

Article

## A Modular Modelling Framework for Hypotheses Testing in the Simulation of Urbanisation

Clémentine Cottineau <sup>1,2,\*</sup>, Romain Reuillon <sup>2</sup>, Paul Chapron <sup>2,3</sup>, Sébastien Rey-Coyrehourcq <sup>2,4</sup> and Denise Pumain <sup>2,5</sup>

<sup>1</sup> Centre for Advanced Spatial Analysis, University College London, London W1T 4TJ, UK

<sup>2</sup> UMR 8504 Géographie-cités and The Institute of Complex Systems (ISC-PIF), Paris 75006, France

<sup>3</sup> Faculty of Geoscience, University of Lausanne, Lausanne CH-1015, Switzerland

<sup>4</sup> UMR 6266 IDEES, CNRS, University of Rouen, Rouen 76000, France

<sup>5</sup> Department of Geography, University Paris 1 Panthéon-Sorbonne, Paris 75005, France;

E-Mails: romain.reuillon@iscpif.fr (R.R.); paul.chapron@gmail.com (P.C.);

reyman64@gmail.com (S.R.-C.); pumain@parisgeo.cnrs.fr (D.P.)

\* Author to whom correspondence should be addressed; E-Mail: c.cottineau@ucl.ac.uk;  
Tel.: +44-20-3108-3999.

Academic Editors: Koen H. van Dam and Rémy Courdier

Received: 25 August 2015 / Accepted: 24 November 2015 / Published: 27 November 2015

---

**Abstract:** In this paper, we present a modelling experiment developed to study systems of cities and processes of urbanisation in large territories over long time spans. Building on geographical theories of urban evolution, we rely on agent-based models to 1) formalise complementary and alternative hypotheses of urbanisation and 2) explore their ability to simulate observed patterns in a virtual laboratory. The paper is therefore divided into two sections : an overview of the mechanisms implemented to represent competing hypotheses used to simulate urban evolution; and an evaluation of the resulting model structures in their ability to simulate—efficiently and parsimoniously—a system of cities (between 1000 and 2000 cities in the Former Soviet Union) over several periods of time (before and after the crash of the USSR). We do so using a modular framework of model-building and evolutionary algorithms for the calibration of several model structures. This project aims at tackling equifinality in systems dynamics by confronting different mechanisms with similar evaluation criteria. It enables the identification of the best-performing models with respect to the chosen criteria by scanning automatically the parameter space along with the space of model structures (the different combinations of mechanisms).

**Keywords:** system of cities; ABM; simulation; former soviet union; equifinality; multimodelling

---

## 1. Introduction

Simulation models of urban systems were first developed in the 1950s and 1960s as a way to understand the complexity of cities and to forecast trends and consequences of planning policies. Several formalisms (thermodynamics, general systems theory, synergetic, microsimulation) were used [1], following fashions as well as opportunities arisen from access to new technologies [2,3]. Each of the methods cited have their specific advantages and drawbacks, that we will not discuss here, but all of them provide a similar opportunity and a similar challenge that is linked to simulation and has not changed as the formalisms evolved. The opportunity that we focus on in this paper is the function of a virtual laboratory that a computerised simulation model enables [4], which is of paramount interest for human and social sciences, where *in vivo* experiments are impossible. By allowing to implement competing and/or complementary hypotheses into generative mechanisms (that is, a set of interaction activities performed by entities (cities) producing emergent patterns [5]) in a model testable against empirical data, this function of virtual laboratory makes the model a framework for the evaluation of the plausibility of different theories. However, simulation as a method and an epistemological way of testing theories gives way to a limitation known since von Bertalanffy [6] as equifinality. It describes the fact that even when a model is performing well, one cannot infer that the underlying combination of mechanisms is the one operating “in real life”, because several models can lead to the same results (many-to-one) and the same process can lead to several qualitatively different results (one-to-many) [7]. The problem of adequacy assessment of the model with real life (also known as ontological adequacy testing, cf. [8]) is twofold. First, many effective (or “real”) processes can lead to the same observed situation within the target system that we aim to model. Second, inadequate mechanisms implemented in the model can simulate in a satisfying way the situation under study. The challenge of identifying the actual causal mechanisms is a recurrent problem in social simulation [9,10], but one that is usually overlooked at the stage of results’ analysis.

We have tried in a previous study [11] to tackle the one-to-many part of the challenge within the model, using an evolutionary algorithm to look for the maximum diversity in patterns produced by a given set of mechanisms in a model of systems of cities. What we present here relies to the multiplicity of possible causes leading to the single historical trajectory observed empirically. We present a multimodelling framework which allows to combine different mechanisms into a modular model evaluated against a unique set of evaluation criteria, enabling for the comparison of the performance of different model structures to simulate urbanisation and the evolution of cities in a territorial system.

More precisely, we are interested in simulating the co-evolution of cities encompassed in a territorial system (typically a nation or a continent) and to reproduce their hierarchical and spatial patterns. In other words, we want to obtain simulations that produce the stylised facts [12] generalised from the empirical observation of various systems of cities: their hierarchy, spacing and functional differentiation [13]. Section 2 presents the patterns we aim to reproduce and the catalogue of theories and mechanisms on

which we draw to compose the model. Section 3 describes the multi-model and its implementation in our study case, the evolution of the Soviet and post-Soviet system of cities (around 2000 cities). Section 4 exposes the results of its exploration through multi-calibration. This exploration aims to explore the performance of different hypotheses in explaining the evolution of Soviet and post-Soviet cities. Section 5 concludes on the study case and the method.

## 2. A Catalogue of Possible Mechanisms of Urbanisation

“It may also be useful to think of complex geographical models as extensions of thought experiments, where the necessary and contingent implications of theories can be examined. Further, admitting that “all models are wrong” is akin to the realisation in post-structural social science that multiple competing accounts of the same settings are possible, and that faced with a diversity of accounts the context and intent of each must be an important element in the evaluation process” ([7] p. 291).

A large bunch of theories in geography, economics and the natural sciences have tried to provide an account of the regularity of urbanisation processes and the structuration of a system of cities, through models and narratives. Keeping in mind the equifinality challenge, this means that we already have a strong theoretical basis and several causal model candidates to confront with empirical regularities. Without being exhaustive on these theories, we try to provide an overview of the mechanisms that have been proposed in the literature. We then expose the kind of results that are achieved by statistical models using empirical data, and present the particularity of our study case.

### 2.1. Competing Theories

To describe the structure of a system of cities, one focuses on three macro properties (patterns) of the system: the hierarchy of city sizes, their regular spacing and the functional differentiation of cities [13]. The first one of these macro properties has fostered the larger body of research and will be our main criterion for stating the performance of a simulation model, whereas the other two remain hard to formalise and to compare over time and space with respect to the availability of urban information at a local level. Consequently, we present in larger details the competing theories aiming at explaining the regularity of city size distribution and its evolution over time, and quickly review theories of location and functional specialisation.

The hierarchical organisation of city sizes represents a “mystery” [14] that has intrigued many researchers, because of its regularity and simplicity of description (the rank-size “rule”) despite the complexity of urban functioning and interactions (for reviews, cf. [15,16]). Auerbach, Lotka and Singer [17–19] are known to be the first to formalise this regularity, leading to the famous rule that the size of a city is a power law function of its rank in the hierarchy, a rule particularized by [20] arguing that the Pareto exponent is expected to be  $-1$ . Alternative mathematical descriptions have been proposed, for instance the lognormal distribution [21–23]. The two distributions, however, are close and tend to coincide in the middle-upper part of the hierarchy [24,25].

Because of this regularity, generative models of power laws and lognormal distributions have logically been considered as candidate explanations [24,25]. The most famous one is Gibrat’s law [23],

which generates a lognormal distribution by a process of “proportional effect”. This model of multiplicative growth allocates randomly average population growth rates from independent distributions, at each short time step, to cities independently from their size. It is thus considered close to a random walk. It results in amplifications of urban growth and decline leading to a lognormal distribution. The Simon model of preferential attachment [26] applied to cities generates power laws of city sizes by simulating an incremental creation of new urban blocks attached to existing urban clusters with a probability dependent on the block size [27]. Those two models are elegantly simple enough to generate hierarchical distributions, but they lack an explaining power to help understand urbanisation processes: “Scaling laws often are viewed as over-identified: they can be generated by a wide range of distinct models. It is essential to select specifications that integrate in the model a significant part of the existing knowledge about towns and cities” ([28] p. 1). Some attempts were made to characterise them in terms of exogenous shocks and contingent local policies (for example in Gabaix’s model [29]), but random growth models seem to miss the causal power of mechanism-based simulation models [5].

Other accounts of urban growth models leading to the observed stylised hierarchical structure and evolution (deterministic in Dimou and Schaffar’s typology [30], by opposition to random models) range into two broad categories: equilibrium models of externalities (such as [31,32]) and evolutionary models of spatial interactions (such as [33–35]). The former type of models formalises a balance of centripetal (sharing, matching, learning in the labour market for example) and centrifugal forces (pollution and congestion) resulting in optimal sizes for cities given the current technology. The latter type of models rely on attractivity and spatial interactions of cities to explain the dynamic of competition and cooperation resulting in a regular hierarchy in the urban system.

Location theories of cities typically build on two concepts to explain the regular spacing and empirical distribution of cities in space: site and/or situation advantages [13,32,36,37]. Site advantages refer to natural features available at an absolute location (natural resources, harbour conditions, *etc.*), whereas situation advantages refer to centrality in a transportation network or an interface relative position for example. All the theories mentioned above rely necessarily on relations with (consistent) territories providing resources or distance constraints for instance, in order to explain the urban co-evolution.

Theories of urban specialisation inherit from trade theories (via comparative and competitive advantages) and theories of product and innovation cycles. They state that cities have different features (size, situation, site, former specialisation) that make them akin to be more competitive and thus specialise in one economic sector in comparison with other cities, or to adopt an innovation sooner or later than their neighbours. Depending on the current cycle of the product, those specialisations affect not only the sectoral composition of the active population and its economic output but also confers an advantage to early adopters and defines further options for specialisation and growth [38,39].

