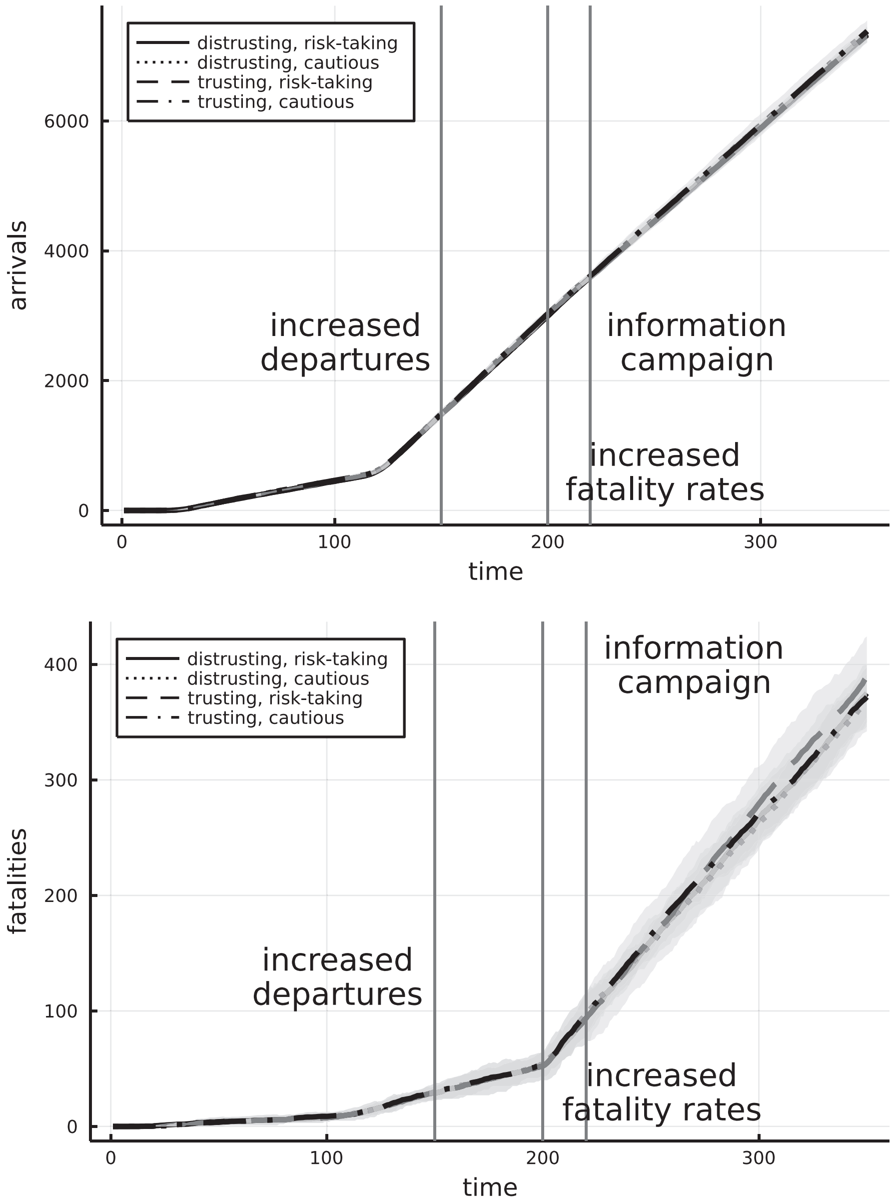


Fig. 9.4 Outcomes of different 'what-if' scenarios for arrivals (top) and deaths (bottom) based on a public information campaign introduced at $t = 210$ in response to the increase in fatalities

An interesting methodological corollary from the comparison of different scenarios is that it is not necessarily the most sophisticated and realistic version of the model that generates the most valuable policy insights: in our case, the calibration of the migration processes to the arrival and departure data in the Risk and Rumours with Reality model version overshadowed the mechanism of information-driven migration decisions, leading to a better-calibrated model, but with smaller role of

Box 9.4: Model as a 'What-If' Tool for Assessing Interventions (Cont.): Example of the Calibrated Routes and Rumours with Reality Model
In this example, we reproduced the results for the 'what-if' assessment of the efficiency of an information campaign, introduced in Box 9.3, for a calibrated version of the empirically grounded model, Routes and Rumours with Reality. A selection of results is shown in Fig. 9.5. The numbers for the original scenario ('plain') and for the one assuming an information campaign are very similar. For the latter scenario, 40 runs generated from the posterior distribution obtained by using Approximate Bayesian Computation are shown (solid grey lines) together with their mean (solid black line), while for the plain scenario, just the mean is presented (dashed black line), for the sake of transparency. For comparison, the (appropriately scaled) numbers from the empirical data are also included on the graph, to demonstrate the fit of the emulator to the real data.

From comparing the results shown in Figs. 9.4 and 9.5 it becomes apparent that the results of the scenario analysis for the calibrated model do not reproduce those for the theoretical version, Risk and Rumours, presented before. The effects that could be seen for the theoretical model disappear once an additional degree of realism is added, with the importance of the decision making mechanism, and the parameters driving it, being dwarfed by the information introduced through the process of model calibration. One tentative interpretation could be that once the model becomes more strongly benchmarked to the reality, the description of the decision processes needs to be more realistic as well. This points to the need for carrying out further enquiries into the nature of the decision processes undertaken by migrants during their journey, enhancing the model by including the possibilities of stopping the journey altogether at intermediate points, returning to the point of departure, travelling via alternative routes or means of transport, and so on.

the underlying behavioural dynamics of the agents and their interactions. Of course, the process of modelling does not have to end here: in the spirit of inductive model-based enquiries, these results indicate the need to get more detailed information both on the mechanisms and on observable features of the migration reality, so that the journey towards further discoveries can follow in a 'continuous ascent' of knowledge, in line with the broad inductive philosophy of the model-based approach.

9.4 Towards a Blueprint for Model-Based Policy and Decision Support

In practice, the identification of the way in which the models can support policy or practice should always start from the concrete needs of the users and decision makers, in other words, from identifying the questions that need answering. Here, the

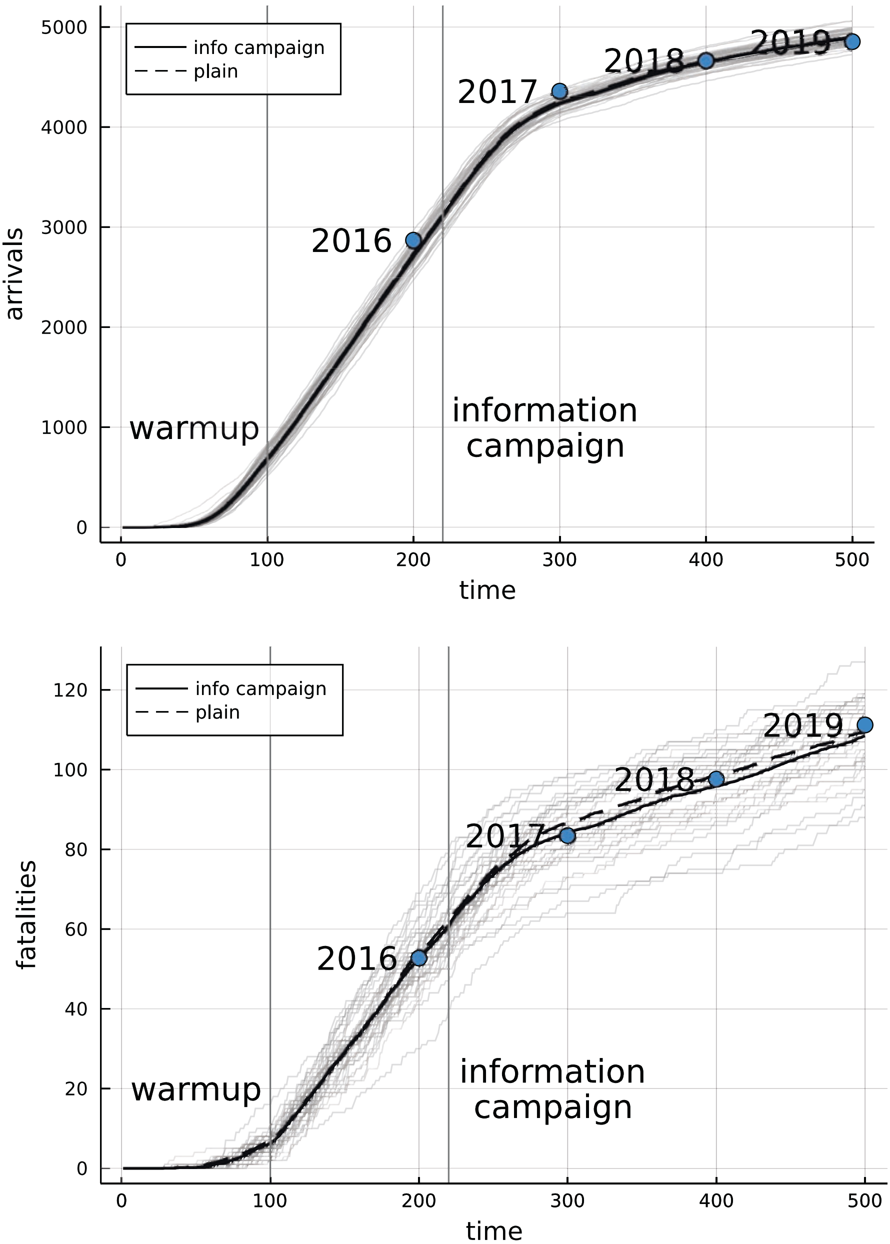


Fig. 9.5 Outcomes of the 'what-if' scenarios for arrivals (top) and deaths (bottom) based on a public information campaign introduced at t = 210, for the calibrated Risk and Rumours with Reality model

policy or practical implications of modelling necessitate formulating the model in the language of the problem, and including all the key features of the problem in the model description (see also Tetlock & Gardner, 2015). The type of problem and the length of the decision horizon will then largely determine the type of model. Coupled with the availability of data and other information, this will enable inferring the types of insights from the modelling exercise. This information will also limit the level of detail in modelling, from relatively arbitrary in data-free models, to limited by the availability and quality of data in empirically grounded ones. Hence, unless there is scope (and resources) for *ad hoc* collection of additional information, the level of reliance on empirical data can be (and often is) outside of the choice of the modeller.

When it comes to the modelling, our recommendation, as argued throughout this book in the spirit of the inductive Bayesian model-based approach, is to start with a simple model and scale it up, adding complexity if needed to answer the question, even in an approximate manner. At this stage, the data should be also brought in, where possible. Once the model produces the results sought, it is then a matter for the decision maker to judge whether the outputs are sufficient for the purpose at hand, and given the data and resource limitations, or if more detail needs adding to the model. The acceptable model version then is used to produce the required outcomes, and – crucially – assess the limitations of the answers offered by the model, as well as residual uncertainty. This broad blueprint for using models to aid policy, operations, interventions, and other types of practical applications is diagrammatically shown in Fig. 9.6.

Of course, a key limitation, present in all modelling endeavours, is the fundamental role of model uncertainty – an effect that has been dubbed the Hawkmoth Effect, analogous to the Butterfly Effect known from the chaos theory (Thompson & Smith, 2019). The Hawkmoth Effect means that even with models that are close

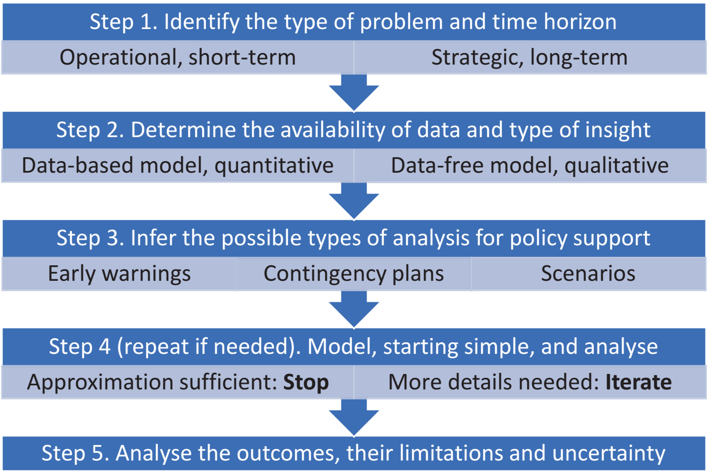


Fig. 9.6 Blueprint for identifying the right decision support by using formal models

to the reality they represent, their results and predictions, especially quantitative (in the short run), but also qualitative (in the long run), can be far off. As any model-based prediction is difficult, and long-term quantitative predictions particularly so (Frigg et al., 2014), the expectations of model users need to be carefully managed to avoid false overpromise.

Still, especially in the context of fundamental and irreducible uncertainty, possibly the most important role of models as decision support tools is to illuminate different trade-offs. If the outputs are probabilistic, and the user-specific loss functions are known, indicating possible losses under different scenarios of over- and underprediction, the Bayesian statistical decision analysis can help (for a fuller migration-related argument, see Bijak, 2010). Still, even without these elements, and even with qualitative model outputs alone, different decision or policy options can be traded off according to some key dimensions: benefits versus risk, greater efficiency versus preparedness, liberty versus security. These are some of the key considerations especially for public policy, with its non-profit nature and hedging against the risk preferable to maximising potential benefits or rewards. At the end of the day, policies, and the related modelling questions, are ultimately a matter of values and public choice: modelling can make the options, their price tags and trade-offs more explicit, but is no replacement for the choices themselves, the responsibility for which rests with decision makers.

Chapter 10
Open Science, Replicability, and Transparency in Modelling

Toby Prike

Recent years have seen large changes to research practices within psychology and a variety of other empirical fields in response to the discovery (or rediscovery) of the pervasiveness and potential impact of questionable research practices, coupled with well-publicised failures to replicate published findings. In response to this, and as part of a broader open science movement, a variety of changes to research practice have started to be implemented, such as publicly sharing data, analysis code, and study materials, as well as the preregistration of research questions, study designs, and analysis plans. This chapter outlines the relevance and applicability of these issues to computational modelling, highlighting the importance of good research practices for modelling endeavours, as well as the potential of provenance modelling standards, such as PROV, to help discover and minimise the extent to which modelling is impacted by unreliable research findings from other disciplines.

10.1 The Replication Crisis and Questionable Research Practices

Over the past decade many scientific fields, perhaps most notably psychology, have undergone considerable reflection and change to address serious concerns and shortcomings in their research practices. This chapter focuses on psychology because it is the field most closely associated with the replication crisis and therefore also the field in which the most research and examination has been conducted (Nelson et al., 2018; Schimmack, 2020; Shrout & Rodgers, 2018). However, the issues discussed are not restricted entirely to psychology, with clear evidence that similar issues can be found in many scientific fields. These include closely related fields such as experimental economics (Camerer et al., 2016) and the social sciences more broadly (Camerer et al., 2018), as well as more distant fields such as biomedical research (Begley & Ioannidis, 2015), computational modelling (Miłkowski et al., 2018), cancer biology (Nosek & Errington, 2017), microbiome research

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_10

(Schloss, 2018), ecology and evolution (Fraser et al., 2018), and even within methodological research (Boulesteix et al., 2020). Indeed, many of the lessons learned from the crisis within psychology and the subsequent periods of reflection and reform of methodological and statistical practices apply to a broad range of scientific fields. Therefore, while examining the issues with methodological and statistical practices in psychology, it may also be useful to consider the extent to which these practices are prevalent within other research fields with which the modeller is familiar, as well as the research fields that the findings of the modelling exercise either relies on, or is applied to.

Although there was already a long history of concerns being raised about the statistical and methodological practices within psychology (Cohen, 1962; Sterling, 1959), a succession of papers in the early 2010s brought these issues to the fore and raised awareness and concern to a point where the situation could no longer be ignored. For many within psychology, the impetus that kicked off the replication crisis was the publication of an article by Bem (2011) entitled "Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect." Within this paper, Bem reported nine experiments, with a cumulative sample size of more than 1000 participants and statistically significant results in eight of the nine studies, supporting the existence of paranormal phenomena. This placed researchers in the position of having to believe either that Bem had provided considerable evidence in favour of anomalous phenomena that were inconsistent with the rest of the prevailing scientific understanding of the universe, or that there were serious issues and flaws in the psychological research practices used to produce the findings.

Further issues were highlighted through the publication of two studies on questionable research practices in psychology, "False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant" by Simmons et al. (2011), and "Measuring the prevalence of questionable research practices with incentives for truth telling", by John et al. (2012). Using two example experiments and a series of simulations, Simmons et al. (2011) demonstrated how a combination of questionable research practices could lead to false-positive rates of 60% or higher, far higher than the 5% maximum false-positive rate implied by the endorsement of $p < 0.05$ as the standard threshold for statistical significance. Specifically, the authors showed that collecting multiple dependent variables, not specifying the number of participants in advance, controlling for gender or the interaction of gender with treatment, or having three conditions but preferentially choosing to report either all three or only two of the conditions, can lead to large increases in the false-positive rates that become even more extreme when several of these research practices are combined. To drive home the point further, Simmons et al. (2011) conducted a real study with 20 undergraduate students and then used the analytical flexibility available to them and the lax reporting standards for statistical analyses to report an impossible finding: that they had 'found' that listening to the song "When I'm Sixty-Four" rather than "Kalimba" led to participants being younger, with the test statistic $F(1, 17) = 4.92$ implying a 'significant' p-value, $p = 0.040$.

Closely following the Simmons et al. (2011) paper, John et al. (2012) published a survey on the research practices of psychologists, finding that the type of practices Simmons et al. (2011) had shown to be highly problematic were commonplace. Responses to the full list of questionable research practices included in the survey varied considerably (see John et al., 2012 for full results for all ten questionable research practices). Some research practices were considered much less defensible, such as outright falsification of data (admitted to by 0.6–1.7% of the sample of researchers, depending on the condition) or making misleading or untrue statements within the paper such as, "In a paper, claiming that results are unaffected by demographic variables (e.g., gender) when one is actually unsure (or knows that they do)", (admitted to by 3.0–4.5% of the sample, depending on condition). Even more commonplace was the benefit of hindsight: the statement, "In a paper, reporting an unexpected finding as having been predicted from the start", was admitted to by 27.0–35.0% of the sample, again depending on condition (John et al., 2012, *passim*).

Other research practices examined in the survey were considered more defensible and were admitted to by a majority of the psychologists surveyed, but can still contribute to massively increased false positive rates prevalent in the literature. For example, 55.9–58.0% of the sample admitted to, "Deciding whether to collect more data after looking to see whether the results were significant", and 63.4–66.5% of the sample admitted to, "In a paper, failing to report all of a study's dependent measures" (*idem*). It is also important to note that these are conservative estimates based on the willingness of individual psychologists to admit that they personally had engaged in questionable research practices, and therefore the actual prevalence of questionable research practices is likely far higher. John et al. (2012) also calculated prevalence estimates based on respondents' answers to questions about the percentage of *other psychologists who have engaged in a questionable research practice* as well as the percentage of those *other psychologists who have engaged in a questionable research practice and would admit to having done so,* and for nearly all of the questionable research practices these estimates were considerably higher than the number who actually made self-admissions within the survey (*idem*).

The publication of a large-scale replication attempt of 100 psychological findings by the Open Science Collaboration (2015) showed the practical extent of the problems highlighted by Simmons et al. (2011) and John et al. (2012). Although 97 of the 100 original studies included for replication reported statistically significant results, only 36 of the replication attempts ended up statistically significant, despite having statistically well-powered designs (with an average power – probability of correctly rejecting a false hypothesis – equal to 0.92), and despite matching the original studies closely, including using original materials wherever possible. Other large-scale replication efforts, including the Many Labs projects within psychology (Ebersole et al., 2016; Klein et al., 2014, 2018), projects in fields such as experimental economics (Camerer et al., 2016), and the social sciences more broadly (Camerer et al., 2018), as well as more distant fields, such as cancer biology (Nosek & Errington, 2017), have highlighted that, to varying extents, there are serious issues with the reliability and replicability of findings published within many scientific areas.

10.2 Open Science and Improving Research Practices

Once the issues outlined above were clearly highlighted, many scholars within psychology decided that reform was necessary, and serious changes within the field needed to be made.[1] Changes to current practices were recommended at several levels of the scientific process, including at the level of individual authors, reviewers, publishers, and funders (Munafò et al., 2017; Nosek et al., 2015; Simmons et al., 2011). Some of the changes to research practice that have been most commonly recommended and widely engaged with by researchers include openly publishing the data and analysis code online, openly publishing study materials online, and the preregistration of study methodology and analysis plans (Christensen et al., 2019).

The change in research practice that has seen the earliest and greatest uptake by researchers is the public sharing of data and/or analysis code (Christensen et al., 2019). Making the data and analysis code underlying research claims openly available has many potential benefits for both science as a whole and for individual researchers who engage in the practice. Benefits to the scientific process from the open sharing of data include: allowing other scientists to re-analyse data to help verify the results and check for errors, providing safeguards against misconduct such as data fabrication, or taking advantage of analytical flexibility, for example, because other scientists can discover that a result is entirely reliant on a specific covariate. It also allows other researchers to reuse the data for a variety of purposes (Tenopir et al., 2011). If data are publicly available, then they may be reanalysed to answer new questions that were not initially examined by the researchers. Without open data, these reanalyses would not be possible and therefore the scientific knowledge would either not be generated at all, or would require the recollection of the same, or highly similar data, leading to waste and inefficiency in the use of resources (usually public funding; Tenopir et al., 2011).

There are also good reasons for individual researchers to publicly post their data even if they are motivated by their own self-interest. Articles with publicly available data have an advantage in the number of citations received (Christensen et al., 2019; Piwowar & Vision, 2013), and willingness to share data are associated with the strength of evidence and quality of the reporting of statistical results (Wicherts et al., 2011). However, even though the uptake of the public posting of data and software code is growing quickly and should be lauded, there are still many problematic areas, such as incomplete data, missing instructions, and insufficient information provided. These issues mean that even when data are publicly shared, independent researchers may still regularly face considerable hurdles and/or not actually be able to analytically reproduce the results reported in the paper (Hardwicke et al., 2018; Obels et al., 2020; Stagge et al., 2019; Wang et al., 2016).

[1] Although it has to be noted that there was also pushback from some scholars – see Schimmack (2020) for further discussion of the responses to the replication crisis.

Another common and rapidly growing area of open science is the public posting of study materials or instruments and experimental procedures (Christensen et al., 2019). Like open data and analysis code, this practice has the benefit of increasing transparency and making it clear to editors, reviewers, and readers of articles, what exactly was done within the study. This increased transparency allows for easier assessment of whether there are potential confounds or other flaws in the study methodology that may have impacted on the conclusions. It also allows for easier assessment of the appropriateness and validity of the stimuli and materials used. Openly sharing materials and procedures also has the additional benefits of making it far easier for other researchers to conduct direct *replications* of the research (i.e., taking the same materials and procedures and collecting new data to independently verify the results), as well as to conduct follow up studies that attempt to conceptually replicate, adapt, or expand on some or all of the aspects of the study without the need to contact the original authors and/or to expend time and resources reproducing or creating new study materials and procedures. These practices are in addition to ensuring the *reproducibility* of the results, which is here understood as ensuring that the software or computer code applied to a given dataset produces the same set of results as reported in the study.[2]

One major change in research practice that has the potential to greatly reduce questionable research practices and improve the quality of science is preregistration: registering the aims, methods and hypotheses of a study with an independent information custodian *before* data collection takes place (Nosek et al., 2018; Wagenmakers et al., 2012). Although preregistration is still currently less common than openly sharing data, code, and materials, the uptake of the practice is increasing rapidly (Christensen et al., 2019). Preregistration has been referred to as 'the cure' for analytical flexibility or 'p-hacking', the practice of fine-tuning analyses until the desired or a publishable result, as measured by the magnitude of p-values, can be obtained (Nelson et al., 2018, p. 519).

When researchers preregister their studies, they need to outline in advance what their research questions and hypotheses are, as well as their plans for analysing the data to answer these questions and verify the hypotheses (Nosek et al., 2018; Wagenmakers et al., 2012). Therefore, if done correctly, preregistration ensures that the analyses conducted are confirmatory, which is a required assumption for null hypothesis significance testing. It also allows both the researchers themselves and other consumers of research products to have much greater confidence that the results can be relied upon, and the false-positive rate has not been greatly inflated through questionable research practices (Simmons et al., 2011). In this way, preregistration is also useful for the researchers conducting the research, as it helps them to avoid biases and misleading themselves (Nosek et al., 2018). Once discovering an unexpected but impactful result in the data, or that controlling for a variable or excluding participants based on a specific criterion leads to a statistically significant

[2] For a broad terminological discussion of replicability and reproducibility, which are terms that still remain far from being unambiguously defined and used, see e.g. National Academies of Sciences, Engineering, and Medicine (2019).

finding that can be published, it can be easy for hindsight bias and wishful thinking to lead researchers to justify these analytical decisions to both themselves and others, and to believe that they predicted or planned them all along (also known as 'hark-ing' – "hypothesising after results are known"; Kerr, 1998).

However, preregistration alone is not likely to solve the problems with research malpractice unless reviewers, editors, publishers, and readers ensure that researchers actually follow their preregistered hypotheses and analysis plans. Registration of clinical trials has been commonplace for some time now, yet published trials still regularly diverge from the prespecified registrations, with publications switching and/or not reporting the primary outcomes listed in trial registries (Goldacre et al., 2019; Jones et al., 2015), and journals showing resistance to attempts to highlight or correct issues when informed of discrepancies between the trial registries and the articles they had published (Goldacre et al., 2019). Going even further than preregistration, a growing number of journals now offer a registered report format in which studies are reviewed based on the underlying research question(s), study design, and analysis plan and can then be given in principle acceptance, meaning that the study will be published regardless of the results provided the authors adhere to the pre-agreed protocols (Chambers 2013, 2019; Nosek & Lakens, 2014; Simons et al., 2014).

In addition to the changes in research practice outlined above, there has also been considerable discussion about the use of statistics within psychology and other scientific fields, including a special issue of *The American Statistician* entitled "Statistical Inference in the 21st Century: A World Beyond $p < 0.05$". Within the special issue, and in various other articles, books, and publications, the contributors have criticised the use of p-values, and particularly the $p < 0.05$ cut-off conventionally used to determine 'statistical significance', as well as the phrase 'statistically significant' itself. Indeed, the editors of *The American Statistician* recommended that the phrase 'statistically significant' no longer be used (Wasserstein et al., 2019).

There is still much disagreement about what new statistical practices should be adopted or how researchers should move forward, with a variety of potential solutions proposed. For example, some have recommended that the $p < 0.05$ threshold be redefined to $p < 0.005$ instead (Benjamin et al., 2018), whereas others have advocated for a shift away from null hypothesis significance testing towards Bayesian analyses and inference (Wagenmakers et al., 2018). At the same time, some other authors, notably Gigerenzer and Marewski (2015), have warned about the idolisation of simple Bayesian measures, such as Bayes Factors. In the same way as had happened with p-values, indolent statistical reporting can occur under the Bayesian paradigm as much as in the frequentist one. Although there is still some disagreement about the possible future directions for statistical analysis and inference, the general guidance provided by the editors of *The American Statistician* – "Accept uncertainty. Be thoughtful, open, and modest." (Wasserstein et al., 2019, p. 2) – provides a direction for future empirical enquiries.

