Submitted to Progress in Human Geography

January 2020

Modelling urban change with cellular automata: Contemporary issues and future research directions

Yan Liu
University of Queensland, Australia

Michael Batty
University College London, UK

Siqin Wang
University of Queensland, Australia

Jonathan Corcoran
University of Queensland, Australia

January 2020

Abstract

The study of land use change in urban and regional systems has been dramatically transformed in the last four decades by the emergence and application of cellular automata (CA) models. CA models simulate urban land use changes which evolve from the bottom-up. Despite notable achievements in this field, there remain significant gaps between urban processes simulated in CA models and the actual dynamics of evolving urban systems. This article identifies contemporary issues faced in developing urban CA models and draws on this evidence to map out four interrelated thematic areas that require concerted attention by the wider CA urban modelling community. These are: (1) to build models that comprehensively capture the multi-dimensional processes of urban change, including urban regeneration, densification and gentrification, in-fill development, as well as urban shrinkage and vertical urban growth; (2) to establish models that incorporate individual human decision behaviours into the CA analytic framework; (3) to draw on emergent sources of 'big data' to calibrate and validate urban CA models and to capture the role of human actors and their impact on urban change dynamics; and (4) to strengthen theory-based CA models that comprehensively explain urban change mechanisms and dynamics. We conclude by advocating cellular automata that embed agent-based models and big data input as the most promising analytical framework through which we can enhance our understanding and planning of the contemporary urban change dynamics.

Corresponding author:
Yan Liu, University of Queensland, School of Earth and Environmental Sciences, Brisbane, Queensland 4072 Australia.
Email: yan.liu@uq.edu.au

Keywords
agent-based modelling (ABM), big data, cellular automata (CA), future research directions, human behaviours, multi-dimensional urban change processes

I Introduction

Cellular automata (CA)[1] were first proposed in 1943 by Stanislav Ulam and John von Neumann, while working on the Manhattan Project at Los Alamos (Ilachinski, 2001). Using rudimentary computing machines, Von Neumann speculated that cellular automata might be good analogies for self-replication, and Ulam provided two-dimensional CA models that exhibited self-replicating generative properties (Dyson, 2012). It was not until the early 1970s when John Conway invented the *Game of Life* board game that more serious efforts to apply CA to real systems first emerged. Although CA can be applied to any system of objects that can be formally replicated, its initial application was to systems whose elements could be arranged on a regular two-dimensional lattice. Tobler (1975, 1979) pioneered this work in the geographic systems domain, relating this to image processing on the one hand and cartographic transformations on the other. Parallel developments from remote sensing (where pixel arrays are the analogue), from fractal geometries, and from raster-based geographic information systems supported the effort in broadening the range of applications to ecological and urban and regional systems. These applications began to expand greatly some two decades after Conway's first demonstration of the *Game of Life* (Batty et al., 1997).

Over the past four decades, the study of urban systems has been dramatically transformed by the emergence and application of CA models designed to simulate urban land use change from a bottom-up perspective. The fundamental elements of urban CA models are: (a) individual spatial units (i.e. cells) defined by their location (i.e. cell space); (b) geometry (i.e. shape of the cells on a grid, which can be regular or irregular); (c) attributes (i.e. the state of cells; for instance, land use type) that evolve through time and over space; and (d) a set of specified rules governing the transition of cell states within their neighbourhood. The broad aim of urban CA modelling is to capture the rules that determine the way in which the state of a cell changes with respect to what happens in its neighbourhood and, collectively, how these changing states generate meaningful patterns that represent possible paths the spatial system being simulated can take in the future.

Notable achievements have been made in advancing CA models to simulate urban systems since the early efforts of Tobler (1979) (Santé et al., 2010; Li and Gong, 2016). These include the shift from arbitrary and fixed cell grid representation to irregular, flexible entity-based representations (O'Sullivan, 2000; Bithell and Macmillan, 2007; Pinto and Antunes, 2010); the extension from a rigid topologically-based neighbourhood definition to a flexible semantic definition adapted to each entity being represented (Moreno et al., 2009; Van Vliet et al., 2009); the use of better quality datasets for CA model calibration (Pontius and Petrova, 2010); new calibration techniques ranging from simple statistical and probabilistic methods to more sophisticated computational intelligence techniques including neural networks, deep learning, machine learning, and heuristic optimisation (Li and Yeh, 2002; Feng et al., 2011; Feng and Liu, 2013); and the ongoing efforts making urban CA models available as software, including the SLEUTH model (Clarke et al., 1997; Silva and Clarke, 2005; Rafiee et al., 2009; Chaudhuri and Clarke, 2013; Rienow and Goetzke, 2015), the CA_MARKOV model (Shirley and Battaglia, 2008; Václavík and Rogan, 2009; Adhikari and Southworth, 2012), iCity (Stevens et al., 2007), Metronamica (Stanilov and Batty, 2011; White et al., 2015), and the more recently developed FLUS-CA model by Liu et al. (2017).

Yet, contemporary practice in urban CA modelling continues to be limited by its persistent over-simplification of urban processes that are both multidimensional and complex (Torrens and O'Sullivan, 2001; Salvati and Serra, 2016). While models (in general) are abstractions and thus simplifications of reality, most urban CA models focus on simulating the spatio-temporal processes of urban expansion. Few attempts to date have been made to capture the wider chaotic dynamics of the urban system that cover a broader spectrum of urban change processes, including urban regeneration, densification, gentrification, inner-city decline, polycentric formation, de-urbanisation, and urban shrinkage, to name but a few. Current urban CA models are also limited in simulating the influences of physical and economic factors, and often overlook the summative impact of human decision-making behavioural factors on urban development. This shortcoming may be due to the inability of CA models to represent entities at the finest disaggregate levels, such as individual decision behaviours that drive the evolution of social systems (O'Sullivan, 2002). Here, we argue that there is a need for CA models to embrace a wide variety of urban change processes that acknowledge the bottom-up processes that are associated with fractal patterns and dynamics, which taken together offer the capacity to model more realistically highly complex urban systems (Batty, 2007, 2013). We also consider that the integration of CA with other individual-based models such as the agent-based modelling (ABM) provides unique opportunities for urban CA modellers to incorporate the underlying human decision factors that drive urban change. Here, CA is considered as a passive system that provides a cellular space which evolves under different transition rules where the rules are invariant through time, while ABMs are founded on agents acting within the cellular landscape, interacting purposively with one another and with the environment defined by cellular space. Agents thus have the potential to change their own behaviours as they interact in the cellular space.

Given the somewhat diverse practices in the development and application of urban CA models due to the lack of a holistic understanding of the multi-dimensional urban change processes as well as the complex human decision behaviours, there is a pressing need for the CA modelling community to coalesce around a series of research themes through which targeted progress can be achieved. The aim of this paper is to address this need by summarising these contemporary issues and challenges, and then charting a future research agenda towards which we can orientate our collective effort in urban CA modelling.

II Contemporary issues in urban CA modelling

A large number of urban models have been developed for different policy fields at various levels of spatial and temporal resolution (Spiekermann and Wegener, 2018), with the broad aim of capturing the dynamics of space, time, and human choice in relation to urban change. According to King and Kraemer (1993), a model should play at least three key roles in a policy context: (1) to clarify the issues in a debate, such as issues pertaining to the interactions and conflicts between different urban activities and land uses; (2) to enforce a discipline of analysis and discourse among stakeholders; and (3) to provide suggestive feedback and advice primarily in the form of what not to do – since the politics of any practical application often conflicts with what a model suggests. By considering these three objectives alongside the abilities of existing urban models, we identify the following four key issues in the contemporary literature of urban CA modelling. These include:

- the restrictions of CA modelling to urban expansion rather than the multi-dimensional processes of urban change, such as urban regeneration, densification and gentrification, in-fill development, sprawl, as well as urban shrinkage and vertical urban growth;
- the lack of factors representing individual human decision behaviours and their collective implications for urban change;
- the minimal effort in drawing on emergent sources of 'big data' to calibrate and validate CA models and to capture the role of human actors and their impact on urban dynamics; and
- the absence of theories in CA modelling that comprehensively explain urban change mechanisms and dynamics.

We discuss each of these in turn.

1 The restrictions of CA modelling to urban expansion rather than the multi-dimensional processes of urban change

Following the processes of urbanisation and suburbanisation which came to dominate urban growth in the post-war years, one extreme and diametrically opposed trend – urban shrinkage – has more recently emerged. This trend is characterised by low fertility rates, outmigration of young families, declining productivity and the lack of a skilled workforce (Martinez-Fernandez et al., 2012). These exist primarily in the older industrial regions of Europe (Northern England, Scotland's Clydeside, Lorraine, and the Rhine-Ruhr region), in large parts of the post-socialist countries in eastern Europe (Großmann et al., 2008; Haase et al., 2016), and in north-eastern United States rustbelt cities such as Buffalo, Cleveland, and Pittsburgh (Wiechmann and Pallagst, 2012). Recent shrinkage has also been observed in some cities in China, Japan, and South Africa (Rieniets, 2009; Long and Wu, 2016).

