1

The Toolkit

Alan G. Wilson

Geographical systems are characterised by locations, activities at locations, interactions between them and the infrastructures that carry these activities and flows. They can be described at a great variety of scales, from individuals, organisations and buildings, through neighbourhoods, to towns and cities, regions and countries. There is an understanding, often partial, of these entities, and in many case this understanding is represented in theories which in turn are represented in mathematical models. We can characterise these models, with geography as a core, as geo-mathematical models.

In this book, our main examples are models that represent elements of the global system covering such topics as trade, migration, security and development aid. We also work with examples at finer scales. We review this set of models, along with some outstanding research questions, in order to demonstrate how they now form, between them, an effective toolkit that can be applied not only to particular global systems but more widely in the modelling of complex systems.

These examples have been developed in the context of an EPSRC-funded complexity science programme with twin foci: developing new tools and applying these to real-world problems. In presenting the 'tools' here, it is useful to be aware of Weaver's distinction between systems of disorganised complexity and systems of organised complexity. Both kinds of systems have large numbers of elements, but in the first, there are only weak interactions between them; in the second, some strong interactions. This distinction relates to that between *fast dynamics* and *slow dynamics* – essentially, between systems that can return rapidly to equilibrium following a change and those that are slower. It also relates to those that, from a mathematical point of view, can be modelled by using averaging procedures of various kinds and those more challenging systems that demand a variety of methods, many still the subject of ongoing research. Roughly speaking, systems involving large numbers of people – those travelling to work in a city, for example – fall into the first category, while those involving complex organisations within an economy or physical structures, such as buildings, fall into the second.

Approaches to Geo-mathematical Modelling: New Tools for Complexity Science, First Edition. Edited by Alan G. Wilson.

Companion Website: www.wiley.com/go/wilson/ApproachestoGeo-mathematicalModelling

All complex systems involve nonlinearities. In the case of systems of organised complexity, as we will see, path dependence and the possibility of phase changes make the mathematical aspects of this kind of research particularly interesting. It is through these mechanisms that new structures can be seen to 'emerge', and hence the current notion of *emergent behaviour*.

We proceed by reviewing the main elements of the toolkit in this introductory chapter, and then we proceed to illustrate their use through a series of applications. The headings that follow illustrate the richness of the toolkit.

- Estimating missing data: bi-proportional fitting and principal components analysis (Part 2)
- Dynamics in account-based models (Part 3)
- Space–time statistical analysis (Part 4)
- Real-time response models (Part 5)
- The mathematics of war (Part 6)
- Agent-based models (Part 7)
- Diffusion models (Part 8)
- Game theory (Part 9)
- Networks (Part 10)
- Integrated models (Part 11).

There are three kinds of research questions that lead us to new tools for handling issues in complexity science: firstly, the development of particular tools; secondly, new applications of these tools; and, thirdly, the development of new combinations of tools.

The first category includes the addition of spatial dimensions to Lotka–Volterra models – with applications in trade modelling and security, the latter offering a new dimension in Richardson models. Other examples include adding depth to our understanding of the dynamics of the evolution of centres, or the new interpretation of spatial interaction as a 'threat' in building models of security.

Probably the potentially most fruitful area – illustrating Brian Arthur's argument on the nature of technological development – is the development of new combinations. This also illustrates 'new applications'. One of the London riots' models, for example, combines epidemiological, spatial interaction and probability sub-models. In developing the Colonel Blotto model, we have added space and the idea of 'threat' in combination with game theory. By adding dynamics to migration, trade and input–output models, and by incorporating development aid, we have created possible new approaches to economic development.

In Part 2, we show how to expand – by estimating 'missing data' – some sets of accounts. Historically, examples of account-based models are Rogers' demographic model, Leontief's input–output model and the doubly constrained journey to work model developed on a bi-proportional basis by Fratar but later set in an entropy-maximising framework. In this part, we present three examples of account-based models in which bi-proportional fitting is used either to make data from different sources consistent or to estimate missing data. These are examples of well-known techniques being used creatively in new situations. We present three examples. Firstly, we take European migration. There are good data at the intercountry level, and in- and out-totals are available at sub-regional scales. We use bi-proportional methods to estimate the missing data. Secondly, we again apply methods to find missing data on international trade. We have data on total intercountry trade, and we have sector data for exports and imports. We use the model to estimate flow data by sector. Thirdly, we use data on input–output accounts which are rich for a subset of countries, and we use a principal components method to estimate missing data. The results of the second and third of these examples are used in building the integrated model described in Chapter 17.