10.3 Implications for Modellers

The above discussion has outlined a series of issues that have occurred within psychology and a variety of other experimental and empirical domains of science, as well as some of the solutions that are already being implemented and potential future directions for further improvements in methodology and statistics. The following section relates these considerations back to the specific domains of computational modelling and simulation, highlighting the relevance of the lessons learned for researchers and practitioners within these domains. There is documented evidence of similar issues occurring within computational modelling, and issues within empirical fields can also impact computation modelling because of the interconnectedness of scientific disciplines.

Many of the issues highlighted above are also relevant for computational modelling, and even in circumstances where a concern is not directly applicable to modelling challenges, there are some analogous concerns (Miłkowski et al., 2018; Stodden et al., 2013). As with the practice of sharing data, analysis code, study materials, and study procedures for empirical studies, clearly and transparently documenting models is vital for other researchers to be able to verify and expand upon existing work. Chapter 7 of this book highlights several existing methods that modellers can use to document or describe simulation models, such as the ODD protocol (Overview, Design concepts, Details; Grimm et al., 2006), or provenance standards, such as PROV (Groth & Moreau, 2013).

Similar to the sharing of data and analysis code, there are often serious issues with attempting to computationally reproduce existing models and simulations even if code is provided. This can happen because of a range of factors, such as the exclusion of important information within publications and failing to properly document model and/or simulation code (Miłkowski et al., 2018). As with sharing data and analysis code for empirical work, transparently sharing documentation and descriptions of computational models has the advantage of allowing other researchers to test and verify the extent to which outputs are dependent on specific modelling choices made in the modelling process, how sensitive the model is to changes in various inputs (see Chap. 5 for more details on sensitivity analysis), and/or the extent to which the results change (or remain consistent) when the model uses different data or is applied in a different context (e.g., if a model of asylum migration from Syria is applied to asylum migration from Afghanistan).

Computational modelling often requires far more decisions regarding design, formalisation, and implementation than standard experimental or empirical work, and in some cases is more exploratory in nature. Therefore, preregistration does not seem like a readily applicable or appropriate format to be transferred to all aspects of computational modelling, although it is certainly still applicable to at least some aspects (e.g., if models are to be compared, it is useful to preregister the models that will be compared as well as how the comparison will be conducted; see Lee et al., 2019 for more information). Nonetheless, there are several strategies that can be used to try and reduce the extent to which modellers have the flexibility to tinker with their models to find the specific settings that produce the desired (publishable) results.

One option here is for modellers to develop and rely on prespecified architectures within their models, such as the BEN (Behavior with Emotions and Norms) architecture, which provides modules that can add aspects such as emotions, personality, and social relationships to agent-based models (Bourgais et al., 2020). Alternatively, independent researchers can recreate a model without referring to or relying on the original model code, which can help to test the extent to which outputs are dependent on modelling choices for which there are a variety of plausible and defensible alternative options (see Silberzahn et al., 2018 for an analogous example with statistical analyses). Reinhardt et al. (2019) have provided a detailed discussion of the processes and lessons learned from implementing the same model in two different modelling languages, one a general-purpose language using discrete-time and the other a domain-specific modelling language using continuous time.

In addition to the open science and methodological concerns within computational modelling, related research practices within psychology and other empirical fields can also have considerable impact on modelling practice because of the interplay between scientific disciplines and how computational models may rely on or be informed by findings from empirical work. Therefore, the tendency for many empirical fields to simply rely on finding 'statistically significant' effects rather than attempt to accurately estimate effect sizes or test them for robustness limits the extent to which these findings can be usefully and easily applied to computational models. Additionally, if a computational model is informed by, or relies on, empirical findings to justify mechanisms and processes within the model (e.g., the decision making of agents within an agent-based model), then if those findings are unreliable and/or based on questionable research practices, this may effectively undermine the whole model.

These limitations once again highlight the advantage of provenance modelling standards, such as PROV (Groth & Moreau, 2013; Ruscheinski & Uhrmacher, 2017), as a format for documenting and describing models. PROV allows information to be stored in a structured format that can be queried, thereby allowing it to be easily seen which entities a model relies on (see Chap. 7). Therefore, if new research highlights issues within the existing literature (e.g., a failed replication within psychology), or new discoveries are made, it is a relatively simple and straightforward task to search PROV information, and discover which models have incorporated this information as an entity, and therefore may have at least some aspects of the model that need to be reconsidered or updated.

This strategy could also be combined with sensitivity analysis (see Chap. 5) to establish the extent to which the model outputs are sensitive to aspects that rely on the entity now called into question, and therefore whether it is necessary to update the model in light of the new information. Additionally, PROV has the potential to contribute to the empirical literature by highlighting specific entities (e.g., research studies) that are commonly featured within models. Such studies may therefore become a high priority for large-scale replication efforts, not only to ensure the reliability and robustness of the findings, but also to identify potential moderators (mediating and confounding variables) and boundary conditions.

The choice of specific tools and solutions notwithstanding, one lesson for modellers that can be learned from the replicability crisis is clear: transparency and proper documentation of the different stages of the modelling process are vital for generating trust in the modelling endeavours and in the results that the models generate. For the results to be scientifically valid, they need to be reproducible and replicable in the broadest possible sense – and documenting the provenance of models is a necessary step in the right direction.

Chapter 11
Conclusions: Towards a Bayesian Modelling Process

Jakub Bijak and Peter W. F. Smith

In the concluding chapter we summarise the theoretical, methodological and practical outcomes of the model-based process of scientific enquiry presented in the book, against the wider background of recent developments in demography and population studies. We offer a critical self-reflection on further potential and on limitations of Bayesian model-based approaches, alongside the lessons learned from the modelling exercise discussed throughout this book. As concluding thoughts, we suggest potential ways forward for statistically-embedded model-based computational social studies, including an assessment of the future viability of the wider model-based research programme, and its possible contributions to policy and decision making.

11.1 Bayesian Model-Based Population Studies: Moving the Boundaries

Given the current state of knowledge, what are the perspectives for computational migration and population modelling? The two intertwined challenges, those of uncertainty and complexity, can be broken down into a range of specific knowledge gaps, dependent on the context and research questions being addressed. The explanatory power of simulation models (for a general discussion, see Franck, 2002 and Courgeau et al., 2016), well suited for tackling the complexity of social processes, such as migration, can be coupled with the statistical analysis aimed at the quantification of uncertainty. Throughout this book, we have argued for the use of modelling and its encompassing statistical analysis as elements of a language for describing and formalising relationships between elements of complex systems. We discuss some of the specific points and lessons next.

The main high-level argument put forward in this book is that model building is – or needs to be – a continuing process, which aims to reduce the complexity of social reality. The formal sensitivity analysis helps retain focus on the important

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_11

aspects, while disregarding those whose impact is only marginal. All the constituting building blocks of this process are therefore important: starting from the computational model itself, and its implementation in a suitable programming language, through empirical data, information on human decision making – which, as in our case, can come from experiments – and the statistical analysis of each model version. All of these elements contribute to our greater ability to understand the model workings, while retaining realism about the degree to which the model remains a faithful description of the reality it aims to represent. The formalisation of model analysis also allows us to explore the model behaviour and outcomes in a rigorous way, while being transparent about the assumptions made. In this way, we can illuminate the micro-level mechanisms (micro-foundations) that generate the population-level processes we observe at the macro scale, while formally acknowledging the different sources of their uncertainty.

Of course, when it comes to representing reality, all models are more likely to hold higher resemblance to the actual processes under specific conditions. To that end, adding more detail and data helps approximate the reality, but this comes at a cost of increased uncertainty. By doing so, the models also run the risk of losing generality, and their nature becomes more descriptive than predictive or explanatory. At the same time, as shown in Chap. 9, there are trade-offs involved in the different purposes of modelling, too: better predictive capabilities of a model can lead to a loss of explanatory power of the underlying mechanisms, if it is dominated by the information used for model calibration.

In such cases, additional effort is required in terms of data collection and assessment, to make sure that the model-based description of an idiosyncratic social process is as accurate as possible. The successive model iterations may then not be strictly embedded within one another, so that the 'ascent' of knowledge, which would be ideally seen in the classical inductive approach, is not necessarily monotonic (Courgeau et al., 2016). Still, even in such cases, the more detailed models can offer more accurate *approximations* of the reality. Formal description of the model-building process, for example by using provenance modelling tools discussed in Chap. 7, can help shed light on that, while keeping track of the developments in the individual building blocks in the successive model versions.

At the same time, such models can retain some ability to generalise their outcomes, although at the price of increased uncertainty. To that end, models can still make some theoretical contributions (Burch, 2018), especially if 'theory' is not interpreted in a strict nomological way, as a set of well-established propositions from which the predictions can be simply deduced (Hempel, 1962). Instead, the models can answer well-posed explanatory questions ('how?') in a credible manner – offering increasingly plausible descriptions of the underlying social mechanisms, as long as their construction follows several iterations of the outlined process, checking the model-based predictions against the observed reality. At the same time, some residual (aleatory) uncertainty always remains, especially in the modelling of social processes, and addressing it requires going beyond models alone.

In the light of the above findings, the modelling processes can also be given novel interpretations. Social phenomena, such as migration, are very complicated and complex inverse problems, which in the absence of an omniscient Laplace's demon – a hypothetical being with the complete knowledge of the world, devoid of the epistemic uncertainty – do not have unique solutions (see Frigg et al., 2014). The scientific challenges of model identifiability are therefore akin to the studies of non-response or missing information, but this time carried out on a space of several possible (and plausible) models. Model choice becomes yet another source of the uncertainty of the description of the process under study, alongside the data, parameters, expert input, and so on. Still, the iterative model construction process advocated throughout this book enables building models of increasing analytical and explanatory potential, which at the same time remain computationally tractable.

This is yet another argument for turning to the philosophy of Bayesian statistical inference: the initial model specification is but a prior in the space of all possible models, and the modelling process by which we can arrive at the increasingly accurate approximations of reality is akin to Bayesian model selection. Of course, there is an obvious limitation here of being restricted to a class of models pre-defined by the modellers' choices and, ultimately, their imagination (see also the discussion of inductive and abductive reasoning in Chap. 2). The inductive process of iterative learning about the dynamics of complex phenomena, besides being potentially Bayesian itself, can also include several other Bayesian elements, describing the uncertainty of different constituting parts, such as individual decisions of agents in the model (and updating of knowledge), model estimation and calibration, and meta-modelling.

The status quo in demography and population studies, on which this work builds, can be broadly described as the domination of empiricism at the expense of more theoretical enquiries (Xie, 2000), with an increasing recognition that some areas of theoretical void can be filled by formal models (see Burch, 2003, 2018). At the same time, recent years have seen promising advances in the demographic and social science methodology. The modelling approaches of statistical demography, including Bayesian ones, hardly existent until the second half of the twentieth century, are now a well-established part of mainstream population sciences (Courgeau, 2012; Bijak & Bryant, 2016), while agent-based and other computational approaches, despite recent advances (Billari & Prskawetz, 2003; van Bavel & Grow, 2016), remain somewhat of a novelty. So far, as discussed in Daniel Courgeau's Foreword, these two modelling approaches have remained hardly connected, and connecting them was one of the main motivations behind undertaking the work presented in this book.

Against this background, our achievements can be seen both at the level of the individual constituent parts of the modelling process, presented in Chaps. 3, 4, 5, 6, and 7, as well as – if still tentatively – the way in which they can coherently work together. To that end, advances made at the level of process development and documentation, together with their philosophical underpinnings, offer a blueprint for constructing empirically relevant computational models for studying population (and, more broadly, social) research questions. The opening up of population and other social sciences for new approaches and insights from other disciplines can be an important step towards moving the boundaries of analytical possibilities for studying the complex and the uncertain social world. However, despite all the advances, some important obstacles on this journey remain, which we discuss next.

11.2 Limitations and Lessons Learned: Barriers and Trade-Offs

From the discussion so far, key challenges for advancing the Bayesian model-based agenda for population and broader social sciences are already clear. The main one relates to putting the different building blocks together in a unified, interdisciplinary modelling workflow. The interdisciplinarity is of lesser concern: most disciplines in social sciences are very familiar and comfortable with the high-level notion of modelling as an approximation of reality, so all that is needed for a successful bridging of disciplinary barriers is willingness to share other perspectives, open communication, and clear definitions of the concepts and ideas so that they can be understood across disciplines.

A much greater challenge lies in the fusion of different building blocks at an operational level: how to include experimental results in the simulation model? How to operationalise data and model uncertainty? How to implement the model in a way that balances computational efficiency with the transparency of code? These are just a few examples of questions that need answering for this approach to reach its full potential. Some possibilities for ideas dealing with these challenges have been proposed throughout this book, but they are just the tip of the iceberg. To develop some of these ideas further, and to come up with robust practical recommendations, a higher-level reflection is needed. Such a synthetic view and advice could be offered, for example, from the point of view of philosophy of science, science and technology studies, or similar meta-disciplines.

Another key challenge relates to the empirical information being too sparse and not exactly well tailored, either for the model requirements, or for answering individual research questions. What is contained in the publicly available datasets is often, at least to some extent, different to what is needed for modelling purposes. This leads to important problems at several levels. First, the models can be only partially identified through data, with many data gaps and free parameters compounding the output uncertainty. Second, the quality of the existing data may be low, with their uncertainty assessment contributing additional errors into the model. Third, the use of proxies for variables that conceptually may be somewhat different (e.g. GDP per capita instead of income, or Euclidean distance between capital cities of origin and destination countries instead of the distance travelled), can introduce additional biases and uncertainty, not all aspects of which may be readily visible even after a thorough quality assessment (see Chap. 4). The operationalisation problem is particularly acute for such variables and concepts as, for example, trust, risk-aversion, or many other psychological traits, for which no standard measures exist.

At the same time, as shown in Chaps. 5 and 8, modelling coupled with a formal sensitivity analysis can provide a way of identifying the data and knowledge gaps, and consequently of filling them with information collected through dedicated means. From the point of view of addressing individual research questions, this can be quite resource-consuming, sometimes prohibitively so, as it requires devoting additional resources in terms of time, labour and money, to the collection of new data. Yet when such data can be generated and deposited in an open-access repository, such activities, whenever possible, can offer positive externalities for a broader research community, with the possible applications of the collected data going beyond a particular piece of research (see Chap. 10). The same holds for tailor-made experiments, for which an additional aspect of the sensitivity analysis involves verifying the impact of psychologically plausible decision rules and mechanisms against the default placeholder assumptions, such as rational choice and maximum utility (Chap. 6).

The interpretation of models as tools to broaden the understanding of the processes at hand, through illuminating the information gaps, feedbacks, unintended consequences, and other aspects of individual-level human decisions and their impact on observed macroscopic, population-level patterns, is one of the many non-predictive applications of formal modelling (Epstein, 2008). In fact, as with the examples presented in this book, the purely predictive uses of models become of secondary importance. There is so much uncertainty in complex social and population processes, that not only proper description of the full extent of this uncertainty becomes difficult, but also any formal decision analysis on the basis of such predictive models would be very limited, and may well be hardly possible.

In the case of complex social processes, even once everything that is potentially known or knowable has been accounted for, and the corresponding epistemic uncertainty, related to imperfect knowledge, has been reduced, the residual uncertainty remains large. Even the most carefully designed and calibrated models still reflect the underlying messy and complex social reality, which is characterised by relatively large and irreducible aleatory uncertainty, related to the intrinsic randomness of the social world. For such applications, the focus of the analysis shifts from exact prediction and the resulting well-defined cost-benefit decision analysis, to aiding the broader preparedness and planning. In this way, the models can play an important role in testing the impact of different scenarios and assumptions, including qualitative ones, in a logically coherent simulated environment (Chap. 9).

The main lessons learned from the model-based endeavours, however, are about trade-offs. Of course, such trade-offs also exist at the level of the model analysis, with changes in some variables having non-trivial impact on others through non-linear relationships and feedback loops. Still, from the methodological point of view, even more important may be the process-level trade-offs, such as between increasing the level of detail and description of the social phenomena (topology of the world, decision processes, agents' memory and learning, and so on), and the computational constraints, including run times, computer memory efficiency.

Every building block of the modelling process includes trade-offs as well. For data, the choice may be between their bias and variance; for experiments, between different levels of cognitive plausibility and less realistic default assumptions; for implementation, between general-purpose and domain-specific languages; for the analysis, between descriptive and more sophisticated analytical tools; and for documentation, between description and formalisation. As in real life, modelling leaves plenty of room for choice, but the model-based process we suggest in this book is designed to help make these choices and their consequences transparent and explicit.

11.3 Towards Model-Based Social Enquiries: The Way Forward

So, in summary, what can formal models and the lessons learned from following an interdisciplinary modelling process potentially offer population and other social scientists? The specific findings and more general reflections reported throughout this book point to important insights that can be generated by modelling, not necessarily limited to the specific research question or questions, but also leading to chance discoveries of some related process features, which can in turn produce new insights or lines of enquiry. In this way, modelling increases not only our understanding of the pre-defined features of the processes, but also the more general characteristics of the process dynamics. This is especially important for such complex and uncertain phenomena as migration flows. At the same time, it is also important to reflect on the practical limitations of furthering the model-based agenda, and health warnings related to the interpretation of the model results.

The key lessons from the work we describe throughout this book are threefold. First, modelling of a complex social phenomenon itself is a process, not a one-off endeavour. The process is iterative, and its aim is an ever-better sequence of approximations of the problem at hand, in line with the inductive philosophical principles of the scientific method, possibly coupled, where needed, with the pragmatic tenets of abductive reasoning (see Chap. 2). Second, the presence of many aspects of the modelling process – as well as of the process being modelled, especially in the social realm – requires true interdisciplinarity and interconnectedness between the different perspectives, rather than working in individual, discipline-specific silos. Third, the formal acknowledgement of uncertainty – in the data, parameters, and models themselves – needs to be central to the modelling efforts. Given the complex and highly structured nature of social problems, Bayesian methods provide an appropriate formal language for describing this uncertainty in different guises. These principles, coupled with a thorough and meticulous documentation of the work, both for legacy purposes and possible replication (see Chap. 10), are the main scientific guidelines for model development and implementation.

At the same time, the impact of models is not limited to the scientific arena. To make the most of the modelling endeavours targeted at practical applications, as argued in Chap. 9, the involvement of the users and other relevant audiences in the modelling process needs amplifying. This in turn requires greater modelling literacy on the part of the model users, next to statistical literacy (Sharma, 2017). The onus on ensuring greater literacy is on modellers, though: the communication of model workings and limitations needs to be specific and trustworthy, and provided at the right level of technical detail for the audience to understand. The levels of trust can be, of course, heightened by following established conventions in modelling (see Chap. 3): carrying out a thorough assessment of the available data (Chap. 4) and a multi-dimensional assessment of uncertainty (Chap. 5); following established ethical principles in gathering information that requires it (Chap. 6); and providing meticulous documentation of the process, for example through ODD and provenance description (Chap. 7). In short, the keys to good communication and effective user involvement are transparency, rigour, and awareness of the limitations of modelling. At the same time, the very purpose of model-building, and any practical uses of the models, are also related to societal values and can have ethical dimensions, which needs to be borne in mind.

There are other practical obstacles related to interdisciplinary modelling. Large and properly multi-perspective modelling endeavours are themselves complex, time-consuming and costly, having to rely on interdisciplinary teams. For communication within teams, a common language needs to be established, ensuring that the joint efforts are targeting shared problems. Even within the best-functioning teams, however, scientific challenges at the connecting points between the disciplines are inevitable (see Chap. 8). At the same time, overcoming them takes time and patience. Some interesting discoveries reported in this book were a result of our evolution in thinking about the modelling process and its components over the course of a five-year project. That there are not too many existing examples of such modelling projects and endeavours, is exactly why such work is both needed, and so difficult at the same time. This is also why large-scale scientific investments, offering funding beyond disciplinary silos, with modelling explicitly recognised as cross-cutting activity, are of crucial importance. They provide the necessary structures to help scientists from different areas connect by making them learn – and speak – the same language: the language of formal models.

Of course, modelling cannot solve all problems faced by population sciences, migration studies, or social enquiries more generally. As argued above, the aleatory uncertainty, some of which is related to human behaviour and agency, remains irreducible: this is in fact a welcome sign of the power of human spirit, free will and imagination. Still, formal models can help us get answers to questions that are more complex and sophisticated – and hopefully also more interesting and relevant – than those allowed by the more traditional social science tools. This is the beginning of a longer journey into the world of modelling, and despite the price that has to be paid for engaging in such activities, this is definitely worth doing, for the sake of exploring new intellectual horizons, designing more robust solutions to practical and policy problems, and ultimately making the social world a bit less uncertain.

Appendices: Supporting Information

Appendix A. Architecture of the Migrant Route Formation Models

Martin Hinsch

This Appendix supplements the information provided in Chap. 3 by providing a basic description of the main elements of the Routes and Rumours model and, by extension, of the Risk and Rumours, as well as Risk and Rumours with Reality models, introduced in Chap. 8.

A1. Model Description

The aim of the model is to investigate the formation of migration routes and how they are affected by the availability and exchange of information. In our model agents attempt to traverse a – for them – unknown landscape, having to rely on either local exploration or communication with other agents to find the best path across. The following gives a general overview of the model. For a more detailed description, as well as the source code, we would like to refer to Hinsch and Bijak (2019), and the links to the online repository with model code and documentation are available at: www.baps-project.eu.

Entities
Entities directly represented in the model are **agents**, **settlements,** and **transport links**.

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7

Agents
The agents represent migrants undertaking a journey from the origin to the destination. At any time, agents are either present at a settlement or a transport link or they have arrived at the destination.

Contacts
Each agent has a list of other agents that it is in contact with (representing their social network), and can exchange information with (see **information** below).

Knowledge
Each agent has a potentially incomplete and inaccurate set of knowledge items concerning the world. Each item describes the properties and topology of a settlement or a transport link.

Settlements
Settlements are located at a specific position on the map and differ in **quality** and **resources**. Settlements are connected among each other by random transport links (see **setup** below).

Transport Links
Links always connect two settlements. The only property of links is **friction**, which subsumes length and difficulty of travel.

Interactions
The only entities to change state over the course of the simulation are the agents. They do that by interacting with cities, links and other agents. Agents can exchange information with agents either in their contact list or present at the same location as them. They can travel along transport links and collect information on their current and neighbouring locations. For more details see Section A2 on model-specific processes below.

Information
Information and how agents use and exchange it is a crucial part of the model. Each item of knowledge an agent has – for example, the quality of a specific settlement – is described by an **estimate** and a level of **certainty**. That is, an agent has an idea of the numerical value of a given property and how certain it is that the value is correct.

For a given agent, these numbers change either when the agent explores its environment or when it exchanges information with other agents. When collecting information from the environment, the estimate becomes more accurate while the certainty increases. Information exchange is a bit more complicated. Generally speaking, the more certain an agent is (i.e. the higher its certainty value) the stronger the effect on the other agent's estimate. At the same time, agents with similar beliefs (i.e. similar values for estimate) will reinforce each other and their certainty will increase, while for very dissimilar beliefs certainty can decrease.

Travel
Agents start out at entry settlements (origin locations) at one edge of the map and attempt to reach exit settlements (destination locations) at the other edge.

Agents decide if and where to go purely based on the subjective information they have available. If an agent does not have enough information to find a route to an exit, it will attempt to improve its local position (if possible) by travelling to an adjacent city that is 'better' than the current one, where quality is determined by the city properties (*quality* and *resources*), the travel distance or effort (i.e. *friction*) and the city's proximity to the exit edge of the map.

If an agent knows enough to find a complete route, it will attempt to travel the route with the lowest costs, where costs are again a function of city properties and travel effort.

Setup

Before the start of the simulation a map of settlements and links is generated and their property values assigned. To generate the topology we use a random geometric graph: all cities are placed at random locations, then cities that are closer than a given threshold are connected with a transport link. In addition, we place a fixed number of *entry* and *exit* settlements at the respective edge of the map and connect them with the nearest 'regular' settlements.

At the beginning of the simulation no agents are present in the simulation. Newly-added agents (see Processes below) start out at entry cities with (dependent on scenario) either no or only rudimentary knowledge of the world and some randomly selected contacts to other agents pre-assigned.

A2. Processes

The model is implemented as an event-based simulation. That means that updates to the model state do not happen in discrete time steps but instead as asynchronous Poisson processes. Therefore, all activities, interactions and state changes are separate processes with specific rates of occurrence.

Most processes are changes of state in single agents. Whether they can apply is usually dependent on whether an agent is travelling (present at a transport link) or not (present at a city). It is important to note that every agent in the population can potentially experience the state change in question at any time that it fulfils the respective conditions.

Departures

The only process happening at the world level is the addition of new agents. Depending on scenario, the departure rate of new agents is either constant or starts out at zero and increases linearly to a fixed value.

Planning

Agents that are not travelling can re-evaluate their travelling plans if they have received enough new information. The rate for planning depends on how out of date an agent's information is.

Exploration

Agents that are currently not travelling can collect information on their current location and neighbouring links and settlements.

Contacts
Agents that are not travelling can add other agents that are present at the same location to their list of contacts. The rate of gaining contacts depends on the number of agents present at the location.

Leaving
Agents that are not travelling can leave. This means they change their location to a transport link and thus become travelling agents. The rate at which agents leave is constant.

Arriving
Agents that are travelling can arrive at the next location (and thus become non-travelling agents). If they arrive at an exit they immediately become inactive (they can still communicate information to their contacts, however). Arrival rates depend on a link's **friction**.

Communication
At any time, agents can exchange information with one of their contacts. The rate at which this happens depends on the number of contacts an agent has.

A3. Illustration

As an illustration of the model's workings and outcome, we provide a visual description of the finding that while clear migration routes emerge in the model, in many scenarios these can be very different from the routes one would expect if agents always found the optimal path (Fig. A.1).

