Asia Pacific Academy of Science Pte. Ltd.

Sustainable Social Development 2024, 2(4), 2625.
https://doi.org/10.54517/ssd.v2i4.2625

Article

Modelling-based analytics for urban grand challenges

Alan Wilson[1,2]

[1] The Alan Turing Institute, London NW1 2DB, UK; awilson@turing.ac.uk
[2] The London Interdisciplinary School, London E1 1EW, UK

CITATION

Wilson A. Modelling-based analytics for urban grand challenges. Sustainable Social Development. 2024; 2(4): 2625. https://doi.org/10.54517/ssd.v2i4.2625

ARTICLE INFO

Received: 14 May 2024
Accepted: 22 July 2024
Available online: 1 August 2024

Abstract: Society faces grand challenges on a number of dimensions, for example: climate change, pandemics, security and geopolitics, and social exclusion. The future development of towns and cities is key to meeting these. The availability of analytic capabilities provides foundations for developing and evaluating alternative policies and plans. An extensive range of models is available, but they have not been well-focused on these kinds of grand challenges. A significant research task, therefore, is to review the modelling developments needed to provide the necessary analytics base. We consider in turn: the building bricks; the challenge of interdependencies and high dimensionality, using Lowry's model as a framework; the integration of the elements into a comprehensive model as a basis for grand challenge analytics; and the challenges of implementation.

Keywords: climate change; pandemics; social inclusion; security; urban modelling

1. Introduction

Six decades of research underpin that part of our knowledge of cities based on mathematical and computer models. Objectives in building these include developing the science of cities and seeking to use models as an important tool in urban planning. At the present time, society faces major challenges, most of which were not understood or well-articulated through the history of urban modelling. We need to relate to, for example, sustainability and climate change, pandemics, geopolitics and security, and social exclusion. Applications of models to date have been largely in the narrower field of conventional city planning. The aim of this paper is to show how the models can be further developed to contribute both to our understanding of the major challenges and to associated policy articulation and planning. There is a considerable knowledge base that has been developed over a long period of time. We aim to show how further steps can be taken to make use of this in relation to grand challenges that are now recognised as urgent.

This is a very large task and to make this feasible, the paper will be focused on the evolution of the author's own endeavours, and the implications for future research. See Wilson [1] for the origins of this thinking, along with accounts of more recent developments in the work of Wilson [2]. It should be possible to transfer the argument, mutatis mutandis, to other modelling approaches and styles.

2. The building bricks

A comprehensive model of a city can be assembled from interacting submodels, aimed at handling the interdependencies that make this enterprise so challenging. The archetypal model is Lowry's [3] model of metropolis, and this continues to provide a conceptual framework for the development of more detailed, comprehensive models.

Subsequent developments have been concerned with dealing with complexity—the high-dimensional nature of the modelling challenge—and introducing explicit dynamics' mechanisms, along with a tremendous amount of empirical work by many researchers, which has led to extensions and refinements, including handling large data sets and developing associated calibration methods.

The main elements of a comprehensive model are shown in broad terms in **Figure 1**. The migration and population (sub)models on the left-hand side represent the demographics of the system; the trade, economy, and infrastructure models on the right hand side represent the economy. These provide the framework within which people live (residence, employment, and use of services) and organisations function (delivering housing, jobs, and products or services). These elements interact through the transport and communications systems. This modelling of flows is important, ranging from migration and trade to journeys to work and to retail and various services. The flows are carried on the transport and telecom infrastructures. (There are some definitional problems: a move to a new house or job within a city is usually referred to as relocation rather than migration; and handling the 'population' involves both individuals and households and possibly wide family structures. The delivery of services such as health and education are sometimes referred to as 'public services' though the provision can be through public or private organisations (or a combination of both). We neglect these refinements in the first instance).

