Agent-based modelling for Urban Analytics: State of the art and challenges

Nick Malleson [a,b,*], Mark Birkin [a,b,c], Daniel Birks [b,d], Jiaqi Ge [a], Alison Heppenstall [b,e], Ed Manley [a,b], Josie McCulloch [a,b] and Patricia Ternes [f]

[a] *School of Geography, University of Leeds, Leeds, UK*
E-mail: n.s.malleson@leeds.ac.uk
[b] *Leeds Institute for Data Analytics, University of Leeds, Leeds, UK*
[c] *Alan Turing Institute, London, UK*
[d] *School of Law, University of Leeds, Leeds, UK*
[e] *School of Social and Political Sciences; MRC/CSO Social and Public Health Sciences Unit, University of Glasgow, Glasgow, UK*
[f] *Research Computing, University of Leeds, Leeds, UK*

Abstract. Agent-based modelling (ABM) is a facet of wider Multi-Agent Systems (MAS) research that explores the collective behaviour of individual 'agents', and the implications that their behaviour and interactions have for wider systemic behaviour. The method has been shown to hold considerable value in exploring and understanding human societies, but is still largely confined to use in academia. This is particularly evident in the field of *Urban Analytics*; one that is characterised by the use of new forms of data in combination with computational approaches to gain insight into urban processes. In Urban Analytics, ABM is gaining popularity as a valuable method for understanding the low-level interactions that ultimately drive cities, but as yet is rarely used by stakeholders (planners, governments, etc.) to address real policy problems. This paper presents the state-of-the-art in the application of ABM at the interface of MAS and Urban Analytics by a group of ABM researchers who are affiliated with the Urban Analytics programme of the Alan Turing Institute in London (UK). It addresses issues around modelling behaviour, the use of new forms of data, the calibration of models under high uncertainty, real-time modelling, the use of AI techniques, large-scale models, and the implications for modelling policy. The discussion also contextualises current research in wider debates around Data Science, Artificial Intelligence, and MAS more broadly.

Keywords: Multi-Agent Systems (MAS) research, Agent-Based Modelling (ABM), Urban Analytics

1. Introduction

Agent-based modelling (ABM) is a subset of the broader field of Multi Agent Systems (MAS), although the terms are often used interchangeably. The characteristics of an agent-based model – such as the use of "interacting, autonomous 'agents'" [50] or the general research 'mindset' that describes a system "from the perspective of its constituent units" [12] – could be equally applied to MAS. But they differ in that ABM research tends to focus more on the behaviour of living things (humans, animals, cells, etc.) which is not necessarily the case in MAS research. The methodology has proven particularly amenable to modelling human systems with an individual agent often representing a human. The features of the individual (human) agents, including autonomy, heterogeneity, goal-directed behaviour, mobility, reactivity, etc. [14,16], allow models to be created that more closely approximate the overall system behaviour than traditional aggregate approaches as they capture the behaviours of the key individual units directly [6,26,70]. This is simultaneously the most appealing and challenging aspect of ABM.

---

*Corresponding author. E-mail: n.s.malleson@leeds.ac.uk.

One area where ABM is gaining particular interest is in the field of *Urban Analytics*, defined as "the practice of using new forms of data in combination with computational approaches to gain insight into urban processes" [74]. The use of data and computational approaches to study cities is not new, but in recent years the quantity and variety of data that have become available have created new opportunities for research that requires new techniques and approaches. These new data sources often span multiple spatial and temporal scales – from individual actions such as using public transport or paying for goods, through to macro-level outcomes such as gentrification or population health – and this, coupled with the fact that cities are inherently complex systems that are driven by the actions of micro-units [5,6], puts ABM in a unique position to model urban systems.

This paper showcases current agent-based modelling research that is being conducted by researchers in the *Urban Analytics* programme (see Section 2) of the Alan Turing Institute, the UK's national centre for AI and data science. It charts the path taken by researchers as they move from small, abstract models of social systems through to all-encompassing, large-scale models that border the realm of *digital twins* (see Section 8).

The paper begins by describing the 'group' more concisely in Section 2, before moving on to discuss:

- the development of techniques used to capture human behaviour in software agents (Section 3);
- new forms of data that underpin the field Urban Analytics and their impact on research (Section 4);
- fundamental work on the validation of agent-based models of human systems (Section 5);
- efforts to stream real-time data into agent-based models (Section 6);
- work that incorporates AI innovations in ABM (Section 7);
- progress towards representative, large-scale models of urban systems (Section 8);
- some of the successful collaborations that have led to these models being used in policy, as well as the barriers that are preventing them from being more widely adopted (Section 9).

The paper concludes with a summary of the future outlook and challenges (Section 10).

Before continuing, one further definition is necessary. The technique of 'microsimulation' is similar to ABM in its focus on individual units ('agents'), yet in a microsimulation model the agents are less likely to interact and may have less nuanced behaviour. There are also differences in the research questions that they are typically used to answer. Microsimulation is often used to generate synthetic populations that can be used to understand the impact of a policy on individuals within a population over time [76,93]. These synthetic populations can subsequently be used in 'dynamic' microsimulations, or in agent-based models. That said, there is no clear delineation between the two approaches, and, as with 'ABM' and 'MAS', the terms are sometimes used interchangeably. In the discussion below we include some research that might be better classified as 'microsimulation' rather than 'ABM', but it is important that these studies are included as the two methodologies are closely linked anyway, and microsimulation research is having a significant impact on the research agenda more generally.

2. Agent-based modelling and Urban Analytics at the Alan Turing Institute

The UK government's ambition to implement a "ten-year plan to make Britain a global AI superpower" [25, p. 8], envisages collaborations between "globally recognised institutes such as The Alan Turing Institute and the high-performing universities which are core to research in AI" [25, p. 16]. It is clear that in the eyes of the UK government, that the Turing has quickly become established as the UK's national institute for data science and artificial intelligence. Opportunities include: a better understanding of the interactions between health inequalities and environmental factors, such as air quality, and how these new understanding can be used to design more targeted and informed interventions [32, p. 24]; how data can be used better understand crime patterns and to guide interventions that reduce crime in a way that is inclusive and unbiased [24, mission 1]; and intelligent mobility solutions that encourage more sustainable and active travel, reducing the need for parking and thereby freeing up land for new homes or green spaces [34, p. 30]. In addressing these challenges, the UK has access to world-class assets through its academic institutions, research council investments, smart city solutions and a diverse ecosystem of business innovation. Furthermore, advanced city modelling and work towards 'digital twins' (discussed in Section 8) can create a platform to capitalise on international opportunities in which the market for intelligent mobility solutions alone is expected to reach $1.4 trillion by 2030 [34, p. 33].

These opportunities will prove the most fruitful in the context of studying cities because they offer the most fertile domains for "the use of AI to address societal, economic and environmental challenges in the UK" [25, p. 14]. This recognition led to the creation of the Urban Analytics programme at the Turing in 2018 to provide a focus for sophisticated analytics of big data in metropolitan areas for improved understanding and policy making. The missions of the programme are oriented towards the realisation of healthier and more liveable cities, in which technology supports enhanced mobility, sustainability and equality of opportunity. The remainder of this paper outlines work by the agent-based modelling researchers who are members of the Urban Analytics group (hereafter 'the group').

