Lecture Notes in Morphogenesis

Series Editor: Alessandro Sarti

Denise Pumain
Romain Reuillon

Urban Dynamics and Simulation Models

Springer

Lecture Notes in Morphogenesis

Series editor

Alessandro Sarti, CAMS Center for Mathematics, CNRS-EHESS, Paris, France
e-mail: alessandro.sarti@ehess.fr

More information about this series athttp://www.springer.com/series/11247

Denise Pumain · Romain Reuillon

Urban Dynamics and Simulation Models

Paul Chapron · Guillaume Chérel
Clémentine Cottineau · Robin Cura
Mathieu Leclaire · Sébastien Rey-Coyrehourcq
Clara Schmitt · Elfie Swerts

Springer

Denise Pumain
Université Paris I Panthéon-Sorbonne
and UMR Géographie-Cités
Paris
France

Romain Reuillon
CNRS, UMR Géographie-Cités
and Institut des Systèmes Complexes
Paris-Ile-de-France
Paris
France

ISSN 2195-1934 ISSN 2195-1942 (electronic)
Lecture Notes in Morphogenesis
ISBN 978-3-319-46495-4 ISBN 978-3-319-46497-8 (eBook)
DOI 10.1007/978-3-319-46497-8

Library of Congress Control Number: 2016956834

© Springer International Publishing AG 2017
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

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

The original version of the bookfrontmatter was revised: For detailed information please see erratum. The erratum to this chapter is available at DOI 10.1007/978-3-319-46497-8_7

Acknowledgements

That research was supported by the advanced grant GeoDiverCity of the European Research Council.

Contents

Authors and Contributors to the Book

Paul Chapron, Ph.D., in Computer Sciences, Post-doctorate at Institut des Systèmes Complexes Paris-Ile-de-France
Guillaume Chérel, Ph.D., in Computer Sciences, Post-doctorate at Inria, France
Clémentine Cottineau, Ph.D., in Geography (University Paris I), Research Associate, Urban Dynamics Lab, Centre for Advanced Spatial Analysis, University College London, UK
Robin Cura, Ph.D., Student in GIScience, CNRS, UMR Géographie-cités, Paris, France
Mathieu Leclaire Research Engineer, Institut des Systèmes Complexes Paris-Ile-de-France
Denise Pumain Professor in Geography, University Paris I Pantheon-Sorbonne, CNRS UMR Géographie-cités, Paris, France
Romain Reuillon Researcher, CNRS, UMR Géographie-cités and Institut des Systèmes Complexes Paris-Ile-de-France
Sébastien Rey-Coyrehourcq, Ph.D., in Geomatics (University Paris I), Research Engineer at University of Rouen, France
Clara Schmitt, Ph.D., in Geography (University Paris I), Seattle Washington, USA
Elfie Swerts, Ph.D., in Geography (University Paris I), Post-doctorate at University of Lausanne, Switzerland

Introduction

Since the turn of twenty-first century, half of the population of the world lives in cities. This is by no means a revolution or a decisive inflexion as it was sometimes suggested by the media, because the world trend in increasing the share of urban population (that ratio is technically called the 'urbanization rate') has been a continuous one for at least two centuries. The major significance of crossing the symbolic 50% threshold is that cities have become the places where the majority of people will henceforth make their living and invent their ways of life. This has already been the case for 50 years for the population of the more developed countries. It is becoming a reality for more and more inhabitants of the world. Before the end of this century more than three quarters of some nine billion humans will live in cities. The consequence is that most societal challenges have now to be confronted and solved within urban environments. Whatever the topic: transportation, pollution, production, consumption, social segregation, poverty, culture, architecture, and esthetics, it must now be thought of as an urban problem, to be managed in an urban milieu and solutions have to be adapted to the specific places.

Although cities are obviously part of many problems (climate change, inequality, etc.), they also represent the only viable solution to such problems (Carter et al. 2015), and they already have started solving them. Indeed, although they were never coordinated by any dedicated institution in a global and continuous way, cities have self-organized throughout the course of history into *systems of cities*.

Because they are connected by multiple networks for exchanging material goods as well as immaterial information,*systems of cities* could figure prominently among the most ingenious invention of human societies. If one would conceive them as an institution produced intentionally, systems of cities would be thought of as successful tools designed for reducing uncertainties of terrestrial environments using distant resources and multiplying innovation from network interactions. From these repeated and increased exchanges, the systems of cities have spontaneously generated significant co-evolution processes between more and more distant cities in expanding political, trade and cultural networks.

This evolution is not linear and involves many discontinuities in space and time: unequal urban development is a rule and barriers to trade and new borders have been erected many times. Neither was it always a virtuous process since most of it has occurred in terms of conflicts, predations, conquest, or competition for domination. Nevertheless their organization in systems can help cities facing the challenges of this century in terms of supplying goods and well-being for a growing population whilst preserving natural resources. Their multiple connections in systems of cities could nowadays be activated for sharing from the bottom up the good practices that are invented locally every day as well as for disseminating from the top down incentives or regulations that are decided by intergovernmental institutions.

Part of this activation could be done in a more conscious and scientific way by using models of urban dynamics. Systems of cities share with other complex systems intrinsic dynamic properties that were developed over history through inter-urban interaction. A well-known one is the shape of the statistical distribution of city sizes, coined as 'rank-size rule' by Zipf (1941) or the lognormal distribution by Gibrat (1931).

This hierarchical structure, its change over time and the dynamics generating it can be modelled, and we suggest here some new methods and a series of application cases. The methods are new because they integrate the latest technological power of distributed computing, allowing for a multiplication of large-scale simulations that was not possible before. The models are new in that they consider and integrate the uncertainties of urban evolutions (past, present or future), as well as the uncertainties of scientific explanations. Considering and modelling the uncertainties of urban complex dynamics allows to draw a fuzzier picture about urbanization than the one exposed in the media, but also one that is concerned with evaluating the probability of different scenarios and the adaptative properties of cities in various contexts.

Our models both encapsulate a major part of the accumulated knowledge from comparative empirical urban studies and ensure a significant progress in simulation models aiming at reconstructing urban dynamics by greatly improving their validation methods. We have tested the ability of these models to simulate past evolution with a growing confidence that they could as well help predicting possible urban future. That is why we think that our methodology could be used in further work to design policies for enhancing our territorial intelligence.

Why Model Cities?

Besides the traditional reasons for modelling complex systems: explain, illuminate core dynamics, bound outcomes to plausible ranges, illuminate core uncertainties, challenge the robustness of prevailing theory, Expose prevailing wisdom as incompatible with available data and Reveal the apparently simple (complex) to be complex (simple))… suggested by Epstein (2008), we see at least four reasons for developing urban models.

1. Because urban places concentrate the major political, economic and environmental problems, as 80% of the world population will be urban before the end

of the century. Any piece of knowledge and tools to produce it is then appreciated to try and understand, explain and forecast future changes, identify challenges and think about problem solving at large scales;
2. Because of our observation that cities never grow in isolation but always through interaction and co-evolution with other cities, we also know about the important consequences of these interaction processes on structures and dynamics that are shared by so many urban systems and be summarized and hierachized as characteristic stylized facts. Considering cities with their interactions therefore seems to be the right approach but it hampers our ability to understand the complex consequences of changes with our human brains. Therefore tractable computer models are necessary for keeping track of numerous interactions;
3. Since systems of cities are complex systems based on nonlinear interactions, simulation models are necessary for exploring the capabilities of regulation in an until now widely self-organized dynamics, especially for building plausible scenarios of urban evolution (Pumain and Sanders 2013);
4. Because models could then become the right tools for providing insights into our capacity to shape scenarios for future evolution. The challenge of monitoring these systems for better resilience and sustainability could be to imagine a possible scenario to shift from rivalry and competition to cooperation and creative emulation in inter-urban interaction.

As a consequence of the complexity of urban dynamics and of the emergency to face the multiple challenges of our world in terms of ecological and numerical transitions, efficient and adapted models of urban dynamics are more than ever needed to explore the possible directions of urban changes. While defining a future agenda for urban research, Batty (2013) insisted on the role of new technologies as tools for interpreting and adapting urban changes: “There is little doubt that getting a grip of what our future cities will be like must be based on the information and communication technologies and urban analytics that make the greatest use of the data and models that use these technologies to study the impact of the same technologies which are changing and transforming the very system of interest that we are studying. The fact that we use these technologies to explore how these very same technologies are changing the system we are studying is part of the paradox that a future cities agenda must grapple with”.

Originality of the Book

The book has an original approach in many ways. First, it provides a broad view about urbanism, exemplifying cities geographical distribution, hierarchical and functional differentiation, structural and dynamic features rather than apparent differences in urban landscapes and morphologies that are usually put forward in the descriptions of urbanization.

Second, it uses lessons from the past to explore the future of cities, with respect to historical contingency and path dependency. We do not rely on equilibrium theories or on hypotheses about optimal urban organization. On the contrary we promote an evolutionary perspective for urban systems.

Third, we remain rooted in empirical observations filtered through a comparative method by dedicating a special care for building standardized data from the diversity of existing official definitions and statistics in different regions of the world. A fourth originality of this book is aiming at developing open reproducible tools for simulation models stemming from the ‘Simpop family’ (Pumain 2012).

And last but not least, an innovative feature is that we develop an integrative view on digital social science, since the authors are part of a multidisciplinary team including computer scientists and geographers who worked together on the same material and with the same tools they built during the 5 years of the GeoDiverCity project.

Complex Cities and Complex Models

The system we study—cities interacting with one another at national and continental scales—is complex. A model (or models) can be used to abstract some of the dynamics and features of the target system, but given that it still allows interactions, nonlinearities and stochasticity, the model is itself a complex object (Amblard and Phan 2007). Therefore it can lead to a diversity of possible trajectories originating from the same simulation starting point.

This property of complex models makes them interesting to represent aspects of the urban world and to identify a range of possible outcomes, but it also challenges the kind of knowledge one can draw from the simulation results. A thorough use of the computer model for explanation, understanding and forecasting thus needs to include and acknowledge this diversity, uncertainty and contingency (the observed trajectory is just ‘one of’ the multiple realizations that were possible at some point).

It is made possible by replicating the experiment in a virtual laboratory, changing initial conditions and repeating simulations, which is impossible ‘in real life’. The function of virtual laboratory is therefore a key advantage of simulation models in urban studies. It is however subject to an uncertainty regarding the validity of the rules and mechanisms modelled, since a large variety of processes can result in similar outcomes. This challenge is known as equifinality (von Bertalanffy 1968), and it is particularly strong in urban modelling at large spatial and temporal scales, because the data we rely on to characterize the system and validate micro-behaviour are rather sparse and limited. Equifinality means that a larger pool of processes to be candidates for the explanation. We think it also requires urban modellers to integrate this issue all along the evaluation protocol as it will determine the quality of knowledge to be extracted from the models for urban prospective and understanding (O’Sullivan 2004; Batty and Torrens 2005).

Indeed, at the scale of systems of cities, we can observe urbanization patterns via historical census data (hierarchical organization, spacing, specialization, etc.) but rarely processes themselves or the actual inter-urban flows. Simulation models constitute interesting tools to try and 'replay history', to see if it corresponds to the observed patterns, and what other situation could have resulted from the same initial conditions. However, there are two levels of adequacy assessments (Rossiter et al. 2010) to check:

1. What other mechanisms are able to produce the same outcome (within the model)
2. Are the processes modelled the ones that happened 'in real life'?

The first question can be addressed by an extensive inclusion of several theories through a multi-modelling approach. The second question remains and refers to broader causality issues in generative explanation (Hedström and Ylikoski 2010; Elsenbroich 2012). If indeed there is no definite way to identify a causal mechanism modelled as the cause in systems of cities, generative mechanisms provide a causal link between interacting entities and macro-patterns. If micro-behaviours are proven robust by empirical and theoretical evidences, it provides an important complementary explanation to usual statistical explanations (Byrne 1998; Goldthorpe 2001), from which causal links are missing.

Book Proposition: Exploring Multiple Parsimonious Models

The answer we provide to the challenge of equifinality in urban simulation is threefold:

1. *Parsimony*. Given the multiple accounts potentially responsible for the same urban outcome, we state as a principle to not start with the most complicated and detailed model, but instead to evaluate the ability of simple mechanisms, and then to combine them into a more complex model. This way, we can better trace the complex effects of each mechanism and theoretical hypothesis, and characterize its necessity for the reproduction of a feature within the model.
2. *Multiplicity*. This principle applies to the causes as well as to the trajectories of urban evolution. Because of the complexity of urban interactions and because of their contingencies, we think that a thorough investigation of cities needs to include multiple accounts of urban growth and interactions (competition and cooperation for example) as well as to look for the diversity of possible trajectories of a system could have reach given random processes and small perturbations.
3. *Extensive Exploration/Evaluation*. The combination of parsimonious modules into models enables to build a large incremental exploration of models, to compare their structure, their parameterization and their ability to simulate past trends as well as diverse (credible) alternatives. Indeed, we think that it is necessary to take advantage of the full extent of working in a virtual

(computational) laboratory. We have developed several generic and open tools and methods to do so. We present at the end of the book the integrated platform in which they take place and how they can be reused in a large variety of simulation contexts.

Book Content

The new generation of Simpop models presented in this book, as well as the approach followed all along their construction and evaluation, has specifically targeted and tackled the equifinality challenge of (urban) modelling. It is presented as a progression from solving elementary generative problems to adapting models for encompassing a variety of urban situations by sharing open tools.

The first chapter presents our empirical knowledge of systems of cities, and ways of summarizing their regular properties. It builds the 'system of reference' upon which model-building can take place. Indeed, by generalizing processes and structural properties of empirical case studies in different spatio-temporal contexts, it specifies the elements that can be forecasted (the total urban growth, the degree of differentiation of city sizes, the spatial balance of growth, etc.) and ways to do it (theories of innovation diffusion, of agglomeration economies, of spatial distribution, etc.). By asking 'is urban future predictable', we question the logics of urban evolution as well as the different levels of uncertainty attached to different aspects of urban growth and interactions. Identifying the possibility of prediction makes the task of modelling interesting. Identifying key features of systems of cities provides a stylized empirical ground to evaluate simulations and study alternative trajectories. Finally, identifying areas of uncertainty as leading to the processes responsible for the urban evolution calls for a multi-modelling approach that tackles equifinality in the virtual laboratory.

The second chapter addresses the first step of the modelling of system of cities. It presents a parsimonious model of the emergence of cities from a homogenous settlement system. It aims to answer a very basic question: are we able to identify a simple set of meaningful mechanisms that reproduces the observed emergence of cities at the scale of thousands of years? The SimpopLocal model is an answer to this question and it raises the challenge of calibration, in order to prove that there exists a set of parameters that are sufficient to model this emergence. It also raises the challenge of formalizing what a good simulation is in terms of long-term urban evolution, in order to automate the search for this parameter set (for which there exists no empirical ways of determination).

The third chapter goes beyond the possibility of finding one way of simulating the emergence of cities. It presents a new method for assessing parameter sensitivity, by looking at the necessity of each mechanism within a given model structure. Indeed, despite the diversity of solutions to the calibration challenge, are some parameters isolated, not interacting with other parameters in the simulated output? Are they all necessary, besides being sufficient? A new method called 'calibration profiling' was developed to validate not only sufficiency of modelled mechanisms but also the necessity of theoretical hypotheses that are behind the construction of the model. It is a progress of social sciences towards the scientific methods (all things being equal), and it allows to increase the parsimony of urban models.

The fourth chapter builds on this quest for parsimony, as it presents an incremental model-building approach to simulate empirical systems of cities. Given the specificity of the system we aim to model, we expect the mechanisms needed to reproduce the observed trajectory to be multiple and interacting in a complex way. Therefore, we have built a framework of hypothesis-testing and implemented modules of mechanisms that we combine and simulate. The combination follows a path of complexification as well as particularisation from any system of cities to a specific case study. The quality of each simulation is evaluated with respect to the populations observed in the corresponding empirical cities. This approach was developed to model the evolution of Soviet and post-Soviet cities from the 1960s on. Its strength is to be transferable at a very low cost to any other national system. A tentative check was performed on Indian cities. We finally show in this chapter that a theoretical-based modular model allows to evaluate and compare the power of different hypotheses to explain urban growth at different periods of time and in different geographical contexts, and therefore suggests a way to account for equifinality in urban models.

The fifth chapter corresponds to an innovative way of exploring simulation models, and especially urban models. It considers a parsimonious structure of mechanisms and looks for the diversity of possible outcomes that the model can reach within a reasonable range of parameters. This means that it explores what the trajectory of a system of cities *could have been*, if we simulate past trends, or *what it could be* in the future, in terms of two or three properties of the system (like its total population, or the degree of inequality of city sizes). We present the algorithm developed to maximize the diversity of a model's output, as well as the kind of knowledge it leads to in an empirical context. For instance, we analyze the alternative pasts of the Soviet system of cities (as modelled within different model structures) and the corresponding parameters and their meaning. In particular, we highlight configurations that result in population growth and configurations that result in population shrinkage, configurations that result in hierarchization or in the equalization of city sizes for each of the demographic regimes at two periods of time (Soviet and post-Soviet eras).

In the last chapter, we present the platform that brings together and enables all the cutting-edge exploration methods in urban simulation. This integrated, innovative and open toolbox for urban modelling is called OpenMOLE.

As an epilogue, we present what could be a world atlas of urban models for global prospective on urban future. We also stress the challenges that hamper its construction so far, especially because of the data challenge that is comparing cities over time and over space. Indeed, each country having (or having not) developed its own way of defining cities and quantifying urban features, there remains a monumental amount of work to collect and harmonize urban data over large period of time, as well as to identify what in each national evolution relates to generic and specific processes. Cumulative modelling could help perform this task, or at least to highlight areas of uncertainties. Our guess is that it will only be achieved by a large collective and interdisciplinary collaboration (between urban and regional specialists, modellers, computer scientists, empirical and theoretical experts, data providers and data analysts) based on open practices (as to data, methods and models).

References

Batty, M.: The New Science of Cities. MIT Press, Massachusetts (2013)
Batty, M., Torrens, P.M.: Modelling and prediction in a complex world. Futures **37**(7), 745–766 (2005)
Byrne, D.S.: Complexity Theory and the Social Sciences: An Introduction. Psychology Press, London (1998)
Carter, A., Tonkiss, F., Gardiner, L.: City talks: Is urban inequality inevitable? Joining Andrew to discuss these questions are Laura Gardiner, Senior Research and Policy Analyst at the Resolution Foundation, and Prof Fran Tonkiss, Director of the Cities Programme at the London School of Economics (2015). URL accessed in october 2016
Elsenbroich, C.: Explanation in agent-based modelling: functions, causality or mechanisms? J. Artif. Soc. Soc. Simulat. **15**(3), 1 (2012)
Epstein, J.M.: Why model? J. Artif. Soc. Soc. Simulat. **11**(4), 12 (2008)
Gibrat, R.: Les inégalités économiques. Librairie du Recueil Sirey, Paris (1931)
Goldthorpe, J.H.: Causation, statistics, and sociology. Eur. Sociol. Rev. **17**(1), 1–20 (2001)
Hedström, P., Ylikoski, P.: Causal mechanisms in the social sciences. Annu. Rev. Sociol. **36**(1), 49–67 (2010)
O'Sullivan, D.: Complexity science and human geography. Trans. Inst. Br. Geogr. **29**(3), 282–295 (2004)
Phan, D., Amblard, F. (eds.): Agent-based Modelling and Simulation in the Social and Human Sciences. The Bardwell Press (2007)
Pumain, D.: Multi-agents system modelling for urban systems: the series of simpop models. In: Heppenstall, A.J., Crooks, A.T., S.L.B.M. (eds.) Agent-based models of geographical systems, pp. 721–738. Springer (2012)
Pumain, D., Sanders, L.: Theoretical principles in inter-urban simulation models: a comparison. Environ. Plan. A 1–43 (2013)
Rossiter, S., Noble, J., Bell, K.: Social simulations: improving interdisciplinary understanding of scientific positioning and validity. Journal of Artificial Societies and Social Simulation **13**(1) (2010)
von Bertalanffy, L.: International library of systems theory and philosophy. General System Theory: Foundations, Development, Applications, G. Braziller (1968)
Zipf, G.K.: National Unity and Disunity. The Principia Press, inc., USA (1941)

Chapter 1
Is Urban Future Predictable?

Abstract Despite uncertainties linked to the increasing speed of technological and societal evolution, important features of future urbanism can be predicted at the regional and global levels and even sometimes for local situations. Comparative urban studies have brought results about universal processes and typical trajectories in the history of urban systems. This analytic description provides the basis for designing robust dynamic models as well as realistic scenarios for exploring a diversity of possible urban futures.

Introduction: Systems of Cities as Adaptive Complex Systems

Cities are places of the world where the majority of human beings are living now and where the largest amounts of physical and societal wealth, skills, and values of humankind are concentrated. Much is expected by many stakeholders about predicting what can be expected from their future evolution. A recent report by the World Bank (2009) underlines the essential role of cities in economic development and technological and social innovations, for the first time recognizing the value of urban concentrations, even in poor countries. The World Bank suggests that interventions be targeted according to the type of city, via a regional, hierarchical typology: metropolitan areas (areas of advanced urbanization) are the most liable to make use of productive investments; intermediate or small cities (intermediate urbanization) and densely populated 'lagging areas' diffuse this growth towards rural areas, while the incipient urbanization in sparsely populated lagging areas means that they do not draw much benefit from the process. Hence this report sets out to explore the geographical diversity of cities with respect to size and regional density, the effects of concentration on city growth and on the ability of cities to diffuse these effects towards their hinterlands.

While facing tremendous challenges from the perspective of environmental change, economic competition, disruptive technological innovation or political conflicts, the resilience of cities is often questioned. The originality of our approach is to assess predictions about possible futures for cities from the knowledge that was constructed about their past dynamics (Pred 1977; Pumain et al. 2015). The knowledge we present here is new because it is inspired not by the observation of isolated cities but by the lessons learnt from their relative situation in systems of cities and

© Springer International Publishing AG 2017

D. Pumain and R. Reuillon, *Urban Dynamics and Simulation Models*,
Lecture Notes in Morphogenesis, DOI 10.1007/978-3-319-46497-8_1

their interaction with other cities within very large sets (of cities) over long time periods. Our major assumption is that cities become more and more interdependent during their evolution from the many exchanges they have with other cities and that much can be learned from that co-evolution within systems of cities which become a strong constraint on the future of individual cities.

Such knowledge is reliable because it is rooted in comparative empirical analysis of huge datasets retracing the evolution of many urban indicators for a very large number of cities over long periods of time in a variety of countries in the world. When using classical statistical methods for international comparisons, we construct a relevant geographical ontology for cities, identifying urban units which retain significance over time, and which can be compared from one country to the next. We choose to define urban units according to concepts that are rooted in geographic theory, including reference to individual space-time budgets and taking into account the variable forms of urbanization and territorial organization in different countries: this means defining and delineating urban entities according to a geo-historical concept. We thus maintain a theoretical approach in our conception of modelling in collaboration with mathematicians, physicists and computer scientists, leading to models of complex systems that have relevance for social sciences. That is why we can design simulation models that reconstruct the most salient features of the dynamics of systems of cities for exploring plausible scenarios of their future evolution.

Cities are complex systems which are embedded in multiple networks conveying people, goods and information (Rozenblat and Melançon 2013). These networks are not randomly distributed on the surface of the earth, the spatial organization of the flows they are conveying follows patterns that are widely governed by the existing relative location of urban concentrations and reinforces them. Indeed, there are positive feedbacks between cities and all kind of flows (goods, people, information...) sustaining them. Spatial interactions in societies shape geographical organization at two main levels: most repeated interactions of daily life occur within cities while the less frequent but recurrent interactions shape regional sub-networks of strongly interdependent cities (Pumain 2006). That two-level organization was coined by Brian Berry in a famous expression: 'cities as systems within systems of cities' (Berry 1964). Our conception is not unlike that view but instead of imagining a nested hierarchy it puts the emphasis on the complexity of this organization, on the new dynamic properties which emerge in the course of the formation of systems of cities, and on the effect of these properties on the itinerary of each individual city.

One distinctive emerging property of this spatial organization is its sustainability, because cities and systems of cities are adaptive. Adaptation means more- or less-coordinated changes in every aspect of social life due to the assimilation by many stakeholders of all kinds of innovations, may they be technological, organizational or cultural (Lane et al. 2009). It is because goods, people and information circulate that, all over the planet, cities are growing and adapting to the societal, technological, economic and cultural changes that they continuously generate, and they become more and more the reference for organizing the life of societies in geographical space. At a micro-level, a city's adaptive behaviour results from its own history, specific relationship with its surrounding environment, and individual and institutional stakeholders' decisions. But at a macro-level, because of their mutual dependencies and competition, cities' evolutions also result from complex interactions with other cities, within multiple networks cities are involved in. Even if these interactions are surely less frequent than those that occur in daily life, their recurrence shape systems of strongly interdependent cities and in the long run, the multiplicity of interurban exchanges and interactions results in the co-evolution of urban systems.

As the nature and scope of interactions between cities vary according to city size and the spatial scale considered, 'systems of cities' are not easily identified and delineated. Systems of cities usually are observed within national boundaries which constrain some of the most frequent societal practices through a common language, a sovereign regulation, a collective identity imagination and a cultural habitus, but the interaction space of some cities, either because of their size or of their specialization, can expand well beyond national borders or be organized in international specialized networks whose size is more or less extensive, for instance in the case of transnational firms (Rozenblat and Pumain 2007). That is why spatial and temporal contexts always have to be carefully analyzed when trying to categorize dynamic processes of urban systems.

Because of the interconnectedness that ensures the interdependence in their evolution, there are common features in the dynamics of systems of cities which can be considered as universal all over the world and that share non-trivial similarities with observed structures in other complex systems (Pumain 2006; Batty 2013). From a comparative statistical analysis of the demographic, economic and physical transformation of cities in many regions of the world we have identified the major invariants in urban systems dynamics as 'stylized facts' that are summarized in our evolutionary theory of urban systems (Pumain 1997) and can be implemented in simulation models.

Despite recent papers still attempting to validate theories of economic convergence in systems of cities (Chauvin et al. 2016), complex systems have spread through economic theory with the new economic geography (Krugman 1991) or evolutionist economic geography (Boschma and Frenken 2006). It is now possible to move beyond the idea that urban systems are in a state of equilibrium, or that they tend towards optimization. Theoretical consideration can be based on the historical and open nature of urban dynamics, and on the heterogeneity of agents at work in the urban environment, whatever the geographical scale. We conceptualize the dynamics of systems of cities as an open evolution that emerges from spatial and societal interactions between cities without considering that it should meet any equilibrium or optimization pattern. Multi-agent models are especially well designed for simulating how a structure emerge and is maintained at the macro-level from the many interactions between a variety of agents at micro-level (Ferber and Perrot 1995). We also prefer using them because spatial network dynamics are still difficult to represent in mathematical models of differential equations (Pumain and Sanders 2013). Multi-agent models allow testing theories by varying the rules and parameters of the simulation model, paving the way for a hypothetico-deductive approach. They are part of constituting a virtual laboratory which is the only possible substitute to scientific experiment in social sciences.

We consider that simulation models are very good tools to test elements of urban theory, which can thus be articulated and hierarchised within a single model. First the model is assessed in its ability to reproduce past evolutions with respect to their main components and geographical expressions. This modelling process does not set out to be realistic at all costs, it is not intended to accurately reproduce local observations, but nor is it on the other hand purely 'theoretical' in the mathematical meaning of the word. The rules and parameters, even if they correspond to abstract categories, are always linked to processes and magnitudes which can be measured, so that the model is calibrated with observations. It is this attention to plausibility as well as the importance of path dependence in urban dynamics which enables us to consider the use of these models not only as tools for reconstructing past dynamics, but also as tools for exploring future scenarios. Comparative urban studies have brought results about universal processes and typical trajectories in the history of urban systems. We rely on these common features and universal evolution trends to construct plausible broad patterns of cities behaviour at a global or regional scale, according to various scenarios.

The first set of stylized facts we examine about systems of cities is linked to their emergence which occurred some 10,000 years ago and created a new way of managing the spatial relationships of societies and their resources with their environment. These facts are the basis of the SimpopLocal model we implement in Chap. 2. We shall then summarize a second set of stylized facts that describes the ordinary dynamics of systems of cities once they are constituted. The simulation models that were implemented to represent these facts will be detailed in Chap. 4. The third set of stylized facts is about which major specific features can be retained for adapting our models to the variety of situations we find in different parts of the world. We shall use these observations to adapt our models to real-word situations as explained in Chap. 4.

1.1 Emergence

We are interested in the emergence of cities not so much for the sake of reconstructing that history but mainly because it implies identifying precisely the processes that distinguish the functioning of urban systems from previous forms of settlement systems. In the history of societies, the appearance of cities marks an essential step leading to an unprecedented multiplication of wealth and human populations, the modification of their way of inhabiting the earth and managing resources. We thus select, in describing this event, the major features that explain the urban development through fundamental dynamic processes, trying to provide an abstract framework for preparing a parsimonious model.

Urbanization in many respects is a universal process, temporally and spatially. Temporally, it has emerged independently in different parts of the world, at different moments in historical times since 10,000 BP but always a few thousand years after the invention of agriculture in the few large regions where it was maintained

(Bairoch 1985). Mesopotamia seems to be the earliest place of emergence followed by Indus valley, South China and later on Meso-America. Due to its connection with the previous invention of agriculture (the so-called 'neolithic revolution'), the emergence of cities was located in regions of subtropical climate in ecosystems where plant and animal species were sufficiently diversified and exploitable (Diamond 1997). Cities appeared when the social organization which enabled a bifurcation towards sedentary productive activities had lived a time long enough for multiplying their densities by a factor of about 100, compared with previous nomadic societies of hunter-gatherers (Marcus and Sabloff 2008). In the mean time, this long process lead to the accumulation of an economic surplus which enabled a societal division of labour. However for many centuries, this process more or less stagnated at a maximum of 10% of the population engaged in activities other than merely food production (Bairoch 1985).

The spontaneous emergence of systems of cities and their success as an innovation that would diffuse after many centuries all over the world can be interpreted in functionalist terms. To face survival and ecological challenges, cities appear as an ingenious societal invention which enables the reduction of the uncertainties linked to resources of their local environment, by expanding connections towards other cities that can provide complementary resources. Indeed, in all regions, cities did not emerge as isolated entities but always as interconnected systems of cities, diffusing and sharing their innovations as well as competing in a permanent rivalry for enlarging their stock of accumulated and potential resources. As a result (according to the demonstration we recall below), these precocious systems of cities were already characterized with a very asymmetric distribution of city sizes following a Zipf's law (Fletcher 1986). The stylized facts that we retain in modelling the emergence of cities are indeed the basic elementary mechanisms of the ordinary functioning of systems of cities. They include the development of innovations that were induced by increasing interactions between people within and between cities, as well as the spatial diffusion of such innovations from place to place, and their positive retroaction on further population concentration and urban development. Unfortunately for that remote period we do not have precise empirical data which would enable calibrating a model. That is why we choose to develop an abstract model using theoretical knowledge about spatial interaction (following a gravity rule), urban growth in systems of cities (see section below) and empirical average orders of magnitude for the maximum city size, population growth rates (including implicitly the effect of innovation on resource and population growth) and duration of the emergence period (Schmitt 2014). The model that is described in detail in Chap. 2 is an attempt at establishing generic conditions for the emergence and maintenance of systems of cities. The only stylized fact on the simulated system of cities that we use as a constraint for calibration is the generation of an urban hierarchy following a Zipf' s law or lognormal distribution, as explained in the next section.

1.2 Generic Dynamic Features of Systems of Cities

Once systems of cities have emerged, their dynamics can be summarized by a few regularities that appear everywhere, whatever the political organization or the economic regime of the societies or the period of their development. Such dynamics are characterized by: - the hierarchical differentiation of city sizes; - a relative meta-stability of urban hierarchies over several decades and sometimes centuries; - a regular quasi-stochastic process of growth sharing between cities due to their interactions that explains both the hierarchical distribution and its meta-stability; - a recurrent process of exogenous shocks by radical innovation which contributes to reinforce the hierarchical structure and may introduce from time to time a qualitative differentiation leading to the specialization of subsets of cities in new roles in the urban system.

1.2.1 The Hierarchical Differentiation of City Sizes

One of the main universal characteristic of systems of cities observed in regional or national territories is their organization in urban hierarchies. Nowadays there are four orders of magnitude between the sizes of the smallest and the largest entities called 'cities' in most of the large countries in the world: small towns cluster a few thousand people (10^3) whereas the gigantic megacities concentrate more than 10^7 and as many as 30 million and more for the largest, Tokyo for instance. Whatever the region of the world the number of cities follows an inverse geometric progression of their size. Table 1.1 gives the approximate number of urban entities in the world (urban agglomerations, i.e. defined by contiguous built-up area) at two dates. To avoid misinterpretation due to the spatial aggregation of cities that would not be engaged in strong interactions, epistemological care must prevail. Indeed, when simulating the evolution of systems of cities, even if a possible benchmark for validating the model is that the final distribution of city sizes is Pareto or lognormal, one essential preliminary assumption to confirm its validity is to ensure that the selected cities within the system are really strongly interacting during the simulated historical period. This strong regularity is observed since the emergence of systems of cities and whatever the political, economic and cultural systems in which cities evolve.

Table 1.1 Number of urban agglomerations in the world according to their size (Source: F. Moriconi-Ebrard (Geopolis) and Population.net)

Number of inhabitants	1950	2010
10^7	2	39
10^6	83	526
10^5	1050	5100
10^4	10800	59000

This hierarchical distribution of city sizes was formalized during the twentieth century first by Felix Auerbach (Auerbach 1913) who noticed the rather constant product between the size of a city and its rank in urban hierarchy and Robert Gibrat (Gibrat 1931) who applied his 'law of proportional effect' (added growth during a short period is proportional to initial city size and randomly redistributed at each time interval) for generating lognormal distributions of city sizes, before the statistician George Kingsley Zipf imposed the Paretian model of the 'rank-size rule' (Zipf et al. 1941) that is mostly satisfying for its descriptive visualization and the slope of the adjusted line being a convenient comparable index of the inequality of city sizes (Pumain et al. 2012). A huge literature about which statistical model would be the best fit has been published leading to significant methodological improvements but very often the results are contradictory and remain uncertain (Nitsch 2005; Favaro and Pumain 2011) mainly because not enough attention is paid to the quality of empirical data (including the consistency of urban definitions over time, as well as the effect of variable number and size threshold of cities being considered (Pumain et al. 2015; Cottineau 2016).

Besides these statistical descriptions summarizing this 'heavy tailed' asymmetric distribution, geographical explanations of the number of cities and towns according to their size were provided the central place theory built by Walter Christaller (Christaller 1933). According to this theory, cities develop through the principle of providing services to a regional population under the constraint of proximity. As a consequence, the bigger the city, the more it provides rare goods and services, to a larger and more distant population. Cities of similar size thus have a regular distance between them, and the population in their sphere of influence is proportional to their size. The corresponding spatial organization is a nested urban hierarchy. Although Christaller mentioned in a perspective that the number of centres should be reduced when the improved transportation systems could enlarge the range of the services provided by the largest cities in the urban hierarchy, this observation was not formalized in the theory and other stylized facts regarding the evolution of urban hierarchies were added later (see below Sect. 1.2.4).

1.2.2 The Meta-Stability of Urban Hierarchies

Many centuries after the emergence of the first systems of cities, all urban regions in the world became connected by more or less regular exchanges and the local urban systems were developing mutual influences and complementarities. The sharp lowering of transportation cost associated to the Industrial Revolution at the end of eighteenth century and the accompanying immense increase in productivity triggered huge population migrations towards urban centres. This started the 'urban transition' (Zelinsky 1971) that has come to an end in the more developed countries since second half of twentieth century but is still operating in the emerging and poor economies. Half of the world population lived in cities at the turn of this century and more than three quarters are expected to become urban by 2050.

Despite the technological, economic and societal turmoil associated with the industrial revolution, the demographic transition and the strong acceleration of urban growth, the urban hierarchies keep a remarkable stability on the long run in the regions of the world where urbanization developed continuously such as Europe or Asia. For instance, the correlation coefficient between the ranks of cities in the European urban hierarchy reaches about 0.8 between the beginning of the eighteenth and the end of the twentieth century (Bretagnolle et al. 2000). Of course this stability only occurs in integrated systems where cities have been interconnected through recurrent exchanges of people, goods and information and evolve under the same set of societal rules. As a proof of this, when putting together all cities of the world and comparing the evolution of their rank, Batty (2006) finds much more contrasted trajectories on his 'rank clock' visualization tool. This was also rather the case in his paper about the demographic trajectories of 200 US cities from 1790 to 2000 (Batty 2003), because of the peculiarity of this system of cities which was expanding very recently compared to Europe and Asia and according to a colonial process that filled space along a moving frontier (Bretagnolle et al. 2008). As very few cities are nowadays 'created' from scratch (even in the past most of them were previously villages that expanded in a progressive way including many fluctuations in their development), and as sudden complete destruction of cities have become very rare after fifteenth century, there are rather strong 'path dependence' effects (Arthur 1994) in the current evolution of urban hierarchies. This can be better understood if one considers the general process of transmission and amplification of growth in systems of cities.