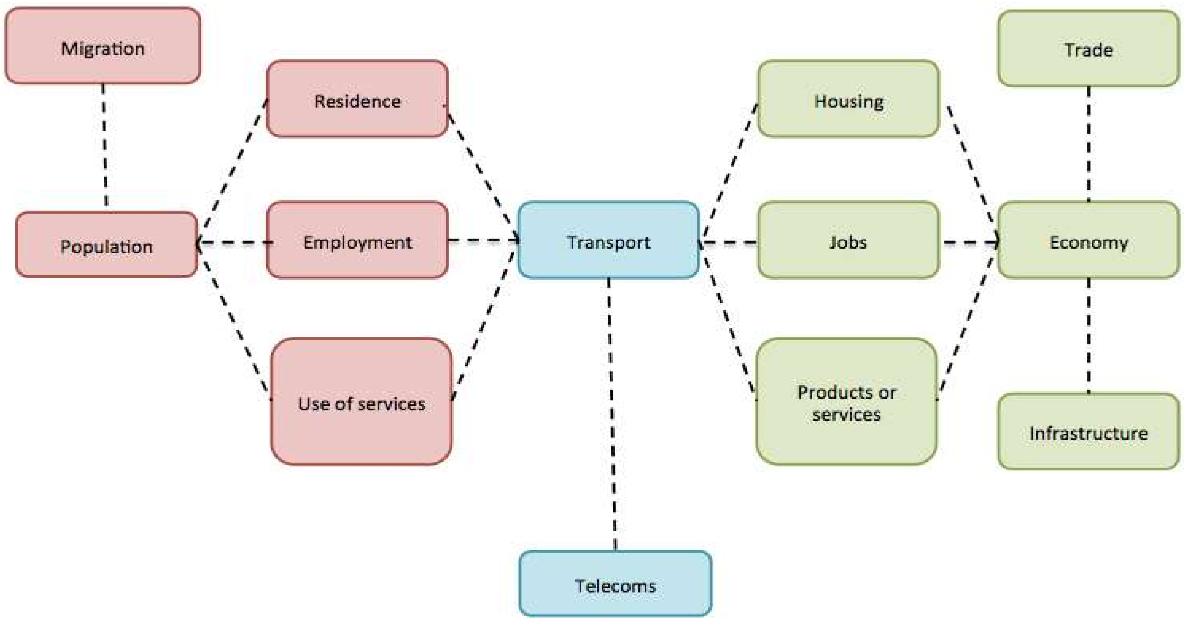

Figure 1. The main elements of a comprehensive model.

The core submodels can be therefore summarised as:

- Demographic (including migration)
- Economic (including trade and infrastructure)
- Residential location
- Housing
- Employment location
- Provision and use of services (retail, health, education, leisure, …)

- Transport
- Telecoms

Some of these have generic model elements:

- Models of flows (spatial interaction)
- Network analysis (the infrastructure, with flows loaded on to links)
- Accounts (flow matrices, demographic and input-output accounts)
- Dynamics

As noted, the core models have been combined into general or comprehensive urban models, with Lowry's [3] simple and brilliant model capturing the main interacting elements as the precursor. Many of the models that have built on the Lowry heritage have neglected full connection to demographic and economic input-output submodels, though there have been notable exceptions to this. They have not been widely used, possibly because of the scale of effort involved in assembling them. See, for example, Echenique et al. [4], Simmonds [5], and Kim et al. [6], in the last case building on Wilson [1].

These models usually function on a steady-state or equilibrium basis, but a dynamic model for a retail system was presented in Harris and Wilson [7]—a method that can be extended to other submodels—in transport networks, for example—see de Martinis et al. [8]. This method has been tested by a number of authors, but it is not in general use. See, for example, Dearden and Wilson [9] on the dynamics of shopping centre development and gentrification in housing. It has been further developed in a new framework by Ellam et al. [10] and Gaskin et al. [11]. This has led to new calibration methods—reflecting the impact of new techniques from data science and AI. The continued growth in computing power has led to further advances, as with the Quant model (see Batty and Milton [12]). Many of the submodels are valuable in their own right for particular applications, and in specific fields such as retail and transport, they have been extensively tested and deployed (see Birkin et al. [13], Boyce and Williams [14]).

More recently, we have seen the beginnings of attempts to widen the scope of the models to contribute to the grand challenges, and we present these in Section 3. We pursue the extent to which these initiatives can be driven further in the rest of the paper. The subsystem models need to be expanded to incorporate, for example, carbon production (see Huber et al. [15]); the spread and control of infection in pandemics (see Spooner et al. [16]); more systematic applications in geopolitics and security (see Guo et al. [17]); and social inclusion, including access to health and education facilities. Modellers have not focused on social inclusion, though of course many other researchers have, notably Dorling [18,19]. The explicit incorporation of dynamics would enable the exploration of phase changes, both to identify future challenges through abrupt change and to target policy levers to support productive change. This enables the use of the models for systematic scenario exploration.

We proceed as follows: in Section 3, we explore how the model system can be developed, handling interdependence and high dimensionality as a preliminary to defining a model system that can provide the analytics basis for exploring grand challenges in Section 4. We conclude in Section 5 with a discussion of the implementation challenges.

3. Interdependencies and high dimensionality: Insights from the Lowry model

3.1. The modelling challenge

In order to develop modelling systems to contribute to the grand challenges, we need to add more depth and detail, and in some cases, new submodels. We particularly need to represent interdependence, as we will demonstrate, and hence expand comprehensive models. We can identify one interdependence chain below by way of illustration to test the potential for developing extended models.

Consider the following real challenge, albeit oversimplified but appropriate for illustration: A housing problem is usually presented as a shortage of houses. However, it can be argued that the underlying problem is an income problem—too many households with insufficient income to find or maintain accommodation. There is an income problem because there is an employment problem. There is an employment problem because there is a skills problem. There is a skills problem because there is an education problem. There is an interdependency chain: education—skills—employment—income—housing. This leads to a further complication: as well as seeking to ensure that the education system does not leave a long tail of under-skilled people, we need to recognise that to change this, it takes time. It is more complicated to model this than to predict the revenue attracted to a new supermarket. However, that is part of our agenda.

The problem of high dimensionality was considered in Wilson [20]—which led to the argument that in a relatively coarse-grained model, 10^{13} variables would be needed for a comprehensive description, something that is obviously not feasible in practice. As we expand our scope, therefore, we will need to make the necessary approximations in as optimal a way as possible. Part of this relates to data availability: we almost always have problems with missing data, and we can use such methods as biproportional fitting to fill in the gaps—see Dennett and Wilson [21] for an example in migration modelling and Caschilli and Wilson [22] on trade.

In the following subsections, we discuss first scales (3.2), and then we set up the Lowry model as a foundational example (3.3). In 3.4, we present the model in formal terms as the basis for developing extensions. This allows us to seek to build extended, comprehensive model-based analytics in Section 4.