3. Modelling human behaviour in agent-based models

One of the most appealing aspects of agent-based modelling (ABM) is its ability to represent human behaviour and, through simulation, understand how these behaviours play out over space and time. Our group has made several advances within this area. [38] is an early example of how behavioural rules for agents (in this case petrol stations) were devised using a combination of approaches including numerical analysis, genetic algorithms and published literature. The behavioural rules were simple and driven by individual petrol stations maximising profit. Whilst this was appropriate for these agents, replicating the behaviour of human agents requires a more sophisticated approach. Members of the group adopted behavioural frameworks, such as the PECS framework, relatively early [51,52] – see [71] and [62] for a detailed overview of ABM behavioural frameworks – to allow a more nuanced set of behaviours (in this case, the daily behaviours and motivations of burglars) to be captured and successfully simulated. More recent efforts have included the use of approaches such as Reinforcement Learning, where positive behaviours are learned through a repeated exposure to an environment. For example, [64] use reinforcement learning in the context of a simple predator-prey model to explore the extent to which the approach can be used to create emergent agent behaviours. Similarly, it may be possible to teach agents how to navigate spaces as if they were human. These "spatial learning" agents may both better reflect the actual behaviours of humans and model their behaviour under changing conditions. Only a handful of simulation models have sought to integrate aspects of spatial cognition and bounded learning [55,58].

Taking a somewhat different approach, work in the group has also explored how ABM can be effectively used to test hypotheses about typically unobserved behaviours such as those associated with crime and offending [9–11]. Given fundamental logistical and ethical challenges associated with studying such behaviours, these models allow researchers unique means to test the veracity of behavioural theories in generating known empirical system outcomes like commonly observed patterns of crime, or provide counterfactual realities against which potential crime prevention interventions can be effectively compared.

The growth in data availability is one of the main foundations for Urban Analytics (see Section 4). Higher resolution data can contain information about micro-level behaviours (e.g. the details about interactions between individuals) and, as new forms of data have appeared, researchers have been able to move from relying purely on literature and theory towards using data to both create and validate behavioural rule sets. For example the pedestrian model of [20] (discussed in Section 5) used survey data, sensor data and focus groups to refine behaviour. These data are, however, limited by the events that they do not contain; how can we simulate behaviours that we are interested in if they are not present in the data? This is a further area where ABM might hold promise; we could use an agent-based model to infer the characteristics of an unobserved behaviour by calibrating it to data on known behaviours [e.g. 35,88]. This is the aim of the new field of Inverse Generative Social Science [89].

Overall, the combination of new, high resolution data coupled with advanced approaches for modelling behaviour – whether through the use of behavioural frameworks, reinforcement learning, or others – is allowing researchers to build behaviours into agent-based models that can be scaled up from models of hundreds of agents to thousands or millions. Section 8 discusses work that moves towards these much larger models.

4. New forms of data

The Urban Analytics field is characterised by the burgeoning availability of novel data sets. These provide opportunities for the creation of new agent-based models (ABMs) in settings that previously would not have been possible, or for more effective calibration and validation to improve their rigour (discussed in more detail in Section 5). For example: mobile telephone records can offer more detailed cross-sectional or longitudinal movement histories [72]; smart ticketing could inform behavioural sensitivities to price or service quality in public transport; smart-phone apps could be used to infer routing behaviour [54]; social media or financial transactions could give broader insight into the motivation for travel decisions. The advantages, and difficulties, in accessing such data have been recognised by learned societies of the major global economies in specific response to the COVID-19 pandemic [63], but continuing obstacles include licensing restrictions and the maintenance of high standards of privacy and data protection. Furthermore, such data may be readily available in major metropolitan areas such as New York or Singapore, but may not exist, or be inaccessible, in provincial or less industrialised regions.

Mobility data – i.e. data that describe the movements of people – are a particularly pertinent example of the use of 'novel' urban analytics data sets that are being used by researchers in our group to inform their ABMs. For example, [56] use origin-destination trip data provided by Transport for London to create a 'hybrid' agent-based model that represents urban traffic flow driven by complex behavioural models and yet balances this with computational capacity to allow the model to be scaled to large areas with large numbers of agents. In related work, [55] use similar transport data to explore how drivers' behaviour and cognition, with respect to route choice, impact the spatial extent and volume of traffic flow within a real-world setting. The results suggest that greater attention should be paid to individual-level models of spatial cognition as these play an important role in predicting urban traffic flow.

The use of mobility data also extends beyond the study of the behaviour of drivers. New pedestrian count (aka 'footfall') data [e.g. see 92] can provide insight into the dynamics of cities, and are being used to build agent-based models of wider urban dynamics. For example, the work of [20] (discussed in Section 5) uses counts of pedestrians to inform the types of agents that are included in their model and to simulate the daily movement dynamics of people in a small town. At a finer spatio-temporal scale, it is also possible to use pedestrian mobility data to simulate the dynamics of crows of people. This is becoming ever more important as global urbanisation trends [83] lead to larger and denser cities. For example, [80] use the trajectories of individual people to create agent-based models that could ultimately be used to manage the movements of crowds in busy urban spaces in real-time (discussed in detail in Section 6).

In summary, there are obvious opportunities offered by these new forms of data. They provide the potential for much more rigorous model calibration or validation (Section 5) and will be key to building larger-scale 'digital twin' style models that can support policy making (Sections 8 and 9). However, there are challenges associated with these 'big' data sets that have been well known for some time [e.g. 44] and care must of course be taken to ensure that the natural biases and other issues around representation are taken into account if they are to be used in modelling.

5. Calibration, validation and uncertainty

Despite ongoing efforts to improve the empirical rigour of ABMs through the use of robust calibration and validation methods [e.g 81], these have remained one of the 'key challenges' for the discipline for a number of years [2,21,37,59]. This is not surprising for two reasons in particular. Firstly, many evaluation approaches require the underlying model to be executed a large number of times – a million runs is not uncommon [e.g. 86] – which is problematic for computationally expensive ABMs. Secondly, the complex, multi-scale nature of the systems that ABMs usually study means that validating a single 'pattern' at a single spatial/temporal scale (as would be typical with an aggregate model) would not be sufficient in an agent-based model where multiple patterns at different scales interact [36], as is the case for most urban phenomena. The degree of uncertainty at one scale is an insufficient measure of the overall uncertainty [36].

Fortunately the emergence of new urban data sets, as discussed in Section 4, offers opportunities for more nuanced calibration and validation of ABMs. As an example, [20] use an agent-based model to study people's daily
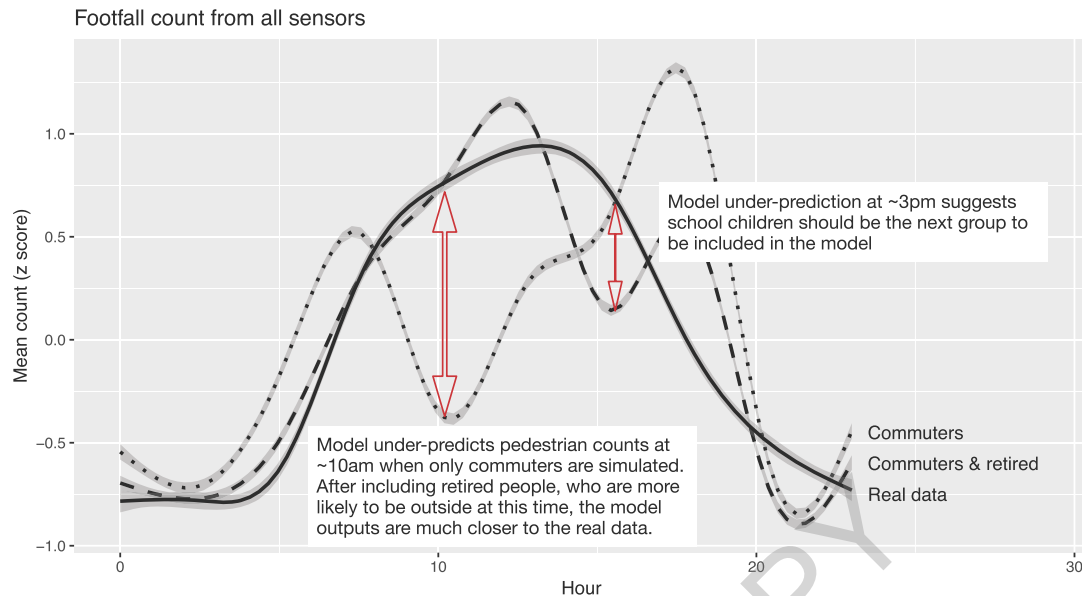
Footfall count from all sensors

