

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Sensor-Driven, Spatially Explicit Agent-Based Models

Francis Oloo

Abstract

Conventionally, agent-based models (ABMs) are specified from well-established theory about the systems under investigation. For such models, data is only introduced to ensure the validity of the specified models. In cases where the underlying mechanisms of the system of interest are unknown, rich datasets about the system can reveal patterns and processes of the systems. Sensors have become ubiquitous allowing researchers to capture precise characteristics of entities in both time and space. The combination of data from in situ sensors to geospatial outputs provides a rich resource for characterising geospatial environments and entities on earth. More importantly, the sensor data can capture behaviours and interactions of entities allowing us to visualise emerging patterns from the interactions. However, there is a paucity of standardised methods for the integration of dynamic sensor data streams into ABMs. Further, only few models have attempted to incorporate spatial and temporal data dynamically from sensors for model specification, calibration and validation. This chapter documents the state of the art of methods for bridging the gap between sensor data observations and specification of accurate spatially explicit agent-based models. In addition, this work proposes a conceptual framework for dynamic validation of sensor-driven spatial ABMs to address the risk of model overfitting.

Keywords: data-driven models, sensor-driven models, dynamic spatial models, spatial simulation models

1. Introduction

Agent-based models (ABMs) are mathematical models that attempt to reveal system-level properties by representing local-level behaviour and interaction of entities that make up the system [1]. Agents include people, animals, robots, vehicles, plants and smart devices that may be linked in a network, etc. ABMs have been applied to investigate systems in ecology [2, 3], human behaviour [4], epidemiology [5–7], public transport [8, 9], diffusion of technology [10], land use change [11], industrial processes, economics and psychology, among other areas.

An important characteristic of agent-based models is their ability to reveal the emergence of system-level patterns from the local-level behaviours and interactions of system components [12]. However, one traditional weakness of ABMs is their over-reliance on existing theories about the system of phenomena of interest [13]. Over-reliance on domain knowledge limits the application of ABMs in situations where knowledge about the system of interest is incomplete. In such

cases, parameter values and behavioural rulesets have to be assumed, thus reducing the plausibility of the models [14]. In addition, in knowledge-driven models, face validation [15] is preferred to statistical validation. Specifically, modelled system behaviours are compared against the qualitative patterns as described in the theories or in the implicit expert knowledge. Moreover, validation is usually implemented as the final step of model specification, hence hindering the dynamic verification of the models during the simulation runs. In very dynamic systems, the model is thus likely to deviate from the real-world scenario unless data about the dynamics of the real world is incorporated into the model during the simulation process.

Due to the limited computational resources and lack of fine-scaled spatial data, ABMs were historically nonspatial, implying that geographic characteristics of the systems of interest were not explicitly specified in the model [16]. As an example, to understand market dynamics, service area of the market of interest may be specified as a Cartesian grid with random cells representing business entities, while consumers are specified as points that move randomly across the modelling surface. Even though such a model can answer generic questions on the consumer behaviour, it may not be able to provide specific insights of the influence of spatial context on the market dynamics.

Advances in sensor technology have made it possible to collect accurate geo-referenced data about entities and systems of interest [17]. Fine-scaled sensor data from remote locations are now available for analysis and visualisation. For instance, spatial entities such as humans, vehicles, buildings, animals and plants can be monitored via sensor data streams, revealing interesting spatial and temporal characteristics of these agents. This rich data can provide behavioural information [18, 19] for specifying accurate data-driven models to study the dynamics of the agents of interest.

The emergence of sensor data has not only heightened the interest in spatial ABMs [20] but has also motivated the specification of data-driven models [21] for accurate environmental monitoring and simulation [22]. At the same time, the dynamic nature of sensor data streams has motivated research to bridge the gap between sensor observations and modelling frameworks as a way of facilitating bidirectional communication between sensor observation networks and environmental monitoring systems.

Unfortunately, the progress in sensor-driven spatial simulation models has been ad hoc, with no standardised methods for incorporating data into spatially explicit models. The existing implementations have aimed to address the needs of various disciplines. For instance, in the sensor community, research in sensor web networks [23] is geared towards improving communication, computation and sensor resource management. On the other hand, in computer science, sensor-based research is geared towards developing methods of pervasive computing [24, 25], artificial intelligence and related areas. Due to the multidisciplinary nature of spatial simulation research, documenting the body of knowledge of sensor-driven simulation modelling is critical for the research community.

Spatial systems are special [26] due to the inherent spatial relationships and temporal characteristics of geographic entities. It is therefore necessary to consider the spatio-temporal context [27, 28] and relationships when modelling and simulating spatial processes. Attempts to introduce data into spatial simulation models must be cognizant of the unique characteristics of the spatial systems. This work synthesises the existing methods for dynamic assimilation of sensor data into spatially explicit ABMs and proposes a potential method to address model overfitting that is common to most data-driven modelling methods.

2. Traditional knowledge-driven models

2.1 Essential building blocks of spatial agent-based models

Patterns are the holy grail of spatial agent-based models [29, 30], implying that reproducing spatial patterns is an important characteristic of spatial agent-based simulation. The three important aspects of spatial systems include agents, spatial context or environment and interactions between agents and their environment.

2.1.1 Spatial agents

Spatial agents include autonomous entities that can be characterised by their geographic attributes. Geographic attributes are critical in linking the agent to a unique spatial location and context of its environment. Distinctive attributes of spatial agents include spatial intelligence and spatial interactions. Spatial intelligence entails awareness of the geographic differences of the environment, hence being able to make autonomous decisions over a geographic space [31]. Moreover, spatial intelligence allows agents to interact with other spatial entities and adapt to spatial realities [32].

In defining the character and behaviour of agents in spatial simulation models, traditional models have ignored empirical data and instead used documented knowledge about the agents or random initialisation of agent characteristics [33]. This raises the question on whether such models are immune to the challenge of path dependence [34] that bedevils most ABMs.

2.1.2 Spatial environment

Initially, the variations in the environment of ABMs were commonly specified as an artificial lattice with random variables [35, 36]. With the improvements in computation, and availability of spatial data in both vector and raster data models, spatial data has been introduced to introduce geographic variability and context the in the environment [37]. In particular, the use of remote sensing products has improved the specification of geographic modelling environments [38].

2.1.3 Spatial interactions

Spatial interaction entails the ability to sense, communicate and respond stimuli from other entities based on their geographic proximity or connections. Interaction is the distinctive attribute of spatial ABMs, differentiating such models from other micro-simulation models. Whereas initially there has been little empirical data to reveal the interaction between agents, in situ sensors are now capable of capturing detailed aspects of agent interactions including proximity, avoidance, competition and spatial linkages. For instance, trajectories of birds in navigation have been used to describe the social interactions and leadership strategies that are adopted by birds [39]. Also, physiological sensors have been used to detect emotional reactions of road users in urban traffic [40, 41]. Similarly, there are portable sensors that can be used to monitor the health of human agents remotely [42]. Data from such sensor deployments can improve the specification of agent interactions and contribute to the accuracy of agent-based models.

2.2 Conventional modelling cycle

Traditionally, the modelling cycle [43] begins by building a conceptual model about the real-world system of interest. The conceptual model is created from

repeated observations of the mechanisms of the real-world system or by relying on documented knowledge about the system. From observations and the domain knowledge, important entities, interactions and patterns are identified. A hypothesis of how the individual level interactions of the agents lead to the emergence of system-level patterns is then formulated. At this level, the use of empirical data is limited to identifying essential entities, interactions and characteristic patterns of the system of focus (**Figure 1**).

Based on the conceptual model, a formal model specification could be undertaken to test a specific hypothesis. Model specification requires the definition of parameters to guide the operation of the model. The choice of the parameters and essential behaviour models depends on the expertise of the modeller and prevailing knowledge of the system under investigation [44]. This in essence means that different modellers can specify different ABMs to test the same hypothesis. It may so happen that different models can confirm the hypothesis, raising the question on true model for addressing the hypothesis in question.

Empirical data is rarely used during model specification; this is both epistemic and strategic. Epistemic in the sense that rather than starting with the data, a plausible model that is founded on sound knowledge should produce data, which is comparable to empirical data from the real world [45]. In addition, agents interact based on their knowledge of their environment and goals and not so much based on their rigorous analysis of data. The limited use of data in ABMs is also strategic to prevent the contamination of model with empirical data, which may ultimately lead to model overfitting.

Upon a successful model specification, the process of verification confirms the logical consistency between the specified behaviour and the known behaviour of the system. The verified model then becomes a candidate for calibration.