From recent reviews of CA-based urban modelling (Aburas et al., 2016; Musa et al., 2017), it is evident that most existing urban models focus on simulating urban expansion and sprawl (Liu, 2008, 2012; Liu et al., 2013; Sakieh et al., 2015; Pérez-Molina et al., 2017), whereas studies of other types of urban transformation such as gentrification, regeneration, and urban shrinkage have been rather limited.

Recent CA modelling concerned with the process of urban regeneration involves the application of ABMs that employ cellular representations and complex model calibration routines. For instance, Jordan et al. (2014) developed an agent-based model to simulate residential mobility and assess the impact of urban regeneration policy on the housing choice behaviours for a residential community in the United Kingdom. Another application by Zheng et al. (2015) simulated land use change in an urban renewal district in Hong Kong by combining the conversion of land use and its effects at the small area level (CLUE-S) model with a Markov prediction model. Their work demonstrates the utility of the modelling framework as a policy tool for scenario analysis of urban renewal, but lacked the capacity for capturing the entire life-cycle of urban regeneration from land evacuation to redevelopment.

Urban gentrification, as a process that is usually accompanied by urban regeneration, has also gained attention in the CA modelling community. O'Sullivan (2000, 2002) simulated urban gentrification at the micro level by applying an irregular graph-based CA architecture, drawing on the principle of proximal space and rent gap theory in residential property markets. Following on from this, Diappi and Bolchi (2006) investigated the gentrification process by applying an urban spatial model of gentrification also based on rent gap theory (Smith, 1987). Their model included behavioural rules for each type of agent such as homeowners, landlords, tenants, and developers, with non-linear interactions between agents at the local level which can then produce different configurations of the system at the macro level. Similarly, Torrens and Nara (2007) developed a hybrid cellular and agent-based automata model that allowed for the representation of co-interactions among fixed and mobile entities in urban settings as well as across multiple spa- tial scales. While this hybrid approach is useful in representing human behaviours in complex adaptive urban systems, it could also benefit from considering more top-down factors such as urban planning and zoning, social biases, and cultural factors in the form of traditional customs and behaviours shared by certain ethnic communities.

Beyond urban expansion, some scholars have also modified the strict CA model by adding constraints and processes that enable the simulation of urban shrinkage (Sante´ et al., 2010; Schwarz et al., 2010). For instance, Haase et al. (2010) developed an agent-based model that computes spatially explicit household pat- terns, housing demand, and residential vacan- cies. This model was applied to simulate urban shrinkage in Leipzig, Germany. An updated version of a joint system dynamics (SD)-CA model and an ABM was developed by Haase et al. (2012) using an integrated dataset of land cover and cadastral data, with specific indicators of urban shrinkage including population decline, change in household structure, housing costs, proximity of growing and declining neighbourhoods, decline in land use density through the expansion of brownfields and the emergence of unintended green spaces, and the concomitant consumption of land for new development. The integration of different modelling approaches enabled the inclusion of data on demographic, socioeconomic, housing, and governance features of urban shrinkage. However, these models capture neither decision-making processes nor the relationships between housing supply and the specific demands of the individual agents, both of which are key issues that should be considered. As such, there is a pressing need to develop models that address the question of how socio-spatial land use change dynamics contribute to urban shrinkage.

Urban development is a process that involves changes in urban form in both the horizontal and vertical geometric dimensions. Since the 1950s, planners have commonly perceived that vertical growth in the form of tall buildings increases urban density and is desirable in eliminating urban sprawl, increasing housing affordability, reducing energy costs for transportation, and distributing resources in a more compact way (Goetz, 2013). Therefore, vertical urban growth represents one of the most important aspects of ‘smart growth’ and ‘sustainable development’, which further transforms the morphology and functioning of cities (Palme and Ram´ırez, 2013). For example, the development of high-rise buildings has also been considered as a symbol of advancement, wealth, and efficiency, which resonates particularly with the politics in developing countries where the transformation of cities is primarily concerned with a large population base (Palme and Ram´ırez, 2013). Vertical urban growth is also reflected in buildings with various functions including commercial, residential, and industrial land uses (Lin et al., 2014). Some scholars also report the negative effects of vertical urban growth; for example, the densification of urban surfaces can decrease urban efficiency (Jaksch et al., 2016), increase pollution (Aristodemou et al., 2018), accelerate urban heat island effects (Santamouris et al., 2015), amplify road traffic noise (Tang and Wang, 2007), and adversely affect resident habits and lifestyles. In fact, there is no longer widespread agreement that the increased compactness of vertical urban growth reduces either urban sprawl or the use of energy in cities.

However, despite the divergent views on vertical urban growth and its impact on urban form, it is only in recent years that CA models of vertical urban growth have emerged (Lin et al., 2014; Koziatek and Dragićević, 2017), though 3D CA models have for some years already been applied in other fields. Despite the fact that CA can be easily extended to incorporate 3D growth, most simulations have only dealt with two-dimensional space, with the exception of early ad hoc examples such as those developed by Batty and Longley (1994), Semboloni (2000), and Semboloni et al. (2004), who generated fractal growth in 3D when exploring the space-filling properties of urban regions but did not pursue the simulations further. However, models that better account for simulating the dynamics of spatio-temporal processes and patterns associated with vertical urban growth need to be further developed. This limitation is not just confined to CA models; all types of urban simulation models focus exclusively on simulating horizontal expansion (Schwarz et al., 2010), despite the global trend of increasing urban density through high-rise living. For instance, in CA-based urban models, the state of an 'urban' cell is typically considered as the final state that will not change regardless of any further development on the cell through multi-functional development (such as high-rise buildings that have multiple functions including retail and residential use), densification, or vertical urban growth. Furthermore, transportation and interactions between cells are usually confined to two dimensions, despite there being significant interaction effects in the third dimension. The advancement of LiDAR and building information models (BIM) representing physical and functional characteristics of places and the vertical mixed use of multi-level buildings has begun to fulfil an analytical need in urban CA modelling (Batty, 2000). Links between 3D representation and CA are being explored using generative grammars such as those embodied in ERSI's CityEngine software (Koziatek and Dragićević, 2017). As a practical example and a pioneer study, Lin et al. (2014) developed a 3D CA model that adopts a linguistic approach to simulate building distribution patterns across space and time. This model combines a series of variables such as population density, building height, and accessibility to transportation nodes. It provides a modelling framework that captures vertical urban growth patterns across space from the city centre through the fringe, periphery, and hinterland. More recently, Koziatek and Dragićević (2017) developed the 3D iCity to model vertical urban development in the Canadian city of Surrey. To further develop applications of this type of 3D urban CA model, we need to confront a series of challenges which include: (1) insufficient data to parameterise the model at the initial stage of model operation; (2) the inability to validate the results over the simulation period; and, (3) the inaccuracies in simulating buildings with heterogeneous heights using representations based on varying neighbourhood cells (Lin et al., 2014; Koziatek and Dragićević, 2017).

2 The lack of factors representing individual human decision behaviours and their collective implications for urban change

Compared to natural or agricultural landscapes, urban systems are strongly influenced by both human and environmental factors. In the context of the urban environment, this can be interpreted as a concern for the degree of harmony between city residents and their everyday urban surroundings (Pacione, 1990). Land use changes are the consequence of various drivers that include the economy, technology, and human behaviours and decisions (Agarwal et al., 2002). Among these drivers, people's behaviours and decisions such as the housing choices of residents and investors, investment decisions by land developers, and urban planning regulations and design have played an overriding role. Haase and Schwarz (2009) provide a general but comprehensive overview of the major components of an urban system, consisting of three components collected through four types of relationships: (1) the impact of the human sphere on land use; (2) feedback of land use on the human sphere; (3) the impact of land use on the environment (including ecosystems); and (4) feedback of the environment on the human sphere (Haase and Schwarz, 2009). The principal driving force for urban change is the way in which the human sphere and the pressure it exerts on urban land use can have environmental consequences, which can in turn affect future human decisions and behaviours. Since cities are places where people deliberately come together to interact, our understanding of the evolution of cities must be enriched by studies of networks, interactions, connections, and transactions between humans and the environment, ranging from individual, local, and regional, to global scales (Batty, 2013).

In the process of coupling human and environmental systems with urban modelling, understanding how humans make decisions is of paramount importance (Gimblett, 2002; An, 2012). Human decisions and subsequent actions change the structure and function of many environmental systems, which in turn influence human decisions and actions. To assess the extent to which current urban CA models have captured the human–environmen- tal dynamics, Table 1 presents a selection of urban CA models that considered (or not) dif- ferent levels of human-decision complexity categorised by Agarwal et al. (2002), as illu- strated in Figure 1.