Finally, political factors and singular policies are usually considered necessary to explain differentiated urban growth (through administrative functions or investment policies targeting specific cities at a given moment in time).

We compose our catalogue with mechanisms that formalise the theories cited above. They fall into five broad classes of generative mechanisms potentially accounting for the emergence of a structured system of cities:

- **Spatial Interactions** and diffusion allow for the exchange of informations, monies, goods and people. It thus makes cities co-evolve in time and adapt collectively to changing economic and innovation cycles through competition and cooperation, resulting in some complementarity of their specialisation. These local interactions and their consequences on the regular organisation of the system as a whole under spatial constraints could be thought of as “complex systems effects”.
- **Size effects** like agglomeration economies and urbanisation externalities illustrate a very direct and self-reinforcing cause for hierarchical differentiation.
- **Site effects** explain the spatial location of growth processes around resource-rich areas for the related innovation cycle.
- **Situation effects** illustrate the importance of the neighbouring relational environment (potential field, network accessibility, *etc.*) on a city’s pattern of growth.
- **Territorial effects** account for some exogenous (policy) shocks and the solidarity of urban trajectories in a common political space (through redistributive processes for example).

## 2.2. Empirical Results from the Literature

We look for statistical evidences of these factors of urbanisation in the empirical literature before turning to our case study experiment, where we propose an implementation and evaluate the power of each of the five mechanisms into an agent-based (multi-) model.

- **Spatial Interactions** are tricky to measure because of the variety and non-commensurability of flows circulating between cities at various temporalities. Until recently, the diffusion of innovations (agricultural techniques [40], telephone lines [41] or newspapers [42]) served as a proxy for these interactions. Since the development of various volumes of high velocity data, actual interactions (like phone calls [43]) have confirmed for example the relevance of the gravity model to describe inter-city interactions.
- **Size effects** on urban growth and differentiation were revealed by a persistent empirical correlation between growth rates and city sizes over long periods of time. All over the 19th century, Robson ([41] p. 79) measured a positive coefficient between the log of English and Welsh cities’ population and their gross ten-year growth rates (from a minimum of +1.47 between 1861 and 1871 to a maximum of +8.53 points in percentage between 1821 and 1831). This correlation is found for French cities as well [44]. The size effect finally relates to the stability of the rank position of large cities (by comparison with the fluctuations of smaller cities).
- **Site effects** were classically approached by estimating the surplus of growth associated with a localised resource (typically, deposits of natural materials, such as coal or gas). In the Soviet urban system of the 1920s–1930s, the location of a city on a coal deposit was associated with a surplus of 1.15 points in percentage of average demographic growth per annum, everything else being equal. A surplus over 0.5 point is observed nowadays (1989–2010) for oil and gas deposits [45]. In the USA, Black and Henderson [46] estimate that the coastal location (e.g., a resource for tourism) is associated with a significantly higher ten-year growth rates of three to five points.
- **Situation effects** can be revealed by the spatial autocorrelation of growth or the co-evolution between transportation networks and urban networks. In the first case, Hernando *et al.* [47] found a

characteristic distance for spatial autocorrelation of growth rates of 215 km for American counties and of 80 km for Spanish cities. As for transportation dynamics, Bretagnolle [48] measured the correlation between accessibility and growth rates for French cities in the last two centuries. She finds that cities that were weakly connected (by any transportation network: road, rail or air) in 1900 and stayed isolated in 2002 grew slower (0.94% on average per annum between 1900 and 2002) than cities that became motorway nodes (1.21%) or multimodal hubs (1.69%). Likewise, well connected cities at the beginning of the period tended to grow faster than the first category.

- **Territorial effects** can be approached empirically by relating political statuses to dynamics of growth. In developing countries, regional capitals were found to grow significantly faster by 0.5 to 1 point of annual average growth rate in the 1960s [49] and the 1990s [50]. In the Former Soviet Union, the regional status of capital has proven important to predict urban growth [51], the coefficient regressed against growth rates over time ranges from +0.24 point between 1989 and 2002 to +1.88 between 1926 and 1939 [45]. Besides, cities that belong to the same territory have shown an increased pattern of synchronicity in their growth and decline trajectories from the 1980s on, suggesting evidence of both political shocks and territorial solidarity in the spatial distribution of urban growth.

Despite the corroboration of theories provided by the numbers cited above, statistical correlations do not prove any causal chain and fail to explain the processes at work. Moreover, they are not adapted to model dynamic, non-linear and complex evolutions and interactions. This is a recurrent problem that has led social scientists to promote a combination of statistical methods with generative (simulation) modelling [52,53].

### 2.3. A Case Study : The Former Soviet Union

We applied such a mixed methodology to the case study of the urban Former Soviet Union (FSU) of the last 50 years [45,54]. This system of cities is mainly interesting to us because it has the reputation of having been a controlled economic, political and social system aiming at reaching equalisation (of city size for example) and yet, while observing its evolution with generic urban models, the urban trends appear very classic [45,55]. By this, we mean for instance that the evolution of the percentage of urban population follows a very classical logistic function of urban transition, that the rank-size slope is close to the value expected with respect to the time of settlement of the different parts of the territory, and that it has increased over time, indicating an increase of inequality of city sizes (despite anti-urban political discourses). The monographic explanation of urban growth in the Former Soviet Union has been dominant in urban geography, and we argue that it might have hidden very generic processes of urbanisation in this system of cities.

Our modelling experiment aims at testing the genericity of the Soviet and post-Soviet urbanisation by simulating generic mechanisms and comparing their results with harmonised historical data. The database used in this study (DARIUS: Demography of Agglomerations in the perimeter of Imperial Russia and the former Soviet Union [56]) consists of open demographic, territorial and site informations on 1929 agglomerations of the FSU between 1840 and 2010.



We analyse and compare the explaining power of generic mechanisms in a simulation of the period before (1959–1989) and the period after (1989–2010) the dislocation of the Soviet Union. The agent-based model framework of this experiment is called MARIUS (Models of Agglomerations in the perimeter of Imperial Russia and the former Soviet Union. For further details, see [54]). In this model, cities interact as collective agents [34] and we evaluate every simulations with three evaluation criteria. The first two criteria are boolean indicators of the “realism” of the microdynamics of the simulation: we only analyse simulations in which the number of cities with a nil wealth and the number of cities producing more output in one year than their stock of wealth is 0. These codified “controlling” criteria help filtering implausible sets of parameters from the exploration, in the manner of patterns in Pattern-Oriented Modelling (POM) [57,58]. Once they are met, we try to minimise the distance  $\delta$  between the simulated population  $P_s$  and the observed population  $P_o$  of each city  $i$  at each date  $t$  for which we have census information, as stated in Equation (1) (cf. [54]).

$$\delta = \sum_t \left( \sum_i (\log(P_{o,i,t}) - \log(P_{s,i,t}))^2 \right) \quad (1)$$

This criterion ensures that a “perfect” model (achieved for  $\delta = 0$ ) would predict the right amount of urban growth and its exact location in the different cities over time. In order to compare different time periods, we normalise  $\delta$  by the number of cities and the number of censuses used in a given period.

### 3. Modular Multimodelling Experiment

The incentive to implement competing and complementary theories into different models evaluated against one another is a recurrent plea in the simulation literature [7,59,60,62,82]. It reveals how tricky its implementation and automatic evaluation might be, besides the epistemological challenge of equifinality and the kind of conclusions one can draw from this confrontation. Indeed, thirty years after the “Automated Modeling System to Explore a Universe of Spatial Interaction Models” by Openshaw [61], there are no standard tools nor formal methodology for theory testing with simulation models. Indeed, Openshaw’s automated way to explore model structures, being a pure optimisation way of discovering model structures, is impressive methodologically but it does not suit our goal of theory testing in a virtual laboratory, because it can result in optimal models that are impossible to interpret. Instead, we think that the first step should be to gather a catalogue of theoretical processes and mechanistic hypotheses working as potential explanations. This usually precedes the mixed-modelling step [63,64] and prevents from endlessly building models “from scratch” [62]; it allows to build on previous work and experience.

Following the ODD protocol [65], we briefly present an overview of the model (O, Section 3.1), its design concepts (D, Section 3.2) and implementation details (D, Section 3.3). More specifically, we distinguish between a baseline model of urban interactions and the hypotheses implementing our catalogue of mechanisms as modular blocks, that can be activated or discarded to compose a family of simulation models (Section 3.3).

#### 3.1. Overview

This family of models aims at simulating the demographic growth of cities of the Former Soviet Union, in terms of magnitude, rhythm and spatial location. The concept of model family points to the

fact that we consider several mechanisms as candidates for explaining the historical processes observed, and we aim to test and compare their ability to simulate the (post-)Soviet urban growth.

Agents in the model are collective entities: cities. We model between 1145 and 1822 such cities (depending on the modelling period) for each simulation. This number corresponds to the number of urban agglomerations recorded at the initial date for the period under enquiry in the Soviet Union [56]. Two periods are considered: a Soviet simulation (from 1959 to 1989, in 30 time steps of one year) and a post-Soviet one (1989 to 2010, based on Russian census dates). A city (low level) is described by a location (absolute and relative, such as the belonging to a region), a population and quantities of production, consumption and total wealth. The system of cities (high-level) is described by the distribution of city sizes.

At each step:

- Each city updates its economic variables based on its current population;
- Cities interact (*i.e.*, exchange product) with other cities according to their supply, demand and distance;
- Each city updates its wealth based on the results of its interactions;
- A simulation step ends when each city updates its population.

This baseline schedule is modified when additional mechanisms are activated in the model structure. We detail these modifications in the following detailed description of additional mechanisms (Section 3.3).