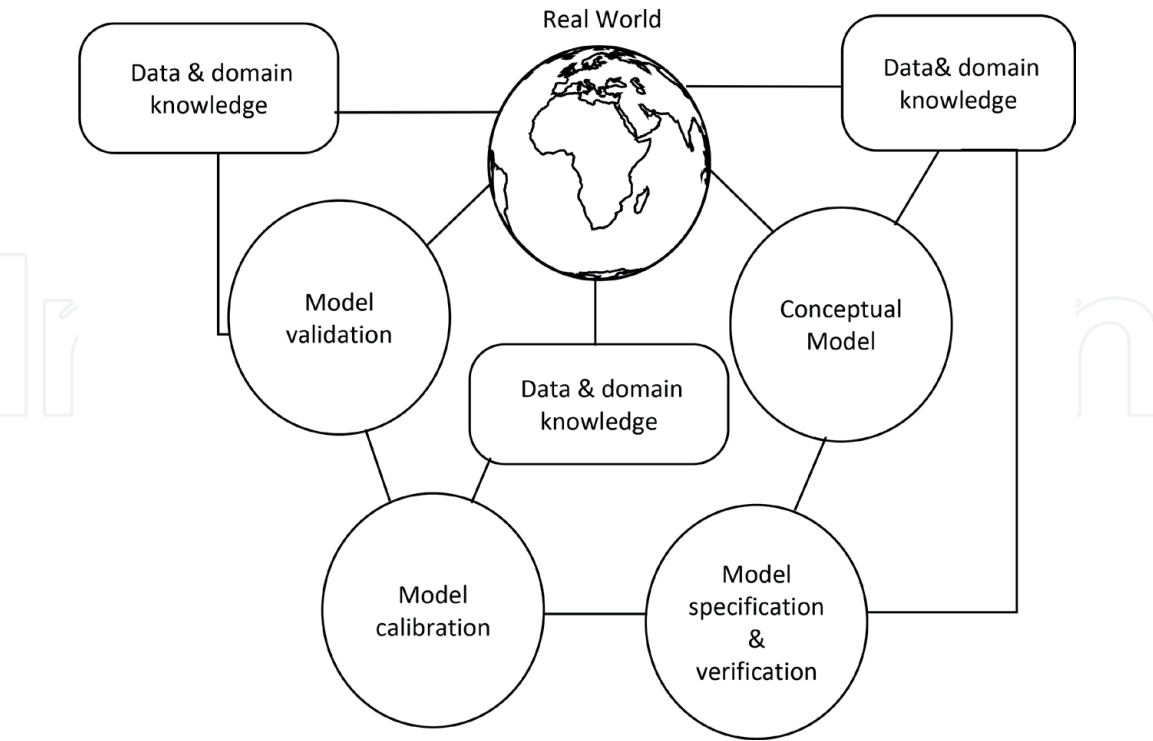


Figure 1. Application of data and domain knowledge in the conventional modelling life cycle. The cycles represent the important steps in the specification of ABMs. The real world hosts the systems of interest. Observations, expert knowledge and data about the systems in the real world provide the foundational concepts for building the conceptual model. Foundational knowledge and patterns from the real world are used for model specification, calibration and validation. A properly validated model can then be used to test hypothesis and to represent the dynamics of a system in the real world.

The calibration step entails comparing a specified model against empirical data to determine the parameter space for accurate simulation of patterns and dynamics in the real system. A popular method for calibration is the use of pattern-oriented modelling (POM) approach [46]. In POM approach, model calibration involves evaluating parameters based on their ability to replicate multiple patterns that are evident in the real world. In traditional modelling frameworks, historical data may be used in calibration and in other components of the agent-based modelling life cycle.

Validation process is usually the last step in the modelling cycle and involves assessing the degree to which a model is an accurate representation of the real-world system for which it is meant to simulate [47]. For validation, qualitative approaches may be adopted to compare the results from the models against patterns that are observable in the real world. In particular the use of face validation which may include animation or graphical representation is usually the first step in traditional ABM validation [48]. Once again, according to POM framework, an accurate model should be able to produce patterns that are inherent in the real world but which are not explicitly defined in the model.

Apart from qualitative methods, statistical methods [49] may also be adopted to validate the models by comparing the statistical variance between the results of the model against empirical. Statistical comparison is suitable for models that produce detailed quantitative state variables that can be compared to related observations from the real world. A properly specified, rigorously calibrated and accurately validated model can then be deployed for simulation to represent the system of interest and to explore the internal operations of systems of interest.

In summary, the rationale for incorporation of data into traditional agent-based models is to ensure quality and credibility throughout all the modelling stages [50, 51].

2.3 Standards for the specification of ABMs

Because of the straightforward manner of specifying knowledge-driven models, such models are simpler to specify and easy to communicate. The publication of standards to guide ABM specification [52] and protocols like transparent and comprehensive ecological modelling (TRACE) documentation [53], pattern-oriented modelling [46] and Overview, Design Concepts and Details (ODD) protocol [54, 55] have greatly contributed to streamlining the process of model specification. In addition, depending on domain knowledge ensures that models are only acceptable when their results confirm the documented knowledge hence helping to weed out models that result in spurious outcome. The multidisciplinary nature of geographic information science avails knowledge from related disciplines including ecology, computer science, geography, environmental science, economics and psychology, which can support the specification of spatially explicit agent-based models. In the reverse direction, properly specified spatial simulation models can support hypothesis testing and representation of dynamics of systems in other disciplines.

2.4 Critiques of knowledge-driven ABMs

In spite of the benefits of knowledge-driven ABMs, there have been critiques of aspects that limit their broad adoption and application. In particular, the process of model specification depends on the knowledge and expertise of the modeller; as such, discovery of patterns and the specification of behavioural rulesets in the model may be arduous task in situations where the system of interest is not well understood [56]. In addition, the ad hoc manner of model specification may result in multiple models for the same system without bringing clarity on the internal workings of the system. Further, lack of modules to actualise rigorous data mining

within the simulation suites has hindered the development of agent-based model that can take advantage of the growing big geospatial data. Moreover, the dependence on domain knowledge and the expertise of individual modellers worsen the gap between modelled examples and the ever-growing data volumes. Individual modellers cannot keep pace with the growth of data, hence necessitating the development of automated methods for model discovery and analysis.

Last but not the least, whereas knowledge-driven models can support the specification of simple models, such models are usually weak in predicting future behaviours of the system [57]. This is more so when the potential effect of various inputs on future states of a system is unknown. As an example, initially, it was possible to model the generic annual behaviour of migratory birds particularly in the wintering months. However, with the reality of human-induced changes to the environment, some birds avoid the long winter journeys and instead find food and warm nesting places around garbage disposal sites in the northern hemisphere [58]. Such specific adaptive behaviours were only detectable through analysis of the empirical trajectories of the birds.

Bridging the gap between the advances in big geospatial sensor data and spatially explicit ABMs requires robust methods for automated pattern detection and model discovery.

2.5 Multi-agent systems and swarm intelligence

Multi-agent systems (MAS) are an extension of single-agent systems and comprise of multiple software agents interacting with each other and their environment to achieve certain goals. Important characteristics of multi-agent systems include communication, collaboration and interaction. In MAS, the agents can either be intelligent or reactive [59]. Intelligent agents are those that are able to logically use knowledge and information at their disposal to make rational decisions. On the other hand, reactive agents respond to the realities of their environment. In multi-agent systems with reactive agents, system-level robustness and complexity emerges from local-level interactions of the constituent agents. Collective intelligence that emerges from MAS is similar to those of swarm intelligence (SI), hence promoting the adoption of MAS in SI [60].

Swarm intelligence has its foundation in the behaviour of natural bio-systems [61]. Specifically, social organisms like bee and ant colonies, flocks of birds and schools of fish have been known to exhibit impressive collective behaviours that may not be directly linked to the capabilities of individual organisms. Swarm intelligence is therefore an attempt to adopt ideas and knowledge from the natural bio-systems to build robust algorithms with application in a number of fields. In particular, in swarm intelligence, software agents are specified to mimic the behaviour of natural systems with the aim of achieving specific goal through the emergence of coherent and functional patterns from the collective behaviour of interacting entities. The particular characteristics of software agents in swarm intelligence include autonomy, interaction, distributed functioning and self-organisation, ensuring that the software agents solve problems at hand without a central control. Swarm intelligence has been employed to build solutions for optimisation, computer network-based search, wireless sensor networks and traffic control, among other areas. Epistemologically, there are two motivations for swarm intelligence [62], the first being to learn about natural system and to understand the emergence of system-level patterns from collective interactions of individual entities of a system. The second motivation is to discover novel algorithms that can be used to solve various engineering, social and computer science problems.

A number of signature algorithms have been developed to actualise swarm intelligence in various applications. The most common of these algorithms include ant colony optimization (ACO), bee colony optimization (BCO) [63] and particle swarm optimization (PSO) [64]. ACO is motivated by the foraging behaviour of ant colonies. Specifically, as individual ants forage for food, they release a chemical known as pheromone when they succeed at finding food. Other members of the colony can detect the pheromone and move to the spot where food has been found. The pheromone evaporates with time. This type of communication between members of a colony ensures an efficient search for food. This model has been applied to simulate swarm intelligence in public transport services [65]. Bee colony optimization algorithms mimic the foraging behaviours of bee colonies where individual bees make characteristic “dances” to alert the members of the colony on the locations of food availability. Other members of the colony can choose to go to this spot by a certain probability. Particle swarm optimization are stochastic optimisation techniques that are inspired by the goal-oriented behaviour of flocking birds [66] that improve the efficiency of their navigation and foraging behaviours through collaboration, cooperation and independent local-level decisions. Particles in a swarm are considered to have limited intelligence and autonomy and exercise simple local-level rules to optimise their flow. PSO has been applied to optimise network-based communication.