1.2.3 A Regular Quasi-stochastic Process of Growth

The hierarchical distribution of city sizes no longer remains a 'mystery' (Krugman 1996) if we consider that it is the outcome of a process of urban growth in which every city has a probability to grow at the same rate in each time interval (Gibrat 1931). This averaging of urban growth rates is explained in first approximation by the expectancy of natural growth to be constant in homogeneous populations at a given period in historical time and by the gravity like process of migrations when geographical space is integrated through regular interactions (Wilson 1970). On the very long run, due to stochastic fluctuations, these interactions differentiate accumulations of wealth and population in the cities according to highly dissymmetrical statistical distributions. Many tests of that theoretical growth process were conducted on longitudinal empirical data that confirm its ability to describe the interurban distribution of urban growth rates, at least as a first approximation (Gibrat 1931; Robson 1973; Pumain 1982; Pumain and Moriconi-Ebrard 1997; Pumain 2006; Bretagnolle et al. 2008; Pumain et al. 2015).

However many other attempts have lead to contradictory results about the accuracy of the model and it is difficult to determine whether these contradictory results are explained by the differences in testing methods, by the geographical and historical context of the country and period under analysis, or, more probably, by the way urban entities are defined in the empirical databases. In any case, even when cities are properly delineated according to a consistent geographical definition and when population growth is measured on long enough time periods (for instance see Robson 1973), it appears that the growth process in integrated systems of cities is stochastic in a first approximation only. 'Three anomalies compared to the stochastic rules of Gibrat's model have been identified: (i) A trend for a positive correlation between city size and urban growth ; (ii) a positive or negative correlation between successive growth rates at some periods of time, over several decades, indicating a persistency of growth impulses in the same locations or a reversal in growth locations; (iii) a heteroscedasticity of growth rates: large cities have smaller growth rates standard deviation than the smaller towns' (Favaro and Pumain 2011).

These observations have the advantage to question a fundamental assumption of the stochastic model, namely the statistical independence between urban units, which contradicts our concept of connected cities, each other informed of their changes and co-evolving through multiple networks linking them (Robson 1973; Pumain 1982; Hernando et al. 2015). Indeed, the 'random' distribution of growth rates rises from these connections that integrate cities in relations of complementarities and competition within a territory where common rules of political, economic and social functioning are shared. As the slight deviations to the stochastic model can be explained by the way innovation waves are propagated in systems of cities, taking them into account would help conciliating both views.

1.2.4 Hierarchical Diffusion of Innovation Waves and Functional Specializations

The functional and hierarchical organization of cities in systems is powered by the regular emergence of all kinds of societal innovations (Duranton and Puga 2001). According to the well-exemplified theory of hierarchical diffusion of innovations by Hägerstrand (1952), there is a higher probability that new technologies or cultural practices will be at first captured by the largest cities. Large cities have more risk capital ready to invest, a higher diversity and level of skills, and a better access to information. It is there that new things can develop and become adapted to societal uses inventing new functionalities to every new artefact and social practice (Lane et al. 2009). As a result, although large cities suffer higher costs (mainly through wages and rents) they take benefits from innovation at its highest point of return, and this explains why they receive the accompanying growth impulse earlier than the other cities of the system that soon will imitate the innovation. On the long run, the repeated incremental growth surplus linked to this delay in the hierarchical transmission of innovation leads to a (very) slight deviation from the stochastic model of growth inducing a small advantage to largest cities in the system. A first consequence is a trend towards growing inequalities in city sizes, exceeding what would be produced by a purely stochastic Gibrat's model.

The growth impulse associated with innovation is more spectacular when, for various reasons, (usually a localized resource necessary to the development of the new product or activity) it concentrates in smaller towns which experience a sudden urban growth. This leads to the specific development of specialized cities (cities of mining industry, of steel industry during the first industrial revolution, tourist cities since the 19^{th} century, as well as university towns, are the most frequent examples). Usually when the associated development remains located in small towns the specialization tends to hamper their further development because it reduces their ability to adapt to further innovation waves, but if the innovation wave brings large quantities of employment and profits in specific cities or regions it may lead to a partial reorganization of the spatial and hierarchical structure of the system of cities, for instance when British industrial cities emerged in nineteenth century, or when industrial revolution contrasted urban evolution in North-East and South-Western France, or when successive waves of innovation were developing the North-East, then the West and South, of the system of cities in the United States, (Bretagnolle et al. 2008) or when mining and oil industries boosted the development of Siberian new towns (Cottineau 2014). The structure of urban systems, understood as the hierarchical and spatial distribution of cities, thus results of social, functional and technical adaptation of cities to the anthropological developments of the territories in which they are rooted. The functional and hierarchical organization of the systems also evolves according to the progress in transportation and communication techniques and networks. That evolution transforms the interactions between cities, and the diffusion of the innovation waves within the systems.

1.3 Variety in the Evolution of Urban Systems

In a multi-scale perspective, it is of great importance to identify levels of system description that are significant, sufficiently abstract to enable comparison and modelling, but sufficiently detailed to return a plausible picture of geographical diversity. Computer modelling focuses even more attention on the need for 'ontological' clarification. The models we produce do have a generalization potential, illustrating a theory of systems of cities. But our approach differs from certain theories such as the New Urban Economy, or experiments generated from a universalistic physical research perspective (West, Bettencourt), or again research in the field of the dynamics of complex systems made up of sub-systems or homogenous agents: it differs in that it does not start from an object, 'the city', apprehended in its singleness. Our approach stipulates that the dynamics that lead to a co-evolution of cities within a system generates the diversity of the cities, and proceed from this diversity, according to a recurrent process throughout the history of the cities in question. It also assumes that while the systems of cities share common static and dynamic properties, they also present a diversity that can be related explicitly to certain conditions attendant upon their spatiotemporal development.

It is appropriate here to define what we mean by diversity. What is meant is 'geo-diversity', which corresponds to attributes that are expressed in spatial and territorial manner, on the levels of both the city and the system of cities. For each city, among numerous possible descriptors, and among those suitable for comparisons and quantitative modelling, we have chosen those generating the most marked differentiations between cities. From multivariate econometric analysis, it is recognized (Pumain 2006) that the main source of urban diversity is of hierarchical nature, because city size (measured in international comparisons either by the number of residents, in which case it varies from 10^3 to 10^7, or on the basis of wealth or the urban product) is strongly correlated with a whole set of indicators—indicators of the level and diversity of equipment and services, of position and outreach of territorial and functional networks, of quality of the human capital, of complexity of social organization and ability to capture innovation. The second important source of urban geo-diversity is the economic orientation, or functional specialization, which is only partially dependent on city size (smaller cities can have an important role, for instance arising from the presence of international institutions, or cutting-edge research). This depends to a large extent on local and regional production factors, and relates to concentrations of activity that have often been formed under the impetus of successive innovation cycles.

Each system of cities, considered for example within a national territory or a sufficiently large continental unit, possesses a hierarchical and functional diversity that corresponds to the relations of complementarities set up within its boundaries; these boundaries remaining relatively impermeable over periods of variable duration, often several centuries. An issue of general importance in complex systems is the evaluation of scaling relationships between the dimensions of the territories and those of the cities, and also the degree of functional specializations. The difficulty in this type of research is to establish equivalences between systems, and the problem of the unsuitability of the socioeconomic nomenclatures available. The permeability of boundaries is not independent from city size (the outreach of its relationships tends to increase with size), and varies over time along with city sizes, which adds a further difficulty which modelling experiments can help to overcome. Other attributes of cities, such as their type of governance, their architectural or cultural style, or their image, also are major sources of diversity, and their effects on city dynamics could be simulated by way of suitable scenarios (this is particularly true for governance). Although these qualitative attributes receive considerable interest in the literature among various disciplines (Panerai et al. 1997; Hannerz 1992; Polèse and Stren 2000; Scott 2002 to quote only a few) that type of urban diversity is not the one we shall focus on in this book.

We concentrate here on the evolution of the hierarchical structure of urban systems. That feature seems to be universal, regardless of the geographical environment, the economic system and sociopolitical organization in which they are rooted. However, beyond the common structures and dynamics, systems of cities in various parts

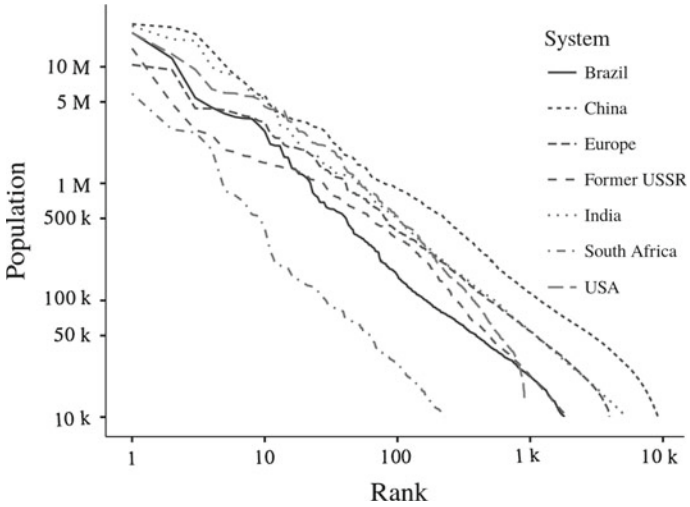


Fig. 1.1 Distributions of city sizes in seven large regions in the world

of the world present singularities resulting from their evolution in historical and political contexts that were sometimes very different. Thus identification of generic evolutionary mechanisms and processes peculiar to each system of cities should enable better prediction of its future evolution. From the observation of the variations of urban hierarchies around the strong attractor that represent the statistical models of Zipf's rule or the lognormal distribution of city sizes (Fig. 1.1), we have identified three different types of systems of cities: ancient systems, 'new world' systems and colonization systems.

1.3.1 A Simplified Typology of Systems of Cities

The diversity of hierarchical differentiations has been widely studied, using information databases that are strictly comparable, by F. Moriconi-Ebrard (Moriconi-Ebrard 1993); other authors, including ourselves, have analyzed functional diversity in different types of economy and for different periods identifying three types of systems of cities, each linked to a specific historical context, differentiated by their spatial and hierarchical distribution.

A first type characterizes the countries and continents where the urbanization process has been ancient and regular, as in Asia and Europe. In this context, cities were formed at times of slow transportation means; they were more or less regularly spaced, at short distances. Urban hierarchies are still not very contrasted and include many small towns. China, India and Europe illustrate that type of structure on Fig. 1.1.

A second type of system of cities characterizes countries of the 'new world', where cities developed more recently, at a time when transportation systems could link more distant places faster, and where development proceeded more from the dynamics of these networks than the provision of services to a less dense or non-existing rural population (Moriconi-Ebrard 1993; Bretagnolle et al. 2008). In these countries urban growth depended first on trade with the outside and it boosted small towns in waves along pioneering fronts (Bretagnolle and Pumain 2010). Growth was less regularly distributed spatially and temporally as in systems of cities of the 'old world'. As a consequence of this particular history of urban development, systems of cities in the new world have a smaller number of cities, stronger contrasts among city sizes and a higher degree of concentration of the urban population (this is true at the macro-scale, even though the local density of urban settlements may be very low as in the United States). Comparing the slopes of rank-size distribution on Fig. 1.1 illustrates this fact since the USA, South Africa and Brazil have values of the exponent well above one whereas China, India and Europe are below (Pumain et al. 2015).

A third type of system of cities is specific to the countries where the urbanization process is ancient but its evolution was disturbed by an 'external' shock during a stage of colonization. As a consequence, there is often a dual pattern including an endogenous system of central places superimposed by one or a few very large metropolises, often maritime ports that ensured the connection with the country from which the colonial power originated and whose growth was disproportionate with the economy of the local territory. This is the case in many African countries, it also happened in India with the exceptional development at the time of the British Empire of large ports as Mumbai, Kolkata and Chennai and the capital Delhi (Swerts 2013). In such systems, the stigma of colonization is thus a 'macrocephaly' of the system of cities which introduces a sharp discontinuity in the distribution of city sizes—this is more rarely encountered in countries having a more continuous history, although it may happen as well in particularly politically centralized smaller countries such as France.

1.3.2 Systematic Variations in the Rhythm of Urban Growth

Another important distinction which has to be made before modelling the evolution of systems of cities is linked to the date of their 'urban transition', which is rather well correlated with the level of economic development they have nowadays. The richest countries in the world have achieved this urban transition. It started in the nineteenth century with industrial revolution and ended a little after the mid-twentieth, reaching urbanization rates (i.e. the proportion of urban in total population) above 70%, around 80% in Japan, North America and Europe and close to 100% in the smallest of 'city states' such as Singapore. Among developing countries, history has been more heterogeneous, since many countries of Hispanic culture in South America were urbanized early with respect to their current level of wealth, whereas on the contrary the urbanization rates remained very low in India and China as in most African countries South of Sahara—South Africa excepted.

There is another difference occurring in demographic and urban transition which is linked to the level of economic development. As a consequence of their delayed urban transition, which coincided with their also delayed demographic transition, the poor countries had access to better health conditions because of globalization and they thus experienced much higher urban growth rates than what was observed in industrialized countries at the time of the industrial revolution. Typical urban growth rates are around 4% per year in many poor countries since the 1950s which is twice the highest mean values observed in UK in nineteenth century for instance (Robson 1973). Such an acceleration of the urbanization processes combined to their historical concomitance with a completely modified context of socio-spatial interactions and economical and ecological constraints may of course introduce new peculiarities in the way systems of cities will evolve in the now developing countries.

Therefore we use these classifications as a basis for selecting systems of cities on which our simulation experiments are conducted, using adapted versions of the generic SIMPOP model in order to provide a better understanding of the particularities of the evolution of each system of cities. We shall summarize briefly in Chap. 2 of this book how the successive versions of the 'Simpop family' of multi-agent models (Pumain et al. 2012) have helped in testing and developing our evolutionary theory of urban systems. We rely on this type of modelling for constructing a 'virtual laboratory' that mixes knowledge from urban experts and computer scientists and is dedicated to the exploration of possible urban evolutions.

1.4 Urban Future: Models and Scenarios

There are many challenges in predicting the future or urban systems. Some are linked with uncertainties about the concrete sources of yet unpredictable technological or cultural innovations or even sudden outbreak of societal conflicts that may induce catastrophic changes and create tipping points in the evolutionary curve of urban development. Such events are clearly the worse limitation to any attempt in predicting urban futures. But even in the perspective of an urban evolution that would not suffer such kind of 'external' perturbation, our simulation exercises meet two types of difficulties; some depends on the choice between different scenarios; others relate to the capacities of our modelling techniques.

1.4.1 Challenges in Building Scenarios About Urban Evolution

The simplest way of drawing a scenario for the future evolution is first to capture the existing dynamics (for instance by calibrating a model on a historically observed evolution) and then to extend it with the integration of highly predictable evolutions that are already calculated by many institutions engaged in future studies and forecasting. It is easy to integrate their previsions in our models, especially those which relate to demographic growth, population and labour force, and the predicted trends for the urban transition, accelerated in China and still slower in India and booming in Africa: urban growth is both necessary for raising wealth and better health and welfare conditions while being facilitated by raising education level or foreign investments, even if possibly collaterally hampered by growing societal inequalities or an outbreak of armed conflicts. But even in this highly probable evolution, it is possible that the aggregated variables we use for simulating the urban evolution do neglect possible sources of tipping points that become difficult to predict when the detailed description of the corresponding processes is left out of the model.

Such failures to predict urban futures were observed with the previous versions of the Simpop model when we could not reconstruct the observed inflexions of past evolution without introducing ad hoc modifications of the ‘ordinary urban dynamics’ that was implemented in the model. For instance, we discovered that the Simpop2 model, while correctly reconstructing the evolution of most European towns and cities over three centuries, was totally unable to predict correctly the size that was actually reached by a few cities of exceptionally large size that dominated this system of cities at least since the Middle Ages (Bretagnolle and Pumain 2010). The existence of ‘primate cities’ (Jefferson 1989) in systems of cities, although rare, is a well-known feature usually included in urban theories and linked to the concentration in one place (usually a political capital) of the urban interactions with other systems of cities that brings very high returns and accelerated growth locally. Simpop2 should have included a special urban function (i.e. defining attributes and interaction rules of a ‘city agent’) corresponding to ‘global cities’ in the model to solve this problem.

Another challenge is to find a way to simulate the tipping points that lead to a switchover in the relative evolution of entire regions of the system of cities. Our previous simulation of the European system of cities demonstrated the difficulty of accommodating the ‘usual’ dynamics of a system of cities to reconstruct the shift in urban growth from cities of the Mediterranean regions towards those of the North Sea around the seventeenth century (Bretagnolle and Pumain 2010). A similar difficulty would arise when trying to develop a scenario taking into account the shift of urban and economic development that occurred in the second half of the twentieth century from the Atlantic toward the Pacific regions of the world. Although, in demographic terms, cities of the Pacific region already win the size contest, it is debatable if and when the same will happen in economic terms. Tuning that in models is not so easy, as illustrated perhaps by the successive adjustments of Chinese policies dealing precisely with that evolution.

Other challenges that are often mentioned in designing scenarios for urban futures may be not so difficult to meet. For instance dealing with the ‘ecological revolution’ that would be made inevitable by climate change and the depletion of resource and implementing its possible effects in the evolution of systems of cities can be relatively easy to imagine because it is probable that it will follow the usual process of diffusion of innovation waves: there will most probably be a ‘top down’ diffusion of new regulations, as for instance those of the COP21 international agreement, and a 'bottom-up' diffusion of locally invented 'best practices' that circulate among cities in a diversity of information and emulation networks. A similar predictive scenario could be imagined about the possible effects of the Internet or the digital revolution: although many often mention their contribution to a 'flattening' of world disparities what is observed is a trend to growing inequalities as usual in the dynamics of systems of cities. Even if local urban bifurcations are still possible with respect to the surge of some highly specialized cities, it must be recalled that such 'economic miracles' do not happen in isolated places as in the example of Silicon Valley indeed close to San Francisco already existing infrastructures, social networks, human skills and capital as well as federal investments (Saxenian 1994; Storper et al. 2015).

1.4.2 Challenges in Model Validation

The reliability of the predictions that can be made using models not only depends on the plausibility of the scenarios they use but it also—if not above all—depends on the quality of these models which are usually tested through calibration on past data. However the methods that were available until recently to ensure the validity of the calibration of a multi-agent model were clearly far from sufficient (Rey-Coyrehourcq 2015). For instance, the first versions of the Simpop model were calibrated simply by trial and error. This method has the advantage of allowing the experimenter to intuitively guide the selection of parameter values depending on the desired effect, but it proves a source of difficulty when the model is complex and subject to many bifurcations: a slight change in the value of a parameter can cause an opposite trend to that desired, which requires the whole process to be tested again after changing the values of one or more other parameters.

This method quickly becomes tedious and time consuming, notwithstanding frustrating as the modeller has no idea of the huge parameter space which is left out of his scope. It is generally based on a calculation designed to minimize an objective function whose stopping criterion does not establish the certainty that it has attained the best possible value. Consequently, once the model was consolidated in its implementation with a first battery of experiments, the determination of the parameter values for considering the model as 'validated' was performed with a hundred simulations (Bura et al. 1996; Sanders et al. 1997; Bretagnolle and Pumain 2010). Indeed, this rather low number of experiments that was completed with each version of the model also is explained by the capacity of available computers at the time that implied a rather long-computation time for each simulation because this type of model uses a huge number of spatial interactions between cities.

We demonstrate in this book how these difficulties can now be overcome or reduced by the development of new ways of building and evaluating simulation models in geographical systems. First, we better assess the elementary processes which characterize the dynamics of systems of cities in a very simplified model designed to simulate the emergence of cities, the SimpopLocal model; second, we develop an automated method for its evaluation that change dramatically the order of magnitude of the number of experiments for validation from some 10^2 to some 10^8 (Chaps. 2 and 3); third, we develop a new way of designing a model that injects step-by-step refinements in the granularity of its description, either by introducing new mechanisms that integrate less general aspects of the dynamics of the urban system under study or by introducing complementary elements of the environmental or political context that may interfere in the evolution of the system of cities (Chap. 4). These developments rely on the construction in parallel of powerful new tools that deeply transform the practice of geographical modelling and enlarge our confidence in its further practical application.

Conclusion

In theory, urban systems have a historical dimension, including path-dependent specifics and contingencies, which makes their future unpredictable, like that of other self-organizing complex systems subject to emergence processes or bifurcations. However, observations of past developments enable to establish a series of 'stylized facts' or regularities in urban dynamics that provide a framework of possible future developments, for periods of several decades with a relatively satisfactory level of confidence.

We shall demonstrate in this book a new way of building models that would be more reliable and useful for simulating the evolution of systems of cities in a variety of geographical and economic conditions. A new start is proposed, relying on our previous experiences, for building and implementing new multi-agent models using the power of new computing methods and considerably enlarged computing capabilities.

References

Arthur, W.B.; Increasing returns and path dependence in the economy. The University of Michigan Press (1994)

Auerbach, F.: Das Gesetz der Bevölkerungskonzentration. Petermanns Geogr. Mitt. **59**, 74–76 (1913)

Bairoch, P.: De Jéricho à Mexico: Villes et économie dans l'histoire. Gallimard, Paris (1985)

Bank, W.: Geography. Reshaping Economic. The World Bank, Washington DC (2009)

Batty, M.: The Emergence of Cities: Complexity and Urban Dynamics. CASA Working Papers 64. http://www.casa.ucl.ac.uk/working_papers/paper64.pdf (2003)

Batty, M.: Hierarchy in Cities and City Systems. In: Pumain, D. (ed.) Hierarchy in the Natural and Social Sciences, pp. 143–168. Springer, Dordrecht, Netherlands (2006)

Batty, M.: The New Science of Cities. MIT Press, Massachusetts (2013)

Berry, B.J.: Cities as systems within systems of cities. Pap. Reg. Sci. **13**(1), 147–163 (1964)

Boschma, R.A., Frenken, K.: Why is economic geography not an evolutionary science? towards an evolutionary economic geography. J. Econ. Geogr. **6**(3), 273–302 (2006)

Bretagnolle, A., Giraud, T., Mathian, H.: La mesure de l'urbanisation aux etats-unis, des premiers comptoirs coloniaux aux metropolitan areas (1790–2000). Eur. J. Geogr, Cybergeo (2008)

Bretagnolle, A., Mathian, H., Pumain, D., Rozenblat, C.: Long-term dynamics of european towns and cities: towards a spatial model of urban growth. Eur. J. Geogr, Cybergeo (2000)

Bretagnolle, A., Pumain, D.: Simulating urban networks through multiscalar space-time dynamics: Europe and the united states, 17th-20th centuries. Urban Stud. **47**(13), 2819–2839 (2010)
Bura, S., Guérin-Pace, F., Mathian, H., Pumain, D., Sanders, L.: Multiagent systems and the dynamics of a settlement system. Geogr. Anal. **28**(2), 161–178 (1996)
Chauvin, J. P., Glaeser, E., Ma, Y., and Tobio, K.: What is different about urbanization in rich and poor countries? cities in brazil, china, india and the united states. J. Urban Econ. (2016)
Christaller, W.: Die zentralen Orte in Süddeutschland: eine ökonomisch-geographische Untersuchung über die Gesetzmässigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen. Fischer, Iena (1933)
Cottineau, C.: L'évolution des villes dans l'espace post-soviétique. Observation et modélisations. Ph.D. Thesis, Université Paris 1 Panthéon-Sorbonne (2014)
Cottineau, C.: Metazipf.(re) producing knowledge about city size distributions. arXiv preprint (2016) arXiv:1606.06162
Diamond, J.: Guns, Germs, and Steel: The Fates of Human Societies. WW Norton & Company, New York (1997)
Duranton, G., Puga, D.: Nursery cities: urban diversity, process innovation, and the life cycle of products. American Economic Review **91**(5), 1454–1477 (2001)
Favaro, J.-M., Pumain, D.: Gibrat revisited: an urban growth model incorporating spatial interaction and innovation cycles. Geogr. Anal. **43**(3), 261–286 (2011)
Ferber, J., Perrot, J.-F.: Les systèmes Multi-agents: Vers une Intelligence Collective. InterEditions, Paris (1995)
Fletcher, R.: Settlement archaeology: world-wide comparisons. World Archaeol. **18**(1), 59–83 (1986)
Gibrat, R.: Les inégalités économiques. Librairie du Recueil Sirey, Paris (1931)
Hagerstrand, T.: The Propagation of Innovation Waves. Royal University of Lund, Sweden (1952)
Hannerz, U.: Cultural Complexity: Studies in The Social Organization of Meaning. Columbia University Press, New York (1992)
Hernando, A., Hernando, R., Plastino, A., Zambrano, E.: Memory-endowed US cities and their demographic interactions. J. R. Soc. Interface **12**(102), 20141185 (2015)
Jefferson, M.: Why geography? the law of the primate city. Geogr. Rev. **79**(2), 226–232 (1989)
Krugman, P.: Increasing returns and economic geography. Technical report, National Bureau of Economic Research (1991)
Krugman, P.: Confronting the mystery of urban hierarchy. J. Jpn. Int. Econ. **10**(4), 399–418 (1996)
Lane, D., Pumain, D., van der Leeuw, S.E., West, G.: Complexity Perspectives in Innovation and Social Change, vol. 7. Springer Science & Business Media, Berlin (2009)
Marcus, J., Sabloff, J.A.: The Ancient City: New Perspectives on Urbanism in The Old and New World. School for Advanced Research, New Mexico (2008)
Moriconi-Ebrard, F.: L'urbanisation du monde depuis 1950. Anthropos, Paris (1993)
Nitsch, V.: Zipf zipped. J. Urban Econ. **57**(1), 86–100 (2005)
Panerai, P., Castex, J., Depaule, J.-C.: Formes urbaines: de l'îlot à la barre. Editions Parentheses, France (1997)
Polèse, M., Stren, R.E.: The Social Sustainability of Cities: Diversity and The Management of Change. University of Toronto Press, Canada (2000)
Pred, A.: City Systems in Advanced Societies. Hutchison, London (1977)
Pumain, D.: La dynamique des villes. Paris: Economica (Coll. Géographie Économique) (1982)
Pumain, D.: Pour une théorie évolutive des villes. Espace géographique **26**(2), 119–134 (1997)
Pumain, D.: Hierarchy in Natural and Social Sciences. Springer, Netherlands (2006)
Pumain, D., et al.: Une théorie géographique pour la loi de Zipf. Région et Développement **36**(2012), (2012)
Pumain, D., Moriconi-Ebrard, F.: City size distributions and metropolisation. Geojournal **43**(4), 307–314 (1997)
Pumain, D., Sanders, L.: Theoretical principles in interurban simulation models: a comparison. Env. Plan. A **45**(9), 2243–2260 (2013)

Pumain, D., Swerts, E., Cottineau, C., Vacchiani-Marcuzzo, C., Ignazzi, C.A., Bretagnolle, A., Delisle, F., Cura, R., Lizzi, L., Baffi, S.: Multilevel comparison of large urban systems. Eur. J. Geogr, Cybergeo (2015)

Rey-Coyrehourcq, S.: Simprocess, a integrated method to validate Agent-based models in geography. Ph.D. Thesis, Université Paris 1 Panthéon-Sorbonne (2015)

Robson, B.T.: Urban Growth. An Approach. Methuen, London (1973)

Rozenblat, C., Melançon, G.: Methods for Multilevel Analysis and Visualisation of Geographical Networks. Springer, Berlin (2013)

Rozenblat, C., Pumain, D.: Firm linkages, innovation and the evolution of urban systems. Cities in Globalization: Practices, Policies, Theories, pp. 130–156. Routdledge, London (2007)

Sanders, L., Pumain, D., Mathian, H., Guérin-Pace, F., Bura, S.: Simpop: a multiagent system for the study of urbanism. Env. Plan. B: Plan. Design **24**(2), 287–305 (1997)

Saxenian, A.L.: Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Harvard University Press, Cambridge, MA (1994)

Schmitt, C.: Modélisation de la dynamique des systèmes de peuplement: de SimpopLocal à SimpopNet. Ph.D. Thesis, Paris 1 (2014)

Scott, A.J.: The Cultural Economy of Cities: Essays on The Geography of Image-Producing Industries. Sage, London (2002)

Storper, M., Kemeny, T., Makarem, N., Osman, T.: The Rise and Fall of Urban Economies: Lessons from San Francisco and Los Angeles. Stanford University Press, California (2015)

Swerts, E.: Les systèmes de villes en Inde et en Chine. Ph.D. Thesis, Université Paris 1 Panthéon-Sorbonne (2013)

Wilson, A.: Catastrophe Theory and Bifurcation: Applications to Urban and Regional Systems. Routledge, London (1970)

Zelinsky, W.: The hypothesis of the mobility transition. Geogr. Rev. **61**(2), 219–249 (1971)

Zipf, G.K., et al.: National unity and disunity. The Principia Press, inc., USA (1941)

Chapter 2
The SimpopLocal Model

Abstract The model is described in detail in order to explain how to implement the stylized facts that are theoretically essential for representing the emergence of cities into a multi-agent system. The model activates rather simple feedback loops between settlement size, innovation and resources enabling settlement growth from local innovation creation and inter-settlement diffusion. The necessary attributes, processes and parameters of the model, that were identified according to a rule of parsimony, are described and their estimated values after simulations are presented.

2.1 Introduction

From the stylized facts summarizing the historical process of the emergence of cities which we have briefly recalled in Chap. 1, we create SimpopLocal to simulate the growth dynamics of agrarian settlements and their possible evolution towards urban settlements under strong environmental constraints that are progressively overcome by successive innovations. In order to ensure the replicability of the model, the source code of SimpopLocal is filed in a public repository (http://iscpif.github.io/simpoplocal-epb/).

2.2 Purpose of SimpopLocal

This exploratory model aims at reproducing a remarkable aspect of the spatial structure of settlements systems, defined in the literature as a major stylized fact: in any system already studied, in any places and in any period of history or prehistory times, the distribution of size (population or spatial extent) is strongly differentiated, including many very small settlements and only few very large settlements according to a rather regular distribution of the Zipf or log-normal type (Fletcher 1986; Liu 1996).

This hierarchical pattern is a structural property (order in the size of entities) at macroscopic level that is particularly resilient over time, whatever the local fluctuations which take place at entity level. The SimpopLocal model is designed for testing

© Springer International Publishing AG 2017

D. Pumain and R. Reuillon, *Urban Dynamics and Simulation Models*,
Lecture Notes in Morphogenesis, DOI 10.1007/978-3-319-46497-8_2

the hypothesis enunciated in the evolutionary theory of urban systems (Pumain 1997) which explains this structural feature from the urban growth process sustained by all kinds of technological and societal innovations and their spatial diffusion among connected settlements.

This model adds to the usual stochastic model of urban growth that is simply proportional to city size and leading to a log-normal distribution (Gibrat 1931) the effect of the spatial interaction which amplifies the growing hierarchical differentiation among settlement sizes that is observed over time in geographical urban systems (Favaro and Pumain 2011).

SimpopLocal is part of the Simpop agent based model family. In comparison of models already developed, SimpopLocal adopts some new original paths. First, it simplifies the way that was used until now to qualitatively discriminate the successive innovation waves that were represented by various urban functions, it captures all of them in a single abstract innovation object.

Second, SimpopLocal makes the process of innovation creation endogenous by linking it with the size of settlement. This more parsimonious version of model building enables the development of better and more systematic exploration and evaluation of ABM. SimpopLocal was initially developed using Netlogo language, and later redeveloped using Scala programming language.

We describe the model following the ODD standard principles (Grimm et al. 2010), in a slightly different order, and without describing the 'design concepts', whose categories are not relevant here.

2.3 Entities, State Variables and Scales

The model represents the evolution of settlement units that are dispersed in an area large enough for sustaining a few thousands population but limited enough in surface for ensuring the possible connection between settlements according to the transportation means that are available at the time. Typically, it could be a region as antique Mesopotamia or Meso America. The landscape of the simulation space is composed of hundreds of settlements. Each settlement is considered as a fixed agent and is described by three attributes: the location of its permanent habitat (x, y), the size of its population P, and the available resources in its local environment.

The amount of available resources R is quantified in units of inhabitants and can be understood as the carrying capacity of the local environment for sustaining a population which depends on the resource exploitation skills that the local population has acquired from inventing or acquiring innovation. This resource exploitation is done locally and sharing or trade is not represented explicitly in the model. Each new innovation created or acquired by a settlement develops its exploitation skills. Contrary to previous more 'realistic' models of the Simpop family, we do not want to consider the nature of innovation by identifying each significant innovation wave as a new urban function. We simplify the model by retaining only the processes of emergence of innovation and their effect on urban growth. The innovation entity is understood here as a large and abstract invention socially accepted which could represent a technical invention, a discovery, a social organization, some new habits or practices ... Each acquisition of innovation by a settlement brings there the possibility to surpass its capacity threshold, and by consequence authorizes a demographic growth. The state variables defined at macro-level are the size distribution of settlements and the slope of the rank size distribution.

2.4 Processes Overview and Scheduling

The scheduling of a simulation of the model is presented on Fig. 2.1 and will be further detailed for each part of the process linking the evolution of innovation, resource and population growth in the settlements.

After the initialization of the settlements, the interaction network is created. Then, at each simulation step, the mechanisms of population growth (*grow population*) and innovation diffusion (*diffuse innovation*) are applied. According to the number of innovation, the impact of these innovations is applied on the settlement's resource extraction efficiency (*apply innovations*). Then, the innovation creation mechanism (*create innovation*) is applied, with the same effect on resource extraction efficiency. This loop is iterated until the stopping criterion is reached: in this case after 4000 steps or if the maximum number of innovation has been reached. We now present each of these mechanisms in detail. Regarding the ODD protocol, these mechanisms would be labelled as the submodels of SimpopLocal.

2.4.1 Population Growth Mechanism

The growth dynamics of a settlement are modelled according to the assumption that its size is dependent on the amount of available resources in the local environment and is inspired by the Verhulst model (Verhulst 1845) or logistic growth.

For this experiment, we assume a continuous general growth trend for population—this may be different in another application of the model. The growth factor r is expressed on an annual basis; thus, one iteration or step of the model simulates one year of demographic growth. The limiting factor of growth R^i_M is the amount of available resource that depends on the number M of innovations the settlement i has acquired by the end of the simulation step t.

P^i_t is the population of the settlement i at the time t:

$$P^i_{t+1} = P^i_t \left[1 + r \left(1 - \frac{P^i_t}{R^i_m} \right) \right] \tag{2.1}$$

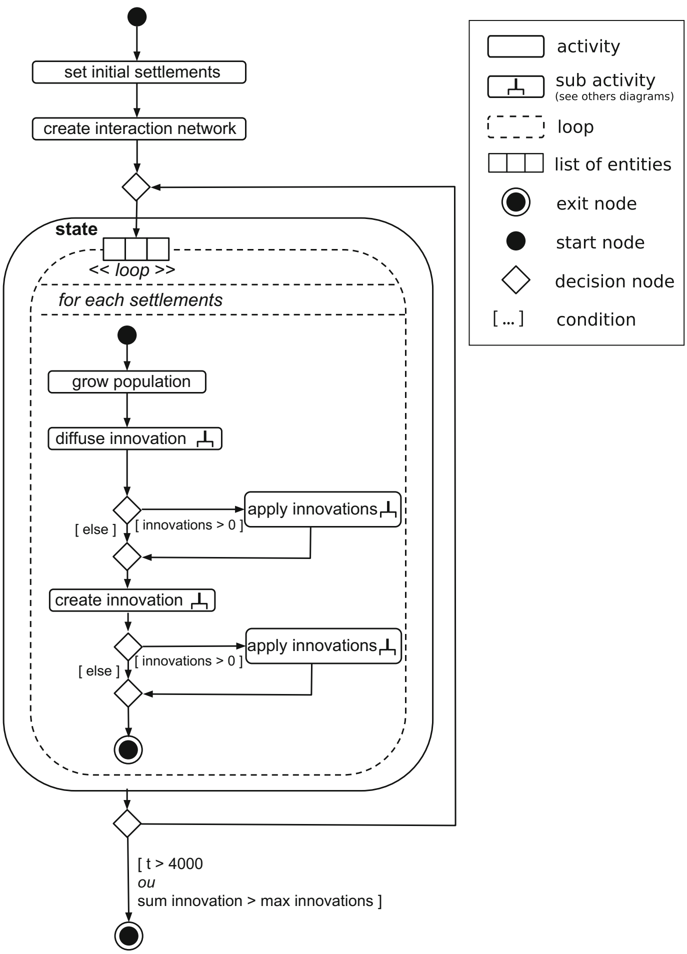


Fig. 2.1 SimpopLocal activity diagramm

2.4.2 Apply Innovation Mechanism

The acquisition of a new innovation by a settlement allows it to overtake its previous growth limitation by enabling a more efficient extraction of resources and thus a gain in population-size sustainability. With the acquisition of innovations, the amount of available resources tends to the maximal carrying capacity R_{max} of the simulation environment:

$$R_M^i \xrightarrow{innovations\ acquisition} R_{max} \tag{2.2}$$

The mechanism of this impact follows the Ricardo model of diminishing returns (which also is a logistic model). The $InnovationImpact$ represents the impact of the acquisition of an innovation and has a decreasing effect on the amount of available resources R^i_{M+1} with the acquisition of innovations:

$$R^i_{M+1} = R^i_M \left[1 + InnovationImpact\left(1 - \frac{R^i_M}{R_{max}}\right)\right] \tag{2.3}$$

2.4.3 Create and Diffuse Innovation Mechanisms

Acquisition of innovations can occur in two ways, either by the emergence of innovation within a settlement or by its diffusion through the settlement system. In both cases, interaction between people inside a settlement or between settlements is the driving force of the dynamics of the settlement system. It is a probabilistic mechanism, depending on the size of the settlement. Indeed, innovation scales superlinearly: the larger the number of innovations acquired the larger the settlement and the higher the probability of innovation. To model the superlinearity of the emergence of innovation within a settlement, we model its probability to be created by a binomial law.