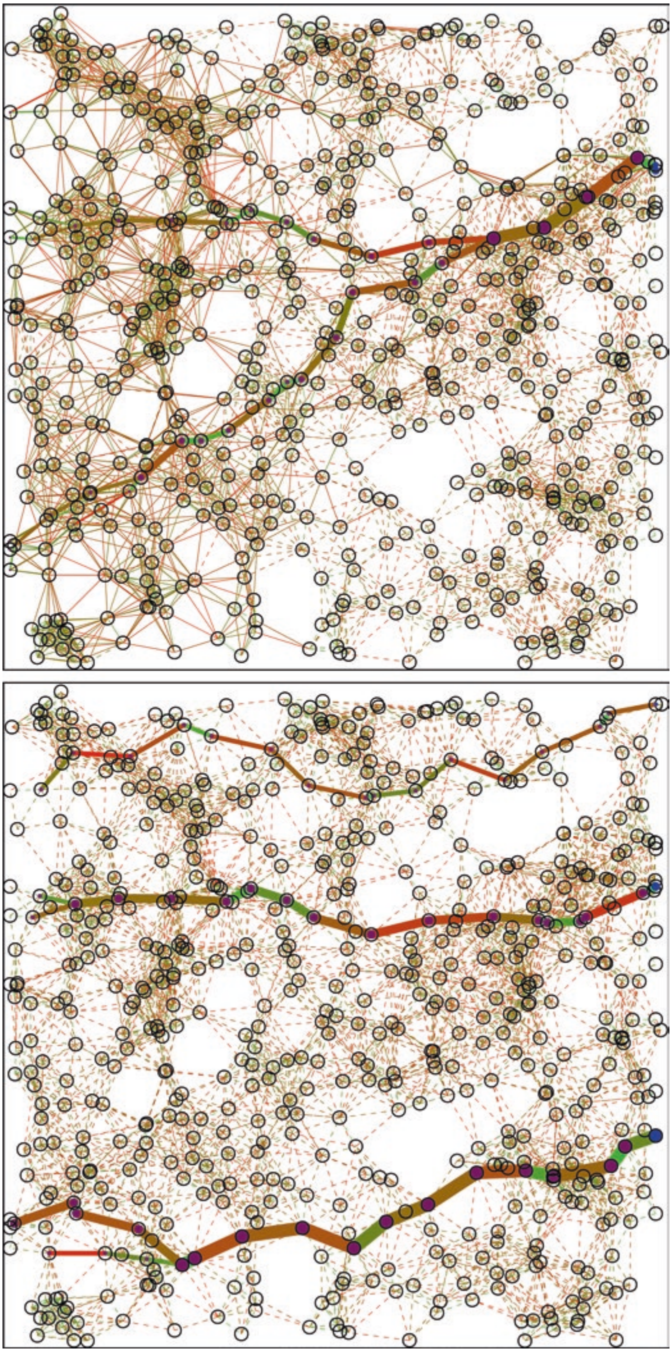


Fig. A.1 Realised (top) and hypothetical optimal (bottom) migration routes with migrants travelling left to right. Circles represent cities, transport links are shown as lines. Links without any traffic are drawn dashed, and lines with traffic are solid. Thickness of the line represents sum traffic over the entire run of the simulation. Source: own elaboration

Appendix B. Meta-Information on Data Sources on Syrian Migration into Europe

Sarah Nurse and Jakub Bijak

This Appendix supplements the information provided in Chap. 4 devoted to building a knowledge base on the data concerning a specific migration flow, together with their uncertainty assessment. In particular, we provide meta-information on selected data sources on Syrian asylum-related migration into Europe in the 2010s, with the view of aiding computational modelling of migration processes.

This Appendix contains two parts. In the first part (B1), we offer summary information on the various data sources that can be used for modelling recent Syrian migration into Europe, together with brief description and quality assessment following a common methodology described in the working paper. Additionally, in the second part (B2), we list key supplementary general sources on migration processes, mechanisms, drivers, or features (numbered with a prefix S) for reference, with basic information on their most important aspects.

For all sources, the information provided includes a broad topic (e.g. populations, routes, or drivers), type of a particular source (registrations, survey, census, operational, review, journalistic, interviews), type of data (quantitative or qualitative, process-related or contextual, and macro-level or micro-level), as well as their temporal and spatial detail. This is accompanied by a brief content description, some general notes, including those justifying the quality assessment, a link, and information on access.

In addition, in the first part (B1), an assessment of the quality of sources is carried out across eight dimensions, wherever relevant: purpose of collection; timeliness of data; trustworthiness; detailed disaggregation; population under study and associated definitions; transparency of the source; its completeness; as well as sample design for surveys. Each of these dimensions, as well as a global summary score, is classified into one of three categories: green, amber – and red, or possibly one of the two mixed ones (green-amber, amber-red) for the in-between ratings. Specific descriptors for assessing data sources according to all the individual criteria are listed in Chap. 4 (see Table 4.1).

As discussed in Chap. 4, the classification and rating are done purely from the point of view of usefulness of the data for modelling, rather than for their own stated purpose, so that for example data on border apprehensions, while of crucial importance for border enforcement purposes, cover only a selected subgroup of the population that would be modelled. By no means should the assessment be therefore interpreted as definitive and valid for all different purposes for which the data may be used.

This version of the meta-inventory presented in this Appendix is current as of 1 May 2021, and any future updated versions are available via an interactive online tool at www.baps-project.eu.

B1. Selected Key Sources of Data on Syrian Migration into Europe

01	UNHCR operational portal			**Topic:** destination population	
Source type: registration	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** daily	**Geography:** 5 countries
Content description: otal cumulative daily numbers of Syrian refugees and asylum seekers registered in Egypt, Iraq, Jordan, Lebanon and Turkey, including breakdown by age group/sex and camp/non-camp.					
Notes: administrative data supporting relief efforts, comprising daily numbers published approximately quarterly, specifically on the Syrian refugees. Limits/caps on registration may under-represent numbers.					
Link: https://data2.unhcr.org/en/situations/syria **Access information:** data series and distributions publicly available for download					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green/amber	
Population and definitions	Transparency	Completeness	Sample design N/A		

02	UNHCR population stocks			**Topic:** destination population	
Source type: registration	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** annual	**Geography:** all countries
Content description: total annual stocks of the UNHCR populations of concern, including refugees, asylum seekers and internally displaced persons, for all countries of origin and destination					
Notes: a by-product of the administrative registration process, with very wide coverage, but small temporal granularity, published with a delay of over a year. Possible undercount: as above.					
Link: http://popstats.unhcr.org/en/persons_of_concern **Access information:** data publicly available for download from an interactive database					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green/amber	
Population and definitions	Transparency	Completeness	Sample design N/A		

03	UNHCR sea & land arrivals			**Topic:** routes and journey	
Source type: registration	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** monthly	**Geography:** 5 countries
Content description: aggregate registration data on sea and land arrivals since 2015 by main European country of arrival in the Mediterranean basin (Greece, Italy, Spain, Cyprus, Malta)					
Notes: monthly data on registered arrivals, published a few months after the reference date. Possible undercount: as above.					
Link: https://data2.unhcr.org/en/situations/mediterranean# **Access information:** data publicly available for download from an interactive database					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green/amber	
Population and definitions	Transparency	Completeness	Sample design N/A		

04	**UNHCR Syrian arrivals**			**Topic:** routes and journey	
Source type: survey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** Jan-Mar 2016	**Geography:** Greece
Content description: socio-demographic characteristics of Syrian migrants, with information on region of origin, route, resources, reason for decisions, access to information and support received.					
Notes: three one-off surveys in Greece, aiming to provide better information on refugees, with sufficient detail for key variables and with methodology (interval sampling) explicitly described.					
Link: https://data2.unhcr.org/en/documents/download/47014 and …/en/documents/details/47162 **Access information:** survey publications and summary results available for download					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green/amber	
Population and definitions	Transparency	Completeness	Sample design		

05	**UNHCR Longing to go Home**			**Topic:** destination population; Routes and journey	
Source type: survey, interviews	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** 2017	**Geography:** Lebanon
Content description: a one-off survey and interviews/focus groups containing a range of information on intentions of refugees in camps in Lebanon, including intentions for moving to third countries					
Notes: the survey aims to measure intentions, based on a limited sample, the details of which have not been presented in the report. Results include basic description and fragments of interviews.					
Link: https://data2.unhcr.org/en/documents/details/63310 **Access information:** survey publication and summary results available for download					
Purpose	**Timeliness**	Trustworthiness	Disaggregation	**Summary rating:** Amber	
Population and definitions	Transparency	Completeness	Sample design		

06	**EASO asylum trends**			**Topic:** destination population	
Source type: registration	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** daily*/monthly	**Geography:** whole EU+
Content description: applications, decisions and pending cases for EU+ countries, total and broken down by citizenship. Figures not yet validated, so may differ from the official Eurostat statistics.					
Notes: administrative data, published with two months' delay. Not validated. Aggregate statistics for EU+ only, with national-level data by receiving country not published for legal reasons.					
Link: https://www.easo.europa.eu/latest-asylum-trends **Access information:** monthly data publicly available, *daily data available for internal EASO purposes					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green/amber	
Population and definitions	Transparency	Completeness	Sample design N/A		

07	**Eurostat asylum data**			**Topic:** destination population	
Source type: registration	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** monthly	**Geography:** EU+ countries
Content description: a range of data on many relevant topics: applications, decisions, pending cases, Dublin statistics, and enforcement including number refused entry by border type and nationality.					
Notes: administrative official statistics on various aspects of asylum and enforcement, with monthly granularity, published regularly. Data subject to quality control before publication.					
Link: https://ec.europa.eu/eurostat/data/database > … > Asylum and managed migration (migr) **Access information:** data publicly available for download from a well-organised interactive database					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green	
Population and definitions	Transparency	Completeness	Sample design N/A		

08	**Eurostat country data**			**Topic:** destination context	
Source type: varies	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** varies	**Geography:** whole EU+
Content description: various data for EU countries on migration factors and drivers, including: migrant integration, economic indicators (including GDP and employment rates), social conditions, and policy.					
Notes: mostly administrative and survey (LFS) data, with clear definitions, but lacking some detail for certain variables of interest e.g. country of birth. Examples: economy and finance – national accounts (GDP); Population & social conditions: demography and migration, Asylum and managed migration, Health, Labour market, Living conditions and welfare, Income, consumption & wealth, Social protection					
Link: https://ec.europa.eu/eurostat/data/database > … > Economy and finance, Population & Social conditions. **Access information:** data publicly available for download from an interactive database					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green	
Population and definitions	Transparency	Completeness	Sample design *where relevant*		

09	**Syrian official statistics**			**Topic:** origin population	
Source type: census, survey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** varies	**Geography:** Syria
Content description: population distributions before conflict e.g. by educational status, marital status, age groups and nationality, sub-national labour force statistics, basic demographic indicators.					
Notes: data from the 2004 census, 2006–12 labour force surveys, and a one-off 2009–10 family health survey, with some limited characteristics of the pre-conflict Syrian population. Meta-information largely unavailable. For surveys, sampling frames unknown. More recent data (e.g. yearbooks) untrustworthy.					
Link: http://cbssyr.sy/index-EN.htm **Access information:** selected publications available for download					
Purpose	**Timeliness**	Trustworthiness	**Disaggregation**	**Summary rating:** Amber/red	
Population and definitions	**Transparency**	**Completeness**	**Sample design *where relevant***		

<table>
<tr><td>10</td><td colspan="3">IOM GMDAC portal</td><td colspan="2">Topic: destination population; context; routes and journey</td></tr>
<tr><td>Source type: various</td><td>Quantitative
Qualitative</td><td>Process
Context</td><td>Macro-level
Micro-level</td><td>Time detail: mainly annual</td><td>Geography: worldwide</td></tr>
<tr><td colspan="6">Content description: a comprehensive data portal of the IOM Global Migration Data Analysis Centre presenting a range of migration-related variables and indicators from a variety of secondary sources (e.g. UN, Eurostat) and data on migrant deaths and disappearances from the Missing Migrants Project (see 12).</td></tr>
<tr><td colspan="6">Notes: provides very easy access to reliable migration-related data. The data are mainly annual; and often lacking detail for some key variables. There is a clear description of sources and methodology. Some estimates (e.g. UN stocks) rely on definitions from national censuses and on interpolations.</td></tr>
<tr><td colspan="6">Link: https://migrationdataportal.org/
Access information: data, metadata and reports available for download from a well-organised database</td></tr>
<tr><td>Purpose</td><td>Timeliness</td><td>Trustworthiness</td><td>Disaggregation</td><td colspan="2" rowspan="2">Summary rating:
Green/amber</td></tr>
<tr><td>Population and definitions</td><td>Transparency</td><td>Completeness</td><td>Sample design N/A</td></tr>
</table>

<table>
<tr><td>11</td><td colspan="3">IOM Missing Migrants: flows</td><td colspan="2">Topic: routes and journey; destination population</td></tr>
<tr><td>Source type: operational</td><td>Quantitative
Qualitative</td><td>Process
Context</td><td>Macro-level
Micro-level</td><td>Time detail: monthly</td><td>Geography: global; Med.</td></tr>
<tr><td colspan="6">Content description: number of coastguard interceptions, with specific focus on Mediterranean crossings (for the Central Mediterranean route, from Libya and Tunisia to Italy/Malta).</td></tr>
<tr><td colspan="6">Notes: data on the maritime interceptions (e.g. for the Central Mediterranean route, obtained from Libyan and Tunisian coastguards) published up to 2019. Recording interceptions rather than people means that a person may be counted several times, making multiple attempts.</td></tr>
<tr><td colspan="6">Link: https://missingmigrants.iom.int/downloads
Access information: data publicly available for download</td></tr>
<tr><td>Purpose</td><td>Timeliness</td><td>Trustworthiness</td><td>Disaggregation</td><td colspan="2" rowspan="2">Summary rating:
Amber</td></tr>
<tr><td>Population and definitions</td><td>Transparency</td><td>Completeness</td><td>Sample design N/A</td></tr>
</table>

<table>
<tr><td>12</td><td colspan="3">IOM Missing Migrants: deaths</td><td colspan="2">Topic: routes and journey; context</td></tr>
<tr><td>Source type: various</td><td>Quantitative
Qualitative</td><td>Process
Context</td><td>Macro-level
Micro-level</td><td>Time detail: daily/monthly</td><td>Geography: global; Med.</td></tr>
<tr><td colspan="6">Content description: numbers of the dead and missing by date, route and location, as recorded in administrative, operational and journalistic sources. Focus on Mediterranean crossings.</td></tr>
<tr><td colspan="6">Notes: minimum estimates of deaths recorded by IOM observers, national authorities and media. Reports information source for each death/event (e.g. boat capsizing). Information published approximately weekly.</td></tr>
<tr><td colspan="6">Link: https://missingmigrants.iom.int/downloads
Access information: data publicly available for download</td></tr>
<tr><td>Purpose</td><td>Timeliness</td><td>Trustworthiness</td><td>Disaggregation</td><td colspan="2" rowspan="2">Summary rating:
Amber</td></tr>
<tr><td>Population and definitions</td><td>Transparency</td><td>Completeness</td><td>Sample design N/A</td></tr>
</table>

13	IOM Displacement Tracker			**Topic:** destination population; flows; drivers: conflict/disasters	
Source type: registration	Quantitative / Qualitative	Process / Context	Macro-level / Micro-level	**Time detail:** monthly/daily	**Geography:** worldwide
Content description: the Displacement Tracking Matrix (DTM) presents data on displaced and returned populations, including some local assessments of shelter/living conditions, and flow monitoring.					
Notes: includes population displacement due to conflict, disaster and other reasons, monitored by IOM. Flow database includes a selection of Southern European countries of arrival.					
Link: https://displacement.iom.int/ (displacement statistics), https://flow.iom.int/ (flows) **Access information:** data available for download from a highly visual interactive database					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green/amber	
Population and definitions	Transparency	Completeness	Sample design N/A		

14	IDMC Global Displacement			**Topic:** destination population; drivers: conflict/disasters	
Source type: various	Quantitative / Qualitative	Process / Context	Macro-level / Micro-level	**Time detail:** monthly/daily	**Geography:** worldwide
Content description: data on persons internally displaced due to conflict, persecution and natural or human-made disasters, compiled by the Internal Migration Monitoring Centre (IDMC). Demographically consistent flow (new displacements) and stock data. Exemplary documentation and meta-information.					
Notes: data based on multiple sources: IOM DTM (see **13** above), augmented by using other collections (e.g. UN OCHA, national governments and humanitarian organisations) and formal risk modelling.					
Link: http://www.internal-displacement.org/ **Access information:** data publicly available for download					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green	
Population and definitions	Transparency	Completeness	Sample design N/A		

15	OECD Migration databases			**Topic:** destination population	
Source type: various	Quantitative / Qualitative	Process / Context	Macro-level / Micro-level	**Time detail:** annual	**Geography:** OECD+
Content description: three databases: *OECD International Migration database* – annual flows and stocks; *Database on Immigrants in OECD countries* (including a few non-OECD) – demographic and labour market characteristics of immigrants; and *Indicators of Immigrant Integration* – national and local measures of employment, education, social inclusion, civic engagement and social cohesion.					
Notes: information from the network of migration correspondents ('Sopemi') from OECD+ countries.					
Link: http://www.oecd.org/migration/mig/oecdmigrationdatabases.htm **Access information:** data series and distributions publicly available for download					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green	
Population and definitions	Transparency	Completeness	Sample design N/A		

16	World Bank Factbook			Topic: flows and impacts	
Source type: registration	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** annual or less	**Geography:** worldwide
Content description: World Bank's Migration and Remittances Factbook dataset includes estimates of bilateral migration flows (once every few years), as well as financial remittance flows (annual).					
Notes: estimates are compiled from a range of national and international sources, down to the level of single bilateral flows of migrants and remittances, where quality aspects may vary by country.					
Link: http://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data **Access information:** migration/remittance matrices and series available for download					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green/amber	
Population and definitions	Transparency	Completeness	Sample design N/A		

17	ILOStat (formerly Laborsta)			Topic: destination population, flows and impacts	
Source type: various	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** annual	**Geography:** worldwide
Content description: comprehensive database of the International Labour Organization, covering different aspects of the labour force, including migration flows and migrant stocks.					
Notes: the estimates are derived from the UN migrant stock data (see also **10** above), Eurostat and OECD statistics, as well as regional sources (e.g. ASEAN), which may vary in quality across countries.					
Link: https://www.ilo.org/ilostat/ **Access information:** data series and interactive query results available for download					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green/amber	
Population and definitions	Transparency	Completeness	Sample design N/A		

18	Frontex apprehensions			Topic: routes and journey	
Source type: operational	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** monthly	**Geography:** EU ext. borders
Content description: administrative/operational data on monthly numbers of 'Illegal border crossings' (i.e. apprehensions) by nationality, route and border type, for sections of EU external borders					
Notes: data collected for border enforcement, and published with two months' delay. Illegal border crossings rather than all border crossings or number of migrants; one migrant may cross multiple times. Sources are published, but limited information on data collection. No way of assessing completeness.					
Link: https://frontex.europa.eu/along-eu-borders/migratory-map/ **Access information:** monthly data freely available for download					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Amber	
Population and definitions	Transparency	**Completeness**	Sample design N/A		

19	Frontex Risk Analysis data			Topic: routes and journey	
Source type: operational	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** monthly	**Geography:** EU ext. borders
Content description: data on detections of illegal border-crossing at/between border crossing points; refusals of entry; asylum applications; detections of illegal stay, facilitators or fraudulent documents.					
Notes: enforcement data, reported monthly and published quarterly, for top ten nationalities in each category (Syrians not always in the top ten). Sources, data collection and completeness: as above.					
Link: https://frontex.europa.eu/publications/?category=riskanalysis **Access information:** publications and reports (EaP-RAN and FRAN) freely available for download					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Amber	
Population and definitions	Transparency	**Completeness**	Sample design N/A		

20	Human Costs of Borders			Topic: routes and journey	
Source type: registrations	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** annual	**Geography:** Mediterranean
Content description: official, state-produced records of people who died while attempting to reach southern EU countries via the Mediterranean, and whose bodies were found in or brought to Europe. Death registration data for 1990–2013 in selected coastal areas of Greece, Italy and Spain.					
Notes: only limited disaggregations available. Clear definitions for inclusion but lacking detail for some key variables. Methodology rigorous and explicitly described. Explicit strategies to achieve completeness but limited to strict definition of bodies found (=minimum confirmed), rather than total death estimates.					
Link: http://www.borderdeaths.org/ **Access information:** data and publications freely available for download					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Amber	
Population and definitions	Transparency	Completeness	Sample design N/A		

21	Displaced persons in Austria			Topic: destination population	
Source type: survey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** Nov-Dec 2015	**Geography:** Austria
Content description: DiPAS: a dedicated survey on socio-economic characteristics, human capital, and attitudes of asylum-seekers, predominantly from Syria, Iraq, and Afghanistan.					
Notes: a one-off academic survey, aimed at better understanding of the asylum seeking population; specifically includes Syrian refugees. Peer reviewed publications on data collection and methodology.					
Link: https://www.oeaw.ac.at/en/vid/research/research-projects/dipas/ **Access information:** only meta-data and publications are freely available for download					
Purpose	**Timeliness**	Trustworthiness	Disaggregation	**Summary rating:** Green/amber	
Population and definitions	Transparency	Completeness N/A	Sample design		

22	**IAB-BAMF-SOEP Survey**			**Topic:** destination population; routes and journey	
Source type: survey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** panel data, 2016–2019	**Geography:** Germany
Content description: a panel survey of refugees and asylum seekers, who arrived in Germany since 1 Jan 2013, with data including reason for migration, costs and risk, experiences of journey and integration.					
Notes: focus on understanding the asylum-seeking population and integration of refugees, including Syrians. Methodology and data published; problems with interviewers clearly described and addressed.					
Link: https://www.diw.de/en/diw_01.c.538695.en/research_advice/iab_bamf_soep_survey_of_refugees_in_germany.html. **Access information:** data and publications freely available for download, for data access, see https://fdz.iab.de/en/FDZ_Individual_Data/iab-bamf-soep.aspx					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Green	
Population and definitions	Transparency	Completeness N/A	Sample design		

23	**Syrian Refugees in Germany**			**Topic:** destination population	
Source type: survey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** Sept-Oct 2015	**Geography:** Germany
Content description: survey of 889 Syrian refugees' opinions including reason for fleeing Syria and views on the conflict, aiming to fill information gaps and give refugees a voice					
Notes: a one-off survey, by an organisation aiming to promote refugee rights, specifically concerned with Syrian refugees. Sample design targeted a number of locations, but with no systematic strategy.					
Link: https://adoptrevolution.org/en/survey-amongst-syrian-refugees-in-germany-backgrounds/ **Access information:** summary data available only in aggregate formats (pdf tables)					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Amber	
Population and definitions	Transparency	Completeness N/A	Sample design		

24	**Flight 2.0 / Flucht 2.0**			**Topic:** routes and journey; information	
Source type: survey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** Apr-May 2016	**Geography:** en route
Content description: survey of refugees' use of mobile devices and information including mobiles, media sources of information and levels of trust during journey to Germany. Report in German.					
Notes: a one-off retrospective survey on asylum seekers housed in reception centres in Berlin, including Syrians, based on a quota sample with main distributions matched with register.					
Link: https://www.polsoz.fu-berlin.de/en/kommwiss/arbeitsstellen/internationale_kommunikation/Forschung/Flucht-2_0/index.html **Access information:** report on methods and key results available.					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Amber	
Population and definitions	Transparency	Completeness N/A	Sample design		

25	**MedMig**			**Topic:** routes and journey; Policy; Information	
Source type: interviews	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** 2015–16	**Geography:** Mediterranean
Content description: interviews with 500 migrants in Italy, Greece, Malta and Turkey during 2015, including reason for migration, experience of violence, use of media/information[(1)], networks, intentions.					
Notes: a one-off study, aiming for academic understanding of the asylum seeking population, including Syrian refugees. Data disaggregated by nationality and arrival location. Methods and results published.					
Link: https://www.compas.ox.ac.uk/project/unravelling-mediterranean-migration-crisis-medmig **Access information:** only publications are available for download					
Purpose	**Timeliness**	Trustworthiness	Disaggregation	**Summary rating:** Green/amber	
Population and definitions	Transparency	Completeness N/A	Sample design		

[(1)] For a related study, see also: http://www.open.ac.uk/ccig/research/projects/mapping-refugee-media-journeys

26	**Evi-Med**			**Topic:** routes and journey	
Source type: mixed survey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** 2016	**Geography:** Mediterranean
Content description: survey of 750 migrants and 45 in-depth interviews across Sicily, Greece and Malta including reason for migration and experience of journey.					
Notes: a one-off survey aimed to provide insights into the situation of asylum seekers, specifically Syrians, and impacts on countries of arrival. Minimal description; number of locations targeted but no systematic strategy. Value added in the description of reception systems in the three countries.					
Link: https://evimedresearch.wordpress.com/ **Access information:** publications and briefings only available for download					
Purpose	**Timeliness**	Trustworthiness	Disaggregation	**Summary rating:** Amber	
Population and definitions	Transparency	Completeness N/A	Sample design		

27	**4Mi**			**Topic:** routes and journey	
Source type: mixed survey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** since 2014	**Geography:** Africa/Europe
Content description: The *Mixed Migration Monitoring Mechanism Initiative* (4Mi) – information from 3522 interviews plus survey data of migrants, smugglers and observers across (East) Africa and Europe.					
Notes: aims to understand various aspects of migrant journeys; data from two phases (2014–17 and 2017 onwards), aggregated by phase. Data lacking some detail, top-tens reported. Does not concern the Syrian population. Methodology explicitly described, heavily reliant on monitor/observer reports.					
Link: https://mixedmigration.org/4mi/ **Access information:** information available via an interactive online interface					
Purpose	Timeliness	Trustworthiness	Disaggregation	**Summary rating:** Amber	
Population and definitions	Transparency	Completeness N/A	Sample design		

28	**IMPALA**			**Topic:** policy	
Source type: legal	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** 1960 onwards	**Geography:** 20 countries
Content description: database of trends in immigration selection, naturalization, illegal immigration policy and bilateral agreements across 20 migrant-receiving OECD countries, across time.					
Notes: aims to understand migration policies and their impact, and specifically includes policy on asylum and other types of forced migration. Public release of data delayed, as of 1 May 2019.					
Link: http://www.impaladatabase.org/ **Access information:** key publications only available for download					
Purpose	Timeliness	Trustworthiness	Disaggregation N/A	**Summary rating:** Green/amber[(1)]	
Population and definitions	Transparency	Completeness N/A	Sample design N/A		

[(1)] Potential rating, once the data are released, based on available meta information and documentation

B2. Supplementary General Sources on Migration Processes, Drivers or Features

S01	**PROMINSTAT**			**Topic:** meta-information	
Source type: review	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** mostly annual	**Geography:** EU countries
Content description: legacy website of an important FP6 project, focusing on providing information and meta-information on "the scope, quality and comparability of statistical data collection on migration in a wide range of thematic fields", including flows, stocks and various characteristics, across Europe. The scope covers registers, counts, censuses and sample surveys. The reports are current as of ca. 2009.					
Link: http://www.prominstat.eu/drupal/node/64 **Access information:** reports and meta-information publicly available for download					

S02	**OpenStreetMap**			**Topic:** routes, origin, destination context	
Source type: map	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** continuously updated	**Geography:** global
Content description: map data built by contributors using aerial imagery, GPS devices, and low-tech field maps to maintain and update data.					
Link: https://www.openstreetmap.org/ **Access information:** open data publicly available for download					

S03	**MAFE Project**			**Topic:** destination population, origin population, routes	
Source type: survey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** 2005–2012	**Geography:** 3+6 countries
Content description: the *Migrations between Africa and Europe* (MAFE) Project contains multi-level surveys carried out at sending and receiving ends of migration from Congo, Ghana and Senegal to 6 EU countries. The survey focuses on migration patterns, routes, drivers, as well as socio-demographic impacts. MAFE data have been used in agent-based modelling of international migration by F Willekens and A Klabunde.					
Link: https://cordis.europa.eu/project/id/217206 **Access information:** data freely available for download for research and educational purposes					

S04	**Mexican Migration Project**			**Topic:** destination population, origin population, routes	
Source type: ethnosurvey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** since 1982	**Geography:** Mexico - US
Content description: the *Mexican Migration Project* (MMP) contains detailed and very rich ethnosurvey-based data, quantitative and qualitative, on Mexican migration to the US, collected in parallel from both sides of the border. In general, ethnosurveys combine a quantitative survey with ethnographic methods, and can therefore provide uniquely detailed insights into the mechanisms driving migration flows.					
Link: https://mmp.opr.princeton.edu/ **Access information:** data freely available for download for research and educational purposes					

S05	**Latin American Migration**			**Topic:** destination population, origin population, routes	
Source type: ethnosurvey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** since 1982	**Geography:** 10 origins-US
Content description: parallel to the MMP, the *Latin American Migration Project* (LAMP) contains detailed ethnosurvey data for migration from 10 Latin American origin countries to the US					
Link: https://lamp.opr.princeton.edu/ **Access information:** data freely available for download for research and educational purposes					

S06	**ICMPD Yearbook**			**Topic:** destination population, routes, flows and impacts	
Source type: secondary	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** annual	**Geography:** C-E Europe
Content description: legacy "Annual Yearbook on Illegal Migration, Human Smuggling and Trafficking in Central and Eastern Europe" produced for 1999–2013 by the International Centre for Migration Policy Development in Vienna, compiling different data from migration and border enforcement authorities.					
Link: http://research.icmpd.org/projects/irregular-migration/yearbook/ **Access information:** publications freely available for download					

S07	**Uppsala Conflict Data**			**Topic:** drivers: conflict/violence	
Source type: journalistic	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** daily-annual*	**Geography:** Worldwide
Content description: a comprehensive database of conflict and organized violence-related fatal events, actors, and the numbers of deaths, based on geocoded news items, with some data going back to 1946. * Time granularity depends on a specific dataset. Events reported for >25 deaths.					
Link: https://www.pcr.uu.se/research/ucdp/ **Access information:** data freely available for download					

S08	**ACLED Conflict Data**			**Topic:** drivers: conflict/violence	
Source type: journalistic	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** daily*	**Geography:** selection**
Content description: comprehensive information on “the dates, actors, types of violence, locations, and fatalities of all reported political violence and protest events”, event-centred and with detailed spatial granularity, updated weekly. * Temporal range differs by region. ** Includes conflict-affected countries from Africa, Middle East, South and South East Asia, Europe, and Latin America.					
Link: https://www.acleddata.com/ **Access information:** data spreadsheets and results of queries publicly available for download					

S09	**PITF Worldwide Atrocities**			**Topic:** drivers: conflict/violence	
Source type: journalistic	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** daily	**Geography:** conflict zones*
Content description: the Political Instability Task Force Worldwide Atrocities database, based on geocoded news items, providing information on conflict/violence events, updated monthly. * Includes Syria.					
Link: http://eventdata.parusanalytics.com/data.dir/atrocities.html **Access information:** data spreadsheets are available for download					

S10	**Global Terrorism Database**			**Topic:** drivers: conflict/violence	
Source type: various	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** daily	**Geography:** worldwide
Content description: database of terrorist events including the date, perpetrator and fatalities, based on a range of secondary open sources, from journalistic accounts, to reports and legal documents.					
Link: https://www.start.umd.edu/gtd/ **Access information:** data available for download after pre-registration					

S11	**“The New Odyssey”**			**Topic:** routes and journey	
Source type: journalistic	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** 2012–15	**Geography:** varies
Content description: a comprehensive book containing interviews, anecdotes and observations of the journey of asylum seekers into Europe (also from Syria) including insights into networks, barriers, strategies and resources.					
Reference: Kingsley P (2016) *The New Odyssey. The Story of Europe's Refugee Crisis*. London: Guardian / Faber & Faber.					