3.2. Scales

There are three dimensions of scale to be considered: spatial, sectoral, and time. The spatial ranges from the global via the national, the city-regional, and the neighbourhood to the micro. Model-based analytics have been developed for all of these. We will largely take the city-region scale as our focus but show links to others—up to the global and down to the micro. Indeed, it is possible to model at two scales.

For sectors, there are classifications available at different levels of granularity for industry and services and for labour and government, for example. These range from the very coarse—for example, working with a 'basic' industry sector and a small number of 'retail' sectors—through to the very fine. However, the desired granularity demands easily useable data which is often only partially available. For example, fine-scale sector data may not be available for fine-scale spatial units. At the finest scale, we can, in principle, seek data on individual organisations.

For time, we can think in broad terms of short, middle, and long runs and seek to choose what is appropriate for the systems we are modelling and their applications. We also have to decide whether to treat time as discrete or continuous—usually discrete, but we sometimes want to incorporate data on individual events in health analytics, for example, where precision is important, for example, in dealing with co-morbidities.

3.3. The Lowry model

To fix ideas, we first present the Lowry model and then convert this into something more formal. This gives us a platform for demonstrating interdependence and for extending the models in the direction of handling grand challenges. First, a recap on Lowry's variables:

A = area of land
E = employment
P = population
c = trip cost
Z = constraints

to which should be added the following to be used as subscripts or superscripts:

U = unusable land
B = basic sector
R = retail sector
H = household sector
k = class of establishment within a sector
m = number of classes of retail establishment
i, j = zones
n = number of zones

A_i^H, for example, is the area of land in zone i that is used for housing. If a subscript or a superscript is omitted, this implies summation so A_i is the total amount of land in i. There are two kinds of economic sectors: basic and retail—the latter further subdivided. Basic employment—and its spatial distribution across zones—is given exogenously. Retail employment is generated by the population. Once this simple principle of building the variables—the region's descriptors—is understood, the model can be presented in twelve equations.

The key land use equation is:

$$A_j^H = A_j - A_j^U - A_j^B - A_j^R \quad (1)$$

This captures some key hypotheses: that land for basic and retail industries can always outbid housing, so this shows land available for housing is a residual. The household sector is represented by:

$$P = f\Sigma_j E_j \quad (2)$$

$$P_j = g\Sigma_i E_i f_{\text{res}}(c_{ij}) \quad (3)$$

$$\Sigma_j P_j = P \quad (4)$$

$$P_j \leq z^H A_j^H \quad (5)$$

This sequence generates the population through employment and begins the process of housing them. Equation (2) calculates the total population as proportional to total employment. Equation (3) allocates this population to zones, *i*. $f_{\text{res}}(c_{ij})$ is a declining function of travel cost from *i* to *j*, thus building in the likelihood that workers live nearer to their workplace. Equation (4) enables *g* in Equation (3) to be calculated as a normalising factor. The fourth equation is particularly interesting and also shows how the model is more complicated than it appears at first sight. z^H is the unit amount of land used for residences, and so this equation is constraining the numbers assigned to zone *i* in relation to land availability. This is one of the subtleties—and part of the trickiness—of the model: the equations have to be solved iteratively to ensure that this constraint is satisfied.

The retail sector is represented by

$$E^{Rk} = a^k P \quad (6)$$

$$E_j^{Rk} = b^k[c^k \Sigma_i P_i f^k(c_{ij}) + d^k E_j] \quad (7)$$

$$\Sigma_j E_j^{Rk} = E^{Rk} \quad (8)$$

$$E_j^{Rk} > z^{Rk} \quad (9)$$

$$A_j^R = \Sigma_k e^k E_j^{Rk} \quad (10)$$

$$A_j^R \leq A_j - A_j^U - A_j^B \quad (11)$$

These six equations determine the amount of employment generated in the retail sector. The total in sector *k* within retail is given by Equation (6), and this is spatially distributed through Equation (7). As with the residential location equation, the function $f^k(c_{ij})$ is a decreasing function of travel cost, indicating that retail facilities will be demanded relatively nearer to residences. c^k converts these units into employment. The term $d^k E_j$ represents the use of retail facilities from the workplace. b^k is a normalising factor which can be determined from Equation (8). Equation (9) imposes a minimum size for retail sector k at a location. (No school for half a dozen pupils for example!) Equations (10) and (11) sort out retail land use, the first calculating a total from a sum of *k*-sector uses—e^k converting employment into land—and the second specifying the maximum amount of retail land—in effect giving 'basic' (which has been given exogenously) priority over retail. In this case, unlike the residential case where P_i was constrained by land availability, retail employment is not so constrained. Lowry argued that, if necessary, retail could 'build upwards'. If A_j^R from Equation (10) exceeds $A_j - A_j^U - A_j^B$, it is reset to this maximum, but employment does not change.

Total employment is then given by:

$$E_j = E_j^B + \Sigma_k E_j^{Rk} \quad (12)$$

This final equation simply adds up the total employment in each zone. The equations are solved iteratively, starting with $E_j^{RK} = 0$.

To develop further model-building insight as a preliminary, we can regroup these equations into four categories: accounting, aggregate relations, spatial interaction (flow) relations, and constraints.