### 3.2. Design Concepts

The family of models includes a weak concept of emergence [66]. This means that no surprising pattern emerges from cities' interactions, but that they organise so as to produce patterns of increasing (or decreasing) [11] levels of hierarchy of their size distribution. Cities adapt to their interaction network by adjusting their (population) size. They sense their environment in two additional mechanisms, when they can benefit from underground resources or regional rural migrations. Their interactions are monetary as cities do not exchange residents: they trade production value (either bilaterally in the baseline model or through the mutual redistribution of taxes in the redistribution mechanism). Cities are part of regional collectives when the redistribution or urban transition mechanisms are activated. No fitness is computed at the city level: we keep this term for the calibration, to qualify the fitness between the simulated size distribution and the historical one over time. The model finally involves neither prediction nor stochasticity.

### 3.3. Implementing Mechanisms as Building Blocks (Details)

The multi-model is composed of the baseline model and additional modules of mechanisms that override the sequence of agents' rules when they are activated. In the following sections, we detail the baseline model and the modular blocks of mechanisms that are added incrementally to the original equations. For further information, the baseline model and the first two additional mechanisms were described and evaluated in details in [54].



### 3.3.1. The Baseline Model

The baseline model relies on the assumption that population and wealth are the basic descriptors of cities and the engine of their co-evolution. It therefore models exclusively size effects of population on wealth and spatial interactions.

At initialization, each city  $i$  of the Former Soviet Union is setup with its historical population  $P_i$  at the beginning date of simulation, and located at its empirical coordinates, enabling site, situation and distance to play in the same geometry as in the target system. An estimated value of wealth  $W_i$  (expressed in a fictive unit) is determined for each city with respect to its size, following Equation (2).

$$W_i = P_i^{populationToWealth} \quad (2)$$

$$populationToWealth \geq 1 \quad (3)$$

This first mechanism is a first possibility of implementing the theoretical hypothesis of agglomeration economies. Wealth is indeed distributed superlinearly for each value of  $populationToWealth$  significantly greater than 1. Time is modelled as discrete steps, each of which represents a time period of one year, during which interactions occur in a synchronous way. At each step:

- **Each city  $i$  updates its economic variables:** a global supply  $S_i$  Equation (4) and a global demand  $D_i$  Equation (7), according to its population  $P_i$  and three parameters ( $economicMultiplier$ ,  $sizeEffectOnSupply$  and  $sizeEffectOnDemand$ ).

$$S_i = economicMultiplier \times P_i^{sizeEffectOnSupply} \quad (4)$$

$$economicMultiplier > 0 \quad (5)$$

$$sizeEffectOnSupply \geq 1 \quad (6)$$

$$D_i = economicMultiplier \times P_i^{sizeEffectOnDemand} \quad (7)$$

$$sizeEffectOnDemand \geq 1 \quad (8)$$

- **Each city interacts with other cities** according to the intensity of their potential  $IP$  Equation (9). For two distinct cities  $i$  and  $j$ , the computation of the interaction potential  $IP_{ij}$  consists in confronting the supply of  $i$  Equation (11) to the demand of  $j$  with an equation borrowed to the gravity model Equation (12).

$$IP_{ij} = \frac{S_i \times D_j}{d_{ij}^{distanceDecay}} \quad (9)$$

$$distanceDecay \geq 0 \quad (10)$$

with  $d_{ij}$  a measure of distance between  $i$  and  $j$ .

Interactions of cities  $i$  and  $j$  based on their potential  $IP_{ij}$  result in a transaction  $T_{ij}$  from  $i$  to  $j$  Equation (13).

$$S_{ij} = S_i \times \frac{IP_{ij}}{\sum_k IP_{ik}} \quad (11)$$

$$D_{ij} = D_i \times \frac{IP_{ji}}{\sum_k IP_{ki}} \quad (12)$$

$$T_{ij} = \min[S_{ij}, D_{ji}] \quad (13)$$

- **Each city updates its wealth**  $W_i$  according to the results of its transactions  $T$  (unsold supply  $US_i$  Equation (15) and unsatisfied demand  $UD_i$  Equation (16)) in which it was committed Equation (14).

$$W_{i,t} = W_{i,t-1} + S_i - D_i - US_i + UD_i \quad (14)$$

$$US_i = S_i - \sum_j T_{ij} \quad (15)$$

$$UD_i = D_i - \sum_j T_{ji} \quad (16)$$

- **A simulation step ends when each city updates its population** according to its new resulting wealth Equation (17) :

$$P_{i,t} = P_{i,t-1} + \frac{W_{i,t}^{wealthToPopulation} - W_{i,t-1}^{wealthToPopulation}}{economicMultiplier} \quad (17)$$

### 3.3.2. Mechanism Increments

- **The mechanism that accounts for interactions benefits** at the intercity level is the one called **bonus**. It

“[...] features a non-zero sum game [...], rewarding cities who effectively interact with others rather than internally. We assume that the exchange of any unit of value is more profitable when it is done with another city, because of the potential spillovers of technology and information [54].”

This bonus  $B_i$  depends on the volume of transactions and the diversity of partners  $J_i$  the city  $i$  has exchanged with Equation (18). It is added to the wealth at the end of each step Equation (19), following Equation (14):

$$B_i = bonusMultiplier \times \frac{(\sum_j T_{ij} + \sum_j T_{ji}) \times J_i}{n} \quad (18)$$

$n$  being the total number of cities in the system (*i.e.*, 1145 for a simulation beginning in 1959, 1822 for a simulation beginning in 1989), and  $J_i$  the number of partners with which  $i$  has engaged in the current simulation step.

$$W_{i,t} = W_{i,t} + B_i \quad (19)$$

- **A mechanism related to situation advantages** is called **fixed costs**. It ensures that the situation of each city in the system is taken into account in its interactions with other cities.

“Every interurban exchange generates a fixed cost (the value of which is described by the free parameter *fixedCost*). This implies two features that make the model more realistic: first, no exchange will take place between two cities if the potential transacted value is under a certain threshold ; second, cities will select only profitable partners and not exchange with every other cities. This mechanism plays the role of a condition before the exchange” [54].

The interaction potential between city  $i$  and city  $j$  will be positive only if the potential value that  $i$  is willing to sell to  $j$  is superior to the fixed value it costs it to negotiate, prepare and send the transaction Equation (20):

$$IP_{ij} = \begin{cases} IP_{ij} & \text{if } S_{ij} > fixedCost, \\ 0 & \text{otherwise.} \end{cases} \quad (20)$$

Therefore, each transaction of a city  $i$  gives way to a fixed cost. Their sum is subtracted from the wealth of city  $i$  at the end of each step Equation (21), following Equation (14):

$$W_{i,t} = W_{i,t} - J_i \times fixedCost \quad (21)$$

- **Site effects** are targeted by the **resource** mechanism: site advantages are particularised in this model by natural resource deposits (more specifically: coal deposits  $C$  on the one hand, and oil and gas deposits  $O$  on the other hand). The assumption is made that if the city  $i$  is located on some coal or oil deposits ( $C_i = 1$  or  $O_i = 1$ ), the city benefits from the advantage granted by the extraction activity. The capacity of extraction depends on the capital (wealth) of the city and takes the form of a wealth multiplier for each resource Equation (22) after Equation (14):

$$W_{i,t} = W_{i,t} \times \left(1 + \begin{cases} coalEffect & \text{if } C_i = 1, \\ oilAndGasEffect & \text{if } O_i = 1 \end{cases}\right) \quad (22)$$

- **Territorial and political effects** are formalised by the **redistribution** mechanism. It allows for a redistribution of wealth between cities of the same territory  $R$  (region or State). To do so, territorial taxes  $tt_k$  are collected in each city  $k_R$ , as a proportion *territorialTaxes* of their wealth. The total amount of taxes collected is  $TT_R$  Equation (23):

$$TT_R = \sum_i tt_{i,R} = \sum_i territorialTaxes \times W_{i,R} \quad (23)$$

From this taxes, the administrative status of the territory  $R$  (denoted by  $CC_{i,R}$ , set to 1 if  $i$  is the capital city of the region and 0 otherwise) allows the capital city to take a share  $CS$  for its administration needs Equation (24):

$$CS_R = capitalShareOfTaxes \times TT_R \quad (24)$$

The rest of the taxes is redistributed to cities of region. Each city  $i_R$  receives a share  $tr_{i,R}$  that is proportionate to its population Equation (25):

$$tr_{i,R} = (TT_R - CS_R) \times \frac{P_{i,R}}{\sum_k P_{k,R}} \quad (25)$$

The balance of the territorial redistribution is added to the wealth of a city Equation (26) after Equation (14):

$$W_{i,t} = W_{i,t} - tt_{i,R} + tr_{i,R} + \begin{cases} CS_R & \text{if } CC_{i,R} = 1, \\ 0 & \text{otherwise} \end{cases} \quad (26)$$

- **Finally, territorial and situation explanations** are mixed in the **urban transition** mechanism. To account for the different opportunities of cities to attract rural migrants in the different regions, we model the evolution of the urban transition curves over time. As shown empirically [45], 100 out of the 108 regions of the Former Soviet Union have followed the scheme of the urban transition. It means that their urbanisation rate  $U_R$  (in %) has followed a logistic function over time  $t$  Equation (27):

$$U_{R,t} = \frac{100}{1 + e^{-urbanisationSpeed \times t}} \quad (27)$$

The parameter *urbanisationSpeed* is not a free parameter, thus it will not be calibrated. It has been empirically determined as the “mean” transition regime that minimises the error of logistic adjustment performed on data depicting the urbanisation rate over time for each region of the post-Soviet Union [45]. This was done to position every region on a single urbanisation curve with respect to the historical time lags their urban transition. Indeed, western regions were already mainly urban in the 1960s whereas some central asian countries (Tajikistan, Kirghizstan) are still dominantly rural nowadays. This imply strong disparities in migration opportunities from the cities. In order to model a generic process of urban transition, we located the different regions at different stages of the same transition curve instead of considering a specific urbanisation curve for each region. The consequence of this initialisation process is that each region will move one step further on the urbanisation curve (leading eventually to 100%) at each simulation step, but that the rural potential to migrate in cities will depend on its current position on the urban transition curve (high potential for weakly urbanised regions, small potential for regions already very urban). The migration potential of each city  $i$  in territory  $R$  is built as a multiplier  $TM_R$  specific to each region for each time step Equation (28):

$$TM_{R,t} = 1 + (1 - U_{R,t} \times ruralMultiplier) \quad (28)$$

This extra population growth is added Equation (29) after Equation (17) and the new urbanisation rates of regions are updated for  $t + 1$  Equation (27):

$$P_{i,R,t} = P_{i,R,t} \times (1 + TM_{R,t}) \quad (29)$$

Because we want to evaluate the contribution of each theoretical mechanism to the simulation of urbanisation and its interactions with other mechanisms, we need the modules to be easily activated and de-activated dynamically in the model. To that extent, we leveraged the mixin-methods system of the Scala programming language [67]. The mixins were first proposed at the beginning of the 90 s in the Jigsaw programming language [68]. They are now adopted in mainstream languages such as Scala. It has been established as a powerful way to perform type-safe dependency injection framework [69], which is a powerful paradigm to achieve modularity. The mixin pattern allows to achieve type safe dependency injection. Mixins are therefore a suited tool to achieve modularity in model implementations, which we use to implement MARIUS in Scala language (to access the source code, cf. [70]).