Table 1 shows that some CA models, particularly those in more recent studies, have comprehensively considered human–environment interactions in the modelling process. Furthermore, the integration of CA and ABM would allow the decision behaviours of various 'agents' to be incorporated and simulated in the cellular space in order to understand the emergent spatial patterns through time (Batty, 2009; Waddell, 2002; Jjumba and Dragićević, 2012).

Nevertheless, the existing literature reveals a somewhat simplified capacity to model complexity in human decision-making using more or less sophisticated measurements or conceptualisations of stochasticity (De Almeida et al., 2003; Feng et al., 2016). These models are yet to capture the fundamental role of humans and their interactions within the built environment which shapes our cities, and there is no unique approach to represent an individual's decision behaviours and how collectively such individual decision behaviours impact on the change in urban form. The critical role of human behaviours in urban models has been generally overlooked for two main reasons: (1) the challenges of developing realistic operational models to incorporate the impact of decision behaviours and other drivers of urban transformation (Elliott and Kiel, 2002), and (2) the difficulty of even observing, but also collecting, verify- ing, and validating data reflecting an individu- al's decision behaviours as dynamic contextual parameters (Crooks et al., 2008). Subsequently, CA models rarely have had the capacity to incorporate the most comprehensive drivers of human behaviours with respect to land use change (Agarwal et al., 2002). The emergent sources of 'big data' in various forms at the individual level could shed new light on how urban modellers might tackle this fundamental issue in urban CA models, but this will require new forms of data concerning decision processes which currently barely exist. Moreover, the lack of effort in using such data in current CA modelling practice needs to be addressed.

3 The minimal effort in drawing on emergent sources of 'big data' to calibrate and validate CA models and to capture the role of human actors and their impact on urban change dynamics

A persistent challenge in urban CA modelling concerns the fitting of urban CA models to data.

Table 1. The levels of modelling human-decision complexity in urban CA models, adapted from Agarwal et al. (2002: 6).

Level	Human-decision complexity	Model components	Examples
1	No human decision-making (only biophysical variables)	Land use (e.g. urban or non-urban, roads, different land use types) Environment (e.g. topography)	Rafiee et al., 2009; Van Vliet et al., 2009; Feng et al., 2011; Liao et al., 2016; Liu and Feng, 2016.
2	Human decision-making assumed to be related deterministically to selected human variables (such as population size, change or density)	Human sphere (e.g. demand rules, population, economy, planning, accessibility via transportation network) Land use (e.g. suitability rules, land use functions)	Verburg and Overmars, 2007; He et al., 2008; Lin et al., 2014.
3	Human decision-making seen as a probability function depending on socioeconomic and/or biophysical variables *without* feedback from the environment to the choice function	Human sphere (e.g. population, household, jobs, employment) Land use (e.g. single-family residential, multi-family residential, commercial, industrial, transportation, public) Environment (e.g. undeveloped land)	Jantz et al., 2010; Haase et al. 2012; Fuglsang et al., 2013; Kamusoko and Gamba, 2015; Rienow and Goetzke, 2015; Sakieh et al., 2015; Berberoğlu et al., 2016; Tian et al., 2016.
4	Human decision-making seen as a probability function depending on socioeconomic and/or biophysical variables *with* feedback from the environment to the choice function	Human sphere (e.g. urban growth, policy simulation and evaluation) Environment (e.g. habitat change and habitat fragmentation)	Dabbaghian et al., 2010.
5	One type of agent whose decisions are modelled overtly with regard to choices made about variables that affect other processes and outcomes	Human sphere (e.g. initial capital, lending amount) Environment (e.g. distance to CBD/toll road) Land use (e.g. land cover/values) Interaction: developer-environment relation (e.g. land find, assess, and decide)	Torrens and Nara, 2007; Jordan et al., 2014; Wahyudi et al., 2019a.
6	Multiple types of agents whose decisions are modelled overtly with regard to choices made about variables that affect other processes and outcomes	Human sphere (e.g. multiple agents, socioeconomic change) Environment (e.g. climate change) Land use (e.g. historical land use change) Interactions between variables via system dynamics and ABMs	O'Sullivan, 2002; Semboloni et al. 2004; Diappi and Bolchi, 2006; Liu et al., 2013; Liu et al., 2017; Wahyudi et al., 2019b.

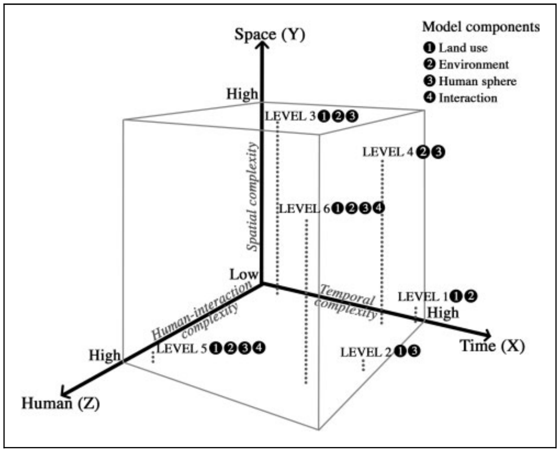

Figure 1. Three dimensions of modelling human-decision complexity in urban CA models with model components, revised from Agarwal et al. (2002: 7).

This is usually referred to as model calibration and validation. Model calibration typically involves determining the parameter values which specify the model's rules to particular applications, while validation involves testing a simulation model with a set of sample data different from those applied to model calibration and then evaluating the output, with the ultimate goal of producing accurate and credible results. In fact, many applications are based on rules that are plausible and appear to reflect rational decision-making, but are rarely fitted and tested in any way (Crooks et al., 2008; An, 2012). The rules can be inferred from vari- ables that define various attributes of cells and neighbourhoods, and then are matched against outcomes using various data mining and multivariate methods. However, these rules tend to be arbitrary unless they are consistently examined and pruned from the list of causes that determine how urban development occurs. Many new methods of multivariate analysis involving neural nets and related decision-making/data-mining techniques can be developed, but there are only a small number of applications so far. However, currently, the calibration of urban CA models is dominated by the type of rules that need to be defined for such models, and this relies more on theory than real world applications. The actual pro- cesses of urban development need to be ana- lysed, after which rules based on these actual processes need to be devised and validated with respect to what the data are saying about urban development. This is a critical subject to broach and requires a systematic review in itself so that we might make progress on model validation and calibration. Verification, which is usually defined as ensuring that the model code is functioning correctly, is also becoming a more significant part of the process of model development as more models with different software requirements and diverse data structures come to characterise ever more complicated CA models.

Historically, research on urban land use change has primarily relied on remotely sensed imagery and aggregated census data or data from small-scale surveys. Over the past decades, however, rapid global change in the digitisation of records, expansion of networks, and the computerisation of societies has created large quantities of data with both spatial and temporal features relevant to urban forms and urban change dynamics (Glaeser et al., 2018). With the globalisation of information and technologies, the world is becoming increasingly connected with virtual, perceived, and real spaces. Computation and analytics driven by the so called 'big data', defined in terms of its large volume, have become essential to tackle fundamental urban issues (Chen et al., 2012; Batty, 2018). Computers have been embedded into almost every conceivable type of objects, and the rise of the Internet of Things (IoTs) is changing how humanity interacts with itself and with the environment that we have built (Hammi et al., 2017). Accordingly, data in cities is being generated, recorded, and stored in unprecedented quantities from sources rang- ing from parking meters to smart-card-based devices, from crowdsourced application users to hot-line callers, and from participatory mapping applications to activities involving citizen science. Big data from diverse real-time streaming offer information that traditional censuses and survey data cannot provide. For example, social-media sites such as Facebook can track user locations to reflect the formation and dissolution of networks in real time, and cell phone companies can map the movements of their customers, which are contrasted with the limited information gleaned from traditional data collection. The emerging sources of big data offer modellers the opportunity to capture and reconstruct the spatial movements and decision behaviours of individuals, making it possible to develop 'big CA' models, that is, CA models at a fine spatial scale that incorporate individual human decisions, their interactions with each other and with the built environment in which they reside, as well as the way these human-environment interactions would collectively reveal the spatial and temporal dynamics of cities. Meanwhile, there also exist some emerging concerns as a consequence of the grow- ing size and complexity of urban big data. These pose a suite of daunting challenges for urban scholars with regard to the substantial investment in time and technical abilities needed to ensure that such data deliver the required information, while also considering the computational expense of employing such data across a variety of spatial and temporal scales.