Accounting

$$A_j^H = A_j - A_j^U - A_j^B - A_j^R \quad (13)$$

$$\Sigma_j P_j = P \quad (14)$$

$$E_j = E_j^B + \Sigma_k E_j^{Rk} \quad (15)$$

$$\Sigma_j E_j^{Rk} = E^{Rk} \tag{16}$$

Aggregate relations

$$P = f\Sigma_j E_j = fE,\ E = \Sigma_j E_j \tag{17}$$

$$E^{Rk} = a^k P \tag{18}$$

$$A_j^R = \Sigma_k e^k E_j^{Rk} \tag{19}$$

Spatial interaction (flow) relations

$$T_{ij} = g\Sigma_i E_i f_{\text{res}}(c_{ij}) \tag{20}$$

$$S_{ij}^k = ck[\Sigma iPifk(c_{ij}) + d^k E_j] \tag{21}$$

extracted from (7), written as

$$E_j^{Rk} = b^k[\Sigma_i S_{ij}^k + d^k E_j] \tag{22}$$

Constraints

$$P_j \le z^H A_j^H \tag{23}$$

$$E_j^{Rk} > z^{Rk} \tag{24}$$

$$A_j^R \le A_j - A_j^U - A_j^B \tag{25}$$

The accounting relations are statements for each main model element: land, population, and employment; these elements are linked first at an aggregate level and then through spatial interaction relationships. The constraints are Lowry's way of handling some of the problems of interdependence.

3.4. The Lowry model in formal terms

The principal endogenous variables are A_i^H, A_i^R, P_i and E_j^{Rk} and we can write them in formal form as follows.

$$A_i^H = A_i^H(P_i, E_i^B, E_i^{Rk}, A_i^U, A_i^R, A_i^B, c_{ij}) \tag{26}$$

$$P_i = P_i(H_i, \{E_j\}, \{c_{ij}\}, A_i^H, A_i^U, A_i^R, A_i^B) \tag{27}$$

$$E_j^{Rk} = E_j^{Rk}(W_j^{Rk}, \{P_i\}, \{c_{ij}\}, E_j^B) \tag{28}$$

together with the constraint equations

$$P_j \le z^H A_j^H \tag{29}$$

$$E_j^{Rk} > z^{Rk} \tag{30}$$

$$A_j^R \le A_j - A_j^U - A_j^B \tag{31}$$

now renumbered. In Equations (27) and (28), we have added structural variables H_i and W_j^{Rk} as measures of housing supply and 'retail' infrastructure respectively. Such terms do not appear in Lowry's model: H_i is (implicitly) taken as proportional to P_i, and W_j^{Rk} similarly to E_j^{Rk}. However, we need to recognise that these structural variables change more slowly than the activity variables, and we will model this explicitly below. Meanwhile, we can add formal equations for them:

$$H_i = H_i(P_i, \{E_j^*\}, \{c_{ij}\}, \ldots) \tag{32}$$

$$W_j^{Rk} = W_j^{Rk}(\{P_i\}, \{c_{ij}\}, \ldots.) \tag{33}$$

It is also helpful to write the interaction Equations (20) and (21), in formal terms:

$$T_{ij} = T_{ij}(P_i, H_i, \{E_j^*\}, \{c_{ij}\}, \ldots) \tag{34}$$

$$S_{ij}^{Rk} = S_{ij}^{Rk}(P_i, \{W_j^{Rk}\}, \{c_{ij}\}, \ldots) \tag{35}$$

The curly brackets in Equations (32)–(35) indicate the vectors or matrices within them. The highly interdependent structure of these equations leaps out here: the endogenous variables in one equation appear as exogenous variables in the others. This is why they always have to be solved iteratively. Of course, this simply represents real-world interdependence.

We now have the framework to show how the models can be extended to contribute to meeting the grand challenges. All the submodels listed in Section 2 are implicit here. The crucial spatial interaction models have been made explicit in Equations (20) and (21) and formally in Equations (34) and (35). These have been developed in highly sophisticated and tested forms. Lowry's retail sectors are based on a very broad definition of retail to mean 'any form of interaction of people using a facility'—and so can be expanded to include education and health. It is immediately clear that the population and sector variables need much more detailed subscripts and superscripts, as indicated in the previous subsection. This can be done but creates, again, as noted, very large multi-dimensional arrays, and a full integration of these disaggregated models with microsimulation methods is called for. See Birkin and Clarke [23] and Smith et al. [24].

4. Towards an extended comprehensive modelling system: Analytics for grand challenges

4.1. Model development

We can now combine the analytics needed for the grand challenges sketched in Section 1 with the building bricks discussed in Section 2 and the insights offered by the Lowry model in Section 3.3, transformed into a formal structure in Section 3.4. The analytics base we are seeking is described as a 'modelling system' because its elements are often valuable in their own terms, and there will be variants of the comprehensive model as different kinds of approximations are made to deal with the challenges of high dimensionality.

Equations (26)–(35) provide a good starting point, first in terms of omissions and weaknesses in our usual comprehensive models: Equation (26) reminds us that we have not always dealt with 'land' effectively; Equations (29)–(31) remind us of the importance of handing constraints. However, in a broad conceptual way, Equations (27) and (28) along with Equations (32) and (33), represent the core of the model building challenge.