If $P_{creation}$ is the probability that the interaction between two individuals of the same settlement is fruitful, that is, leads to the creation of an innovation, and N the number of possible interactions, then, by the binomial law, the probability of the emergence of at least one innovation $P(m_{creation} > 0)$ can be calculated and then used in a random drawing:

$$\begin{aligned} P(m_{creation} > 0) &= 1 - P\left(m_{creation=0}\right), \\ &= 1 - \left[\frac{N!}{0!(N-0)!} * P^0_{creation} * (1 - P_{creation})^{N-0}\right], \\ &= 1 - (1 - P_{creation})^N \end{aligned} \tag{2.4}$$

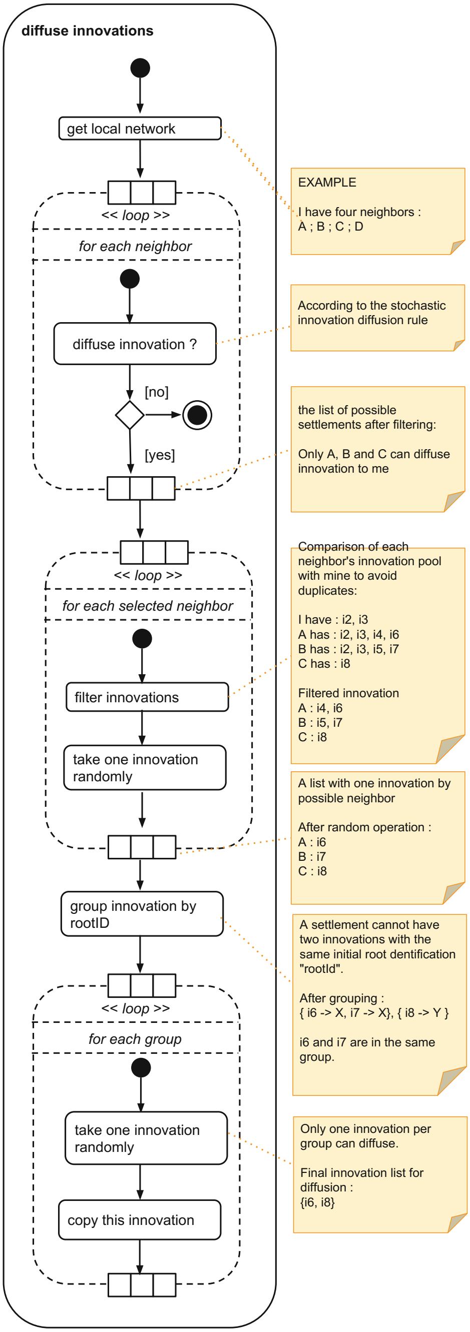
If the size of the settlement is P^i_t then the number N of possible interactions between individuals of that settlement is:

$$N = \frac{P^i_t\left(P^i_t - 1\right)}{2} \tag{2.5}$$

The diffusion of an innovation between two settlements depends on both the size of populations and the distance between them.

If $P_{diffusion}$ is the probability that the interaction of two individuals of two different settlements is fruitful—that is, leads to the transmission of the innovation—and K is the number of possible interactions, then, by the binomial law, the probability of diffusion of at least one innovation $P(m_{diffusion} > 0)$ can be calculated and used in a random drawing:

Fig. 2.2 Diffuse innovation mechanism activity diagram



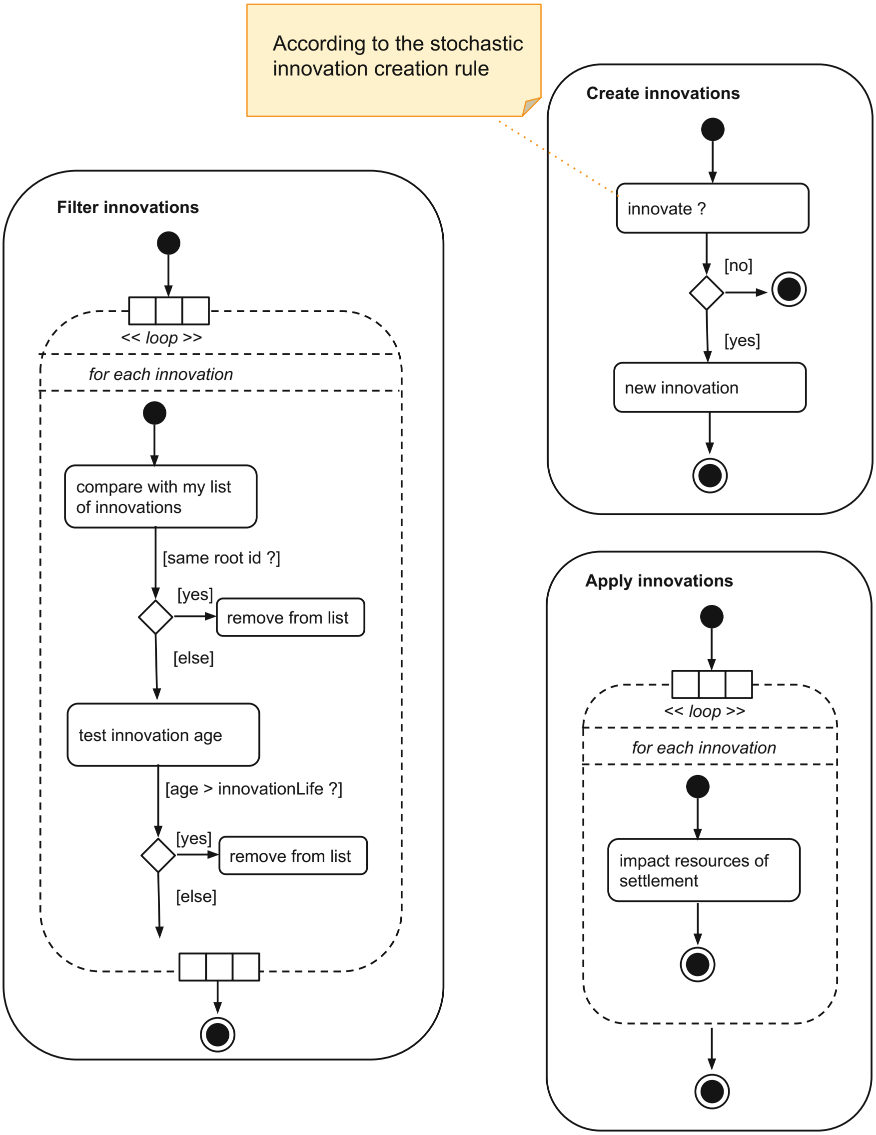


Fig. 2.3 Creation, filter and apply innovation mechanisms activity diagrams

$$P(m_{diffusion} > 0) = 1 - (1 - P_{diffusion})^K \tag{2.6}$$

But in this case, the size K of the total population interacting is a fraction of the population of the two settlements i and j which is decreasing by a factor

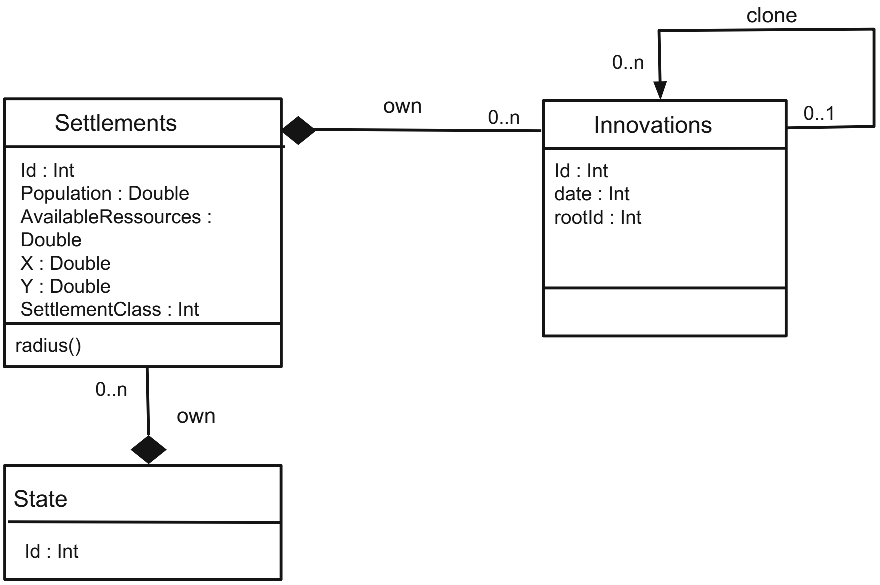


Fig. 2.4 UML class diagramm of SimpopLocal

$DistanceDecay$ with the distance D_{ij} between the settlements, as in the gravity model:

$$K = \frac{P_t^i P_t^j}{2D_{ij}^{DistanceDecay}} \tag{2.7}$$

The process of population growth and the process of innovation creation and diffusion are reiterated throughout the simulation (Figs. 2.2 and 2.3). Because of the two positive feedbacks that operate on resource and population growth through the creation of innovation, the model is able to generate a very rapid expansion of settlements: that is, an escalation of settlement growth. The simplest way to avoid situations where too many innovations are created, which would lead to huge time-consuming simulations, is to decide to stop the simulation when it reaches an arbitrary number of, say, 10,000 innovations. Finally, on Fig. 2.4 the UML class diagram of the model is illustrated.

2.5 Initial Conditions

The initial configuration we have chosen to keep in every experiment has therefore been defined the following rules to ensure a good representation of common structures of settlement systems: The size of settlements follows a log-normal distribution.

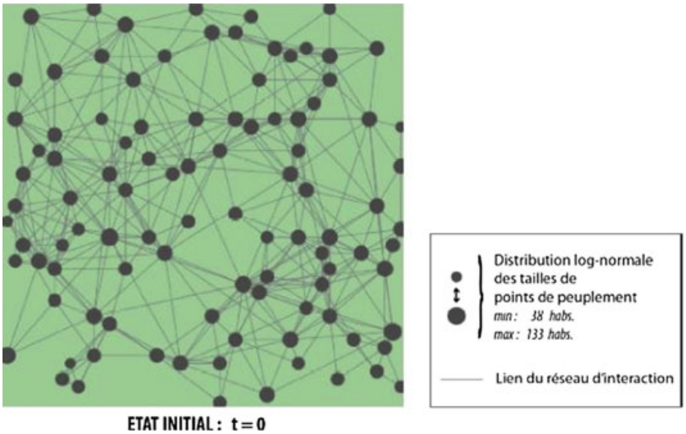


Fig. 2.5 Agents (settlements) and attributes at initial configuration in SimpopLocal

100 settlements size are initialized with a random demographic size vary from 38 to 133 inhabitants. The spatial repartition of the settlements assumes a Christallerian pattern (according to Christaller,[1] a theoretical pattern of central places has regular spacing distances between settlement nodes of the same size category, the largest settlement nodes having larger spacing distances than the smaller) (Fig. 2.5).

The network interlinking the settlements enables spatial interactions according to the same Christallerian logic: small settlement nodes establish less remote connections than the largest. Furthermore, at initial state, population of settlements is considered at equilibrium regarding the available resources: the initial amount of resource of each settlement is considered equal to its initial population.

2.6 Input

Although relatively parsimonious as a multi-agent system, SimpopLocal has a dozen of parameters that have to be estimated for calibrating the model. Some can be empirically evaluated with the help of historical data and knowledge, while it is very difficult to give values to others (Table 2.1). Those regarding the initial spatial distribution and organization of settlements in the landscape can be approximated. The log-normal distribution of the settlement sizes and the central place theory of Christaller for the geographical distribution of locations are models that are widely used by

[1]Christaller's central place theory (1933) considers cities as centres serving services of different levels of rarity to a regional resident population according to a hierarchy of range and size of the centres. As residents minimize the cost of access to services centres would exhibit regular patterns on an homogenous plain. The simplest pattern is made of nodes located at the summits of hexagons that are embedded in hexagons of larger areas designed for larger centres of the next upper level.

Table 2.1 Parameters of the SimpopLocal model (Source: Schmitt (2014), p. 167)

Parameters	Description
R_{max}	Maximum carrying capacity of a settlement (measured in number of residents)
r_{growth}	Mean growth rate as in verhulst model
$InnovationImpact$	Impact of any innovation on available resources
$P_{Creation}$	Probability of creating an innovation in one settlement
$P_{Diffusion}$	Probability of diffusing innovation between settlements
$DistanceDecay$	Deterrent effect of distance on innovation diffusion
$InnovationLife$	Time during which an innovation may diffuse
$MaxInnovation$	Total number of innovation generated before end of simulation

archaeologists to describe their spatial data (Archaeomedes 1998; Johnson 1977; Sanders 2012) including Neolithic archaeological sites (Liu 1996).

In SimpopLocal, the mean density of that landscape and the average size of each settlement are representative of the usual orders of magnitude presented in these works. A hundred settlements are distributed according to these two theories and each settlement is initially composed of some 80–400 inhabitants. Several scholars agree that an average annual growth of 0.02% is representative of the growth of agrarian settlements in the Neolithic times (Bairoch 1985; Renfrew and Poston 1979). The length of time required for a transition from agrarian to urban settlements is estimated according to (Bairoch 1985; Marcus and Sabloff 2008) to about three thousand years. We choose to operate our simulations on a four thousand years time period for settlements ranging from one hundred inhabitants up to about ten thousand inhabitants.

Because of a lack of empirical data, five parameters cannot be empirically approximated and have to be estimated through simulation:

- $Pcreation$, the probability that an innovation emerges from the interaction between two individuals of a same settlement.
- $Pdiffusion$, the probability that an innovation is transmitted between two individuals of different settlements. We consider that the probability of diffusion is greater than the probability of creation, which means that copying is easier than inventing (Pennisi 2010) in the model.
- $InnovationImpact$, the impact of the acquisition of innovation on the growth of settlements.
- $DistanceDecay$, the deterrent effect of distance on diffusion.
- R_{max}, the maximum carrying capacity of the landscape of each settlement (measured in number of inhabitants).

2.7 Running the Model for Parameter Estimates: Calibration

The principle of parsimony that led the development of SimpopLocal was applied as well in designing a way for estimating possible values for each parameter of the model. This original estimation process that leads as well to a huge qualitative improvement in the validation process of the hypothesis of the model will be examined in detail in Chap. 3. We only mention here which general line was followed in order to make understandable the results of simulation that are recalled below. As we lack of observed measurements for determining the possible values of most of parameters of the model, our method of estimation is not exactly a 'calibration' exercise. It consists in determining which global subsets of parameters values are leading to an emergence of a system of cities whose characteristics match at best the stylized facts that were identified in Chap. 1. A kind of machine learning method is necessary in order to identify through many possible behaviours of the model, the one which is able to correctly reproduce the expected results of simulation. The values we will get for the parameters are thus not measured in absolute units, they are abstract estimations and their significance relies on these measurements taken as a whole, each parameter remaining associated to the others having to be considered in relative terms. As two parameters involve probability distributions, the model is stochastic, therefore the two techniques (by trial and error and by full plan) usually used to calibrate a model are not suited to calibrate the SimpopLocal model (or any multi-agents model in general). We adopted innovative methods of automatic exploration of the patterns in the behavioural space of parameters which are developed on the simulation platform OpenMole. In Chap. 3 we shall explain in detail how genetic algorithms and grid computing are used to explore in a comprehensive way the parameter space, as it was roughly defined at first by a plausible but large enough variation domain for each parameter (Table 2.2).

We briefly retrace here in a vocabulary that is accessible to non-specialists of computing science how the method is working. An important first step in calibration is to define an objective function which is identified from the stylized facts describing the period. It includes three quantifiable elements that must be obtained at the end of simulation for the simulated system of cities:

- A log-normal distribution of settlement size
- The size of the largest aggregate settlement of about 10,000 inhabitants
- A total of 4,000 simulation time steps (equivalent to some 4,000 years).

The first requirement of the objective function reflects the essential hierarchical property of any system of cities; the second acknowledges that in the political and technological conditions of the time, groups of resident population over 2,000 inhabitants were very rare and a concentration of 10,000 could represent a major political and economic capital of a kingdom or empire; the third condition is constraining the model to be contained in a domain of growth regime for settlements that is plausible in demographic terms for the post-Neolithic period: at that time, rapid urban growth

Table 2.2 Variation domain for parameters after 500 millions simulations with SimpopLocal and their precision with calibration profile algorithm (Source: Schmitt (2014), p. 203)

Parameters	Initially assumed variation domain	Variation domain inside Pareto Front	Calibration validated domain
R_{max}	[1; 40000]	[9500; 11500]	[10090; 10465]
$InnovationImpact$	[0; 2]	[6.10.10^{-3}; .10^{-2}]	[7.7.10^{-3}; 8.4.10^{-3}]
$P_{Creation}$	[0; 1]	[4.0.10^{-7}; 2.1.10^{-6}]	[1.1.10^{-6}; 1.3.10^{-6}]
$P_{Diffusion}$	[0; 1]	[3.10^{-7}; 1.8.10^{-6}]	[6.7.10^{-7}; 6.9.10^{-7}]
$DistanceDecay$	[0; 4]	[0.2; 1.1]	[0.66; 0.75]

rates were hampered by insufficient resource, high mortality rates due to bad sanitary conditions and frequent catastrophes due to natural hazards or devastation caused by war. Meeting these objectives entails contradictory dynamic trends. Moreover, the need of (at least) a hundred replications of the simulations using the same set of parameter values to handle the stochasticity and the wide range of variation attributed a priori to five unknown parameters led us to use an evolutionary algorithm to solve this multi-objective optimization problem as well as massively distributed computing for the exploration of the entire parameter space.

2.8 Simulation Results and Return on Observations

In total, 500 million of model runs were conducted to achieve the calibration of parameters presented in Table 2.2. This table shows next to each parameter in a first column the hypothetical variation domain initially assumed, which was designed deliberately very wide, and in a second column the possible interval of values as it was reduced from the simulations to calibrate the model. This result does not lead to a single value but provides a range of possible values for each parameter, because a 'Pareto front' establishes compromise between values that ensures the multi-objective optimization function (that is explained with more details in Chap. 3). A third column shows which precision gain was realized for each parameter by using a more powerful algorithm that calculates the model's sensitivity to variations of a parameter at a time, all things being equal as to changes in the others. In addition, we must remember that the values presented in Table 2.2 are not independent measurements, they are connected and so it is their entire configuration that must be adapted when the model will be applied for a calibration on empirical, historical or archaeological situation (Figs. 2.6 and 2.7).

Another novelty of this experience is that the method for exploring the behaviour of the model is also a validation: we can establish to what extent the assumptions chosen to implement the mechanisms of the model are both necessary and sufficient to achieve the desired result—of course within the framework of the description

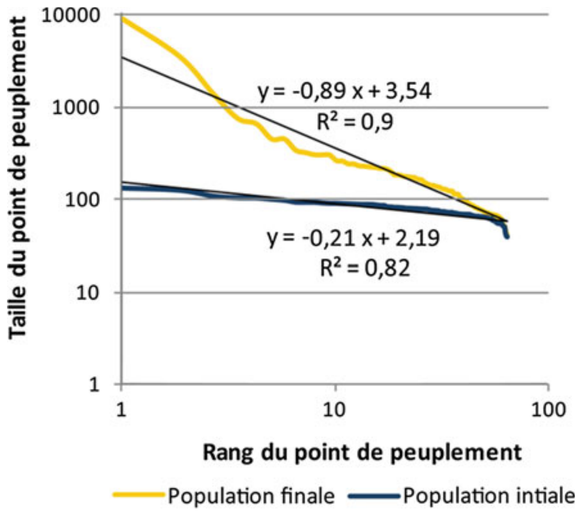


Fig. 2.6 Initial and simulated distribution of city sizes. *Source* Schmitt 2014

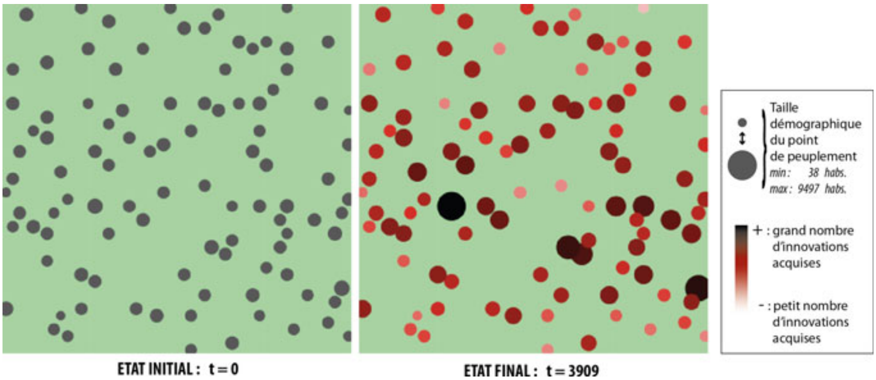


Fig. 2.7 Comparing initial and simulated settlement distributions. *Source* Schmitt 2014

chosen for the selected model. This second result comes from the development of a method called 'calibration profiles' (described in detail in Chap. 3) that calculates for each parameter the effect of its variation on the quality of the model, as fitted by the objective function, ceteris paribus about the changes of other parameters. That particular method led to the rejection of $InnovationLife$ parameter (lifetime of innovation) which was not sufficiently constraining the development of the model. The method also contributed to clarify the role of parameters through reducing their domain of variation to a more precise interval that is both necessary and sufficient to achieve the desired change (last column of Table 2.2).

These results provided by the analysis of calibration profiles provide valuable feedback on the modelling assumptions and urban theory that oversees the development of the model. The urban evolutionary theory that guided the construction of SimpopLocal model and more generally all models of the Simpop family insists on the concept of system, that is to say, relationship and interaction between the elementary entities (cities, villages or settlements points) that are components of the system. Yet, this is the first time that the exploration of the simulation model demonstrates that the mechanisms of interaction between the entities of the system are essential to the production of an evolution similar to that of real systems, in the context of the Simpop family of modelling (Pumain and Robic 2012).

Without these mechanisms describing spatial interaction, which in SimpopLocal are controlled by $Pdiffusion$ and $DistanceDecay$, so without diffusion of innovation between cities, according to a gravity principle, it is not possible to generate urban growth dynamics that are representative of dynamics actually observed in real systems. These results also show the importance of the role of space in structuring and organizing the settlement system: without the effect of $DistanceDecay$ parameter, which reduces the frequency of interaction with distance, changes in the simulated system are no longer representative of actual system developments. These first evaluations of mechanisms will also be useful for the next versions and applications of the SimpopLocal model. If this model is too simplified to be fully compliant with current or former real settlement systems in its first abstract and parsimonious version, the concepts generating the simulated processes can be reflected in certain contexts (i.e. the proto-historic cities, for example) or archaeological theories such as 'peer polity interaction' (Renfrew 1975).

Christopher Renfrew noticed how frequently the first small states were not born in isolation but in cluster, with strong similarities in terms of size, social structure, material culture, etc. He also observed that political entities comparable in size and organization (as the first forms of state organization) tended to emerge in the same areas and evolve simultaneously. Moreover, archaeological evidence suggests that these changes did not emanate from a single source of innovation, but emerged contemporaneously in several interacting units. According to these remarkable observations, our theory do confirm the central explaining role of the mechanism of exchange between settlement sites and describe the interaction processes as essential in urban development and social change.

The SimpopLocal model, whose dynamic is grounded in social and spatial interaction, could be used as a core model for testing this theory by simulation.

Perspectives of the application of complexity theory and methodological means to construct models (agent-based modeling) underline possible implications for the study of some theoretical issues of scientific research in archeology. The process of model building on the basis of theoretical concepts itself reveals gaps in our data. Within the archeological record we lack data for some processes which must be supplemented with estimates (Turchin and Gravilets 2009).

References

Archaeomedes,: Des oppida aux métropoles. Anthropos, Paris (1998)

Bairoch, P.: De Jéricho à Mexico: villes et économie dans l'histoire. Gallimard, Paris (1985)

Favaro, J.-M., Pumain, D.: Gibrat revisited: an urban growth model including spatial interaction and innovation cycles. Geogr. Anal. **43**, 261–286 (2011)

Fletcher, R.: Settlement Archaeology: World-Wide Comparisons. World Archaeol. **18**(1), 59–83 (1986)

Gibrat, R.: Les inégalités économiques. Librairie du Recueil Sirey, Paris (1931)

Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F.: The ODD protocol: A review and first update. Ecol. Model. **221**(23), 2760–2768 (2010)

Johnson, G.A.: Aspects of regional analysis in archaeology. Annu. Rev. Anthropol. **6**, 479–508 (1977)

Liu, L.: Settlement patterns, chiefdom variability, and the development of early states in North China. J. Anthropol. Archaeol. **15**(3), 237–288 (1996)

Marcus, J., Sabloff, J .A.: The ancient city: New perspectives on urbanism in the old and new world. Resident Scholar book, vol. 2005. School for Advanced Research, Santa Fe (2008)

Pennisi, E.: Conquering by Copying. Science **328**(5975), 165 (2010)

Pumain, D.: Pour une théorie évolutive des villes. Espace géogr. **26**(2), 119–134 (1997)

Pumain, D., Robic, M.-C.: Les Quarante Glorieuses de l'Espace géographique. L'Espace géogr. **1**(41), 1–3 (2012)

Renfrew, C.: Trade as action at a distance: questions of integration and communication. In: Sabloff, J.A., Lamberg-Karlovsky, C.C. (eds) A School of American Research Book, 25 cm.Papers from a seminar held in Santa Fe, Oct. 28-Nov. 2, 1973, pp. 3–60. School of American Research. University of New Mexico Press, Albuquerque (1975)

Renfrew, C., Poston, T.: Discontinuities in the endogenous change of settlement pattern. In: Renfrew, C., Cooke, K.L. (eds.) Transformations: Mathematical Approaches to Culture Change, pp. 437–461. Academic Press, New York (1979)

Sanders, L.: Regards croisés de géographes, économistes et archéologues sur la hiérarchie des systémes de peuplement: de l'empirie aux systèmes complexes. Région et Développement, 36 (2012)

Schmitt, C.: Modélisation de la dynamique des systèmes de peuplement: de SimpopLocal à SimpopNet. Ph.D. thesis, Université Paris 1 Panthéon-Sorbonne (2014)

Turchin, P., Gavrilets, S.: Evolution of complex hierarchical societies. Social Evolution & History **8**(2), 167–198 (2009)

Verhulst, P.F.: Recherches mathématiques sur la loi d'accroissement de la population. Nouveaux Mémoires de l'Academie Royale des Sciences et Belles-Lettres de Bruxelles **18**, 1–41 (1845)

Chapter 3
Evaluation of the SimpopLocal Model

Abstract The SimpopLocal model exposes 6 free parameters that cannot be set using empirical data. This chapter presents how to evaluate SimpopLocal in spite of these degrees of freedom. A first evaluation establishes whether the model has the capacity to produce acceptable dynamics. To achieve this evaluation, the quality of the simulated dynamics is made explicit using a quantitative analysis. Based on this quantitative evaluation, an automated calibration algorithm is designed using a state-of-the-art multi-objective genetic algorithm. The results show that the model is able to produce acceptable dynamics. A second evaluation exposes the contribution of each free parameter to the capacity of the model to produce these acceptable dynamics. A novel sensitivity analysis algorithm called calibration profile is then applied. The results of this analysis show that the model can be simplified by removing one superfluous mechanism and one superfluous parameter and that all the remaining mechanisms are mandatory in the model and all the remaining parameters can be better constrained by narrowing down their definition domains.

3.1 Quantitative Evaluation

As exposed in the previous chapter, the SimpopLocal model includes 6 free parameters. To evaluate this model, a design of experiment in the space of its parameters should be carried out. In order to find out if the model works as expected the space of parameters should be explored extensively. To design the exploration of SimpopLocal several aspects have to be taken into account: 1/ which stopping criterion should be used? 2/ how to measure the quality of the computed solution? 3/ how to sample the space of parameters?

3.1.1 Stopping Criterion

In order to automate the exploration and to test a large number of configurations, a stopping criterion should be established. To do so, two kinds of constraints have to be taken into account: thematic constraints and technical ones.

© Springer International Publishing AG 2017

D. Pumain and R. Reuillon, *Urban Dynamics and Simulation Models*,
Lecture Notes in Morphogenesis, DOI 10.1007/978-3-319-46497-8_3

SimpopLocal is an attempt to identify some of the mechanisms at works during the transition from a system of small settlements before agriculture advent (80–400 inhabitants per settlement), to a system of urban settlements (up to 7000 or 10,000 inhabitants for the bigger settlement). We cannot capture this complex transition into a single realist and unique history, therefore we decide to focus on the capacity of SimpopLocal to generate plausible dynamics with carefully chosen constraints (parameters, mechanisms, initial condition). To do so, the initial size and organization of system of settlements are set using common values and knowledge taken from specialized studies on this subject. We choosed to study the growth of 100 initial settlements with population size generated using the widely used log-normal and we used the central place theory (Christaller 1933) to distribute them geographically (Archaeomedes 1998; Johnson 1977; Liu 1996; Sanders 2012). Then we simulate the urbanization of the initial agrarian system of settlements in about four thousand years (Bairoch 1985; Marcus and Sabloff 2008). Based on this empirical values, we decided that SimpopLocal should be evaluated given its first 4000 simulation steps, which matches 4000 years of evolution of the system of cities.

On the technical aspect, we choose to represent innovations as autonomous objects into SimpopLocal. Choosing an object representation easily ensures the tracking of each innovation diffusion during simulation. Even more important, the acquisition process ensures that an innovation does not already exist in this settlement before recopy (i.e. adoption by the same settlement). One way to prevent recopy of an already existing innovation consist to store into each innovation the identification number of original innovation (before any copy). The number of innovation is therefore only growing during a run. These innovations are represented as objects which consume memory and increase the computation complexity of the model. For high values of the parameter `pCreation` this number of innovations can get arbitrary high. It would slow down the model execution and fill the computer memory. To avoid this situation we decided to establish a technical stopping criterion, by stopping the simulation when the number of innovation reaches 10,000 innovations. With this limit the model runs in approximately 1 second per simulation. This stopping criterion is purely technical and we seek to produce acceptable dynamics despite this computational limitation.

3.1.2 Expectations

Now that stopping criterion have been established, we can define a design of experiments to explore the parameter space of SimpopLocal. The widely used full factorial design of experiment is unpractical in our case. Indeed, using 10 levels for each of the 6 free parameters of SimpopLocal would produce 1 million parameter sets to evaluate. We will see bellow that this quantity of computation is affordable using modern distributed computing architecture, however checking 1 million dynamics visually is impossible.

The evaluation should be automated. The exploration of the space of parameters should directly produce a small set of parameter values which produce "expected" dynamics. We should first quantify what is an expected dynamic. To do so, we design three objectives to evaluate a single run of SimpopLocal (the lower the objective the better the dynamic):

- The objective of distribution, which quantifies the ability of the model to produce settlement size distributions that fit a log-normal distribution. To compute this objective, we evaluate the outcome of each simulations using a 2-sample Kolmogorov–Smirnov test (the deviation between the simulated distribution and a theoretical log-normal distribution having the same mean and standard deviation). Two criteria are reported, with value 1 if the test is rejected and 0 otherwise: the likelihood of the distribution (the test returns 0 if p-value >5%) and the distance between the two distributions (the test returns 0 if D-value <D[1]). In order to summarize those tests in a single quantified evaluation, we add the results of the two tests (the result of the test may be 0, 1 or 2 depending on the fit of the settlement size distribution to a log-normal).
- The objective of population, which quantifies the ability of the model to generate large settlements. The outcome of one simulation is tested by computing the deviation between the size of the largest settlement and the expected value of 10,000 inhabitants:
  $|(\textit{population of largest settlement} - 10{,}000)/10{,}000|$.
- The objective of simulation duration, which quantifies the ability of the model to generate expected configurations in a suitable length of time (in simulation steps). The duration of one simulation is tested by computing the deviation between the number of iterations of the simulation and the expected value of 4000 simulation steps:
  $|((\textit{simulation duration} - 4000)/4000|$.

3.1.3 Handling the Stochasticity

SimpopLocal is a stochastic model, meaning that its outputs are probability distributions. To estimate the quality of the dynamics produced by a stochastic model for a given set of parameters, the model should be run several times or replicated using independent random number streams. The results of the replications are independent realizsations of the output random variates of the model. The quantitative expectations for this model should then be expressed as descriptive statistics on the output distributions.

In our case we want to find suitable dynamics which are robust to stochasticity. It means that we seek parameter values such as the model dynamics gets as close as possible to the three previously defined objectives as often (for as many realizations

[1]Computed with $\alpha = 1.36$.

of the dynamics) as possible. To take into account the stochasticity of the model we define the three new evaluation objectives as follows:

- the aggregated distribution objective: the mean of the distribution objective among the replications,
- the aggregated population objective: the median of the population objective among the replications,
- the aggregated duration objective: the median of the duration objective among the replications

Note that the scale of these objectives are independent from the number of replications. It means that an evaluation based on n replications can be quantitatively compared with another based on m replication. This property will be useful for the following of this chapter.

3.2 Automated Calibration

3.2.1 Optimization Heuristic

Now that we have defined a quantitative evaluation of the model dynamics, we can sample the input parameter space to test if the model is able to produce suitable dynamics. Several methods are available to sample the space of parameters. They can can be split in two categories:

- the a priori samplings methods sample the space of parameters once and for all and then evaluate the model for each of the sampled points. In this category, the regular lattice is often used by modellers. Other samplings, with better space coverage are available such as the Latin Hypercube Sampling[2] or the Sobol Sequence[3] (for a full review on parameter space sampling report to Kleijnen 2007). These methods are simple to carry on and often allow rigorous statistical analysis of the results. However, they might be inefficient at finding acceptable dynamics when very few knowledge is available on the possible range of the input parameters.
- the iterative samplings take into account the already computed evaluations in order to generate more samples. This category contains instance calibration processes based on optimization algorithms (Stonedahl 2011), approximate Bayesian computation (Beaumont 2010; Lenormand et al. 2012), Calibration Profiles (Reuillon et al. 2015), Pattern Search Exploration (Chérel et al. 2015).

For SimpopLocal we have chosen to calibrate it through an iterative process based on a genetic algorithm. Since we have 3 objectives we used the well-established NSGA2 multi-objective optimization algorithm (Deb et al. 2000) using the 6 free

---

[2] https://en.wikipedia.org/wiki/Latin_hypercube_sampling.

[3] https://en.wikipedia.org/wiki/Sobol_sequence.

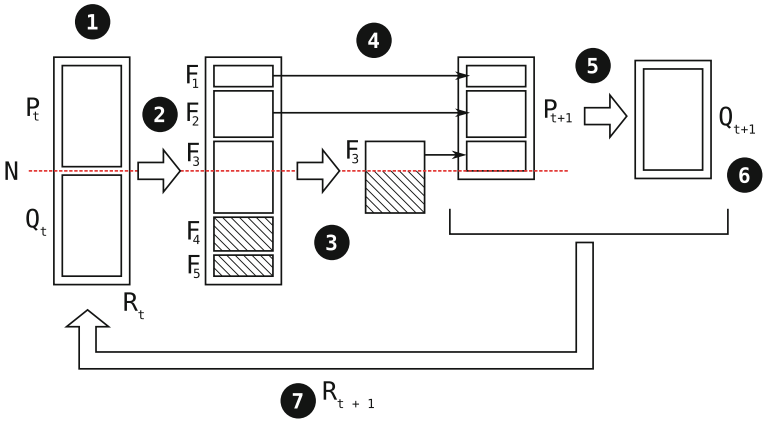


Fig. 3.1 NSGA 2 step by step procedure, inspirated by original schema from Deb et al. (2000)

parameters of SimpopLocal as the genome of the algorithm and the 3 objectives as the multi-objective fitness of the algorithm.