Fig. 1. Comparing average pedestrian counts in two ABMs (one with just commuters, the other with commuters and retired people) to real pedestrian counts recorded using Wi-Fi counters. Adapted from [20].

movements in the town of Otley, UK. In the study area, sensors had been installed to count the number of mobile phones detected in their vicinity over a short time window. Individual phones were not tracked, so the data were anonymous, but the pedestrian counts themselves can give insight into patterns of mobility in the town, showing how many people are in different parts of the town at a high spatio-temporal resolution. [20] created a model that attempted to simulate the daily activities of all individuals in the area, using data from the 2011 census, as they undertook the most common activities (working, shopping, schooling, etc.). The model outputs were compared to the raw data to both calibrate the model parameters but also to determine whether categories of individuals were absent from the model. As illustrated in Fig. 1, the model significantly under-predicted the pedestrian count when only commuter agents were accounted for. The inclusion of retired people brought the model outcomes closer to the real data, and it was expected that including school age children would also improve the accuracy of the output.

The group are also conducting methodological work that leverages techniques from related fields to support the evaluation of ABMs. Efforts to reduce uncertainty in real-time (discussed in Section 6) will help in some situations, but not necessarily for more general calibration/validation challenges. For more general applications, efforts to adapt techniques from the field of Uncertainty Quantification are proving useful. Recent work by [61] proposes a framework for better understanding the uncertainties associated with an agent-based model and to take these into account when calibrating a model. The authors propose the use of History Matching [19] to rule out *implausible* models and thus reduce the size of the parameter space that needs to be searched during calibration. Calibration is achieved using Approximate Bayesian Computation (ABC) to provide credible intervals over which the given parameters could have created the observed data [22,82]. The framework therefore allows for the explicit identification and quantification of the various sources of uncertainty in the model, and these are taken into account during calibration or validation.

Related work is also exploring the potential of probabilistic modelling [23,33]; an approach to model building that provides an elegant framework for representing all forms of uncertainty, updating beliefs about the target system in response to new data, and making predictions. A *fully-probabilistic* agent-based model, where the agents' variables are represented with probability distributions (i.e. random variables) rather than deterministic variables as is usual, could transform the way that uncertainty is captured and modelled. Initial research has attempted to use bespoke probabilistic programming libraries to simulate a crowd of people in a toy model [3] and novel attempts are also underway to develop probabilistic agent-based frameworks that allow for data assimilation [78] and the calculation

of the maximum a-posteriori probability (the mode of a posterior distribution) [79]; all of which are useful advances in the development of probabilistic ABM.

With respect to computational efficiency, often there is little that can be done to speed up the runtime of an agent-based model. Ultimately the models require large numbers of agents to make decisions and to interact, and these processes are computationally expensive. That said, recent work by the group involved re-implementing a large-scale COVID microsimulation [76] to reduce its computational complexity using distributed computation technologies such as OpenGL. The gains were astonishing; the time required to execute the model on a desktop PC went from 2+ hours in the original python implementation to under 5 seconds. However, it is worth noting that this was possible because the interactions between individuals were limited, so the improvements are unlikely to be so dramatic when applied to a more traditional agent-based model. In these cases, work on emulators of ABMs might prove useful; for example see [27] who adapt a Quantile Gaussian Process to emulate their model of disease spread. Ongoing, as yet unpublished, work is exploring the opportunities for emulating agent-based models using existing frameworks developed at the Alan Turing Institute (e.g. see [13]).

6. Real-time ABM

Despite important innovations in methods to calibrate and validate ABMs (see Section 5), there remains a fundamental methodological barrier that agent-based modelling has yet to overcome; there are no established mechanisms for incorporating data into ABMs in *real time* [47,90,91]. Models are typically calibrated using historical data and then used to make predictions. This is entirely appropriate in many circumstances, such as where systems are in equilibrium, but there are cases where the calibrated model and the underlying system will diverge as they evolve independently. The dynamic re-calibration of an agent-based model's parameter values in response to new data is not uncommon [e.g. see 65,66], but adjusting parameter values does not help if a running model has diverged in real-time. In those cases the model *state* needs updating. Fortunately, a body of work has developed in fields that require the integration of real-time data to make more accurate short-term predictions, such as in meteorology [43], under the banner of 'data assimilation'. Broadly, data assimilation is starting to be recognised as an important barrier and opportunity for ABMs [2,37], and some effort in the group has been devoted to solving data assimilation challenges for ABM. Ultimately, if successful, data assimilation will allow ABMs to make better use of many of the 'novel' data sets that are becoming available, adapting in real-time as the social systems that they are tasked with simulating evolve (see Fig. 2).

Inspired by efforts that demonstrated data assimilation is possible for relatively simple ABMs [91], the majority of the work has focused on the technical challenges of adapting existing data assimilation methods in the context of simulating crowds of people. Crowding was chosen as a test case because crowds exhibit complex behaviour and are driven by the behaviours of individual people, so are an ideal target for simulation with ABM, and because high quality data are readily available for algorithm testing and validation (e.g. [94]). Initial work [53] focused on the use of Particle Filters as these are non-parametric so well suited to systems that have non-linear and non-Gaussian behaviour [49]. However, a drawback with the Particle Filter is that the number of individual model instances ('particles') that are required to represent the vast space of possible model trajectories grows exponentially with the size and complexity of the system (i.e. the number of agents and the number of agent-level parameters). Many thousands or millions of particles may be required to successfully represent even relatively small crowds. Whilst there are improvements that can be made to the vanilla Particle Filter to improve its efficiency, alternative methods such as the Unscented Kalman Filter [17] are also being investigated. In addition, work into an entirely new means of defining agent-based models through the use of operators derived from quantum field theory – which is a field ideally suited to representing inherently uncertain systems – may prove transformational [78].