This takes us back to **Figure 1**: the left-hand side of the diagram represents the population and associated activities; the right-hand side represents the economy. To recap on Section 2, we have detailed demographic models that are usually applied at an aggregate scale, and similarly, good economic models—based on input-output accounts—are also applied at an aggregate scale. These provide envelopes for the more granular multi-zone models. For both the demographic and the economic models, we need fine-grained specifications of people (and households) and sectors, and there are obvious links to the activity models: life expectancy in relation to education, for example, and employment through the skills base, represented in both supply and demand terms in the input-output model.

This implies the need for disaggregation, which we now pursue. In Section 3, we have no disaggregation of person type, and the economic sectors are divided into 'basic' and a set of 'retail'. A starting point, therefore, is to define a superscript, n, to represent person type, and m to represent economic sectors. The latter replaces 'basic' and 'Rk'. This releases k to be used as a superscript for housing type. We can then rewrite the formal model equations in these terms (and this defines the disaggregated variables in an obvious way).

$$A_i^H = A_i^H(\{P_i^n\}, \{E_i^m\}, A_i^U, A_i^R, A_i^B, \{c_{ij}^{mn}\}) \quad (36)$$

$$P_i^n = P_i^n(H_i^k, \{E_j^m\}, \{c_{ij}^{mn}\}, A_i^H, A_i^U, A_i^R, A_i^B) \quad (37)$$

$$E_j^m = E_j^m(W_j^m, \{P_i\}, \{c_{ij}^{mn}\}, \ldots), m\varepsilon\{\text{retail}\} \quad (38)$$

$$P_j^n \le z^H A_j^H \quad (39)$$

$$E_j^m > z^m, m\varepsilon\{\text{retail}\} \quad (40)$$

$$A_j^R \le A_j - A_j^U - A_j^B \quad (41)$$

$$H_i^k = H_i^k(\{P_i^n\}, \{E_j^*\}, \{c_{ij}\}, \ldots) \quad (42)$$

$$W_j^m = W_j^m(\{P_i^n\}, \{c_{ij}\}, \ldots), m\varepsilon\{\text{retail}\} \quad (43)$$

$$T_{ij}^{mn} = T_{ij}^{mn}(P_i^n, H_i^k, \{E_j^*\}, \{c_{ij}^{mn}\}, \ldots), m\varepsilon\{\text{retail}\} \quad (44)$$

$$S_{ij}^{mn} = S_{ij}^{mn}(P_i^n, \{W_j^m\}, \{c_{ij}^{mn}\}, \ldots), m\varepsilon\{\text{retail}\} \quad (45)$$

We assume an upper-tier envelope that will handle, for example, migration (Rees and Wilson [25}) and the economy through an input-output model (see Zhang et al. [26]). We also need to recognise that further disaggregation may well be required. In relation to person types, for example, the superscript n may be a 'vector' of superscripts representing age, sex, education, skills, employment, income, and so on. For example, define P_{ij}^{dakmse} where i = residential location, j = workplace location (or a service location—probably not both needed at the same time), d = demographics, a = income, k = house type, m = employment sector, s = skill level, and e = education status.

Equation (37), along with Equation (44), when assembled into its various submodels, represents the population and associated activities, usually anchored in residence and work and with a strong interaction between these—the journey to work. It is also important to include those who don't work away from home: those in regular employment but, post-pandemic, working from home; the retired; the unemployed; and those committed to domestic work. These latter categories will be an important element of the social inclusion challenge and are usually neglected in current models. It is then possible to enumerate the set of activities, usually modelled as spatial flows from home or work origin to a facility—such as retail, education, and health. We also seek to model the impact of these activities, such as the development of skills through education and in the treatment of disease, for example. As we noted earlier, many of these impacts are at future times, which implies in modelling a cross-section at each point in time. We need to take past histories into account, something that is not usually done, though the modelling of co-morbidities in health may be a pointer for future work on a broader basis. See Pagliara et al. [27] for a review of residential location models. Equation (38), along with Equation (43), works on a similar basis, but representing organisations in the economy, usually aggregated into sectors—though there is a challenge here in linking micro and meso/macro theory—see Pagliara et al. [28].

The interaction Equations (44) and (45) provide the basis for transport analysis and planning, which is a well-developed field of study in engineering. These models provide the key to providing generalised cost terms that figure throughout, handling congestion in networks, for example. See Boyce and Williams [14] for a brilliant review of the history of these models and an account of the current state of play.

The model system represented by Equations (36)–(45) is, essentially, handling time either through steady state or comparative static equilibrium assumptions. It is important to move towards explicit dynamics, particularly so that we can identify four-phase changes. The dynamic retail model of Harris and Wilson [7] provides a starting point. In formal terms, for a discrete time period, take the W-retail structure from Equation (43).

$$\Delta W_j^m = \Delta W_j^m(D_j^m - KW_j^m)W_j^m \tag{46}$$

where,

$$D_j^m = \sum_{in} S_{ij}^{mn} \tag{47}$$

Identification of phase changes is a particularly important task for the analytics of grand challenges—both to identify potential catastrophes and possible routes to 'good' outcomes. For a detailed account of progress in incorporating these dynamic concepts into a comprehensive model, see Dearden and Wilson [9].

There are submodels available for each element of the Equations (36)–(47) system—embracing demography, economics, residential and housing, basic industry, 'retail'—people-driven services, employment, transport and telecoms. This formal system could be expanded to add the detail that these offers. It remains essential that these are combined into a general model, particularly in relation to grand challenges so as not to lose track of the interdependencies.