Apart from the main algorithms for swarm intelligence, other algorithms which are motivated by natural systems have been tested in multi-agent systems and later adapted for swarm intelligence; these include genetic algorithms, neural networks, re-enforced learning and simulated annealing. Apart from serving as a test bed for nature inspired algorithm, MAS also provide a platform for specifying, modelling and simulating natural systems, thus contributing to the knowledge that is then ultimately adapted in swarm intelligence. The emergence of sophisticated sensors has made it possible to embed sensor in systems of interest. The sensor data can then be used to specify multi-agent models of the system allowing biologists and computer scientists to learn the behaviours of these systems, hence making it possible to simulate these to improve the algorithms for swarm intelligence [67].

3. Foundations of data-driven agent-based models

3.1 Influence of data in the character of agent-based models

There are three broad motivations for specifying agent-based models including testing hypothesis about a particular system, representing the dynamics of a system and predicting the potential future states of a system. Empirical data has traditionally been used in ABMs to characterise agents in the model, for model initialisation and for validation [68]. Injecting data into agent-based models can influence the purpose of the models. Consequently, three general types of agent-based models with distinct roles depending on the degree to which data is used to aide their specification emerge. The three categories include generator models, mediator models and predictive models (**Figure 2**).

3.1.1 Generator models

Generator models are the most common types of agent-based models and have their foundations in generative social sciences [69]. These models rely heavily on the domain knowledge and the expertise of the modellers to specify behaviour rules and model structures. Consequently, such models are predominantly used for

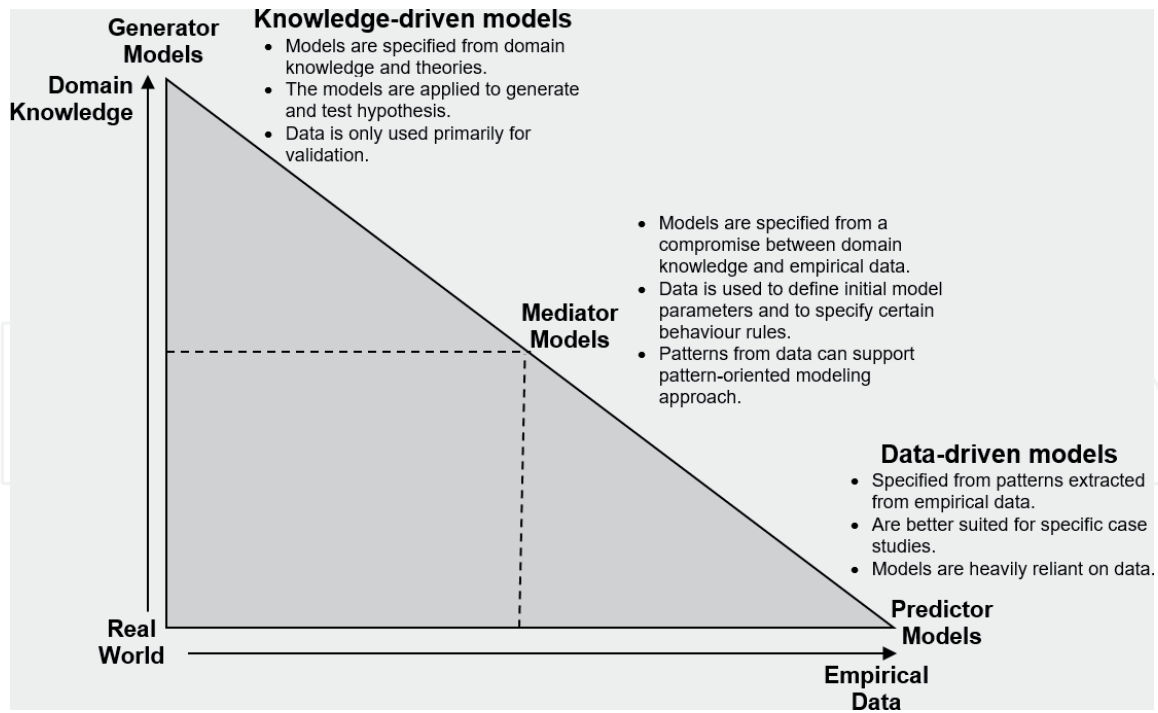


Figure 2.

Categories of agent-based models depending on the degree of data used in the models. Generator models rely heavily on the domain knowledge and are suited for “generating” system scenarios based on variations of input parameters. Mediator model combine both domain knowledge and patterns extracted from data to improve the understanding of a system. Data-driven models are heavily reliant on data and perform better as prediction models.

generating and testing different hypotheses [70]. Generator models aim to demonstrate or “generate” a scenario based on the foundational theories of the dynamic of a system of interest. The models may require minimal data to support initialisation, calibration and validation. By relying on domain knowledge and ingesting marginal data, such models are generic and can replicate related systems. However, the models cannot be relied on to reveal very detailed dynamics of the systems.

3.1.2 Mediator models

The second category of models are the mediator models that move beyond hypothesis testing and attempt to create a better understanding of the system of interest, hence attempting to explain the dynamics of a system. In these models, empirical data provide additional interesting patterns and parameters for model specification. Validation step then confirms whether the models can replicate the patterns that are apparent in the empirical data. Such models can be used to evaluate the implications of empirical research on formal theories [71]. A distinction between mediator models and the generator models is that the modellers do not require complete knowledge about the systems of interest. Specification of accurate models can be achieved by combining partial knowledge of the systems with important characteristics of the system as captured in data. As an example, crowd behaviour in an enclosed building can be studied from video data [72] and used to improve models that represent the behaviour of crowd agents.

3.1.3 Predictor models

Whereas knowledge-driven models can be generic and be applicable to test broad system characteristics, they have not been particularly strong in predicting very specific and detailed aspects in the future system characteristics. In contrast,

models that are fuelled by rich datasets are likely to perform better as predictive models [7, 73]. The rich data supports the understanding of the respective system by revealing useful inputs and systemic behaviours that can be used in specifying the model structure [74]. For instance, a fire model that is trained with accurate spatial data on the vegetation characteristics, climatic variables and other contextual information regarding fire dynamics in a particular locality is likely to predict future fire scenarios better than a model that is based on the general understanding of fire dynamics [75].

Domain knowledge provides a good starting point for specifying realistic models for hypothesis testing and for representing behaviours of systems. However, the lack of solid foundational knowledge should not be a handicap for the specification of accurate agent-based models. In an emerging field like geographic information science, the process of knowledge discovery should continue in tandem with the advances in methods that can facilitate infusion of rich data into agent-based models. This can create a mutually beneficial feedback between knowledge-driven models and data-driven model. Importantly, there are concepts in spatial science that are yet to be defined in a crisp manner. The development of data-rich and spatially explicit simulation models can therefore contribute towards building the understanding of some concepts in spatial science, particularly those that concern spatial behaviours [76].

Models that entirely depend on historical knowledge and static datasets may be limited by their failure to appreciate the dynamic conditions of spatio-temporal systems [77] that can only be revealed by capturing the data in near real time. In addition, spatial simulation models which rely primarily on conventional spatial data models may be limited in capturing all the necessary spatial processes [38]. It is therefore important to augment the spatial data models with sensor data streams or other ambient positioning methods that can capture the multiple dimensions of spatial phenomena and processes. Moreover, an understanding of dynamic spatial processes requires the specification of data-driven models that can combine both spatial data models and spatial process models. Sensor data streams can capture dynamic spatial events [78] and associated processes, hence supporting a tighter link between dynamic data and dynamic spatially explicit agent-based models.

3.2 Dynamic data-driven simulation models

In the last two decades, there have been attempts to achieve dynamic data-driven simulation models (DDDABM). This is more so in systems that are characterised by dynamic spatial and temporal behaviours [79]. Advances in sensor capabilities are a major driver of the attempts to actualise dynamic data-driven application systems (DDDAS). In particular, miniaturisation of the sensors, improvement in computational power and developments in telecommunication have led to the growth of robust sensor web networks that can be adopted to address questions in various spatial domains. Importantly, the growth of geosensor networks has made it possible for sensors to capture not only the geographic locations of entities but also the behavioural characteristics of such entities [80, 81]. For example, there are sensors that can capture both the location and multidimensional acceleration of animals, hence revealing their energy use during different activities [82]. The sensor measurements can be related to animal behaviours in different settings, hence allowing for the behaviour of animals to be documented remotely. Another example includes the possibility of capturing location, mobility characteristics and fuel consumption in vehicles, hence linking the mobility patterns to energy use efficiency and safety [83].

Within the wireless sensor networks (WSN), a common approach for actualising dynamic data driven simulation has been to specify sensors as software agents

within the model [84]. Such an architecture allows the sensor data to influence the specification of the agent-based model, while the output from the simulation influences the sensor measurement strategies and network configuration. Moreover, agent-based specification of sensor nodes allows for optimisation of the network resources and promotes energy efficiency within the WSN [85]. The bidirectional feedback between wireless sensor networks and the software agents is mutually beneficial both for the efficiency of sensor data collection and for the accuracy of the simulation models. There are three general approaches for actualising data-driven agent-based simulation. These include decoupled data integration, dynamic unidirectional data integration and dynamic bidirectional data assimilation.