As represented on Fig. 3.1, the NSGA-2 algorithm computes the evolution of a population of size $2 * N$ called R. At the first iteration of NSGA-2 (I_0), the algorithm is initialized a population (P_0) of randomly generated solutions. Starting from I_1 the algorithm loops until convergence, performing the following steps:

- In step 2, the population (R_t) is ranked using **non-dominated sorted (NDS) algorithm**.[4] This algorithm uses the Pareto dominance to compute so-called fronts (a definition of Pareto dominance, and a detailed example of front computation is given later in this section). It groups individuals of the population by Pareto front $F_{1...n}$ using a multi-objective fitness. The individuals that belongs to front F_1 are dominated by no individual in the populations R_t. The individuals of front F_2 are dominated by no individual in the population $R_T - F_1$, the population where the individuals of F_1 are excluded ... and so forth.
- In step 3, the NSGA-2 algorithm computes a new population from R_T by selecting the best individual of R_T. The algorithm first adds all the individual of F_1, then the ones of F_2, the ones of F_3 ... It stops just before when adding an additional front in the population make it bigger than N individuals.
- In step 4, the algorithm adds a sub-part of the next front in order to complete the new population (it should reach a size of exactly N individuals). The individuals of a front cannot be discriminated by their objective values (by definition they constitute compromise solutions), therefore NSGA-2 uses another ranking based

[4] Invented by Goldberg (1989) but first implemented by Deb in NSGA (Deb et al. 2000).

on a diversity metric called the crowding distance operator. This selection based on a diversity metric helps maintaining a diversity of solutions in the population and not to converge too early in a local minimum. This population of size N is called P_{t+1}.
- In step 5, a set N parameter values (or genomes) is generated is generated by recombining and mutating individuals taken at random from P_{t+1}. This set is called Q_{t+1}.
- In step 6, this new offspring population Q_{t+1} is evaluated by running the fitness function. Each individual gain a new vector of value which contain evaluation for each objective function.
- In step 7, the new Q_{t+1} and already existing population P_{t+1} are merged into population R_{t+1}. Some convergence criterion is then tested. If the convergence has not been reached, the algorithm go to steps 2 otherwise it stops and returns R_{t+1}.

In this type of algorithm the best individuals are preserved in the R_t population used for fitness evaluation. This property is called **elitism** in evolutionary algorithms literature.

NSGA-2 heavily relies on the computation of the successive Pareto fronts. A Pareto front captures a group of individuals who are **non-dominated** by other individuals in a population. Classic definition of dominance say that "an element x_1 dominates (is preferred to) an element $x_2 (x_1 \dashv x_2)$ if x_1 is better than x_2 in at least one objective function and not worse with respect to all other objectives" (Weise 2011).

For instance in the Table 3.1, if we consider that best individuals are individuals which minimize value on objective function f_1 and f_2, we can see that I_d is better than I_c on f_1, but I_c is better than I_d on f_2. Therefore they are compromise solutions: none of them dominates the other. The set of individuals which are not dominated by any other individual of the population constitutes the Pareto front of the population. In this example, the Pareto front (which contain all non-dominated individuals) is $\{e, d, c, b, a\}$.

At step 4 of the algorithm, the Non-Dominated Sorting algorithm (NDS) computes successive fronts $F_{i...n}$ by iteratively removing non- dominated individuals $\{I_{\varnothing}\}$ from population P. In the example of Fig. 3.2, the front F_2 is computed by removing all non-dominated individuals (i.e. individuals in F_1) from population and then computing the Pareto front of this population. In this example we remove $\{e, d, c, b, a\}$, so the new Pareto front of P is equal to $\{f, h, j, k, l\}$.

3.2.2 Adaptation of NSGA2 to a Stochastic Model

A problem when using genetic algorithms to calibrate simulation models is that some of them do not cope well with stochasticity. This is especially the case for algorithms of type $\mu + \lambda$ (such as NSGA2), which preserve best solutions between

Table 3.1 Example of fitness value computed using NDS fitness algorithm and a population of individuals evaluated on two objective function f_1, f_2

Individuals	f_1	f_2	Dominated by	Fitness value
a	3.5	1	∅	1
b	3	1.5	∅	1
c	2	2	∅	1
d	1	3	∅	1
e	0.5	4	∅	1
f	0.5	4.5	{e}	2
g	1.5	4.5	{d, e, f, h}	3
h	1.5	3.5	{d}	2
i	2	3.5	{c, d, h}	3
j	2.5	3	{c, d}	2
k	3.5	2	{a, b, c}	2
l	4.5	1	{a}	2
m	4.5	2.5	{a, b, c, k, l}	3
n	4	4	{a, b, c, d, e, h, i, j, k, o}	5
o	3	4	{b, c, d, e, h, i, j}	4

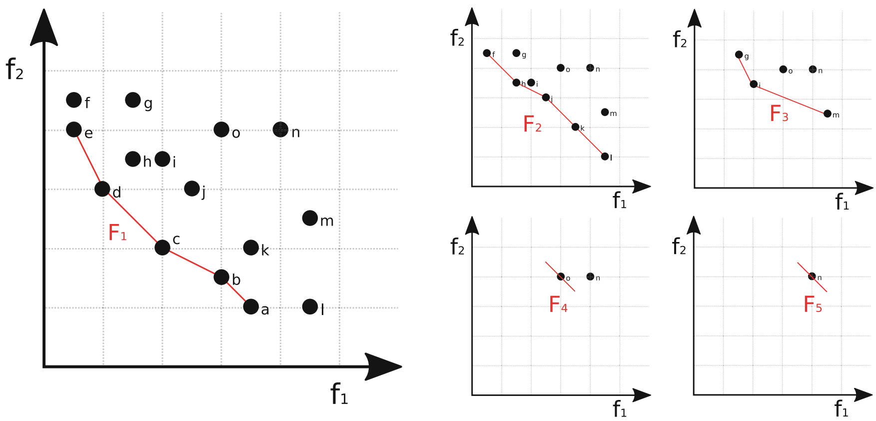


Fig. 3.2 Building steps example for NDS algorithm front computation with two objective function optimization. Algorithm produces five fronts, F_1 to F_5

the generations. In that kind of optimization the value of a solution is only estimated and not computed exactly. They can therefore be overvalued or undervalued (the quality of a solution is estimated with a significantly greater or a lower value than the one that would have been estimated given an infinite number of replications). Undervalued solutions are not very problematic for $\mu + \lambda$ genetic algorithms, they might be discarded instead of being kept, but the algorithm has a chance to retry a very similar solution later on. Conversely, the overvalued solution are very problematic for genetic algorithms, since the genetic algorithm might keep overvalued solution in the population of good solutions (because they have falsely been evaluated as good solutions) and generates new offspring solutions from them. This behaviour can greatly slow down the convergence of the calibration algorithm and even make it converge toward set of parameters producing very unstable output dynamics which are very likely to produce false positive good solutions.

To reduce the influence of the fitness fluctuation, the most commonly used approach is called "resampling". It consists in running several replications for each fitness evaluation. The computed quality for a set of parameters is then an estimation given a finite number of replications of the fitness computation. However, to limit the computation time taken to evaluate the quality of a single set of parameters during the calibration process, the number of replications is generally limited to a level which constitutes a compromise between the computation time taken to evaluate one set of parameters and an acceptable level of noise for the quality. Any number of replications, even very high, still implies that some solutions are overvalued with a non-negligible probability given that the fitness function is evaluated millions of times.

Other methods have been developed to optimize stochastic functions using genetic algorithms. Some of them are based on using the history of the genetic algorithm to estimate the probability distribution of the fitness (Sano and Kita 2002), others are based on the differences between the parents and the offspring (Tanooka et al. 1999) and others propose to use a partial order based on statistical tests (Rudolph 2001) ... Even if these methods seem statistically sound they complicate significantly the optimization algorithm, they are often based on some assumptions that are hard or impossible to verify (such as the invariance of the noise distribution over the fitness space) and they add parameters to the algorithm that are difficult to tune.

To overcome these limitations we have developed an auto-adaptive strategy to handle stochastic fitness functions in NSGA2. It is loosely related to the idea of resampling, for which only the best solutions are more precisely evaluated (presented in Branke 1998). In our method, called "stochastic resampling" we propose to evaluate the individual with only 1 replication and then to resample the individuals of the population with a fixed probability at each generation of the evolutionary algorithm. For instance, at each generation 90% of the individual offspring genomes are new genomes and 10% of the offspring genomes are already evaluated genomes randomly taken in the current population for which the algorithm computes one additional replication. The replications of each individual are stored in a vector of replications. The fitness of an individual is computed using (for instance) the median of each objective stored in the replication vector. The intuition is that in $\mu + \lambda$

genetic algorithms, best individuals survive several generations and therefore are the most likely to be resampled given that each individual has a fixed chance of being resampled at each generation. However, this fixed probability of resampling is not sufficient by itself to get an auto-adaptive algorithm. With this mechanism alone, well- evaluated solutions are very likely to be replaced by overvalued ones (new solution with a few “lucky” replications). To compensate this bias, we add technical objectives in NSGA2 in order to maximize the number of samples of a solution to the multi-objective optimization problem. Therefore, the number of replications is taken into account in the Pareto compromise elitism of NSGA2: solutions with many replications are kept even if some solutions are better on the other objectives but have been evaluated with less replications. By doing so, we let the multi-objective optimization algorithm handle the compromise between the quality of the solutions and their robustness. This method adds only two new parameters: 1/ the probability of resampling an individual at each generation 2/ the max number of samples for an individual to limit the memory used to store an individual. We propose to store the sample in a FIFO with a fixed size, therefore new samples are always taken into account even if the maximum number of replications has been reached for a given individual. This method has been implemented in the library for evolutionary computing: MGO[5] and has not been published yet.

3.2.3 Experimental Setup

To carry on the huge computation load required by the calibration of a stochastic multi-agent model using a genetic algorithm, we distributed it on the EGI,[6] a word-wide computation grid. To do so we used the framework OpenMOLE for distributed numerical experiments on simulation models[7] (this framework is described in more detail in the Chap. 6).

A classical way to distribute genetic algorithm is the technique known as the ‘island model’ (Belding 1995). The classical island model consists of instantiating permanent islands (isolated instances of an evolutionary algorithm) on many computers and organizing the migration of solutions between those islands. The EGI grid is a worldwide batch system on which organizing direct communications between islands running on multiple execution nodes is very challenging. Thus, we adapted the classical island model proposed in OpenMOLE to still benefit from the EGI architecture.

[5] https://github.com/openmole/mgo.

[6] http://www.egi.eu.

[7] http://www.openmole.org.

In this adapted version of the island model, a central population of 200 solutions is maintained on a central computer that orchestrates the submission of the computing jobs on the grid. Each job computes the evolution of the population of an island, which is an independent instance of NSGA2 started on a snapshot of the central population of 200 individuals at the time of submission. The 'island job' life cycle is managed by the EGI. Each job is submitted to the EGI and starts running when a slot becomes available on one of the data centres aggregated by the grid. When it starts running it is configured to run for 15 min. Using this distribution scheme, 1000 concurrent jobs are maintained (submitted + running) on the grid at any time. The OpenMOLE script for this experiment is exposed here.[8]

We executed 20,000 thousand islands of 15 min on the grid, after which we observed that the genetic algorithm is converged. An evolutionary algorithm is declared 'converged' when it makes no further improvements in the search for good solutions. One of the best metrics for measuring the convergence of the multi-objective optimization algorithm that is currently available is the stagnation of the hypervolume. The hypervolume measures the volume of the dominated portion of the objective space and its stagnation indicates that the algorithm has converged. To test if it is the case for our calibration we used the library MGO[9] to compute the evolution of the hypervolume. We considered only the solutions that are robust enough (estimated by the stochastic resampling strategy of the genetic algorithm with the maximum number of replications: 100 replications) and we used reference points (nadir) with the coordinates: distribution = 2.0, population = 2.0, simulation duration = 2.0 (the script to compute the hypervolume is available online[10]). Figure 3.3 shows the evolution hypervolume of the Pareto with the number of executed islands. It stagnates after 7000 islands have been executed.

3.2.4 Results

At the end of the evolution we get a file containing: 200 parameter values, the value of the three objectives for each of this points and the number of the samples (or replications) which have been taken into account in the computation of the objective values. In the stochastic resampling strategy candidate solutions are first evaluated with few replications and then promising solutions are resampled (evaluated with more replications), therefore in the resulting file not all solutions have been evaluated with the maximum number of 100 replications. We decide to consider only the most robust solutions in our result analysis (119 solutions among the 200 solutions proposed by the algorithm have been evaluated based on 100 replications). Among these robust solutions, 27 of them produce low (>0.1) objective values for the each of

[8] https://github.com/Geographie-cites/spinger-simpoplocal.

[9] https://github.com/openmole/mgo.

[10] https://github.com/Geographie-cites/springer-simpoplocal.

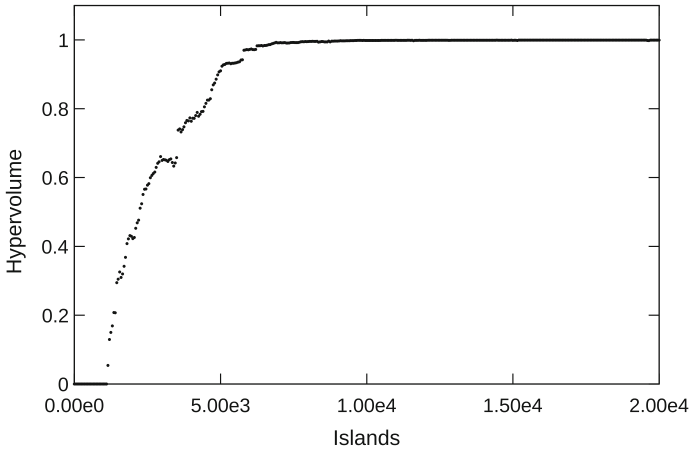


Fig. 3.3 Hypervolume of the Pareto front

the 3 objectives. The fact that the 3 objectives can take be fulfilled altogether means that these objectives are not mutually exclusive (they are compatible).

The figure exposes a run of the model for the set of parameters: *rMax* = 10134.5564655276, *innovationImpact* = 0.0100011820347467, *distanceDecay* = 0.937264929710603, *pCreation* = 0.00000119999472951903, *pDiffusion* = 0.000000879838251240765, *innovationLife* = 1529 whose fitness has been evaluated to *ksValue* = 0.015, *deltaPop* = 0.0129611448, *deltaTime* = 0.002875.

For this set of parameters, the evolution corresponds to what is expected from the model: a progressive and continuous process of hierarchical organization of the settlement system (the slope of the linear fit of the rank–size distribution shifts from 0.2 to 0.9 in 4000 years for a maximum reached size of about 10,000 inhabitants).

Further analysis show that this result is quite robust to stochasticity as shown by the low variability of the recorded final state from one simulation to another exposed on Fig. 3.4.

3.3 Calibration Profiles

In the previous section we have used automatic calibration process based on multi-objective genetic algorithms (Schmitt 2014). Nevertheless, this method produces a reduced set of candidate parameter values that represent optimal trade-off with regard to several model quality criteria. The result of the calibration process is thus solely that the model *can* reproduce the data with a given precision. It does not say anything about how often parameter sets lead to realistic behaviours, and how each parameter will change the behaviour of the model. For instance, it is often interesting to know when some parameter values would prevent the system to reach a realistic behaviour, rather than only knowing a singles set of "optimal" parameter values.

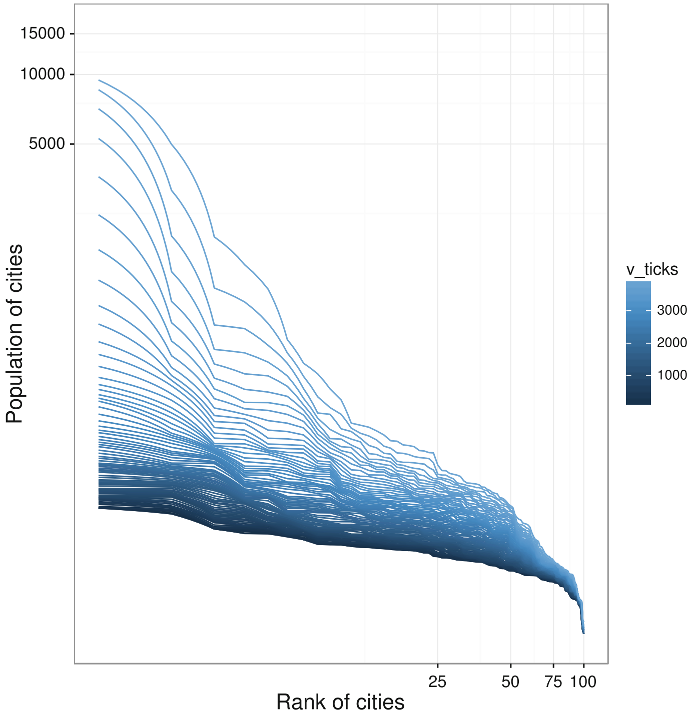


Fig. 3.4 Evolution of the rank-size distribution during a simulation of one of the best calibrated parameter settings

3.3.1 Algorithm

To compute a more global view of the parameter space, we have therefore designed a novel method that exposes the sensitivity of a single parameter on the calibration of a model independently of the other parameters (Reuillon et al. 2015). Given a function which computes a single scalar value depicting a calibration error for the model, the calibration profile algorithm computes the lowest calibration error that can possibly be obtained when the value of a given parameter is fixed and the others are free (Figs. 3.5 and 3.6). It computes this minimal error for many values of the parameter under study. The value of the parameters are sampled all along its domain of definition to produce a so-called *calibration profile*. For each sample value, the value of the remaining parameters are optimized in order to find the lowest possible calibration error. The profile can then be drawn on a 2-dimensional chart that depicts the influence of the parameter under study on the model calibration.

To produce such a profile, a naive approach would consist in executing an entire calibration algorithm for each value of the parameter under study. Current automated calibration algorithms are too computationally intensive to make this approach tractable in practice. To tackle this problem we have designed an algorithm which

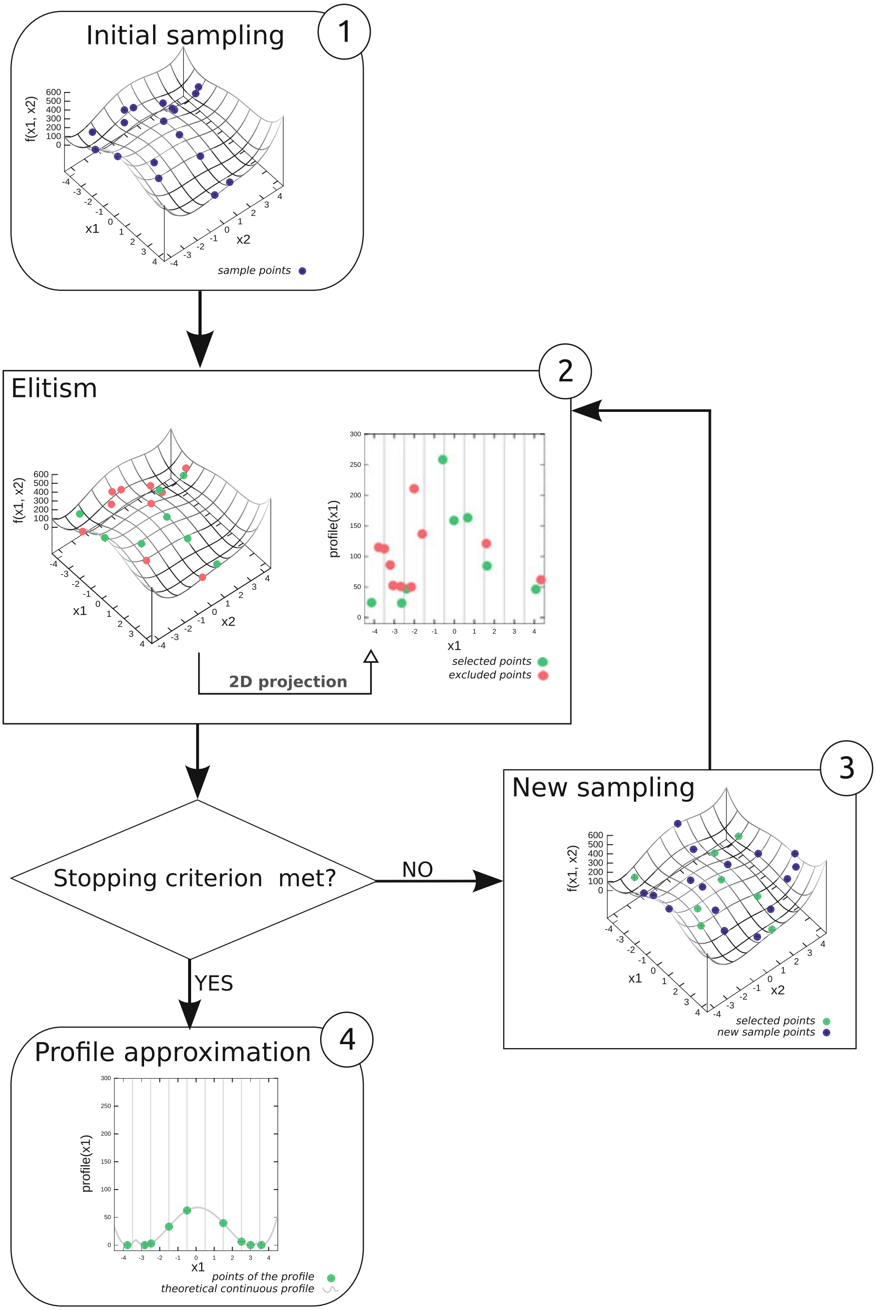


Fig. 3.5 Illustration of the calibration profile (CP) algorithm (Reuillon et al. 2015)

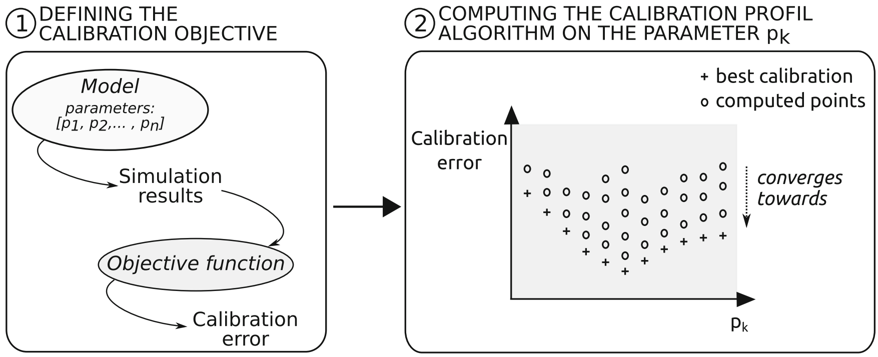


Fig. 3.6 High level representation of the calibration profile (CP) method (Reuillon et al. 2015)

computes the numerous points composing a calibration profile altogether (this algorithm has been inspired by the recently published MOLE method (Mouret 2013; Clune et al. 2013), which computes two dimensional maps of phenotype landscapes using evolutionary algorithms).

This new algorithm has been designed in the framework of evolutionary algorithms. The Fig. 3.5 illustrates the progress of this algorithm through the example of the computation of a 9-points profile along x_1 of a function $f(x_1, x_2)$. In this example, the function f represents a 2-parameters model: x_1 and x_2 and $f(x_1, x_2)$ represents the calibration error of the model. In the step 1 the algorithm randomly samples points (random values of x_1 and x_2) and computes $f(x_1, x_2)$. In the step 2, the algorithm divides the definition domain of x_1 in 9 disjoint even intervals (called niche) and keeps only the sampled point with the lowest $f(x_1, x_2)$ in each of these niches (this constitutes the elitism stage). The points that have been selected constitute a first approximation of the calibration profile of the model f along x_1. In step 3 new samples (x_1, x_2) are generated by mutating the points in the current approximation of the calibration profile and $f(x_1, x_2)$ is evaluated for each of these new points. The newly evaluated samples are merged with the existing ones and the algorithm iterates to the step 2. This iteration stops once a given stopping criterion is met after step 2. The projection of the last selected points along x_1 constitutes an approximation of the theoretical continuous profile (step 4). A detailed description of the algorithm can be found in Reuillon et al. (2015).

The calibration profile is a $\mu + \lambda$ genetic algorithm. It suffers from the same problem regarding stochasticity as the ones described in the previous section. To overcome this shortcoming, we have adapted the "stochastic resampling" strategy to this algorithm (described in the previous section). The deterministic version of CP keeps one single individual for each niche (or interval). To enable the stochastic resampling for CP, we changed this algorithm and made it keep a Pareto front in each niche (by applying the elitism strategy of NSGA2 in each niche). This Pareto front constitutes a compromise between maximizing the number of replications while minimizing the calibration error. Each Pareto front (in each niche) converges towards solutions which are both good and properly evaluated (robust to stochasticity).

3.3.2 Guide of Interpretation

A calibration profile is a 2D curve with the value of the parameter under study represented on the X-axis and the lowest possible calibration error on the Y-axis. To ease the interpretation of the profiles we propose to define an acceptance threshold on the calibration error: under this acceptance threshold the calibration error is considered sufficiently satisfying and the dynamics exposed by the model acceptable, over this acceptance threshold the calibration error is considered too high and the dynamics exposed by the model are considered unacceptable.

The computed calibration profiles may take very diverse shapes depending on the effect of the parameter of the model dynamics, however some of this shapes are recurrent. The most typical shapes are shown on the Fig. 3.7. They have been discriminated according to the variation of the values of the profile compared to the threshold value:

- The shape 1 is exposed when a parameter is restricting with respect to the calibration criterion and when the model is able produce acceptable dynamics only for a specific range of the parameter. In this case a connected validity interval can be established for the parameter.
- The shape 2 is exposed when a parameter is restricting with respect to the calibration criterion, but the validity domain of the parameter is not connected. It might mean that several qualitatively different dynamics of the model meet the calibration requirement. In this case model dynamics should be observed directly

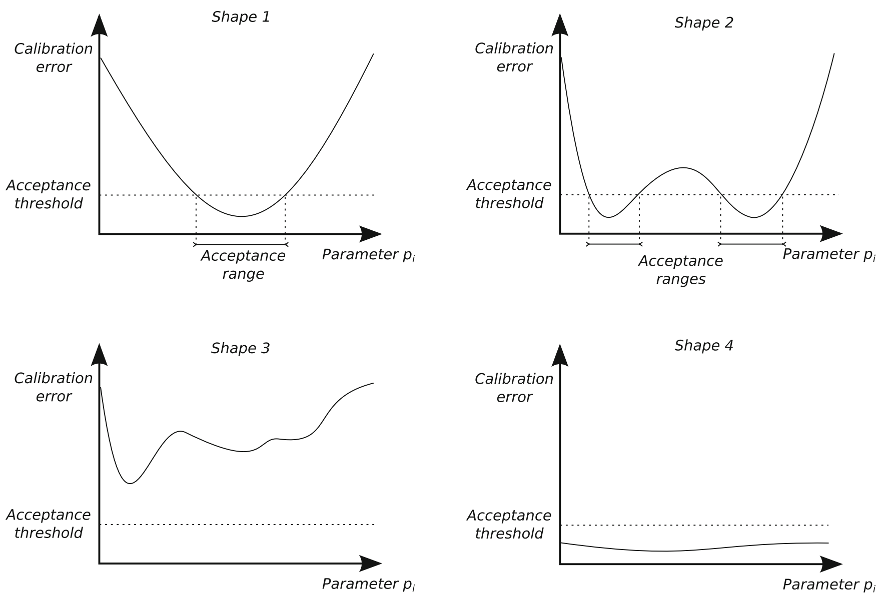


Fig. 3.7 Calibration profile (CP) algorithm (Cottineau et al. 2015)

to determine if the different kinds of dynamics are all suitable or if some of them are mistakenly accepted by the calibration objective.

- The shape 3 is exposed when the model is not possible to calibrate. The profile doesn't expose any acceptable dynamic according to the calibration criterion. In this case, the model should be improved or the calibration criterion should be adapted.
- The shape 4 is exposed when a parameter does not restrict the model dynamics with regards to the calibration criterion. The model can always be calibrated whatever the value of the parameter is. In this case this parameter constitutes a superfluous degree of liberty for the model since its effect can always be compensated by a variation on the other parameters. In general it means that this parameter should be fixed, that a mechanism of the model should be removed or that the model should be reduced by expressing the value of this parameter in function of the value of the other parameters.

3.3.3 Result Analysis

The calibration profile algorithm makes it possible to evaluate the impact of each parameter of SimpopLocal individually on the capacity of the model to produce acceptable dynamics. In Sect. 3.2, we have shown that the 3 objectives used for the calibration (distribution, duration and population) can be fulfilled altogether. To apply the calibration profiles to SimpopLocal, we consider an aggregated evaluation function f defined as the maximum value over the 3 objectives ($f = max(distribution, duration, population)$). If this value is low then a model presents acceptable dynamics. To analyse the produced results, we have established that only input parameters leading to values of f of less than 0.1 of error are considered valid. Indeed, the empirical data and theoretical knowledge that led to the definition of the objective function are not precise enough to justify a more thorough analysis of the model. This threshold is largely exceeded for some parameter values, however rendering this threshold of acceptability explicit enables the definition of credible bounds for each of the free parameter of the model. These bounds define a validity domain of each of the parameters.[11]

The Fig. 3.8 exposes the calibration profile for each of the parameter of SimpopLocal. Several interesting conclusions can be drawn by interpreting them:

- the profile for innovation life exposes that this profile has no significant impact on the capacity of the model to produce acceptable dynamics. This parameter pilots the innovation deprecation mechanism. When $innovationLife = 4000$ that the innovation deprecation time is longer than the simulation time. The facts that we can calibrate the model for this particular value indicates that the deprecation

[11] Note that the profile algorithm iteratively refines the computed profiles from high values toward lower ones through through an iterative process, therefore the proposed bounds are more restrictive than the exact ones.

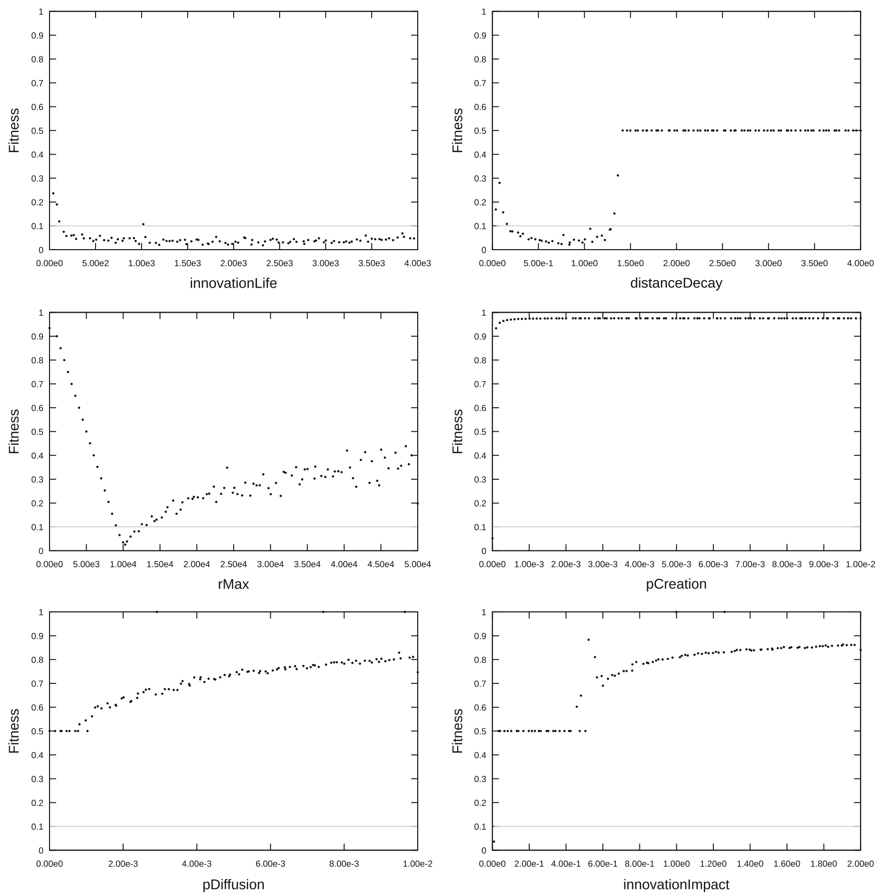


Fig. 3.8 Calibration profiles

mechanism is useless in order to reproduce acceptable dynamics and that the model can be simplified by removing it.

- The profile for *distanceDecay* exposes that the the model cannot be calibrated when the value of this parameter is lower than 0.15 or greater than 1.30. The parameter *distanceDecay* introduces a decaying effect of the distance between settlements on the diffusion of innovations from one settlement to another. When it is high, settlements are isolated from each other (and unable to exchange innovations), when it is low settlements can all exchange equally independently of their respective distances. This profiles shows that the spatial heterogeneity implemented in the model is mandatory to produce acceptable dynamics.
- The profile for *rMax* exposes that the model only produces acceptable dynamics when rMax is close to 10,000 and a best value for $rMax = 10{,}277$. The calibration objective for the size of the biggest settlement has been set to 10,000 inhabitants.

When *Rmax* is lower than 10,000 it is by construction impossible for the model to reach the 10,000 population objective. Low values of this parameter cannot achieve acceptable calibration errors. Surprisingly the calibration algorithm is not acceptable dynamics when *Rmax* is above 10,500. It indicates that this mechanism is necessary to produce acceptable dynamics and reaching settlements of a size matching empirical evidences.
- The 3 other profiles exposes only 1 or even 0 values bellow the acceptability threshold. It means that there have not been observed at the right scale (the parameter range is too broad). In the calibration experiment in Sect. 3.2, we have observed that the model exposes acceptable dynamics when they are close to 0. Therefore, we reduced the range of exploration of these 3 parameters and compute new calibration profiles.

The Fig. 3.9 exposes profiles with narrowed ranges for the 3 parameters *pCreation*, *pDiffusion* and *innovationImpact*. From this 3 curves we can deduce that:

- the validity domain of *pCreation* is between $0.4\ 10^{-6}$ and $2.2\ 10^{-6}$. For low values of *pCreation* the model cannot be calibrated. The innovations are generated too slowly to engender a sufficient growth (whatever the value of the other parameters can be). For high values of *pCreation* the system races and growth is too fast.
- the validity domain of *pDiffusion* is between $0.2\ 10^{-6}$ and $2.1\ 10^{-6}$. A very interesting aspect of this profile is that low values of pDiffusion prevent the model from producing acceptable dynamics. Noticeably when $pDiffusion = 0$ it disables entirely the diffusion mechanism of the model. At this particular point the calibration error is unsatisfying, thus this profile shows that the diffusion of innovation mechanism of the model is mandatory in order to produce realistic behaviours.
- the validity domain of *innovationImpact* is between $6\ 10^{-3}$ and $1.2\ 10^{-2}$. Under the lower bound, the impact on the settlement growth of the innovation is too low and the dynamic is too slow to reach the calibration objectives. On the contrary, when *innovationImpact* is too high the growth of the settlements is too fast to match credible dynamics (whatever the value of the other parameters can be).

3.4 Conclusion

This chapter exposes the evaluation the SimpopLocal model through a novel methodology built on top of the quantitative evaluation of the model dynamics. This methodology is generic and can be reused to other models as long as a quantitative evaluation of the model dynamics can be designed. Furthermore, all the algorithms are available in a reusable form in the free and open-source platform OpenMOLE,[12] which is presented in the Chap. 6 of this book.

The evaluation work presented in this chapter is often perceived as taking place after the modelling process, once the model is finished. On the contrary, we believe that this evaluation work should be carried all along with modelling process, from

[12] www.openmole.org.