In Part 3, we describe an account-based trade model integrated with dynamic adjustment mechanisms on both prices and economic capacity. These mechanisms are rooted in Lotka–Volterra types of equations. In this case, therefore, we are demonstrating the power of integrating different 'tools'. Given the complexities of this task, what is presented is a demonstration model. The spatial Lotka–Volterra type of dynamics, as represented in the retail model, can be seen as a more general archetype of centre dynamics. In this part, we explore the dynamics of such models in more depth.

In Part 4, we discuss different statistical approaches to hypothesis testing for spatial and space–time patterns of crime and other events. Methods for examining point processes are presented in the case of insurgency and piracy, and for riots at the area level. We discuss methods for identifying regularities in observed patterns (e.g. spatial statistics, the K-S test, Monte Carlo methods and simulated annealing), methods for testing theories of those patterns (logistic and conditional logit models) and statistical models that may be used to describe and potentially predict them (self-exciting point process models).

Part 5 provides further examples of combinations. Epidemiology provides the model of propensity to riot; spatial interaction modelling answers the 'Where?' question; and we have a third model of probability of arrest. We present two alternative models of the riots: one illustrates the deployment of discrete-choice spatial models, and the other uses an agent-based approach. The differences between such computational approaches and the mathematical models are explored

Models of war, in a broad sense, have a long history. Richardson's 'arms race' model is an excellent example. This model can be seen as a special case of Lotka–Volterra dynamics. Space is not explicit in the original, and this is clearly a critical feature if such models are to be used strategically. In Part 6, therefore, we extend this model to incorporate space.

Agent-based simulations are widely used across a vast range of disciplines, yet the fundamental characteristics of their behaviour are not analysed in a systematic and mathematically profound manner. In Part 7, we present a toolkit of mathematical techniques, using both the rules that govern multi-agent simulations and time series data gathered from them, to deduce equations that describe the dynamics of their behaviour and to predict rare events in such models. In certain cases, the methods employed also suggest the minimal interventions required to prevent or induce particular behaviours.

In Part 8, we introduce diffusion models. These have a long history in disciplines such as physics but less so in the social sciences. We first, following Medda, show how Turing's model of morphogenesis – with two interacting processes generating spatial structure – can be adapted to urban dynamic structural modelling. We also add, in Appendix B, some mathematical explorations of a different kind of diffusion: the control of insect populations, which are also rooted in Lotka–Volterra mathematics.

In Part 9, we invoke concepts from game theory to present a framework for the analysis of situations in which limited resources must be efficiently deployed to protect some region from attacks that may arise at any point. We discuss how the mathematical techniques described may be applied to real-world scenarios such as the deployment of police to protect retail centres from rioters and the positioning of patrol ships to defend shipping lanes against pirate attacks.

We promote Colonel Blotto to Field Marshal in the game of that name by adding space more effectively.

Graph-based analyses, focusing on the topological structure of networks, provide crucial insight into the kinds of activities that occur within them. Studies of small-scale spatial networks have demonstrated conclusive predictive capacity with respect to social and economic factors. The relationship between multiple networks at a global scale, and the effect of one on the structure of another, is discussed in Part 10. It presents models of two different but mutually interdependent kinds of networks at an international scale. The first is a global analysis of the structure of international transportation, including roads, but also shipping, train and related networks, which we intend to be the first such study at this scale. We assess the applicability of centrality measures to graphs of this scale by discussing the comparison between measures that include geometrical properties of the network in space, with strictly topological measures often used in communications networks. The second is the economic structure of national industry and trade, as expressed in recorded input–output structure for individual countries. Flows expressed in these represent a non-spatial or trans-spatial network that can be interpreted by similar measures, to understand both comparative differences and similarities between nations, and also a larger picture of economic activity. The two networks will be analysed as a coupled system of both physical goods through space and non-spatial economic transactions. Due to the relative stability of these networks over time, their use as a background for modelling activity makes them useful as a predictive tool. Visualisations of this network will indicate points most susceptible to shock from economic or physical events, or areas with the potential for greatest impact from investment or aid.

In the concluding chapter which makes up Part 11, we review the progress made in geo-mathematical modelling and discuss ongoing research priorities.