3.2.1 *Decoupled data integration*

In this first approach, data is decoupled from the simulation and is only introduced sparingly to influence various steps in the modelling workflow. For instance, data may be used in specifying the initial conditions of agents, defining the initial model parameters, supporting calibration and validation of spatial agent-based models [56]. For this kind of approach, archival data in the form of surveys [13] or historical movement trajectories of agents may be adopted. The data provides the main characteristics of the agents and possibly also the transition probabilities from one state to the next. However, since such models are delinked from the real world and only make little use of historical data, they may fail to reflect dynamic characteristics of the real world [86]. In addition, apart from using the data for validation, the data may not influence the structure of the model [87].

3.2.2 *Dynamic unidirectional data integration*

The second approach entails a unidirectional flow of data from measuring systems to the simulation model. The data may capture the characteristics of the agents and be used to influence the dynamic behaviour of agents. For instance, taxi probe data may be gathered and be used to learn about agent characteristics and to implement a traffic-related simulation model [88]. However, the results from the simulation are not transferred to the measuring system to influence the data collection strategies. In addition, it is not necessary for the data to be recorded in real time. Data provides a means of extracting patterns that can then be used in ABM specification [72]. As an example, trajectories of animals, with precise spatial and temporal attributes can be used to infer patterns that may not be apparent in the domain knowledge [89]. The patterns from data can then be used to specify agent characteristics and to improve the model structure. An advantage of dynamic sensor data of this type lies in the repeated measurements of such data, which reveal the evolution of agent behaviour in space and time [90]. However, the simulation results are not compared to the real-time dynamics of the systems of interest; hence the model may still deviate from the reality.

3.2.3 *Dynamic bidirectional data assimilation*

In dynamic bidirectional data-driven models, real-time or near real-time sensor observations provide the empirical input that influence the simulation in real-time [21]. In the simplest form, data from the real world may only influence the characterisation of the modelling environment. For instance, real-time temperature and wind characteristics may be used to influence the environment of a model on fire dynamics [91, 92].

At the advanced level of dynamic data-driven simulation, output from the model can be used to influence the sensor measurement strategies. For example, when modelling the influence of a hurricane, sensors in areas which are characterised by minimal intensity and impact of the hurricane both in the real world and in the model can be shut down or slowed down, while the frequency of data collection of sensors in high priority areas can be increased [93]. The bidirectional feedback improves both the data collection strategy and the accuracy of the models [94].

To facilitate the bidirectional communication between sensors and simulation models, dynamic data-driven systems adopt a three-step process consisting of sensing, prediction and adapting [95]. During sensing, sensors measure or record the entities of interest; simulation models then predict the probable change in the state of the entity. Finally, the sensing system is adapted to capture and validate the simulated state of the entities. This generic approach provides the foundational concepts for specification of dynamic data-driven agent-based model.

3.3 Dynamic data-driven agent-based models

Dynamic data-driven agent-based models remain one of the exemplary specifications of sensor data-driven ABMs. In this implementation, dynamic sensor data streams improve the specification of multi-agent systems, allowing models to benefit from the real-time behaviour and interactions between agents in a real-world setting. The main components of this framework include (i) the sensor measuring, which observe entities in the real world, providing a mirror of the happenings in the system of interest; (ii) data management system; (iii) modelling or simulation platform; and (iv) visualisation and dynamic communication suite. In some instances, the modelling platform may also serve as the visualisation interfaces.

Whereas the initial DDDABM implementations were ad hoc and relied on standards developed within the project or on widely recognised standards within computer science and engineering, the latter adaptations have utilised well-established standards by the Open Geospatial Consortium (OGC) to promote standardised discovery of sensor resources, documentation of sensor observations and uncertainties and transfer of outputs from modelling workflows [96]. Data management strategy can either be loosely coupled, distributed or centralised or adopt a complex negotiation between a distributed and centralised data management strategy.

3.4 Types of dynamic sensor data-driven applications for simulation

In an attempt to bridge the gap between models and real-world systems, different approaches have been proposed or adopted to incorporate dynamic sensor data into spatially explicit ABMs. The purpose of data integration influences the methods for the integration and the extent to which data is used in the models. Some of the generic implementations include the following: (i) data-driven calibration of agent-based models, (ii) adaptive optimisation of model parameters, (iii) service-oriented architecture in geosimulation, (iv) agent parallelisation and dynamic visualisation, (v) dynamic data-driven multi-agent systems (DDDMAS) and (vi) adaptive discovery of models from sensor data streams.

While these categories may not be conclusive, they cover the main attributes of sensor data-driven spatial simulation models as will be specified in the following sections.

3.4.1 Data-driven calibration of spatial agent-based models

Conventionally, calibration of agent-based models aims to achieve two purposes. The first purpose is to find a robust and comprehensive list of parameters that can simulate the intended behaviours of a model. The second aim is to find optimal parameter ranges that can replicate the intended behaviours. Calibration is therefore an important step in model specification as it provides an idea of the essential parameters that affect agent behaviour while also providing the sensitivity ranges of these parameters. A properly calibrated model captures the essential dynamics of a system and can contribute towards achieving an accurate representation of the real world.

Traditionally, calibration of ABMs relies on historical data. However, in time-dependent and contextually sensitive systems like most of the spatial systems, historical data may not capture all the dynamics of a variant system. Consequently, calibration of models with historical data may cause the models to deviate from real-world realities. This is more so when the agents in the model face dynamics and situations that were absent in the historical data. Incorporating data from the real world during the model run can therefore provide a means of fine-tuning the model parameters to be reflective of the realities in the real-world scenarios [97]. For instance, when simulating road traffic, historical data may not have captured traffic jams that result from emergencies on the road, incorporating real-time data of such incidences when they occur can provide the necessary input to fine-tune the model parameters and to ensure the currency of model results.

To achieve dynamic calibration, it is important to have a systematic means of comparing model states against the real-world states. A proper scheduling scheme and tightly coupled link between the real-world schedule and the model schedule can help in deciding the calibration points. For instance, the schedule of sensor data collection and collation should be synchronised with the model time to allow for comparison between sensor observations and simulation results. It is therefore important to have an observation model from the sensor observation system that is comparable to the simulated results.

Dynamic calibration is achieved through methods of data assimilation which combine the state of the system as observed in the real world with the results from a simulation model in order to produce an improved prediction [98]. In particular, particle filter (PF) methods [99], for instance, Kalman filter (KF), have been used to assimilate data from pedestrian counts into a pedestrian simulation model [100]. In another example, Sequential Monte Carlo (SMC) method was used to assimilate sensor into building occupancy simulation [101].

For data assimilation, a proper sampling scheme allows for a randomised selection of data from the real world, and assimilating these with a sample of the simulation results to provide an updated state of model dynamics. Data assimilation improves the accuracy of the model as model parameters are updated to be in harmony with the patterns in the real world. However, models that are heavily reliant on data assimilation for the calibration of model parameters may run the risk of overfitting the model parameters to the data and therefore reduce the replicability of the models in data-deficient scenarios. Consequently, other methods which promote cross-validation [102] have been proposed as they go beyond dynamic calibration.

3.4.2 Adaptive optimisation and validation of model parameters

Discovery of representative parameters remains an outstanding challenge in the specification of data-driven spatial simulation models. This is more so when the velocity and volume of data collection outstrip the knowledge domain of the system

of interest. When rich-annotated sensor data is available, multiple parameters may be inferred from the data. However, not all the parameters may be useful or adequately robust for representing the system of interest. Identifying robust and representative set of parameters for capturing the behaviour of the system of interest becomes a challenge. In addition, finding optimal parameter space for simulating the real-world system accurately can be challenging. Consequently, discovery and optimisation of parameters has been another aspect of dynamic data-driven simulation. Statistical methods including Markov chain [103] and its variants have been employed to discover initial parameters that may influence the dynamics of a model.

Evolutionary methods are suitable for dynamic optimisation of model parameters. In particular, genetic algorithms that borrow from biology have been employed to optimise parameters in data-driven ABMs [104, 105]. This has particularly been possible because of the adaptive nature of genetic algorithms which allows them to learn from data and to improve the specification of model parameters.

It is possible to implement dynamic calibration and optimisation of model parameters from a centralised data management system. However, the dynamic nature of sensor resources requires a service-oriented architecture to facilitate dynamic discovery, analysis and communication of sensor resources using open and standardised protocols. Consequently, the development of various sensor resource management standards within OGC has promoted development of service-oriented architectures including Sensor Observation Service (SOS), Sensor Web Enablement (SWE) and other Geosensor Network Services that facilitate the discovery, access and computation on sensor resources in a standardised way. As a result, there have also been advances in sensor-oriented geosimulation frameworks.

3.4.3 Service-oriented geosimulation framework

In service-oriented geosimulation frameworks, sensor resources are specified as services that can be accessed and used in the model to achieve specific goals [106]. The adoption of sensor-oriented architecture in a dynamic data-driven ABM begins by considering Agents-as-a-Service (AaaS) [107]. In the approach, different aspects of the sensor data collection, management and computation system can be viewed in terms of their functionality [108]. The functionality defines the agency of these sensor network resources. For instance, sensor nodes whose role is to measure environmental characteristics exemplify measuring service hence can be specified as measuring agents.