S12	**Coming to the UK**			**Topic:** information	
Source type: interviews	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** one-off (2006)	**Geography:** UK
Content description: Gilbert A and Koser K (2006) Coming to the UK: What do Asylum-Seekers Know About the UK before Arrival? *Journal of Ethnic and Migration Studies*, 32(7) 1209–25: interviews with 87 asylum seekers from Afghanistan, Columbia, Kosovo and Somalia about how much they knew about the UK before arrival.					
Link: https://www.tandfonline.com/doi/figure/10.1080/13691830600821901 **Access information:** publication access available for JEMS subscribers					

S13	**GLMM Syrian migration**			**Topic:** destination population	
Source type: administrative	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** 2010–13	**Geography:** Gulf States
Content description: Gulf Labour Markets and Migration report on Syrian Refugees in the Gulf until 2013, by F De Bel-Air, reviewing selected annual official data from Gulf countries on Syrian migration.					
Link: https://gulfmigration.org/media/pubs/exno/GLMM_EN_2015_11.pdf **Access information:** publication freely available for download					

S14	**RRE Life in Limbo**			**Topic:** destination populations	
Source type: interviews	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** one-off (2016)	**Geography:** Greece
Content description: a Refugee Rights Europe publication reporting on a dedicated survey carried out amongst asylum seekers in Greece, containing potentially relevant process and contextual information.					
Link: http://refugeerights.org.uk/reports/ > Life in Limbo (and other reports) **Access information:** all publications freely available for download					

S15	**Fortress Europe blog**			**Topic:** routes and journeys	
Source type: journalistic	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** 1988–2016	**Geography:** Mediterranean
Content description: compilation of news reports on migrant deaths at European borders, with individual dates reported, with the aim to fill information gaps. Publicly available news reports, varying journalistic standards, sometimes including. Sometimes includes specifically Syrians. Content in Italian.					
Link: http://fortresseurope.blogspot.com/ **Access information:** information freely available					

S16	**IMPIC**			**Topic:** policy	
Source type: legal	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** 1980–2010	**Geography:** OECD
Content description: *Immigration Policies in Comparison*: a legacy database, comparing migration policies of 33 OECD countries, aimed at better understanding migration policies and their impact.					
Link: http://www.impic-project.eu/ **Access information:** dataset freely available, also for quantitative analysis					

S17	**Migration Policy Centre**			**Topic:** routes and journey; drivers: conflict/violence; policy	
Source type: secondary	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** intra-month	**Geography:** Syria
Content description: contextual information of the Migration Policy Centre, with the timeline of the Syrian conflict and policy responses, based on journalistic accounts and legal documents					
Notes: information collated on events related to Syrian migration, with a selection of individual dates reported, ending in 2016; based on publicly-available news reports of varying journalistic standards					
Link: http://syrianrefugees.eu/ **Access information:** information available via an interactive online interface					

S18	**IOM Impact Evaluation study**			**Topic:** information	
Source type: RCT survey	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** Oct-Nov 2018	**Geography:** Senegal
Content description: a one-off impact evaluation study, employing a survey-based randomized control trial (RCT) amongst the participants of IOM information and intervention programmes, aiming to assess the efficiency of peer-to-peer information campaigns about the reality between prospective migrants from Senegal.					
Link: https://publications.iom.int/books/migrants-messengers-impact-peer-peer-communication-potential-migrants-senegal-impact. **Access information:** individual-level data are not publicly available, but the report and the accompanying technical annex contain aggregate results tables.					

S19-S20	**Global flow estimates**			**Topic:** flows	
Source type: stock data	Quantitative Qualitative	Process Context	Macro-level Micro-level	**Time detail:** five-yearly	**Geography:** global
Content description: two sets of global migration flow estimates (five-year transitions), linked to two articles on deriving migration flow estimates consistent with population and migrant stock data from the UN: [1] Abel GJ and Sander N (2014) Quantifying Global International Migration Flows. *Science*, 343(6178), 1520-1522, and [2] Azose JJ and Raftery AE (2019) Estimation of emigration, return migration, and transit migration between all pairs of countries. *Proceedings of the National Academy of Sciences, USA*, 116(1), 116-122.					
Links: https://science.sciencemag.org/content/343/6178/1520.abstract (Abel & Sander 2014) http://download.gsb.bund.de/BIB/global_flow/ (database for Abel & Sander 2014) https://www.pnas.org/content/116/1/116 (Azose and Raftery 2019, including data) **Access information:** open source data and publications available via the links above					

Appendix C. Uncertainty and Sensitivity Analysis: Sample Output

Jakub Bijak and Jason Hilton

This Appendix supplements information provided in Chap. 5, by offering some additional detail on the statistical analysis of the first version of the model, as well as including selected result tables. In particular, the contents include: the results of the initial pre-screening of model inputs, following the Definitive Screening Design of Jones and Nachtsheim (2011, 2013); outputs of the uncertainty and sensitivity analysis, carried out after fitting a Gaussian Process (GP) emulator on the reduced set of inputs in the GEM-SA package (Kennedy & Petropoulos, 2016); and sets of results – predictions of model outputs for the most important input pairs, carried out for three additional output variables, supplementing the results reported in Chap. 5. The pre-screening, uncertainty and sensitivity analysis have been carried out for four outputs: the mean share of time, in which the agents follow their route plan (*mean_freq_plan*), standard deviation of the number of visits over all links (*stdd_link_c*), correlation of the number of passages over links with the optimal scenario (*corr_opt_links*) and standard deviation of traffic between replicate runs (*prop_stdd*).

To start with, Table C.1 offers brief information about selected software packages for carrying out experimental design analysis, emulation, sensitivity and uncertainty analysis, and model calibration. In terms of the results of the analysis, Table C.2 includes detailed results of the model pre-screening exercise, described in Sect. 5.2. The initial set of 17 parameters of potential interest is analysed with respect to how much they contribute – individually and jointly – to the overall variance of the model output. The model construction, including a description of variables, is described in more detail in Chap. 3 and Appendix A. More specific information about the model architecture is provided in Appendix B, and the Julia code for reproduction and replication purposes is available from the online repository: https://github.com/mhinsch/RRGraphs_CT also hyperlinked from the project website www.baps-project.eu (as of 1 August 2021).

The pre-screening has been done in GEM-SA, with two separate sets of results obtained by using different random seed (the second one labelled RSeed2), as well as in R, by using a standard analysis of variance (ANOVA) routine. The GP emulators for the pre-screening have been fitted based on a Definitive Screening Design space of 37 points, with ten replicates at each design point for three outputs (*mean_freq_plan*, *stdd_link_c*, *corr_opt_links*), and one replicate per point for *prop_stdd*. For each output, the precise numerical results differ somewhat between the three pre-screening attempts (GEM-SA, RSeed2 and ANOVA), but the qualitative conclusions are the same: they all point to the same sets of key inputs for each output variable, mostly concentrated on variables related to information transfer and errors (see Chap. 5).

The results for uncertainty, sensitivity and emulator fit are reported in Table C.3, for two sets of assumptions on the input priors: normal and uniform, with qualitative results (i.e. the key variables of influence) largely remaining robust to the prior

Table C.1 Selected software packages for experimental design, model analysis, and uncertainty quantification

Software	Description	URL
R packages	**R packages related to uncertainty quantification**	https://cran.r-project.org/
lhs	Package for creating Latin hypercube samples	…/package=lhs/
AlgDesign	Package for creating different (algorithmic) experimental designs, including factorial ones	…/package=AlgDesign/
DiceKriging	Package for estimating and analysing computer experiments with non-Bayesian kriging models	…/package=DiceKriging/
rsm	Package for generating response surface models, creating surface plots	…/package=rsm/
tgp	Treed GPs: package for a general, flexible, non-parametric class of meta-models	…/package=tgp/
BACCO	Toolkit for applying the Kennedy and O'Hagan (2001) framework to emulation and calibration	…/package=BACCO
gptk	GP Toolkit: package for a range of GP-based regression model functions	…/package=gptk/
GEM-SA	Gaussian Emulation Machine for Sensitivity Analysis (see Kennedy & Petropoulos, 2016)	http://www.tonyohagan.co.uk/academic/GEM
Gaussian Processes	A repository of links to various GP-related routines, mainly in Matlab, Python and C++	http://www.gaussianprocess.org/
UQLab	Comprehensive, general-purpose software for uncertainty quantification, based on Matlab	https://www.uqlab.com/

Source: own elaboration. Links current as of 1 February 2021

specification. The heatmaps of means and standard deviations of the emulator-based predictions are shown in Figs. C.1, C.2 and C.3 for three outputs (*stdd_link_c*, *corr_opt_links* and *prop_stdd*), with the fourth one (*mean_freq_plan*) reported in Chap. 5 (Fig. 5.5). For each output except for *prop_stdd*, the emulators are fitted for six replicates at each Latin Hypercube Sample design point, with 65 points in total, whereas for *prop_stdd*, the design sample is limited to 65 points, given the cross-replicate nature of this output.

Table C.2 Pre-screening for the Routes and Rumours (data-free) version of the migrant route formation model: Shares of variance explained under the Definitive Screening Design, per cent

	mean_freq_plan			stdd_link_c			corr_opt_links			prop_stdd[a]		
Input\output	GEM-SA	RSeed2	ANOVA	GEM-SA	RSeed2	ANOVA	GEM-SA	RSeed2	ANOVA	GEM-SA	Rseed2	ANOVA
p_keep_contact	1.3269	1.1639	1.1105	2.6553	1.8634	2.2230	1.0616	0.7624	0.0647	0.5734	1.8076	0.2523
p_drop_contact	0.4707	0.8016	0.4519	**12.2928**	**11.7525**	**11.6759**	0.2352	0.8298	0.0602	1.3993	**5.6776**	2.3173
p_info_mingle	**6.0680**	3.3479	**5.5556**	0.5464	0.5644	0.3367	0.3170	1.0963	0.0074	0.8622	1.5881	0.0000
p_info_contacts	**18.4436**	**9.1519**	**15.2193**	**12.9457**	**11.0136**	**11.3739**	0.6759	0.9392	0.0289	3.0982	**7.0421**	**6.5203**
p_transfer_info	**59.6102**	**41.2814**	**67.8690**	**20.6097**	**18.5105**	**20.1560**	0.1713	0.9565	0.0054	4.6738	**9.7578**	**6.4172**
Error	0.2538	0.5488	0.0669	**36.4927**	**36.6574**	**35.0669**	**76.4010**	**14.7351**	**68.0989**	**20.1476**	**43.4140**	**51.6134**
p_find_links	0.4081	0.5613	0.1191	0.2482	0.9514	0.0002	2.6256	1.3189	2.8085	2.3934	3.8855	3.3441
p_find_dests	1.4957	0.6918	0.6128	0.1820	0.2583	0.0479	0.7313	0.8168	0.5810	0.8978	1.5653	0.1153
speed_expl_stay	0.5536	0.6144	0.3574	1.0826	1.0858	0.7769	1.3740	0.8992	0.7412	1.1358	2.2779	1.2338
speed_expl_move	0.2003	0.4141	0.0003	0.5533	0.6358	0.2910	2.4765	0.8633	0.2852	1.6236	2.8415	2.4769
qual_weight_x	0.2386	0.5018	0.0050	0.9840	1.1857	0.7549	0.3081	1.0565	0.1810	1.5166	**5.9727**	3.4927
qual_weight_res	0.4842	0.6208	0.1180	0.2899	0.4654	0.1639	0.8429	0.9435	0.7635	0.8720	2.4428	0.2459
path_weight_frict	0.8701	0.8619	0.5510	1.5559	1.4843	1.3565	0.2935	0.9555	0.0574	1.0483	2.7944	1.7892
weight_traffic	0.5218	0.6555	0.2951	0.2848	0.3933	0.0462	0.5493	0.9833	0.3406	1.0284	1.8629	0.0022
costs_stay	0.2331	0.4304	0.0480	0.2707	0.8047	0.0238	0.2927	0.9534	0.0633	0.8261	3.5357	0.1534
costs_move	2.0486	0.4799	0.0025	2.6870	3.5960	1.6021	0.3941	0.9004	0.2902	2.4457	1.6440	0.7625
ben_resources	1.4450	0.5813	0.1549	0.3710	0.5254	0.0953	4.4356	1.1965	0.0963	1.1298	1.8900	0.0177
Interactions	4.7815	**6.5045**	0.0000	**5.7987**	**7.5184**	0.0000	**6.3422**	**12.4944**	0.0000	**10.9863**	0.0000	0.0000
Total % explained:	**99.45**	**69.21**	**92.54**	**99.85**	**99.27**	**85.99**	**99.53**	**42.70**	**74.47**	**56.66**	**100.00**	**80.75**

Notes: for each output, the sensitivity was assessed three times: two times in GEM-SA (Kennedy & Petropoulos, 2016), under two different random seeds, and through ANOVA in R. The values **in bold** correspond to the inputs with high (>5%) share of the variance attributed to individual variables

[a]The experiments were run on 37 Definitive Screening Design points: for *prop_stdd* one repetition per point, for all other outputs ten per point Source: own elaboration in GEM-SA and R

Table C.3 Key results of the uncertainty and sensitivity analysis for the Routes and Rumours (data-free) version of the migration model

Input assumptions:	Normal prior				Uniform prior			
Sensitivity analysis								
Input\Output	*mean_freq_plan*	*stdd_link_c*	*corr_opt_links*	*prop_stdd*[a]	*mean_freq_plan*	*stdd_link_c*	*corr_opt_links*	*prop_stdd*[a]
p_drop_contact	0.2906	4.4151	0.8502	**5.0949**	0.4802	**5.7799**	1.0706	**5.2907**
p_info_mingle	**6.9762**	4.8300	**9.5387**	**9.4807**	9.6796	**5.8638**	**8.5181**	**11.3356**
p_info_contacts	**9.3481**	0.3196	3.2030	**6.1303**	**9.5604**	0.3077	3.7471	2.8154
p_transfer_info	**71.7956**	**22.2017**	**44.4823**	**30.7951**	**66.7805**	**16.1100**	**39.0571**	**21.6262**
Error	0.0990	**27.5070**	**18.7827**	**7.1979**	0.1130	**24.9533**	**17.6223**	**7.6150**
Exploration	0.7538	3.0526	2.3338	4.0429	0.5882	3.6425	2.9926	4.1407
Interactions	**8.6676**	**26.8109**	**16.6100**	**33.0812**	**9.5531**	**28.1409**	**19.7976**	**39.5982**
Residual	2.0692	10.8631	4.1992	4.1771	3.2450	15.2020	7.1946	7.5782
Total % explained	**97.9308**	**89.1369**	**95.8008**	**95.8229**	**96.7550**	**84.7980**	**92.8054**	**92.4218**
Uncertainty analysis								
Mean of expected code output	0.4296	50.5192	0.3219	0.0173	0.4130	53.1535	0.3024	0.0178
Variance of expected code output	0.0000	1.2563	0.0000	0.0000	0.0000	1.2931	0.0000	0.0000
Mean total variance in code output	0.0068	278.7480	0.0141	0.0000	0.0080	348.9120	0.0161	0.0001
Fitted sigma^2	1.1363	1.4992	1.6505	4.1507	1.1363	1.4992	1.6505	4.1507
Nugget sigma^2	0.0094	0.0203	0.0194	0.2307	0.0094	0.0203	0.0194	0.2307
RMSE	0.0069	3.7551	0.0199	0.0085	0.0069	3.7551	0.0199	0.0085
RMSPE (%)	2.72%	4.77%	7.89%	74.51%	2.72%	4.77%	7.89%	74.51%
RMSSE (standardised)	1.5894	1.9498	1.6701	1.7812	1.5894	1.9498	1.6701	1.7812

[a]The experiments were run on 65 Latin Hypercube Sample design points: for prop_stdd one repetition per point, for all other outputs six per point. The values **in bold** denote inputs with high (>5%) share of attributed variance. Source: own elaboration in GEM-SA (Kennedy & Petropoulos, 2016)

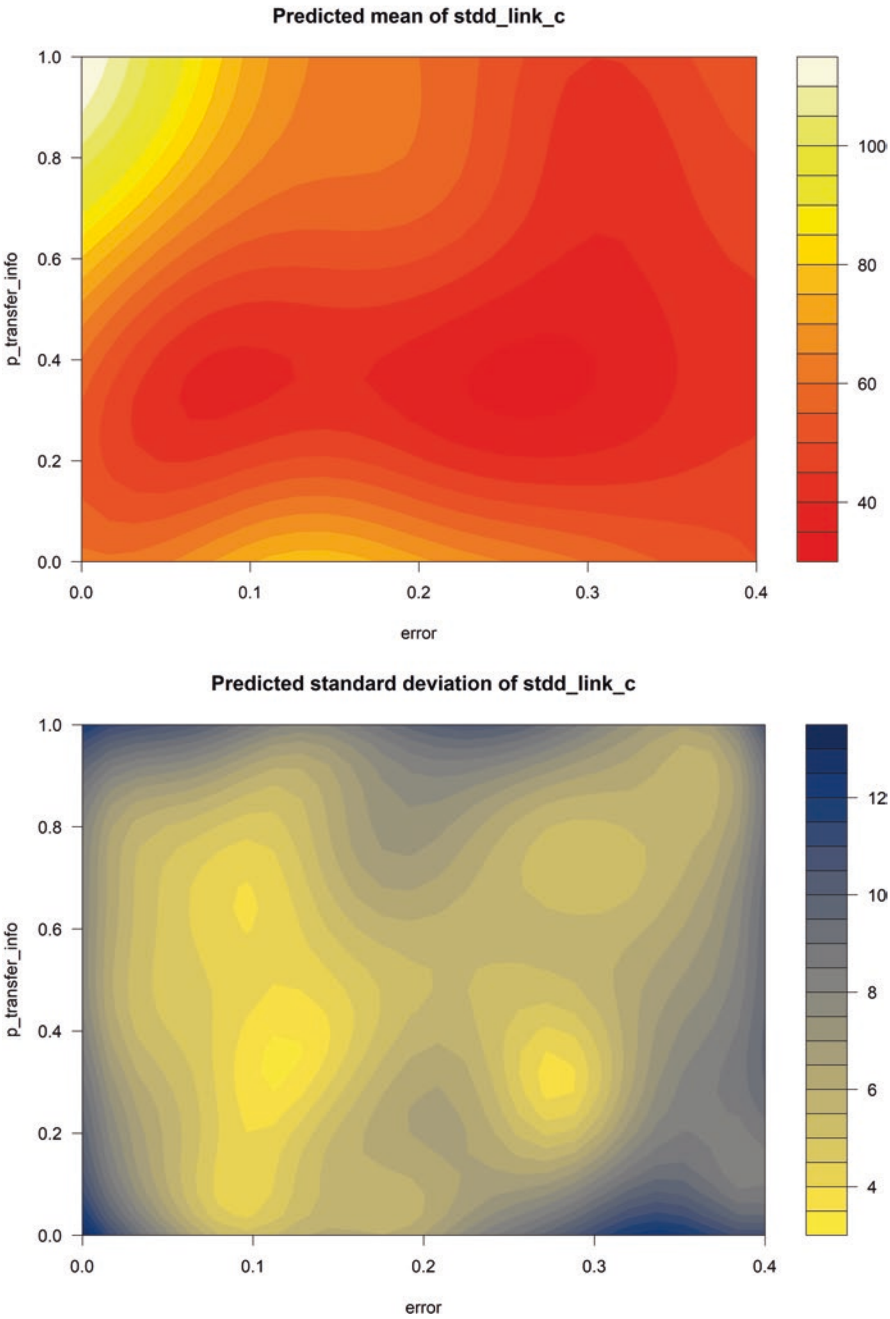
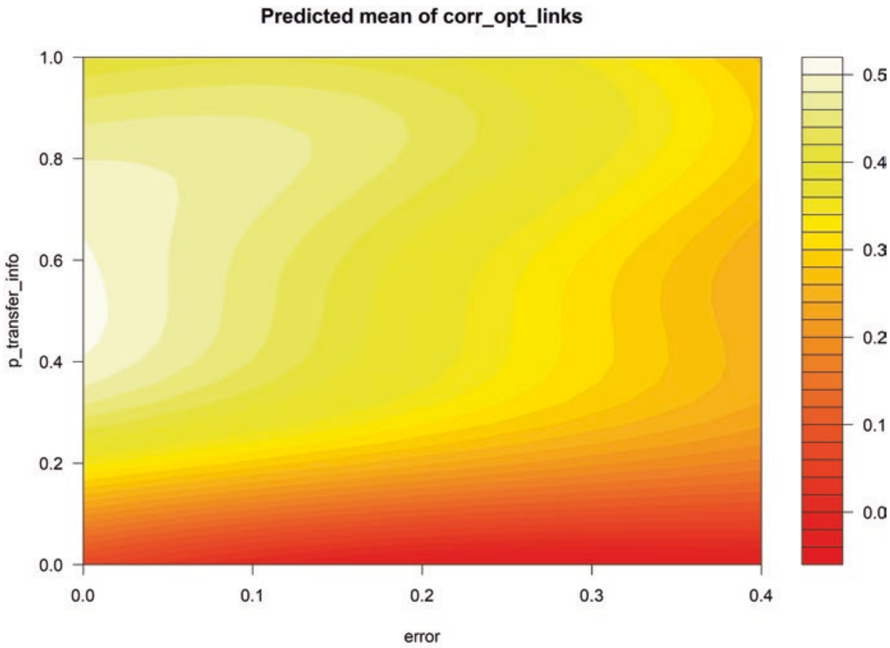


Fig. C.1 Estimated response surface of the standard deviation of the number of visits over all links vs two input parameters, probabilities of information transfer and information error: mean (top) and standard deviation (bottom). Source: own elaboration



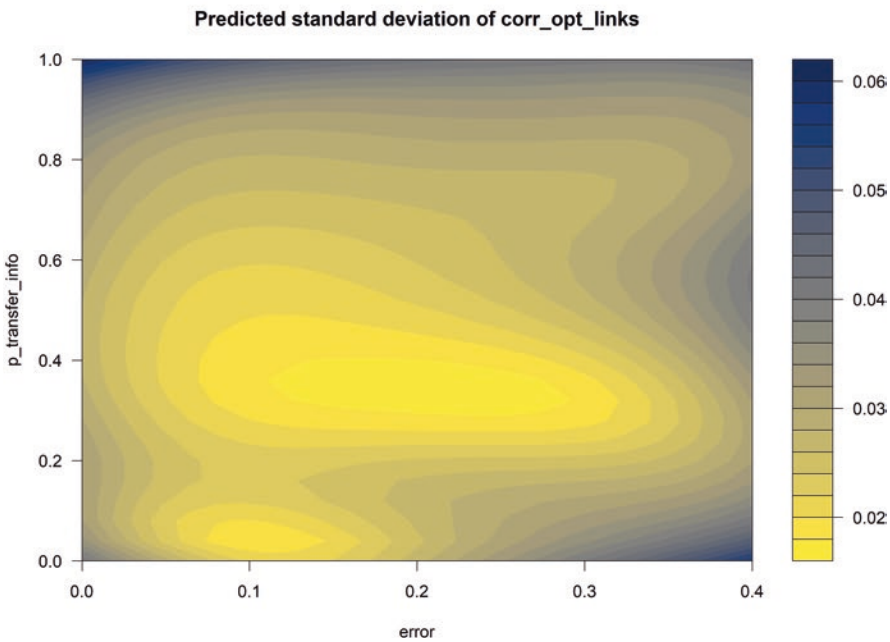


Fig. C.2 Estimated response surface of the correlation of the number of passages over links with the optimal scenario vs two input parameters, probabilities of information transfer and information error: mean (top) and standard deviation (bottom). Source: own elaboration