### 3.4. Technical Modular Implementation

The mixin-methods concept [71,72] generalizes object-oriented programming to enable feature-oriented programming [73]. It makes the definition of class hierarchy more flexible [74]. Variations of isolated features (such as the different update functions of city wealth and population in our model) are defined in separate modules and mixed-in with each other at the object instantiation point, also called mixin application. The advantage of mixins (or trait) over classical object-oriented programming is the possibility of defining numerous variations of several features without increasing exponentially the number of specialized class.

For instance, let us consider a class *C* implementing two methods **a** and **b**. We could think of class *C* as the interaction potential function, **a** representing the basic implementation of Equation (9) and **b** the fixed cost selection of Equation (20). To define alternative implementations of those methods in the classical object-oriented paradigm, one would implement subclasses of *C* and override the implementations of **a** and **b** in each subclass. For instance, in the Listing 1, the class *C1* specialises *C* and defines implementations for the methods **a** and **b**.

Listing 1: Object oriented specialisation.

```
abstract class C extends A with B {
  def a(x: Double): Double
  def b(x: Double): Double
  def c(x: Double) = /* Compute something using a and b */
}

class C1 extends C {
  def a(x: Double) = /* Some implementation of a */
  def b(x: Double) = /* Some implementation of b */
}
```

This pattern achieves a very low level of code reusability. Let **a** and **b** have 10 possible implementations, then 100 specialised implementations of *C* would be required. The mixins method solves the problems of combinatorial explosion of the number of implementations by delaying the entanglement of the class components at the instantiation site. Listing 2 exposes an implementation based on mixins providing alternative implementations of **A** and **B** and the corresponding parameter specifications. The implementation choice is delayed until the instantiation point (last lines of Listing 2) at which a mixin is defined.

Listing 2: Mixin in Scala.

```
trait A {
  def a(x: Double): Double
}

trait B {
  def b(x: Double): Double
```

```
}

trait C extends A with B {
  def c(x: Double) = /* Compute something using a and b */
}

// Implementation 1 of trait A
trait A1 extends A {
  def a(x: Double) = /* Some implementation */
}

// Implementation 2 of trait A
trait A2 extends A {
  // Parameter p0 used in this version of a
  def p0: Double
  def a(x: Double) = /* Some implementation using p0 */
}

// Implementation 1 of trait B
trait B1 extends B {
  def b(x: Double) = /* Some implementation */
}

// Implementation 2 of trait B
trait B2 extends B {
  def b(x: Double) = /* Some implementation */
}

val instance1 =
  new C with A1 with B2 {}

val instance2 =
  new C with A2 with B2 {
    // Value for parameter p0 of trait A2
    def p0 = 1.0
  }
```

Scala traits expressiveness can be leveraged to implement modular evolutive type-safe modelling frameworks proposing alternative model features, feature composability and the formalisation of the feature dependencies. The implementation of alternative behaviours in several traits provides the isolation of model component implementations and explicit dependencies between these components.



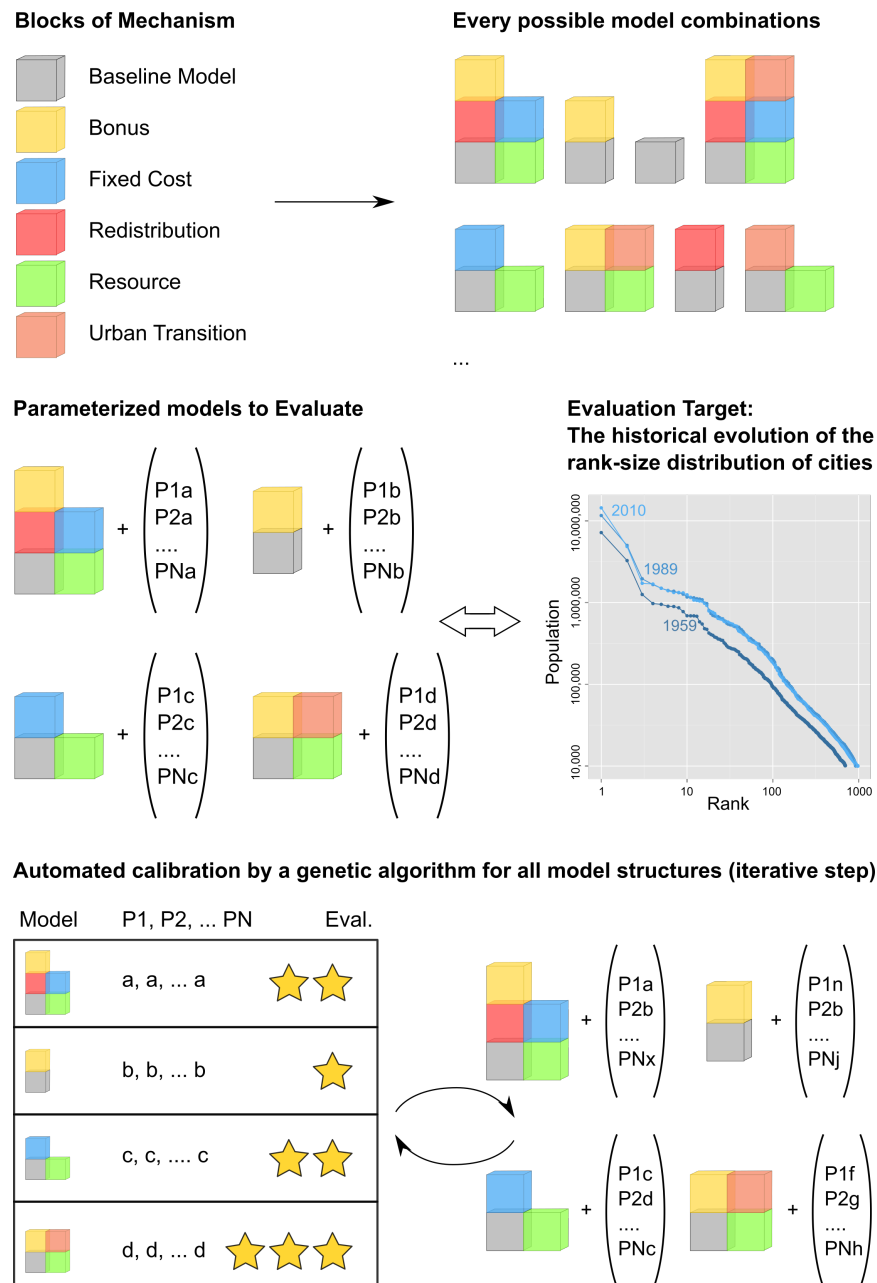
Each component defines free parameters that have to be set at the model instantiation site, otherwise it won't compile.

In this experiment, we defined each alternative model mechanism in a particular trait. The executable model has to be composed by picking the traits we wanted to test. In order to evaluate concurrently all the alternative mechanisms, we generated all the possible models (or combination of mechanisms). It was achieved using a code generation algorithm which produces all the possible models implementations by generating a Scala source code containing all the possible traits combinations, such as the one shown on Listing 3.

Listing 3: Example of generated code.

```
def model(index: Int , parameters: Seq[Double]) =
  index match {
    case 0 =>
      new Model with T11 with T21 with ... {
        def p0 = parameters(0)
        def p1 = parameters(1)
        ...
      }
    case 1 =>
      new Model with T11 with T22 with ... {
        def p0 = parameters(0)
        def p1 = parameters(1)
        ...
      }
    case 2 => ...
  }
```

This generated source code implements a single function encapsulating all the model alternatives. This function takes two arguments: an index of the implementation that shall be executed and a vector of parameters to set for the model (for this work, we calibrated only double precision floating points values). Note that the vector has a fixed size which does not depends on the model instantiated. A given model implementation generally does not use all the parameters: instead it will use only some of them and ignore the others (Figure 1).



**Figure 1.** Multi-Calibration Protocol. This schema proposes a simplified graphic description of the modular model building (**top row**) together with the empirical evaluation process (**middle row**) and the iterative genetic calibration of the model structures (**bottom row**).

### 3.5. Calibrating a Multi-Model

“Whatever changes occur in the institutional, political and social context of computational models, the question of how to learn from models remains. It is clear that assessment of the accuracy of a model as a representation must rest on argument about how competing theories are represented in its workings, with calibration and fitting procedures acting as a check on reasoning” ([7] p. 291).

In order to evaluate the capacity of the models to reproduce the historical trajectory of urbanisation in the Former Soviet Union, we rely on an automated calibration. This procedure is part of the model evaluation [75,76,78]. Its aim is to find the values of parameters for which the model results match the fitness criteria (here:  $\delta$ , the lowest possible distance to the data, once the realism criteria are met). If the model can be calibrated, then it is shown that the mechanisms included in the model are sufficient to simulate the urban trajectory (that does not prove yet that they are all necessary). If there are no parameter value for which the fitness criterion is met, then the combination of mechanisms is not sufficient to model urbanisation under the current implementation.