Specification of sensor resources as services allows the elements of sensor network to be represented in the models as software agents. The behaviour and operations of sensor software agents can be simulated in parallel to other agents of interest in the system under analysis. For instance, in a hydrological network whose aim is to observe and analyse nutrient and sediment load in a catchment. Different sensors, for measuring environmental and hydrological characteristics, can be specified as agents in the model. Entities of interest, which may include water particles and sediment, can also be specified as autonomous agents. Specification of various sensor components as service agents in the model also allow for agent characteristics like autonomy, intelligence, interaction and adaptability to be included. Such agent characteristics can enhance the efficiency in the use of the sensor network resources and the versatility of the sensors in the model.

The service agents can provide the link between the real world and the simulation environment [109]. Specifically, the service agents capture information from the real world and execute the initial network level computations before relaying the processed information to fine-tune the specification of the model world while also providing information for dynamic calibration of the models. At the same time,

specifying sensors as service agents in the model also makes it possible to influence the behaviour of such sensor agents, hence promoting a bidirectional communication between the model and the sensing system. This characteristic makes it possible to manipulate the sensor behaviour from the model.

In spite of the positive attributes of adopting a service-oriented geosimulation, challenges emerge in communication, computation, visualisation and data management, necessitating the refinement of the service-oriented approach and the development of other paradigms like agent parallelisation and dynamic visualisation.

3.4.4 Agent parallelisation and dynamic visualisation

Parallelisation improves efficiency in spatial explicit ABMs with thousands of agents and multiple interconnected tasks [110]. As an example, incorporation of sensor data into geosimulation models may require distributed data management, exploratory data analysis, pattern extraction, dynamic calibration, analysis of the model results and complex communication between different model components. The multiple tasks, particularly when the velocity of the data streams is high and the volume of the data is big, can limit the efficiency of the intended model. Consequently, parallelisation can improve the efficiency of modelling operations. Within sensor-driven agent-based systems, two common types of parallelisation in spatial ABM include agent parallelisation and environment parallelisation [111]. For models with multiple sub-models, a third type of parallelisation is known as task parallelisation.

3.4.4.1 Agent parallelisation

Agent parallelisation entails separating, distributing and simulating the behaviour of various agents in different cores. Individual cores keep track of agent properties and spatial locations. In ecology, agent parallelisation has been implemented to simulate predator-prey models [112].

3.4.4.2 Environment parallelisation

Environment parallelisation involves breaking an expansive modelling world into multiple smaller spatial units or tiles and distributing the small units to different cores. Simulation can then proceed in each core. One challenge in this kind of setup is in simulating mobile agents that move extensively across the area of study.

3.4.4.3 Task parallelisation

Task parallelisation involves breaking down modelling tasks into different modular operations that can be performed in parallel in different cores [113]. For instance, an agent-based model can be broken down into sub-models that can run concurrently on parallelised cores. This kind of setup can also help in solving scheduling questions and can improve efficiency of simulation.

Important components of an effective parallelisation include distributed data management system, high-performance geosimulation environment which includes modules for specification of agency and a dynamic geo-visualisation platform [114]. The performance of the parallelisation scheme can be leveraged on open standards that facilitate distributed database management system, efficient communication [115], high-performance geosimulation and cyberGIS [116].

3.4.5 *Dynamic data-driven multi-agent systems (DDDMAS)*

Dynamic data-driven multi-agent systems are a modification of dynamic data-driven applications systems [21]. The initial motivation of DDAS was to support the implementation of dynamic environmental monitoring systems incorporating different application systems with real-time data from the system of interest. An important attribute of the DDAS is the possibility of bidirectional communication between sensors and models, which allows sensors to provide data from the real world for assimilation into models, hence improving the reliability of the models. On the other hand, simulation results influence the sensor measurement strategies.

In DDDMAS, the concepts from DDAS are adopted in a multi-agent system to improve the specification and accuracy of multi-agent models [117]. In particular, sensors capture individual agent characteristics, hence facilitating the specification of agents. In addition, other sensors can capture environmental characteristics, hence ensuring that the environment in which the agents interact is dynamic and representative of the reality. On the other hand, the model outcomes influence sensor measurement strategies by promoting priority sensor deployment depending on the scenarios in the model. In spatial models, sensor network components and other entities in the model can be represented as autonomous, which can be identified by their unique geographic characteristics [118].

Apart from placing the sensors on the environment, on-body sensors [119] can provide both the contextual and physiological characteristics of agents that may be important in understanding ambient behaviours of the simulated agents. To get the best out of the sensor agents, sensors should not only be measuring devices but must also be cognitive [120]. Cognitive sensor agents can have a mental state which may include intelligence, computational ability and decision-making components [20]. Other attributes that such cognitive agents may have include self-organisation, learning and adaptability. These attributes allow the sensors to gather information (both from the environment and from the models), analyse such information and make autonomous decisions that improve the data collection strategies and facilitate the specification of accurate models.

3.4.6 *Dynamic discovery of models from sensor data*

The most advanced level of dynamic data-driven simulation entails the discovery of rulesets and algorithms that make up accurate simulation models. The process of model specification can be arduous especially when there is vague knowledge about the system of interest. Automated discovery of robust algorithms, which are capable of representing the dynamics of a system of interest, is therefore a giant leap in the epistemology of agent-based simulations [121].

The essential building blocks of agent-based models are the entities, interactions and contextual information that influence entity decisions and interactions. Data containing detailed characteristics of the entities, their interactions and the contextual information in the environment where they operate may provide an avenue for discovering behavioural models of the agents, hence facilitating automated model specification.

Capturing the cognitive characteristics of humans and animals remains a challenge both technically and due to ethical reasons. However, recent advances in biosensor technology have made it possible to capture nonintrusive physiological characteristics which can then be related to the emotional and mental state of humans [122] and animals. Data on such cognitive characteristics of agents can facilitate in understanding and specifying the motivation of agents. In robotics and unmanned aerial vehicles (UAV), sensors can also be used to capture information

for building the intelligence of the robotics and of the UAVs [123]. Such agents therefore need an additional capability of learning, hence building their knowledge beyond the hard-wired artificial intelligence. The learned knowledge can improve swarm intelligence in UAVs, safety in self-driving cars and efficiency in adaptive industrial processes.

Because of the dynamic nature of data and complexities of the spatial environments, understanding of the agent decisions and the emergence of system-level characteristics requires an automated model discovery. One suggestion for generating spatial rulesets for multi-agent systems is the global-to-local programming approach [124]. The approach attempts to decompose a programming task into individual simple spatial dimensions and then generate candidate rulesets for each dimension. The dimensions may include configuration, local rules, timing, patterns and robustness. Genetic algorithm can then be used to combine and evolve the candidate sub-models resulting in a robust rulesets that can simulate the multi-agent system of interest [121].

Other implementations involve implementing methods from machine learning to discover an initial population of algorithms from a solution space [125]. The initial population of algorithms can then be optimised using genetic algorithms to produce the most efficient combination of algorithms that can simulate the system of interest. The result is an adaptive ruleset, which is not handicapped by the domain knowledge but that emerges based on the richness of solution space. The richness of the solution space depends on the diversity of data from various sensor data streams. Automated discovery of models can reduce the time spent in model specification and result in behaviours that can be described mathematically, hence improving the conceptualisation of agent behaviours. Consequently, such modelling workflows can contribute to automated knowledge discovery.

As has been outlined in this section, tremendous progress has been made to facilitate dynamic data integration into agent-based models. The progress is bound to shorten modelling cycle and to improve accuracy of ABMs by ensuring the fidelity of the models to the dynamic sensor observations in the real world. However, dependence on data may come with the challenge of model overfitting. Similarly, unless proper flexibility is allowed in the parameter estimation and model discovery, data-driven models can end up as “black box” models, which, even though may lead to accurate results, do not allow users to understand how the optimised parameters and adaptive algorithms emerge. In order to contribute to addressing the challenge of model overfitting, we see potential solutions in leveraging the specification of sensor-driven spatially explicit models on well-established guidelines like pattern-oriented modelling, service-oriented architecture, parallelisation and optimisation of various model components through evolutionary algorithms. In the following section, a conceptual framework for sensor-driven spatially explicit model is provided.

4. Framework for dynamic sensor-driven spatially explicit agent-based models

An accurate, spatially explicit, agent-based model should aim at replicating all the essential patterns of the system by simulating the local behaviour of agents. Pattern or behaviour detection is therefore an important component of data-driven simulation models [126]. Consequently, in order to specify accurate models, the modelling workflow requires a module to facilitate pattern extraction in order to discover multi-scale patterns from the sensor data streams. The patterns can drive dynamic calibration and validation of the model. Because of the velocity and dynamic nature

of sensor observations, bridging the gap between sensor observations and model specification necessitates the processes of calibration and validation to be closer and tied tightly to the simulation processes. This is in contrast to the conventional methods where specification, calibration and validation are sequential steps that are implemented at separate times. The challenge thus is to decide on a suitable pattern-oriented modelling strategy in which the patterns from sensor data streams are separated into specification, calibration and validation patterns. **Figure 3** provides the conceptual framework for sensor-driven spatial simulation model.