4.2. Analytics for grand challenges

We start with the initial assumption that the submodels listed in Section 4.1 are in place, along with comprehensive models that integrate these for tackling interdependence. This assumption will not typically be satisfied, of course: there are no organisations that are big enough to carry the full range. What is common to each case, as we can anticipate from the focus of the paper, is that responses will demand radical change in the organisation of our towns and cities, and the analytics base through modelling provides a means of assessing what is needed and what is possible. In each case, we need to calculate new kinds of indicators for sustainability analytics which can be used to guide us towards potentially effective plans. It is worth emphasising that analysis alone—modelling—will not solve problems and meet these grand challenges: inventiveness in creating plans and policies is essential.

The analysis above, taking insights from Lowry's original comprehensive model, warns us that we will need to develop our current models to fill some gaps—notably to handle land more effectively (and associated questions of pricing and rent) and to understand the dynamics more deeply, particularly being able to identify phase changes that might be the basis of sustainability policies.

We proceed on the basis that an enhanced modelling system is in place, take each of the four grand challenge areas in turn, and note the state of play and the opportunities for further development of grand challenge analytics. We introduce the first three and then consider the social inclusion and associated housing challenges in more detail.

Climate change:

A starting point is to seek an indicator based on carbon production. This needs estimates that relate to activities in buildings and transport flows. When we have estimates of carbon production at a fine scale, we can add these to our models to calculate carbon production for different kinds of towns and cities and then to pursue the questions: what is a reasonable and achievable target for carbon reduction, and how can this be done? We can conjecture that production will be a function of urban form—residential densities, housing, the organisation of workplaces, the ways in which services are accessed and used, and links strongly to transport and telecoms. At the present time, in relation to UK cities, for example, densities of new housing remain low, and its relation to workplaces and the availability of public transport means that car trips, on average, are lengthening. In other words, the intensity of carbon production is increasing and taking us further away from reduction targets. We can use our model systems, with enhanced indicators, to explore alternative forms and organisation and it will become very clear that radical policies are needed. This kind of work is well under way—for example, through Micael Wegener's team in Germany—see for example, Huber et al. [15].

Demographic modelling and analysis will be important at both intra- and international scales, particularly in relation to migration flows. Modelling the economy through input-output models can be used to measure the effects of 'greening' policies and what this means for employment.

Pandemics:

COVID has shown that we need to be better prepared to handle future pandemics. In this case, urban models have been enhanced through integration with epidemiological modelling which facilitates the understanding of spread and the policies needed for control. This approach has been much extended by Spooner et al. [16]—this is essentially an 'add-on'.

Security, geopolitics, and war:

Models are being developed that can generate probabilities of conflict, intra-nationally and particularly internationally [17,29,30]. These applications are based on a new use of the spatial interaction concept: as 'threat'. The issues link to demographics, particularly migration, driven by climate-driven change—energy, food, and water shortages, for example. Effective and efficient international trade is important [31], again in relation to energy and food security, and this has generated models of other kinds of conflict, notably piracy [32]. This area provides a new challenge for modelers and new opportunities for analysis and planning in police and defense agencies.

Social inclusion:

This is an important area with huge potential, especially, as noted earlier, modelers lack of application to social exclusion. Our core models, as described in earlier sections, are well developed and potentially effective in their ability to represent the well-being of people (households) and organisations (the economy, private and public; mixed). Accessibility indicators are particularly important in relation to employment (and hence incomes), education, and health. We can work through this example in more detail using our illustrative interdependency chain and the associated formal model: education—skills—employment—income—housing. Although the stimulus for this is related to housing challenges, there are obviously wider implications for social inclusion.

If we assume that the core models are in place, we can extract from these to follow the chain through: education will be one of the sectors, m; the population vector will be disaggregated to reflect skills levels and incomes (associated with age and gender); employment in all sectors will be disaggregated by skills needed and wage. The handling of time will be critical. For illustration, let us assume an annual cycle denoted by $t = 1, 2, 3, \ldots$ and that the model system we are building will run through this time sequence. The shift from t to $t + \Delta t$ will be achieved through dynamic models such as Equations (46) and (47), with equivalent equations for housing. In terms of these structural variables, they will, at least in part, be specified exogenously through actions and plans. This might include investment in education provision to increase skills, and we would model the impacts of such investment through the skills element of the demographic model. Similarly, the developing employment distribution would be partly generated by the model (for the 'retail' sectors) and would be partly exogenous. This system would generate changes in income distribution over time, and these would become inputs into the housing model.

The key equations in the model for this case, in sequence, extracted from Equations (36)–(47) above and re-ordered as appropriate, follow. A brief commentary is added to show how this facilitates a walk along the chain.

$$E_j^m = E_j^m(W_j^m, \{P_i\}, \{c_{ij}^{mn}\}, \ldots), m\varepsilon\{\text{retail}\} \quad (48)$$

$$W_j^m = W_j^m(\{P_i^n\}, \{c_{ij}\}, \ldots), m\varepsilon\{\text{retail}\} \quad (49)$$

$$\Delta W_j^m = \Delta W_j^m(D_j^m - KW_j^m)W_j^m \quad (50)$$

to record investment in m = education through W_j^m and the take-up through D_j^m.