The first challenge lies in the nature of urban big data. The agglomeration of disparate big data sources spreads across a city as a digital skin woven by the 'IoTs', communication network, monitoring individuals, organisations, and governments (Rabari and Storper, 2014). Among these, three major big data sources are thought to be valuable to urban sciences (Arribas-Bel, 2014). These include:

- bottom-up data collected by individual mobile/computer users;
- intermediate data aggregated and created by compiling primary source data; and
- top-down data released by government, public organisations, and the private sector.

Collectively, these data are not only archived in volumes and across different formats, but are also dynamic and continuously generated and updated, with significant differences in quality, coverage, accuracy, and timeliness. Data acquisition requires a comprehensive understanding of their sources and structures, and more importantly, a clear mind for managing the appropriate information from the vast ocean of data for research purposes.

The considerable value of urban big data is demonstrated when it can be linked and fused with other data sources, though this in itself is a challenge, that is, how to manipulate models that contain big data. It is critical to realise that the promise of big data is equally poised with numerous difficulties across many dimensions (Dong and Srivastava, 2013). Big data manipulation involves multiple stages – from data extraction, filtering, cleaning, formatting, re-structuring, and integrating – before being ready for analysis. This procedure is usually time-consuming and costly, requiring suitable computational facilities (e.g. software, interface, and storage) and computer skills (e.g. data mining, programming and statistical analysis) (Labrinidis and Jagadish, 2012). In particular, big data that can be used for urban modelling are usually geospatially explicit, which requires geographic thinking, methods, and spatial analytical and visualisation skills. Furthermore, data privacy is another key issue since a large body of big data are user-generated via social media, fostering citizen engagement in urban activities but also exposing personal information with the potential for use across unappropriated purposes if without protection (Perera et al., 2015).

When data are ready for analysis, the challenge becomes how to organise existing data in meaningful ways to allow modelling and comparison of hidden patterns and relationships among dynamic urban systems. Data-driven urban modelling, particularly where the data tend to be big, is considerably more sophisticated than the simpler skills we used in the past for locating, identifying, analysing, and citing data (Wu et al., 2014). What needs to be primarily kept in mind is the purpose of modelling, rather than adapting studies to data that are available. An opportunistic way of conducting data-driven research may lead to interesting observations, but often bypasses ideas not meaningful to tackling real-world issues (Hashem et al., 2015). Past studies have revealed how big data can help to improve the planning of smart cities in at least four respects (Lim et al., 2018): preventive administration (e.g. civil complaint and crime prevention), operational management (e.g. trash collection and traffic control), network development (e.g. bus service scheduling and Wi-Fi hotspot optimisation), and information diffusion (e.g. pollution monitoring and intelligent navigation). Thus, changing city operations and improving the lives of city residents should be the ultimate goal of using big data in urban modelling, where urban research should combine both academic rigor and practical knowledge.

4 The absence of theories in CA modelling that comprehensively explain urban change mechanism and dynamics

Contemporary sciences usually invoke theory that enables scientific predictions to be compared with reality, with respect to the variations in the phenomena of interest across time and space (Batty, 2009). To deal with the complexity of such realities, theories are translated into a form that enables them to be represented as mathematical or logical models, with computers acting as the laboratory in which the simulation of reality takes place (Haase et al., 2012). Given a good theory, an urban model would be constructed, validated, and then used as a vehicle for refining the theory through 'what if' style experiments and sensitivity testing (Batty, 1976; Crooks et al., 2008). However, the overt role of theory has faded in many contexts as urban models embed theory within themselves (Crooks et al., 2008); that is, theory is often derived as the model is constructed. For example, UrbanSim as a simulation platform for supporting planning and analysis of urban development has been developed by adopting a micro-simulation strategy that directly represents the choices of households, businesses, developers, and governments (representing policy inputs) in the real estate market in a way that is 'behaviourally natural and intuitive' and can be understood by non-technical stakeholders (Waddell, 2011: 217). Because models come to control all of the inputs and parameters, most social systems cannot be represented in the form of a theory that guarantees any measure of closure. Earlier generations of urban models have demonstrated the extent to which they can be validated in terms of the goodness-of-fit of the model to an existing system, but the difficulties and failures in model validation are features that exist for both the theory involved and data from the observed reality (Crooks et al., 2008).

A large number of urban models have been developed over the last half century, and most can generally be classified into two categories: top-down aggregate models or bottom-up disaggregate models (Tan et al., 2014). Top-down models, such as economic equilibrium models, are often constructed by breaking down a system to gain insight into its compositional subsystems in a reverse engineering fashion; such a modelling approach is typically based on traditional macroeconomic theories and is unable to deal with micro-level decision-making or with small-scale social and environmental problems

(Itami, 1994). On the other hand, the bottom-up modelling approach is formulated by piecing together sub-systems or its components to formulate more complex systems; this modelling approach has become more dominant in the field of urban modelling with the ongoing development of algorithms to represent the dynamics of urban systems. CA models, as one of the most popular bottom-up approaches, can capture changes in urban morphology through simple and flexible transition rules (Sante´ et al., 2010; Feng and Liu, 2016). However, CA models inevitably have the common weakness of all bottom-up models: modellers have done little to link their models to urban theories which exist at a more aggregate scale (Torrens and O'Sullivan, 2001). They have also faced difficulties in incorporating human decision behaviours (Haase et al., 2010; Arsanjani et al., 2013) and in capturing the macro-scale social and economic driving forces of urban change (Han et al., 2009).

The theoretical orientations of many CA and ABMs remain implicit and hidden, often covered by ad hoc assumptions about modelling structure, process, and software interfacing (Crooks et al., 2008). Moreover, their processes, although explicit in these models, are almost invisible with respect to observation and data. In many cases, the development of an urban CA model is only an additional application of some simple structures that are adjusted for a local context (Couclelis, 2002). In the increasingly diverse array of CA modelling applications, such models are considered generic; they can be applied and fitted to data and processes in any particular field, subject to use for particular purposes, and hence largely independent of theory and practice (Batty, 2007). Although a theory is not necessarily required to guide this kind of urban modelling, many facets of theorising and thinking should be brought to bear on model construction, and to understand the mechanisms of urban change dynamics. For instance, location theory, which has long been used in the study of urban spatial structure, is reflected in the equilibrating micro-economics of the individual and the firm. This theory provides important insights into residential development, and there is clear potential for invoking such theory in any CA or ABM that captures human behaviours. In short, the scope of urban models and theory is now considerably wider than in the past, and urban CA models must respond to this increasing complexity by explicitly embracing the most appropriate theories pertaining to economic and social decision-making. This is a major challenge given that it requires theory that deals not only with the static structure of the space but also its tem- poral dynamics, though this has been slow in progressing over the recent decades.

Developing urban models based on explicit theory has two benefits for modellers. First, it simplifies the process of identifying driving factors – physical/environment, institutional, or human decision behaviour factors – on urban change dynamics, particularly with regard to modelling the complexity of human decision-making behaviours. Second, the theory contributes to stakeholders' involvement with respect to an in-depth understanding of causes and feedbacks of changes in human behaviours. Such model applications could well be used to cumulatively modify human behaviours (Davis et al., 2015). This benefit is critical when stakeholders need to encourage city residents to behave in a certain way. A good example is the positive influence of transit-oriented development that increases the amount of residential, business, and leisure space within walking distance of public transport, thereby enhancing accessibility and promoting healthy community. It is thus necessary that urban modellers pay more attention to develop models in accordance with theories of human behaviour. On the other hand, some stakeholders such as politicians may also prefer a 'black box' approach, whereby a result is generated from a model that fully explores a theory. In this case, CA models engaging Artificial Neural Networks (ANNs) as a machine learning approach, for example, would appear more appropriate (Li and Yeh, 2002).

III Towards a future research agenda

Responding to current challenges in urban CA modelling, we now map out four interrelated thematic areas that we argue require concerted attention by the urban CA modelling community. These are: (1) to build models that comprehensively capture the multi-dimensional processes of urban change, including urban regeneration, densification and gentrification, in-fill development, as well as urban shrinkage and vertical urban growth; (2) to establish models that incorporate individual human decision behaviours into the CA analytic framework; (3) to draw on emergent sources of big data to calibrate and validate urban CA models and to capture the role of human actors and their impact on urban change dynamics; and (4) to strengthen theory-based CA models that comprehensively explain urban change mechanisms and dynamics.

1 Modelling multi-dimensional processes of urban change

As discussed in the previous sections, few urban CA models can simulate multiple urban forms and processes. Although some of these gaps are being addressed to a certain extent by a number of researchers, such as O'Sullivan (2002), Diappi and Bolchi (2006), and Haase et al. (2010, 2012), the challenges faced by developers of urban CA models producing single urban form outputs lie in how to combine the multiple background processes to produce a range of urban forms and phenomena – from urban regeneration, gentrification, densification, vertical growth, to urban shrinkage. In order to overcome the weaknesses and limitations in current scholarship, future research needs to consider the following.