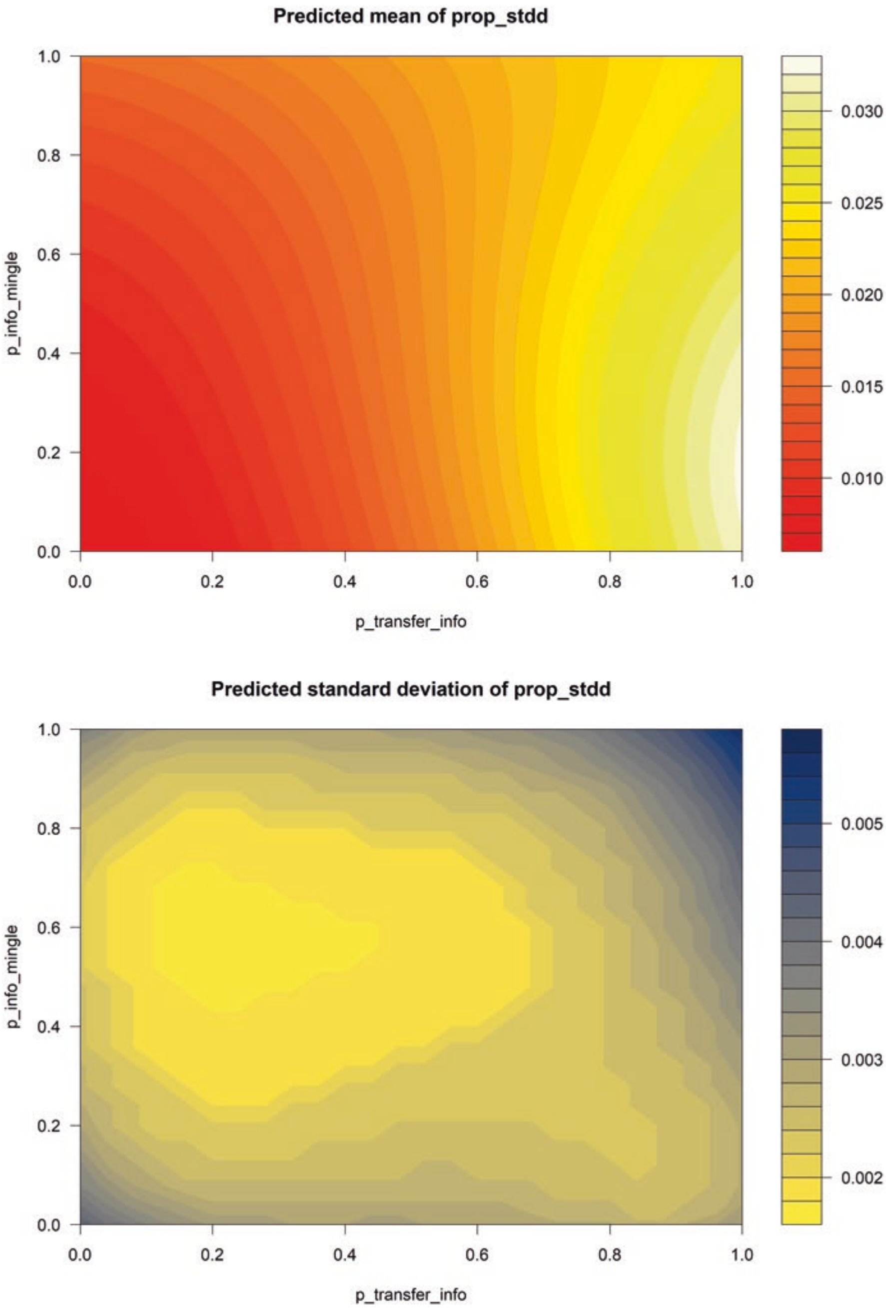


Fig. C.3 Estimated response surface of the standard deviation of traffic between replicate runs vs two input parameters, probabilities of information transfer and of communication with local agents: mean (top) and standard deviation (bottom). Source: own elaboration

Appendix D. Experiments: Design, Protocols, and Ethical Aspects

Toby Prike

This Appendix supplements information provided in Chap. 6, by offering more detailed information on the preregistration of the individual research hypotheses (for a broader discussion of the need for preregistration in the context of experimental psychology and tools for ensuring the reproducibility and replicability of results, see e.g. Nosek et al., 2018 and Chap. 10 in this book), number of participants, and ethical issues for the experiments reported in the chapter. This Appendix covers in more detail the first three experiments presented in Chap. 6, that is, the elicitation of the prospect curves and utility functions in a discrete-choice framework, enquiries into subjective probabilities and risk attitudes, and their relationships with the source of information received, as well as the conjoint analysis of migration drivers.

In terms of organisation and execution, live, lab-based experiments carried out in controlled conditions on undergraduate participants recruited from the University of Southampton were only conducted for the first experiment, on eliciting the prospect curves. For that experiment, the sample size was 150 participants. The online experiments, for all three studies reported in Chap. 6 and in this Appendix, were implemented in Qualtrics and executed via the Amazon Mechanical Turk (the first two experiments) and Prolific environments (the third one),[1] with specific details discussed separately for each experiment. For these three online experiments, related to eliciting the information related to prospect theory, subjective probability questions, and conjoint analysis of migration drivers, their sample sizes were equal to 400, 1000 and 1000 participants, respectively.

The links below provide more specific information: the Open Science Framework links include the study preregistrations, anonymised data, and analysis code for the individual studies, while the experimental links offer a way of taking part in 'dry run' experiments, with no data being collected.

D.1. Prospect Theory and Discrete Choice Experiment

Experiment Link:
https://southampton.qualtrics.com/jfe/form/SV_e9uicjzpa30RDeu
Open Science Framework Link:https://osf.io/vx4d9/

Because the research in this study involved participants making choices between gambles, there was the potential that it could cause harm or distress to some

[1] See https://www.mturk.com/ and https://www.prolific.co/ (as of 1 June 2021).

participants, especially in the context of possible problem gambling. However, the exposure to gambling within this study was fairly mild, and it is likely that participants regularly receive greater exposure to gambling-related themes in their everyday lives (e.g., via television advertisements).

To minimise the risk that exposure to gambling might cause harm or distress to participants, the advertisement and participant information sheet clearly outlined that the study involved making choices between gambles. We also recommended that participants did not participate if they had a history of problem gambling and/or believed that participating in this study was likely to cause them distress or discomfort. Additionally, we provided links to relevant support services on both the participant information sheet and the debriefing sheet. Finally, we screened participants for problem gambling using the Brief BioSocial Gambling Screen, developed by the Division on Addiction at Cambridge Health Alliance,[2] and any participants who answered 'yes' to a related question, indicating that they are at risk of problem gambling, were redirected to a screen indicating that they were ineligible to participate in the study and noting that the screening tool is not diagnostic.

This study has received approval from the University of Southampton Ethics Committee, via the Ethics and Research Governance Online (ERGO) system, submissions number 45553 (lab-based version of the experiment) and 45553.A1 (amendment extending the research to an online study, via the Amazon Mechanical Turk platform). The lab-based data collection took place in November 2018, and the online collection in May and June 2019.

D.2. Eliciting Subjective Probabilities

Experiment Link:
https://southampton.qualtrics.com/jfe/form/SV_20kQsSP0cyi6o06
Open Science Framework Link:https://osf.io/3qrs8

In this study, the salience of the topics (risk involved in migration and travel during a pandemic) in the public consciousness, and the general, high-level formulation of the individual tasks, questions and responses, without specific recourse to individual experience, meant that the ethical issues were minimal. Any residual issues were controlled through an appropriate research design, participant information and debriefing, which can be seen under the experiment link above. This study has received approval from the University of Southampton Ethics Committee, via the Ethics and Research Governance Online (ERGO) system, submission number 56865. Given that the timing of data collection coincided with the COVID-19 pandemic of 2020, the experiments were carried out exclusively online, via Amazon Mechanical Turk. The data collection took place in June 2020.

[2] See the version cited on https://www.icrg.org/resources/brief-biosocial-gambling-screen (as of 1 February 2021).

D.3. Conjoint Analysis of Migration Drivers

Experiment Link:
https://southampton.qualtrics.com/jfe/form/SV_2h4jGJH1PA9qJsq
Open Science Framework Link:https://osf.io/ayjcq/

In this study, we asked about aspects of a country that influence its desirability as a migration destination. Because the migration drivers and countries were included at an abstract level and without specific recourse to individual experience, the ethical issues were minimal. Any residual issues were controlled through an appropriate research design, participant information and debriefing, which can be seen under the experiment link above. This study has received approval from the University of Southampton Ethics Committee, via the Ethics and Research Governance Online (ERGO) system, submission number 65103. Given that the timing of data collection coincided with the COVID-19 pandemic, the experiments were carried out exclusively online, via the Prolific platform. The data collection took place in October 2021.

Appendix E. Provenance Description of the Route Formation Models

Oliver Reinhardt

This Appendix contains supplementary information for Chap. 7, with particular focus on explaining the provenance graph shown in Fig. 7.3, which depicts a sketch of the provenance of the whole research project and the broader model development process (for an early version of the provenance graph, see Bijak et al., 2020). Tables E.1 and E.2 in this Appendix shortly describe the entities and activities shown on the provenance graph, referring to the corresponding parts of this book and to outside sources with more detailed information, where relevant.

Thus, the structure of the provenance graph presented in Fig. 7.3 in Chap. 7 roughly reflects the key components of the model-building process and its constituting elements, with the model development in the middle panel, surrounded by model analysis, data collection and assessment, psychological experiments, and policy scenarios. The modelling panel shows five iterations of model development (m1 to m5) resulting in five successive model versions (M1 to M5), each improving on the previous one with respect to the degree of realism and usefulness, in line with the (classical) inductive philosophical tenets of the model-based research programme (Chap. 2).

The model-building process additionally includes the re-implementation of the model in the domain-specific modelling language ML3 (m2′, resulting in the model version M2′, and later M3′) discussed in Chap. 7. The data panel mentioned above that depicts the collection and assessment of the relevant data (see Chap. 4). Here, only those data that ended up being used in the modelling work are included. Next to the data, the policy-relevant scenarios described in Chap. 9 are shown. The model analysis panel, in turn, shows the simulation experiments and analysis that were conducted on the successive model versions. Finally, the bottom panel presents the parallel work on psychological experiments (see Chap. 6), with three phases of experiments discussed in Sects. 6.2, 6.3 and 6.4. Of those, the second experiment – on eliciting subjective probabilities and the role of information sources – ended up being used in the model (versions M4 and M5).

At this level of detail, the provenance model does not document the model development in detail (as does the meta-modelling and sensitivity example in Fig. 7.2 in Chap. 7), but gives a broad overview of the simulation study and model-building process as a whole. In a digital version of the provenance model, the modellers and users might be able to zoom in to specific processes or areas of the graph, in order to see them in more detail. In that vein, Fig. 7.2 then becomes a zoomed-in version of a2, with M3 and S1 in Fig. 7.3 corresponding to M and S in Fig. 7.2.

Table E.1 Entities in the provenance model presented in Fig. 7.3

Entity	Description
A16	Methodology of (Abdellaoui et al., 2016)
AF	Data assessment framework (Sect. 4.4)
AR	Probability distribution representing bias and variance of data on sea arrivals in Italy
B09	Review of the role of source used to inform ex2 study design (Briñol & Petty, 2009)
B17	Previous quality assessment frameworks in the literature, e.g., (Bijak et al., 2017)
C20	Review of migration drivers used to inform ex3 study design (Czaika & Reinprecht, 2020)
DT	IOM displacement tracker data (see Appendix B – Source 13)
DTA	Assessment of IOM displacement tracker (see Appendix B – Source 13)
EWS	Model-based early warning system (Box 9.1)
F2	Flight 2.0 data (see Appendix B – Source 24)
F2A	Assessment of flight 2.0 (see Appendix B – Source 24)
H15	Conjoint analysis paper used to inform ex3 study design (Hainmueller et al., 2015)
ID	Probability distributions representing bias and variance of data on interceptions by Libyan and Tunisian coastguards and deaths in the Central Mediterranean
K01	Methodology of Kennedy and O'Hagan (2001)
M1	Initial model version (grid-based, discrete time) (Bijak et al., 2020)
M2	Second model version (graph-based, discrete time) (Bijak et al., 2020)
M2'	Reimplementation of M2 in ML3 (Reinhardt et al., 2019)
M3	Routes and Rumours (graph-based, discrete event) (Sect. 3.3)
M3'	ML3 version of routes and Rumours (Sect. 7.2)
M4	Risk and Rumours (Chap. 8, Sect. 8.3)
M4'	Version of M4 including the proposed intervention (Box 9.3)
M5	Risk and Rumours with reality (Chap. 8, Sect. 8.4)
M5'	Calibrated risk and Rumours with reality (using ABC) (Sect. 8.4)
M5"	Calibrated risk and Rumours with reality (using GP) (Sect. 8.4)
M5"'	Version of M5 including the proposed intervention (Chap. 9, Sect. 9.3)
MM	IOM missing migrants data (see Appendix B – Sources 11/12)
MMA	Assessment of IOM missing migrants (see Appendix B – Sources 11/12)
NSR	Non-scientific reports about migration route formation (e.g., Kingsley, 2016; Emmer et al., 2016)
OSM	OpenStreetMap city locations via OpenRouteService (see Appendix B – S02)
PI	Proposed intervention: Public information campaign (Box 9.3)
PT	Prospect theory (Kahneman & Tversky, 1979) as the theoretical foundation of ex_1
R1	OpenScienceFramework repository for ex1 (preregistration, data, code): https://osf.io/vx4d9/

(continued)

Table E.1 (continued)

Entity	Description
R2	OpenScienceFramework repository for ex2 (preregistration, data, code): https://osf.io/ws63f/
R3	OpenScienceFramework repository for ex3 (preregistration, data, code): https://osf.io/ayjcq/
RF	Risk functions derived from the subjective probabilities (Box 6.1)
RQ1	Research question: Does information exchange between migrants play a role in the formation of migration routes? (Box 3.1)
RQ2	Research question: How do risk perception and risk avoidance affect the formation of migration routes? (Chap. 8)
RQ3	Research question: In a realistic scenario, can more information lead to fewer fatalities? (Chap. 9, Sect. 9.3)
RW	Relative weights of migration drivers
S1	Sensitivity information about all 17 parameters of the routes and Rumours model (Box 5.1)
S2	Sensitivity information about the routes and Rumours model (Box 5.3)
S3	Sensitivity information about the risk and Rumours model (Table 8.2)
S4	Sensitivity information about the risk and Rumours with reality model (Table 8.3)
SCI	Scenario inputs (Box 9.2)
SCO	Scenario outcomes (Box 9.2)
SIO	Simulated intervention outcomes (Box 9.3)
SIO’	Simulated intervention outcomes (Box 9.4)
SP	Subjective probabilities elicited in the second experiment (Sect. 6.3)
SR	Scientific reports about migration route formation, e.g., (Massey et al., 1993; Castles, 2004; Alam & Geller, 2012; Klabunde & Willekens, 2016; Wall et al., 2017)
SU1	Survey (demonstration link: https://sotonpsychology.eu.qualtrics.com/jfe/form/SV_e4FTbu1MidTCsyW)
SU2	Survey (demonstration link: https://sotonpsychology.eu.qualtrics.com/jfe/form/SV_41PZg9XavyKFNl3)
SU3	Survey (demonstration link: https://sotonpsychology.eu.qualtrics.com/jfe/form/SV_cMzaslXJ47MrErk)
U2	Uncertainty information about the routes and Rumours model (Box 5.3)
U3	Uncertainty information about the risk and Rumours model (Table 8.2)
U4	Uncertainty information about the risk and Rumours with reality model (Table 8.3)
UF	Utility functions the first experiment (Sect. 6.2)
W19	Paper on interpreting verbal probabilities used to inform ex2 study design (Wintle et al., 2019)

Table E.2 Activities in the provenance model presented in Fig. 7.3

Activity	Description
a1	Preliminary screening of the routes and Rumours model on all 17 model parameters (Box 5.1)
a2	Uncertainty and sensitivity analysis of the Routes and Rumours model (Box 5.2)
a3	Uncertainty and sensitivity analysis of the Risk and Rumours model (Sect. 8.3)
a4	Uncertainty and sensitivity analysis of the Risk and Rumours with Reality model (Chap. 8, Sect. 8.4)
cal1	Calibrating M5 using ABC (Sect. 8.4)
cal2	Calibrating M5 using GP (Sect. 8.4)
da1	Assessing the flight 2.0 data
ar	Deriving the arrival probability, AR
da2	Assessing the IOM Missing Migrants data
da3	Assessing the IOM Displacement Tracker data
daf	Designing the data quality assessment framework (Chap. 4)
ex1	Designing and conducting of the first round of experiments (Sect. 6.2)
ex2	Designing and conducting of the second round of experiments (Sect. 6.3)
ex3	Designing and conducting of the third round of experiments (Sect. 6.4)
g1	Identifying a knowledge gap in M3
g2	Identifying a knowledge gap in M4
id	Deriving the probability of death, ID
m1	Creating the initial model version (Bijak et al., 2020)
m2	Creating the second model version, Routes and Rumours (Bijak et al., 2020)
m2'	Re-implementing M2 in ML3 (Reinhardt et al., 2019)
m3	Bringing M2 and M2' into alignment
m4	Extending the routes and Rumours model by including risk, leading to the risk and Rumours model (Chap. 8, Sect. 8.2)
m4'	Integrating the proposed policy intervention into M4 (Box 9.3)
m5	Adding geography of and data about the Mediterranean crossing in the risk and Rumours model, to become risk and Rumours with reality (Chap. 8, Sect. 8.4)
m5'	Integrating the proposed intervention into M5 (Box 9.4)
rf	Deriving the risk function, RF (Box 6.1)
sc1	Calibrating a model-based early warning system (Box 9.1)
sc2	Simulating the scenarios (Box 9.2)
sc3	Simulating the policy intervention (Box 9.3)
sc3'	Simulating the policy intervention with a calibrated model (Box 9.4)

Glossary

Listed below are non-technical, general-level, intuitive explanations of some of the key terms appearing throughout the book. While they are no substitute for more formal definitions, which can be found elsewhere in this book (and in the wider literature), and which can vary between scientific disciplines, we hope that they will help our interdisciplinary readership share our understanding of the key concepts.

Abduction An approach of making inferences to the 'best explanation,' in an attempt to formulate plausible explanations between the observed phenomena and to unravel the mechanisms that might have contributed to observed outcomes. In the context of agent-based modelling, some elements of model construction can be seen as abductive (Chap. 2).

Agency An all-encompassing term with many possible interpretations, but in the context of this book understood as the ability of **agents**, representing people, institutions, or other decision-making units, to react to all aspects of a situation – including their own internal state and the state of their environment – in surprising and essentially unpredictable ways (Chaps. 2 and 3).

Agent-based model computer simulation, with a population of simulated agents following individual-level rules of behaviour and interacting with one another and with their environment, leading to the emergence of observable properties at the macroscopic level (Chaps. 2 and 3).

Asylum migration The movement of an individual or individuals from their country of origin, for the purpose of seeking international protection from persecution, as set out in the 1951 UN Convention Relating to the Status of Refugees and the 1967 Protocol (Chap. 4).

Attitude An evaluation that an individual makes regarding an object such as a viewpoint, topic, idea, or person. Attitudes are usually developed through experience with, or related to, the object. Attitudes can vary in strength (be weak or strong) and can be positive, negative, or ambivalent (Chap. 6).

Bayesian methods Methods of statistical inference based on the work of Thomas Bayes and on his famous 1763 theorem, whereby the *prior knowledge* about unknown events, features of the world, model parameters, or models, gets updated in the light of new data (evidence) to produce *posterior knowledge*. Bayesian methods rely on the subjective definition of probability and, by treating all unknown quantities as random, offer a coherent description of uncertainty (Foreword; Chaps. 1, 2, and 11).

Calibration A process of aligning model outputs with the empirical observations (data) through changing the relevant model parameters (inputs). In the context of statistical, typically Bayesian methods of uncertainty quantification, the process may involve full statistical inference about the probability distributions of the parameters (Chap. 5).

Causality Informally, a situation where phenomenon A precedes phenomenon B in time, and the occurrence of phenomenon A makes the occurrence of phenomenon B more likely in different contexts, assuming that A and B do not share a common cause themselves (Chaps. 2 and 3).

Cognition Thoughts and other mental processes that occur within a person's brain. Within psychology, often used to distinguish from behaviour that focuses on people's external actions in the world. Some common areas of cognition include memory, learning, language, and *metacognition* – thinking about thinking (Chap. 6).

Complexity Another all-encompassing term with many possible interpretations, here interpreted as a feature of a given system indicating how difficult it is to understand it (Chaps. 2 and 3).

Data Empirical information collected through observations, reports, or responses in experiments or in a real-world context. Sources may collect and publish data for administrative or operational purposes, to further our understanding through research, or in journalistic pursuits (Chap. 4).

Decision Reaching a conclusion or resolution, and selecting a specific option or alternative from those available, following a thought process. For example, looking outside the window before leaving home and then deciding to not take an umbrella because the weather looks good, or choosing between several potential holiday destinations and deciding to travel to the Greek Islands (Chap. 6).

Domain-Specific Language After van Deursen et al. (2000), programming languages that are "focused on, and usually restricted to, a particular problem domain" to solve the specific problems in that domain more easily, rather than being designed as general-purpose tools (Chap. 7).

Emulator (meta-model) A statistical model of an underlying complex, computational model, designed to approximate the model dynamics and illuminate the often opaque relationships between model inputs and outputs. The specification of emulators may vary, from simple regression models to the commonly used Gaussian processes (Chap. 5).

Experiment (psychology) Research design, in which the researcher has full control over the independent variable of interest and therefore can randomly assign participants to different levels of the independent variable, allowing for causal claims to be made about the impact of the independent variable on measured outcomes – dependent variables (Chap. 6).

Experiment (simulation) Following Cellier (1991), "an experiment is the process of extracting data from a system by exerting it through its inputs," and "a simulation is an experiment performed on a model." Throughout this book, we refer to the process of experimenting on a model as a simulation experiment (Chap. 7).

Experimental design A range of statistical methods, at the first step in planning an experiment, aimed at setting up the experiments (natural, computational, or other), and running them in such a way, for specific values of inputs, to maximise the resulting information gains (Chap. 5).

Induction In the classical sense, dating back to Francis Bacon (1620), the backbone of the scientific method, relying on *inducing* the various formal principles guiding the phenomena under study, without which these phenomena would not come about in the same form as they do (Courgeau et al., 2016). An alternative, modern meaning, associated with John Stuart Mill, is that of a method of scientific reasoning through making inferences based on generalised observations (Chap. 2).

Information In the context of models discussed in this book, knowledge of any part of the migration process (such as job prospects in destination countries, or how to access resources at a stop-off point) that may influence an individual's decisions. Information may be transferred between individuals or received from other external sources (Chaps. 3, 4, 8 and 9).

Language A set of words, usually a subset of all words constructed from the symbols of an alphabet (Hopcroft & Ullman, 1979). In a typical programming language, the words are sequences of (Unicode) characters, representing the alphabet. Character sequences that form legal programs are words of the language. However, this definition does not restrict the words to be character sequences (Chap. 7).

Migration The movement of an individual or individuals from their place of origin or residence. This movement can take place within a country/region (internal) or involve crossing international borders (international), and can be defined by a specified duration of stay at the destination (Chaps. 2 and 4).

Model In the widest sense, a well-described – either formally or in the form of a physical instance – entity that can be used to infer or demonstrate the consequences of a set of conditions, where these conditions are assumed to capture a relevant part of a phenomenon of interest.

Network Generally, a structure consisting of entities and links between them. In the context of agent-based modelling, often specifically a *social* network of individuals (agents) and contacts between them.

Probability Formal measure of uncertainty, bounded between zero and one, which can have either objective or epistemic interpretation (Courgeau, 2012). In the former case, linked with classical (frequentist) statistics, probability is usually related to the frequency of events, and in the latter case, typically associated with Bayesian inference, can be a subjective measure of belief, or a logical relationship (Foreword; Chaps. 1, 2, 5 and 6).

Provenance After Groth and Moreau (2013), "information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness" (Chap. 7).

Quality (of data) An expert assessment based on a range of criteria relating to aspects of the data collection, content, reporting and relevance to the purpose for which they are to be used (Chap. 4).

Replicability The practice of repeating an experiment (or study) to collect new data from a new set of participants. Replications can be conducted by the same researchers as those who conducted the original study, but confidence in the replicability of a study is usually greater if the replication is conducted by independent researchers (Chaps. 6 and 10).

Reproducibility The ability of researchers to recreate an aspect of a study (e.g., a statistical analysis or a computational model) based on the materials provided by the original authors within a publication, as well as any supplementary datasets, analysis code, or other materials that can be accessed (Chaps. 7 and 10).

Risk Circumstances in which the outcomes are not known, but may be represented in terms of a probability of two or more possible outcomes occurring. For example, when tossing a fair coin there is an approximately 50% chance each of it landing on heads or tails. Therefore, betting on heads or tails is a risky decision. Risk can be contrasted with **uncertainty**, where the probabilities of potential outcomes are unknown (or unknowable). The term *risk* is also often used to refer to uncertain events that may have negative outcomes (Chaps. 2, 6 and 8).

Semantics A function that maps the words of a language to some other set, e.g., a class of abstract machines or a class of stochastic processes. The element of the other set to which a word is mapped is interpreted as the "meaning" of the word (Chap. 7).

Sensitivity The extent to which the model results (outputs) change when the individual parameters or inputs – or their combinations – change. The sensitivity analysis can be local, around some specific parameter values, or global, across the whole parameter space (Chaps. 5 and 8).

Simulator According to Zeigler et al. (2019) "any computation system (such as a single processor, a processor network, the human mind, or more abstractly an algorithm), capable of executing a model to generate its behavior" (Chap. 7).

Syntax The set of rules that defines which of the words constructed from an alphabet are elements of the language. The syntax therefore defines the subset of words that make up the language (Chap. 7).

Topology Informally, the spatial structure of something (an object, fragment of the physical or simulated world, and so on), looking solely at connections between and relative positions of its constituting elements and ignoring their sizes and exact distances (Chaps. 3 and 8).

Uncertainty The state of imperfect knowledge about the world (epistemic uncertainty), as well as its intrinsic randomness (aleatory uncertainty), leading to unpredictability. Some forms of uncertainty are measurable (quantifiable) as **risk** by using statistical models relying on probability theory and, typically, Bayesian methods of inference. The *uncertainty analysis* measures how much uncertainty in model outputs is induced by the inputs (Chaps. 2, 5 and 8).

Utility A way of representing the value of something in terms of its usefulness or importance, rather than simply focusing on explicit value. For example, money may have different levels of utility depending on who is receiving it and when: $1000 has more utility if received now to pay bills and buy food, and relatively less utility if received in three months' time, when there are no extra bills to be paid, even though the actual monetary amount has not changed (Chap. 6).

References

Abdellaoui, M., & Kemel, E. (2014). Eliciting prospect theory when consequences are measured in time units: "Time is not money". *Management Science, 60*, 1844–1859.