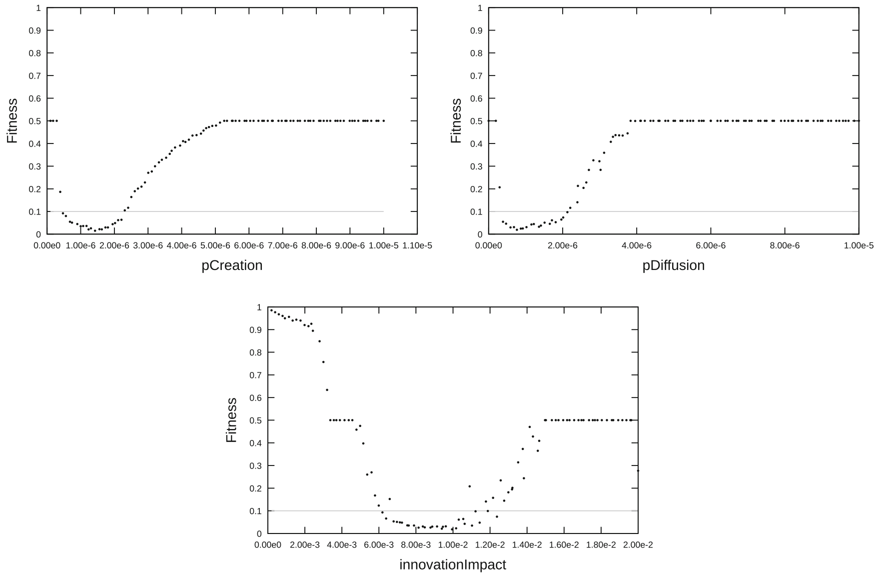


Fig. 3.9 Narrowed calibration profiles

the very early stages. From a conceptual point of view, it has the great advantage of modelling the expectation along with the model mechanisms. From a technical point of view, the quantitative evaluation can guide the modelling choices. Indeed by evaluating face-to-face candidate mechanisms, it is possible to determine which ones are the best fitted to reproduce such or such aspect of the expected dynamic. The next chapter extends this evaluation methodology and proposes an modelling framework guided by the evaluation process.

References

Archaeomedes: Des oppida aux métropoles. Anthropos, Paris (1998)

Bairoch, P.: De Jéricho à Mexico: villes et économie dans l'histoire. Gallimard, Paris (1985)

Beaumont, M.A.: Approximate Bayesian computation in evolution and ecology. Annu. Rev. Ecol. Evol. Syst. **41**(1), 379–406 (2010)

Belding, T.C.: The distributed genetic algorithm revisited. In: *Proceedings of the Sixth International Conference on Genetic Algorithms*, pp. 114–121 (1995)

Branke, J.: Creating robust solutions by means of evolutionary algorithms. Parallel Problem Solving from Nature — PPSN V: 5th International Conference Amsterdam, The Netherlands September 27–30, 1998 Proceedings, pp. 119–128. Springer, Berlin (1998)

Chérel, G., Cottineau, C., Reuillon, R.: Beyond corroboration: strengthening model validation by looking for unexpected patterns. PLoS ONE **10**(9), 1–28 (2015)

Christaller, W.: Die Zentralen Orte in Süddeutschland. Fischer, Jena (1933)

Clune, J., Mouret, J.-B., Lipson, H.: The evolutionary origins of modularity. Proc. R. Soc. Lond. B: Biol. Sci. **280**(1755), 20122863 (2013)

Cottineau, C., Chapron, P., Reuillon, R.: Growing models from the bottom up. An evaluation-based incremental modelling method (EBIMM) applied to the simulation of systems of cities. J. Artif. Soc. Soc. Simul. (JASSS), **18**(4), 9. (2015) doi:10.18564/jasss.2828. http://jasss.soc.surrey.ac.uk/18/4/9.html

Deb, K., Agrawal, S., Pratap, A., Meyarivan, T.: A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. Lect. Notes Comput. Sci. **1917**, 849–858 (2000)

Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning, p. 412. Addison-Wesley Press, Reading (1989)

Johnson, G.: Aspects of regional analysis in archaeology. Annu. Rev. Anthr. **6**, 479–508 (1977)

Kleijnen, J.: Design and Analysis of Simulation Experiments. Springer Publishing Company, Incorporated, New York (2007)

Lenormand, M., Jabot, F., Deffuant, G.: Adaptive approximate Bayesian computation for complex models (2012)

Liu, L.: Settlement patterns, chiefdom variability, and the development of early states in north China. J. Anthropol. Archaeol. **15**(3), 237–288 (1996)

Marcus, J., Sabloff, J.A.: The Ancient City: New Perspectives on Urbanism in the Old and New World. Resident Schoolar Book, vol. 2005. School for Advanced Research, Santa Fe (2008)

Mouret, J.-B.: An algorithm to create phenotype-fitness maps. In: *Proceedings of the Thirteenth International Conference on the Simulation and Synthesis of Living Systems (ALIFE 13)*, pp. 593–594 (2013)

Reuillon, R., Schmitt, C., Aldama, R.D., Mouret, J.: A new method to evaluate simulation models: the calibration profile (cp) algorithm. J. Artif. Soc. Soc. Simul. **18**(1), 12 (2015)

Rudolph, G.: A partial order approach to noisy fitness functions. In: Congress on Evolutionary Computation, Seoul, Korea, pp. 318–325. Press (2001)

Sanders, L.: Regards croisés de géographes, économistes et archéologues sur la hiérarchie des systémes de peuplement: de l'empirie aux systémes complexes. Région et Développement, **36**, (2012)

Sano, Y., Kita, H.: Optimization of noisy fitness functions by means of genetic algorithms using history of search with test of estimation. In: *Proceedings of the 2002 Congress on Evolutionary Computation, 2002. CEC'02*, vol. 1, pp. 360–365 (2002)

Schmitt, C.: Modélisation de la dynamique des systémes de peuplement: de SimpopLocal à SimpopNet. Ph.D. thesis, Université Paris 1 Panthéon-Sorbonne (2014)

Stonedahl, F.J.: Genetic algorithms for the exploration of parameter spaces in agent-based models. Ph.D. thesis, Northwestern University of Illinois (2011)

Tanooka, K., Tamaki, H., Abe, S., Kitamura, S.: A continuous age model of genetic algorithms applicable to optimization problems with uncertainties. In: IEEE International Conference on Systems, Man, and Cybernetics. IEEE SMC'99 Conference Proceedings, vol. 1, pp. 637–642 (1999)

Weise, T.: Global Optimization Algorithms - Theory and Application, 3rd edn. (2011). www.it-weise.de

Chapter 4
An Incremental Multi-Modelling Method to Simulate Systems of Cities' Evolution

Abstract Explaining the evolution of urban systems at large spatio-temporal scales is uneasy. Processes are frequently unobserved empirically and equifinality is a challenge for any generative explanation models. We try to address the causation challenge in urban modelling by proposing a multi-modelling framework for the comparison of several model structures. Each structure represents a combination of mechanisms translating alternative or complementary hypotheses about the processes at play. This approach implies that the conception, implementation and evaluation of the model(s) integrate a diversity of mechanisms. Their contribution to the explanation of urbanization is evaluated in time and space by confronting models to empirical data through an interactive visualization platform. We argue that multi-modelling can provide an alternative way to account for the possible causes generating observed patterns, between traditional approaches such as 1/simple models focusing on a single cause (as is often the case for proving a theory) or 2/very complex models including all possible mechanisms at once (as it might prevent from distinguishing their individual contribution).

4.1 Introduction

Given the complexity of 'real' urban systems, our plea in the introduction was for parsimonious but fully explored models, and for multiple models which account for the equifinality playing in the model (several mechanisms can produce the observed pattern) as well as in reality (without historical and detailed information on the actual processes at work, we only have theoretical hypotheses on how to model cities). These reasons argue for a model-building framework which allows multiple modules to be assembled and combined, each of which represents a particular hypothesis as to the urban dynamics needed to model a particular system of cities at a particular time. This framework should allow the different structures of models to be evaluated in the same way and consequently compared in their ability to simulate an observed trajectory of cities' growth. Because of the diversity of possible factors and theories for explaining the diversity of cities' trajectories, we need a visualization that displays the same kind of information for different model structures and city attributes. The interactive

© Springer International Publishing AG 2017

D. Pumain and R. Reuillon, *Urban Dynamics and Simulation Models*,
Lecture Notes in Morphogenesis, DOI 10.1007/978-3-319-46497-8_4

feature provided by an online (automated) application serves this purpose quite well, and makes it possible for the user to explore different aspects of model results with the same visualization design, thus easing the process of model comparison. Moreover, the exploration of a variety of results and model outputs is guided by a selection of representations used in geography to compare systems of cities, to analyze models and their residuals.

This chapter describes a multi-modelling method that was developed and applied to the system of Soviet and post-Soviet cities (but could be later transferred to any system of cities for which we have historical data). We first present the theoretical, methodological and technical framework of multi-modelling (Sect. 4.2), before detailing the resulting family of models that was developed to simulate Soviet and post-Soviet city growth (Sect. 4.3). Section 4.4 presents the results obtained and the geographical knowledge that one can draw from such a methodology. Section 4.5 describes the application built to explore and communicate these results interactively and online. Section 4.6 concludes.

4.2 Methodological and Technical Framework for Multi-modelling Systems of Cities

In this section, we review the theories competing for the explanation of the evolution of systems of cities (or their stylized facts). We also review previous attempts to build multi-modelling frameworks and expose our own approach.

4.2.1 Complementary and Competing Theories

As stated in the first chapter, systems of cities give rise to very robust regularities over time and space. For instance, the Zipf's distribution of city sizes has been described and studied for almost a century (Lotka 1925 and Nitsch 2005 for a meta-analysis). This 'mystery' (Krugman 1996a) has fostered a wide range of possible explanations, from random processes to economic, social and geographical rationality. If we focus on causal mechanisms (thus excluding random generative models), we can identify five broad categories of explanations, reflecting one or several theories to account for the evolving sizes, locations and functional specialization of cities within a given system (Pumain et al. 2006; Schmitt et al. 2015; Cottineau et al. 2015b):

1. Spatial interactions and the diffusion of innovations (Pumain 1997; Pumain et al. 2006; Pumain 2006) explain the stability of the distribution of city sizes and the functional differentiation by formalizing exchange mechanisms of competition and cooperation between cities that diffuse social, political and economic innovations in a way that gives an advantage to large cities, thus explaining their fastest growth, rank inertia and inner diversity on the long term.

2. Size effects comprise the theories of agglomeration economies and diseconomies (for a review, cf. Rosenthal and Strange 2001). They explain the existence of cities of different sizes by the different possible equilibria between centripetal and centrifugal economic forces. Centripetal forces refer to matching, sharing, learning and sorting advantages of large cities (Duranton and Puga 2004). Centrifugal forces usually refer to congestion and pollution externalities of population agglomerations (Krugman 1996b).

3. Site effects explain the location of cities, and the spatial distribution of growth which is due to an easy access to some localized resources, may they be natural deposits (oil, river, seaside, climate) or social amenities (patrimonial sites, creative atmosphere). The causal mechanism translating this principle is very simple: cities which are located near advantageous resources attract more people and create specific products at a lesser cost, and therefore tend to grow faster.

4. Situation effects such as the one used in location theories (Reynaud 1841; Christaller 1933; Ullman 1941) explain the regular spacing of cities, their size and specialization by looking at the relative accessibility in the system. For instance, hub locations on transportation networks provide advantageous locations for urban growth, as well as large cities because they provide a larger access to a larger pool of products.

5. Territorial effects finally differentiate cities according to the political territories they belong to and look at factors of common evolution enhanced by public policies (fiscal redistribution for example) and shared habits (with respect to natality, for example). It also explains the particular evolution of capital cities by their specific function in the system (Preston 1979; Brockerhoff 1999; Bretagnolle and Pumain 2010).

Theoretically, there is a simple reason why we should try and combine different theoretical (partial) explanations into a unique model: it is to evaluate the explaining power of different hypotheses and of their combination on an empirical case study (Martin 2015). By allowing different accounts to play in the same simulation, we can compare and order different theories, we identify equifinality for the ones performing equally, we spot areas or periods for which some theories work better than others—thus characterizing the genericity and specificity of different hypotheses—and finally we build a composite theory made of existing complementary mechanisms (Thiele 2015).

4.2.2 A Methodology for Implementing Multi-models

Methodologically, there are examples of complexification of the models proposed by agent-based modellers. The pioneers (Epstein and Axtell 1996) indeed proposed a modelling framework of the Anasazis that started from a simple model and added supplementary mechanisms of individual interactions (trade, reproduction, etc.). This incremental approach has been applied later in geographical (Conte et al. 2012) and ecological (Grimm and Railsback 2012) models. At earlier stages of the modelling process, we also have examples of organized reviews of the literature aimed at formalizing the pool of competing theories to account for the pattern to simulate, in organization science (Contractor et al. 2000) and health studies (Auchincloss and Roux 2008; Galster 2012).

However, we do not know of many attempts to combine model structures in the same framework and thorough explorations of multiple model structures against empirical data. Indeed, this is recurrent plea in the literature (cf. Batty and Torrens 2005) for which we provide a proposition. The only example we know of this kind of approach is the pioneer one of S. Openshaw (1983; 1988). His 'model-crunching' method produced a way to select efficient model structures of spatial interactions. However, his pool of alternatives was restricted to different mathematical forms of relating spatial interactions to masses and distances between geographical zones, and led to model structures that were not always meaningful and interpretable. We propose a framework which builds on a consistent set of causal mechanisms drawn from the theoretical literature on systems of cities and which will produce models that we can interpret and use for understanding, explaining and predicting urban systems dynamics.

An initial set of mechanisms is implemented in a programming language that enables their combination (in our case, Scala). A model structure is a certain combination of mechanisms, i.e. the core mechanisms plus additional activated mechanisms. All models in the same family are initialized with the same empirical data and are evaluated with the same measures. Those measures can refer to stylized facts to reproduce (for example, a rank-size distribution of city populations) or to empirical patterns (the actual growth of cities and their hierarchical differentiation).

4.2.3 Exploiting the Results of a Family of Models

To assess the characteristics (performance vs. data, equifinality property, genericity or specificity degree) of the mechanisms of our composite theory, we need to calibrate all model structures with the same criteria—or objective function—(Sect. 4.2.3.1). This requires to identify measures of what a good simulation is with respect to its distance to the observed patterns and empirical data, but also to control for unrealistic dynamics (Sect. 4.2.3.2). The different model structures are then compared according to this measure, and analyzed according to the values of parameters for which the best simulation is obtained. The systematic combination of mechanisms allows to estimate the explaining power of a single mechanism (everything else being equal), this explaining power being measured as how much it reduces the distance to the empirical pattern (Sect. 4.2.3.3).

4.2.3.1 Quantitative Measures to Define a Good Simulation at the Micro-Geographical Level

In order to compare simulated systems with empirical systems of cities with respect to the spatial and hierarchical distribution of growth over time, we compute a measure which sums the distance between the simulated population and the observed population for each city of the system. We sum this distance for each time steps for which we can compare simulated populations with empirical ones (typically, a census year):

$$\delta = \sum_t \left(\sum_i (log(P_{o,i,t}) - log(P_{s,i,t}))^2 \right) \tag{4.1}$$

We use logarithms to compare the impact of relative differences in small cities with differences in large cities, and use the power 2 to give a larger weight to large discrepancies in the sum of distances. Finally, we normalize this index by t the number of time steps for which we can compare simulated populations with empirical ones and by n the number of cities simulated, in order to compare systems with different sizes and simulations of different historical lengths.

We assess the quality of a simulation by looking at how small δ is, considering it has passed micro-behaviour validity tests. We control for unrealistic micro-geographical dynamics by checking for each simulation that there is no city with no wealth and that no city produces and consumes more during a step than the wealth it accumulated over time (for more details, cf. Cottineau et al. 2015a). Taking these three criteria into consideration during the calibration process filters the parameters space of a given model structure, excluding portions which lead to unrealistic behaviours of the model during simulation. The minimization of the distance δ (given the two boolean controls) represents the objective function of the (multi-) calibration.

4.2.3.2 The Multicalibration Procedure

The different modules of the model were combined and calibrated using mixin methods (Steyaert et al. 1993, Lucas and Steyaert 1994, Prehofer 1997) in the object-oriented programming framework of the Scala language. These methods allow the implementation of different alternatives for a single trait (in our case: a mechanism of city interaction or growth) and the generation of a source code containing all the possible combinations (and their dependencies in terms of parameters and variables). To run one of the possible implementations of the model, one has to specify an index referring to the corresponding combination, and a vector containing values for all the possible parameters, even when the given mechanism combination does not make use of some of them. Given this functional way of implementation, the multicalibration thus corresponds to the calibration procedure described in Chap. 3, with an additional parameter corresponding to the model index: therefore the genome of a model defined as a combination of mechanisms contains the vector of all parameter values and the index driving the composition of the model. Models are run on the European Grid Infrastructure and evaluated with respect to the fitness function described in Chap. 3. The only difference with the single-model calibration is that we want results for each possible value of the model index parameter. Therefore, the elitism specification of the calibration algorithm has been transformed to keep the best individuals of each subpopulation (models run with a specific index). The top 50 best performing sets of parameters were kept. The mutation specification of the calibration algorithm has also been tuned to favour a fast convergence: the model index has a 10% chance of mutation. This feature facilitates the exchange of efficient solutions between the different model combinations.

Combining the baseline model with five additional mechanisms for two different time periods (1959–1989 and 1989–2010) resulted in 64 different model implementations (64 values for the model index), approximately 72 million evaluations of which were drawn the best 3200 parameter sets evaluated during the multicalibration (50 for each model instantiation). This database is the one we use to analyze the family of models in the next section.

4.2.3.3 Analyzing the Calibrated Models of the Family

There are three types of analyses that can be drawn from the multicalibrated family of models.

- First, we propose to interpret the overall **performance of the different model structures**, by looking at the shortest distance to the observed pattern obtained for each parsimonious model (the core model plus one additional mechanism). This performance can also be measured as the average distance reduction reached by any model that contains this mechanism compared to models which do not. It means that mechanisms and the theories they formalize can be compared and ordered according to these two criteria, for each spatio-temporal simulation and between territories and time periods. For example, if site effects produce systematically better simulations for a time period, but not in the next one, this process can be said a good candidate for explaining the empirical urban dynamics in the first time span, but another range of explanation might be more relevant to understand the subsequent period.
- Second, for a given structure of model, we propose to interpret the meaning of the **calibrated values of parameters** that give the best simulation. That way, we gain an insight into the strength of different processes and can compare them in different systems (in time or in space).
- Finally, we propose to study the **residuals**, i.e. the cities that cannot be modelled in a satisfactory way with the given structure of model. The magnitude of deviation and the location of those cities tell us about the singularity of their trajectories, that we can try to explain further with by historical events or supplementary explanations. This last analysis is of crucial importance for the geographer as it reveals the areas

of the observed urban evolution that 'resist modelling' (Durand-Dastès 2001) and that suggest the singularity of the realized trajectory of the system.

As we will see in the following sections, such residual trajectories are of particular importance in the Soviet urbanization. However, some cities of this system can be simulated with generic mechanisms and thus the modelling process helps us disentangle the dynamics of cities that are common to other systems and the trajectories of cities that one can only understand if one knows about the history of the Soviet Union.

4.3 A Family of Models of (Post-) Soviet Cities: MARIUS

The application of our multi-modelling framework on the case of (post-) Soviet cities relies on the Simpop principles for modelling cities (Bura et al. 1996; Sanders 2005; Bretagnolle and Pumain 2010) and on a harmonized urban database of 1929 urban agglomerations and their populations over the twentieth and twenty-first centuries (Cottineau 2014a, b). As in the Simpop models, we consider cities as collective agents and model time with 1-year steps. The MARIUS contribution brings up a new way to categorize mechanisms, to order them ex ante given their specificity to the case study (Sect. 4.3.1), as well as a reusable open-source modular implementation (Sect. 4.3.2).

4.3.1 Ordering Possible Causes of Evolution from the Most Generic to the Most Specific

We reviewed five classes of explanation that could account for the regular features of systems of cities in Sect. 4.2.1. They describe systems of cities in general. In the study of a particular system of cities, we expect the realization of general processes to take a particular twist, but we can also expect: 1/other processes to take place, for example political and economic processes shaping the overall geography and affecting cities and 2/the different theoretical processes to appear at different levels of importance in the empirical mix. Indeed, the Soviet and post-Soviet cities exhibit some of the general features of systems cities: a hierarchy of city sizes that follows a power law, the increase of size inequality between cities over time, the spacing of cities in the inhabited space, a specialization of functions and economic interactions. However, we identified empirically the territorial immensity, the importance of subsurface resources and the planned nature of some of the economic interactions during the Soviet Union to be particular and singular features affecting the location and growth rates of specific cities (compared to the generic structure predicted, cf. Cottineau (2014b). With this particular knowledge in mind, we identified and ordered the mechanisms that we think are at play in the evolution of Soviet and post-Soviet cities. We also distinguished between mechanisms as to those which imply interac-

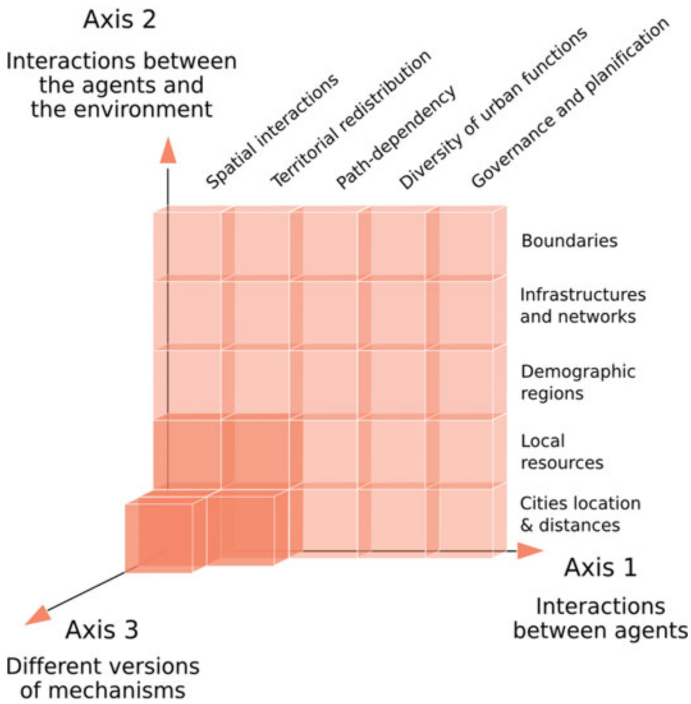


Fig. 4.1 Ordering possible causes of urbanization in the (post-) Soviet case

tions between cities and those which include interactions between cities and their environment. There might be several ways of implementing each of them, so we organized our path of model particularization into three axes (Fig. 4.1).

Axis 1 comprises mechanisms of interurban interactions. The first of this kind, that we think the most generic yet important to model cities in the post-Soviet space, corresponds to the theory of spatial interactions. The second one refers to territorial effects and consists in a fiscal redistribution between cities of the same political region. Other mechanisms, more and more specific to the system under study, include the path dependency and lock-in of interactions' networks, the economic specialization (and monopoles for example) and the planning policy.

We place on axis 2 the mechanisms that formalize rules of interactions between city agents and their broader geographical environment, such as the specification of spatial interactions by actual distances, site effects encompassed in the extraction of localized resources (a general mechanism, yet of particular importance for understanding the contemporary economy and location of growth in Russia and Central Asia). The imperial construction of the Soviet Union makes us consider that the different demographic regions and their differentiated paces of urbanization played a particular role of situation and territorial effects in the trajectories of Soviet and post-Soviet cities. Situation effect mechanisms such as the accessibility by transportation networks might be of singular relevance to the huge territory of the Soviet Union. Finally, the role of open and closed boundaries of the system appears very singular to this case, but a strong amplifier of territorial effects during the Soviet period, compared to other systems of cities.

The third axis of alternative implementations of the same conceptual processes has been exploited here by allowing different mechanisms to represent spatial interactions, with different levels of complexity.

4.3.2 Implementing Modular Mechanisms

The different implementations of mechanisms have been described in detail in (Cottineau et al. 2015b). In this chapter, we will only outline the main features of the mechanisms actually implemented and evaluated as part of the multicalibration.

4.3.2.1 Size Effects and Spatial Interactions: The Baseline Model

The baseline model includes basic features of cities: their population initialized at the empirical starting point of the simulation, and a wealth estimated as a power law of this population, with a parameter populationToWealth ranging from 1 (no economic size effect) to 2 (larger cities are wealthier). In this baseline model, each city produces and consumes as a power law function of its population at each time steps, with two parameters sizeEffectOnSupply and sizeEffectOnDemand ranging from 1 (no productive/consumptive size effect) to 2 (larger cities are increasingly productive/consumptive per capita) and a normalizing parameter economicMultiplier. Each city then proceeds to an estimation of potential exchanges of value with other cities based on their respective size and distance, following a gravity model of distance exponent distanceDecay ranging from 0 (no distance effect) to 2 (the interaction between cities decreases faster than proportionately with the distance between them). It then shares its supply (/demand) between potential city clients (/city providers) and updates its wealth by adding the amount produced during the current time step, subtracting the total demand, adding unsatisfied demand and subtracting the unsold supplies during the external exchange round. The conclusive operation of a simulation step involves translating the wealth differential into a population gain (or loss), using a power law of exponent wealthToPopulation between 0 and 2.

This simple baseline model has proved unsatisfactory by itself to model the evolution of Soviet cities but other implementations of spatial interactions were shown necessary and sufficient to do so (Cottineau et al. 2015b). Such implementations included a transactional bonus mechanism and a fixed cost of transaction mechanism.

4.3.2.2 Spatial Interactions: The Bonus Mechanism

The bonus mechanism models positive externalities[1] of external exchanges of cities (compared to internal production for internal consumption within a city). It simply adds to the wealth update a term B_i, which is a positive function of the volume traded by a city i to all its urban partners, and the number of cities with which it interacted (relatively to the total number of cities n). When this mechanism is activated, it creates an alternative implementation of the spatial interactions baseline model (cf. axis 3 of Fig. 4.1).

4.3.2.3 Spatial Interactions and Situation Effects: The Fixed Cost Mechanism

The fixed cost mechanism complements the spatial interactions baseline model by including a condition on the realization of exchanges between cities after the computation of interaction potentials. The new rule states that this potential needs to exceed a value fixedCost because each exchange generates transaction costs (Spulber 2007). If the trading potential between two cities is under this value, because of their small size and/or large distance, they will not interact. Otherwise, they will share their supply and demand over the remaining set of potential partners as in the baseline model. During the wealth updating step, each city will subtract the value of fixedCost as many times as the number of transactions it was involved in. When this mechanisms is activated (and when it is activated along with the bonus mechanism), it creates an alternative implementation of the spatial interactions baseline model (cf. axis 3 of Fig. 4.1).

4.3.2.4 Site Effects: The Resource Mechanism

Site effects in MARIUS are understood as subsurface resources. Natural deposits are long known to be favoured locations of growth (Reynaud 1841), but their abundance in the Soviet area makes it a relevant choice for explaining the spatial distribution of growth. Resources can be of two types: coal and hydrocarbons. The location of deposits is initialized empirically from observed patterns, and cities with access to each of these resources are given an extracting advantage that depends on their total wealth (a proxy for the capital they can invest in extracting the resource locally). This mechanism thus has two parameters: coalEffect translates the percentage of wealth added at each time step for cities located on coal deposits (by comparison with cities located elsewhere), and oilAndGasEffect translates the percentage of wealth added at each time step for cities located on oil and gas deposits (by comparison with cities located elsewhere). Both range from -1 (the site has negative externalities on cities' wealth) to 1 (the site has positive externalities on cities' wealth), with 0 corresponding

[1] Accounting for knowledge spillovers for example.

to the absence of site effects. This mechanism represents the first increment that we think specific to the Soviet system with respect to the interactions between cities and their environment (cf. axis 2 of Fig. 4.1).

4.3.2.5 Territorial Effects: The Redistribution Mechanism

Territorial effects are implemented in MARIUS as a redistribution of wealth within regions and within countries. At the beginning of a simulation step, cities of the same territory mutualize a share territorialTaxes (from 0 to 1) of their wealth. From this amount, the capital city raises a share capitalShareOfTaxes (from 0 to 1) to sustain its administrative duty. The remaining amount of money is redistributed to every city according to its size (in population). The balance of this redistribution is included in the update of wealth at the end of the simulation step. This mechanism represents the first increment that we think specific to the Soviet system with respect to the interactions between cities (cf. axis 1 of Fig. 4.1).

4.3.2.6 Territorial and Situation Effects: The Urban Transition Mechanism

In this second increment relating to the interaction between cities and their environment, we formalize uneven opportunities of rural immigration for cities of different regions by modelling a logistic curve of the urbanization rate and locating each region on this curve given its level of urbanization at the initial date of the simulation.[2] At each time step, the region moves from one unit on the relative urbanization time, and reach a higher urbanization rate. The cities which belong to each region have an extra growth of population due to rural migration that is a negative function of the urbanization rate. This function is normalized by a parameter ruralMultiplier which possibly ranges from 0 (no migration) to 1 (the population is doubled by rural migrants).

All these increments are combined into 64 model structures that we have calibrated over two periods of time: the Soviet stable era (1959–1989) and the post-Soviet transition (1989–2010).

4.4 Geographical Insights on (Post-) Soviet City Growth from Multi-modelling

By looking at mechanisms' performance, corresponding parameters and residual trajectories, we hope to understand better the probable drivers of urbanization before

[2]The logistic curve was estimated on empirical urban and regional demographic data between 1959 and 2010 for 108 regions of the Former Soviet Union (cf. Cottineau 2014b).

and after the crash of the USSR, and to compare the power of different theories in this explanation. Simulations of the Soviet period correspond to models of 30 steps from 1959 to 1989, with an initialization of 1145 cities at their empirical population in 1959 and 3 census check-up dates (1970, 1979 and 1989) for the evaluation. In other words, in Eq. (4.1), $t = 3$ and $n = 1145$. Simulations of the post-Soviet period correspond to models of 21 steps from 1989 to 2010, with an initialization of 1822 cities at their empirical population in 1989 and 2 census check-up dates (2002 and 2010) for the evaluation. In other words, in Eq. (4.1), $t = 2$ and $n = 1822$.

All the results presented come from the multicalibration of the 64 model structures, evaluated with the open database DARIUS on post-Soviet agglomerations,[3] after 400,000 generations of a generic algorithm which objective function was to minimize the distance δ while meeting the microdynamics controls, using parallel computing through OpenMOLE.[4] These results can be explored and replicated within an online application called VARIUS.[5] The point of this section is more about the geographical insights and knowledge that are gained through multi-modelling.

4.4.1 Mechanisms' Performance

From the pool of 64 model structures calibrated for each time period, we first look at the best performance achievable (the controls for realistic dynamics being met) and the corresponding model structure for the given period. Between 1959 and 1989, this best performing model corresponds to a complete model (all mechanisms are active), minus the resource mechanism. The normalized distance to empirical data amounts to 0.0123. Between 1989 and 2010, the best performing model corresponds to a complete model (all mechanisms are active), minus the bonus mechanism. It amounts to a normalized distance to data of 0.0041. These results confirm the intuition that the differentiated urbanization process might more probably be the consequence of a **mix of effects (or partial explanations)**—site, situation, size, territory and interaction—than the result of a single mechanism. Thus, more complete models simulate better the trajectory of all cities in the system (they also have more degrees of freedom during the calibration, some mechanisms balancing others). These first results also show that the **dynamics of post-Soviet cities are on average three times easier to model than the trajectory of Soviet cities**. Is this evidence of some 'normalization' of the economic and political system or does it only attest the low population growth (and even demographic shrinkage on most of the post-Soviet territory) of the last 20 years? We cannot say at this point. However, we can observe that the location of resources and the way we modelled site effects do not help reproduce cities' trajectories before the transition (as the best model performs without this mechanism). This reinforces the empirical impression of diversity of trajectories in neighbouring locations at this

[3]http://dx.doi.org/10.6084/m9.figshare.1108081.

[4]http://openmole.org/.

[5]http://shiny.parisgeo.cnrs.fr/VARIUS.

period Cottineau 2014b. However, this mechanism of resource extraction seems to be an important candidate for explaining trajectories during the next period. However, bonified interactions between cities tend to increase simulated deviations from the observed urban trajectories. We see two interpretations to this result: 1/ technically, the bonus parameter permits to model larger demographic growth (cf. Sect. 4.3.2.2), and is thus not required at a time of demographic shrinkage) and 2/ the absence of spillover 'bonuses' from interurban exchange might mirror the new localism of post-Soviet urban strategies, which are less prone to deal with distant, uncertain and costly suppliers from within the Former Soviet Union, but also rely more on the wealth from subsurface resources and/or international partners (Europe, China, Middle East, etc.).

Another way to look at the **performance of single mechanisms** is to compare model structures composed of the baseline model plus a single additional mechanism. For the first period, the best performing parsimonious such model structure involves the **Urban Transition** mechanism. In this case, the normalized distance to empirical δ amounts to 0.0142, which is just over the best performing complete model (δ = 0.0123), but in the same range of performance. For the post-Soviet period, the best parsimonious model performs 25% worse ($delta = 0.0052$) than a more complete model ($delta = 0.0041$), with the **Resource** mechanism only. This indicates a possible shift in the main drivers of differentiation of urban trajectories before and after the transition. During the late Soviet Union, difference in rural migrant potentials would be the most important criterion to distinguish fast growing cities from more steady relative trajectories. Territorial and temporal lags in the urban transition would have been the important determinants of the evolution of cities in the different parts of the former empire. In the post-Soviet New Independent States, on the other hand, the access to important resources such as oil and gas would explain much better the contrasted destinies of population growth and economic dynamics of cities.

Overall, we then see on Fig. 4.2 that the alternative implementations of the spatial interaction mechanism (Bonus and Fixed Cost, appearing as the Bonus_true and Cost_true in the bars of the Fig. 4.2.) contribute significantly to the reduction of the distance to observed trends, as well as the mechanisms complexifying the environment with which cities interact: Resources and Urban Transition (although differently for the two periods). The Redistributive mechanism is not significant (at a threshold of 0.5% of statistical error) in this average reduction. Finally, everything being equal with respect to the structure of the model, applying it to the latest period gives much better simulations.

We cannot make further comments on the modelled dynamics that simulate (post-) Soviet cities most satisfactorily without looking at the values of the parameters calibrated for a given model structure.

4.4.2 Parameter Values

To simplify the analysis, let us focus on the best performing models for each period, with a mix of four additional mechanisms each. The study of their calibrated parame-

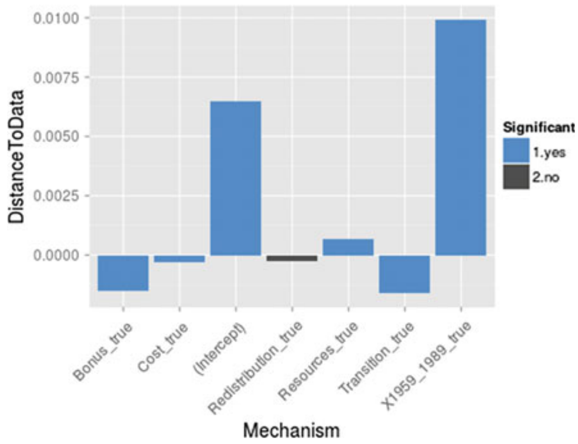


Fig. 4.2 Estimating the contribution of each mechanism to a 'good simulation'. Each bar plots the coefficient estimated during a regression of the normalized distance delta to empirical delta. The intercept gives the average delta for the baseline model during the period 1989–2010. The other bars correspond to the specific contribution of each mechanism, or the average surplus of distance for the period 1959–1989. The *colour* of the bar indicates if the coefficient estimated is significant or not in the regression ($pvalue < 0.005$ in *blue*)

Parameter	Associated Mechanism	1959-1989	1989-2010
populationToWealthExponent	Baseline Model	1.009626528	1.045500
economicMultiplier	Baseline Model	0.459641548	0.5981786
sizeEffectOnSupply	Baseline Model	1.000997749	1.008078
sizeEffectOnDemand	Baseline Model	1.073854144	1.000000
distanceDecay	Baseline Model	0.139061393	0.3090559
wealthToPopulationExponent	Baseline Model	0.388314458	0.6477288
bonusMultiplier	Bonus	79.713516669	-
fixedCost	Fixed Cost	0.370626676	100.0000
oilAndGazEffect	Resource	-	0.03939965
coalEffect	Resource	-	0.001998449
territorialTaxes	Redistribution	0.008163108	0.09916132
capitalShareOfTaxes	Redistribution	0.993792729	0.2643902
ruralMultiplier	Urban Transition	0.019761433	0.001722130

Fig. 4.3 Calibrated parameters of best performing model structures for two periods

ters (Fig. 4.3) reveals insightful variations of the effect of the different mechanisms needed to simulate two sets of very different historical urban dynamics.

The higher value of the parameter populationToWealthExponent for the initialization of the second period (i.e. superlinear scaling of wealth with population compared to the linear relation of the precedent period) indicates a higher economic inequality between cities with respect to their size at the beginning the post-Soviet era, which is necessary to simulate observed trajectories under the modelled assumptions.