A related challenge that the group are addressing, that came to light as they moved from toy to empirical models [80], is that data assimilation algorithms cannot, by default, optimise *categorical* variables. This type of variable is extremely common in ABM. Therefore attempts are underway to include categorical variables through innovations such as the Reversible Jump, which may allow multiple filters to try to find the 'correct' categorical parameter values [18], or by resampling only some model variables [80]. As Section 10 will discuss, the ultimate aim of these efforts is to allow for real-time agent-based models of dynamic urban systems, drawing on diverse data sources and making accurate short-term predictions to aid policy making.
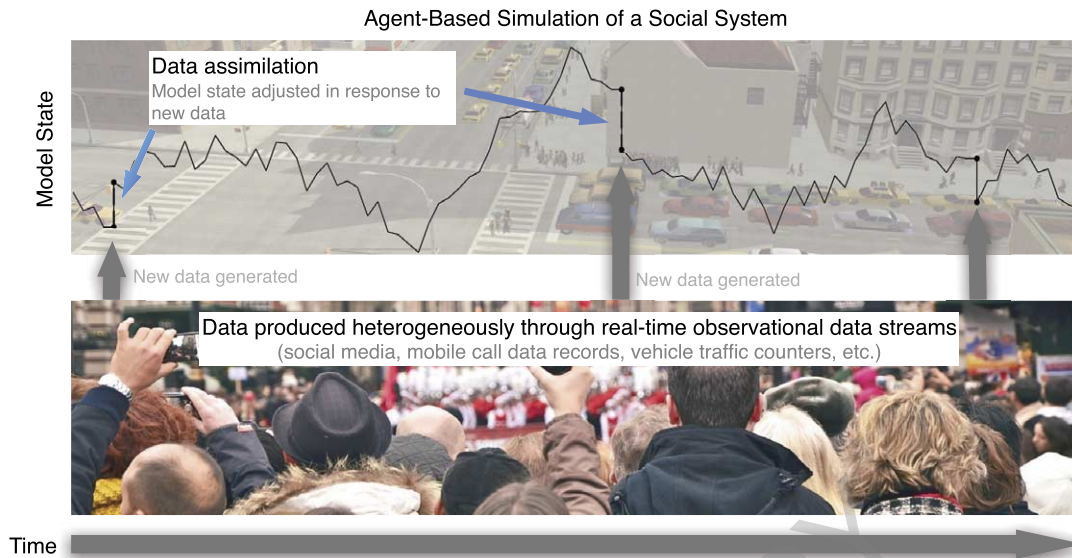
Agent-Based Simulation of a Social System

Fig. 2. An example of a 'real-time' agent based model. The model simulates the state of an urban system in real time and draws on up-to-date data to update its state as the urban system evolves.

7. AI in ABM

The emergence of novel approaches in AI presents ABM research with new opportunities for modelling agent behaviour (as discussed in Section 3). These are particularly pertinent for the study of cities where, as well offering opportunities, AI raises new questions as computational agents mix with humans in creating urban dynamics.

The use of AI approaches in modelling agent decisions has become established over the last decade. Statistical classifiers, such as logistic regression, have been a commonly used approach for simulating decisions [45], while computational classification methods, such as neural networks, have been used in simulating vehicle movements [15] and segregation dynamics [39], among other scenarios. As 'deep learning' approaches have risen in popularity, these methods have enabled even more data to be included in modelling agent decisions [85]. Recent developments on modelling trajectories reflects the potential for these approaches within ABM [40]. While Deep Reinforcement Learning (DRL), in part due to its implementation within the Unity ML-Agents library [42], has extended to integration of expansive 3D visual inputs within agent decisions. DRL represent a particularly promising avenue for modelling agent behaviours in unseen scenarios and environments, but there is a risk of intractability in agent behaviour. Regardless of the methodology, the increased availability of data, combined with exponentially increasing computational power, enables a greater array of features to be combined in predicting individual behaviours. Nevertheless, questions remain around the trade-offs between the specificity of the behavioural models, simulation scale, computational resources, and the role of theory [57] in recreating and predicting future urban dynamics.

Within the swathe of ABMs of cities, one AI technique that has seen little exploitation is Multi-Agent Reinforcement Learning (MARL). Extending the RL framework to multiple agents enables recreation of mutual and social learning, norm formation, and cooperative behaviours, which have implications across various urban problems. However, the technique suffers from challenges relating to the non-stationarity in other agents and environments, which remains an open research challenge in AI more generally [67]. This area represents one for future exploration and is an ideal candidate for research at the Alan Turing Institute, due to the interdisciplinary nature of the organisation.

Perhaps the strongest contribution ABM can make in relation to AI (as opposed to vice versa) is around understanding the potential future implications of AI for cities. AI is increasingly intertwined with the city. The roll-out of autonomous vehicles (AVs) on city streets is already underway; autonomous lorries, buses, and even aviation vehicles are anticipated over the coming years. While digital technologies have a large impact on the production of urban dynamics (e.g. commuting patterns, shopping habits) that until recently have seemed stable. There is potential

for these technologies to prompt a reconsideration of how we design public spaces, roads [69], housing and land use, and beyond. These technologies have potential implications for social outcomes too, which must be carefully considered and anticipated. The extent to which these technologies will change society remains an open question, and ABM is an important method for exploring these scenarios [60]. At present, the group is undertaking projects that use ABM to explore the potential impacts of AVs on pedestrian dynamics, on elderly mobility, and uptake by different demographic segments of society. Each of these scenarios have important implications for reinforcing or counteracting social inequalities.

8. Large scale ABMs and 'digital twins'

This paper has outlined various efforts that our group is engaging in – such as improving models of agent behaviour, using novel data, advancing calibration and validation methods, incorporating innovations in AI, etc. – to improve the methodological rigour of ABM and to usefully apply ABMs in policy-relevant contexts. If these innovations are successful then they pave the way for the use of ABM as part of much broader computational tools that can simulate human systems over relatively large spatial areas and time periods. This section will outline some of the ways that the group have been exploring the development of larger ABMs before discussing policy-relevant work specifically in Section 9.

The simulation of transportation systems, in particular, has been shown to be possible at increasingly large scales. Using methods similar to those presented in [56], frameworks such as MATSim[1] enable millions of mobile agents to be constructed and simulated to predict mobility dynamics up to country-level scales. Within our group, the CoSMoNorth project is building a simulation of the north of England, an area covering millions of people, for policy testing at scale. On top of the obvious technical/computational challenges associated with such a large simulation, we also hope to develop methods to allow the model to be regularly updated with relevant data, possibly in real time (i.e. through data assimilation as discussed in Section 6).

The group also develops spatial ABMs and microsimulations on long temporal scales (10–50 years). These models allow us to address different types of research questions such as the impact of a socio-economic shocks on an urban area and its neighbourhoods [28,29], or the impact of Brexit and changing international relations on food security of cities [30,31]. These larger-scale transitions and the mid- to long-term ramifications can be hard to pick up in fine-scale data of everyday activities, and require different types of data such as regional or national demographic and economic data, as well as data on migration and social networks. They may also require the researchers to look at different types of behaviour and decision making. For example, the models may look at decisions to buy houses, change jobs, or make investments for retirement, which requires different and perhaps more complex decision making models than those that are appropriate for everyday decisions. There have also been attempts at integrating ABMs at *different* scales to combine multiple systems such as land-use, residential and transport systems [68], using frameworks such as SimMobility [1]. Behaviours that occur at different temporal scales (e.g. daily transport mode choice and annual residential choice) can be integrated via the same agents making the decisions.

Given the range of distinct efforts towards building large-scale models, the Urban Analytics programme has begun to invest in a common platform that combines the simulation of demographic and socio-economic profiles of individuals and households [48] with detailed activity patterns across small geographical areas [7]. The integration of agent-based models with both microsimulation [e.g. 48] and spatial interaction models [e.g. 7] is not entirely novel, having been explored previously e.g. in the context of urban housing markets [41], local labour markets [4], and retail behaviour [77]. However the Turing is now supporting for the first time the integration of microsimulation, spatial interaction and agent-based models at scale within the framework of a 'digital twin' of a city with real policy challenges and mechanisms for intervention [76].