Abdellaoui, M., Bleichrodt, H., L'Haridon, O., & van Dolder, D. (2016). Measuring loss aversion under ambiguity: A method to make prospect theory completely observable. *Journal of Risk and Uncertainty, 52*, 1–20.

Ahmed, M. N., Barlacchi, G., Braghin, S., Calabrese, F., Ferretti, M., Lonij, V., Nair, R., Novack, R., Paraszczak, J., & Toor, A. S. (2016). A multi-scale approach to data-driven mass migration analysis. In *The Fifth Workshop on Data Science for Social Good*. SoGood@ECML-PKDD.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action control. From cognition to behavior* (pp. 11–39). Springer.

Akgüç, M., Liu, X., Tani, M., & Zimmermann, K. F. (2016). Risk attitudes and migration. *China Economic Review, 37*, 166–176.

Alam, S. J., & Geller, A. (2012). Networks in agent-based social simulation. In A. Heppenstall, A. Crooks, L. See, & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 199–216). Springer.

Andrianakis, I., Vernon, I. R., McCreesh, N., McKinley, T. J., Oakley, J. E., Nsubuga, R. N., Goldstein, M., & White, R. G. (2015). Bayesian history matching of complex infectious disease models using emulation: A tutorial and a case study on HIV in Uganda. *PLoS Computational Biology, 11*(1), e1003968.

Angione, C., Silverman, E., & Yaneske, E. (2020). Using machine learning to emulate agent-based simulations. Mimeo. *arXiv*. https://arxiv.org/abs/2005.02077. (as of 1 August 2020)

Apicella, C., Norenzayan, A., & Henrich, J. (2020). Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior, 41*(5), 319–329.

Arango, J. (2000). Explaining migration: A critical view. *International Social Science Journal, 52*, 283–296.

Arellana, J., Garzón, L., Estrada, J., & Cantillo, V. (2020). On the use of virtual immersive reality for discrete choice experiments to modelling pedestrian behaviour. *Journal of Choice Modelling, 37*, 100251.

Ariely, D. (2008). *Predictably irrational. The hidden forces that shape our decisions*. Harper Collins.

Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist, 63*(7), 602–614.

Attema, A. E., Brouwer, W. B., & L'Haridon, O. (2013). Prospect theory in the health domain: A quantitative assessment. *Journal of Health Economics, 32*, 1057–1065.

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7

Attema, A. E., Brouwer, W. B., L'Haridon, O., & Pinto, J. L. (2016). An elicitation of utility for quality of life under prospect theory. *Journal of Health Economics, 48*, 121–134.

Axtell, R., Epstein, J., Dean, J., et al. (2002). Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. *Proceedings of the National Academy of Sciences of the United States of America, 99*(Suppl. 3), 7275–7279.

Azose, J. J., & Raftery, A. E. (2015). Bayesian probabilistic projection of international migration. *Demography, 52*(5), 1627–1650.

Bacon, F. (1620). *Novum organum*. J. Bill. English translation by J Spedding, RL Ellis, and DD Heath (1863) in *The Works* (Vol. VIII). Taggard and Thompson.

Bakewell, O. (1999). Can we ever rely on refugee statistics? *Radical Statistics, 72*, art. 1. Accessible via: www.radstats.org.uk/no072/article1.htm (as of 1 February 2019)

Baláž, V., & Williams, A. M. (2018). Migration decisions in the face of upheaval: An experimental approach. *Population, Space and Place, 24*, e2115.

Baláž, V., Williams, A. M., & Fifekova, E. (2016). Migration decision making as complex choice: Eliciting decision weights under conditions of imperfect and complex information through experimental methods. *Population, Space and Place, 22*, 36–53.

Banks, D., & Norton, J. (2014). Agent-based modeling and associated statistical aspects. In *International Encyclopaedia of the Social and Behavioural Sciences* (2nd ed., pp. 78–86). Oxford University Press.

Banks, D. L., Rios Aliaga, J. M., & Rios Insua, D. (2015). *Adversarial risk analysis*. CRC Press.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives, 27*, 173–196.

Barbosa Filho, H. S., Lima Neto, F. B., & Fusco, W. (2013). Migration, communication and social networks – An agent-based social simulation. In R. Menezes, A. Evsukoff, & M. C. González (Eds.), *Complex networks. Studies in computational intelligence* (Vol. 424, pp. 67–74). Springer.

Barker ER and Bijak J (2020) . *Conceptualisation and analysis of migration uncertainty: Insights from macroeconomics* (QuantMig project deliverable D9.1). University of Southampton. Via https://www.quantmig.eu

Barth, R., Meyer, M., & Spitzner, J. (2012). Typical pitfalls of simulation modeling – Lessons learned from armed forces and business. *Journal of Artificial Societies and Social Simulation, 15*(2), 5.

Bauermeister, G.-F., Hermann, D., & Musshoff, O. (2018). Consistency of determined risk attitudes and probability weightings across different elicitation methods. *Theory and Decision, 84*(4), 627–644.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London, 53*, 370–418.

BBC News. (2015). *Syrian journey: Choose your own escape route*. Accessible via: https://www.bbc.co.uk/news/world-middle-east-32057601 (as of 1 February 2021).

Beaumont, M. A., Cornuet, J.-M., Marin, J.-M., & Robert, C. P. (2009). Adaptive approximate Bayesian computation. *Biometrika, 96*(4), 983–990.

Begley, C. G., & Ioannidis, J. P. A. (2015). Reproducibility in science. *Circulation Research, 116*(1), 116–126.

Bélanger, A., & Sabourin, P. (2017). *Microsimulation and population dynamics. An introduction to Modgen 12*. Springer.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology, 100*(3), 407–425.

Ben-Akiva, M., de Palma, A., McFadden, D., Abou-Zeid, M., Chiappori, P.-A., de Lapparent, M., Durlauf, S. N., FosgerauM, F. D., Hess, S., Manski, C., Pakes, A., Picard, N., & Walker, J. (2012). Process and context in choice models. *Marketing Letters, 23*, 439–456.

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C. D., Clyde, M., Cook, T. D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., … Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour, 2*(1), 6–10.

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2014). Julia: A fresh approach to numerical computing. *SIAM Review, 59*(1), 65–98.

Bijak, J. (2010). *Forecasting international migration in Europe: A Bayesian view*. Springer.

Bijak, J., & Bryant, J. (2016). Bayesian demography 250 years after Bayes. *Population Studies, 70*(1), 1–19.

Bijak, J., & Czaika, M. (2020). *Assessing uncertain migration futures – A typology of the unknown* (QuantMig project deliverable D1.1). University of Southampton and Danube University Krems. Via https://www.quantmig.eu

Bijak, J., & Koryś, I. (2009). Poland. In H. Fassman, U. Reeger, & W. Sievers (Eds.), *Statistics and reality: Concepts and measurements of migration in Europe* (pp. 195–216). AUP.

Bijak, J., & Lubman, S. (2016). The disputed numbers. Search of the demographic basis for studies of Armenian population losses, 1915–1923. In A. Demirdjian (Ed.) *The Armenian Genocide Legacy* (pp. 26–43). Palgrave.

Bijak, J., & Wiśniowski, A. (2010). Bayesian forecasting of immigration to selected European countries by using expert knowledge. *Journal of the Royal Statistical Society, Series A 173*(4): 775–796.

Bijak, J., Kupiszewska, D., Kupiszewski, M., Saczuk, K., & Kicinger, A. (2007). Population and labour force projections for 27 European countries, 2002-2052: Impact of international migration on population ageing. *European Journal of Population, 23*(1), 1–31.

Bijak, J., Hilton, J., Silverman, E., & Cao, V. D. (2013). Reforging the wedding ring: Exploring a semi-artificial model of population for the UK with Gaussian process emulators. *Demographic Research, 29*, 729–766.

Bijak, J., Forster, J. J., & Hilton, J. (2017). *Quantitative assessment of asylum-related migration: A survey of methodology* (Report for the European Asylum Support Office). EASO.

Bijak, J., Disney, G., Findlay, A. M., Forster, J. J., Smith, P. W. F., & Wiśniowski, A. (2019). Assessing time series models for forecasting international migration: Lessons from the United Kingdom. *Journal of Forecasting, 38*(5), 470–487.

Bijak, J., Higham, P., Hilton, J. D., Hinsch, M., Nurse, S., Prike, T., Reinhardt, O., Smith, P. W., & Uhrmacher, A. M. (2020). Modelling migration: Decisions, processes and outcomes. In *Proceedings of the Winter Simulation Conference 2020* (pp. 2613–2624). IEEE.

Billari, D. C. (2015). Integrating macro- and micro-level approaches in the explanation of population change. *Population Studies, 65*(S1), S11–S20.

Billari, F., & Prskawetz, A. (Eds.). (2003). *Agent-based computational demography: Using simulation to improve our understanding of demographic behaviour*. Plenum.

Billari, F. C., Fent, T., Prskawetz, A., & Scheffran, J. (Eds.). (2006). *Agent-based computational modelling. Applications in demography, social, economic and environmental sciences*. Physica-Verlag.

Billari, F., Aparicio Diaz, B., Fent, T., & Prskawetz, A. (2007). The "Wedding–Ring". An agent-based marriage model based on social interaction. *Demographic Research, 17*(3), 59–82.

Bishop, Y. M., Fienberg, S. E., & Holland, P. W. (1975/2007). *Discrete multivariate analysis: Theory and practice* (Reprint ed.). Springer.

Bocquého, G., Deschamps, M., Helstroffer, J., Jacob, J., & Joxhe, M. (2018). *Risk and refugee migration*. Sciences Po OFCE Working Paper hal-02198118. Paris, France.

Bohra-Mishra, P., & Massey, D. S. (2011). Individual decisions to migrate during civil conflict. *Demography, 48*(2), 401–424.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America, 99*(Suppl 3), 7280–7287.

Borjas, G. J. (1989). Economic theory and international migration. *International Migration Review, 23*(3), 457–485.

Bortolussi, L., De Nicola, R., Galpin, V., Gilmore, S., Hillston, J., Latella, D., Loreti, M., & Massink, M. (2015). CARMA: Collective adaptive resource-sharing Markovian agents. *Electronic Proceedings in Theoretical Computer Science, 194*, 16–31.

Boulesteix, A.-L., Groenwold, R.H.H., Abrahamowicz, M., Binder, H., Briel, M., Hornung, R., Morris, T.P., Rahnenführer, J., and Sauerbrei, W. for the STRATOS Simulation Panel. (2020). Introduction to statistical simulations in health research. *BMJ Open 10*: e039921.

Boukouvalas, A., & Cornford, D. (2008). *Dimension reduction for multivariate emulation* (Technical report NCRG/2008/006. Neural Computing Research Group). Aston University.

Bourgais, M., Taillandier, P., & Vercouter, L. (2020). BEN: An architecture for the behavior of social agents. *Journal of Artificial Societies and Social Simulation, 23*(4), 12.

Bourgeois-Pichat, J. (1994). *La dynamique des populations. Populations stables, semi stables, quasi stables*. Institut national d'études démographiques, Presses Universitaires de France.

Brenner, T., & Werker, C. (2009). Policy advice derived from simulation models. *Journal of Artificial Societies and Social Simulation, 12*(4), 2.

Briñol, P., & Petty, R. E. (2009). Source factors in persuasion: A self-validation approach. *European Review of Social Psychology, 20*(1), 49–96.

Bryant, J., & Zhang, J. (2018). *Bayesian demographic estimation and forecasting*. CRC Press.

Bryson, J. J., Ando, Y., & Lehmann, H. (2007). Agent-based modelling as scientific method: A case study analysing primate social behaviour. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences, 362*(1485), 1685–1698.

Budde, K., Smith, J., Wilsdorf, P., Haack, F., & Uhrmacher, A. M. (2021). Relating simulation studies by provenance – Developing a family of Wnt signaling models. *PLoS Computational Biology, 17*(8), e1009227.

Budescu, D. V., Por, H.-H., Broomell, S. B., & Smithson, M. (2014). The interpretation of IPCC probabilistic statements around the world. *Nature Climate Change, 4*, 508–512.

Burch, T. (2003). Demography in a new key: A theory of population theory. *Demographic Research, 9*, 263–284.

Burch, T. (2018). *Model-based demography. Essays on integrating data, technique and theory* (Demographic Research Monographs, Vol. 14). Springer.

Burks, A. W. (1946). Peirce's theory of abduction. *Philosophy of Science, 13*(4), 301–306.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science, 351*(6280), 1433.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., … Wu, H. (2018). Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. *Nature Human Behaviour, 2*(9), 637–644.

Carling, J., & Collins, F. (2018). Aspiration, desire and drivers of migration. *Journal of Ethnic and Migration Studies, 44*(6), 909–926.

Carling, J., & Schewel, K. (2018). Revisiting aspiration and ability in international migration. *Journal of Ethnic and Migration Studies, 44*(6), 945–963.

Casini, L., Illari, P., Russo, F., & Williamson, J. (2011). Models for prediction, explanation and control: Recursive Bayesian networks. *Theoria, 26*(1), 5–33.

Castles, S. (2004). Why migration policies fail. *Ethnic and Racial Studies, 27*(2), 205–227.

Castles, S., de Haas, H., & Miller, M. J. (2014). *The age of migration: International population movements in the modern world* (5th ed.). Palgrave.

Cellier, F. E. (1991). *Continuous system modeling*. Springer.

Ceriani, L., & Verme, P. (2018). *Risk preferences and the decision to flee conflict* (Policy research working paper no. 8376). World Bank.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgement. *Journal of Personality and Social Psychology, 66*, 460–473.

Chaloner, K., & Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science, 10*(3), 273–304.

Chambers, C. D. (2013). Registered reports: A new publishing initiative at cortex. *Cortex, 49*(3), 609–610.

Chambers, C. (2019). The registered reports revolution: Lessons in cultural reform. *Significance, 16*(4), 23–27.

Channel 4 News. (2015). *Two billion miles*. Accessible via: http://twobillionmiles.com/ (as of 1 February 2021).

Christensen, K., & Sasaki, Y. (2008). Agent-based emergency evacuation simulation with individuals with disabilities in the population. *Journal of Artificial Societies and Social Simulation, 11*(3), 9.

Christensen, G., Dafoe, A., Miguel, E., Moore, D. A., & Rose, A. K. (2019a). A study of the impact of data sharing on article citations using journal policies as a natural experiment. *PLoS One, 14*(12), e0225883.

Christensen, G., Wang, Z., Paluck, E. L., Swanson, N., Birke, D. J., Miguel, E., & Littman, R. (2019b). Open Science practices are on the rise: The state of social science (3S) survey. *MetaArXiv*. Preprint.

Cimellaro, G. P., Mahin, S., & Domaneschi, M. (2019). Integrating a human behavior model within an agent-based approach for blasting evacuation. *Computer-Aided Civil and Infrastructure Engineering, 34*, 3–20.

Clark, R. D., III, & Maass, A. (1988). The role of social categorization and perceived source credibility in minority influence. *European Journal of Social Psychology, 18*, 381–394.

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *The Journal of Abnormal and Social Psychology, 65*(3), 145–153.

Cohen, J. E., Roig, M., Reuman, D. C., & GoGwilt, C. (2008). International migration beyond gravity: A statistical model for use in population projections. *Proceedings of the National Academy of Sciences of the United States of America, 105*(40), 15268–15274.

Coleman, J. S. (1986). Social theory, social research, and a theory of action. *American Journal of Sociology, 91*(6), 1309–1335.

Collier, N., & Ozik, J. (2013). *Repast Simphony Batch runs getting started*. Available from https://repast.sourceforge.net/docs/RepastBatchRunsGettingStarted.pdf (as of 1 January 2021).

Collins, F. L. (2018). Desire as a theory for migration studies: Temporality, assemblage and becoming in the narratives of migrants. *Journal of Ethnic and Migration Studies, 44*(6), 964–980.

Collins, A. J., & Frydenlund, E. (2016). Agent-based modeling and strategic group formation: A refugee case study. In *Proceedings of the Winter Simulation Conference 2016* (pp. 1289–1300). IEEE.

Collins, A. J., Etemadidavan, S., & Pazos-Lago, P. (2020). A human experiment using a hybrid agent-based model. In *Proceedings of the Winter Simulation Conference 2020*. IEEE.

Commons, M. L., Nevin, J. A., & Davison, M. C. (Eds.). (2013). *Signal detection: Mechanisms, models, and applications*. Psychology Press.

Conte, R., & Paolucci, M. (2014). On agent-based modeling and computational social science. *Frontiers in Psychology, 5*(668), 1–9.

Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., Loreto, V., Moat, S., Nadal, J.-P., Sanchez, A., Nowak, A., Flache, A., San Miguel, M., & Helbing, D. (2012). Manifesto of computational social science. *European Physical Journal Special Topics, 214*, 325–346.

Courgeau, D. (1985). Interaction between spatial mobility, family and career life cycle: A French survey. *European Sociological Review, 1*(2), 139–162.

Courgeau, D. (2007). *Multilevel synthesis. From the group to the individual*. Springer.

Courgeau, D. (2012). *Probability and social science: Methodological relationships between the two approaches*. Springer.

Courgeau, D., Bijak, J., Franck, R., & Silverman, E. (2016). Model-based demography: Towards a research agenda. In J. Van Bavel & A. Grow (Eds.), *Agent-based modelling in population studies: Concepts, methods, and applications* (pp. 29–51). Springer.

Cox, D. R. (1958/1992). *Planning of experiments*. Wiley.

Cressie, N. (1990). The origins of kriging. *Mathematical Geology, 22*, 239–252.

Crisp, J. (1999). *Who has counted the refugees? UNHCR and the politics of numbers* (New issues in refugee research, No. 12). UNHCR.

Cusumano, E., & Pattison, J. (2018). The non-governmental provision of search and rescue in the Mediterranean and the abdication of state responsibility. *Cambridge Review of International Affairs, 31*(1), 53–75.

Cusumano, E., & Villa, M. (2019). *Sea rescue NGOs: A pull factor of irregular migration?* (Migration policy Centre policy brief 22/2019). European University Institute.

Czaika, M. (2014). Migration and economic prospects. *Journal of Ethnic and Migration Studies, 41*, 58–82.

Czaika, M., & Reinprecht, C. (2020). *Drivers of migration: A synthesis of knowledge* (IMI working paper no. 163). University of Amsterdam.

Czaika, M., Bijak, J., & Prike, T. (2021). Migration decision-making and its four key dimensions. *The Annals of the American Academy of Political and Social Science*, forthcoming.

David, N. (2009). Validation and verification in social simulation: Patterns and clarification of terminology. In F. Squazzoni (Ed.), *Epistemological aspects of computer simulation in the social sciences* (Lecture Notes in Artificial Intelligence, 5466) (pp. 117–119). Springer.

Davies, O. L., & Hay, W. A. (1950). The construction and uses of fractional factorial designs in industrial research. *Biometrics, 6*(3), 233–249.

de Castro, P. A. L., Barreto Teodoro, A. R., de Castro, L. I., & Parsons, S. (2016). Expected utility or prospect theory: Which better fits agent-based modeling of markets? *Journal of Computational Science, 17*, 97–102.

De Finetti, B. (1974). *Theory of probability* (Vol. 2). Wiley.

de Haas, H. (2010). Migration and development: A theoretical perspective. *International Migration Review, 44*(1), 227–264.

De Jong, G. F., & Fawcett, J. T. (1981). Motivations for migration: An assessment and a value-expectancy research model. In G. F. De Jong & R. W. Gardener (Eds.), *Migration decision making: Multidisciplinary approaches to microlevel studies in developed and developing countries* (pp. 13–57). Pergamon.

de Laplace, P.-S. (1780). Mémoire sur les probabilités. *Mémoires de l'Académie Royale des Sciences de Paris, 1781*, 227–332.

De Nicola, R., Latella, D., Loreti, M., & Massink, M. (2013). A uniform definition of stochastic process calculi. *ACM Computing Surveys, 46*(1), 1–35.

DeGroot, M. H. (2004). *Optimal statistical decisions. Wiley classics* (Library ed.). Chichester: Wiley.

Dekker, R., Engbersen, G., Klaver, J., & Vonk, H. (2018). Smart refugees: How Syrian asylum migrants use social media information in migration decision making. *Social Media & Society, 4*(1), 2056305118764439.

Devroye, L. (1986). *Non-uniform random variate generation*. Springer.

Di Paolo, E. A., Noble, J., & Bullock, S. (2000). Simulation models as opaque thought experiments. In *ALife 7 conference proceedings* (pp. 497–506). MIT Press.

Diaz, B. A., Fent, T., Prskawetz, A., & Bernardi, L. (2011). Transition to parenthood: The role of social interaction and endogenous networks. *Demography, 48*(2), 559–579.

Disney, G., Wiśniowski, A., Forster, J. J., Smith, P. W. F., & Bijak, J. (2015). *Evaluation of existing migration forecasting methods and models* (Report for the Migration Advisory Committee). Centre for Population Change.

D'Orazio, M., Di Zio, M., & Scanu, M. (2006). *Statistical matching: Theory and practice*. Wiley.

Douven, I. (2017). Abduction. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of philosophy*. (Summer 2017 ed.). Available via https://plato.stanford.edu/archives/sum2017/entries/abduction (as of 1.10.2018)

Drogoul, A., Vanbergue, D., & Meurisse, T. (2003). Multi-agent based simulation: Where are the agents? In J. S. Sichman, F. Bousquet, & P. Davidsson (Eds.), *Multi-agent-based simulation II. Lecture Notes in Computer Science*, (Vol. 2581, pp. 1–15). Springer.

Dunsch, F., Tjaden, J., & Quiviger, W. (2019). *Migrants as messengers: The impact of peer-to-peer communication on potential migrants in Senegal. Impact evaluation report*. International Organization for Migration.

Dustmann, C., Fasani, F., Meng, X., & Minale, L. (2017). *Risk attitudes and household migration decisions* (IZA discussion papers no. 10603). Institute for the Study of Labor (IZA).

EASO. (2016). *The push and pull factors of asylum-related migration. A literature review* (Report by Maastricht University and the global migration data analysis Centre (GMDAC) for the European Asylum Support Office). EASO.

Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., Baranski, E., Bernstein, M. J., Bonfiglio, D. B. V., Boucher, L., Brown, E. R., Budiman, N. I., Cairo, A. H., Capaldi, C. A., Chartier, C. R., Chung, J. M., Cicero, D. C., Coleman, J. A., Conway, J. G., … Nosek, B. A. (2016). Many labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology, 67*(Special Issue: Confirmatory), 68–82.

Edmonds, B., Le Page, C., Bithell, M., Grimm, V., Meyer, R., Montañola-Sales, C., Ormerod, P., Root, H., & Squazzoni, F. (2019). Different modelling purposes. *Journal of Artificial Societies and Social Simulation, 22*(3), 6.

Elson, M., & Quandt, T. (2016). Digital games in laboratory experiments: Controlling a complex stimulus through modding. *Psychology of Popular Media Culture, 5*(1), 52–65.

Emmer, M., Richter, C., & Kunst, M. (2016). *Flucht 2.0: Mediennutzung durch Flüchtlinge vor, während und nach der Flucht*. Institut für Publizistik, FU Berlin.

Entwisle, B., Williams, N. E., Verdery, A. M., Rindfuss, R. R., Walsh, S. J., Malanson, G. P., Mucha, P. J., et al. (2016). Climate shocks and migration: An agent-based modeling approach. *Population and Environment, 38*(1), 47–71.

Epstein, J. M. (2008). Why model? *Journal of Artificial Societies and Social Simulation, 11*(4), 12.

Epstein, J. M., & Axtell, R. (1996). *Complex adaptive systems. Growing artificial societies: Social science from the bottom up*. MIT Press.

Erdal, M. B., & Oeppen, C. (2018). Forced to leave? The discursive and analytical significance of describing migration as forced and voluntary. *Journal of Ethnic and Migration Studies, 44*(6), 981–998.

Euler, L. (1760). Recherches générales sur la mortalité et la multiplication du genre humain. *Histoire de l'Académie Royale des Sciences et des Belles Lettres de Berlin, 16*, 144–164.

European Commission. (2015). *Communication from the commission to the European Parliament, the council, the European economic and social committee and the Committee of the Regions* Commission work programme 2016. COM(2015)610 final. European Commission.

European Commission. (2016). *Fact sheet: Reforming the common European asylum system: Frequently asked questions*. European Commission, 13 July 2016. Accessible via: http://europa.eu/rapid/press-release_MEMO-16-2436_en.htm (as of 26 September 2019)

European Commission. (2020). *Communication from the commission to the European Parliament, the council, the European economic and social committee and the Committee of the Regions on a new pact on migration and asylum.* COM(2020)609 final. European Commission.

Ewald, R., & Uhrmacher, A. M. (2014). SESSL: A domain-specific language for simulation experiments. *ACM Transactions on Modeling and Computer Simulation (TOMACS), 24*(2), 1–25.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global Evidence on Economic Preferences. *The Quarterly Journal of Economics, 133*(4), 1645–1692.

Fang, K.-T., Li, R., & Sudjianto, A. (2006). *Design and modeling for computer experiments*. CRC.

Farooq, B., Cherchi, E., & Sobhani, A. (2018). Virtual immersive reality for stated preference travel behavior experiments: A case study of autonomous vehicles on Urban roads. *Transportation Research Record, 2672*(50), 35–45.

Feldman, R. H. L. (1984). The influence of communicator characteristics on the nutrition attitudes and behavior of high school students. *Journal of School Health, 54*, 149–151.

Felleisen, M. (1991). On the expressive power of programming languages. *Science of Computer Programming, 17*(1), 35–75.

Fisher, R. A. (1926). The arrangement of field experiments. *Journal of the Ministry of Agriculture of Great Britain, 33*, 503–513.

Fisher, R. A. (1935/1971). *The design of experiments*. Macmillan.

FitzGerald, D. S. (2015). The sociology of international migration. In C. B. Brettell & J. F. Hollifield (Eds.), *Migration theory, talking across disciplines* (3rd ed., pp. 115–147). Routledge.

Flake, J. K., & Fried, E. I. (2020). Measurement Schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science, 3*(4), 456–465.

Foresight. (2011). *Migration and global environmental change: Future challenges and opportunities. Final project report*. Government Office for Science.

Fowler, M. (2010). *(with R Parsons) Domain-specific languages*. Addison-Wesley.

Franck, R. (Ed.). (2002). *The explanatory power of models*. Kluwer Academic Publishers.

Frank, U., Squazzoni, F., & Troitzsch, K. G. (2009). EPOS-epistemological perspectives on simulation: An introduction. In F. Squazzoni (Ed.), *Epistemological aspects of computer simulation in the social sciences* (Lecture Notes in Artificial Intelligence, 5466) (pp. 1–11). Springer.

Fraser, H., Parker, T., Nakagawa, S., Barnett, A., & Fidler, F. (2018). Questionable research practices in ecology and evolution. *PLoS One, 13*(7), e0200303.

Fricker, T. E., Oakley, J. E., & Urban, N. M. (2013). Multivariate Gaussian process emulators with nonseparable covariance structures. *Technometrics, 55*(1), 47–56.

Frigg, R., Bradley, S., Du, H., & Smith, L. A. (2014). Laplace's demon and the adventures of his apprentices. *Philosophy of Science, 81*(1), 31–59.