In the conceptual model, there are three important layers in the dynamic simulation life cycle. The three are the observation layer, exploratory analysis layer and the simulation layer.

4.1 Observation layer

The observation layer specifies the data collection and management strategy. In particular, the layer specifies the sensor-driven observation experiment and the associated sensor and network infrastructure that facilitate accurate, complete and efficient data collection and preprocessing. The preprocessing step may include spatial and temporal sampling of the sensor observations to capture only the important attributes of the agents of interest. In order to address the spatial questions that are the focus of spatial simulation modelling, observations should include both the spatial characteristics such as the location and time and other agents and environment-specific data. Consequently, standards from OGC Geosensor Network Services can

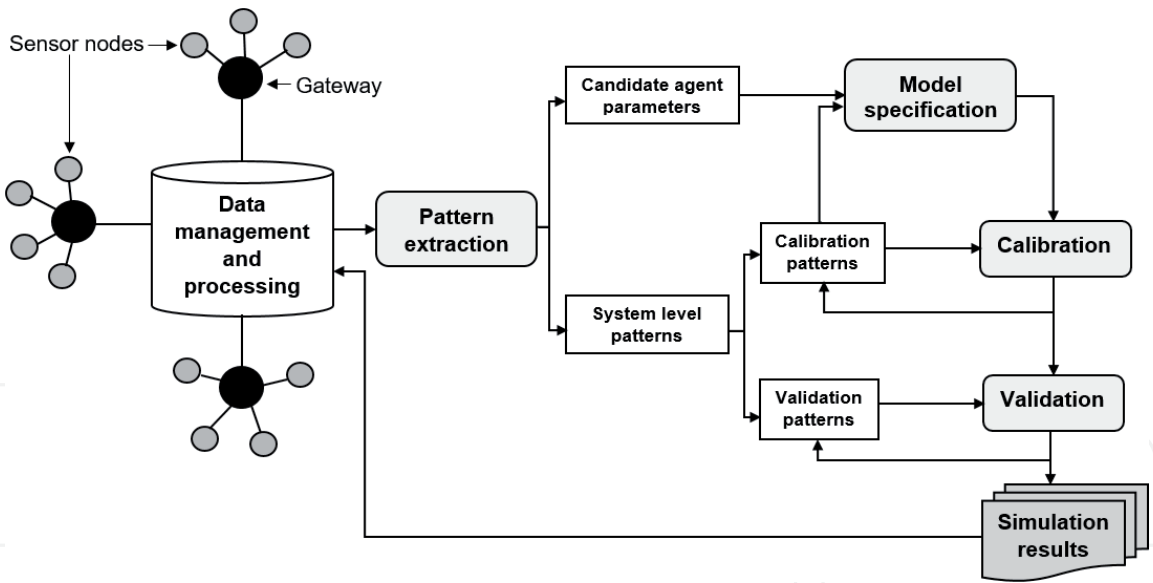


Figure 3. Proposed conceptual model for dynamic data-driven spatially explicit ABM. Geosensors capture data on the dynamics of the agents including their behaviour characteristics, space use and interactions. Preliminary on-site computation can be executed within the sensor networks before the data is parsed onto a data management and processing unit. Apart from storing the data in an efficient manner, the processing unit provides applications that can facilitate communication between the simulation model and the sensor network. The second process in the workflow involves exploratory data analysis and pattern extraction which results in an exhaustive list of system-level patterns and potential parameters that can be used to create a population of solutions for specifying the ABM. In addition, exploratory analysis reveals the tentative temporal steps at which it may be possible to identify micro level patterns of the system of interest. The patterns are separated into calibration and validation categories. The calibration patterns are hard-wired into the model through evolutionary optimisation of the candidate parameters. A specified model is calibrated dynamically and iteratively during the model. Once an adequate number of calibration steps have been executed, dynamic validation is initiated and is based on the reserve patterns that were not initially hard-wired into the model. The simulation outcome resulting from the validation stage is transferred to the data processing unit for assimilation with “fresh” data from the sensor observations. The process is cyclic and continues until the model accurately reproduces all the patterns in the data.

be adopted to guide the sensor selection and data collection processes. In addition, open standards that promote interoperability and transfer of sensor data and other resources should be encouraged. In particular, the use of Observation and Measurement (O&M) specification can facilitate both the documentation of data and uncertainties associated with the data. This is important for communicating the provenance of uncertainty throughout the modelling cycle.

For the data management, a distributed spatio-temporal database [114] is preferable when the study area is expansive and when there may be a need to carry out on-site quality assessment of the data from various sensor networks. Otherwise, a centralised Sensor Web Enablement (SWE) platform allowing for seamless discovery and manipulation and transfer of data and resources through standardised OGC compliant specifications is the most reliable. Examples of agent-oriented middleware for decentralised dynamic data collection include Sensomax [127], SenseWare [128] and MAPS [129]. Standardised data management systems facilitate characterisation of agent behaviours, multi-tasking and bidirectional communication between different components of the simulation workflow and the sensor nodes.

Apart from the geosensor data, additional spatial data from standard GIS data models and remote sensing products can be incorporated into the data management system to boost the characterisation of the environment in which the agents operate. For instance, when simulating dynamics of environmental changes, spatial data including human population and settlement, land use characteristics, topography, accessibility, vegetation indices, land surface temperature (LST), fire occurrence, night-time light, aerosols etc. can be combined with the in situ sensor data to provide a rich characterisation of the modelling world.

4.2 Exploratory analysis layer

The strength of data-driven models lies in the robust discovery of distinctive spatial and temporal patterns in the sensor data streams. Such patterns may be indicative of the essential processes and dynamics of the system of interest. The exploratory analysis is therefore a critical stage where statistical and machine-learning methods are applied to extract multi-scale patterns and other important characteristic parameters which may facilitate an accurate specification and simulation of the system behaviours. In situations where some knowledge has been documented concerning the system under study, then such information can guide and improve the pattern extraction process.

Statistical methods including multi-scale clustering and classification have been employed to reveal clusters in the data. For instance, in animal movement, Expectation–Maximization Binary Clustering (EMBC) [130] method has been applied to detect specific spatial and temporal navigation behaviours of birds. In human mobility trajectory analysis, DBSCAN clustering method has been applied to find traffic patterns [131]. Similarly, in flocking and swarm behaviour models, Spatial Clustering Algorithm Through Swarm Intelligence (SPARROW) clustering method has been used [132]. In addition, spatio-temporal data analysis methods including Bayesian spatio-temporal partitioning and clustering methods can be implemented to reveal the variation in behaviour of agents and the dynamics of the system in both space and time. Apart from statistical methods, machine-learning methods including convolutional neural networks (CNN), artificial neural network (ANN) and deep learning have been applied to reveal patterns. The use of mathematical and computational methods has the advantage that resulting patterns can be explicitly defined, hence building a mathematical or computational conceptualisation of such patterns [125]. The patterns can also provide a hint of agent processes that are inherent in the systems of interest.

In addition to the patterns, exploratory analysis process identifies essential parameters behind the patterns and processes of the system, allowing for specification of model parameters and identification of potential behaviour characteristics. Parameters are independent variables that influence the local-level behaviour of the agents. Related to the parameters, the exploratory analysis should also identify appropriate simulation schedules to facilitate the replication of all the necessary multi-scale patterns. An appropriate scheduling scheme also ensures the efficiency of the computation by informing a realistic temporal scale for the model and limiting unnecessary iteration of model runs.

Further, the exploratory analysis should also identify potential variables to be specified as the state variables of the agents. State variables are the agent-specific characteristics that vary dynamically in the model. State variables are essential as they provide a way of comparing the simulated agents against real-world agents while also providing a means of understanding how the local agent variables contribute to the multi-scale patterns. As an output from the exploratory process, a modeller should have an extensive list of potential model parameters and patterns that are essential for understanding the dynamics of the system. It is at this point that patterns should clearly be separated into the calibration and the validation patterns in preparation for their use in the dynamic simulation process.

4.3 Simulation layer

The simulation layer entails dynamic model specification, calibration and validation steps. As opposed to the conventional static ABMs, the specification, calibration and validation steps of a dynamic sensor-driven model can be implemented dynamically and iteratively during a single runtime and may run concurrently in a parallelised system.

In the model specification stage, the first step is to decide on a mechanism of combining or reducing the population of parameters into a robust set that can drive the essential behaviour of the agents. Evolutionary computation methods have been effective particularly in optimising ABM parameters [133]. One common example of evolutionary method for data-driven simulation is genetic algorithms [134]. In genetic algorithm, a random combination of the parameters can be created for each agent to provide the initial solution space [135]. The solution space evolves chromosomal crossover and mutation, which are critical operators of a genetic algorithm.

Further, to generate robust parameters, a proper fitness function should be derived to provide a basis of comparing the performance of the simulated agents against their real-world counterparts. A simple approach involves deriving fitness function a function of the variance between state variables and empirical agent characteristics. However, such a fitness function may increase the risk of model overfitting as the simulated agents are forced to replicate specific stepwise processes that are captured in the data. Consequently, deriving a fitness as a function from patterns may relax the focus from the state variables to the flexible multi-scale patterns. In addition, the combination of multiple patterns in defining fitness function can lead to generic and robust fitness functions. This is because patterns are generic spatial and temporal footprint that can be observed and described in the data.