A major challenge with integrating distinct models into a combined framework is the increasing complexity and potential loss of transparency in the integrated models. Individual-level models are already complex, so there is the danger of creating an "integronster" [87] when the coupling of different modules leads to internal inconsistency

---

[1]https://matsim.org/

(e.g. inconsistent data frequency) that may be difficult to reconcile or even identify. In addition, there is a consensus in some fields (such as economics) that medium-term predictions are the most challenging because short-term predictions are unlikely to be radically different from the current system state and in the long-term systems reach equilibrium. This will be a particular challenge when we try to integrate short- and long-term models to create an overarching 'digital twin'.

9. Towards ABMs as policy tools

Cities are famously now home to more than two-thirds of population of the world [83], and as such engines of prosperity, opportunity and social progress. They also provide a canvas for environmental externalities (e.g. pollution, congestion, noise), deprivation, health inequalities and crime. Such outcomes – both positive and negative – are the product of complex social and economic interactions. This is one key challenges that modellers must face in supporting policy; models – like the systems they seek to emulate – are 'messy' and typically do not produce simple predictions. Predictions are often highly uncertain and it is crucial that these uncertainties are communicated properly [75,84], but this uncertainty in outcomes can be challenging for practitioners. Conversely, by representing individuals directly, ABMs can be very accessible and allow practitioners to inform and critique in ways that some other analytical models do not. Hence the demand for comprehensive and effective city modelling tools is escalating within national and local governments, and commercial organisations (ranging from infrastructure providers to retailers), in the UK and internationally [8,73].

In a recent and ongoing project, our group has worked with partners in the Alan Turing Institute, Urban Observatory Network, Connected Places Catapult (now home to the UK Digital Twin Hub) and the Department for Transport to establish a prototype of a digital twin for mobility in UK cities. Drawing on real-time data feeds from three UK cities (Birmingham, Manchester and Newcastle), this demonstrator has already delivered capabilities for decision-support and policy evaluation relating to AI-based traffic simulation, real-time monitoring and forecasting of black carbon, and configuration of low traffic neighbourhoods (LTNs). For example, in relation to LTNs, travel demand algorithms can integrate historical commuting data with real-time mobility flows to provide a 360° view of movement patterns by purpose which are sensitive to diurnal rhythms with fluctuating regional and seasonal cycles. Onto this base of travel demand and behaviour, microscopic simulations can then predict and evaluate the impact of traffic reduction measures such as traffic calming, pedestrianisation or access restrictions. In ongoing and future work, the ambition is to provide widespread access to models for any local authority or government department seeking better intelligence and augmented intelligence for enhanced design of urban mobility, infrastructure and land-use.

Recent research by the group funded under the Turing's AI for Science and Government programme is developing ABMs capable of supporting UK Police in better understanding the dynamics of resourcing associated with responding to and investigating crime. Combining spatial micro-simulations that draw on a combination of open source and protected crime data, realistic scenarios of crime-related police demand in any of England and Wales's 43 police force areas can be generated. An agent-based model of police resource allocation that encodes expert insights from operational police officers will allow police to explore how changes in the nature and distribution of crime may affect police agencies ability to respond, and in turn, evaluate the effectiveness of potential new resourcing strategies. Ultimately, these models seek to improve decision-making and support better understanding of complex resourcing problems faced by police [46].

Co-production of modelling systems for augmented decision-making will be another critical success factor for agent-based policy models of the future. This will require collaboration across disciplines and between academic institutions, but most importantly it will rest on the engagement of organisations from business, the third sector and especially government, with the authority and ambition to exploit next generation technologies. Further methodological development can be expected to yield improvements in handling uncertainty; propagation of errors through complex or linked systems; assimilation of data which are varied, fast-moving and voluminous; and improved representation and parametrisation of human decision-making. Advances of this kind – many of which our group are currently working on, as outlined throughout this paper – can only help to boost the appeal of models in a policy context. Crucially, the alignment to a broader programme of innovation in AI and Data Science within the Turing

will provide a platform for addressing key challenges for example in controlling uncertainty and the propagation of errors, assimilation of massive data sets and machine learning of agent behaviours which are vital in bridging the remaining gap between algorithmic sophistication and augmentation of policy response to real world problems.

10. Outlook and challenges

This paper has outlined the historical and ongoing research priorities of ABM researchers in the Urban Analytics group of the Alan Turing Institute. It discussed the challenges associated with modelling human behaviour in software agents; outlined how new forms of data that characterise the field of Urban Analytics more broadly are being used to improve the group's ABMs; presented fundamental work on calibration and validation of ABMs, particularly under high levels of uncertainty; outlined preliminary work towards *real-time* agent based models of social systems; presented progress towards large-scale models; and, finally, discussed efforts to use ABMs as key policy tools.

The outlook for ABM in this area is undoubtedly bright. Only a decade ago models were largely constrained to simulating relatively narrow scenarios – i.e. over small spatial areas, short time periods, or restricted to a narrow set of behaviours – but recent work is delivering models that can operate in much larger contexts. They are being used to simulate systems as diverse as transport [56], crime [10,51] and farming [30,31], and work towards 'digital twins' at the Turing aspires to include ABMs alongside other modules that are already working together [76]. In the long run, efforts to better track uncertainty in ABMs [61], coupled with closer collaboration with policy makers through the Turing, will begin to foster the use of ABMs as important policy-making tools.

Despite this optimism, it is important to recognise that ABM in general faces many of the hurdles that were identified a decade ago [2,21,37] and work must aim to address these. The main challenges, as discussed throughout the paper, include issues of scale (how models that work at different spatio-temporal scales should be aligned), model rigour (including how to understand and communicate uncertainty in models and their outputs), computational complexity (as ever, ABMs continue to be extremely computationally expensive) and considerations around ethics and data privacy as models become closer proxies of the real systems that they are simulating.

As members of a research group, the authors have promoted awareness and relevant capabilities both as individuals and as a team, but also importantly through collaboration within the Urban Analytics programme. The Turing is ideally placed to lobby for improved access to data; to promote collaborations between disciplines and exploit channels for engagement in government and business; and to accelerate the development of methods and their translation into practice. The future for agent-based models to support urban analytics and the associated policy challenges is bright, and as a national programme in AI and data science the potential is vast.

Acknowledgements

References

[1] M. Adnan, F.C. Pereira, C.M.L. Azevedo, K. Basak, M. Lovric, S. Raveau, Y. Zhu, J. Ferreira, C. Zegras and M. Ben-Akiva, Simmobility: A multi-scale integrated agent-based simulation platform, in: *95th Annual Meeting of the Transportation Research Board Forthcoming in Transportation Research Record*, 2016.

[2] L. An, V. Grimm, A. Sullivan, B.L. Turner II., N. Malleson, A. Heppenstall, C. Vincenot, D. Robinson, X. Ye, J. Liu, E. Lindkvist and W. Tang, Challenges, tasks, and opportunities in modeling agent-based complex systems, *Ecological Modelling* **457** (2021), 109685. doi:10.1016/j.ecolmodel.2021.109685.

[3] L. Archer, N. Malleson and J. Ward, Probabilistic Programming for Dynamic Data Assimilation on an Agent-Based Model, 2019.