Frontex. (2018). *Risk analysis for 2018*. Frontex.

Frydenlund, E., Foytik, P., Padilla, J. J., & Ouattara, A. (2018). Where are they headed next? Modeling emergent displaced camps in the DRC using agent-based models. In *Proceedings of the Winter Simulation Conference 2018*. IEEE.

Frydman, R., & Goldberg, M. D. (2007). *Imperfect knowledge economics*. Princeton University Press.

Fujimoto, R. M. (2000). *Parallel and distributed simulation systems* (Wiley series on parallel and distributed computing). Wiley.

Gabrielsen Jumbert, M. (2020). *The "pull factor": How it became a central premise in European discussions about cross-Mediterranean migration*. Available at: www.law.ox.ac.uk/research-subject-groups/centre-criminology/centreborder-criminologies/blog/2020/03/pull-factor-how (as of 1 February 2021).

GAO. (2006). *Darfur crisis: Death estimates demonstrate severity of crisis, but their accuracy and credibility could be enhanced* (Report to congressional requesters GAO-07-24). US Government Accountability Office.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2013). *Bayesian Data Analysis* (3rd ed.). CRC Press/Chapman and Hall.

Ghanem, R., Higdon, D., & Owhadi, H. (2019). *Handbook of uncertainty quantification.* Living reference work. Online resource, available at https://link.springer.com/referencework/10.1007/978-3-319-11259-6 (as of 1 November 2019).

Gibson, J., & McKenzie, D. (2011). The microeconomic determinants of emigration and return migration of the best and brightest: Evidence from the Pacific. *Journal of Development Economics, 95*, 18–29.

Gigerenzer, G. (2008). *Rationality for mortals: How people cope with uncertainty*. OUP.

Gigerenzer, G., & Marewski, J. N. (2015). Surrogate science: The idol of a universal method for scientific inference. *Journal of Management, 41*(2), 421–440.

Gilbert, N., & Ahrweiler, P. (2009). The epistemologies of social simulation research. In F. Squazzoni (Ed.), *Epistemological aspects of computer simulation in the social sciences* (Lecture Notes in Artificial Intelligence, 5466) (pp. 12–28). Springer.

Gilbert, N., & Tierna, P. (2000). How to build and use agent-based models in social science. *Mind and Society 1*(1): 57–72.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry, 81*(25), 2340–2361.

Gillespie, D. T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics, 115*(4), 1716–1733.

Ginot, V., Gaba, S., Beaudouin, R., Aries, F., & Monod, H. (2006). Combined use of local and ANOVA-based global sensitivity analyses for the investigation of a stochastic dynamic model: Application to the case study of an individual-based model of a fish population. *Ecological Modelling, 193*(3–4), 479–491.

Godfrey-Smith, P. (2009). Models and fictions in science. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 143*(1), 101–116.

Goldacre, B., Drysdale, H., Dale, A., Milosevic, I., Slade, E., Hartley, P., Marston, C., Powell-Smith, A., Heneghan, C., & Mahtani, K. R. (2019). COMPare: A prospective cohort study correcting and monitoring 58 misreported trials in real time. *Trials, 20*(1), 118.

Graunt, J. (1662). *Natural and political observations mentioned in a following index, and made upon the bills of mortality*. Tho. Roycroft for John Martin, James Allestry, and Tho. Dicas.

Gray, J., Bijak, J., & Bullock, S. (2016). Deciding to disclose – A decision theoretic agent model of pregnancy and alcohol misuse. In J. Van Bavel & A. Grow (Eds.), *Agent-based modelling in population studies: Concepts, methods, and applications* (pp. 301–340). Springer.

Gray, J., Hilton, J., & Bijak, J. (2017). Choosing the choice: Reflections on modelling decisions and behaviour in demographic agent-based models. *Population Studies, 71*(Supp), 85–97.

Grazzini, J., Richiardi, M.G., & Tsionas, M. (2017). Bayesian estimation of agent-based models. *Journal of Economic Dynamics and Control, 77*(1), 26–47.

Greenwood, M. J. (2005). Modeling Migration. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (pp. 725–734). Elsevier.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand, T., & DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science, 310*(5750), 987–991.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., … DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling, 198*(1–2), 115–126.

Grimm, V., Augusiak, J., Focks, A., Frank, B. M., Gabsi, F., Johnston, A. S. A., Liu, C., Martin, B. T., Meli, M., Radchuk, V., Thorbek, P., & Railsback, S. F. (2014). Towards better modelling and decision support: Documenting model development, testing, and analysis using TRACE. *Ecological Modelling, 280*, 129–139.

Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A. S. A., Milles, A., Nabe-Nielsen, J., Polhill, J. G., Radchuk, V., Rohwäder, M.-S., Stillman, R. A., Thiele, J. C., & Ayllón, D. (2020). The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation, 23*(2), 7.

Groen, D. (2016). Simulating refugee movements: Where would you go? *Procedia Computer Science, 80*, 2251–2255.

Groen, D., Bell, D., Arabnejad, H., Suleimenova, D., Taylor, S. J. E., & Anagnostou, A. (2020). Towards modelling the effect of evolving violence on forced migration. In *Proceedings of the Winter Simulation Conference 2019* (pp. 297–307). IEEE.

Groth, P., & Moreau, L. (2013). *PROV-overview – An overview of the PROV family of documents. Technical report*. World Wide Web Consortium.

Grow, A., & Van Bavel, J. (2015). Assortative mating and the reversal of gender inequality in education in Europe: An agent-based model. *PLoS One, 10*(6), e0127806.

Gurak, T., & Caces, F. (1992). Migration networks and the shaping of migration systems. In M. M. Kritz, L. L. Lim, & H. Zlotnik (Eds.), *International migration systems: A global approach* (pp. 150–176). Clarendon Press.

Hafızoğlu, F. M., & Sen, S. (2012). Analysis of opinion spread through migration and adoption in agent communities. In I. Rahwan, W. Wobcke, S. Sen, & T. Sugawara (Eds.), *PRIMA 2012: Principles and practice of multi-agent systems* (Lecture Notes in Computer Science) (pp. 153–167). Springer.

Hahn, U., Harris, A. J. L., & Corner, A. (2009). Argument content and argument source: An exploration. *Informal Logic, 29*, 337–367.

Hailegiorgis, A., Crooks, A., & Cioffi-Revilla, C. (2018). An agent-based model of rural households' adaptation to climate change. *Journal of Artificial Societies and Social Simulation, 21*(4), 4.

Hainmueller, J., Hopkins, D. J., & Yamamoto, T. (2014). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis, 22*(1), 1–30.

Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences, 112*(8), 2395–2400.

Hainy, M., Müller, W. G., & Wynn, H. P. (2014). Learning functions and approximate Bayesian computation design: ABCD. *Entropy, 16*(8), 4353–4374.

Hanczakowski, M., Zawadzka, K., Pasek, T., & Higham, P. A. (2013). Calibration of metacognitive judgments: Insights from the underconfidence-with-practice effect. *Journal of Memory and Language, 69*, 429–444.

Hardwicke, T. E., Mathur, M. B., MacDonald, K., Nilsonne, G., Banks, G. C., Kidwell, M. C., Hofelich Mohr, A., Clayton, E., Yoon, E. J., Henry Tessler, M., Lenne, R. L., Altman, S., Long, B., & Frank, M. C. (2018). Data availability, reusability, and analytic reproducibility: Evaluating the impact of a mandatory open data policy at the journal Cognition. *Royal Society Open Science, 5*(8), 180448.

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review, 60*(1), 126–142.

Harris, A. J. L., Hahn, U., Madsen, J. K., & Hsu, A. S. (2016). The appeal to expert opinion: Quantitative support for a Bayesian network approach. *Cognitive Science, 40*(6), 1496–1533.

Hassani-Mahmooei, B., & Parris, B. (2012). Climate change and internal migration patterns in Bangladesh: An agent-based model. *Environment and Development Economics, 17*, 763–780.

Haug, S. (2008). Migration networks and migration decision making. *Journal of Ethnic and Migration Studies, 34*(4), 585–605.

Heard, D., Dent, G., Schiffeling, T., & Banks, D. (2015). Agent-based models and microsimulation. *Annual Review of Statistics and Its Application, 2*, 259–272.

Hébert, G. A., Perez, L., & Harati, S. (2018). An agent-based model to identify migration pathways of refugees: The case of Syria. In L. Perez, E.-K. Kim, & R. Sengupta (Eds.), *Agent-based models and complexity science in the age of geospatial big data* (pp. 45–58). Springer.

Hedström, P. (2005). *Dissecting the social: On the principles of analytical sociology*. Springer.

Hedström, P., & Swedberg, R. (Eds.). (1998). *Social mechanisms. An analytical approach to social theory*. Cambridge University Press.

Hedström, P., & Udehn, L. (2011). Analytical sociology and theories of the middle range. In P. Hedström & P. Bearman (Eds.). *The Oxford handbook of analytical sociology* (online). https://doi.org/10.1093/oxfordhb/9780199215362.013.2

Hedström, P., & Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology, 36*, 49–67.

Heiland, F. (2003). The collapse of the Berlin Wall: Simulating state-level east to west German migration patterns. In F. C. Billari & A. Prskawetz (Eds.), *Agent-based computational demography. Using simulation to improve our understanding of demographic behaviour* (pp. 73–96). Kluwer.

Heller, C., & Pezzani, L. (2016). *Death by rescue: The lethal effects of the EU's policies of non-assistance at sea*. Goldsmiths University of London.

Hempel, C. G. (1962). Deductive-nomological vs. statistical explanation. In *Scientific explanation, space, and time. Minnesota studies in the philosophy of science* (Vol. 3, pp. 98–169). University of Minnesota Press.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature, 466*(7302), 29.

Henzinger, T., Jobstmann, B., & Wolf, V. (2011). Formalisms for specifying markovian population models. *International Journal of Foundations of Computer Science, 22*(04), 823–841.

Herzog, T. N., Scheuren, F. J., & Winkler, W. E. (2007). *Data quality and record linkage techniques*. Springer.

Higdon, D., Gattiker, J., Williams, B., & Rightley, M. (2008). Computer model calibration using high-dimensional output. *Journal of the American Statistical Association, 103*(482), 570–583.

Higham, P. A., Zawadzka, K., & Hanczakowski, M. (2015). Internal mapping and its impact on measures of absolute and relative metacognitive accuracy. In *The Oxford handbook of Metamemory*. OUP.

Highhouse, S. (2007). Designing experiments that generalize. *Organizational Research Methods, 12*(3), 554–566.

Hilton, J. (2017). *Managing uncertainty in agent-based demographic models*. PhD Thesis, University of Southampton.

Hilton, J., & Bijak, J. (2016). Design and analysis of demographic simulations. In J. Van Bavel & A. Grow (Eds.), *Agent-based modelling in population studies: Concepts, methods, and applications* (pp. 211–235). Springer.

Himmelspach, J., & Uhrmacher, A. M. (2009). What contributes to the quality of simulation results? In: *2009 INFORMS Simulation Society research workshop* (pp. 125–129). Available via http://eprints.mosi.informatik.uni-rostock.de/346/ (as of 1 February 2021).

Hinsch, M., & Bijak, J. (2019). *Rumours lead to self-organized migration routes*. Paper for the Agent-based Modelling Hub, Artificial Life conference 2019, Newcastle. Available via www.baps-project.eu (as of 1 August 2021).

Hobcraft, J. (2007). Towards a scientific understanding of demographic behaviour. *Population – English Edition, 62*(1), 47–51.

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial. *Statistical Science, 14*(4), 382–417.

Holland, J. H. (2012). *Signals and boundaries*. MIT Press.

Hooten, M. B., Johnson, D. S., & Brost, B. M. (2021). Making recursive Bayesian inference accessible. *American Statistician, 75*(2), 185–194.

Hooten, M., Wikle, C., & Schwob, M. (2020). Statistical implementations of agent-based demographic models. *International Statistical Review, 88*(2), 441–461.

Hopcroft, J. E., & Ullman, J. D. (1979). *Introduction to automata theory, languages, and computation*. Addison-Wesley.

Hovland, C., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *The Public Opinion Quarterly, 15*, 635–650.

Hughes, C., Zagheni, E., Abel, G. J., Wiśniowski, A., Sorichetta, A., Weber, I., & Tatem, A. J. (2016). *Inferring migrations: Traditional methods and new approaches based on Mobile phone, social* *media, and other big data. Feasibility study on inferring (labour) mobility and migration in the European Union from big data and social media data* (Report for the European Commission). Publications Office of the EU.

Hugo, G., Abbasi-Shavazi, M. J., & Kraly, E. P. (Eds.). (2018). *Demography of refugee and forced migration* (International studies in population, Vol. 13). Springer.

IOM. (2021). *Missing migrants: Mediterranean*. IOM GMDAC. Accessible via: https://missing-migrants.iom.int/region/mediterranean? (as of 9 February 2021)

Isernia, P., Urso, O., Gyuzalyan, H., & Wilczyńska, A. (2018). *A review of empirical surveys of asylum-related migrants*. Report, European Asylum Support Office.

Jacobs, S. (1991). John Stuart mill on induction and hypotheses. *Journal of the History of Philosophy, 29*(1), 69–83.

Jaeger, D. A., Dohmen, T., Falk, A., Huffman, D., Sunde, U., & Bonin, H. (2010). Direct evidence on risk attitudes and migration. *The Review of Economics and Statistics, 92*(3), 684–689.

Jager, W. (2017). Enhancing the realism of simulation (EROS): On implementing and developing psychological theory in social simulation. *Journal of Artificial Societies and Social Simulation, 20*(3), 14.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*(5), 524–532.

Johnson, B. R. (2010). Eliminating the mystery from the concept of emergence. *Biology and Philosophy, 25*(5), 843–849.

Jones, B., & Nachtsheim, C. J. (2011). A class of three-level designs for definitive screening in the presence of second-order effects. *Journal of Quality Technology, 43*(1), 1–15.

Jones, B., & Nachtsheim, C. J. (2013). Definitive screening designs with added two-level categorical factors. *Journal of Quality Technology, 45*(2), 121–129.

Jones, C. W., Keil, L. G., Holland, W. C., Caughey, M. C., & Platts-Mills, T. F. (2015). Comparison of registered and published outcomes in randomized controlled trials: A systematic review. *BMC Medicine, 13*(1), 282.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291.

Kamiński, B. (2015). Interval metamodels for the analysis of simulation input-output relations. *Simulation Modelling Practice and Theory, 54*, 86–100.

Kashyap, R., & Villavicencio, F. (2016). The dynamics of son preference, technology diffusion, and fertility decline underlying distorted sex ratios at birth: A simulation approach. *Demography, 53*(5), 1261–1281.

Kemel, E., & Paraschiv, C. (2018). Deciding about human lives: An experimental measure of risk attitudes under prospect theory. *Social Choice and Welfare, 51*, 163–192.

Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society B, 63*(3), 425–464.

Kennedy, M. C., & Petropoulos, G. P. (2016). GEM-SA: The Gaussian emulation machine for sensitivity analysis. In G. P. Petropoulos & P. K. Srivastava (Eds.), *Sensitivity analysis in earth observation modelling* (pp. 341–361). Elsevier.

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review, 2*(3), 196–217.

Kersting, K., Plagemann, C., Pfaff, P., & Burgard, W. (2007). Most likely heteroscedastic Gaussian process regression. In Z. Ghahramani (Ed.), *Proceedings of the 24th International Conference on Machine Learning, Corvallis, OR, 2007*. Association for Computing Machinery.

Keyfitz, N. (1971). Models. *Demography, 8*(4), 571–580.

Keyfitz, N. (1972). On Future Population. *Journal of the American Statistical Association, 67*(338), 347–363.

Keyfitz, N. (1981). The limits of population forecasting. *Population and Development Review, 7*(4), 579–593.

Kim, J. K., & Shao, J. (2014). *Statistical methods for handling incomplete data*. CRC Press/ Chapman & Hall.

King, R. (2002). Towards a new map of European migration. *International Journal of Population Geography, 8*(2), 89–106.

Kingsley, P. (2016). *The new Odyssey: The story of Europe's refugee crisis*. Faber & Faber.

Kirk, P. D. W., Babtie, A. C., & Stumpf, M. P. H. (2015). Systems biology (un)certainties. *Science, 350*, 386–388.

Klabunde, A. (2011). What explains observed patterns of circular migration? An agent-based model. In *17th International Conference on Computing in Economics and Finance* (pp. 1–26).

Klabunde, A. (2014). Computational economic modeling of migration. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2470525 (as of 1 February 2021).

Klabunde, A., & Willekens, F. (2016). Decision making in agent-based models of migration: State of the art and challenges. *European Journal of Population, 32*(1), 73–97.

Klabunde, A., Zinn, S., Leuchter, M., & Willekens, F. (2015). *An agent-based decision model of migration, embedded in the life course: Description in ODD+D format* (MPIDR working paper WP 2015-002). Max Planck Institute for Demographic Research.

Klabunde, A., Zinn, S., Willekens, F., & Leuchter, M. (2017). Multistate modelling extended by behavioural rules: An application to migration. *Population Studies, 71*(Supp), 51–67.

Kleijnen, J. P. C. (1995). Verification and validation of simulation models. *European Journal of Operational Research, 82*(1), 145–162.

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Bahník, Š., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., Brumbaugh, C. C., Cemalcilar, Z., Chandler, J., Cheong, W., Davis, W. E., Devos, T., Eisner, M., Frankowska, N., Furrow, D., Galliani, E. M., … Nosek, B. A. (2014). Investigating variation in replicability. *Social Psychology, 45*(3), 142–152.

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., … Nosek, B. A. (2018). Many labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science, 1*(4), 443–490.

Kniveton, D., Smith, C., & Wood, S. (2011). Agent-based model simulations of future changes in migration flows for Burkina Faso. *Global Environmental Change, 21*, S34–S40.

Kok, L. D. (2016). *Forecasting violence induced human mobility flows: Introducing fear to the decision model. Steps towards establishing a conceptual framework of violence induced human mobility* (Report for Intergovernmental Consultations on Migration, Asylum and Refugees). IGC.

Köster, T., Warnke, T., & Uhrmacher, A. M. (2020). Partial evaluation via code generation for static stochastic reaction network models. In *Proceedings of the 2020 ACM SIGSIM conference on principles of advanced discrete simulation, Association for Computing Machinery, Miami, FL, Spain, SIGSIM-PADS '20* (pp. 159–170).

Kovera, M. B. (2010). Confounding. In N. J. Salkind (Ed.), *Encyclopedia of research design*. Sage.

Kozlov, M. D., & Johansen, M. K. (2010). Real behavior in virtual environments: Psychology experiments in a simple virtual-reality paradigm using video games. *Cyberpsychology, Behavior and Social Networking, 13*(6), 711–714.

Kritz, M., Lim, L. L., & Zlotnik, H. (Eds.). (1992). *International migration systems: A global approach*. Clarendon Press.

Kulu, H., & Milevski, N. (2007). Family change and migration in the life course: An introduction. *Demographic Research, 17*(19), 567–590.

Lattner, C., & Adve, V. (2004). LLVM: A compilation framework for lifelong program analysis & transformation. In *Proceedings of the international symposium on code generation and optimization: Feedback-directed and runtime optimization. CGO '04*. IEEE.

Law, A. (2006). *Simulation modeling and analysis* (4th ed.). McGraw-Hill.

Lazega, E., & Snijders, T. A. B. (Eds.). (2016). *Multilevel network analysis for the social sciences. Theory, methods and applications*. Springer.

Lee, E. S. (1966). A theory of migration. *Demography, 3*(1), 47–57.

Lee, M. D., Criss, A. H., Devezer, B., Donkin, C., Etz, A., Leite, F. P., Matzke, D., Rouder, J. N., Trueblood, J. S., White, C. N., & Vandekerckhove, J. (2019). Robust modeling in cognitive science. *Computational Brain & Behavior, 2*(3), 141–153.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology, 66*(1), 799–823.

Leurs, K., & Smets, K. (2018). Five questions for digital migration studies: Learning from digital connectivity and forced migration in(to) Europe. *Social Media & Society, 4*(1), 205630511876442.

Lieberoth, A. (2014). Shallow gamification: Testing psychological effects of framing an activity as a game. *Games and Culture, 10*(3), 229–248.

Liepe, J., Filippi, S., Komorowski, M., & Stumpf, M. P. H. (2013). Maximizing the information content of experiments in systems biology. *PLoS Computational Biology, 9*, e1002888.

Lin, L., Carley, K. M., & Cheng, S.-F. (2016). An agent-based approach to human migration movement. In *Proceedings of the Winter Simulation Conference 2016* (pp. 3510–3520). IEEE.

Lipton, P. (1991/2004). *Inference to the best explanation* (1st/2nd ed.). Routledge.

Little, R. J. A., & Rubin, D. B. (2020). *Statistical analysis with missing data* (3rd ed.). Wiley.

Lomnicki, A. (1999). Individual-based models and the individual-based approach to population ecology. *Ecological Modelling, 115*(2–3), 191–198.

Lorenz, T. (2009). Abductive fallacies with agent-based modelling and system dynamics. In F. Squazzoni (Ed.), *Epistemological aspects of computer simulation in the social sciences* (Lecture Notes in Artificial Intelligence, 5466) (pp. 141–152). Springer.

Lovreglio, R., Ronchi, E., & Nilsson, D. (2016). An evacuation decision model based on perceived risk, social influence and Behavioural uncertainty. *Simulation Modelling Practice and Theory, 66*, 226–242.

Lucas, R. E., Jr. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy, 1*, 19–46.

Lutz, W. (2012). Demographic metabolism: A predictive theory of socioeconomic change. *Population and Development Review, 38*(Suppl), 283–301.

Lynch, S. M. (2007). *Introduction to applied Bayesian statistics and estimation for social scientists*. Springer.

Mabogunje, A. L. (1970). Systems approach to a theory of rural-urban migration. *Geographical Analysis, 2*(1), 1–18.

MacKay, D. J. C. (1992). Bayesian interpolation. *Neural Computation, 4*(3), 415–447.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide* (2nd ed.). Erlbaum.

Maddux, J. E., & Rogers, R. W. (1980). Effects of source expertness, physical attractiveness, and supporting arguments on persuasion: A case of brains over beauty. *Journal of Personality and Social Psychology, 39*, 235–244.

Marin, J. M., Pudlo, P., Robert, C. P., & Ryder, R. J. (2012). Approximate Bayesian computational methods. *Statistics and Computing, 22*(6), 1167–1180.

Masad, D., & Kazil, J. L. (2015). Mesa: An agent-based modeling framework. In K. Huff & J. Bergstra (Eds.), *Proceedings of the 14th Python in science conference* (pp. 51–58).

Massey, D. S. (2002). A synthetic theory of international migration. In V. Iontsev (Ed.), *World in the mirror of international migration* (pp. 142–152). MAX Press.

Massey, D. S., & Zenteno, R. M. (1999). The dynamics of mass migration. *Proceedings of the National Academy of Sciences of the United States of America, 96*(9), 5328–5335.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., & Taylor, J. E. (1993). Theories of international migration: Review and appraisal. *Population and Development Review, 19*(3), 431–466.

Mauboussin, A. & Mauboussin, M. J. (2018). If you say something is "likely," how likely do people think it is? *Harvard Business Review*, July 3. https://hbr.org/2018/07/if-you-say-something-is-likely-how-likely-do-people-think-it-is

McAlpine, A., Kiss, L., Zimmerman, C., & Chalabi, Z. (2021). Agent-based modeling for migration and modern slavery research: A systematic review. *Journal of Computational Social Science, 4*, 243–332.

McAuliffe, M., & Koser, K. (2017). *A Long way to go. Irregular migration patterns, processes, drivers and decision making*. ANU Press.

McGinnies, E., & Ward, C. D. (1980). Better liked than right: Trustworthiness and expertise as factors in credibility. *Personality and Social Psychology Bulletin, 6*, 467–472.

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics, 21*(2), 239–245.

Merton, R. K. (1949). *Social theory and social structure*. The Free Press.

Miłkowski, M., Hensel, W. M., & Hohol, M. (2018). Replicability or reproducibility? On the replication crisis in computational neuroscience and sharing only relevant detail. *Journal of Computational Neuroscience, 45*(3), 163–172.

Mintz, A., Redd, S. B., & Vedlitz, A. (2006). Can we generalize from student experiments to the real world in political science, military affairs, and international relations? *Journal of Conflict Resolution, 50*(5), 757–776.

Mironova, V., Mrie, L., & Whitt, S. (2019). Risk tolerance during conflict: Evidence from Aleppo, Syria. *Journal of Peace Research, 56*(6), 767–782.

Mol, J. M. (2019). Goggles in the lab: Economic experiments in immersive virtual environments. *Journal of Behavioral and Experimental Economics, 79*, 155–164.

Morgan, S. P., & Lynch, S. M. (2001). Success and future of demography. The role of data and methods. *Annals of the New York Academy of Sciences, 954*, 35–51.

Moussaïd, M., Kapadia, M., Thrash, T., Sumner, R. W., Gross, M., Helbing, D., & Hölscher, C. (2016). Crowd behaviour during high-stress evacuations in an immersive virtual environment. *Journal of the Royal Society Interface, 13*(122), 20160414.

MUCM. (2021). *Managing uncertainty in complex models*. Online resource, via https://mogp-emulator.readthedocs.io/en/latest/methods/meta/MetaHomePage.html (as of 1 March 2021).

Müller, B., Bohn, F., Dreßler, G., et al. (2013). Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol. *Environmental Modelling & Software, 48*, 37–48.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*(1), 0021.

Naivinit, W., Le Page, C., Trébuil, G., & Gajaseni, N. (2010). Participatory agent-based modeling and simulation of Rice production and labor migrations in Northeast Thailand. *Environmental Modelling & Software, 25*(11), 1345–1358.

Napierała, J., Hilton, J., Forster, J. J., Carammia, M., & Bijak, J. (2021). Toward an early warning system for monitoring asylum-related migration flows in Europe. *International Migration Review*, forthcoming.

Naqvi, A. A., & Rehm, M. (2014). A multi-agent model of a low income economy: Simulating the distributional effects of natural disasters. *Journal of Economic Interaction and Coordination, 9*(2), 275–309.

National Academies of Sciences, Engineering, and Medicine. (2019). *Reproducibility and replicability in science*. The National Academies Press.

Neal, R. M. (1996). *Bayesian learning for neural networks*. Springer.

Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's renaissance. *Annual Review of Psychology, 69*(1), 511–534.

Noble, J., Silverman, E., Bijak, J., et al. (2012). Linked lives: The utility of an agent-based approach to modeling partnership and household formation in the context of social care. In *Proceedings of the Winter Simulation Conference 2012*. IEEE.

North, M. J., Collier, N. T., Ozik, J., Tatara, E. R., Macal, C. M., Bragen, M., & Sydelko, P. (2013). Complex adaptive systems modeling with repast Simphony. *Complex Adaptive Systems Modeling, 1*(1), 3.

Nosek, B. A., & Errington, T. M. (2017). Reproducibility in cancer biology: Making sense of replications. *eLife, 6*, e23383.