First, it is critical that models incorporate various types of drivers and constraints associated with urban transformation so that these can be applied across multiple metropolitan contexts. By doing so, complex transition rules need to be considered to determine how, when, and to what extent a given land parcel might be developed through time. For example, when modelling urban regeneration and gentrification, one often needs to change the order in which the transition rule sets are implemented so that the various effects of regeneration and gentrification can be incorporated over time. We argue that urban growth and decline as simulated in CA models is not symmetric; the way the transition rules are implemented differs for modelling growth and modelling decline.

Second, the selection of parameters should consider qualitative factors such as land ownership and land lease, land use density change, and land parcel shape by adjusting the CA model's parameter settings. For instance, the choice of the neighbourhood type and size, which has significant impact on the global behaviour of a CA model, can be adjusted contingent on varying perceptions of the relative merits of other neighbourhoods in the wider urban system.

Third, 3D urban modelling needs to be enhanced, especially for modelling vertical growth. To this end, LiDAR datasets are a promising source of information for 3D urban modelling, but collecting and processing these datasets is both time consuming and costly. Land use attributes at the parcel level are being collected slowly by various national mapping agencies, but policymakers should provide more facilities for universities and the private sector to collect and process the data so that it can be used for urban modelling – without the need for intensive pre-processing. As a result, creating a platform for data sharing can provide new opportunities for urban modellers to model in 3D. Scholars also need to develop more complex transition rules for 3D modelling; to achieve this, close collaboration between programmers and urban modellers is critical.

2 Incorporating human decision behaviours into the CA modelling framework

The integration of CA with other types of models has been suggested by various scholars (Hewitt et al., 2014; Musa et al., 2017). This integration has only been partially addressed through participatory modelling by obtaining feedback from stakeholders to calibrate the CA model and enhance its performance (Hewitt et al., 2014). To close this gap, ABMs are an unparalleled tool for modelling human decisions. However, the main challenges of ABMs arise from their complexity in implementing and designing the real-world rules for the relevant agents, as noted earlier. Often, these rule sets are generated using plausible hypotheses but are never tested due to lack of observational data (Wahyudi et al., 2019a). The CA-ABM hybrid approach needs to consider top-down (as well as bottom-up) concepts to address issues such as urban planning and zoning, transportation, social biases, and cultural factors, which may all be reflected in micro-scale survey data or other primary data sources (Torrens and Nara, 2007).

Moreover, both CA and ABM have limited geographic functionalities when considered in isolation (Torrens and Benenson, 2005). In contrast, the integration of CA and ABM offers a powerful spatial approach to modelling complex geographic systems that are affected by physical and human factors at multiple scales ranging from the individual to the metropolitan region, as tested by Batty (2007) and Torrens and Benenson (2005). An integrated CA-ABM model would provide unique opportunities for urban modellers to address various types of urban transformation with regard to human decisions and preferences. The integration with other types of models to develop extended suites of urban models has been attempted particularly in 'Metronamica' (Research Institute for Knowledge Systems, 2013) and the more recent 'GeoDynamiX' (Flemish Research and Technology Organisation, 2018), but these tend to reflect a loose coupling across different spatial scales, and further work remains to be done.

The development of urban models that can capture human decision behaviours and their interactions with the built environment requires micro-scale spatial and social survey data, which can be furnished in part by the use of public participatory GIS mapping (McCall, 2003; Aburas et al., 2016) which incorporates the knowledge of stakeholders – residents, land developers, and urban planners – who are key drivers in the urban transformation process. The emergent sources of big data from government, social media, citizen science, and other location-based services and devices can also serve as excellent input for urban modellers to understand and model urban change dynamics.

3 Drawing on emergent sources of big data to calibrate and validate urban CA models and capture the role of human actors and their impact on urban change dynamics

The emergence of open and new data available from various sources has presented significant opportunities for research in the urban sciences. Entering into the new era of big data, ever-increasing quantities at near real time will ultimately change the ways in which human agents interact with each other and with the urban space they occupy and transform; these pose new challenges to urban modellers and researchers (Batty, 2018), and much effort should be devoted to conquer the aforementioned big data challenges. Considering that these challenges are interrelated since information creation, data collection, manipulation, analysis, and modelling are interdependent activities, we propose the following four considerations for future work of modelling human-environmental interactions and urban change dynamics.

The first consideration is to be problem-oriented and to clarify the purpose of urban modelling in order to identify the requirements of data, analytic system, and other elements involving in the modelling process (Lim et al., 2018). This direction should guide us through the whole procedure from data acquisition to building models.

Second, more advanced computing paradigms and methods need to be developed to retrieve, store, manipulate, integrate, and analyse such large data volumes across multiple sources. In particular, interdisciplinary approaches combined with complex spatio-temporal analysis and models are needed for transformative innovation and effective and timely solutions to urban problems (Croitoru et al., 2013). These methods should fuse the bottom-up user-contributed information to more traditional top-down data sources so that we can move in between short-term snapshots and long-term planning to address real world issues (Crooks et al., 2016).

Third, the quality and integrity of user-generated data should be controlled and improved by paying more attention to the security of the virtual environment (Perera et al., 2015). The privacy of information providers should be protected and enforced by strict laws and regulations – from the time data are captured in mobile or computer terminals to the point at which data are securely extracted and stored. Only by doing so can IoT solutions gain users' confidence and in turn provide trusted data (Hammi et al., 2017).

The fourth consideration is to promote data sharing and minimize conflicts between data-related stakeholders, including citizens, visitors, local governments, and commercial companies (Lim et al., 2018). The art of using big data for modelling urban dynamics lies in effective matchmaking among the concerns of all urban activity participants who are the data contributors as well as the beneficiaries by the creation of useful contents from big data. Urban modelling serves as a platform absolving information from these participants and optimising solutions to benefit them (Dong et al., 2015).

Bearing these four considerations in mind, we have the potential to study, test, and refine ideas and theories pertaining to diverse urban problems at various spatial and temporal scales, and to open up a richer context for advancing urban modelling, and eventually pave the way for the systematic implementation of new technologies in the computational urban sciences.

4 Strengthening theory-based CA models that comprehensively explain urban change mechanisms and dynamics

Choosing relevant theory should be the most important consideration even though this can be a challenging task for modellers, especially given the large number of theories, many of which have similar or overlapping constructs. Model developers may draw upon specific theories either at the beginning of the design process or after conducting preliminary research to indicate which theories are likely to be most relevant. CA models can accommodate urban morphology and land use theory by defining appropriate transition rules, model structures, and relevant parameters. However, while contextual or environmental variables are relatively straightforward to consider (Wu and Webster, 2000), it is far less likely that we can develop CA models based on full theories of human behaviours since they require an in-depth understanding of the causes and feedbacks of changes in such behaviours (Davis et al., 2015).

There is a need for methods that can better represent the effects of human behaviours on the urban development process. Related methods for extracting appropriate rule sets have been employed by De Almeida et al. (2003) and Feng et al. (2016), but much remains to be done to incorporate various decision models commonly used in ABM, including micro-economic models, space theory-based models, psychosocial and cognitive models, institution-based models, experience- or preference-based decision models (rules of thumb), participatory agent-based models, empirical- or heuristic-based models, and evolutionary programming methods (An, 2012). Some machine learning models such as ANNs have also been developed (see, for example, Li and Yeh, 2002), but these models need to avoid the many traps of autocorrelation and multi-collinearity, along with a host of ad hoc pattern-matching features that are essentially spurious in terms of the way spatial systems function and evolve. Furthermore, CA models based on human behaviour theory need to incorporate the probability-of-occurrence with the conversion cost, neighbourhood conditions, and competition among the different agents (Liu et al., 2017; Wahyudi et al., 2019b). The interactive coupling of the top-down system dynamic demands and the bottom-up CA approach would likely enhance the model's capability for long-term stochastic simulation, reflecting the real city comprised of chaotic human behaviours.

IV Conclusion

This article acknowledges the fundamental transformations which have been brought to urban modelling via the bottom-up perspective through cellular automata models. Despite notable achievements in this field over a number of decades, we argue that there remain at least four pressing issues faced by CA modellers in contemporary urban modelling practices. We draw out these issues into a dis- cussion mapping out four interrelated thematic areas that require concerted attention by the urban CA modelling community, which include: (1) building models that comprehen- sively capture the multi-dimensional processes of urban change; (2) establishing urban CA models that incorporate individual human decision behaviours; (3) drawing on emergent sources of big data to calibrate and validate urban CA models and to capture the role of human actors and their impact on urban change dynamics; and (4) strengthening theory-based CA models that comprehensively explain urban change mechanisms and dynamics.