[4] D. Ballas and G. Clarke, GIS and microsimulation for local labour market analysis, *Computers, Environment and Urban Systems* **24**(4) (2000), 305–330. doi:10.1016/S0198-9715(99)00051-4.

[5] M. Batty, *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*, The MIT Press, 2007. ISBN 0-262-52479-1 978-0-262-52479-7.

[6] M. Batty, Building a science of cities, *Cities* **29**(1) (2012), S9–S16. doi:10.1016/j.cities.2011.11.008.

[7] M. Batty and R. Milton, A new framework for very large-scale urban modelling, *Urban Studies* **58**(15) (2021), 3071–3094. doi:10.1177/0042098020982252.

[8] M. Birkin, G. Clarke, J. Corcoran and R. Stimson (eds), *Big Data Applications in Geography and Planning: An Essential Companion*, Edward Elgar Publishing, Northampton, 2021. ISBN 978-1-78990-978-4.

[9] D. Birks and T. Davies, Street network structure & crime risk: An agent-based investigation of the encounter and enclosure hypotheses, *Criminology* **55**(4) (2017), 900–937. doi:10.1111/1745-9125.12163.

[10] D. Birks, M. Townsley and A. Stewart, Generative explanations of crime: Using simulation to test criminological theory, *Criminology* **50**(1) (2012), 221–254. doi:10.1111/j.1745-9125.2011.00258.x.

[11] D. Birks, M. Townsley and A. Stewart, Emergent regularities of interpersonal victimization: An agent-based investigation, *Journal of Research in Crime and Delinquency* **5**(1) (2014), 119–140. doi:10.1177/0022427813487353.

[12] E. Bonabeau, Agent based modeling: Methods and techniques for simulating human systems, *Proceedings of the National Academy of Sciences* **99**(90003) (2002), 7280–7287. doi:10.1073/pnas.082080899.

[13] L. Bowler and N. Malleson, Urban-Analytics/Stochastic-Gp: Version 1, 2022. doi:10.5281/ZENODO.6475056.

[14] C. Castle and A. Crooks, Principles and Concepts of Agent-Based Modelling for Developing Geospatial Simulations, Series, Centre for Advanced Spatial Analysis, London, 2006, UCL Working Paper.

[15] L. Chong, M.M. Abbas and A. Medina, Simulation of driver behavior with agent-based back-propagation neural network, *Transportation Research Record* **2249**(1) (2011), 44–51. doi:10.3141/2249-07.

[16] C. Cioffi-Revilla, *Introduction to Computational Social Science, Texts in Computer Science*, Springer, London, 2014. ISBN 978-1-4471-5661-1 978-1-4471-5660-4.

[17] R. Clay, L.-M. Kieu, J.A. Ward, A. Heppenstall and N. Malleson, Towards real-time crowd simulation under uncertainty using an agent-based model and an unscented Kalman filter, in: *Advances in Practical Applications of Agents, Multi-Agent Systems, and Trustworthiness*, Y. Demazeau, T. Holvoet, J.M. Corchado and S. Costantini, eds, The PAAMS Collection, Vol. 12092, Springer International Publishing, Cham, 2020, pp. 68–79. ISBN 978-3-030-49777-4 978-3-030-49778-1. doi:10.1007/978-3-030-49778-1_6.

[18] R. Clay, J.A. Ward, P. Ternes, L.-M. Kieu and N. Malleson, Real-time agent-based crowd simulation with the reversible jump unscented Kalman filter, *Simulation Modelling Practice and Theory* **113** (2021), 102386. doi:10.1016/j.simpat.2021.102386.

[19] P.S. Craig, M. Goldstein, A.H. Seheult and J.A. Smith, Pressure matching for hydrocarbon reservoirs: A case study in the use of Bayes linear strategies for large computer experiments, in: *Case Studies in Bayesian Statistics*, Springer, 1997, pp. 37–93. doi:10.1007/978-1-4612-2290-3_2.

[20] T. Crols and N. Malleson, Quantifying the ambient population using hourly population footfall data and an agent-based model of daily mobility, *GeoInformatica* (2019). doi:10.1007/s10707-019-00346-1.

[21] A. Crooks, C. Castle and M. Batty, Key challenges in agent-based modelling for geo-spatial simulation, *Computers, Environment and Urban Systems* **32**(6) (2008), 417–430. doi:10.1016/j.compenvurbsys.2008.09.004.

[22] K. Csilléry, M.G. Blum, O.E. Gaggiotti and O. François, Approximate Bayesian computation (ABC) in practice, *Trends in ecology & evolution* **25**(7) (2010), 410–418. doi:10.1016/j.tree.2010.04.001.

[23] C. Davidson-Pilon, *Bayesian Methods for Hackers: Probabilistic Programming and Bayesian Inference*, Addison-Wesley, New York, 2016. ISBN 978-0-13-390283-9.

[24] Department for Digital, Culture, Media & Sport, *National Data Strategy*, UK Government, London, UK, 2020.

[25] Department of Digital, Culture, Media and Sport, *National AI Strategy*, Office for Artificial Intelligence, London, UK, 2021. ISBN 978-1-5286-2894-5.

[26] J. Epstein and R. Axtell, *Growing Artificial Societies: Social Science from the Bottom Up*, Brookings Institution Press, 1996.

[27] A. Fadikar, D. Higdon, J. Chen, B. Lewis, S. Venkatramanan and M. Marathe, Calibrating a stochastic, agent-based model using quantile-based emulation, *SIAM/ASA Journal on Uncertainty Quantification* **6**(4) (2018), 1685–1706. doi:10.1137/17M1161233.

[28] J. Ge and B.A. Furtado, Simulating urban transition in major socio-economic shocks, in: *2021 Winter Simulation Conference (WSC)*, IEEE, 2021, pp. 1–10.

[29] J. Ge, J.G. Polhill, T. Craig and N. Liu, From oil wealth to green growth-an empirical agent-based model of recession, migration and sustainable urban transition, *Environmental modelling & software* **107** (2018), 119–140. doi:10.1016/j.envsoft.2018.05.017.

[30] J. Ge, J.G. Polhill, J.I. Macdiarmid, N. Fitton, P. Smith, H. Clark, T. Dawson and M. Aphale, Food and nutrition security under global trade: A relation-driven agent-based global trade model, *Royal Society open science* **8**(1) (2021), 201587. doi:10.1098/rsos.201587.

[31] J. Ge, J.G. Polhill, K.B. Matthews, D.G. Miller and M. Spencer, Not one brexit: How local context and social processes influence policy analysis, *PloS One* **13**(12) (2018), e0208451. doi:10.1371/journal.pone.0208451.

[32] Geospatial Commission, *Unlocking the Power of Location: The UK's Geospatial Strategy*, Cabinet Office, London, UK, 2020.

[33] Z. Ghahramani, Probabilistic machine learning and artificial intelligence, *Nature* **521**(7553) (2015), 452–459. doi:10.1038/nature14541.

[34] Government Office for Science, *The Future of Mobility*, UK Government, London, UK, 2019.

[35] R. Greig and J. Arranz, Generating agent based models from scratch with genetic programming, in: *The 2021 Conference on Artificial Life*, MIT Press, Online, 2021.

[36] V. Grimm, E. Revilla, U. Berger, F. Jeltsch, W.M. Mooij, S.F. Railsback, H.-H. Thulke, J. Weiner, T. Wiegand and D.L. DeAngelis, Pattern-oriented modeling of agent-based complex systems: Lessons from ecology, *Science* **310**(5750) (2005), 987–991. doi:10.1126/science.1116681.