Nosek, B. A., & Lakens, D. (2014). Registered reports: A method to increase the credibility of published results. *Social Psychology, 45*(3), 137–141.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science, 348*(6242), 1422.

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America, 115*(11), 2600–2606.

Nowak, A., Rychwalska, A., & Borkowski, W. (2013). Why simulate? To develop a mental model. *Journal of Artificial Societies and Social Simulation, 16*(3), 12.

NRC [National Research Council]. (2000). *Beyond six billion. Forecasting the World's population*. National Academies Press.

Nubiola, J. (2005). Abduction or the logic of surprise. *Semiotica, 153*(1/4), 117–130.

O'Hagan, A. (2013). *Polynomial Chaos: A tutorial and critique from a Statistician's perspective mimeo*. University of Sheffield. Via http://tonyohagan.co.uk/academic/pdf/Polynomial-chaos.pdf (as of 1 November 2019)

Oakley, J., & O'Hagan, A. (2002). Bayesian inference for the uncertainty distribution of computer model outputs. *Biometrika, 89*, 769–784.

Oakley, J. E., & O'Hagan, A. (2004). Probabilistic sensitivity analysis of complex models: A Bayesian approach. *Journal of the Royal Statistical Society B, 66*(3), 751–769.

Obels, P., Lakens, D., Coles, N. A., Gottfried, J., & Green, S. A. (2020). Analysis of open data and computational reproducibility in registered reports in psychology. *Advances in Methods and Practices in Psychological Science, 3*(2), 229–237.

Öberg, S. (1996). Spatial and economic factors in future south-North Migration. In W. Lutz (Ed.), *The future population of the world: What can we assume today?* (pp. 336–357). Earthscan.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.

Ozik, J., Wilde, M., Collier, N., & Macal, C. M. (2014). Adaptive simulation with repast Simphony and swift. In *Lecture Notes in Computer Science*. Springer.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika, 41*(1), 100–115.

Pascal, B. (1654). *Traité du triangle arithmétique, avec quelques autres traités sur le même sujet*. Guillaume Desprez.

Pascal, B. (1670). *Pensées*. Editions de Port Royal.

Peck, S. (2012). Agent-based models as fictive instantiations of ecological processes. *Philosophy, Theory and Practice in Biology, 4*(3), 1–2.

Peirce, C. S. (1878/2014). Deduction, induction and hypothesis. In C. De Waal (Ed.), *Illustrations of the logic of science* (pp. 167–184) [original text from *Popular science monthly* 13, 470–482, ibid].

Peng, D., Warnke, T., & Uhrmacher, A. M. (2015). Domain-specific languages for flexibly experimenting with stochastic models. *Simulation Notes Europe, 25*(2), 117–122.

Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology, 41*, 847–855.

Pilditch, T. D., Madsen, J. K., & Custers, R. (2020). False prophets and Cassandra's curse: The role of credibility in belief updating. *Acta Psychologica, 202*, 102956.

Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ, 1*, e175.

Poile, C., & Safayeni, F. (2016). Using computational modeling for building theory: A double-edged sword. *Journal of Artificial Societies and Social Simulation, 19*(3), 8.

Polhill, J. G., Sutherland, L.-A., & Gotts, N. M. (2010). Using qualitative evidence to enhance an agent-based modelling system for studying land use change. *Journal of Artificial Societies and Social Simulation, 13*(2), 10.

Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies, 47*(11), 1451–1458.

Poole, D., & Raftery, A. E. (2000). Inference for deterministic simulation models: The Bayesian melding approach. *Journal of the American Statistical Association, 95*(452), 1244–1255.

Pope, A. J., & Gimblett, R. (2015). Linking Bayesian and agent-based models to simulate complex social-ecological systems in semi-arid regions. *Frontiers in Environmental Science, 3*, art. 55.

Popper, K. R. (1935). *Logik der Forschung*. Julius Springer Verlag, Wien [(1959) *The logic of scientific discovery*. Hutchinson].

Popper, K. R. (1982). *The open universe. An argument for indeterminism*. Hutchinson.

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology, 34*, 243–281.

Poulain, M., Perrin, N., & Singleton, A. (Eds.). (2006). *Towards harmonised European statistics on international migration*. Presses Universitaires de Louvain.

Preston, S. H., & Coale, A. J. (1982). Age structure/growth, attrition and accession: A new synthesis. *Population Index, 48*(2), 217–259.

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology, 14*(2), 154–166.

Rad, M. S., Martingano, A. J., & Ginges, J. (2018). Toward a psychology of *Homo sapiens*: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences, 115*(45), 11401.

Raftery, A. E., Givens, G. H., & Zeh, J. E. (1995). Inference from a deterministic population dynamics model for bowhead whales. *Journal of the American Statistical Association, 90*(430), 402–416.

Rahmandad, H., & Sterman, J. D. (2012). Reporting guidelines for simulation-based research in social sciences. *System Dynamics Review, 28*(4), 396–411.

Railsback, S. F., Lytinen, S. L., & Jackson, S. K. (2006). Agent-based simulation platforms: Review and development recommendations. *Simulation, 82*(9), 609–623.

Ranjan, P., & Spencer, N. (2014). Space-filling Latin hypercube designs based on randomization restrictions in factorial experiments. *Statistics & Probability Letters, 94*, 239–247.

Rao, A. S., & Georgeff, M. P. (1991). Modeling rational agents within a BDI architecture. In *Proceedings of the second international conference on principles of knowledge representation and reasoning, KR'91, Cambridge, MA* (pp. 473–484). Morgan Kaufmann.

Ravenstein, E. G. (1885). The laws of migration. *Journal of the Statistical Society of London, 48*(2), 167–227.

Raymer, J., Wiśniowski, A., Forster, J. J., Smith, P. W. F., & Bijak, J. (2013). Integrated modeling of European migration. *Journal of the American Statistical Association, 108*(503), 801–819.

Read, S. J., & Monroe, B. M. (2008). Computational models in personality and social psychology. In R. Sun (Ed.), *The Cambridge handbook of computational psychology* (pp. 505–529). Cambridge University Press.

Reichlová, N. (2005). *Can the theory of motivation explain migration decisions?* (Working papers IES, 97). Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies.

Reinhardt, O., & Uhrmacher, A. M. (2017). An efficient simulation algorithm for continuous-time agent-based linked lives models. In *Proceedings of the 50th Annual Simulation Symposium, International Society for Computer Simulation, San Diego, CA, USA, ANSS '17*, pp 9:1–9:12.

Reinhardt, O., Hilton, J., Warnke, T., Bijak, J., & Uhrmacher, A. (2018). Streamlining simulation experiments with agent-based models in demography. *Journal of Artificial Societies and Social Simulation, 21*(3), 9.

Reinhardt, O., Uhrmacher, A. M., Hinsch, M., & Bijak, J. (2019). Developing agent-based migration models in pairs. In *Proceedings of the Winter Simulation Conference 2019* (pp. 2713–2724). IEEE.

Reinhardt, O., Warnke, T., & Uhrmacher, A. M. (2021). A language for agent-based discrete-event modeling and simulation of linked lives. In *ACM Transactions on Modeling and Computer Simulation.* (under review).

Richiardi, M. (2017). The future of agent-based modelling. *Eastern Economic Journal, 43*(2), 271–287.

Rieger, M. O., Wang, M., & Hens, T. (2017). Estimating cumulative prospect theory parameters from an international survey. *Theory and Decision, 82*(4), 567–596.

Rogers, A., & Castro, L. J. (1981). *Model migration schedules* (IIASA Report RR8130). IIASA.

Rogers, A., Little, J., & Raymer, J. (2010). *The indirect estimation of migration: Methods for dealing with irregular, inadequate, and missing data*. Springer.

Romanowska, I. (2015). So you think you can model? A guide to building and evaluating archaeological simulation models of dispersals. *Human Biology, 87*(3), 169–192.

Rossetti, T., & Hurtubia, R. (2020). An assessment of the ecological validity of immersive videos in stated preference surveys. *Journal of Choice Modelling, 34*, 100198.

Roustant, O., Ginsbourger, D., & Deville, Y. (2012). DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodelling and optimisation. *Journal of Statistical Software, 51*(1), 1–55.

Ruscheinski, A., & Uhrmacher, A. (2017). Provenance in modeling and simulation studies – Bridging gaps. In *Proceedings of the Winter Simulation Conference 2017* (pp. 872–883). IEEE.

Ruscheinski, A., Wilsdorf, P., Dombrowsky, M., & Uhrmacher, A. M. (2019). Capturing and reporting provenance information of simulation studies based on an artifact-based workflow approach. In *Proceedings of the 2019 ACM SIGSIM conference on principles of advanced discrete simulation* (pp. 185–196). Association for Computing Machinery.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion, 30*(4), 344–360.

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior, 69*, 371–380.

Salecker, J., Sciaini, M., Meyer, K. M., & Wiegand, K. (2019). The nlrx R package: A next-generation framework for reproducible NetLogo model analyses. *Methods in Ecology and Evolution, 10*(11), 1854–1863.

Saltelli, A., Tarantola, S., & Campolongo, F. (2000). Sensitivity analysis as an ingredient of modeling. *Statistical Science, 15*(4), 377–395.

Saltelli, A., Chan, K., & Scott, E. M. (2008). *Sensitivity Analysis*. Wiley.

Sánchez-Querubín, N., & Rogers, R. (2018). Connected routes: Migration studies with digital devices and platforms. *Social Media & Society, 4*(1), 1–13.

Santner, T. J., Williams, B. J., & Notz, W. I. (2003). *The design and analysis of computer experiments*. Springer.

Sargent, R. G. (2013). Verification and validation of simulation models. *Journal of Simulation, 7*(1), 12–24.

Sawyer, R. K. (2004). Social explanation and computational simulation. *Philosophical Explorations, 7*(3), 219–231.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology, 1*(2), 143–186.

Schelling, T. C. (1978). *Micromotives and macrobehavior*. Norton.

Schimmack, U. (2020). A meta-psychological perspective on the decade of replication failures in social psychology. *Canadian Psychology/Psychologie Canadienne, 61*(4), 364–376.

Schloss, P. D. (2018). Identifying and overcoming threats to reproducibility, replicability, robustness, and generalizability in microbiome research. *MBio, 9*(3), e00525-18.

Schmolke, A., Thorbek, P., DeAngelis, D. L., & Grimm, V. (2010). Ecological models supporting environmental decision making: A strategy for the future. *Trends in Ecology & Evolution, 25*(8), 479–486.

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition and Emotion, 14*(4), 433–440.

Sechrist, G. B., & Milford-Szafran, L. R. (2011). "I depend on you, you depend on me. shouldn't we agree?": The influence of interdependent relationships on individuals' racial attitudes. *Basic and Applied Social Psychology, 33*, 145–156.

Sechrist, G. B., & Young, A. F. (2011). The influence of social consensus information on intergroup attitudes: The moderating effects of Ingroup identification. *The Journal of Social Psychology, 151*, 674–695.

Ševčíková, H., Raftery, A. D., & Waddell, P. A. (2007). Assessing uncertainty in urban simulations using Bayesian melding. *Transportation Research Part B, 41*(6), 652–669.

Sharma, S. (2017). Definitions and models of statistical literacy: A literature review. *Open Review of Educational Research, 4*(1), 118–133.

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology, 69*(1), 487–510.

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., Bahník, Š., Bai, F., Bannard, C., Bonnier, E., Carlsson, R., Cheung, F., Christensen, G., Clay, R., Craig, M. A., Dalla Rosa, A., Dam, L., Evans, M. H., Flores Cervantes, I., … Nosek, B. A. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science, 1*(3), 337–356.

Silveira, J. J., Espíndola, A. L., & Penna, T. J. P. (2006). Agent-based model to rural-Urban migration analysis. *Physica A: Statistical Mechanics and its Applications, 364*, 445–456.

Silverman, E. (2018). *Methodological investigations in agent-based modelling, with applications for the social sciences* (Methodos series, vol. 13). Springer.

Silverman, E., Bijak, J., Hilton, J., Cao, V. D., & Noble, J. (2013). When demography met social simulation: A tale of two modelling approaches. *Journal of Artificial Societies and Social Simulation, 16*(4), 9.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366.

Simon, M. (2019). Path dependency and adaptation: The effects of policy on migration systems. *Journal of Artificial Societies and Social Simulation, 22*(2), 2.

Simon, M., Schwartz, C., Hudson, D., & Johnson, S. D. (2016). Illegal migration as adaptation: An agent based model of migration dynamics. In *2016 APSA Annual Meeting & Exhibition*.

Simon, M., Schwartz, C., Hudson, D., & Johnson, S. D. (2018). A data-driven computational model on the effects of immigration policies. *Proceedings of the National Academy of Sciences, 115*(34), E7914–E7923.

Simons, D. J., Holcombe, A. O., & Spellman, B. A. (2014). An introduction to registered replication reports at perspectives on psychological science. *Perspectives on Psychological Science, 9*(5), 552–555.

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science, 12*(6), 1123–1128.

Singleton, A. (2016). *Migration and asylum data for policy-making in the European Union – The problem with numbers* (CEPS paper no. 89). Centre for Europe and Policy Studies.

Sisson, S. A., Fan, Y., & Beaumont, M. (2018). *Handbook of approximate Bayesian computation*. Chapman and Hall/CRC.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy, 70*(5), 80–93.

Smajgl, A., & Bohensky, E. (2013). Behaviour and space in agent-based modelling: Poverty patterns in East Kalimantan, Indonesia. *Environmental Modelling and Software, 45*, 8–14.

Smaldino, P. E. (2016). Models are stupid, and we need more of them. In R. Vallacher, S. Read, & A. Nowak (Eds.), *Computational models in social psychology* (pp. 311–331). Psychology Press.
Smith, R. C. (2013). *Uncertainty quantification: Theory, implementation, and applications*. SIAM.
Smith, C. (2014). Modelling migration futures: Development and testing of the rainfalls agent-based migration model – Tanzania. *Climate and Development, 6*(1), 77–91.
Smith, C., Wood, S., & Kniveton, D. (2010). Agent based modelling of migration decision making. In *Proceedings of the European workshop on multi-agent systems (EUMAS-2010)* (p. 15).
Sobol', I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation, 55*(1–3), 271–280.
Sokolowski, J. A., Banks, C. M., & Hayes, R. L. (2014). Modeling population displacement in the Syrian City of Aleppo. In *Proceedings of the Winter Simulation Conference 2014* (pp. 252–263). IEEE.
Spiegelhalter, D. J., & Riesch, H. (2011). Don't know, can't know: Embracing deeper uncertainties when analysing risks. *Philosophical Transactions of the Royal Society A, 369*(1956), 4730–4750.
Stagge, J. H., Rosenberg, D. E., Abdallah, A. M., Akbar, H., Attallah, N. A., & James, R. (2019). Assessing data availability and research reproducibility in hydrology and water resources. *Scientific Data, 6*(1), 190030.
Stan Development Team. (2021). *Stan modeling language users guide and reference manual*. Retrieved from http://mc-stan.org/index.html (as of 1 February 2020).
Stark, O. (1991). *The migration of labor*. Basil Blackwell.
Stark, O., & Bloom, D. E. (1985). The new economics of labor migration. *American Economic Review, 75*(2), 173–178.
Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—Or vice versa. *Journal of the American Statistical Association, 54*, 30–34.
Stillwell, J., Bell, M., Ueffing, P., Daras, K., Charles-Edwards, E., Kupiszewski, M., & Kupiszewska, D. (2016). Internal migration around the world: Comparing distance travelled and its frictional effect. *Environment and Planning A, 48*(8), 1657–1675.
Stodden, V., Guo, P., & Ma, Z. (2013). Toward reproducible computational research: An empirical analysis of data and code policy adoption by journals. *PLoS One, 8*(6), e67111.
Strevens, M. (2016). How idealizations provide understanding. In S. R. Grimm, C. Baumberger, & S. Ammon (Eds.), *Explaining understanding: New essays in epistemology and the philosophy of science* (pp. 37–49). Routledge.
Suhay, E. (2015). Explaining group influence: The role of identity and emotion in political conformity and polarization. *Political Behavior, 37*, 221–251.
Suleimenova, D., & Groen, D. (2020). How policy decisions affect refugee journeys in South Sudan: A study using automated ensemble simulations. *Journal of Artificial Societies and Social Simulation, 23*(1), 17.
Suleimenova, D., Bell, D., & Groen, D. (2017). Towards an automated framework for agent-based simulation of refugee movements. In *Proceedings of the Winter Simulation Conference 2017* (pp. 1240–1251). IEEE.
Suriyakumaran, A., & Tamura, Y. (2016). Asylum provision: A review of economic theories. *International Migration, 54*(4), 18–30.
Swets, J. A. (2014). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. Psychology Press.
Tabeau, E. (2009). *Victims of the Khmer rouge regime in Cambodia, April 1975 to January 1979: A critical assessment of existing estimates and recommendations for court*. Expert report, Extraordinary Chambers of the Courts of Cambodia.
Tack, L., Goos, P., & Vandebroek, M. (2002). Efficient Bayesian designs under heteroscedasticity. *Journal of Statistical Planning and Inference, 104*(2), 469–483.
Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American Economic Review, 100*, 557–571.

Tavaré, S., Balding, D. J., Griffiths, R. C., & Donnelly, P. (1997). Inferring coalescence times from DNA sequence data. *Genetics, 145*(2), 505–518.

ten Broeke, G., van Voorn, G., & Ligtenberg, A. (2016). Which sensitivity analysis method should I use for my agent-based model? *Journal of Artificial Societies and Social Simulation, 19*(1), 5.

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS One, 6*(6), e21101.

Tetlock, P. E., & Gardner, D. (2015). *Superforecasting: The art and science of prediction*. Random House.

Thompson, E. L., & Smith, L. A. (2019). Escape from model-land. *Economics: The Open-Access, Open-Assessment E-Journal, 13*(40), 1–17. https://www.econstor.eu/handle/10419/204779

Tipping, M. E. (2001). Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research, 1*, 211–244.

Tobin, S. J., & Raymundo, M. M. (2009). Persuasion by causal arguments: The motivating role of perceived causal expertise. *Social Cognition, 27*(1), 105–127.

Todaro, M. P. (1969). A model of labor migration and Urban unemployment in less developed countries. *The American Economic Review, 59*(1), 138–148.

Troitzsch, K. G. (2017). Using empirical data for designing, calibrating and validating simulation models. In W. Jager, R. Verbrugge, A. Flache, G. de Roo, L. Hoogduin, & C. Hemelrijk (Eds.), *Advances in social simulation 2015* (pp. 413–427). Springer.

Tsvetkova, M., Wagner, C., & Mao, A. (2018). The emergence of inequality in social groups: Network structure and institutions affect the distribution of earnings in cooperation games. *PLoS One, 13*(7), e0200965. https://doi.org/10.1371/journal.pone.0200965

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297–323.

UN. (2016). *New York declaration for refugees and migrants* (Resolution adopted by the general assembly on 19 September 2016. A/RES/71/1). United Nations.

UNHCR. (1951/1967). *Text of the 1951 convention relating to the status of refugees; text of the 1967 protocol relating to the status of refugees; resolution 2198 (XXI) adopted by the United Nations general assembly with an introductory note by the Office of the United Nations High Commissioner for refugees*. UNHCR.

UNHCR. (2021). *UNHCR refugee statistics*. Via: https://www.unhcr.org/refugee-statistics/ (as of 1 February 2021).

Van Bavel, J., & Grow, A. (Eds.). (2016). *Agent-based modelling in population studies: Concepts, methods, and applications*. Springer.

Van der Vaart, E., Beaumont, M. A., Johnston, A. S. A., & Sibly, R. M. (2015). Calibration and evaluation of individual-based models using approximate Bayesian computation. *Ecological Modelling, 312*, 182–190.

Van Deursen, A., Klint, P., & Visser, J. (2000). Domain-specific languages: An annotated bibliography. *Sigplan Notices, 35*(6), 26–36.

Van Hear, N., Bakewell, O., & Long, K. (2018). Push-pull plus: Reconsidering the drivers of migration. *Journal of Ethnic and Migration Studies, 44*(6), 927–944.

Vernon, I., Goldstein, M., & Bower, R. G. (2010). Galaxy formation: A Bayesian uncertainty analysis. *Bayesian Analysis, 5*(4), 619–669.

Vogel, D., & Kovacheva, V. (2008). *Classification report: Quality assessment of estimates on stocks of irregular migrants* (Report of the Clandestino project). Hamburg Institute of International Economics.

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7*(6), 632–638.

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian infer-

ence for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review, 25*(1), 35–57.
Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge University Press.
Wakker, P., & Deneffe, D. (1996). Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management Science, 42*(8), 1131–1150.
Wall, M., Otis Campbell, M., & Janbek, D. (2017). Syrian refugees and information Precarity. *New Media & Society, 19*(2), 240–254.
Waltemath, D., Adams, R., Bergmann, F. T., Hucka, M., Kolpakov, F., Miller, A. K., Moraru, I. I., Nickerson, D., Sahle, S., Snoep, J. L., & Le Novère, N. (2011). Reproducible computational biology experiments with SED-ML – The simulation experiment description markup language. *BMC Systems Biology, 5*(1), 198.
Wang, S., Verpillat, P., Rassen, J., Patrick, A., Garry, E., & Bartels, D. (2016). Transparency and reproducibility of observational cohort studies using large healthcare databases. *Clinical Pharmacology & Therapeutics, 99*(3), 325–332.
Warnke, T., Klabunde, A., Steiniger, A., Willekens, F., & Uhrmacher, A. M. (2016). ML3: A language for compact modeling of linked lives in computational demography. In *Proceedings of the Winter Simulation Conference 2015*. IEEE.
Warnke, T., Reinhardt, O., Klabunde, A., Willekens, F., & Uhrmacher, A. M. (2017). Modelling and simulating decision processes of linked lives: An approach based on concurrent processes and stochastic race. *Population Studies, 71*(Supp), 69–83.
Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a world beyond "p<0.05". *The American Statistician, 73*(Sup1), 1–19.
Weintraub, E. R. (1977). The microfoundations of macroeconomics: A critical survey. *Journal of Economic Literature, 15*(1), 1–23.
Weisberg, M. (2007). Three kinds of idealization. *Journal of Philosophy, 104*(12), 639–659.
Werth, B., & Moss, S. (2007). Modelling migration in the Sahel: An alternative to cost-benefit analysis. In S. Takahashi, D. Sallach, & J. Rouchier (Eds.), *Advancing social simulation: The first world congress* (pp. 331–342). Springer Japan.
Wicherts, J. M., Bakker, M., & Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLoS One, 6*(11), e26828.
Wilensky, U. (1999). *NetLogo*. https://ccl.northwestern.edu/netlogo (as of 1 February 2021).
Wilensky, U. (2018). *NetLogo 6.0.3 user manual: BehaviorSpace guide*. Available from: https://ccl.northwestern.edu/netlogo/docs/behaviorspace.html (as of 1 February 2021).
Willekens, F. (1990). Demographic forecasting; state-of-the-art and research needs. In C. A. Hazeu & G. A. B. Frinking (Eds.), *Emerging issues in demographic research* (pp. 9–66). Elsevier.
Willekens, F. (1994). Monitoring international migration flows in Europe: Towards a statistical data base combining data from different sources. *European Journal of Population, 10*(1), 1–42.
Willekens, F. (2009). Continuous-time microsimulation in longitudinal analysis. In A. Zaidi, A. Harding, & P. Williamson (Eds.), *New Frontiers in microsimulation modelling* (pp. 413–436). Ashgate.
Willekens, F. (2018). Towards causal forecasting of international migration. *Vienna Yearbook of Population Research, 16*, 1–20.
Willekens, F., Massey, D., Raymer, J., & Beauchemin, C. (2016). International migration under the microscope. *Science, 352*(6288), 897–899.
Willekens, F., Bijak, J., Klabunde, A., & Prskawetz, A. (Eds.). (2017). The science of choice: An introduction. *Population Studies, 71*(Supp), 1–13.
Williams, A. D., & Baláž, V. (2011). Migration, risk, and uncertainty: Theoretical perspectives. *Population, Space and Place, 18*(2), 167–180.
Williams, A. M., & Baláž, V. (2014). Mobility, risk tolerance and competence to manage risks. *Journal of Risk Research, 17*(8), 1061–1088.

Williamson, D., & Goldstein, M. (2015). Posterior belief assessment: Extracting meaningful subjective judgements from Bayesian analyses with complex statistical models. *Bayesian Analysis, 10*(4), 877–908.

Wilsdorf, P., Dombrowsky, M., Uhrmacher, A. M., Zimmermann, J., & van Rienen, U. (2019). Simulation experiment schemas – Beyond tools and simulation approaches. In *Proceedings of the 2019 Winter Simulation Conference*. IEEE.

Wilsdorf, P., Haack, F., Budde, K., Ruscheinski, A., & Uhrmacher, A. M. (2020). Conducting systematic, partly automated simulation studies Unde Venis et quo Vadis. In *17th International Conference of Numerical Analysis and Applied Mathematics, 020001* (AIP conference proceedings 2293(1)). AIP Publishing LCC.

Wintle, B. C., Fraser, H., Wills, B. C., Nicholson, A. E., & Fidler, F. (2019). Verbal probabilities: Very likely to be somewhat more confusing than numbers. *PLoS One, 14*, e0213522.

Wipf, D. P., & Nagarajan, S. S. (2008). A new view of automatic relevance determination. In J. C. Platt, D. Koller, Y. Singer, & S. T. Roweis (Eds.), *Advances in neural information processing systems 20* (pp. 1625–1632). Curran Associates.

Xie, Y. (2000). Demography: Past, present and future. *Journal of the American Statistical Association, 95*(450), 670–673.

Xie, Q., Lu, S., Cóstola, D., & Hensen, J. L. M. (2014). An arbitrary polynomial Chaos-based approach to analyzing the impacts of design parameters on evacuation time under uncertainty. In D. Nilsson, P. van Hees, & R. Jansson (Eds.), *Fire safety science–proceedings of the eleventh international symposium* (pp. 1077–1090). International Association for Fire Safety Science.

Yang, L., & Guo, Y. (2019). Combining pre- and post-model information in the uncertainty quantification of non-deterministic models using an extended Bayesian melding approach. *Information Sciences, 502*, 146–163.

Zaidi, A., Harding, A., & Williamson, P. (Eds.). (2009). *New Frontiers in microsimulation modelling*. Routledge.

Zawadzka, K., & Higham, P. A. (2015). Judgments of learning index relative confidence, not subjective probability. *Memory & Cognition, 43*(8), 1168–1179.

Zeigler, B., & Sarjoughian, H. S. (2017). *Guide to modeling and simulation of systems of systems* (Simulation foundations, methods and applications) (2nd ed.). Springer.

Zeigler, B. P., Muzy, A., & Kofman, E. (2019). *Theory of modeling and simulation* (3rd ed.). Academic.

Zelinsky, W. (1971). The hypothesis of the mobility transition. *Geographical Review, 61*(2), 219–249.

Zinn, S. (2012). A mate-matching algorithm for continuous-time microsimulation models. *International Journal of Microsimulation, 5*(1), 31–51.

Zolberg, A. R. (1989). The next waves: Migration theory for a changing world. *International Migration Review, 23*, 403–430.

Index

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7

T

U

V

W