Adopting a pattern-oriented modelling approach ensures that the process of model specification, calibration and validation is driven by the patterns that are inherent in the dynamic data. Consequently, this reduces the risk of tying the model parameters to the data, hence limiting the chance of model overfitting. In addition, since validation patterns are not explicitly hard-coded into the model, rigorously validated data-driven models can help in explaining the agent dynamics that lead to multi-scale patterns. The results of a properly calibrated and dynamically validated

model can be parsed to the central data management and processing unit for data assimilation and to influence the behaviour of the observation and measurement layer in cases where this is necessary. The cyclic communication between sensor observations and assimilation of simulation results bridges the gap between sensor data measurements and model specification and facilitates a mutually beneficial feedback between sensing unit and simulation model.

5. Outlook and potential applications of sensor data-driven spatially explicit ABMs

Advances in sensor technology, particularly the miniaturisation and ubiquity of sensors have led to an exponential growth in the diversity of the fine-scaled data, which can facilitate model specifications. In geographic information science, in situ sensor data provide accurate measurements of spatial entities and augment other data from earth observation workflows in characterising the environment in which agents interact. Sensor data therefore plays an important role in capturing the dynamics that cause spatial and temporal patterns. Accurate sensor data contributes towards understanding local-level interactions of humans, animals, firms, smart appliances and traffic, and the role of such interactions in global environmental changes. Similarly, the application of sensors has been a major driver in pervasive geographic information systems [136, 137] including in indoor environments and in the internet of things (IoT), technologies that are relevant for smart building and facility management.

However, advances in tools and software to support dynamic spatially explicit ABM specification have not been in tandem to the progress in sensor observation systems. Common ABM software including NetLogo, SWARM, MASON and Repast can handle only desktop-based geospatial data models. GAMA [138], which has the most extensive suite of tools for geospatial data handling and manipulation, does not have an equally extensive suite of APIs that can support dynamic data injection from sensor data streams, while FRAME and Repast HPC, which even though can support specification and simulation of distributed ABM, are not open and widely accessible. Moreover, most implementations of dynamic sensor-driven ABMs have been implemented to meet the objectives of specific projects and mainly in the computer science community and in the sensor (or geosensor) community. It is therefore important that modellers and practitioners in spatial simulation should develop reliable tools that can allow ABMs to be fed with rich sensor data streams from the systems of interest.

Potential areas of application of dynamic sensor-driven spatially explicit include animal ecology, human mobility studies and particularly in understanding mobility patterns and use of urban environment, energy use, indoor positioning systems, fire behaviour modelling, tourism research military applications, smart agriculture, environmental monitoring and in automation of industrial processes.

Epistemologically, the emergence of methods for data-driven ABMs raises questions on the place of conventional ABMs. In particular, do the data-driven models radically change the epistemological underpinnings of traditional ABM modelling framework? In other words, can accurate models be specified without relying on the domain knowledge and expertise of the modellers? To this question, a cautious approach should be encouraged. Whereas data-driven models are promising particularly in specifying models for theory-poor systems, a hybrid approach that starts from the domain knowledge and augments such knowledge with well-structured data-driven methods can improve reliability of agent-based simulation. Domain knowledge can provide the foundational understanding of a system of interest,

while rich and dynamic data can provide a means of discovering detailed local-level patterns and parameters of the system. In addition, results from data-driven models can augment domain knowledge. In a nutshell, data should help in defining the crisp concepts and in discovering hidden characteristics of the systems when these are not apparent in the domain knowledge. Consequently, as the data continues to grow in scale, accuracy and volume, while methods for big data analysis become more robust, data-driven models can be expected to grow and augment knowledge discovery in theory-poor domains. Sensor-driven spatially explicit ABMs therefore have an important role to play in understanding and representing dynamic spatial processes.

6. Conclusion

The aim of this paper was to trace and document the progress in the methods for specifying data-driven ABMs for spatial systems. In particular, the focus here has been on models that are fed with data from dynamic sensor data streams. It is clear from the documentation that advances in sensor and wireless communication technology have contributed immensely to the growth of data-driven ABMs. Data has been used in initialising, calibrating and validating models. However, traditionally, historical data has been fed into the models only sparingly without considering the dynamic changes in the real world. Though the conventional ABMs have been effective in generating hypothesis and representing dynamics of knowledge-rich systems, they have not been very applicable in addressing questions in complex and adaptive spatial systems whose internal dynamics are yet to be well understood. Moreover, the weakness of ABMs in predicting future states of systems persists. Designing accurate models for such systems can therefore be leveraged on the rich sensor data streams.

In this work, we proposed a framework for pattern-oriented, sensor-driven and spatially explicit ABM. In the framework, the steps of model specification, calibration and validation are implemented dynamically during the model run and are facilitated by patterns that can be derived dynamically from sensor data. This approach could contribute towards addressing the challenges of model overfitting that face most data-driven models. By validating models based on validation patterns that are not explicitly hard-coded into the model, the framework ensures that model parameters are not tied merely to the data but that the parameters and behaviours of the model can replicate patterns that are evident in the data. Most importantly, to promote efficient communication and management of sensor resources, we propose a service-oriented framework where sensor and network components are represented in the model as software agents in parallel to agents representing other real-world entities. This kind of arrangement allows well-known standards like the OGC standards of sensor specification to be applied in the modelling process, hence promoting discoverability and interoperability of sensor and model resources.

The main limitation of this review was in the fact that we did not include a prototype to demonstrate the practical application of the framework. However, examples can be seen in 4D-SAS [114] and DDDMAS application. Future research should include developing open and efficient tools that can promote distributed processing, simulation and visualisation of sensor-driven ABMs. Moreover, as behaviour specification has been one of the daunting tasks in typical ABM specification, automate discovery of algorithms from dynamic sensor data streams remains an exciting area of research that requires additional research. Robust methods for automated model discovery will improve the efficiency of data-driven spatial simulation models.

Epistemologically, sensor data-driven models raise important questions on the role of data in the specification of spatial simulation models. As geographic information science is an emerging field. It is our view that data-driven spatial simulation models will not only rely on the domain knowledge but will also contribute to methods of knowledge discovery in the field. As such, specification of ABMs can no longer merely rely on the domain knowledge but must be leveraged on the big data resources that are emerging from various advances in technology and computation. However, caution should be taken to allow a systematic development of data-driven methods in spatial simulation. Presently, we recommend a hybrid approaches that combine both domain knowledge and data-driven methods. Such models could be improved by relying on patterns that are extracted dynamically from sensor data streams.

Acknowledgements

This research was funded by the Austrian Science Fund (FWF) through the Doctoral College GIScience at the University of Salzburg (DK W 1237-N23).

Author details

Francis Oloo^{1,2}

1 Doctoral College of GIScience (DK GIScience), University of Salzburg, Salzburg, Austria

2 The Technical University of Kenya, Nairobi, Kenya

*Address all correspondence to: oloofrank@gmail.com

IntechOpen

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Bonabeau E. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*. 2002;**99**(Suppl 3):7280-7287
- [2] Janssen MA, Ostrom E. Empirically based, agent-based models. *Ecology and Society* 2006;**11**(2):37. Available from: <http://www.ecologyandsociety.org/vol11/iss2/art37/>
- [3] McLane AJ et al. The role of agent-based models in wildlife ecology and management. *Ecological Modelling*. 2011;**222**(8):1544-1556
- [4] An L. Modeling human decisions in coupled human and natural systems: Review of agent-based models. *Ecological Modelling*. 2012;**229**:25-36
- [5] Ajelli M et al. Spatiotemporal dynamics of the Ebola epidemic in Guinea and implications for vaccination and disease elimination: A computational modeling analysis. *BMC Medicine*. 2016;**14**(1):130
- [6] Alderton S et al. A multi-host agent-based model for a zoonotic, vector-borne disease. A case study on Trypanosomiasis in Eastern Province, Zambia. *PLoS Neglected Tropical Diseases*. 2016;**10**(12):e0005252
- [7] Venkatramanan S et al. Using data-driven agent-based models for forecasting emerging infectious diseases. *Epidemics*. 2018;**22**:43-49
- [8] Fang Z et al. Agent-based simulation analysis on the effect of an LNG terminal on a port transport system. In: *CICTP 2014: Safe, Smart, and Sustainable Multimodal Transportation Systems—Proceedings of the 14th COTA International Conference of Transportation Professionals*. 2014
- [9] Fernández-Isabel A, Fuentes-Fernández R. Analysis of intelligent transportation systems using model-driven simulations. *Sensors (Switzerland)*. 2015;**15**(6):14117-14141
- [10] Jensen T, Chappin ÉJL. Automating agent-based modeling: Data-driven generation and application of innovation diffusion models. *Environmental Modelling and Software*. 2017;**92**:261-268
- [11] Castella JC, Verburg PH. Combination of process-oriented and pattern-oriented models of land-use change in a mountain area of Vietnam. *Ecological Modelling*. 2007;**202**(3-4):410-420
- [12] Hanappi G. Agent-based modelling. History, essence, future. *PSL Quarterly Review*. 2017;**70**(283):449-472
- [13] Bruch E, Atwell J. Agent-based models in empirical social research. *Sociological Methods & Research*. 2015;**44**(2):186-221
- [14] MacBride RAF. *Applying Systems Science Methods to Risk-Based Disease Management and Population Oral Health*. Los Angeles: University of California; 2018
- [15] Drchal J, Čertický M, Jakob M. Data driven validation framework for multi-agent activity-based models. In: *International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Springer; 2015
- [16] Arifin S, Davis GJ, Zhou Y. Modeling space in an agent-based model of malaria: Comparison between non-spatial and spatial models. In: *Proceedings of the 2011 Workshop on Agent-Directed Simulation*. Society for Computer Simulation International; 2011
- [17] Nittel S. A survey of geosensor networks: Advances in dynamic environmental monitoring. *Sensors*. 2009;**9**(7):5664-5678