[37] A. Heppenstall, A. Crooks, N. Malleson, E. Manley, J. Ge and M. Batty, Future developments in geographical agent-based models: Challenges and opportunities, *Geographical Analysis* **53**(1) (2021), 76–91. doi:10.1111/gean.12267.

[38] A.J. Heppenstall, A.J. Evans and M.H. Birkin, Genetic algorithm optimisation of an agent-based model for simulating a retail market, *Environment and Planning B: Planning and Design* **34** (2007), 1051–1070. doi:10.1068/b32068.

[39] G. Jäger, Replacing rules by neural networks a framework for agent-based modelling, *Big Data and Cognitive Computing* **3**(4) (2019), 51. doi:10.3390/bdcc3040051.

[40] F.T. Johora, H. Cheng, J.P. Müller and M. Sester, An agent-based model for trajectory modelling in shared spaces: A combination of expert-based and deep learning approaches, in: *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '20, International Foundation for Autonomous Agents and Multiagent Systems*, Richland, SC, 2020, pp. 1878–1880. ISBN 9781450375184.

[41] R. Jordan, M. Birkin and A. Evans, An agent-based model of residential mobility, *Computers, Environment and Urban Systems* **48** (2014), 49–63. doi:10.1016/j.compenvurbsys.2014.06.006.

[42] A. Juliani, V.-P. Berges, E. Teng, A. Cohen, J. Harper, C. Elion, C. Goy, Y. Gao, H. Henry, M. Mattar et al., Unity: A general platform for intelligent agents, 2018, arXiv preprint. arXiv:1809.02627.

[43] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, 2003.

[44] R. Kitchin, Big data and human geography opportunities, challenges and risks, *Dialogues in Human Geography* **3**(3) (2013), 262–267. doi:10.1177/2043820613513388.

[45] A. Klabunde and F. Willekens, Decision-making in agent-based models of migration: State of the art and challenges, *European Journal of Population* **32**(1) (2016), 73–97. doi:10.1007/s10680-015-9362-0.

[46] J. Laufs, K. Bowers, D. Birks and S.D. Johnson, Understanding the concept of 'demand' in policing: A scoping review and resulting implications for demand management, *Policing and Society* **31**(8) (2021), 895–918. doi:10.1080/10439463.2020.1791862.

[47] D.J.B. Lloyd, N. Santitissadeekorn and M.B. Short, Exploring data assimilation and forecasting issues for an urban crime model, *European Journal of Applied Mathematics* **27**(Special Issue 03) (2016), 451–478. doi:10.1017/S0956792515000625.

[48] N. Lomax, A.P. Smith, L. Archer, A. Ford and J. Virgo, An open-source model for projecting small area demographic and land-use change, *Geographical Analysis* **54**(3) (2022), 599–622. doi:10.1111/gean.12320.

[49] Y. Long and X. Hu, Spatial partition-based particle filtering for data assimilation in wildfire spread simulation, *ACM Transactions on Spatial Algorithms and Systems* **3**(2) (2017), 1–33. doi:10.1145/3099471.

[50] C.M. Macal and M.J. North, Tutorial on agent-based modelling and simulation, *Journal of Simulation* **4**(3) (2010), 151–162. doi:10.1057/jos.2010.3.

[51] N. Malleson, A. Heppenstall and L. See, Crime reduction through simulation: An agent-based model of burglary, *Computers, Environment and Urban Systems* **34**(3) (2010), 236–250. doi:10.1016/j.compenvurbsys.2009.10.005.

[52] N. Malleson, A. Heppenstall, L. See and A. Evans, Using an agent-based crime simulation to predict the effects of urban regeneration on individual household burglary risk, *Environment and Planning B: Planning and Design* **40**(3) (2013), 405–426. doi:10.1068/b38057.

[53] N. Malleson, K. Minors, L.-M. Kieu, J.A. Ward, A. West and A. Heppenstall, Simulating crowds in real time with agent-based modelling and a particle filter, *Journal of Artificial Societies and Social Simulation* **23**(3) (2020), 3. doi:10.18564/jasss.4266.

[54] N. Malleson, A. Vanky, B. Hashemian, P. Santi, S.K. Verma, T.K. Courtney and C. Ratti, The characteristics of asymmetric pedestrian behavior: A preliminary study using passive smartphone location data, *Transactions in GIS* (2018). doi:10.1111/tgis.12336.

[55] E. Manley and T. Cheng, Exploring the role of spatial cognition in predicting urban traffic flow through agent-based modelling, *Transportation Research Part A: Policy and Practice* **109** (2018), 14–23. doi:10.1016/j.tra.2018.01.020.

[56] E. Manley, T. Cheng, A. Penn and A. Emmonds, A framework for simulating large-scale complex urban traffic dynamics through hybrid agent-based modelling, *Computers, Environment and Urban Systems* **44** (2014), 27–36. doi:10.1016/j.compenvurbsys.2013.11.003.

[57] E. Manley, S.W. Orr and T. Cheng, A heuristic model of bounded route choice in urban areas, *Transportation Research Part C: Emerging Technologies* **56** (2015), 195–209. doi:10.1016/j.trc.2015.03.020.

[58] E.J. Manley, S.W. Orr and T. Cheng, A heuristic model of bounded route choice in urban areas, transportation research part C: Emerging, *Technologies* **56** (2015), 195–209. doi:10.1016/j.trc.2015.03.020.

[59] S. Manson, L. An, K.C. Clarke, A. Heppenstall, J. Koch, B. Krzyzanowski, F. Morgan, D. O'Sullivan, B.C. Runck, E. Shook and L. Tesfatsion, Methodological issues of spatial agent-based models, *Journal of Artificial Societies and Social Simulation* **23**(1) (2020), 3. doi:10.18564/jasss.4174.

[60] M. Mazur and E. Manley, Exploratory models in a time of big data, *Interdisciplinary Science Reviews* **41**(4) (2016), 366–382. doi:10.1080/03080188.2016.1257196.

[61] J. McCulloch, J. Ge, J.A. Ward, A. Heppenstall, J.G. Polhill and N. Malleson, Calibrating agent-based models using uncertainty quantification methods, *Journal of Artificial Societies and Social Simulation* **25**(2) (2022), 1. doi:10.18564/jasss.4791.

[62] B. Müller, F. Bohn, G. Dreßler, J. Groeneveld, C. Klassert, R. Martin, M. Schlüter, J. Schulze, H. Weise and N. Schwarz, Describing human decisions in agent-based models – ODD + d, an extension of the ODD protocol, *Environmental Modelling & Software* **48** (2013), 37–48. doi:10.1016/j.envsoft.2013.06.003.

[63] W. Nature, Rules for pandemic data access that everyone can trust, *Nature* **594**(7861) (2021), 8–8. doi:10.1038/d41586-021-01460-7.

[64] S. Olmez, D. Birks and A. Heppenstall, Learning complex spatial behaviours in ABM: An experimental observational study, 2022, arXiv.

[65] F. Oloo, K. Safi and J. Aryal, Predicting migratory corridors of white storks, ciconia ciconia, to enhance sustainable wind energy planning: A data-driven agent-based model, *Sustainability* **10**(5) (2018), 1470. doi:10.3390/su10051470.