Size effects on yearly production and consumption behaviours are almost insignificant for the two periods (sizeEffects $\sim$1). The exception relates to consumption during the Soviet Union, which appears superlinear with population: large cities generate a higher demand per capita during this period under the modelled assumptions.

The reducing effect of distance on potential interactions (distanceDecay) is low in this huge country, compared to empirical estimations on France and the UK (Fotheringham 1981; Baccaïni and Pumain 1998) but doubles over time, from 0.14 to 0.31, suggesting a decrease of large distance transactions under market conditions. This hypothesis is strengthened by the disproportionate increase of the value of the fixedCost parameter. As it represents the threshold under which potential interactions are not profitable, it renders a picture of exchanges limited to large volumes between large and neighbouring cities. This framework fits with the descriptions of metropolization and localism within the New Independent States after the transition and under globalization processes.

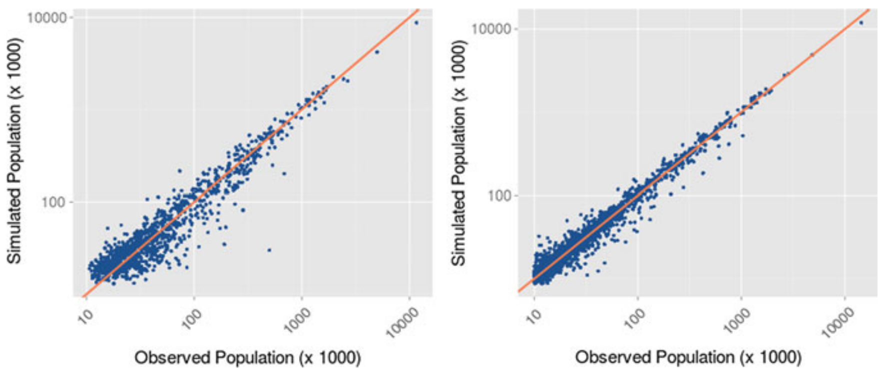
The resource effect which is significant in post-Soviet urbanization patterns is due to oil and gas deposits, generating a surplus of growth equivalent to 4 points in percentage to cities located there, every year between 1989 and 2010. By comparison, coal cities benefit from twenty times less boosting effect, a consequence of the obsolete economic cycle of this resource.

Redistribution is almost absent from the first period's model ($<1\%$ of wealth is taxed, transferred directly to the capital city) but constitutes an important factor of equalization thereafter.

Finally, rural migration appears ten times less important to explain urban trajectories in the second period.

4.4.3 Residual Trajectories

We end the analysis of the best performing models at each time period by looking at the cities that resist modelling, i.e. the urban trajectories which the implemented mechanisms do not succeed in simulating. In particular, we look at the global disper-



(a) 1959-1989 (b) 1989-2010

Fig. 4.4 Global dispersion of residuals for each period

(a) 1959-1989

(b) 1989-2010

Fig. 4.5 Spatial distribution for maximum residuals for each period

sion of residuals (Fig. 4.4), their spatial distribution (Fig. 4.5) and singular trajectories (Fig. 4.6).

The global dispersion of residuals indicate that the most recent period is best simulated by the model, and that in general, the most striking outliers correspond

1959-1989			
City	**Republic**	**Observed Population**	**Simulated Population**
Naberezhnye Tchelny	Russia	500,000	30,000
Volgodonsk	Russia	191,000	35,000
Chajkovskij	Russia	86,000	20,000
Balakovo	Russia	197,000	53,000
Bratsk	Russia	285,000	82,000
...	...	...	...
Zaozernyj	Russia	16,000	52,000
Kizel	Russia	37,000	108,000
Cheremhovo	Russia	74,000	215,000
Gremjachnsk	Russia	21,000	56,000
Gornoaltajsk	Russia	46,000	113,000
1989-2010			
City	**Republic**	**Observed Population**	**Simulated Population**
Mirnyja	Russia	41,000	12,000
Sertolovo	Russia	48,000	16,000
Beineu	Kazakhstan	32,000	11,000
Govurdak	Turkmenistan	76,000	28,000
Serdar / Gyzylarbat	Turkmenistan	98,000	37,000
...	...	...	...
Sovetabad	Uzbekistan	11,000	32,000
Zhanatas	Kazakhstan	21,000	50,000
Krasnozavodsk	Russia	13,000	31,000
Gagra	Georgia	11,000	25,000
Nevelsk	Russia	12,000	25,000

Fig. 4.6 Cities that deviate most from their simulated trajectories

to cities which are much more populated ’in real life’ than what the model predicts. In the model for 1959–1989 for example, such cities include Naberezhnye Chelny or Volgodonsk, respectively, 16 and 5 times bigger than expected. These cities were indeed flagship industrial projects of the Stalin era, in the automobile and power industries.

On the other hand, negative residuals like Sovetabad or Zhanatas in the most recent model correspond to cities which shrank or grew less than expected given their locations, attributes and predicted interactions. In the post-Soviet Uzbekistan or Kazakhstan, they can be cities deserted by Russian migrants after the crash of the USSR. The models of urban interactions of the MARIUS family are thus not designed to simulate such paths and historical events.

4.5 VARIUS: A Visual Aid to Model Composition and Interpretation

The challenge of analyzing and communicating processes and results of geographical modelling, especially in the context of multi-modelling, calls for effective methods of visual representation (Batty et al. 2011). Indeed, we want to describe the urban evolution and to explore the adequacy of several simulation models to reproduce this evolution. These aims imply handling large spatio-temporal quantitative datasets and comparing their features with the realized trajectory of the empirical system of Soviet and post-Soviet cities. Visual representations, such as graphics and maps, seem to be the simplest and most powerful way to do so (Tufte 2001), and the ideal basis for argumentation and geographical interpretation. Indeed, the visual representation of models performance and the distribution of spatial residuals provide supplementary elements for face validation of the simulation models (Hermann 1967), beyond quantitative measures used in the automated calibration process.

VARIUS (http://shiny.parisgeo.cnrs.fr/VARIUS/) provides a platform for interactive exploration of models which complements MARIUS model building, and allows to share and open the exploration of simulated urban trajectories online. Indeed, any user can run the combination of mechanisms and parameterization of their choices, and visualize the resulting urban evolution. Opening the black box of model building is necessary for collective model validation and can be eased by the provision of predesigned tools for exploration, besides the open-licencing of data, models and codes[6].

4.5.1 Building the Model Online

The first part of VARIUS application ('What happened?/Census data') consists in a quick presentation of quantitative evidence about the system to simulate that will help selecting the most relevant set of mechanisms to analyze. Basically, it represents the content of the DARIUS database, and represents urban demographic structures in time and space.

- The first interactive map shows the population of all cities in the Former Soviet Union at the chosen date (left chooser, the right slider adjusting the size of circles, Fig. 4.7). This interactive map therefore shows the spatial and hierarchical distributions of cities in the post-Soviet space from the first Russian census in 1897 to the last in 2010 censuses are like transversal photographs of urbanization taken at irregular points in time. To study an evolution between these photographs, the second map provides a more dynamic approach.
- The second interactive map proposes two choosers representing the starting (T) and the stopping date ($T + P$) of a period P under investigation. It relies on the computation of the average annual population growth rate g of cities i during this period P (Eq. 4.2). Average annual growth rates ease the interpretation and comparison of growth and shrinkage trends across irregular periods (typically, intercensus intervals). As a result, maps produced in this section reveal the spatial and hierarchical distribution of growth (red) and shrinkage (blue).

[6] https://github.com/ClementineCttn/VARIUS.

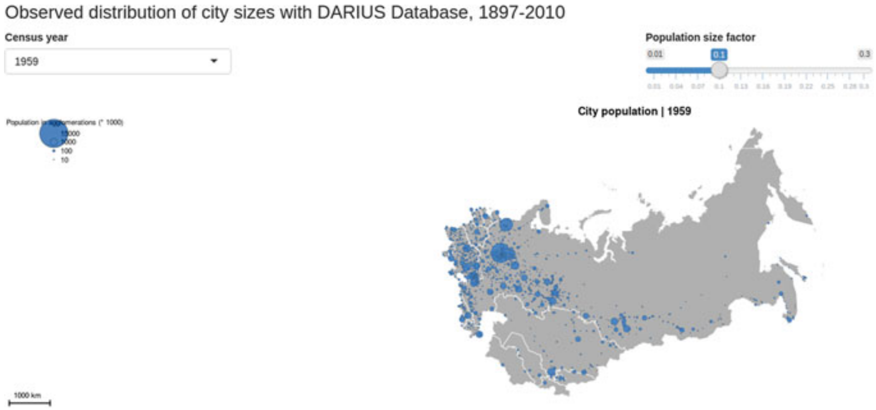


Fig. 4.7 Exploration of the empirical data. Size of *circles* can be adjusted (*top right slider*) and the year of census for which the data are displayed is selected via the *top left* menu

$$g_{i,T,P} = \left(\sqrt[p]{\frac{Population_{i,T+P}}{Population_{i,T}}} \right) * 100 \tag{4.2}$$

- A third interactive map represents the spatial distribution of fixed categories of cities (statuses, access to resources, locations). Seven attributes are available to display: the status of national and regional capital, 342 mono-industry towns as defined by the Russian Federation in 2013[7]; the location in areas of coal and hydrocarbon extraction; accessibility by rail and by air, and absolute east/west location.

In the part ‘How to Simulate it’, VARIUS provides tools to analyze the modular structure of the MARIUS family of models, and estimates the contribution of different mechanisms and their combination to the reduction of discrepancies between observed and simulated urban trajectories. Two approaches are offered to this estimation. First, a linear regression of the fitness measure performed on all calibrated models allows the users to identify by themselves the type of model that they want to run, activating or deactivating the mechanisms that they find interesting for their performance (or underperformance). The second approach yields an optimal combination of mechanisms for simulating the observed evolution in the FSU, given the number of mechanisms to combine (if one seeks parcimony above all). The platform then goes on to providing a parameterization board to run the model online.

[7] http://www.veb.ru/common/upload/files/veb/br/mono/list342.pdf.

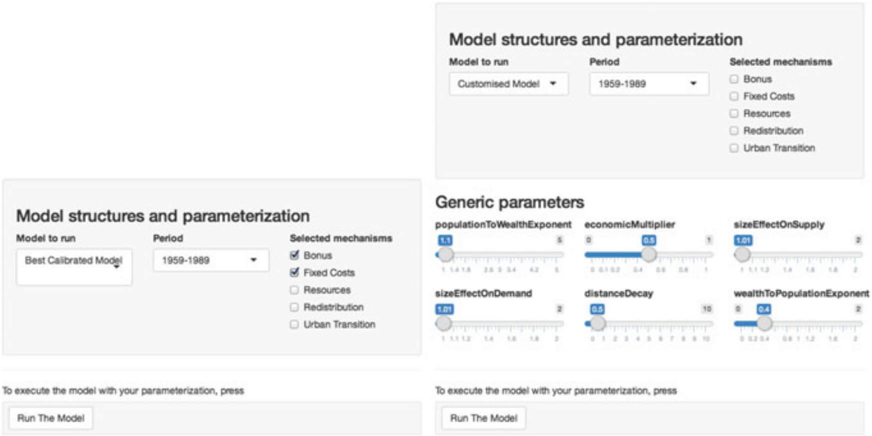


Fig. 4.8 Parameterization board for running any MARIUS model

4.5.2 *Running the Model Online*

The central part of the application gives access to the model itself, by making it possible to run it online, given a particular set of mechanisms and parameters. The mechanism structure and parameter values necessary to execute the model are defined in the tab 'Run a MARIUS model', and can be set automatically as a result of previous calibration, or explored manually by the user. By choosing the option 'Best calibrated model' (Fig. 4.8) in this tab, the user focuses on best performing models only, and explores their performance according to time periods (before or after the dissolution of the USSR) and different mechanisms' combinations. If one does not focus on precalibrated models but seeks to explore the effect of single parameters on the model's behaviour and simulated patterns, it is possible to run a 'Customized model' instead and to define manually its parametrization. With this option, the user can choose a value for each parameter in the intervals considered realistic and interesting (the ones used for the calibration process). For the baseline model, six parameters need to be defined, whereas additional mechanisms include one or two parameters each.

4.5.3 *Analyzing Results Online or 'How Close Are We?'*

The analyzis of model simulations is proposed at three scales: the macro geographic level of urban hierarchy, the micro level of cities, and a meso level of categories of cities, based on their function or status.

- At the macroscale, two visualizations help to explore the quality of a simulation of city sizes. In the first one, cities are ordered by population and plotted against their ranks in the system in a Zipf tradition. The simulated hierarchy of cities (described by the slope of the curve and its deviation from a straight line) can be compared with empirical observation of over time (in grey). Being able to reproduce this pattern is a basic requirement for any model to achieve. It means that the distribution of city sizes match the observed one at the last date, but also the evolution towards an increasing unevenness over time. The second representation displays the dispersion of residuals (as in Fig. 4.4), that is the distribution of simulation errors for each city at the last step of simulation. A model with a perfect fit would be characterized by a distribution of blue dots along the orange line, meaning that every city's simulated population equals its observed population. This limit case is not one we hope to achieve with parsimonious models. However, models are considered good enough if they exhibit a small and symmetrical dispersion around this line. This representation helps spotting outlier cities that the model is not suited to simulate (the dots significantly away from the identity line).
- Outliers are the object of representation of the next tab in the application, at the micro level of cities. Indeed, the maps plotted here show cities whose simulated trajectories deviate most from the empirical one. Using appropriate thresholds and looking at the spatial distribution of these cities at the different points in time, the user is given a glimpse of the spatial and hierarchical pattern of residuals. This sometimes gives way to hypotheses (e.g. the growth of large cities is underestimated by the simulation) that can be tested in the final section.
- At a meso level of groups of cities, this multiple regression aims to profile residuals according to the attributes of cities. We consider this last section as the beginning of a new reflection to complement the model with new mechanisms. The intuition for these new mechanisms would come from the observation of a semi-general feature non-included in the current modules but shared by a large group of cities displaying the highest residuals.

This application is therefore a communication tool for the work done on MARIUS as much as a basis for a work-in-progress regarding model building.

4.6 Conclusion

Because the family of models is designed as a modular framework and because methods were developed to handle modular models in the evaluation processes, the expansion of the model via new mechanisms or its transfer to different urban contexts is made easier and more straightforward, reducing the development cost to the new mechanisms to implement only.

This is a great step forward in the conception of the family of simulation models for several reasons.

In terms of model design, the contribution of a component can be soundly confronted to the others: the introduction of a new component in the family of models leads to new explorations of models dynamics, producing measures that assess both the performance of the new mechanism, but also the performance of the rest of the mechanisms when they interact with it. Each mechanism addition thus reinforces the confidence we have in the mechanisms of the model family, refining the conditions under which they perform best, in an entirely tractable and comparable way.

Furthermore, exploring the models structures allows orienting the model design either towards parsimony or specificity, which is a great support when it comes to strengthening a composite theory where several stylized facts may interfere depending on the chosen level of granularity.

Finally, the fact that a family of models produces comparable and validated trajectories (with respect to their mechanisms, parameterization, and the data against which they are calibrated) would make prospective outcomes, such as territorial policies scenarii, more directly interpretable. This knowledge thus enables to develop policies that are especially adapted to local situations.

References

Auchincloss, A.H., Roux, A.V.D.: A New Tool for Epidemiology: The Usefulness of Dynamic-Agent Models in Understanding Place Effects on Health. Am. J. of epidem. **168**(1), 1–8 (2008)

Baccaïni, B., Pumain, D.: Les migrations dans le systéme des villes françaises de 1982 á 1990. Population **53**(5), 947–977 (1998)

Batty, M., Vargas-Ruiz, C., Smith, D., Serras, J., Reades, J., Johansson, A.: Visually intelligible land use transportation models for the rapid assessment of urban futures. CASA Working Paper vol. **163** (2011)

Batty, M., Torrens, P.M.: Modelling and prediction in a complex world. Futures **37**(7), 745–766 (2005)

Bretagnolle,.A., Pumain, D.: Simulating urban networks through multiscalar space-time dynamics: Europe and the united states, 17–20th centuries. Urban Stud. **47**(13), 2819–2839 (2010). 00043

Brockerhoff, M.: Urban growth in developing countries: a review of projections and predictions. Popul. Dev. Rev. **25**(4), 757–778 (1999). 00067

Bura, S., Guérin-Pace, F., Mathian, H., Pumain, D., Sanders, L.: Multiagent systems and the dynamics of a settlement system. Geogr. Anal. **28**(2), 161–178 (1996). 00160

Christaller, W. (1933). *Die zentralen Orte in Süddeutschland: eine ökonomisch-geographische Untersuchung über die Gesetzmässigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen*. Jena, Fischer. 00316

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.: Manifesto of computational social science. Eur. Phys. J.: Spec. Topics **214**, 325–346 (2012)

Contractor, N., Whitbred, R., Fonti, F., Hyatt, A., O'Keefe, B., Jones, P.: Structuration theory and self-organizing networks. In: Organization Science Winter Conference, Keystone, CO (2000) 00017

Cottineau, C.: DARIUS Database (2014a)

Cottineau, C.: L'évolution des villes dans l'espace post-soviétique. Observation et modélisations. PhD thesis, Université Paris 1 Panthéon-Sorbonne (2014b) 00001

Cottineau, C., Chapron, P., Reuillon, R.: Growing models from the bottom up. an evaluation-based incremental modelling method (ebimm) applied to the simulation of systems of cities. Journal of Artificial Societies and Social. Simulation **18**(4), 9 (2015a)

Cottineau, C., Reuillon, R., Chapron, P., Rey-Coyrehourcq, S., Pumain, D.: A modular modelling framework for hypotheses testing in the simulation of urbanisation. Systems **3**(4), 348 (2015b)

Durand-Dastès, F.: *Les concepts de la modélisation en analyse spatiale*. Paris, Hermés, Lavoisier (2001) 00000

Duranton, G., Puga, D.: Micro-foundations of urban agglomeration economies. Handbook Reg. Urban Econ. **4**(2063–2117), 01533 (2004)

Epstein, J.M., Axtell, R.L.: Growing Artificial Societies: Social Science from the Bottom Up. MIT press, Cambridge (1996)

Galster, G.C.: The mechanism(s) of neighbourhood effects: Theory, evidence, and policy implications. In Neighbourhood effects research: New perspectives. Springer pp. 23–56 (2012)

Fotheringham, A.S.: Spatial structure and distance-decay parameters. Ann. Assoc. Am. Geogr. **71**(3), 425–436 (1981). 00302

Grimm, V., Railsback, S.F.: Pattern-oriented modelling: a multi-scope for predictive systems ecology. Philos. Trans. R. Soc. Lond. Ser. B, Biol. Sci. **367**(1586), 298–310 (2012)

Hermann, C.F.: Validation problems in games and simulations with special reference to models of international politics. Behav. Sci. **12**(3), 216–231 (1967)

Krugman, P.: Confronting the mystery of urban hierarchy. J. Japn. Int. Econ. **10**(4), 399–418 (1996a). 00228

Krugman, P.: Urban concentration: the role of increasing returns and transport costs. Int. Reg.Sci. Rev. **19**(1–2), 5–30 (1996b). 00333

Lotka, A.J.: Elements of Physical Biology. Williams & Wilkins Company, Baltimore (1925)

Lucas, C. Steyaert, P.: Modular inheritance of objects through mixin-methods. In Proceedings of the 1994 Joint Modular Languages Conference (JMLC); Ulm, German, 28–30 September, pp. 273–282 (1994)

Martin, R.: Rebalancing the spatial economy: the challenge for regional theory. Terri. Polit. Gov. **3**(3), 235–272 (2015). 00001

Nitsch, V.: Zipf zipped. J. Urban Econ. **57**(1), 86–100 (2005)

Openshaw, S.: From data crunching to model crunching: the dawn of a new era. Environ. Plann. A **15**(8), 1011–1012 (1983)

Openshaw, S.: Building an Automated Modeling System to Explore a Universe of Spatial Interaction Models. Geographical Analysis **20**(1) (1988)

Prehofer, C.: Feature-oriented programming: A fresh look at objects. In ECOOP'97 Object-Oriented Programming; Springer: Berlin and Heidelberg, German, pp. 419–443 (1997)

Preston, S.H.: Urban growth in developing countries: a demographic reappraisal. Popul. Dev. Rev. **5**(2), 195–215 (1979). 00231

Pumain, D.: Pour une théorie évolutive des villes. Espace géographique **26**(2), 119–134 (1997). 00138

Pumain, D. (ed.): Hierarchy in Natural and Social Sciences. Methodos Series. Springer, Berlin (2006)

Pumain, D., Paulus, F., Vacchiani-Marcuzzo, C., Lobo, J.: An evolutionary theory for interpreting urban scaling laws. Cybergeo: Eur. J. Geogr. (2006) 00000

Reynaud, J.: Villes. Encyclopédie nouvelle **8**, 670–687 (1841). 00034

Rosenthal, S.S., Strange, W.C.: The determinants of agglomeration. J. Urban Econ. **50**(2), 191–229 (2001)

Sanders, L.: Intelligence artificielle et agents collectifs: le modèle EUROSIM. In: SAGEO 2005, Colloque International de Góomatique et d'Analyse Spatiale, pp. XX. Collection, Actes Avignon, CD-ROM (2005). ISBN 2-910545-067

Schmitt, C., Rey-Coyrehourcq, S., Reuillon, R., Pumain, D.: Half a billion simulations: evolutionary algorithms and distributed computing for calibrating the SimpopLocal geographical model (preprint). Planning and Design, Environment and Planning B (2015)

Spulber, D.F.: Global Competitive Strategy, vol. 00091. Cambridge University Press, Cambridge (2007)

Steyaert, P., Codenie, W., D'Hondt, T., de Hondt, K., Lucas, C., van Limberghan, M.: Nested mixin-methods in Agora. In ECOOP'93?Object-Oriented Programming; Springer: Berlin and Heidelberg, German, pp. 197–219 (1993)

Thiele, J.C. (2015). Replicating and breaking models: good for you and good for ecology. *Oikos*, (January):1–6

Tufte, E.R.: The Visual Display of Quantitative Information. Graphics Pr, 2nd edn (2001)

Ullman, E.L.: A theory of location for cities. Am. J. Soc. **46**(6), 853–864 (1941)

Chapter 5
Using Models to Explore Possible Futures (Contingency and Complexity)

Guillaume Chérel, Clémentine Cottineau and Romain Reuillon

> *Such an approach accepts that history and geography could always be 'otherwise', i.e., that the present is by no means guaranteed by the past; thus, to know a society and a geography is to know how it could be different than it is.*
> *Barney Warf, 1993*

Abstract This chapter considers models of urban systems as virtual laboratories to explore possible trajectories that a 'real' geographical system could have taken instead of the evolution path that is observed historically. It therefore builds on the concept of historical contingency, and regards simulation modelling as an opportunity to explore historical contingency *in silico*. This approach is illustrated by an experiment performed on a model of systems of cities applied to the urbanization of the (Former) Soviet Union, MARIUS, using a new algorithm seeking to maximize diversity in a model's outcomes. The discovery of possible trajectories of the target system through the model provides insight into the singularity of the realized trajectory, and can be used for prediction to define a range of possible outcomes resulting from a simulation of the model, not based on predefined scenarios but on the maximum diversity allowed by the model within reasonable parameter bounds.

Keywords Alternative · Trajectories · Model · Outcomes · Phase space · Search · Patterns

Up until now, we have focussed on the processes of urbanization and hierarchization leading to empirically observed patterns and trajectories (cf. Chaps. 2–4). In the last chapter, multi-modelling was introduced as a way to try and disentangle the probable causes of a particular urban evolution. Different combinations of mechanisms were simulated and compared to empirical data about the Soviet Union, allowing us to

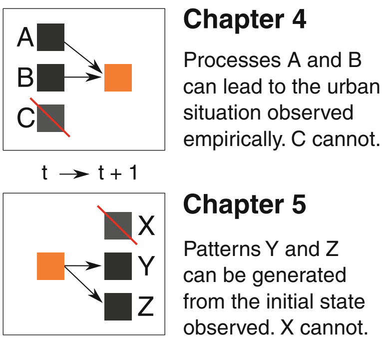
G. Chérel · R. Reuillon (✉)
Geographie-cites - UMR 8504, 13 rue du Four, 75006 Paris, France
e-mail: romain.reuillon@iscpif.fr

C. Cottineau · R. Reuillon
Institut des Systèmes Complexes, 113 Rue nationale, 75013 Paris, France

© Springer International Publishing AG 2017

D. Pumain and R. Reuillon, *Urban Dynamics and Simulation Models*,
Lecture Notes in Morphogenesis, DOI 10.1007/978-3-319-46497-8_5

Fig. 5.1 Equifinality in Chaps. 4 and 5



order and eventually discard theoretical candidates for the explanation of this urban evolution (Fig. 5.1). However, urban systems are path-dependent Arthur 1994 rather than ergodic, therefore it is necessary to test the sensitivity of our models to the variations of parameter values and to look at the variety of configurations that can be generated from the same structure of models. This chapter provides a proposition to study this property of agent-based models with an empirical example. We rely on the MARIUS model to present an algorithm of pattern searching that explores the diversity of outcomes possibly generated by the model with different sets of parameters. This search is not just a robustness test of simulation models, it gives us insights into alternative trajectories of a given urban system following a given set of rules. In other words, it helps us figure out what alternative pasts, presents and futures the system could have, informing prediction given complexity Batty and Torrens 2005.

5.1 Models as Artefacts of Historically Contingent Complex Systems

Simulation models are designed so as to abstract 'real' systems, by simplifying some of their dynamics, resulting in a operable scientific device: a virtual laboratory. Within this virtual laboratory, experiments can be replicated - which is impossible with the 'real system' - and the knowledge thus produced is being made reproductible. Furthermore, beyond a certain degree of complexity, heterogeneous entities, non-linear interactions and stochasticity are likely to generate a huge variety of trajectories from a unique initial state (Amblard 2003). Therefore, conducting simulation experiments with a model allows to explore the diversity of its possible trajectories by varying its initial conditions and parameterization, and consequently, when moving back to the 'real' system, to grasp how the system could have behaved otherwise. This is particularly interesting for geographical systems, such as systems of cities, whose observable regularities result from both particular historical contingencies and generic dynamics (Pumain 2003).

The historical trajectory of a city is the result of a contingent mix of individual events and political decisions, of local opportunities and constrains as well as transactions with higher geographical scales (the region, the State, the world), of random facts and unpredictable micro-decisions. If we go back to our case study of the Soviet Union, the trajectories of Moscow and Saint Petersburg over the years illustrate the multiple influence of individuals (Peter the Great's choice to establish a new capital city as a window towards Europe in the early 1700s for example), of collectives (bolsheviks, the members of the nomenklatura, oligarchs), of regional, national and international affairs, the importance of Saint Petersburg as a port city in a continental empire, the centrality of Moscow in the transportation networks, etc. At the individual level, it could even seem like nothing could be explained of their trajectory without knowing the full content of their historical chain of events. However, it is the contribution of social science - and urban geography in particular - to analyse and compare trajectories so as to draw some patterns, to highlight regularities and to generalize them into a theory of urban evolution.

A large amount of work in urban geography has revealed regularities in cities growth and linked urban attributes to patterns of growth and decline, looking at elements of absolute and relative location, economic specialization and systems' dynamics. However, from the early work of spatial economists in the nineteenth century to the most recent accounts from complexity science, these mechanistic claims are always accompanied by an emphasis on individual perturbations and the unpredictability of individual trajectories, unlike earlier positivist and deterministic views of geography. The present study builds on this ground and acknowledges both the fact that cities present regularities which allow to build mechanistic models of their evolution, but also that the situation observed empirically is just one of many other realizations of its initial potential (Warf 1993; Byrne 1998; Pumain 2003). What other realizations could have happened alternatively? Besides one or two educated guess, it is usually hard to tell from qualitative or analytical studies.

What this contribution from generative simulation adds to the acknowledgement of historical contingency in urban trajectories is a method to try and draw a full picture of the other possible outcomes, based on a set of common rules of urban evolution. This picture helps understanding what the system could have become under close initial conditions. It also provides a richer basis for predictive accounts. Model simulations enable to 'replaying history' with slight changes in parameters or initial conditions, and thus it allows to explore the diversity of a model's outputs to qualitatively delineate different regimes in the simulated trajectories. We used the MARIUS model and the set of rules described in Chap. 4 to simulate urban evolutions. Then we designed an algorithm to detect different patterns in the simulated trajectories and to maximize their trajectories. The following sections present the algorithm, the alternative trajectories of the system modelled at different periods and tries to trace back the values of parameters that achieved different outcomes to inform our knowledge of possible futures, pasts and presents of the former Soviet Union to help us understand it better.

5.2 A Method to Foster Diversity in a Model Outcomes

Fostering diversity in the outputs of a model requires to define some adapted measures to distinguish between the trajectories, and to guide the exploration of the model behaviours towards an increase of the diversity of the resulting trajectories with respect to these measures. The Pattern Search Exploration (PSE) algorithm was developed for this precise purpose, by adapting principles from Novelty Search in genetic algorithm (Chérel et al. 2015).

5.2.1 The Pattern Space Exploration Algorithm: Principles and Implementation

The PSE algorithm relies on the Popper falsifiabilty principle, considering a model as a general hypothesis on the way a system behaves. From such a general hypothesis are derived some more concrete hypothesis realization: simulations, who may be seen as associations of input (parameters, environment) and output values (results). Results are then confronted to observational data (provided they are available), described and interpreted in terms of patterns[1], allowing the model user to distinguish between simulations corroborating the data and simulations who contradict data. So as to validating a model, one should as well ensure that a model is able to corroborate the observed data, by producing the expected patterns as looking at simulation which may falsify it. The information gained by discovering unexpected patterns is then used in the iterative process of model refining. In this process, a model is designed, the pattern space explored, unexpected patterns (in particular falsifiers) are used to revise the model assumptions, exploration of the pattern space starts over and the loop goes on until the model is sufficiently satisfying (Cottineau et al. 2015b).

Some of the patterns a model may produce are unexpected, even to the eye of the modeller. An unexpected pattern can either be a falsifier of the model if it is contradicted by empirical observation, or a prediction of the model if it is validated by it. To search for the unexpected patterns in simulations results is to search for both at the same time.

Looking for unexpected patterns in the output space is a different kind of parameter space exploration than the one presented in Chap. 3, as there is no fitness function to guide the exploration: we do not know what we are looking for.

[1] A pattern denotes a description of a simulation run. For example a simulation of a spatially explicit population dynamics model can be described in terms of population growth and spatial aggregation. All distinct combinations of values for these two variables are as many distinct patterns (Chérel et al., 2015).

PSE algorithm addresses this issue by projecting simulation outputs in a specific space - the pattern space- determined by modellers, and using Novelty Search to guide the exploration of parameter space in relation to the parameter space. The goal of the exploration is now to

5.2.2 *Evolutionary Methods for Parameter Space Exploration*

As in any parameter space exploration of computer-based models of complex systems, several difficulties arise: the high dimensionality of the parameter space, non-linearity between inputs and outputs, and stochasticity. High dimensionality (tens or hundreds of parameters) makes the parameter space prohibitively huge to try out all possible parameter values (or a regular sampling in the case of real valued parameters). Non-linearity entails that interesting output values may correspond to very small regions of the parameter space, making them difficult to look for. Stochasticity makes it necessary to replicate the simulations. For models which take up to minutes to run a simulation, a clever exploration method to reduce the number of simulations is necessary.

Some methods try to tackle the problem of dimensionality of the parameter space with *a priori* samplings, such as the well known Latin hypercube sampling or LHS.[2] Others try to solve it in an adaptive way, by choosing on the base of previously collected information (Gramacy et al. 2004; Castro et al. 2013). To summarize, these methods solve the problem of the parameter space size by not exploring the parameter space exhaustively, sometimes using heuristics based on the irrelevance of some areas of the parameter space given their particular objectives.

We do not have any constraint about the parameter space. This transfers the problem of dimensionality from the parameter space to the pattern space. While it may be hard to reduce the number of free parameters of a model of complex systems, modellers are free to choose what variables to measure on a simulation, i.e. how many dimensions the pattern space has. Typically, a two- or three-dimensional space has both the advantage of being tractable and easily visualized.

Evolutionary methods are well suited for this setting because they can perform a search through the parameter space based on information gathered in the pattern space by simulation, and do so without any *a priori* knowledge of the relationship between the parameter space and the pattern space. Existing evolutionary calibration methods are all based on *a priori* ideas of the patterns one would like to obtain. The *a priori* target pattern can be based on data to reproduce an observed scenario. It can also be designed to represent an expected behaviour and used to assess the models ability to reproduce it (such as the V-shaped formation of flocks in Stonedahl and Wilensky (2011). An objective -or fitness- function is designed in terms of a distance

[2] see examples of sampling methods and comparisons in Saltelli et al. (2008), Kleijnen et al. (2005).

between a model run and the target pattern, and evolutionary methods are used to tune the parameters to minimize it. Our problem is different in that, rather than targeting a particular pattern, we want to find the various ones a model can produce.

5.2.3 Novelty Search

Novelty Search (Lehman and Stanley 2010; 2011) does not rely on an objective function. It generates successive populations of solutions to a problem by breeding the most novel individuals of the current population. The novelty of current individuals is measured with respect to the current population and an archive of past individuals, and reflects the novelty of patterns (not parameter values). Novelty Search continually drives the exploration of the parameter space towards areas which produce novel patterns.

Pattern Space Exploration adapts Novelty Search to the exploration of computer-based models of complex systems. In standard Novelty Search (Lehman and Stanley 2011), the novelty measure is based on the average distance of an individual to its k nearest neighbours in the pattern space. In PSE, novelty relies on a 'hit map' counter. The hit map is a discretization of the pattern space into regular cells, and the number of individuals which belong to each cell are counted throughout the exploration. The novelty measure of an individual is based on the inverse value of its corresponding cell count. The less a cell has received individuals before, the more novel is an individual standing on it. This variation offers several advantages for our purpose.

First, defining a distance measure in a multidimensional space is not a trivial task, in particular when the different dimensions represent unrelated variables. The hit map approach alleviates this problem by suppressing the distance measure, and relying instead on the discretization of the pattern space which is done independently for each dimension. Second, computing the k nearest neighbours is computationally costly, and constitutes a serious bottleneck in an evolutionary algorithm because it cannot be parallelized (while the individuals evaluations are independent and can be). The only operations needed with the hit map are accessing or incrementing the counter at a given cell, and those are done in constant time. Third, the hit map naturally offers a measure of the exploration progress, the volume discovered in the pattern space, approximated by counting the number of cells which counters are positive.

5.2.4 PSE Algorithm

As with all evolutionary algorithms, PSE generates new individuals through a combination of genetic inheritance from parent individuals and mutation. PSE selects the parents according to the rarity of their patterns compared to the pattern population and to the previous generations.

PSE features the common components of evolutionary algorithms:

- a population of individuals, each of which encodes a particular element of the search space,
- a selection method to select individuals from which a new population will be created,
- a crossover and mutation mechanisms to create a new individual from previously selected ones,
- an elitism mechanism to filter all the individuals (newly created and the old population) to be kept and form the new population.

As the algorithm progresses, unwanted individuals are discarded and forgotten as new ones are added to the population. An archive is used to store information about past populations.

Algorithm 1 General Scheme of an evolutionary algorithm as in Chérel et al. (2015)

Generate μ genotypes by sampling the genotypic space
Evaluate them
while stop condition not reached **do**
  for $i = 1$ **to** λ , the number of population to create **do**
    Select 2 individuals from the population (selection)
    Create a new genotype by crossover from the 2 selected individuals' genotypes (crossover)
    Mutate the new genotype (mutation)
  end for
  Compute the offsprings phenotypes (evaluation)
  Filter the individuals (parents and offsprings) according to their phenotypes (elitism)
end while

The algorithm starts by generating individuals with random genomes representing parameter values.

For each individual, a simulation is run and the pattern it produces is measured (evaluation), resulting in as many patterns as generated individuals. At the elitism stage, individuals are filtered to keep only those that are significantly different from one another with respect to their patterns. Pairs of individuals are selected based on how rare their patterns are relatively to the current population and past ones (rarity is PSEs equivalent of the novelty measure). Every element of the genome of each pair is recombined to form a new genome, which is then further mutated. A new generation of individuals is thus created, and the loop continues (Fig. 5.2). The main idea is that by selecting parents whose patterns were rarely observed previously, we increase the chances to find patterns yet undiscovered.

An individual is described by a genome and a phenotype. Its genome encodes an element of the search space. The phenotype of the individual is computed given its genotype through evaluation. The phenotype thus reflects the behaviour of the individual. In PSE, a genotype encodes a value for each parameter. An evaluation corresponds to a simulation run with the parameter values taken from the genome. A phenotype encodes a pattern measured on the simulation output.