We suggest that the continued growth in CA modelling is in part contingent on tackling these four challenges in order to remain at the vanguard of urban modelling. In tandem with progressing these four themes is the need to develop, implement, and train urban planners and policymakers in the use of CA-based model outputs. This is indeed most important to ensuring that urban CA modelling moves from an approach that is arguably the purview of a relatively select academic community to one that has a place in mainstream policy and practice. The first component of setting urban CA modelling on this path towards mainstream use is the development of user-friendly tools that are embedded within familiar computing environments, of which the iCity model (Stevens et al., 2007) is an excellent example in this regard.

To sum up, the current global interest in sustainable urban development has highlighted the need to deepen our understanding of the processes that underpin urban transformations – and land use and urban models here play a vital role. However, there persists a major void between the real features of our urban systems and the relevant representativeness of these features with the current array of urban CA models. The purpose of this study has been to chart these deficits and to map out a future research agenda for both urban CA models and the CA modelling community. To this end, we encourage scholars to concentrate their efforts on developing multi-dimensional, dynamic, and vector-based models to simulate the holistic processes of urban change. We also highlight the need to better capture the role of human behaviours and decisions by drawing on theories that can facilitate more rigorously grounded modelling outcomes. We advocate cellular automata that embed agent-based models and big data input as the most promising analytical framework through which we can enhance our understanding and planning of contemporary urban change dynamics.

Acknowledgements

We thank the reviewers for the constructive comments on an early version of this article.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by an Australian Research Council Discovery grant [DP170104235].

ORCID iD

Yan Liu https://orcid.org/0000-0002-1612-779X

Note

1. Cellular automata (CA) is usually formulated as a discrete model comprised of a cellular lattice which is often but not necessarily regular, where each cell is classified as one of a number of defined states. The state of each cell evolves over a number of discrete time steps and transition between states is controlled by a set of predefined rules which are based on the states of the nearest neighbouring cells, in strict applications of CA. This modelling approach differs from Agent- Based Modelling (ABM), which simulates the actions and interactions of autonomous agents (in the forms of either individual or collective entities such as organisations or groups), with a view to assess their effects on a given system (Liu et al., 2019). The distinction between CA and ABM is often blurred when agents are located on cells and move between them, and when cells contain agents who act to change the state of the cells in question.

References

Aburas MM, Ho YM, Ramli MF and Ash'aari ZH (2016) The simulation and prediction of spatio-temporal urban growth trends using cellular auto- mata models: A review. *International Journal of Applied Earth Observation and Geoinformation* 52: 380–389.

Adhikari S and Southworth J (2012) Simulating forest cover changes of Bannerghatta National Park based on a CA-Markov model: A remote sensing approach. *Remote Sensing* 4(10): 3215–3243.

Agarwal C, Green GM, Grove JM, Evans TP and Schweik CM (2002) *A review and assessment of land-use change models: Dynamics of space, time and human choice. General Technical Report NE-297*. Newton Square, PA: US Department of Agriculture, Forest Service, Northeastern Research Station.

An L (2012) Modeling human decisions in coupled human and natural systems: Review of agent-based models. *Ecological Modelling* 229: 25–36.

Aristodemou E, Boganegra LM, Mottet L, Pavlidis D, Constantinou A, Pain C, Robins A and ApSimon H (2018) How tall buildings affect turbulent air flows and dispersion of pollution within a neighbourhood. *Environmental Pollution* 233: 782–796.

Arribas-Bel D (2014) Accidental, open and everywhere: Emerging data sources for the understanding of cities. *Applied Geography* 49: 45–53.

Arsanjani JJ, Helbich M and de Noronha Vaz E (2013) Spatiotemporal simulation of urban growth patterns using agent-based modeling: The case of Tehran. *Cities* 32: 33–42.

Batty M (1976) *Urban Modelling*. Cambridge: Cambridge University Press.

Batty M (2000) The new urban geography of the third dimension. *Environment and Planning B: Planning and Design* 27: 483–484.

Batty M (2007) *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-based Models, and Fractals*. Cambridge, MA: MIT Press.

Batty M (2009) Urban modeling. In: Kitchin R and Thrift N (eds) *International Encyclopedia of Human Geography*. Oxford: Elsevier, 51–58.

Batty M (2013) *The New Science of Cities*. Cambridge, MA: MIT Press.

Batty M (2018) Artificial intelligence and smart cities. *Environment and Planning B: Urban Analytics and City Science* 45(1): 3–6.

Batty M and Longley P (1994) *Fractal Cities: A Geometry of Form and Function*. San Diego, CA: Academic Press.

Batty M, Couclelis H and Eichen M (1997) Urban systems as cellular automata. *Environment and Planning B: Planning and Design* 24(2): 159–164.

Berberoğlu S, Akın A and Clarke KC (2016) Cellular automata modeling approaches to forecast urban growth for Adana, Turkey: A comparative approach. *Landscape and Urban Planning* 153: 11–27.

Bithell M and Macmillan WD (2007) Escape from the cell: spatially explicit modelling with and without grids. *Ecological Modelling* 200(1–2): 59–78.

Chaudhuri G and Clarke K (2013) The SLEUTH land use change model: A review. *Environmental Resources Research* 1(1): 88–105.

Chen H, Chiang RH and Storey VC (2012) Business intelligence and analytics: From big data to big impact. *MIS Quarterly* 36(4): 1165–1188.

Clarke KC, Hoppen S and Gaydos L (1997) A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B: Planning and Design* 24(2): 247–261.

Couclelis H (2002) Modeling frameworks, paradigms, and approaches. In: Clarke KC, Parks BE and Crane MP (eds) *Geographic Information Systems and Environmental Modelling*. Upper Saddle River, NJ: Prentice-Hall.

Croitoru A, Crooks AT, Radzikowski J and Stefanidis A (2013) GeoSocial gauge: A system prototype for knowledge discovery from geosocial media. *International Journal of Geographical Information Science* 27(12): 2483–2508.

Crooks A, Castle C and Batty M (2008) Key challenges in agent-based modelling for geo-spatial simulation. *Computers, Environment and Urban Systems* 32(6): 417–430.

Crooks AT, Croitoru A, Jenkins A, Mahabir R, Agouris P and Stefanidis A (2016) User-generated big data and urban morphology. *Built Environment* 42(3): 396–414.

Dabbaghian V, Jackson P, Spicer V and Wuschke K (2010) A cellular automata model on residential migration in response to neighborhood social dynamics. *Mathematical and Computer Modelling* 52(9–10): 1752–1762.

Davis R, Campbell R, Hildon Z, Hobbs L and Michie S (2015) Theories of behaviour and behaviour change across the social and behavioural sciences: A scoping review. *Health Psychology Review* 9(3): 323–344.

De Almeida CM, Batty M, Monteiro AMV, Câmara G, Soares-Filho BS, Cerqueira C and Pennachin CL (2003) Stochastic cellular automata modeling of urban land use dynamics: Empirical development and estimation. *Computers, Environment, and Urban Systems* 27(5): 481–509.

Diappi L and Bolchi P (2006) Gentrification waves in the inner-city of Milan. In: Leeuwen JP and Timmermans HJ (eds) *Innovations in Design & Decision Support Systems in Architecture and Urban Planning*. Dordrecht: Springer, 187–201.

Dong XL and Srivastava D (2013) Big data integration. In: *2013 IEEE 29th International Conference on Data Engineering (ICDE)*, 1245–1248.

Dong X, Li R, He H, Zhou W, Xue Z and Wu H (2015) Secure sensitive data sharing on a big data platform. *Tsinghua Science and Technology* 20(1): 72–80.

Dyson G (2012) *Turing's Cathedral: The Origins of the Digital Universe*. New York: Penguin Books.

Elliott E and Kiel LD (2002) Exploring cooperation and competition using agent-based modeling. *Proceedings of the National Academy of Sciences* 99(3): 7193–7194.

Feng YJ and Liu Y (2013) A heuristic cellular automata approach for modelling urban land-use change based on simulated annealing. *International Journal of Geographical Information Science* 27(3): 449–466.

Feng YJ and Liu Y (2016) Scenario prediction of emerging coastal city using CA modeling under different environmental conditions: a case study of Lingang New City, China. *Environmental Monitoring and Assessment* 188: 540. DOI: 10.1007/s10661-016-5558-y.

Feng YJ, Liu Y and Batty M (2016) Modeling urban growth with GIS based cellular automata and least squares SVM rules: A case study in Qingpu- Songjiang area of Shanghai, China. *Stochastic Envi- ronmental Research and Risk Assessment* 30(5): 1387–1400.

Feng YJ, Liu Y, Tong X, Liu M and Deng S (2011) Modeling dynamic urban growth using cellular automata and particle swarm optimization rules. *Landscape and Urban Planning* 102(3): 188–196.

Flemish Research and Technology Organisation VITO (2018) GeoDynamiX land use solutions. Available at: http://www.geodynamix.eu/ (accessed 9 June 2019).