In order to calibrate all the models at once, we designed a variant of the NSGAI genetic algorithm that includes a niching mechanism [79]. Niching methods aim at preserving suboptimal solutions to preserve diversity. Our niching algorithm divided the population of parameterised models (that is, one model with one vector of parameter values) into sub-populations, each sub-population containing one model alternative (every model in the sub-population have the same mechanism structure). The genome of each individual (an individual corresponds to a vector of parameter values and structure index that defines the model under evaluation) contains two parts. The first part is an integer value that corresponds to the index of the model alternative on which the genome is evaluated. The second part is a vector of double values containing the values of all the parameters for the model.

In order to evaluate a genome, we designed a fitness function. This function calls the generated function described above, runs the model and evaluates its dynamics using the fitness function described in Section 2.3 (*i.e.*, the criteria of realism of micro-dynamics and the distance  $\delta$  between simulated and observed population data for each city at each census date, Figure 1).

In NSGAI, the elitism operation preserves the best individuals among the whole population. The evaluation algorithm we applied has the exact same elitism strategy for each sub-population. No global elitism strategy was performed; instead, we kept the 50 best-performing individuals in each sub-population (or model combination of mechanisms). In order to speed up the convergence of the algorithm, we also tweaked the mutation operation: it had a 10% chance of mutating the “model index” part of the genome. The new value for the “model index” was drawn uniformly among the possible model indexes. This allowed to periodically test on other models, some parameter values that were performing well on a given model.

We then distributed this algorithm on the European Grid Infrastructure (EGI) using the technique known as the “island model” in the same way as we described in [76]. We ran 200,000 jobs of 2 h. One model execution being of 40 s on average, this experiment corresponds to approximately 72 million evaluations of parameterised models. The set we analyse in the following section corresponds, for each of the 32 different structures of models (64 in fact because the model can be instantiated for two different time periods: 1959–1989 and 1989–2010), to the 50 vectors of parameter values that resulted in the smallest distance to empirical data, so altogether 3200 performant parameter sets associated with a model structure and a performance measure (Figure 1). Instantiated model structures can be for example: baseline model + 1959–1989, baseline model + fixed cost + 1989–2010, baseline model + urban transition + resources + 1989–2010, baseline model + redistribution + urban transition + 1959–1989, baseline model + fixed cost + bonus + resources + 1959–1989, *etc.*. Those are part of the 64 instantiated model structures for which 50 parameter sets were analysed.

#### 4. Results: Hypothesis Testing to Explain Urbanisation in the Former Soviet Union

The analysis of calibration results consists in relating the performance of these calibrated models (in terms of distance between simulated and observed growth hierarchically and spatially) to their structure (their mix of mechanism) and the values of parameters associated with each activated mechanism. We detail each analysis below, and invite the reader to replicate them using the interactive application VARIUS, built to explore those results online [77]. We present the results of this exploration in the form of three questions at the macro, meso and micro scale of the city-system.

- 1 Which is the most parsimonious model to simulate the evolution of cities before and after the collapse of the Soviet Union?** A way to answer this question is to restrict the set of results to the five model structures that correspond to the a mix of two mechanisms: the baseline model + one additional mechanism (for example: resource extraction). That leaves us with  $50 \times 5 = 250$  parameter sets and 250 performance values  $\delta$  for each time period analysed. We look at the lowest distance to empirical data achieved in this set of results, and identify the corresponding model structure (the additional mechanism involved) and parameter values as the best performing ones.

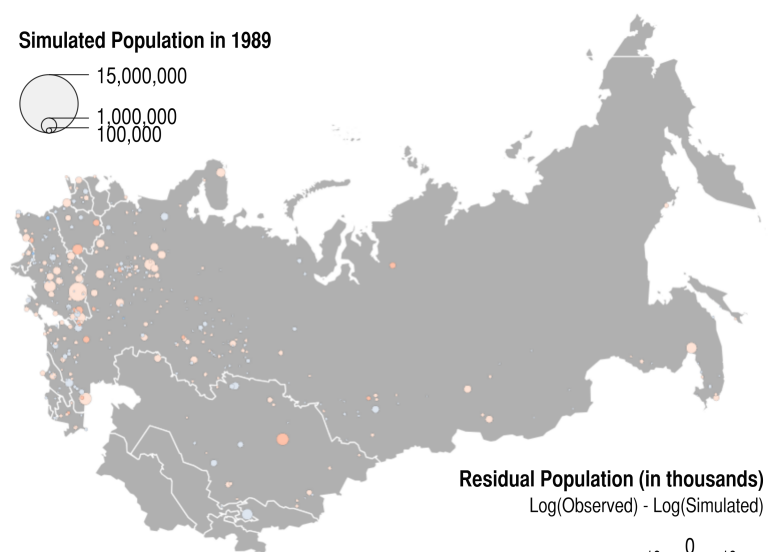
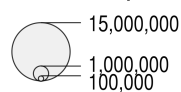
For simulations starting in 1959 up to 1989 (the Soviet Union actually came to an end in December 1991 but the last Census of the Soviet Union was performed in 1989), the most performant parameterised model with only one additional mechanism is composed of the baseline model plus the mechanism of urban transition (cf. Table 1). It is characterised by significant economies of agglomeration ( $sizeEffectOnSupply = 1.05$ ) but a linear function of demand with size. The rural multiplier is equal to 3.5% and allows to simulate fast urbanising regions of Siberia and Central Asia (cf. Figure 2). However, the population of a majority of cities is under-estimated in the simulation, especially in the upper part of the hierarchy, around Moscow and in eastern Ukraine (or more generally in the Western part of the Former Soviet Union. See question 3 for an hypothesis as to why that might be.).

**Table 1.** Parameter values of the best performing simple model before the political transition.

Parameter Name	Value	Mechanism
economicMultiplier	0.002193758	Baseline
populationToWealth	1.000184755	Baseline
sizeEffectOnSupply	1.053943022	Baseline
sizeEffectOnDemand	1.000000000	Baseline
wealthToPopulation	0.203567639	Baseline
distanceDecay	1.872702086	Baseline
ruralMultiplier	0.034975771	UrbanTransition
Normalized $\delta$	n cities	Time steps
0.01423387	1145	30

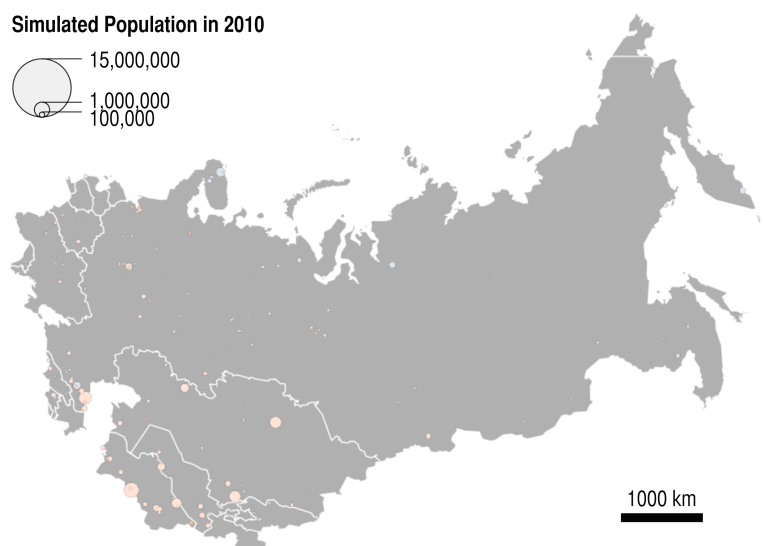
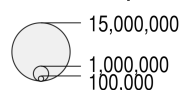
### Baseline Model + Urban Transition (1959-1989)

Simulated Population in 1989



### Baseline Model + Resources (1989-2010)

Simulated Population in 2010



**Figure 2.** Spatial distribution of residuals for the parsimonious models. Only cities largely over- and under-estimated by the simulation are mapped in this figure. The size of the circles indicate the population of the city as simulated at the end of the period. Dark colours indicate large discrepancy: overestimation of the population in blue (negative residual) and underestimation in orange (positive residual). N.B.: To reproduce those maps on the application VARIUS, run the corresponding model and choose a residual cutoff of 0.3.

After the (political) transition, the best calibrated model with one additional mechanism includes site advantages for cities. This model has two additional parameters and meets the evaluation criteria (per city per census) twice better than the best model for the previous period (0.005 vs. 0.01, cf. Tables 1 and 2). The analysis of the parameters fits the empirical observations of faster growing cities located on oil and gas deposits (*oilAndGasEffect* = 0.02), and declining cities in the Donbass and Kuzbass coal regions (*coalEffect* = −0.01). This model's specifications include very low size effects and a very uneven

wealth distribution amongst cities at initialisation ( $populationToWealth = 1.12$ ). The residuals are distributed roughly symmetrically (Figure 3), but without any mechanism of urban transition, the model clearly underestimates the post-1989 growth of all the rapidly growing cities of Central Asia.

**Table 2.** Parameter values of the best performing simple model after the political transition.

Parameter Name	Value	Mechanism
economicMultiplier	0.502616330	Baseline
populationToWealth	1.124963276	Baseline
sizeEffectOnSupply	1.002982515	Baseline
sizeEffectOnDemand	1.000808442	Baseline
wealthToPopulation	0.699943763	Baseline
distanceDecay	1.475836151	Baseline
oilAndGazEffect	0.017066495	Resources
coalEffect	−0.011792670	Resources
Normalized $\delta$	n cities	Time steps
0.005180008	1822	21



**Figure 3.** Distribution of residuals for the parsimonious models. (a) Urban Transition model. 1959–1989; (b) Resource Extraction model. 1989–2010.

To summarise, situation effects and territorial effects seem to be the dominant candidates for explaining the specific part of the trajectories of cities in the Soviet era, while site effects seem to have taken over since 1991, the transition to capitalism and the rise of oil prices in the world markets. Moreover, the better fit of the latter model could be an indication of the “normalisation” of the urban processes in the post-Soviet space, compared to a more singular pattern of Soviet urbanisation.