[66] F. Oloo and G. Wallentin, An adaptive agent-based model of homing pigeons: A genetic algorithm approach, *ISPRS International Journal of Geo-Information* **6**(1) (2017), 27. doi:10.3390/ijgi6010027.

[67] G. Papoudakis, F. Christianos, A. Rahman and S.V. Albrecht, Dealing with non-stationarity in multi-agent deep reinforcement learning, 2019, https://arxiv.org/abs/1906.04737. doi:10.48550/ARXIV.1906.04737.

[68] S. Saeedi, Integrating macro and micro scale approaches in the agent-based modeling of residential dynamics, *International Journal of Applied Earth Observation and Geoinformation* **68** (2018), 214–229, https://www.sciencedirect.com/science/article/pii/S0303243418301764. doi:10.1016/j.jag.2018.02.012.

[69] O.T. Sargoni and E. Manley, A sequential sampling model of pedestrian road crossing choice, in: *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on GeoSpatial Simulation, GeoSim '20*, Association for Computing Machinery, New York, NY, USA, 2020, pp. 10–19. ISBN 9781450381611. doi:10.1145/3423335.3428163.

[70] T. Schelling, Models of segregation, *The American Economic Review* **59**(2) (1969), 488–493.

[71] M. Schlüter, A. Baeza, G. Dressler, K. Frank, J. Groeneveld, W. Jager, M.A. Janssen, R.R.J. McAllister, B. Müller, K. Orach, N. Schwarz and N. Wijermans, A framework for mapping and comparing behavioural theories in models of social-ecological systems, *Ecological Economics* **131** (2017), 21–35. doi:10.1016/j.ecolecon.2016.08.008.

[72] A. Sevtsuk and C. Ratti, Does urban mobility have a daily routine? Learning from the aggregate data of mobile networks, *Journal of Urban Technology* **17**(1) (2010), 41–60. doi:10.1080/10630731003597322.

[73] W. Shi, M.F. Goodchild, M. Batty, M.-P. Kwan and A. Zhang (eds), *Urban Informatics*, The Urban Book Series, Springer, Singapore, Singapore, 2021, 9789811589836. ISBN 9789811589829. doi:10.1007/978-981-15-8983-6.

[74] A.D. Singleton, S.E. Spielman and D.C. Folch, *Urban Analytics, Spatial*, Analytics and GIS Series, SAGE, Los Angeles, 2018. ISBN 978-1-4739-5862-3 978-1-4739-5863-0.

[75] D. Spiegelhalter, Risk and uncertainty communication, *Annual Review of Statistics and Its Application* **4**(1) (2017), 31–60. doi:10.1146/annurev-statistics-010814-020148.

[76] F. Spooner, J.F. Abrams, K. Morrissey, G. Shaddick, M. Batty, R. Milton, A. Dennett, N. Lomax, N. Malleson, N. Nelissen, A. Coleman, J. Nur, Y. Jin, R. Greig, C. Shenton and M. Birkin, A dynamic microsimulation model for epidemics, *Social Science & Medicine* **291** (2021), 114461. doi:10.1016/j.socscimed.2021.114461.

[77] C. Sturley, A. Newing and A. Heppenstall, Evaluating the potential of agent-based modelling to capture consumer grocery retail store choice behaviours, *The International Review of Retail, Distribution and Consumer Research* **28**(1) (2018), 27–46. doi:10.1080/09593969.2017.1397046.

[78] D. Tang, Data assimilation in agent-based models using creation and annihilation operators, 2019, arXiv:1910.09442 [cs].

[79] D. Tang, Finding the maximum-a-posteriori behaviour of agents in an Agent-Based Model, 2020, arXiv:2005.02096 [cs].

[80] P. Ternes, J.A. Ward, A. Heppenstall, V. Kumar, L.-M. Kieu and N. Malleson, Data assimilation and Agent-Based Modelling: Towards the incorporation of categorical agent parameters, *Open Research Europe* **1** (2021), 131. doi:10.12688/openreseurope.14144.1.

[81] J.C. Thiele, W. Kurth and V. Grimm, Facilitating parameter estimation and sensitivity analysis of Agent-Based Models: A cookbook using NetLogo and 'R', *Journal of Artificial Societies and Social Simulation* **17**(3) (2014). doi:10.18564/jasss.2503.

[82] B.M. Turner and T. Van Zandt, A tutorial on approximate Bayesian computation, *Journal of Mathematical Psychology* **56**(2) (2012), 69–85. doi:10.1016/j.jmp.2012.02.005.

[83] United Nations, 68→Live in Urban Areas by 2050, Says UN, un.org, 2018.

[84] M.B.A. van Asselt and J. Rotmans, Uncertainty in integrated assessment modelling, *Climatic Change* **54** (2002), 75–105. doi:10.1023/A:1015783803445.

[85] S. van der Hoog, Deep learning in (and of) Agent-Based Models: A prospectus, 2017, https://arxiv.org/abs/1706.06302. doi:10.48550/ARXIV.1706.06302.

[86] E. van der Vaart, M.A. Beaumont, A.S.A. Johnston and R.M. Sibly, Calibration and evaluation of individual-based models using approximate Bayesian computation, *Ecological Modelling* **312** (2015), 182–190. doi:10.1016/j.ecolmodel.2015.05.020.

[87] A. Voinov and H.H. Shugart, 'Integronsters', integral and integrated modeling, *Environmental Modelling & Software* **39** (2013), 149–158. doi:10.1016/j.envsoft.2012.05.014.

[88] T.M. Vu, C. Buckley, H. Bai, A. Nielsen, C. Probst, A. Brennan, P. Shuper, M. Strong and R.C. Purshouse, Multiobjective genetic programming can improve the explanatory capabilities of mechanism-based models of social systems, *Complexity* **2020** (2020), 1–20. doi:10.1155/2020/8923197.

[89] T.M. Vu, C. Probst, J.M. Epstein, A. Brennan, M. Strong and R.C. Purshouse, Toward inverse generative social science using multi-objective genetic programming, in: *Proceedings of the Genetic and Evolutionary Computation Conference*, ACM, Prague Czech Republic, 2019, pp. 1356–1363. ISBN 978-1-4503-6111-8. doi:10.1145/3321707.3321840.

[90] M. Wang and X. Hu, Data assimilation in agent based simulation of smart environments using particle filters, *Simulation Modelling Practice and Theory* **56** (2015), 36–54. doi:10.1016/j.simpat.2015.05.001.

[91] J.A. Ward, A.J. Evans and N.S. Malleson, Dynamic calibration of agent-based models using data assimilation, *Royal Society Open Science* **3**(4) (2016). doi:10.1098/rsos.150703.

[92] A. Whipp, N. Malleson, J. Ward and A. Heppenstall, Estimates of the ambient population: Assessing the utility of conventional and novel data sources, *ISPRS International Journal of Geo-Information* **10**(3) (2021), 131. doi:10.3390/ijgi10030131.

[93] G. Wu, A. Heppenstall, P. Meier, R. Purshouse and N. Lomax, A synthetic population dataset for estimating small area health and socio-economic outcomes in Great Britain, *Scientific Data* **9**(1) (2022), 19. doi:10.1038/s41597-022-01124-9.

[94] B. Zhou, X. Wang and X. Tang, Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents, in: *2012 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Providence, RI, 2012, pp. 2871–2878. ISBN 978-1-4673-1228-8 978-1-4673-1226-4 978-1-4673-1227-1. doi:10.1109/CVPR.2012.6248013.