Fuglsang M, Münier B and Hansen HS (2013) Modelling land-use effects of future urbanization using cellular automata: An Eastern Danish case. *Environmental Modelling & Software* 50: 1–11.

Gimblett HR (2002) *Integrating Geographic Information Systems and Agent-Based Modeling Techniques for Simulating Social and Ecological Processes*. Oxford: Oxford University Press.

Glaeser EL, Kominers SD, Luca M and Naik N (2018) Big data and big cities: The promises and limitations of improved measures of urban life. *Economic Inquiry* 56(1): 114–137.

Goetz A (2013) Suburban sprawl or urban centres: tensions and contradictions of smart growth approaches in Denver, Colorado. *Urban Studies* 50(11): 2178–2195.

Großmann K, Haase A, Rink D and Steinführer A (2008) Urban shrinkage in East Central Europe? Benefits and limits of a cross-national transfer of research approaches. In: Nowak M and Nowosielski (eds) *Declining Cities/Developing Cities: Polish and German perspectives*. Poznan, Poland: Instytut Zachodni, 77–99.

Haase A, Bernt M, Großmann K, Mykhnenko V and Rink D (2016) Varieties of shrinkage in European cities. *European Urban and Regional Studies* 23(1): 86–102.

Haase D and Schwarz N (2009) Simulation models on human-nature interactions in urban landscapes: a review including spatial economics, system dynamics, cellular automata and agent-based approaches. *Living Reviews in Landscape Research* 3(2): 1–45.

Haase D, Haase A, Kabisch N, Kabisch S and Rink D (2012) Actors and factors in land-use simulation: The challenge of urban shrinkage. *Environmental Modelling & Software* 35: 92–103.

Haase D, Lautenbach S and Seppelt R (2010) Modeling and simulating residential mobility in a shrinking city using an agent-based approach. *Environmental Modelling & Software* 25(10): 1225–1240.

Hammi B, Khatoun R, Zeadally S, Fayad A and Khoukhi L (2017) IoT technologies for smart cities. *IET Networks* 7(1): 1–13.

Han J, Hayashi Y, Cao X and Imura H (2009) Application of an integrated system dynamics and cellular automata model for urban growth assessment: A case study of Shanghai, China. *Landscape and Urban Planning* 91(3): 133–141.

Hashem IAT, Yaqoob I, Anuar NB, Mokhtar S, Gani A and Khan SU (2015) The rise of 'big data' on cloud computing: Review and open research issues. *Information systems* 47: 98–115.

He C, Okada N, Zhang Q, Shi P and Li J (2008) Modelling dynamic urban expansion processes incorporating a potential model with cellular automata. *Landscape and Urban Planning* 86(1): 79–91.

Hewitt R, Van Delden H and Escobar F (2014) Participatory land use modelling, pathways to an integrated approach. *Environmental Modelling & Software* 52: 149–165.

Ilachinski A (2001) *Cellular Automata: A Discrete Universe*. Singapore: World Scientific.

Itami RM (1994) Simulating spatial dynamics: Cellular automata theory. *Landscape and Urban Planning* 30(1–2): 27–47.

Jaksch S, Franke A, Österreicher D and Treberspurg M (2016) A systematic approach to sustainable urban densification using prefabricated timber-based attic extension modules. *Energy Procedia* 96: 638–649.

Jantz CA, Goetz SJ, Donato D and Claggett P (2010) Designing and implementing a regional urban modeling system using the SLEUTH cellular urban model. *Computers, Environment and Urban Systems* 34(1): 1–16.

Jjumba A and Dragićević S (2012) High resolution urban land-use change modeling: Agent iCity approach. *Applied Spatial Analysis and Policy* 5(4): 291–315.

Jordan R, Birkin M and Evans A (2014) An agent-based model of residential mobility: Assessing the impacts of urban regeneration policy in the EASEL district. *Computers, Environment and Urban Systems* 48: 49–63.

Kamusoko C and Gamba J (2015) Simulating urban growth using a Random Forest-Cellular Automata (RF-CA) model. *ISPRS International Journal of Geo-Information* 4(2): 447–470.

King JL and Kraemer KL (1993) Models, facts, and the policy process: the political ecology of estimated truth. In: Goodchild MF, Parks BO and Steyaert LT (eds) *Environmental Modeling with GIS*. New York: Oxford University Press, 353–360.

Koziatek O and Dragićević S (2017) iCity 3D: A geosimualtion method and tool for three-dimensional modeling of vertical urban development. *Landscape and Urban Planning* 167: 356–367.

Labrinidis A and Jagadish HV (2012) Challenges and opportunities with big data. *Proceedings of the VLDB Endowment* 5(12): 2032–2033.

Li XC and Gong P (2016) Urban growth models: Progress and perspective. *Science Bulletin* 61(21): 1637–1650.

Li X and Yeh AG-O (2002) Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science* 16(4): 323–343.

Liao J, Tang L, Shao G, Su X, Chen D and Xu T (2016) Incorporation of extended neighborhood mechanisms and its impact on urban land-use cellular automata simulations. *Environmental Modelling & Software* 75: 163–175.

Lim C, Kim KJ and Maglio PP (2018) Smart cities with big data: Reference models, challenges, and considerations. *Cities* 82: 86–99.

Lin J, Huang B, Chen M and Huang Z (2014) Modeling urban vertical growth using cellular automata: Guangzhou as a case study. *Applied Geography* 53: 172–186.

Liu X, Liang X, Li X, Xu X, Ou J, Chen Y, Li S, Wang S and Pei F (2017) A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landscape and Urban Planning* 168: 94–116.

Liu Y (2008) *Modelling Urban Development with Geographical Information Systems and Cellular Automata*. New York: CRC Press.

Liu Y (2012) Modelling sustainable urban growth in a rapidly urbanising region using a fuzzy-constrained cellular automata approach. *International Journal of Geographical Information Science* 26(1): 151–167.

Liu Y and Feng YJ (2016) Simulating the impact of economic and environmental strategies on future urban growth scenarios in Ningbo, *China. Sustainability* 8(10): 1045.

Liu Y, Corcoran J and Feng YJ (2019) Cellular automata. In: Kobayashi A (ed.) *International Encyclopedia of Human Geography*, *2nd Edition*. Elsevier.

Liu Y, Kong X, Liu Y and Chen Y (2013) Simulating the conversion of rural settlements to town land based on multi-agent systems and cellular automata. *PLoS One* 8(11): e79300.

Long Y and Wu K (2016) Shrinking cities in a rapidly urbanizing China. *Environment and Planning A: Economy and Space* 48(2): 220–222.

Martinez-Fernandez C, Audirac I, Fol S and Cunningham-Sabot E (2012) Shrinking cities: Urban challenges of globalization. *International Journal of Urban and Regional Research* 36(2): 213–225.

McCall MK (2003) Seeking good governance in participatory-GIS: A review of processes and governance dimensions in applying GIS to participatory spatial planning. *Habitat International* 27(4): 549–573.

Moreno N, Wang F and Marceau DJ (2009) Implementation of a dynamic neighborhood in a land-use vector-based cellular automata model. *Computers, Environment and Urban Systems* 33(1): 44–54.

Musa SI, Hashim M and Reba MNM (2017) A review of geospatial-based urban growth models and modelling initiatives. *Geocarto International* 32(8): 813–833.

O'Sullivan D (2000) *Graph-based Cellular Automaton Models of Urban Spatial Processes*. PhD thesis, Centre for Advanced Spatial Analysis (CASA), University College London.

O'Sullivan D (2002) Toward micro-scale spatial modeling of gentrification. *Journal of Geographical Systems* 4(3): 251–274.

Pacione M (1990) Urban liveability: A review. *Urban Geography* 11(1): 1–30.

Palme M and Ram´ırez J (2013) A critical assessment and projection of urban vertical growth in Antofagasta, Chile. *Sustainability* 5(7): 2840–2855.

Perera C, Ranjan R, Wang L, Khan SU and Zomaya AY (2015) Big data privacy in the internet of things era. *IT Professional* 17(3): 32–39.

Pérez-Molina E, Sliuzas R, Flacke J and Jetten V (2017) Developing a cellular automata model of urban growth to inform spatial policy for flood mitigation: A case study in Kampala, Uganda. *Computers, Environment and Urban Systems* 65: 53–65.

Pinto NN and Antunes AP (2010) A cellular automata model based on irregular cells: Application to small urban areas. *Environment and Planning B: Planning and Design* 37(6): 1095–1114.

Pontius RG Jr and Petrova SH (2010) Assessing a predictive model of land change using uncertain data. *Environmental Modelling & Software* 25(3): 299–309.

Rabari C and Storper M (2014) The digital skin of cities: Urban theory and research in the age of the sensored and metered city, ubiquitous computing and big data. *Cambridge Journal of Regions, Economy and Society* 8(1): 27–42.