**2 Which are the mechanisms (and mechanisms’ interactions) that are essential to model the Soviet and post-Soviet urbanisation patterns?** To address this question, we statistically analyse the results of the multicalibration (3200 sets of parameters, the best 50 of each model structure

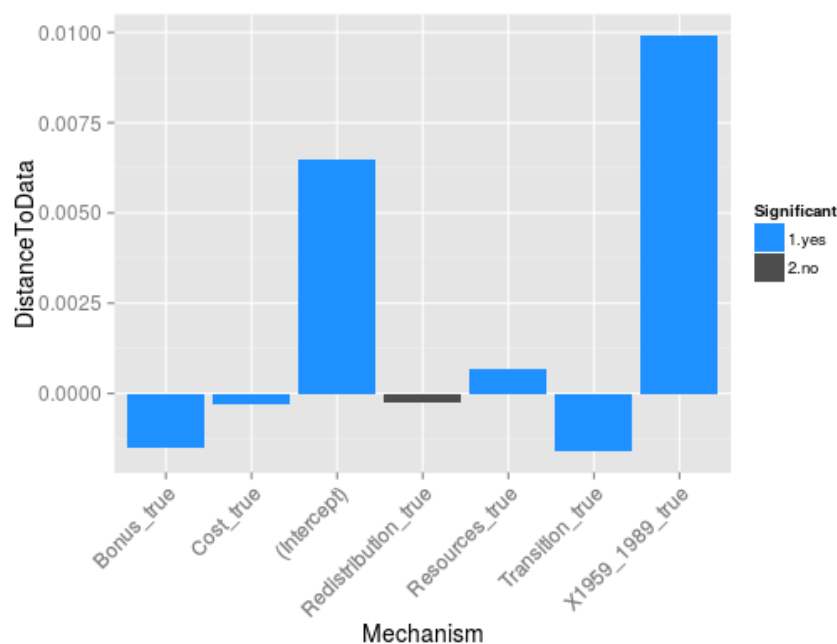


for each time period) to evaluate the contribution of each mechanism (everything else being equal in the model structure) to the reduction of distance between simulated and observed demographic data for each city. More specifically, we regress the distance to data  $\delta$  against the mechanism composition of the model, following Equation (30):

$$\delta_i = a + b_1 * BONUS_i + b_2 * FIXEDCOST_i + b_3 * RESOURCE_i + b_4 * REDISTRIBUTION_i + b_5 * URBANTRANSITION_i + b_6 * PERIOD_i + \epsilon_i \quad (30)$$

with  $\delta$  the distance between simulated and observed trajectories,  $i$  one of the 3200 parameterised models considered,  $BONUS_i = 0$  if the model includes the bonus mechanism, *etc.*,  $PERIOD_i = 0$  if the model is run between 1989 and 2010,  $PERIOD_i = 1$  if the model is run between 1959 and 1989.

The estimated coefficients ( $b_1, b_2, b_3, b_4, b_5, b_6$ ) associated with each mechanism correspond to the average contribution of the mechanism to the distance  $\delta$  when activated, everything else being equal. They are represented in Figure 4, along with the intercept ( $a$ ), that is the average  $\delta$  of the baseline model between 1989 and 2010.



**Figure 4.** Contribution of mechanisms to a simulation that reduces the distance to observed data. N.B.: Estimated coefficients ( $a = "(intercept)"$ ,  $b_1 = "Bonus"$ ,  $b_2 = "Cost"$ ,  $b_3 = "Resources"$ ,  $b_4 = "Redistribution"$ ,  $b_5 = "Transition"$ ,  $b_6 = "X1959 - 1989"$ ) are considered significant for  $p$ -values inferior to 0.005. A negative coefficient indicates that the mechanism contributes to reducing the distance to data, on average, when activated in the model (the mechanism is thus called “essential” to simulate the urban hierarchical evolution of the Former Soviet Union). A positive coefficient indicates that the mechanism increases the distance to data, on average, when activated in the model (and is therefore harmful to a realistic simulation of cities growth). A non-significant coefficient can indicate two things: either the mechanism is neutral in the simulation, either its contribution to reducing the distance to data depends on the presence of other mechanisms.

We find that on average, the bonus, fixed costs and urban transition mechanisms tend to reduce the simulation error significantly (cf. Figure 4). The transition and bonus mechanisms are identified to be the most effective ones. By contrast, the resource mechanism is correlated significantly with a change in the evaluation criteria but tends to increase the error when it is activated (compared with model structures without this mechanism). This counter-intuitive result might be linked to the weak influence of resource extraction for the first period of simulation, when there is a diversity of urban trajectories within resource-rich regions. The redistribution mechanisms does not appear significant on average in this analysis, as it plays a minor role in the reduction of error when activated.

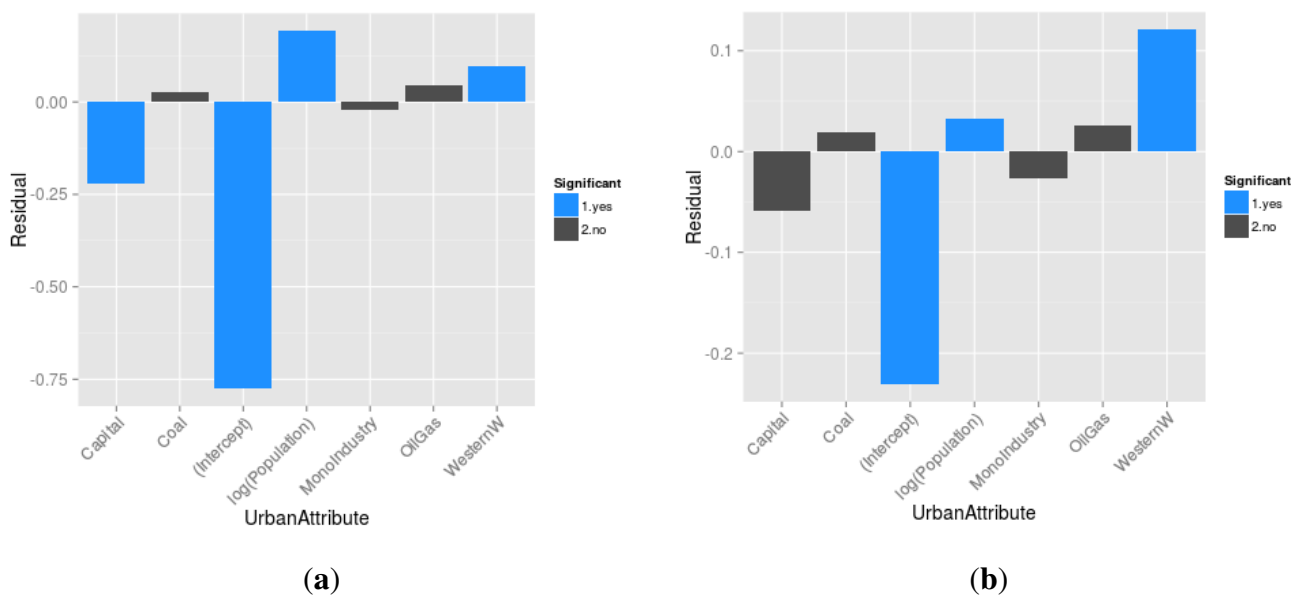
Finally, we confirm the observation that our models work better to simulate the urban evolution following the dismantlement of the Soviet Union. Indeed, as shown in Figure 4, the value of normalized  $\delta$  is greater by 0.01 point on average when the model is specified for the first period. This could be an expression of “normalisation” or simplification of the urban processes in the post-Soviet space after the political and economic transition of the 1990s.

**3 What are the cities that resist modelling? In other words, what are the cities that are too specific to be modelled by any of the mechanisms implemented?** To answer this last question, we statistically analyse the difference between the log of the population observed and the log of the population simulated at the last evaluation Census date for cities included in the simulation, with respect to their locational and functional attributes. The models for which we present the results below contains all the implemented mechanisms, and are applied to the two periods of enquiry.

For the two periods, we find that our models persistently and significantly under-estimate the growth of the largest and most western cities of the (Former) Soviet Union, everything else being equal (cf. Figure 5). Moreover, capital cities appear to have grown less historically than what we can predict with a complete model of the period 1959–1989. The other urban attributes included in the regressions (natural resources and mono-functional specialisation) do not seem associated with any systematic pattern of over- or under-estimation.

The difficulty to reproduce the trajectory of the largest cities has been encountered for a comparable model of system of cities (Simpop2, see [80]) and solved by the exogenous introduction of innovations to account for the creative features and higher probability of adoption of new technology and functions by the largest cities.

The under-estimation of growth in the western part of the territory might be due to its integration within a larger area (the Eastern Europe) during the periods under study: our hypothesis here is that the centrality of western (post-)Soviet cities would then be minored in our model because it does not take into account the interactions with east-European cities (Warsaw, Prague, Bratislava, *etc.*) which formed altogether an economic system (even though the integration was always stronger within the FSU).



**Figure 5.** Profiles of residuals. Estimated coefficients are considered significant for  $p$ -values inferior to 0.01. These graphs should be read similarly to Figure 4: negative coefficients indicate that, on average, cities with the attribute under consideration are overestimated by the simulation, whereas positive coefficients indicate average underestimation of the population of cities with the attribute in the simulation. The attribute “log(population)” is the only quantitative one and the coefficient indicate the increase of residual value for an increase of 1 of the log of population (in thousands). Finally the intercept coefficient indicates the average residual value of cities with none of the attribute considered and a population of 1000. (a) Complete model. Simulation 1959–1989; (b) Complete model. Simulation 1989–2010.

Finally, some individual cities appear as clear outliers of the model, and could correspond to a profile that is too specific to be modelled by any generic mechanism. For the first period, (cf. Table 3), the most obvious examples of singular trajectories are the cities which grew much faster than what was expected from their site, situation or interaction attributes. Indeed, Naberezhnye Tchelny, Volgodonsk, Toljatti or Bratsk owe their sudden development to political decisions to implement flagship projects: automobile industry mega-plants in Naberezhnye Tchelny (trucks) and Toljatti (cars), energy production sites in Volgodonsk (atomic power) and Bratsk (hydroelectric power station). These economic policies of the 1950s and 1960s led those cities to be four times as populated 30 years later than what was expected from their interactions, resource or regional characteristics.