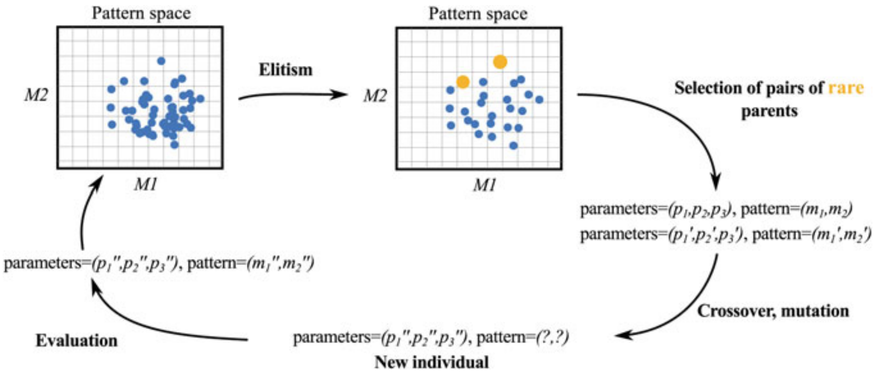


Fig. 5.2 PSE principle

For sake of simplicity, we will consider that the genotypic and phenotypic spaces are multidimensional real spaces, respectively, $\mathbb{R}^K$ and $\mathbb{R}^M$, where K is the number of distinct parameters of a model, and M is the number of descriptives variables used to qualify a pattern, i.e. the dimension of the pattern space.

PSE tends to overcome other existing approaches such as Sobol, Latin hypercube sampling or Grid sampling, in terms of coverage of the pattern space, see Chérel et al. (2015) for details.

5.3 Application to Systems of Cities

Although this method is generic and applicable to any model from which one would want to explore the pattern space, its application to the simulation of geographical systems is of paramount interest to explore the possible futures at a point in time as well as the alternative past and present patterns which could have occurred given the same initial conditions and rules of interaction. It has been developed for the MARIUS family of models described in Chap. 4. By applying the PSE exploration method to this model, the range of its possible trajectories can be explored and the resulting alternative situations can be discussed in the same terms as the trajectory of the best performing parameterization of the model (when calibrated against empirical data).

PSE approach implies the identification of order parameters, that is the measures which define the pattern space in which we look for diverse alternatives (Sect. 5.3.1) to compare with the historical trajectory. We then compare the possible trajectories identified with PSE to the calibrated one (Sect. 5.3.2).

5.3.1 *Order Parameters from Empirical Observation of Urban Systems Evolution Over Time*

Systems of cities are defined by their interactions within a coherent space (Cristelli et al. 2012), usually a country or subcontinent. The system cities form and its evolution over decades of urbanization can be described by it size, its growth over time, its spatial distribution and its hierarchical organization (Pumain 1997). Over short time periods, the number and spatial configuration of cities are relatively stable. We therefore focus on growth and the hierarchical distribution to characterize the state of a system of cities at a given point in time.

5.3.1.1 Population Growth

Growth relates to the variation of population living in cities. Its variation is strongly linked to overall processes of demographic and urban transition, and was observed to be generally comprised between -1 and $+5$ percent of the existing population on average per year, with a mean of 0 to 1 percent (Pumain et al. 2015). The growth indicator which we use in this experiment is simply given by the absolute difference between final and initial population of the system.

In order to distinguish monotonic population growth regimes from more complex ones, the number of sign inversion of the growth rate is also monitored during a simulation.

5.3.1.2 Hierarchisation

The hierarchical distribution refers to the fact that cities are distributed very unevenly according to their size, giving an heavy-tail distribution best described by a power law (known as Zipf's when the exponent value equals -1). The exponent of the power law approximating the size distribution is an indicator of size inequality.

It varies empirically from -0.8 to -1.5 between countries with higher size equality and countries where the city size hierarchy is the strongest, using a consistent definition of cities as built-up areas at the end of the twentieth century (Moriconi-Ebrard 1993).

Combining growth and hierarchy into qualitative patterns, one expects four types of outcomes of urbanization between two dates:

- hierarchisation in a growth context (e.g. Europe in the twentieth century)
- hierarchisation in a shrinkage context (e.g. Russia in the 2000s)
- equalization in a growth context (e.g. China in the 1960–1970s)
- equalization in a shrinkage context.

The complete pattern space explored with PSE was in fact three-dimensional: ($Growth$, $Hierarchisation$, $number\ of\ inversions$), but our focus is on the

Table 5.1 Parameter ranges for MARIUS 1, and calibrated values

Parameter	Min	Max	Calibrated 1959–1989	Calibrated 1989–2010
economicMultiplier	0	100	0.343809442	0.1278786
sizeEffectOnSupply	1	2	1.0017563880	1.0000000
sizeEffectOnDemand	1	2	1.0792607803	1.0141498
wealthToPopulationExponent	0	2	0.3804356044	0.4023026
distanceDecay	0	10	0.6722631615	0.6434053
populationToWealthExponent	1	10	1.0866012754	1.0060258
bonusMultiplier	0	1000	197.9488907791	0.0000000
fixedCost	0	1000	0.2565248068	0.1266480

Table 5.2 Parameter ranges for MARIUS 2, and calibrated values

Parameter	Min	Max	Calibrated 1959–1989	Calibrated 1989–2010
economicMultiplier	0	100	0.667886409	0.6081329
sizeEffectOnSupply	1	2	1.001053646	1.006575
sizeEffectOnDemand	1	2	1.006845651	1.000165
wealthToPopulationExponent	0	2	0.362268806	0.6451355
distanceDecay	0	10	0.000000000	0.2994255
populationToWealthExponent	1	10	1.038185914	1.036641
bonusMultiplier	0	1000	36.448303315	1.602549
fixedCost	0	1000	0.032413850	100
oilAndGazEffect	−1	1	0.003653162	0.04216174
coalEffect	−1	1	−0.011317100	0.0001835161
territorialTaxes	0	1	0.667721301	0.1006380
capitalShareOfTaxes	0	1	0.997001708	0.2347508
ruralMultiplier	0	1	0.018644683	0.0009140980

first two dimensions. The number of inversions completes the interpretation of the overall growth values, by depicting alternating regimes of growth and shrinkage.

5.3.2 Parameter Space and Pattern Space

MARIUS models are combinations of generative mechanisms (cf. Chap. 4). For the purpose of this exploration, we explored the pattern space of two combinations of them, for the two time periods of 1959–1989 and 1989–2010.

The first instantiated model is the most parsimonious combination which still produces a satisfactory reproduction of the Soviet urbanization (Cottineau et al. 2015a). Let us call this structure of model MARIUS 1. It has two additional mechanisms (bonified interactions and fixed costs) and eight free parameters in total, comprised in the bounds given in Table 5.1.

Table 5.3 Parameter ranges for MARIUS 1, and calibrated values

Pattern descriptor	Step	Min	Max	Empirical Value
Rank-Size slope	0.1	-10	0	[−0.8;1.2]
Population difference	5000	$-0.83544 * 10^8$	$5 * 10^10$	$[-10^7;+10^9]$
Inversion count	1	-0	3	[0;2]

The step column gives the discretization step size. Min and max values give the theoretical bounds for each observable. Artificial bounds (inside square brackets) were set on some parameters to focus the exploration on regions of interest. Calibrated values are from Cottineau et al. (2015a), reproducing the hierarchisation and growth of Soviet cities during the late 20th.

The second instantiated model (MARIUS 2) is the complete combination of mechanisms with free parameters in total, comprised in the bounds given in Table 5.2.

Pattern space is discretized according to Table 5.3

Growth has been upper bounded artificially to 500 millions, the inversion count was upper bounded to 3, and the rank–size slope (hierarchisation) was lower bounded to -10. Simulation whose measures overcome these bounds are assumed to have a value equal to the bound. Preliminary experiments not shown here revealed that population growth can get far above 500 million, and hierarchisation can get far below -10, and the number of inversions up to 28 (the maximum that can happen in the 30 steps of the simulation).

PSE has been run for a total of approximately 36 million model evaluations, distributed on 5000 computing units. One model evaluation takes about 40 s. The whole run took 400, 000 h in cumulative computation time and had been drastically reduced to about one week, thanks to the OpenMOLE platform computation distribution method, the degree of parallelism of PSE method and the computational capabilities of the EGI grid.

5.3.3 Results

Experiments reveal that PSE was more efficient than other methods to reveal distinct patterns in the pattern space of MARIUS. Moreover, the first experiments revealed the presence of extreme patterns which can be identified as falsifiers (Chérel et al. 2015) (cf. Fig. 5.3 left). Here we have linked patterns and parameters to try and identify to restrict the pattern space to a significant one.

For a model structure including two additional mechanisms (Bonus and Fixed Costs) between 1959 and 1989, we find that (cf. Fig. 5.4):

- the parameters in charge of the conversion rate between economic growth and population growth (economicMultiplier, wealthToPopulationExponent and bonusMultiplier) need to be limited to 1 to prevent from extreme patterns of growth and hierarchisation (bottom right line).

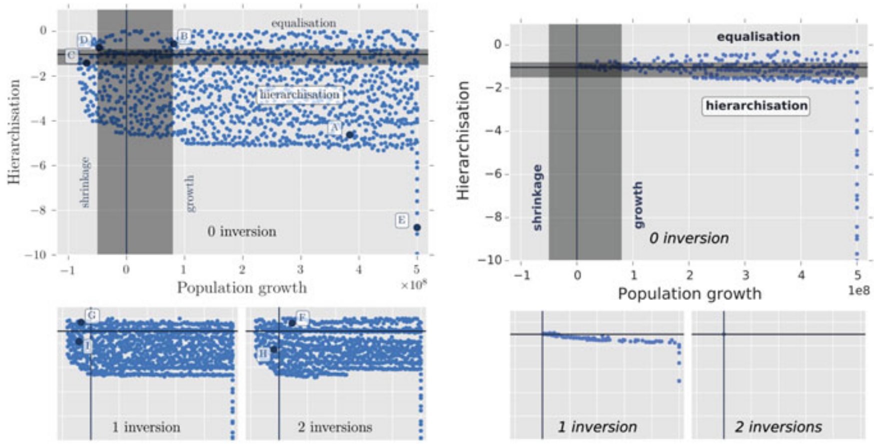


Fig. 5.3 New experiment with restricted parameter bounds and complete model. On the *left* large bounds, MARIUS + Bonus + Fixed Costs + 1959–1989. On the *right* restricted bounds, MARIUS complete + 1959–1989

- similarly, when distance affects interactions too strongly (distanceDecay > 2), we find patterns of growth and hierarchisation empirically unusual (bottom right quadrant).
- parameters of wealth initialization and updating need to be capped from 1.5–1.6 to avoid extreme patterns of growth and equalization (top right quadrant).
- Large values of sizeEffectOnSupply (>1.6) and low values of bonusMultiplier (<0.1) leed to extreme patterns of hierarchisation with global population shrinkage (bottom left quadrant).

With these new bounds (on the right) and a complete model of the same period, we find a much restricted room for alternative pasts (cf. Fig. 5.3). We find that initial conditions might not have allowed a fluctuated growth of population (no pattern for 2 inversions), and that an inversion of growth would necessarily have led to further inequality of city sizes (no pattern with the complete model and 1 inversion lead to a rank-size slope above the initial one). Finally, this period and parameterization lead to growing patterns only. The real alternative revealed by this experiment is the process of hierarchisation or equalization of city size associate with growth. Figure 5.3 shows that both patterns were accessible in 1959 given the modelled mechanisms, even though hierarchisation was the path observed during the period.

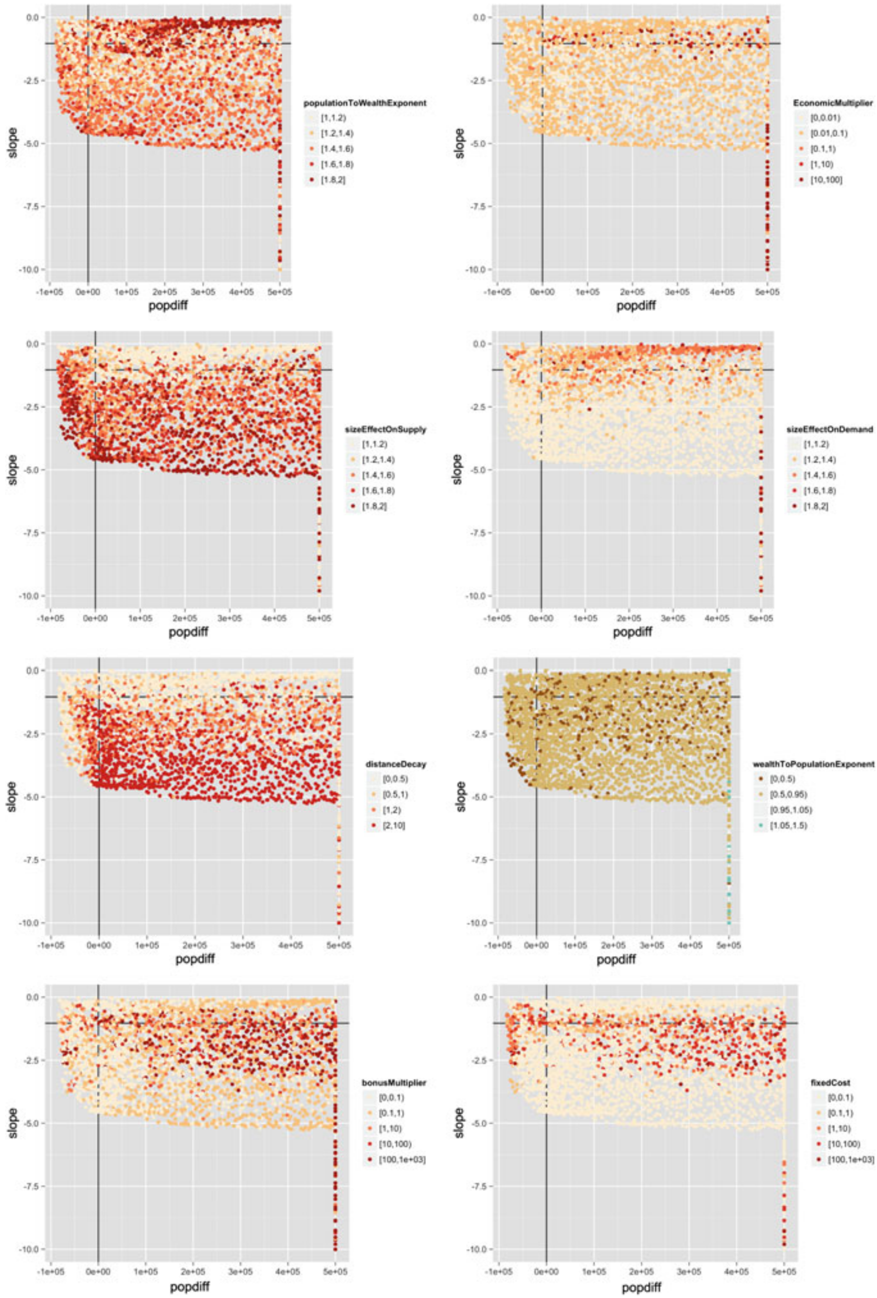


Fig. 5.4 Values of parameters for all patterns

5.4 Conclusion: Acknowledging Historical Contingency for the Prediction of Potential Urban Futures

Exploring pattern space shed light on unexpected patterns that deepen our understanding of the model general behaviour by revealing what it is able to produce inside the whole parameter space and not only around the sole calibrated parameters set. By looking at the corresponding parameterization that have led the model to exhibit some unexpected patterns, we increase the model completeness and correction. Correction is increased by discovering falsifiers and others problematic cases, who help to revise the model assumptions such as the range of possible values for each parameter, or the unexpected interactions between a subset of particular parameters and mechanisms (e.g. the linear production and the superlinear consumption of pattern G (cf. Fig. 5.3 left). Completeness of the model is increased by looking at the plausible trajectories and the pattern to be found within and around the realism span of the pattern space. Aside from clarifying the degree of contingency in the system observed evolution, it also strengthens our belief that the mechanisms of MARIUS are sound candidates to explain and reproduce systems of cities evolution dynamics.

The reason for this is that with complex systems, we may have data about the emerging patterns but not so much about the internal mechanisms, and this is the case with city systems (it is easier to know the population size in time through censuses than quantities related to MARIUSs parameters). The very purpose of complex systems modelling is sometimes to capture internal mechanisms that are not directly observable. In such case, it is easier to say if an observed global pattern is realistic rather than if a parameter value is.

Prediction of complex system is tricky if not impossible (Batty and Torrens 2005). At the very least, attempts at predicting futures should allow a span of possibilities and take contingency of social events into account.

The kind of methods presented here should be of particular relevance to historical systems, which present a high degree of contingency, but these methods are applicable to any other complex systems, and integrated in the OpenMOLE toolbox (cf. Chap. 6).

References

Amblard, F.: Comprendre le fonctionnement de simulations sociales individus-centrées: Application à des modèles de dynamiques dopinions. Ph.D. thesis, Université Blaise Pascal (2003)

Arthur, W.B.: Urban systems and historical path dependence, In: Arthur, W.B., Kenneth J.A. Increasing returns and path dependence in the economy. Ann Arbor, University of Michigan Press (1994). http://www.jstor.org/stable/10.3998/mpub.10029

Batty, M., Torrens, P.M.: Modelling and prediction in a complex world. Futures **37**(7), 745–766 (2005)

Byrne, D.S.: Complexity Theory and the Social Sciences: An Introduction. Psychology Press, Hove (1998)

Castro, P., Petit, E., Farjallah, A., Jalby, W.: Adaptive sampling for performance characterization of application kernels. Concurr. Comput. Pract. Exp. **25**(17), 2345–2362 (2013)
Chérel, G., Cottineau, C., Reuillon, R.: Beyond corroboration: Strengthening model validation by looking for unexpected patterns. PLoS ONE **10**(9), 1–28 (2015)
Cottineau, C., Chapron, P., Reuillon, R.: Growing models from the bottom up. an evaluation-based incremental modelling method (ebimm) applied to the simulation of systems of cities. Journal of Artificial Societies and Social. Simulation **18**(4), 9 (2015a)
Cottineau, C., Reuillon, R., Chapron, P., Rey-Coyrehourcq, S., Pumain, D.: A modular modelling framework for hypotheses testing in the simulation of urbanisation. Systems **3**(4), 348–377 (2015b)
Cristelli, M., Batty, M., Pietronero, L.: There is more than a power law in Zipf. Sci. Rep. **2**, 812 (2012). EP –. Article
Gramacy, R.B., Lee, H.K., Macready, W.G.: Parameter space exploration with gaussian process trees. In: Proceedings of the Twenty-First International Conference on Machine Learning, p. 45. ACM (2004)
Kleijnen, J.P., Sanchez, S.M., Lucas, T.W., Cioppa, T.M.: State-of-the-art review: a users guide to the brave new world of designing simulation experiments. INFORMS J. Comput. **17**(3), 263–289 (2005)
Lehman, J., Stanley, K.: Efficiently evolving programs through the search for novelty. In: Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation, pp. 837–844. ACM (2010)
Lehman, J., Stanley, K.O.: Abandoning objectives: Evolution through the search for novelty alone. Evol. Comput. **19**(2), 189–223 (2011)
Moriconi-Ebrard, F.: L'urbanisation du monde depuis 1950 (1993)
Pumain, D.: Pour une théorie évolutive des villes. Espac. Geogr. **26**(2), 119–134 (1997)
Pumain, D.: Une approche de la complexité en géographie. Géocarrefour **78**(1), 25–31 (2003)
Pumain, D., Swerts, E., Cottineau, C., Vacchiani-Marcuzzo, C., Ignazzi, A., Bretagnolle, A., Delisle, F., Cura, R., Lizzi, L., Baffi, S.: Multilevel comparison of large urban systems. Cybergeo: European Journal of Geography, vol. **705**, (2015). doi:10.4000/cybergeo.26730
Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S.: Global Sensitivity Analysis: The Primer. Wiley, New York (2008)
Stonedahl, F., Wilensky, U.: Finding forms of flocking: Evolutionary search in abm parameter-spaces. In: Bosse, T., Geller, A., Jonker, C.M. (eds.) Multi-Agent-Based Imulation XI: International Workshop, MABS 2010, Toronto, Canada, 11 May 2010, Revised Selected Papers, pp. 61–75. Springer, Berlin (2011)
Warf, B.: Postmodernism and the localities debate: ontological questions and epistemological implications. Tijdschrift voor Economische en Sociale Geografie **84**(3), 162–168 (1993)

Chapter 6
An Innovative and Open Toolbox

6.1 Introduction

This chapter introduces the platform, OpenMOLE. It is a generic tool used to run the different methods, which is presented in detail in the previous chapters. To simplify the comprehension, we focus on a simple model, but which does not concern the city modelling. However the principles are the same.

The modelling process can be seen as an iterative process, in which specific knowledge is injected and series of issues has to be discussed. How does each input participate in the production of outputs? Are all inputs necessary to generate all the expected dynamics? What are the robustness intervals for the inputs? What are all the possible dynamics of the model? Answering these questions helps in getting a better understanding of the model under construction and a better idea of *what my model is*? OpenMOLE has been thought to answer these questions. It exposes a workflow formalism in which the model is the centre of attention. Numerical experiments can be designed from simple parameter exploration to high level methods dealing with calibration, sensitivity analysis, scenario reproduction.

This chapter presents the central concepts and the OpenMOLE formalism with the example of a simple but stochastic complex-system model. In the first part, we explain how to run a piece of program exposing this stochastic model with OpenMOLE, then we show how to do replications on it, how to explore the input space of parameters according to a Latin Hypercube Sampling (LHS). Finally, we expose three advanced methods: The first one is an evolutionary process, which aims at finding an optimal set of input parameters to simulate a given output (or reproducing a scenario). The second one provides with the validity of the input ranges in the context of the previous scenario reproduction. The third one produces a map of output diversity.

This Ant model has been chosen to serve as a didactic example. It is simple to explain its rules, yet it belongs to the category of complex systems. It is a real-world model getting a minimal set of inputs and outputs, so that the OpenMOLE methodology tools can be easily understood. However, in OpenMOLE, a model can

© Springer International Publishing AG 2017

D. Pumain and R. Reuillon, *Urban Dynamics and Simulation Models*,
Lecture Notes in Morphogenesis, DOI 10.1007/978-3-319-46497-8_6

be viewed as a black box so that it is quite simple to transfer the following methods to an other model.

This chapter does not explain how to instal OpenMOLE, how to launch and how to handle the OpenMOLE application. The reason for this omission is that such instructions are provided and updated on the OpenMOLE website.[1]

6.2 The Ant Model

We propose to study a Netlogo model picked up from the Netlogo Library.[2] However, no skills in Netlogo programming are required. As embedded models in OpenMOLE are encapsulated and can be viewed as a black boxes, the following OpenMOLE scripts can be used for any other language.

The Ant model was created by (Ury Wilensky 1997 and 1999) (Fig. 6.1). The NetLogo's website describes this model as follows: *In this project, a colony of ants forages for food. Though each ant follows a set of simple rules, the colony as a whole acts in a sophisticated way. When an ant finds a piece of food, it carries the food back to the nest, dropping a chemical as it moves. When other ants sniff the chemical, they follow the chemical towards the food. As more ants carry food to the nest, they reinforce the chemical trail.*

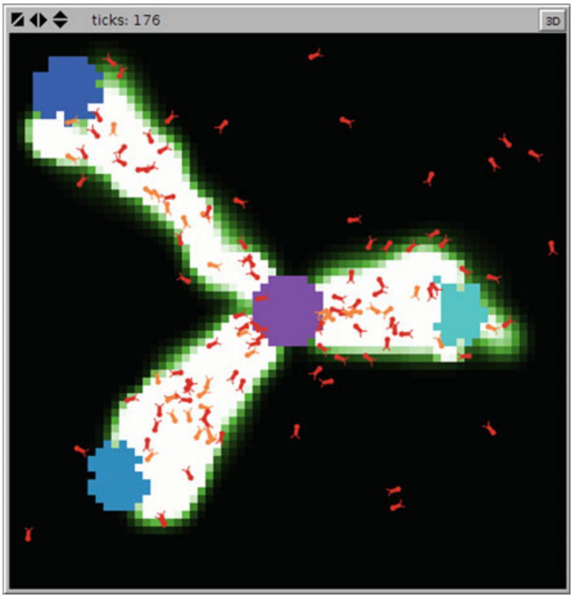


Fig. 6.1 The Netlogo ants model

[1] https://www.openmole.org/.

[2] http://ccl.northwestern.edu/netlogo/models/Ants.

In this experiment, three food spots are set in the ants living area. The experiment consists in testing the impact of the three model inputs on the time required by the ants to consume the three food spots.

The tree inputs of the model are

- the number of ants,
- the evaporation rate of the chemical,
- the diffusion rate of the chemical.

We modified the source code so that we can obtain the food extinction time for each spot

Listing 1 final-ticks-food1, 2, 3 represents the needed number of ticks, measured in simulation steps or ticks (final-ticks-food) to consume the spots 1, 2, 3

```
 to compute-fitness
  if ((sum [food] of patches with [food-source-number = 1] = 0)
     and (final-ticks-food1 = 0)) [
    set final -ticks-food1 ticks ]
  if ((sum [food] of patches with [food-source-number = 2] = 0)
     and (final-ticks-food2 = 0)) [
    set final -ticks-food2 ticks ]
  if ((sum [food] of patches with [food-source-number = 3] = 0)
     and (final-ticks-food3 = 0)) [
    set final -ticks-food3 ticks ]
end
```

This model is stochastic. At each time step an ant, which is not sniffing the chemical, can go in any direction randomly. As a consequence, we need to repeat a given experiment (set with given input values) several times to ensure that any pattern generated is robust. Therefore we need to initialize a Random Number Generator by means of a *seed* value.

6.3 Embed the Model in OpenMOLE

The first operation is to run the Netlogo model on the OpenMOLE platform.

OpenMOLE can run executions on High Performance Computating environments. It implies that we need to ensure that any code embedded by the platform can be ported from one machine to another. This depends on the language with which the model is coded. In the case of the Netlogo language, it is straightforward, since Netlogo runs on the Java Virtual Machine, which has been designed to be portable. Otherwise a packaging operation, based of the Care software[3] would be necessary to ensure that all the required libraries at runtime are embedded. We chose not to expose this packaging operation here to focus on methods. However, this operation is simple and well supported in OpenMOLE.

[3] http://reproducible.io/.

An experiment is described in OpenMOLE as a workflow. A workflow is composed of Tasks, which can be chained and ordered by means of another concept: the Transitions. Let us introduce a couple of OpenMOLE concepts:

A Task is an atomic execution component, which can be run concurrently. They tasks have been designed so that they have no interfering side effects. Therefore they can be safely dispatched on several threads, processes or computers. A task can carry a programm, which will be executed at runtime. It receives values (Val) as inputs from the workflow and can produce other values (Val) as outputs.

A Val is a typed Value. It can represent a Double, an Integer, a String, a File (and even Java defined class).

A Transition defines a precedence link between two Tasks. It is always run locally, unlike the Tasks, which can be run on remote environments. It makes the Vals travel from one Task to another.

We first design a very simple workflow containing only one Task (carrying the Netlogo model). We also map the inputs and the outputs of the Netlogo model to some Vals set as inputs and outputs of the Task. Thus we can assign values to the Netlogo model inputs thanks to the mapped Val. In the general case, Task inputs are set with the values of Vals arriving from the workflow by means of a Transition (what we do later in the chapter). But, for now, we just build a very simple workflow composed of one single Task, so that no Transition can feed the Task with any Val. That is why, the input values of the Task are assigned manually.

So we first define seven Vals corresponding to four inputs: the population of ants, the evaporation rate, the diffusion rate, the seed for the RNG as well as maxsteps, which represents the maximum of steps in the Netlogo code. We also define three outputs: the extinction time for the resource spots 1, 2 and 3.

Listing 2 4 Vals for the inputs and 3 Vals for the outputs

```
 // Define the input variables
val population = Val[Double]
val diffusion = Val[Double]
val evaporation = Val[Double]
val seed = Val[Int]
val maxsteps = Val[Int]

// Define the output variables
val food1 = Val[Double]
val food2 = Val[Double]
val food3 = Val[Double]
```

We define a NetlogoTask, containing the nlogo source, the launching instructions, the input/output mapping, as well as some manual initialization for the inputs.

Listing 3 Set of the task carrying the model

```
// Define the NetlogoTask
val cmds = Seq("random-seed ${seed}", "run-to-grid")
val ants =
  NetLogo5Task(workDirectory / "ants.nlogo", cmds) set (
```

```
    name := "ants",
    // Map the OpenMOLE variables to NetLogo variables
    netLogoInputs += (population, "gpopulation"),
    netLogoInputs += (diffusion, "gdiffusion-rate"),
    netLogoInputs += (evaporation, "gevaporation-rate"),
    netLogoInputs += (maxsteps, "gmax-steps"),
    netLogoOutputs += ("final-ticks-food1", food1),
    netLogoOutputs += ("final-ticks-food2", food2),
    netLogoOutputs += ("final-ticks-food3", food3),
    // The seed is used to control the initialisation of the
        random number generator of NetLogo
    inputs += seed,
    outputs += (population, diffusion, evaporation, maxsteps),
    // Define default values for inputs of the model
    //seed := 42,
    population := 125.0,
    maxsteps := 2000
)
```

Our first workflow is almost ready! We are just not able to visualize the produced outputs. Indeed, a Task has no side effect, so that it cannot display the value it produces. A Task can be viewed as a portable function, which maps an input value to an output value, nothing more. That is why, we introduce the following concept:

A Hook can be plugged on a Task to perform an action upon completion of the task it is attached to. The action is done locally, once the Task execution is back from an eventual remote host. There exists different kinds of Hooks, among which the AppendToCsvHook to append a Val value at the end of a given CSV file or a ToStringHook to display a Val value.

We need the latter to display the values of food1, food2 and food3. As these three Vals are provided as outputs, plugging a ToStringHook on the Task that produces them will result in their displaying when they are produced by the Task.

Listing 4 Hook plugging

```
//Define a workflow with one Task, hooked by the ToStringHook
ants hook ToStringHook()
```

With these final two lines, the workflow can be run and produces the following output:

Listing 5 Hook displaying th waiting times for the extinction of the 3 food spots

```
{food1=746.0, food2=1000.0, food3=2109.0}
```

6.4 Do Repetitions

An interesting thing is to replicate this stochastic model to get a mean value for the outputs. To do so, we introduce a new kind of Task, a new kind of Transition and a new concept.

A Sampling is a tool for exploring a space of parameters. The term parameter is understood in a very broad acceptation in OpenMOLE. It may deal with numbers, files, random streams, images, etc. There exists a lot of different ways to explore a space of parameters. An exhaustive list of the available Samplings in openmole is given on the https://www.openmole.org/ website.

An Exploration Task is a special Task, whose only setting is a Sampling. Its only goal is to compute the Sampling it carries and to generate all the parameter sets produced by the sampling. It is always followed by a special Transition:

An Exploration Transition links an ExplorationTask to another Task. It creates one new execution stream by sample point in the Sampling of the ExplorationTask. Exploration transitions are represented by the symbol $- <$ (Fig. 6.2).

To carry out the replications on our model, we want to pick up n values from a uniform distribution of integers. Let's admit, we just need 10 repetitions. Then the Exploration Task carrying this Sampling is defined by:

Listing 6 The definition of the Exploration Task and the new workflow statement

```
val replications =
  ExplorationTask (
  seed in UniformDistribution[Int]() take 10) set (
  name := "Replicate ants",
  (inputs, outputs) += (diffusion, evaporation),
  diffusion := 10.0,
  evaporation := 10.0
)

replications -< (ants hook ToStringHook())
```

Listing 7 Results for 10 repetitions

```
{food1=625.0, food2=1311.0, food3=1900.0}
{food1=546.0, food2=1109.0, food3=2574.0}
{food1=526.0, food2=1233.0, food3=2063.0}
{food1=790.0, food2=1214.0, food3=1901.0}
{food1=714.0, food2=1205.0, food3=2133.0}
{food1=534.0, food2=1067.0, food3=2035.0}
{food1=748.0, food2=1338.0, food3=2149.0}
{food1=908.0, food2=1148.0, food3=1821.0}
{food1=682.0, food2=1149.0, food3=1829.0}
{food1=905.0, food2=1315.0, food3=1771.0}
```

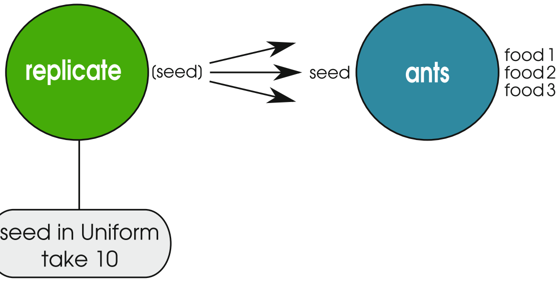


Fig. 6.2 10 different seeds are generated and given as input to 10 instances of the ants Task. Each of them provides food1, food2 and food3

6.5 Automatic Workload Distribution

In the previous section, we generated 10 computation streams. They are independent from one another since they do not require any information from another stream. So that we can easily take advantage of the parallelism with OpenMOLE.

OpenMOLE allows the distribution of computation on servers, on clusters (PBS, OAR, SGE, Slurm, Condor), or on the EGI grid. After having provided with your login/password or your ssh private key or your Grid certificate to the platform depending on what technology you use (see the application documentation for the details on https://www.openmole.org/), delegating the workload on these environments is straightforward. All we need to do is to create the required environment and to specify the Task you want to delegate on it.

Listing 8 Definition of a computational environment (PBS, local multi-core, EGI Grid, ...) and assignment to the ants Task, so that the latter will be deported on the previously defined environment at runtime

```
val env = new PBSEnvironment("myLogin", "PBSmachineName")
// val env = LocalEnvironment(10) to take advantage of the
    cores of your own personal emachine
// val env = EGIEnvironment("vo.complex-systems.eu") for
    accessing the Grid VO vo.complex-systems.eu
// etc.

explore -< ants hook ToStringHook() on env
```

6.6 Expose the Variability of the Model

We can use the previous workflow to highlight the variability of the Ants model and to well understand why it is so important to do repetitions on such stochastic models. Let us set both *diffusion* and *evaporation* to 25.0. Then let us do 100,000 repetitions to have an idea of the variability of the model response. The following graphs show

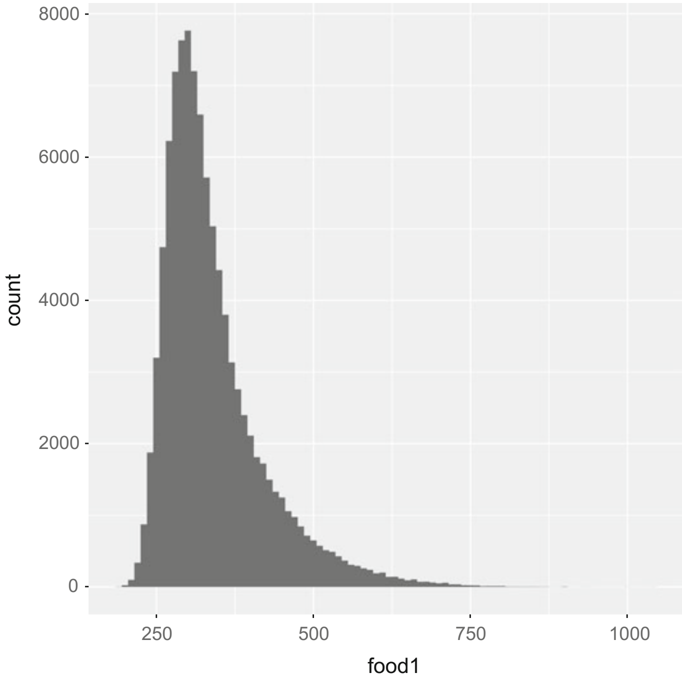


Fig. 6.3 Distribution for the spot food1 for *diffusion* = 25.0 and *evaporation* = 25.0

the required time to consume each food spot for the same input parameters (Figs. 6.3, 6.4 and 6.5).

6.7 Aggregate the Results

We now want to aggregate all the streams and compute a median value on them. To do so, we need a new kind of Transition, which is the counterpart of the Exploration one and merges all the streams generated by the Exploration into one array: **the Aggregation Transition** (represented by $> -$). We then plug another Task onto this transition to perform the median value from that array. For this, we use a Task called a ScalaTask, which can execute some Scala[4] code.

Listing 9 A Task for computing the median values

```
val medFood1 = Val[Double]
val medFood2 = Val[Double]
val medFood3 = Val[Double]
```

[4]http://www.scala-lang.org/.