Rafiee R, Mahiny AS, Khorasani N, Darvishsefat AA and Danekar A (2009) Simulating urban growth in Mashad City, Iran through the SLEUTH model (UGM). *Cities* 26(1): 19–26.

Research Institute for Knowledge Systems (2013) *Metronamica documentation*. Maastricht, The Netherlands. Available at: http://www.metronamica.nl/ (accessed 9 June 2019).

Rieniets T (2009) Shrinking cities: Causes and effects of urban population losses in the twentieth century. *Nature and Culture* 4(3): 231–254.

Rienow A and Goetzke R (2015) Supporting SLEUTH – enhancing a cellular automaton with support vector machines for urban growth modeling. *Computers, Environment and Urban Systems* 49: 66–81.

Sakieh Y, Amiri BJ, Danekar A, Feghhi J and Dezhkam S (2015) Simulating urban expansion and scenario prediction using a cellular automata urban growth model, SLEUTH, through a case study of Karaj City, Iran. *Journal of Housing and the Built Environment* 30(4): 591–611.

Salvati L and Serra P (2016) Estimating rapidity of change in complex urban systems: A multidimen- sional, local-scale approach. *Geographical Analysis* 48(2): 132–156

Santamouris M, Cartalis C, Synnefa A and Kolokotsa D (2015) On the impact of urban heat island and global warming on the power demand and electricity consumption of buildings: A review. *Energy and Buildings* 98: 119–124.

Sante´ I, Garc´ıa AM, Miranda D and Crecente R (2010) Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and Urban Planning* 96(2): 108–122.

Schwarz N, Haase D and Seppelt R (2010) Omnipresent sprawl? A review of urban simulation models with respect to urban shrinkage. *Environment and Planning B: Planning and Design* 37(2): 265–283.

Semboloni F (2000) The dynamic of an urban cellular automata model in a 3D spatial pattern. *National Conference Aisre, Proceedings, XXI, Regional and Urban Growth in a Global Market*. Palermo, Italy.

Semboloni F, Assfalg J, Armeni S, Gianassi R and Marsoni F (2004) CityDev, an interactive multi-agents urban model on the web. *Computers, Environment and Urban Systems* 28(1–2): 45–64.

Shirley LJ and Battaglia LL (2008) Projecting fine resolution land-cover dynamics for a rapidly changing terrestrial–aquatic transition in Terrebonne Basin, Louisiana, USA. *Journal of Coastal Research* 24(6): 1545–1554.

Silva EA and Clarke KC (2005) Complexity, emergence and cellular urban models: Lessons learned from applying SLEUTH to two Portuguese metropolitan areas. *European Planning Studies* 13(1): 93–115.

Smith N (1987) Gentrification and the rent gap. *Annals of the Association of American Geographers* 77(3): 462–465.

Spiekermann K and Wegener M (2018) Multi-level urban models: Integration across space, time and policies. *Journal of Transport and Land Use* 11(1): 67–81.

Stanilov K and Batty M (2011) Exploring the historical determinants of urban growth patterns through cellular automata. *Transactions in GIS* 15(3): 253–271.

Stevens D, Dragićević S and Rothley K (2007) iCity: A GIS–CA modelling tool for urban planning and decision making. *Environmental Modelling & Software* 22(6): 761–773.

Tan R, Liu Y, Liu Y, He Q, Ming L and Tang S (2014) Urban growth and its determinants across the Wuhan urban agglomeration, central China. *Habitat International* 44: 268–281.

Tang UW and Wang ZS (2007) Influences of urban forms on traffic-induced noise and air pollution: Results from a modelling system. *Environmental Modelling & Software* 22(12): 1750–1764.

Tian G, Ma B, Xu X, Liu X, Xu L, Liu X, Xiao L and Kong L (2016) Simulation of urban expansion and encroachment using cellular automata and multi-agent system model: A case study of Tianjin Metropolitan Region, China. *Ecological Indicators* 70: 439–450.

Tobler WR (1975) Linear operators applied to areal data. In: Davis JC and McCullaugh MJ (eds) *Display and Analysis of Spatial Data*. New York: John Wiley & Sons, 14–37.

Tobler WR (1979) Cellular geography. In: Gale S and Olsson G (eds) *Philosophy in Geography*. Dordrecht: Springer, 379–386.

Torrens PM and Benenson I (2005) Geographic automata systems. *International Journal of Geographical Information Science* 19(4): 385–412.

Torrens PM and Nara A (2007) Modeling gentrification dynamics: A hybrid approach. *Computers, Environment and Urban Systems* 31(3): 337–361.

Torrens PM and O'Sullivan D (2001) Cellular automata and urban simulation: Where do we go from here? *Environment and Planning B: Planning and Design* 28(2): 163–168.

Václavík T and Rogan J (2009) Identifying trends in land use/land cover changes in the context of post-socialist transformation in central Europe: A case study of the greater Olomouc Region, Czech Republic. *GIScience & Remote Sensing* 46(1): 54–76.

Van Vliet J, White R and Dragićević S (2009) Modeling urban growth using a variable grid cellular automaton. *Computers, Environment and Urban Systems* 33(1): 35–43.

Verburg PH and Overmars KP (2007) Dynamic simulation of land-use change trajectories with the Clue-S model. In: Koomen E, Stillwell J, Bakema A and Henk JS (eds) *Modelling Land-Use Change: Progress and Applications*. Dordrecht: Springer, 321–337.

Waddell P (2002) UrbanSim: Modeling urban development for land use, transportation, and environmental planning. *Journal of the American Planning Association* 68(3): 297–314.

Waddell P (2011) Integrated land use and transportation planning and modeling: Addressing challenges in research and practice. *Transport Reviews* 31(2): 209–229.

Wahyudi A, Liu Y and Corcoran J (2019a) Simulating the impact of developers' capital possession on urban development across a megacity: An agent-based approach. *Environment and Planning B: Urban Analytics and City Science*. DOI: https://doi.org/10.1177/2399808319875983.

Wahyudi A, Liu Y and Corcoran J (2019b) Generating different urban land configurations based on heterogeneous decisions of private land developers: An agent-based approach in a developing country context. *ISPRS International Journal of GeoInformation* 8: 229. DOI: 10.3390/ijgi8050229.

White R, Engelen G and Uljee I (2015) *Modeling Cities and Regions as Complex Systems: From Theory to Planning Applications*. Cambridge, MA: MIT Press.

Wiechmann T and Pallagst KM (2012) Urban shrinkage in Germany and the USA: A comparison of transformation patterns and local strategies. *International Journal of Urban and Regional Research* 36(2): 261–280.

Wu F and Webster CJ (2000) Simulating artificial cities in a GIS environment: urban growth under alternative regulation regimes. *International Journal of Geographical Information Science* 14(7): 625–648.

Wu X, Zhu X, Wu GQ and Ding W (2014) Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering* 26(1): 97–107.

Zheng HW, Shen GQ, Wang H and Hong J (2015) Simulating land use change in urban renewal areas: A case study in Hong Kong. *Habitat International* 46: 23–34.

Author biographies

Yan Liu is Associate Professor in Geographical Information Science within the School of Earth and Environmental Sciences at The University of Queensland, Australia. She is a Quantitative Human Geographer with research focus on cities and urban analytics, spatial modelling and geo- simulation, and the application of GIS in spatial planning, policy analysis and spatially integrated social geographical studies. She convenes the Spa- tio-Temporal Analytics Research Laboratory (STAR Lab) at UQ and is the co-editor of the *Jour- nal of Spatial Science.*

Michael Batty is Bartlett Professor of Planning at the Centre for Advanced Spatial Analysis, Univer- sity College London, UK. His research work involves the development of computer models of cities and regions, and he has published many books and articles in this area. His books *Cities and Complexity* (MIT Press, 2005) and *The New Science of Cities* (MIT Press, 2013) both won the Alonso Prize of the Regional Science Association in 2010, and 2017. His most recent book is *Inventing Future Cities* (MIT Press, 2018). He is editor of the journal Environment and Planning B: Urban Analytics and City Science.

Siqin Wang holds a PhD in Human Geography within the School of Earth and Environmental Sciences at The University of Queensland, Australia. Her research interests focus on urban geography, migration, 3D visualisation, spatial analysis and urban modelling. Her publications cover topics including migration and residential mobility, residential pathways, big data application in transport planning, neighbour complaints, and ageing issues in the interdisciplinary field of urban geography and sociology.

Jonathan Corcoran is Professor in Human Geography within the School of Earth and Environmental Sciences at The University of Queensland, Australia, and the director of the Queensland Centre for Population Research. His research interests lie in the fields of population geography, spatial science and regional science. His publications cover topics including human mobility and migration, human capital, and social problems, each of which has a focus on quantitative meth- ods. He is the co-editor of *Australian Population Studies* and Secretary of the Applied Geography Commission.