For the second period, (cf. Table 4), Astana is a good example of a similar singular trajectory that we would not aim to simulate with generic urbanisation mechanisms, as it owes its booming growth to the decision of the Kazakh newly independent State to locate its headquarters in this city more central to the country (compared to Almaty). On the contrary, Baikonyr, also in Kazakhstan, has suffered from the cuts in the space industry (non-predictable at the urban level of our mechanisms). Other shrinking cities like Aleksandrovsk-Sahalinsk, Krasnozavodsk or Uglegorsk would require more detailed mechanisms of demographics (lack of birth and emigration) and economic cycles to be simulated adequately.

**Table 3.** Observed and Simulated populations of urban outliers in 1989.

<b>Positive Residuals</b>		
<b>City</b>	<b>Observed Pop.</b>	<b>Simulated Pop.</b>
Naberezhnye Tchelny	500,000	30,000
Volgodonsk	191,000	36,000
Chajkovskij	86,000	19,000
Toljatti	685,000	158,000
Bratsk	285,000	73,000
Balakovo	197,000	52,000
Tihvin	71,000	20,000
Chervonograd	72,000	21,000
Obninsk	111,000	32,000
Saryjorsk	174,000	53,000
<b>Negative Residuals</b>		
<b>City</b>	<b>Observed Pop.</b>	<b>Simulated Pop.</b>
Zaozernyj	16,000	54,000
Gremjachnsk	21,000	56,000
Atakent/Ilich	15,000	38,000
Kizel	37,000	88,000
Cheremhovo	74,000	172,000
Ilanskij	18,000	42,000
Gornoaltajsk	46,000	102,000
Volchansk	15,000	32,000
Zujevka	16,000	35,000
Taldykorgai	138,000	296,000

**Table 4.** Observed and Simulated populations of urban outliers in 2010.

<b>Positive Residuals</b>		
<b>City</b>	<b>Observed Pop.</b>	<b>Simulated Pop.</b>
Mirnyja	41,000	12,000
Sertolovo	48,000	16,000
Beineu	32,000	11,000
Govurdak	76,000	28,000
Serdar/Gyzylarbat	98,000	37,000
Bayramaly	131,000	53,000
Sarov	92,000	39,000
Turkmenabat/Tchardjou	427,000	185,000
Astana/Tselinograd	613,000	278,000
Dashougouz	275,000	126,000

**Table 4.** *Cont.*

Negative Residuals		
City	Observed Pop.	Simulated Pop.
Sovetabad	11,000	33,000
Zhanatas	21,000	50,000
Krasnozavodsk	13,000	31,000
Gagra	11,000	25,000
Nevelsk	12,000	26,000
Arkalyk	28,000	59,000
Chyatura	14,000	28,000
Aleksandrovsk Sahalinsk	11,000	21,000
Uglegorsk	10,000	20,000
Baikonyr	36,000	67,000

To summarise, there are particular types of urban trajectories that are not simulated well by the model because of its simplicity, and trajectories that are too specific to be modelled. We find that the exploration of our models, their calibration and the analysis of residuals has helped to identify those cities and to suggest some missing mechanisms.

## 5. Discussion

“Despite the fact that the experience of individual cities has become more varied internationally (at least within what might be called the mature economies) there is stronger evidence of a predictable pattern of change, determined by common causal factors, than might be expected given the diversity and variety of cities” ([15] p. 1342).

Systems of cities have attracted a lot of attention from social modellers because of the regularity of their patterns. Instead of regarding the profusion of competing theories as a source of confusion (or suspicion) [81], we proposed a framework to integrate complementary accounts of urbanisation processes into a modular agent-based model. This work of synthesis and testing within a virtual laboratory has been made possible by its automation and the extensive use of computation resources to calibrate sixty-four models structures with empirical data on almost two thousands cities in the Former Soviet Union over the last fifty years. The model provides a basis for comparison of the different theories that can be augmented by new or alternative implementations of mechanisms. It could also be applied to different systems of cities (in space or time).

In the present study, we showed that the multi-modelling approach has helped identify and order the mechanisms that most probably generated the urban pattern or Soviet and post-Soviet urbanisation: situation effects of urban transition before 1991 and resource extraction afterwards. The intuitive mechanism of redistribution, however, has proved insignificant. This method was finally useful to spot the most singular trajectories of cities that we interpret with empirical and monographic knowledge to refine the model.

As signaled in [57], the full exploration of the parameter space of the model carries the risk of included implausible combinations that can lead to misleading conclusions, especially because the models are quite generic. Another limit to this work lies in the fact that the exploration of the parameter space and the exploration of the model structure are performed alongside, although the same model structure can lead to a large diversity of patterns [11]. A current direction of this work is indeed to look at the output space produced by different model structures, and to compare their potential diversity in modelling alternative trajectories.

## 6. Conclusions

To conclude, modelling experiments perform a radical compression of reality and an extreme simplification of individual trajectories, events and persons into synthetic aggregates. The ontological adequacy of the model to real life is therefore necessarily evaluated at an aggregated level that creates the problem of equifinality. No simulation model can elude this question, but everything should be tried to reduce its impact on what can be learned from the modelling experience.

## Acknowledgments

This work has been funded by the ERC Advanced Grant Geodiversity, the University Paris 1 Panthéon-Sorbonne and the Institut des Systèmes Complexes Paris Île-de-France. The results obtained in this paper were computed on the biomed and the vo.complex-system.eu virtual organization of the European Grid Infrastructure (<http://www.egi.eu>). We thank the European Grid Infrastructure and its supporting National Grid Initiatives (France-Grilles in particular) for providing the technical support and infrastructure, and Volker Grimm for stimulating reviewing suggestions.

## Author Contributions

C.C., R.R., P.C., S.R. and D.P. conceived the experiment. C.C., R.R. and P.C. implemented the model, and ran the experiments. C.C. collected the data and analysed the results.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Batty, M. Fifty years of urban modeling: Macro-statics to micro-dynamics. In *The Dynamics of Complex Urban Systems*; Springer: Heidelberg, German, 2008; pp. 1–20.
2. Heppenstall, A.J.; Crooks, A.T.; See, L.M.; Batty, M. *Agent-Based Models of Geographical Systems*; Springer Science & Business Media: London, UK, 2012.
3. Pumain, D.; Sanders, L. Theoretical principles in inter-urban simulation models: A comparison. *Environ. Plann. A* **2013**, *45*, 2243–2260.
4. Carley, K.M. (1999). On Generating Hypotheses Using Computer Simulations. DTIC Document. Available online: <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA458065> (accessed on 25 November 2015).



5. Hedström, P.; Ylikoski, P. Causal mechanisms in the social sciences. *Annu. Rev. Sociol.* **2010**, *36*, 49–67.
6. Von Bertalanffy, L. *General System Theory: Foundations, Development, Applications*; Braziller: New York, NY, USA, 1968.
7. O’Sullivan, D. Complexity science and human geography. *Trans. Inst. Br. Geogr.* **2004**, *29*, 282–295.
8. Rossiter, S.; Noble, J.; Bell, K.R.W. Social simulations: Improving interdisciplinary understanding of scientific positioning and validity. *J. Artif. Soc. Soc. Simul.* **2010**, *13*, 1–10.
9. Elsenbroich, C. Explanation in agent-based modelling: Functions, causality or mechanisms? *J. Artif. Soc. Soc. Simul.* **2012**, *15*, 1–9.
10. Grüne-Yanoff, T. The explanatory potential of artificial societies. *Synthese* **2009**, *169*, 539–555.
11. Chérel, G.; Cottineau, C.; Reuillon, R. Beyond Corroboration: Strengthening Model Validation by Looking for Unexpected Patterns. *PLoS ONE* **2015**, *10*, e0138212.
12. Kaldor, N. Capital accumulation and economic growth. In *The Theory of Capital*; Macmillan: London, UK, 1961.
13. Pumain, D. Pour une théorie évolutive des villes. *Espace Géogr.* **1997**, *26*, 119–134, in French.
14. Krugman, P. Confronting the mystery of urban hierarchy. *J. Jpn. Int. Econ.* **1996**, *10*, 399–418.
15. Cheshire, P. Trends in sizes and structures of urban areas. *Handb. Reg. Urban Econ.* **1999**, *3*, 1339–1373.
16. Pumain, D. *Hierarchy in the Natural and Social Sciences*; Methodos Series; Springer: Heidelberg, German, 2006.
17. Auerbach, F. Das gesetz der bevölkerungskonzentration. *Petermanns Geogr. Mitt.* **1913**, *59*, 73–76.
18. Lotka, A.J. *Elements of Physical Biology*; William & Wilkins: Baltimore, MD, USA, 1925.
19. Singer, H.W. The “courbe des populations.” A parallel to Pareto’s Law. *Econ. J.*, **1936**, *46*, 254–263.
20. Zipf, G.K. *Human Behavior and the Principle of Least Effort*; Addison-Wesley: Cambridge, MA, USA, 1949.
21. Berry, B.J. City size distributions and economic development. *Econ. Dev. Cult. Chang.* **1961**, *9*, 573–588.
22. Eeckhout, J. Gibrat’s law for (all) cities. *Am. Econ. Rev.* **2004**, *5*, 1429–1451.
23. Gibrat, R. *Les Inégalités Économiques*; Sirey: Paris, France, 1931, in French.
24. Clauset, A.; Shalizi, C.R.; Newman, M.E. Power-law distributions in empirical data. *SIAM Rev.* **2009**, *51*, 661–703.
25. Mitzenmacher, M. A brief history of generative models for power law and lognormal distributions. *Internet Math.* **2004**, *1*, 226–251.
26. Simon, H.A. On a class of skew distribution functions. *Biometrika* **1955**, *42*, 425–440.
27. Krugman, P. *The Self-Organizing Economy*; Blackwell: Oxford, UK, 1996.
28. Pumain, D. (2004). Scaling laws and urban systems. Santa Fe Institute, WP, (04-02). Available online: <http://samoa.santafe.edu/media/workingpapers/04-02-002.pdf> (accessed on 25 November 2015).
29. Gabaix, X. Zipf’s law for cities: An explanation. *Q. J. Econ.* **1999**, *114*, 739–767.