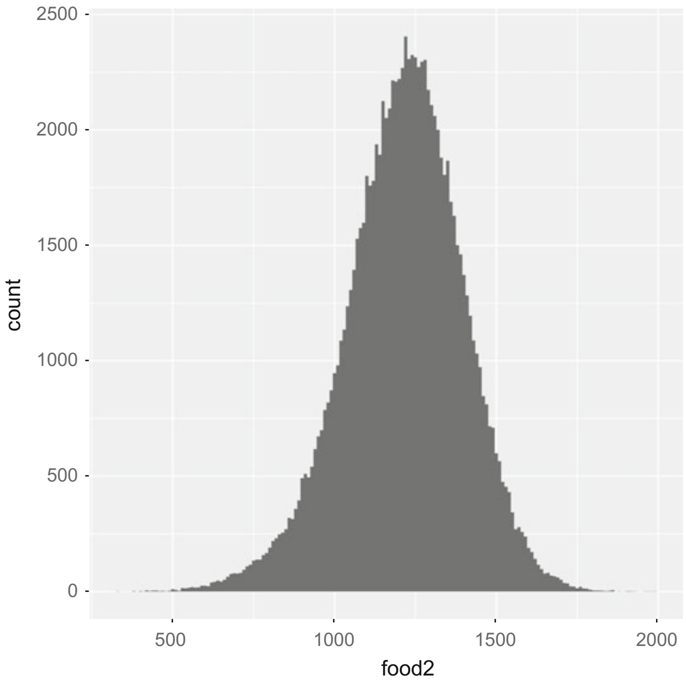


Fig. 6.4 Distribution for the spot food2 for *diffusion* = 25.0 and *evaporation* = 25.0

```
val medians =
  ScalaTask ("""
    import math.abs

    val medFood1 = food1.median
    val medFood2 = food2.median
    val medFood3 = food3.median""") set (
    name := "medians",
    inputs += (food1.array, food2.array, food3.array),
    outputs += (medFood1, medFood2, medFood3)
  )
```

The workflow becomes

Listing 10 A Task for computing the median values

```
replications -< ants >- (medians hook ToStringHook())
```

The resulting workflow can be represented by Fig. 6.6.
The output given by the Hook set on the median Task gives:

Listing 11 Median values for food1, food2 and food3 for 10 repetitions of ants

```
{avgFood1=649.5, avgFood2=1250.0, avgFood3=1979.0}
```

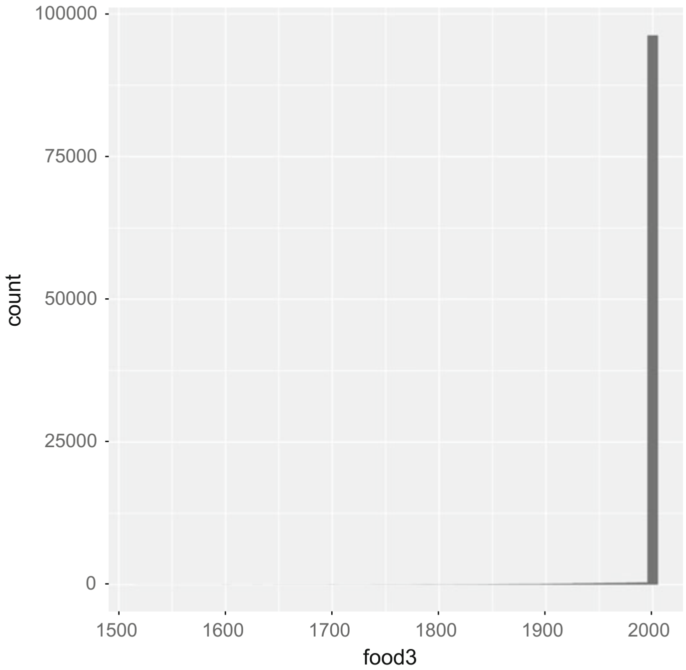


Fig. 6.5 Distribution for the spot food3 for *diffusion* = 25.0 and *evaporation* = 25.0

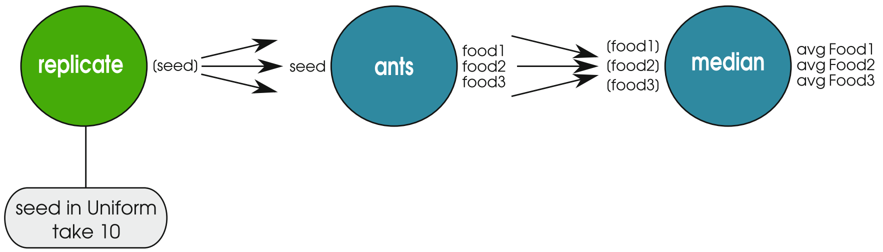


Fig. 6.6 Values generated for food1, food2 and food3 by each of the 10 ants instances are merged into 3 arrays ([food1], [food2] and [food3]) by means of an Aggregation Transition and are processed by the median Task, which provides median values for each array (avgFood1, avgFood2 and avgFood3)

6.8 Explore the Space of Parameters

We now explore the parameter space composed by the evaporation rate and the diffusion rate values to test their individual and combined effects on the time for consuming the food spots. We do not study the impact of the population size since it seems clear that the bigger the population, the faster the food spots will be eaten. The population is thus set arbitrarily to 125.0. To perform the sampling of parameter values, we use a Latin Hypercube Sampling[5] of size 100. It means that a sampling of 100 couples (diffusion, evaporation) is generated. We evaluate each couple 100 times, leading to 10,000 executions of the model. It implies some modifications in the script.

First, we need to build a new ExplorationTask to carry out the LHS sampling. This exploration will be executed before the replication one, so that we can calculate a median value on 100 repetitions for each sample generated by the LHS.

Listing 12 A LHS sampling is carried out by a TaskExploration and build 100 couples (evaporation, diffusion)

```
val sampling =
  LHS(
   500,
   diffusion in (10.0, 100.0),
   evaporation in (10.0, 100.0)
  )

val exploration = ExplorationTask(sampling)
```

As shown in the Fig. 6.7, diffusion and evaporation are propagated in the workflow through the replicate Task and then directly to the median Task. Indeed, we need these values to be stored at the end of the workflow with medians of food extinction times. This way, we can pair the outputs to the inputs used to generate them.

To do so, we add these two parameters as input and as output of the replicate Task and we add a Transition between the replicate Task and the median Task (which takes also these two parameters as inputs). At this point, we need to introduce two new concepts.

The Capsule: carries a Task and several Slots.

A Slot is a synchronization point for all the Transitions arriving on it. It guarantees that all the Transition transmissions are completed before starting the Task carried by the Capsule. When a Task is created, a Capsule is automatically generated to carry it. Sometimes, we need to create it manually to keep a reference on it.

In our case, we need to create the Capsule of the replicate Task in order to build two Transitions: one to the ants Task and the other to the median Task. On the other hand, we need to create manualy the Slot of the median Task to make a synchronization point between the Transitions arriving from the ants Task and the replicate Task. The Fig. 6.7 and the Listing 13 give an overview of this technical rearrangement.

Listing 13 The full script of the experiment

```
val seed = Val[Int]
val population = Val[Double]
val diffusion = Val[Double]
val evaporation = Val[Double]
val maxsteps = Val[Int]

// Define the output variables
```

[5] https://en.wikipedia.org/wiki/Latin_hypercube_sampling.

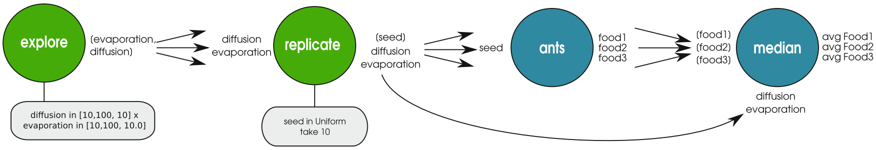


Fig. 6.7 An Exploration Task is designed to make vary diffusion and evaporation. It generates an array of couples (diffusion, evaporation), which combines all different possible combinations of the two variables. These values are transmitted to the replicate Task and then to the median Task, so that they can be stored in a file thanks to the Hook set on the median Task

```
val food1 = Val[Double]
val food2 = Val[Double]
val food3 = Val[Double]

val medFood1 = Val[Double]
val medFood2 = Val[Double]
val medFood3 = Val[Double]

// Define the NetlogoTask
val cmds = Seq("random-seed ${seed}", "run-to-grid")
val ants =
  NetLogo5Task(workDirectory / "ants.nlogo", cmds) set (
    name := "ants",
    // Map the OpenMOLE variables to NetLogo variables
    netLogoInputs += (population, "gpopulation"),
    netLogoInputs += (diffusion, "gdiffusion-rate"),
    netLogoInputs += (evaporation, "gevaporation-rate"),
    netLogoInputs += (maxsteps, "gmax-steps"),
    netLogoOutputs += ("final-ticks-food1", food1),
    netLogoOutputs += ("final-ticks-food2", food2),
    netLogoOutputs += ("final-ticks-food3", food3),
    // The seed is used to control the initialisation of the
        random
number generator of NetLogo
    inputs += seed,
    outputs += (population, diffusion, evaporation, maxsteps),
    // Define default values for inputs of the model
    //seed := 42,
    population := 125.0,
    maxsteps := 2000
  )

val replications =
   ExplorationTask (
    seed in UniformDistribution[Int]() take 100) set (
    name := "Replicate ants",
    inputs += (diffusion, evaporation),
    outputs += (diffusion, evaporation),
    diffusion := 10.0,
    evaporation := 10.0
  )
```

```
val medians =
  ScalaTask ("""
    import math.abs

    val medFood1 = food1.median
    val medFood2 = food2.median
    val medFood3 = food3.median""") set (
    name := "medians",
    inputs += (food1.array, food2.array, food3.array),
    outputs += (medFood1, medFood2, medFood3)
  )

val sampling =
  LHS(
    100,
    diffusion in (10.0, 100.0),
    evaporation in (10.0, 100.0)
  )

val exploration = ExplorationTask(sampling)

val storeHook = AppendToCSVFileHook(workDirectory /
    "result.csv")

exploration -< Strain(replications -< ants >- medians) hook
    storeHook
```

The output of this experiment, stored in the result/result.csv file gives an exploration of 100 different sets of parameters, each having been repeated 100 times. Using this method, we can find the best input couple, which leads to the scenario we aim at simulating. For instance, we may be interested in producing the following real-world experiment: the spots 1, 2 and 3 are emptied in respectively 250, 400 and 800 seconds. So, we are looking for the lowest distance between the simulated output and the expected output, which can for instance be expressed as the sum $\mid 250 - \text{avgFood1} \mid + \mid 400 - \text{avgFood2} \mid + \mid 800 - \text{avgFood3} \mid$

The closest simulation to this target gives the minimal sum of 197. Of course the best score is reached if the experiment reproduces exactly the real case (meaning a sum of 0). The input values associated are presented in the following table:

diffusion	**evaporation**	**Sum of differences**
37.8	10.0	197.0

Well, we find *one* solution. 100 simulations (with 100 repetitions for each, i.e. 10,000 runs) might seem like a large-scale experiment but a continuous two-dimensional problem may produce a lot of heterogeneity in the output space. Is there a better solution to this problem and to which extent is it better? What are the validity intervals for the inputs? What does the output space of parameter look like? So many questions we try to answer with evolutionary methods.

6.9 Optimization with Genetic Algorithms

In a genetic algorithm, an individual carries a genome, which is a set of genes (values for each input parameters). Evaluating an individual means executing a model simulation with the parameter values in the genome and performing the desired measures on the model output. The set of measured values constitutes what we will call here a pattern. Each simulation thus generates a pattern. When the model is stochastic, we can take the mean or median pattern of several simulation replications with the same parameter values. In the end, an individual is composed of the genome and the associated pattern.

Back to our ants optimization problem, the objective here is to find the closest pattern to a experimentally measured pattern (value 250, 400 and 800 for avgFood1, avgFood2, avgFood3 respectively). This problem is also called a calibration problem. To do so, we use the multi-criteria optimization genetic algorithm NSGA2 available in OpenMOLE and used for the calibration of SimpopLocal, cf. Chap. 3. It takes the following parameters as inputs:

- mu: the number of individuals to be randomly generated in order to initialize the population,
- objectives: the objectives to minimise,
- genome: the sequence of model input parameters on which the optimization is done, with the associated lower and upper bounds,
- replication: the repetition strategy

Listing 14 The NSGA2 settings in OpenMOLE

```
// Execute the workflow
// Define the population (10) and the number of generations
    (100).
// Define the inputs and their respective variation bounds.
// Define the objectives to minimize.
// Assign 1 percent of the computing time to reevaluating
// parameter settings to eliminate over-evaluated individuals.
val nsga2 =
  NSGA2(
    mu = 50,
    genome = Seq(
      diffusion in (0.0, 99.0),
      evaporation in (0.0, 99.0)),
    objectives = Seq(deltaFood),
    replication = Replication(seed = seed, aggregation =
        Seq(median))
  )
```

The variable foodTimesDifference is a new Val, representing the sum of absolute differences between the experimental time to reproduce and the simulated times.

We also need to cope with the distributed computation. OpenMOLE offers several approaches to tackle this problem. Among them, we are here interested in the steady-state approach. This algorithm begins with *n* individuals and launches a maximal number of evaluations as long as there are available computing units. When an evaluation is over, it is integrated in the population and a new individual is generated and evaluated on the computing unit that has just been freed. This method uses all computing units continuously and is recommended in a cluster environment.

Listing 15 Distribution in OpenMOLE with the steady approach

```
val evolution =
  SteadyStateEvolution(
    algorithm = nsga2,
    evaluation = ants -- objective,
    parallelism = 10,
    termination = 100
  )
```

We feed SteadyGA with the evolution method that was described above (*nsga*2) and the piece of workflow to be evaluated (*evaluation*). The *parallelism* parameter specifies how many evaluation are concurrently submitted to the execution environment and *termination* is the termination criterion; here it runs for 100 generations (note that this parameter can also be set as a duration (10 h for example)). SteadyGA launches new evaluations as long as current evaluations are below this value.

SteadyGA returns two variables called in our example puzzle and ga. The second contains information on the current evolution and allows to define hooks that save the current population into csv file or to print the current generation. The following code provides 2 Hooks to (i) save the population corresponding to each generation into a file results/population#.csv, where # is replaced by the number of the generation and (ii) to display in console the generation number:

```
// Define a hook to save the Pareto frontier
val savePopulationHook = SavePopulationHook(evolution,
    workDirectory / "results")
```

When we launch this OpenMOLE workflow, the evolution will progressively produce parameter values having the best fitness, i.e. for which the model is closest to experimental values. We show the evolution of the distance between simulation and experimental measures between successive evaluations in the following Fig. 6.8.

In this table is presented the best result at the end of 800 evaluations.

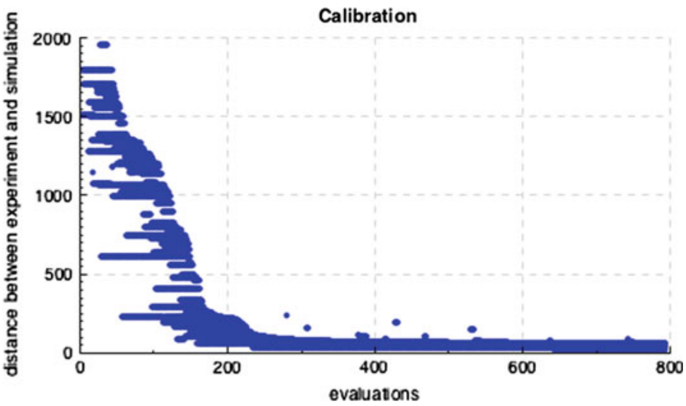


Fig. 6.8 Evolution of the distance between the experimental values and the model. It converges in less than 300 hundred evaluations

diffusion	evaporation	Sum of differences
71.17	5.61	15.5

This result is better than the one obtained with the LHS exploration method. The sum of differences is more than 6 times lower in almost less than half the evaluation time. It is also interesting to notice that the input value are in a completely different regions of space: (71.17, 5.61) versus (13.27, 10.18). It demonstrates how the Genetic Algorithm is faster and more efficient in this kind of optimization problem. The difference between the two methods would be even greater in higher dimensionality problems.

6.10 Sensitivity Analysis with the Profiles Method

The method we now present focuses on the impact of the different parameters in order to better understand how they contribute to the model overall. In our Ants example, we calibrated the model to reproduce a set of notional experimental measurements. We would like to know whether the model can reproduce this pattern for other parameter values. It may be that the model cannot reproduce the experimental measurements if a crucial parameter is set to a value other than the one found by the calibration process. On the other hand, another parameter may prove not to be essential at all; that is, the model may be able to reproduce the experimental measurements whatever its value. To establish the relevance of our model parameters, we will investigate the parameters' profiles for the model and for the targeted pattern.

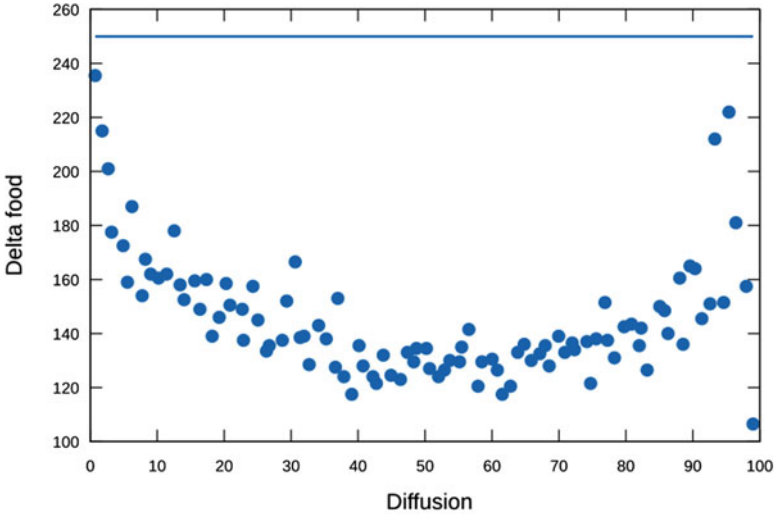
We first establish the calibration of the evaporation parameter. Specifically, we would like to know whether the model can reproduce the targeted pattern with different evaporation rates. We divide the parameter interval into *nX* intervals of the same size, and we apply a genetic algorithm to search for values for other the parameters (the ants model only takes two parameters, so that the dispersal parameter is the only one to be varied), which, as done previously in the calibration, minimise the distance between the measurements produced by the model and the ones observed experimentally. In the calibration case, we kept the best individuals of the population whatever their parameter values. This time, we still keep the best individuals, but we keep at least one individual for each interval division of the profiled parameter (in this case, the evaporation parameter). Then, we repeat the process with the dispersal parameter.

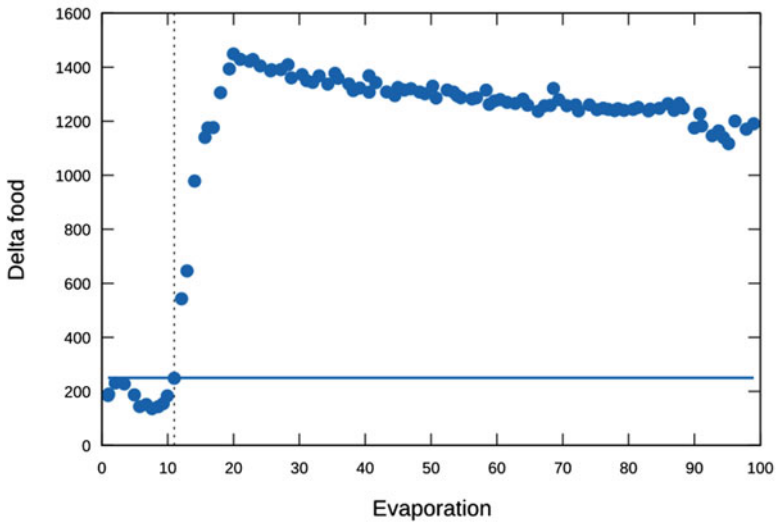
To set a profile for a given *variable* in OpenMOLE, the GenomeProfile evolutionary method is used:

```
def profile(variable: Val[Double]) = {
  val profile =
    GenomeProfile (
      x = variable,
      nX = 100,
      genome = Seq(
        diffusion in (0.0, 99.0),
        evaporation in (0.0, 99.0)),
      objective = deltaFood,
      replication = Replication(seed = seed)
    )

  // Calibration profile of 1000 points for the parameter
  val evolution = SteadyStateEvolution(
    algorithm = profile,
    evaluation = ants -- objective,
    termination = 20000
  )
```

The arguments *genome*, *termination*, *objective* have the same role as the calibration workflow. The argument *objective* is in this instance not a sequence but a single objective to minimise. The argument *x* specifies the index of the parameter to be profiled, i.e. its position within the inputs sequence, indexing starting at 0. *nX* is the size of the of the interval in the parameter range discretisation.





When the diffusion rate is set to any value above 10, the model is able to reproduce experimental measures rather accurately. A refined profile within the interval [0; 20] may be useful to give a more precise picture of the change in the influence of the parameter. Model performance is on the contrary strongly sensitive to the evaporation parameter, as values over 10 lead to a strong increase in minimal fit. When running the model with a diffusion rate of 21 and evaporation rate of 15, we observe that the ants are not able to build a sufficiently stable pheromone path between the nest and furthest food pile, which increases the time needed to exploit it in a considerable way.

6.11 Validation, Testing Output Diversity

Knowing that a model can reproduce an observed phenomenon does not ensure its validity. By validation, we mean that we can trust it to explain the phenomenon in other experimental conditions and that its predictions are valid with other parameter values. We have already established that one way to test a model is to search for the variety of behaviours it can exhibit. The discovery of unexpected behaviours, if they disagree with the experimental data or the direct observation of the system it represents, provides us with the opportunity to revise the assumptions of the model or to correct bugs in the code. This principle also holds for the absence of expected pattern discovery, which reveals the inability of the model to produce such patterns. As we test a model and as we revise it, we can move toward a model we can trust to explain and predict a phenomenon.

One might wonder, for instance, if in our ant colony model the closest food source is always exploited before the furthest. Accordingly, we decide to compare the different patterns that the model generates, looking specifically at the amount of time the model requires to drain the closest and the furthest food sources.

As in the previous experiment, we consider a task that runs 10 replications of the model with the same given parameter values and that provides, as its output, the median pattern described in two dimensions by the variables medFood1, the time in which the closest food source was exhausted, and medFood3, the time in which the furthest food source was exhausted.

To search for diversity, we use the PSE (Pattern Space Exploration) method (Chérel et al. 2015). As with all evolutionary algorithms, PSE generates new individuals through a combination of genetic inheritance from parent individuals and mutation. PSE (inspired by the novelty search method) selects the parents whose patterns are rare compared to the rest of the population and to the previous generations. In order to evaluate the rarity of a pattern, PSE discretises the pattern space, dividing this space into cells. Each time a simulation produces a pattern, a counter is incremented in the corresponding cell. PSE preferentially selects the parents whose associated cells have low counters. By selecting parents with rare patterns, we have a better chance to produce new individuals with previously unobserved behaviours.

In order to use PSE in OpenMOLE, the calibration utilized in the previous section is run with a different evolution method. We used to provide the following parameters:

- genome: the model parameters with their minimum and maximum bounds,
- objectives: the objectives measured for each simulation and within which we search for diversity,
- parallelism and termination have the same meaning as in the calibration example.

Here is the OpenMOLE code used for the PSE

```
val pse =
  PSE(
    genome = Seq(
      diffusion in (0.0, 99.0),
```

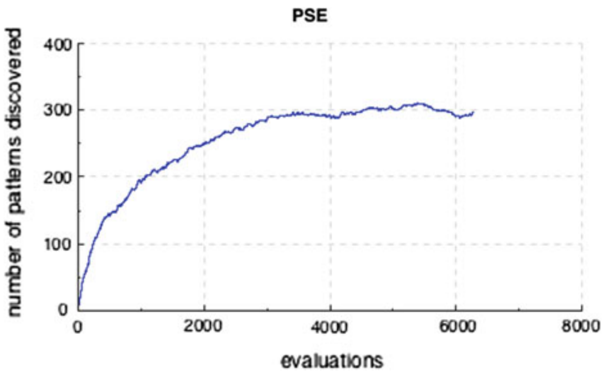
```
    evaporation in (0.0, 99.0)),
  objectives = Seq(
    food1 in (0.0 to 4000.0 by 50.0),
    food3 in (0.0 to 4000.0 by 50.0)),
  replication = Replication(seed = seed)
)

val evolution =
 SteadyStateEvolution(
   algorithm = pse,s
   evaluation = ants,
   parallelism = 10,
   termination = 1000000
 )
```

As the exploration progresses, new patterns are discovered. The following figure gives the number of known patterns (the number of cells with a counter value greater than 0) with respect to the number of evaluations.

When this number stabilizes, PSE is no longer making new discoveries. One has to be careful when interpreting this stabilization. The absence of new discoveries can mean that all the patterns that the model can produce have been discovered, but it is possible that other patterns exist but that PSE could not reach them.

The following figure shows the set of patterns discovered by PSE when we interrupt the exploration after it stabilizes.



The first observation that can be made is that all patterns have indeed been discovered: in every pattern, the closest food source has been drained before the furthest one. Further, there seems to be minimum and maximum bounds on the time period during which the nearest food source is consumed.

These observations give us starting points for further reflections on the collective behaviour of the ants. For instance, is the exploration of the closest food source systematic? Could there be ant species that explore further food sources first? If we found such a species, we would have to wonder which mechanisms make it possible and revise the model to take them into account. This illustrates how the discovery of the different behaviours the model is able to produce can lead us to formulate new hypotheses of the system under study, to test them and to revise the model, thus enhancing our understanding of the phenomenon.

References

Chérel, G., Cottineau, C., Reuillon, R.: Beyond corroboration: Strengthening model validation by looking for unexpected patterns. PLoS ONE **10**(9), 1–28 (2015)

OpenMOLE scientific workflow, distributed computing and parameter tuning. https://www.openmole.org/

Wilensky, U.: NetLogo Ants model. Center for connected learning and computer-based modeling, Northwestern institute on complex systems, Northwestern University, Evanston, IL (1997). http://ccl.northwestern.edu/netlogo/models/Ants

Wilensky, U.: NetLogo. Center for connected learning and computer-based modeling, Northwestern institute on complex systems, Northwestern University, Evanston, IL. (1999). http://ccl.northwestern.edu/netlogo/

Erratum to: Urban Dynamics and Simulation Models

Denise Pumain and Romain Reuillon

Erratum to:
D. Pumain and R. Reuillon, *Urban Dynamics and Simulation Models*, Lecture Notes in Morphogenesis, DOI 10.1007/978-3-319-46497-8

The original version of the book was inadvertently published without including an author name “Paul Chapron” in the title page and the list of “Authors and Contributors to the Book” in the book front matter. The book front matter has been updated with these changes.

The updated original online version for this bookfrontmatter can be found at DOI 10.1007/978-3-319-46497-8

D. Pumain (✉)
Université Paris I Panthéon-Sorbonne and UMR Géographie-Cités, Paris, France
e-mail: pumain@parisgeo.cnrs.fr

R. Reuillon
CNRS, UMR Géographie-Cités and Institut des Systémes Complexes Paris-Ile-de-France, Paris, France

© Springer International Publishing AG 2017
D. Pumain and R. Reuillon, *Urban Dynamics and Simulation Models*,
Lecture Notes in Morphogenesis, DOI 10.1007/978-3-319-46497-8_7

Knowledge Accelerator' in Geography and Social Sciences: Further and Faster, but Also Deeper and Wider

Arnaud Banos

With the development and spread of complex systems theories, methods and tools, large expectations rise. Justifying the positioning of the FuturICT European project as a 'knowledge accelerator', Dirk Helbing for example "*calls for a significant shift in the research and educational focus of academic institutions. Specifically, one needs to develop a better, holistic understanding of the global, strongly coupled and interdependent, dynamically complex systems that humans have created. It is required to push complexity science towards practical applicability, to invent a novel data science, to create a new generation of socially interactive, adaptive ICT systems, and to develop entirely new approaches for systemic risk assessment and integrated risk management*" (Helbing 2012).

Although one can hardly contest this statement in its general expression, two precisions need to be highlighted. First, while building complexity science is indeed a timely initiative, it would not make sense without reinforcing in the same movement the 'traditional' disciplines it is connected to. Second, 'accelerating' knowledge will not help much if this knowledge is not rooted in vivid theoretical, methodological and empirical basis. Therefore, the idea of accelerating knowledge necessarily goes with the idea of deepening and widening it. And in my opinion, this is one of the striking characteristics of GeoDiverCity project, as revealed by the various chapters constituting this book.

A New Fruitful Practice in Interdisciplinarity

First of all, let me stress a key issue: GeoDiverCity is a truly interdisciplinary project. It is definitely not enough to bring together disciplines so that interdisciplinarity emerges. Acculturation is a key factor of success and it involves time, much more time than the current 'project-driven research' may provide. And it is definitely a strength of GeoDiverCity to be rooted in the fertile ground of Géographie-cités' historic legacy.

Moreover—and this will be my second point —this book proves that empowering geographers and, beyond, researchers in human and social sciences, is possible: the level of autonomy they can reach in the field of data processing and modelling has never been so high. However, this does not mean that they become independent

© Springer International Publishing AG 2017

D. Pumain and R. Reuillon, *Urban Dynamics and Simulation Models*,
Lecture Notes in Morphogenesis, DOI 10.1007/978-3-319-46497-8

and can overcome the specialists in these very specialized, fast-evolving scientific fields. Quite the contrary! Such empowerment repositions the interactions between these communities and interdisciplinary working methods at a level that is much more fruitful. Specialists in data and models on the one hand are no longer seen as 'human–machine interfaces' and experts from the social sciences, on the other hand, are not any more apprehended as suppliers or data entry problems and humanized aids to the interpretation of model outputs. It is the entire interdisciplinary process that evolves to richer interactions. Social sciences researchers have for now the capability to build, analyze and experiment on their own methods and models.

By doing so, they expand their area of expertise with a renewed enthusiasm and energy and can interact much more efficiently and with relevant experts in these fields. These experts, in turn, earn more demanding colleagues, capable of posing problems to a more advanced level and actually contributing to the construction of methods, models and even tools needed for their resolution. This 'co-construction' is, for me, the spearhead of a vibrant and productive interdisciplinary research.

Modelling as a Necessary Experimental Method for Complex Social Problems

In this context, the activity of modelling holds a very special place for me. First because behind this term hides an incredible diversity of practices, partly irreducible to one another, yet equally relevant and to a large extent complementary. Second, a model can be a great mediator, even a catalyst of disciplinary and interdisciplinary collaborations. Its genesis involves creating a shared ontology and therefore the gradual creation of a common language.

Furthermore, the experimental 'if-then' approach a simulation model offers singularly broadens our space of possibilities and gives us new ways to question the world in which we live. For these reasons, modelling can be seen as fundamentally as a learning process. Indeed, being by definition iterative and interactive, modelling implies repeated interactions between the model developed and the vision of the phenomenon gradually built, from the preliminary conceptual model to its implementation and the systematic exploration of its behaviour in often multi-dimensional and large parameter spaces.

This last point is of specific relevance in the context of this book, as several chapters deal with this difficult issue. There is no alternative: the behaviour of each model must be known accurately. However, a common criticism against the kind of individual-based and path dependant models developed in GeoDiverCity concerns precisely the difficulty we face when characterizing their behavior.

The parallel with system dynamics is often formulated. For such models, mathematical tools exist and permit, in certain cases, to accurately characterize the behavior of the set of coupled differential equations defined. This limitation is less and less true today and widening and opening access to adapted computing resources significantly extend the range of possibilities for 'less conventional' models. Of course, parsimony is still required as it is wise to minimize the number of parameters included in any model, but more generally, today, the complexity of models can be adapted more closely to the problems addressed by social sciences.

Chapters 3, 4 and 5 illustrate this main output, and they all rely on the collaborative platform OpenMOLE that allows defining and realizing the design of experiments, sensitivity analysis, automatic calibration and pattern searching strategies needed.

New Tools for Model Validation and Reproducibility

These chapters also illustrate a major issue we face when modelling complex systems: it is indeed often impossible to offer unique and optimal solutions to complex problems. Complex systems models are characterized by the existence of multiple solutions, none of which are in general better than the others, if we stick to the only principles that governed the construction of the model. The challenge is then to identify, by simulation, a set of solutions defined by an optimality criterium. Finding such solutions is by no means a simple task, especially with the kind of multi-objective approaches adopted within GeoDiverCity. One of the great outputs of this research project resides for me in its definition and diffusion of methods and technologies favouring the reproduction of such multi-objective explorations.

More generally, this book provides new evidences that reproducibility can and should be a leading principle. While it is widely accepted that models are intended to be replicated by others in order to be validated or refuted, this requirement is not nearly as widespread as it should. The models developed or used in geography and social sciences in general, rarely reach this reproducibility criteria. The reasons are multiple and not necessarily negative: strong and personal commitment to the models that have sometimes taken years to be developed; sake of keeping the keys to some trade secrets developed long-term; desire to keep a step ahead of competitors in a field of increasingly competitive activity; self-censorship-related sometimes to approximate and non-optimal implementations by non-specialists.

Whatever the reasons, one should not forget the cumulative nature of knowledge construction in science in general, and in the social sciences more specifically (Pumain 2005). Sharing and diffusing models, methods and protocols is at least as important as diffusing results and once again GeoDiverCity proves to be exemplary in that perspective.

A New Incremental Multi-modelling Method

Coming back to the process of model building development, I would also like to emphasize a few other qualities of the enormous amount of work achieved here. First of all, as one can imagine by reading the various chapters of this book based on models and methods of increasing complexity, developing a model is not at all a simple or linear process.

There is even a form of path dependency and locking-in effect in this process: the first choices often play a decisive role and once the basic foundations of the model are built and progressive refinements are added, it becomes more and difficult to turn back. The collective investment is therefore too often limited to a single model, not necessarily the best one that could have been developed, and which, nevertheless, continues to exist and be used in laboratories for years, even though common sense would recommend that alternative models should be developed, all competing for the best model position. GeoDiverCity once again goes a step further: the MARIUS framework proposed underpasses this strong limitation, based as it is on an original incremental protocol.

Another characteristic of the modelling approach proposed is that is both theoretically and empirically based, a combination not that easily reached. When you start modelling, it is indeed easy and tempting to be confined to a relatively theoretical level, by eliminating the many constraints of data. Funnily enough, it is also very tempting and easy to stick to data. Theory driven approaches are frequently opposed to data driven ones, although none of the two is able to claim its full independency. GeoDiverCity does not fall in that trap and firmly anchors its models in both dimensions: while they are intended to reproduce observed processes and structures and are closely linked to them, their validation does not only rely on the direct comparison of model outputs with observation data, a questionable approach when dealing with path-dependant processes.

A Successful Tale for ERC Projects

All in all, this book sheds an optimistic light on the current state of public research in France. When coupled with long-term investments in disciplinary and interdisciplinary works, short-term project research proves to be very efficient. In that sense, the ERC project GeoDiverCity indeed played its accelerating role within its hosting institution Géographie-cités. Would the two components have been dissociated, I am not sure the results would have been that much convincing.

Moreover, the whole approach defined during those years and described within this collective book paves the way towards more experimentally grounded social sciences, even though the experimental framework we are talking about is merely computational. I would not dare claiming after the Nobel Prize Herbert Simon that "the social sciences are the real hard sciences" (cited in Squazzoni 2012), but I do believe they belong to this family. Experimentation is therefore an important accelerator factor, provided our capacity to handle it in an adapted way. Dissociating and isolating processes and structures is not an easy task and modelling and simulation can help a lot.

Interdisciplinarity, in its richest and deeper acceptation, is the 'voie royale' in that direction. Social scientists continue to receive little or no training in mathematics, computer science and more generally in modelling. Dissemination and sharing are therefore a responsibility of those who have the opportunity to acquire such skills. Interestingly, the vectors of this diffusion cannot be reduced to formal languages. First because such restriction is the best way to strengthen the structural asymmetry social sciences experiment.

Second, because mathematics and generally formal languages are not necessarily universal by essence. The mathematician Jean-François Colonna[1] recalls "the difficulty, if not impossibility for mathematics to construct naïve representations of the objects under study".

[1]http://www.lactamme.polytechnique.fr/Mosaic/descripteurs/AQuoiServentLesMathematiques.01.html.

Reinventing our methods, diversifying and adapting them to the problems, the contexts and the people we work with, this is the key challenge GeoDiverCity engages all of us to take up.

References

Helbing, D.: Introduction: the futurict knowledge accelerator towards a more resilient and sustainable future. Eur. Phys. J. Special Top. **214**(1), 5–9 (2012)

Pumain, D.: Cumulativité des connaissances. Revue Européenne des Sci. Soc. XLII **I**(131), 5–12 (2005)

Squazzoni, F.: Agent-Based Computational Sociology. Wiley, Chichester (2012)