$$P_i^n = P_i^n(H_i^k, \{E_j^m\}, \{c_{ij}^{mn}\}, A_i^H, A_i^U, A_i^R, A_i^B) \quad (51)$$

$$S_{ij}^{mn} = S_{ij}^{mn}(P_i^n, \{W_j^m\}, \{c_{ij}^{mn}\}, \ldots), m\varepsilon\{\text{retail}\} \quad (52)$$

which will pick up access to education and training, the take-up, and the impact of increasing skill levels.

$$E_j^m = E_j^m(W_j^m, \{P_i\}, \{c_{ij}^{mn}\}, \ldots), m\varepsilon\{\text{retail}\} \quad (53)$$

$$D_j^m = \sum_{in} S_{ij}^{mn} \quad (54)$$

which will give us employment for all m, and then we can look at housing provision, which will feed back into Equation (55).

$$H_i^k = H_i^k(\{P_i^n\}, \{E_j^*\}, \{c_{ij}\}, \ldots) \quad (55)$$

Equation (56) gives the journey to work.

$$T_{ij}^{mn} = T_{ij}^{mn}(P_i^n, H_i^k, \{E_j^*\}, \{c_{ij}^{mn}\}, \ldots), m\varepsilon\{\text{retail}\} \quad (56)$$

and the remaining equations check out the constraints:

$$A_i^H = A_i^H(\{P_i^n\}, \{E_i^m\}, A_i^U, A_i^R, A_i^B, \{c_{ij}^{mn}\}) \quad (57)$$

$$P_j^n \leq z^H A_j^H \quad (58)$$

$$E_j^m > z^m, m\varepsilon\{\text{retail}\} \quad (59)$$

$$A_j^R \leq A_j - A_j^U - A_j^B \quad (60)$$

The hard work then begins in two respects: to develop the model system as indicated; and to integrate it into planning and policy development to respond to the social inclusion challenge, with appropriate creativity and invention to chart ways forward.

5. Concluding comments: The challenges of implementation

We can take for granted computing power, high levels of visualisation skills, and abundant data. We have excellent account-based models of demographics and economics; we can model the functioning of most subsystems, though these skills are not universally applied; we have the beginnings of an understanding of dynamics with its implications of path dependence and phase changes. We could deal with the high-dimensionality challenge through the full integration of microsimulation.

To confront the grand challenges that we have used as examples, we have the modelling tool kit, which could be used at least for short-run analysis, but there is very little related modelling in practice. The modelers' agenda has been too narrow. So, what can be done? A central institute may be needed. This could lead to the development of best practices and the articulation of the research front line. However, government agencies would have to play their part with substantial in-house integrated units, connecting planning and policy development to analytics through inventive design. Britton Harris once remarked, "Policy, design, and analysis each involve different kinds of thinking. It is rare to find all three in the same room at the same time!" This highlights the organisational challenge. Universities would have to play their part with seriously large interdisciplinary centres, each substantially bigger than typical core departments—so a major restructuring is called for. There would be major battles as disciplines went into defensive mode.

Local government should have cross-council units that would construct an intelligent information system for a city or region, incorporating basic data through a GIS, policies and plans, and a model-based analysis and evaluation system. A best-practices system is needed to chart the way. The UK Department of Levelling Up, Communities, and Local Government should lead a cross-department 'joining-up' initiative. Universities should take interdisciplinary challenges more seriously, and some, at least, should use 'cities and regions' as a case study of what can be achieved. They would need to invest themselves, but they would also need partners among research councils, consultants, and government agencies. A new kind of cooperative structure is needed. There are significant implications for education: new courses will be needed for both young and mature students to generate an expanding skilled workforce in modelling and planning. The modelling community should break out of silos, extend horizons, and work in major public services (and the private sector), possibly re-invigorating operational research enroute. This is non-trivial and, with some notable exceptions, is not what we have now. The policy and planning communities should embrace intelligent analysis.

This all turns on a commitment to ambition in both the sciences and in policy and planning. The science needs to become 'big science'—exciting in its own terms but able to contribute to solving some of society's biggest problems. This is not essentially a funding problem; universities and government agencies could reorganise and commit to existing resources. It is a cultural problem with many dimensions. For example, academic researchers, in inefficient silos, have become accustomed to working on 'small' problems, often on 'toy' systems, because they are manageable; policymakers and planners typically have short-term perspectives and do not have the backgrounds to see that the science can help. They are also unambitious in that part of their own territory they should be good at: inventing and developing radical plans for problems that need radical solutions. These cultures are deep-seated, but perhaps dreaming the utopian dream could be the beginning of something new.

The paper aims to articulate for the urban modelling community how existing systems can be expanded to contribute to the analysis of grand challenges and the planning of responses. Although the argument is limited to one style of modelling, the conclusions can easily be translated to other approaches. This provides the basis for an important and urgent ongoing research programme.

Conflict of interest: The author declares no conflict of interest.