30. Dimou, M.; Schaffar, A. Les théories de la croissance urbaine. *Rev. D'écon. Politique* **2011**, *121*, 179–207, in French.
31. Henderson, J.V. The sizes and types of cities. *Am. Econ. Rev.* **1974**, *64*, 640–656.
32. Krugman, P. Urban concentration: The role of increasing returns and transport costs. *Int. Reg. Sci. Rev.* **1996**, *19*, 5–30.
33. Allen, P.M.; Sanglier, M. Urban evolution, self-organization, and decisionmaking. *Environ. Plan. A* **1981**, *13*, 167–83.
34. Bura, S.; Guérin-Pace, F.; Mathian, H.; Pumain, D.; Sanders, L. Multiagent systems and the dynamics of a settlement system. *Geogr. Anal.* **1996**, *28*, 161–178.
35. Weidlich, W.; Haag, G. *Concepts and Models of a Quantitative Sociology. The Dynamics of Interacting Populations*; Springer: Berlin, Germany, 1983.
36. Christaller, W. *Die zentralen Orte in Süddeutschland: Eine ökonomisch-geographische Untersuchung über die Gesetzmässigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen*. Verlag Gustav Fischer: Jena, German, 1933.
37. Ullman, E. A theory of location for cities. *Am. J. Sociol.* **1941**, *6*, 853–864.
38. Duranton, G.; Puga, D. Nursery cities: Urban diversity, process innovation, and the life cycle of products. *Am. Econ. Rev.* **2001**, *91*, 1454–1477.
39. Pumain, D.; Paulus, F.; Vacchiani-Marcuzzo, C.; Lobo, J. An evolutionary theory for interpreting urban scaling laws. *Cybergeo Eur. J. Geogr.* **2006**, *343*, 1–22.
40. Hägerstrand, T. *Innovation Diffusion as a Spatial Process*; University of Chicago Press: Chicago, IL, USA, 1968.
41. Robson, B.T. *Urban Growth: An Approach*; Routledge: London, UK, 1973.
42. Pred, A.R. *The Growth and Development of Systems of Cities in Advanced Economies*; Lund Studies in Geography: Lund, Sweden, 1973.
43. Krings, G.; Calabrese, F.; Ratti, C.; Blondel, V.D. Urban gravity: A model for inter-city telecommunication flows. *J. Stat. Mech. Theory Exp* **2009**, *7*, L07003.
44. Guérin-Pace, F. Rank-size distribution and the process of urban growth. *Urban Stud.* **1995**, *32*, 551–562.
45. Cottineau, C. L'évolution des Villes Dans L'espace Post-Soviétique. Observation et Modélisations. Ph.D. Thesis, Université Paris 1 Pantheon-Sorbonne, Paris, France, 2014.
46. Black, D.; Henderson, V. Urban evolution in the USA. *J. Econ. Geogr.* **2003**, *3*, 343–372.
47. Hernando, A.; Hernando, R.; Plastino, A.; Zambrano, E. Memory-endowed US cities and their demographic interactions. *J. R. Soc. Interface* **2015**, *12*, 20141185.
48. Bretagnolle, A. Vitesse et processus de selection hierarchique dans le systeme des villes francaises. In *Données urbaines*, 4; Pumain, D., Mattei, M.-F., Eds.; Anthropos: Paris, Franch, 2003; pp. 309–322.
49. Preston, S.H. Urban growth in developing countries: A demographic reappraisal. *Popul. Dev. Rev.* **1979**, *5*, 195–215.
50. Bockerhoff, M. Urban growth in developing countries: A review of projections and predictions. *Popul. Dev. Rev.* **1999**, *25*, 757–778.

51. Harris, C.D. *Cities of the Soviet Union: Studies in Their Functions, Size, Density, and Growth*; Published for Association of American Geographers: Chicago, IL, USA, 1970.
52. Byrne, D. *Complexity Theory and the Social Sciences*; Routledge: London, UK, 1998.
53. Goldthorpe, J.H. Causation, statistics, and sociology. *Eur. Sociol. Rev.* **2001**, *17*, 1–20.
54. 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)* **2015**, *18*. Available online: <http://jasss.soc.surrey.ac.uk/18/4/9.html> (accessed on 25 November 2015).
55. Cottineau, C.; Slepukhina, I. The Russian urban system: Evolution under transition? In *International and Transnational Perspectives on Urban Systems*; MIT Press: Cambridge, MA, USA, forthcoming.
56. Cottineau, C. (2014). DARIUS Database. Harmonised Database of Cities in the Post-Soviet Space, 1840–2010, Figshare. Available online: <http://dx.doi.org/10.6084/m9.figshare.1108081> (accessed on 25 November 2015).
57. Jakoby, O.; Grimm, V.; Frank, K. Pattern-oriented parameterization of general models for ecological application: Towards realistic evaluations of management approaches. *Ecol. Model.* **2014**, *275*, 78–88.
58. Wiegand, T.; Revilla, E.; Knauer, F. Dealing with uncertainty in spatially explicit population models. *Biodivers. Conserv.* **2004**, *13*, 53–78.
59. Batty, M. Models again: Their role in planning and prediction. *Environ. Plan. B Plan. Des.* **2015**, *42*, 191–194.
60. Batty, M.; Torrens, P.M. Modelling and prediction in a complex world. *Futures* **2005**, *37*, 745–766.
61. Openshaw, S. From Data Crunching to Model Crunching: The Dawn of a New Era. *Environ. Plan. A* **1983**, *15*, 1011–1012.
62. Thiele, J.C.; Grimm, V. Replicating and breaking models: Good for you and good for ecology. *Oikos* **2015**, doi:10.1111/oik.02170.
63. 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. Epidemiol.* **2008**, *168*, 1–8.
64. Contractor, N.; Whitbred, R.; Fonti, F.; Hyatt, A.; O’Keefe, B.; Jones, P. Structuration theory and self-organizing networks. In Proceedings of the Organization Science Winter Conference, Keystone, CO, USA, January 2000.
65. Grimm, V.; Berger, U.; Bastiansen, F.; Eliassen, S.; Ginot, V.; Giske, J.; Goss-Custard, J.; Grand, T.; Heinz, S.; Huse, G.; *et al.* A standard protocol for describing individual-based and agent-based models. *Ecol. Model.* **2006**, *198*, 115–126.
66. Chalmers, D.J. Strong and weak emergence. In *The Re-Emergence of Emergence*; Oxford University Press: Oxford, UK, 2006; pp. 244–256.
67. Schärli, N.; Ducasse, S.; Nierstrasz, O.; Black, A.P. Traits: Composable units of behaviour. In *ECOOP 2003-Object-Oriented Programming*; Springer: Berlin and Heidelberg, German, 2003; pp. 248–274.
68. Bracha, G. The Programming Language Jigsaw: Mixins, Modularity and Multiple Inheritance. Ph.D. Thesis, The University of Utah, Salt Lake City, UT, USA, 1992.

69. Wampler, D.; Payne, A. *Programming Scala: Scalability = Functional Programming + Objects*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2014.
70. Simpuzzle. Open-licensing source code of the MARIUS Model. *GitHub Repository*; 2015, Available online: <https://github.com/ISCPIF/simpuzzle> (accessed on 25 November 2015).
71. 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 1994; pp. 273–282.
72. Steyaert, P.; Codenie, W.; D'Hondt, T.; de Hondt, K.; Lucas, C.; van Limberghen, M. Nested mixin-methods in Agora. In *ECOOP'93?Object-Oriented Programming*; Springer: Berlin and Heidelberg, German, 1993; pp. 197–219.
73. Prehofer, C. Feature-oriented programming: A fresh look at objects. In *ECOOP'97 Object-Oriented Programming*; Springer: Berlin and Heidelberg, German, 1997; pp. 419–443.
74. Steyaert, P.; de Meuter, W. A marriage of class-and object-based inheritance without unwanted children. In *ECOOP'95 Object-Oriented Programming*; Springer: Berlin and Heidelberg, German, 1995; pp. 127–144.
75. Grimm, V.; Railsback, S.F. Pattern-oriented modelling: A “multi-scope” for predictive systems ecology. *Philos. Trans. R. Soc. B Biol. Sci.* **2012**, *367*, 298–310.
76. Schmitt, C.; Rey-Coyrehourcq, S.; Reuillon, R.; Pumain, D. Half a billion simulations: Evolutionary algorithms and distributed computing for calibrating the SimpopLocal geographical model. *Environ. Plan. B Plan. Des.* **2015**, *42*, 300–315.
77. VARIUS. Open-source and Interactive application to explore MARIUS Models and visualise their results. *Geographie-cites*; 2015, Available online: <http://shiny.parisgeo.cnrs.fr/VARIUS> (accessed on 25 November 2015).
78. Thiele, J.C.; Kurth, W.; Grimm, V. Facilitating Parameter Estimation and Sensitivity Analysis of Agent-Based Models: A Cookbook Using NetLogo and “R”. *J. Artif. Soc. Soc. Simul.* **2014**, *17*, 1–45.
79. Mahfoud, S.W. Niching methods for genetic algorithms. *Urbana* **1995**, *51*, 62–94.
80. Bretagnolle, A.; Pumain, D. Simulating urban networks through multiscalar space-time dynamics: Europe and the united states, 17th–20th centuries. *Urban Stud.* **2010**, *47*, 2819–2839.
81. Martin, R. Rebalancing the spatial economy: The challenge for regional theory. *Territ. Politics Gov.* **2015**, *3*, 235–272.
82. Openshaw, S. Building an automated modeling system to explore a universe of spatial interaction models. *Geogr. Anal.* **1988**, *20*, 31–46.