References

1. Wilson AG. Entropy in urban and regional modelling. London: Pion; 1970.
2. Wilson A. Boltzmann, Lotka and Volterra and spatial structural evolution: an integrated methodology for some dynamical systems. Journal of The Royal Society Interface. 2008; 5(25): 865-871. doi: 10.1098/rsif.2007.1288
3. Lowry IS. ‘A model of metropolis’, Memorandum RM-4035-RC, (Santa Monica: Rand Corporation). Available online: https://www.rand.org/content/dam/rand/pubs/research_memoranda/2006/RM4035.pdf (accessed on 11 March 2024).
4. Echenique MH, Flowerdew ADJ, Hunt JD, et al. The MEPLAN models of Bilbao, Leeds and Dortmund. Transport Reviews. 1990; 10(4): 309-322. doi: 10.1080/01441649008716764
5. Simmonds DC. The design of the DELTA land-use modelling package. Environment and Planning, B. 2011; 26: 665-684.
6. Kim TJ, Boyce DE, Hewings GJD. Combined Input-Output and Commodity Flow Models for Interregional Development Planning: Insights from a Korean Application. Geographical Analysis. 1983; 15(4): 330-342. doi: 10.1111/j.1538-4632.1983.tb00791.
7. Harris B, Wilson AG. Equilibrium Values and Dynamics of Attractiveness Terms in Production-Constrained Spatial-Interaction Models. Environment and Planning A: Economy and Space. 1978; 10(4): 371-388.
8. de Martinis V, Pagliara F, Wilson A. The Evolution and Planning of Hierarchical Transport Networks. Environment and Planning B: Planning and Design. 2014; 41(2): 192-210. doi: 10.1068/b39102
9. Dearden J, Wilson A. Explorations in Urban and Regional Dynamics. Abingdon: Routledge; 2015.
10. Ellam L, Girolami M, Pavliotis GA, et al. Stochastic modelling of urban structure. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2018; 474(2213): 20170700. doi: 10.1098/rspa.2017.0700
11. Gaskin T, Pavliotis G, Girolami M. Neural parameter calibration for large-scale multi-agent models. arXiv. 2002; arXiv:2209.13565v1.
12. Batty M, Milton R. A new framework for very large-scale urban modelling. Urban Studies. 2021; 58(15): 3071-3094. doi: 10.1177/0042098020982252
13. Birkin M, Clarke G and Clarke M. Retail geography and intelligent network planning, Chichester, Wiley, 2002
14. Boyce DE, Williams HCWL. Forecasting urban travel, Cheltenham and Northampton. Ma.: Edward Elgar; 2015.
15. Huber F, Spiekermann K, Wegener M. Cities and climate change: a simulation model of the Ruhr area 2050. Available online: http://www.corp.at (accessed on 11 March 2024).
16. Spooner F, Abrams JF, Morrissey K, et al. A dynamic microsimulation model for epidemics. Social Science & Medicine. 2021; 291: 114461. doi: 10.1016/j.socscimed.2021.114461
17. Guo W, Gleditsch K, Wilson A. Retool AI to forecast and limit wars. Nature. 2018; 562(7727): 331-333. doi: 10.1038/d41586-018-07026-
18. Dorling D. Peak inequality: Britain’s ticking time bomb. Policy Press, Bristol; 2018.
19. Dorling D. Inequality and the 1%. In: Verso Books. London and New York; 2019.
20. Wilson AG. A generalised representation for a comprehensive urban and regional model. Computers, Environment and Urban Systems. 2007; 31(2): 148-161. doi: 10.1016/j.compenvurbsys.2006.05.001
21. Dennett A, Wilson A. A Multilevel Spatial Interaction Modelling Framework for Estimating Interregional Migration in Europe. Environment and Planning A: Economy and Space. 2013; 45(6): 1491-1507. doi: 10.1068/a45398
22. Caschili S, Wilson AG. Test of bi-proportional fitting procedure applied to international trade. In: Wilson A (editor). Geo-mathematical modelling. Wiley, Chichester; 2016. pp. 26-32.
23. Birkin M, Clarke M. Synthesis—A Synthetic Spatial Information System for Urban and Regional Analysis: Methods and Examples. Environment and Planning A: Economy and Space. 1988; 20(12): 1645-1671. doi: 10.1068/a201645

24. Smith A, Lovelace R, Birkin M. Population synthesis with Quasi random integer sampling. Journal of Artificial Societies and Social simulation. 2017; 45: 1.
25. Rees PH, Wilson AG. Spatial population analysis. London: Edward Arnold. New York: Academic Press; 1977.
26. Zhang B, Rees G, Solomon G, Wilson A. Input-output analytics for urban systems: explorations in policy and planning. In: Proceedings of the 31st Annual Geographical Information Science Research UK Conference (GISRUK2023); 19-21 April 2023; Glasgow, Scotland.
27. Pagliara F, Preston J, Simmonds D, et al. Residential location choice: models and applications. Springer, Heidelburg; 2010.
28. Pagliara F, de Bok M, Simmonds D, et al. Employment location in cities and regions: models and applications. Heidelberg: Springer; 2012.
29. Baudains P, Fry HM, Davies TP, et al. A dynamic spatial model of conflict escalation. European Journal of Applied Mathematics. 2015; 27(3): 530-553. doi: 10.1017/s0956792515000558
30. Davies T, Fry H, Wilson AG, Bishop SR. A mathematical model of the London riots and their policing. Nature Scientific Reports. 2013; 3. doi: 10.1038/srep01303
31. Medda FR, Caravelli F, Caschili S, et al. Collaborative approach to trade: enhancing connectivity in sea- and land-locked countries. Springer, Heidelberg; 2017.
32. Marchione E, Johnson SD, Wilson A. Agent-based models of piracy. In: Wilson A (editor). Geo-mathematical modelling. Wiley, Chichester; 2006. pp. 217-236.