- [18] Bergmann JH et al. Exploring the use of sensors to measure behavioral interactions: An experimental evaluation of using hand trajectories. *PLoS One*. 2014;**9**(2):e88080
- [19] Harari GM et al. Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspectives on Psychological Science*. 2016;**11**(6):838-854
- [20] Niazi MA, Hussain A. A novel agent-based simulation framework for sensing in complex adaptive environments. *IEEE Sensors Journal*. 2011;**11**(2):404-412
- [21] Darema F. Dynamic data driven applications systems: New capabilities for application simulations and measurements. In: *International Conference on Computational Science*. Springer; 2005
- [22] Ha SW et al. An environmental monitoring system for managing spatiotemporal sensor data over sensor networks. *Sensors*. 2012;**12**(4):3997-4015
- [23] Bröring A et al. New generation sensor web enablement. *Sensors*. 2011;**11**(3):2652-2699
- [24] Saha D, Mukherjee A. Pervasive computing: A paradigm for the 21st century. *Computer*. 2003;**36**(3):25-31
- [25] Want HWGR, Schmidt A. Pervasive computing. *IEEE*. 2006;**5**:25-33
- [26] Anselin L. What is special about spatial data? *Alternative Perspectives on Spatial Data Analysis* (89-4). 1989
- [27] Liu Y et al. Towards a rich-context participatory cyber environment. *International Workshop on Grid Computing environment*. Reno, NV, USA; 2007
- [28] Hart G, Dolbear C. What's so special about spatial? In: *The Geospatial Web*. Springer; 2009. pp. 39-44
- [29] Gotelli NJ et al. Patterns and causes of species richness: A general simulation model for macroecology. *Ecology Letters*. 2009;**12**(9):873-886
- [30] O'Sullivan D, Perry GL. *Spatial Simulation: Exploring Pattern and Process*. West Sussex, UK: John Wiley & Sons; 2013
- [31] Rodrigues A, Raper J. Defining spatial agents. In: Taylor, Francis Editors. London, United Kingdom: *Spatial Multimedia and Virtual Reality*; 1999:111-129
- [32] Itami RM. *Mobile Agents with Spatial Intelligence*. Oxford, UK: Oxford University Press; 2002
- [33] Sajjad M et al. A data-driven approach for agent-based modeling: Simulating the dynamics of family formation. *Journal of Artificial Societies and Social Simulation*. 2016;**19**(1):9
- [34] Brown DG et al. Path dependence and the validation of agent-based spatial models of land use. *International Journal of Geographical Information Science*. 2005;**19**(2):153-174
- [35] Poza DJ et al. Mesoscopic effects in an agent-based bargaining model in regular lattices. *PLoS One*. 2011;**6**(3):e17661
- [36] Fonoberova M et al. An agent-based model of urban insurgence: Effect of gathering sites and Koopman mode analysis. *PLoS One*. 2018;**13**(10):e0205259
- [37] Filatova T et al. Spatial agent-based models for socio-ecological systems: Challenges and prospects.

Environmental Modelling & Software. 2013;**45**:1-7

systems: Lessons from ecology. Science. 2005;**310**(5750):987-991

[38] Brown DG et al. Spatial process and data models: Toward integration of agent-based models and GIS. Journal of Geographical Systems. 2005;**7**(1):25-47

[47] Ormerod P, Rosewell B. Validation and Verification of Agent-Based Models in the Social Sciences. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009

[39] Nagy M et al. Hierarchical group dynamics in pigeon flocks. Nature. 2010;**464**(7290):890

[48] Niazi MA, Hussain A, Kolberg M. Verification & validation of agent based simulations using the VOMAS (virtual overlay multi-agent system) approach. arXiv preprint arXiv:1708.02361. 2017

[40] Vhaduri S et al. Estimating drivers' stress from GPS traces. In: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM; 2014

[49] Windrum P, Fagiolo G, Moneta A. Empirical validation of agent-based models: Alternatives and prospects. Journal of Artificial Societies and Social Simulation. 2007;**10**(2):8

[41] Al-Husain L, Kanjo E, Chamberlain A. Sense of space: Mapping physiological emotion response in urban space. In: Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication. ACM; 2013

[50] Augusiak J, Van den Brink PJ, Grimm V. Merging validation and evaluation of ecological models to 'evaludation': A review of terminology and a practical approach. Ecological Modelling. 2014;**280**:117-128

[42] Pharow P et al. Portable devices. In: Sensors and Networks: Wireless Personalized eHealth Services. Amsterdam, Netherlands: IOS Press; 2009

[51] Grimm V et al. Towards better modelling and decision support: Documenting model development, testing, and analysis using TRACE. Ecological Modelling. 2014;**280**:129-139

[43] Grimm V, Railsback SF. Individual-based Modeling and Ecology. Princeton NJ, USA: Princeton University Press; 2013

[52] Müller B et al. Standardised and transparent model descriptions for agent-based models: Current status and prospects. Environmental Modelling & Software. 2014;**55**:156-163

[44] Barnes DJ, Chu D. Guide to Simulation and Modeling for Biosciences. London, UK: Springer; 2015

[53] Grimm V, Schmolke A. How to read and write TRACE documentations. Leipzig, Germany: Helmholtz Centre for Environmental Research-UFZ. 2011 Available from: <http://cream-itn.eu/creamwp/wp-content/uploads/Trace-Guidance-11-03-04.pdf>

[45] Marschak J. Economic measurements for policy and prediction. In: Economic Information, Decision, and Prediction. Dordrecht, Netherlands: Springer; 1974. pp. 293-322

[46] Grimm V et al. Pattern-oriented modeling of agent-based complex

[54] Grimm V et al. The ODD protocol: A review and first update. Ecological Modelling. 2010;**221**(23):2760-2768

- [55] Grimm V et al. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*. 2006;**198**(1-2):115-126
- [56] Arroyo J et al. Re-thinking simulation: A methodological approach for the application of data mining in agent-based modelling. *Computational and Mathematical Organization Theory*. 2010;**16**(4):416-435
- [57] Conte R, Paolucci M. On agent-based modeling and computational social science. *Frontiers in Psychology*. 2014;**5**:668
- [58] Flack A et al. Costs of migratory decisions: A comparison across eight white stork populations. *Science Advances*. 2016;**2**(1):e1500931
- [59] Ilie S, Bădică C. Multi-agent approach to distributed ant colony optimization. *Science of Computer Programming*. 2013;**78**(6):762-774
- [60] Ilie S, Bădică C. Multi-agent distributed framework for swarm intelligence. *Procedia Computer Science*. 2013;**18**:611-620
- [61] Beekman M, Sword GA, Simpson SJ. Biological foundations of swarm intelligence. In: *Swarm Intelligence*. Berlin, Heidelberg, Germany: Springer; 2008. pp. 3-41
- [62] Bonabeau E, Corne D, Poli R. Swarm intelligence: The state of the art special issue of natural computing. *Natural Computing*. 2010;**9**(3):655-657
- [63] Mishra BSP, Dehuri S, Wang G-N. A state-of-the-art review of artificial bee colony in the optimization of single and multiple criteria. *International Journal of Applied Metaheuristic Computing (IJAMC)*. 2013;**4**(4):23-45
- [64] Ahmad R et al. A multi-agent based approach for particle swarm optimization. In: 2007 International Conference on Integration of Knowledge Intensive Multi-Agent Systems. IEEE; 2007
- [65] Bertsimas D, Jaillet P, Martin S. Online vehicle routing: The edge of optimization in large-scale applications. *Operations Research*. 2019;**67**(1):143-162
- [66] Bai Q. Analysis of particle swarm optimization algorithm. *Computer and Information Science*. 2010;**3**(1):180
- [67] Bürkle A, Segor F, Kollmann M. Towards autonomous micro uav swarms. *Journal of Intelligent & Robotic Systems*. 2011;**61**(1-4):339-353
- [68] Hassan S et al. Stepping on earth: A roadmap for data-driven agent-based Modelling. In: *Proceedings of the 5th Conference of the European Social Simulation Association (ESSA08)*. 2008
- [69] Epstein JM. Agent-based computational models and generative social science. *Complexity*. 1999;**4**(5):41-60
- [70] Heath B, Hill R, Ciarallo F. A survey of agent-based modeling practices (January 1998 to July 2008). *Journal of Artificial Societies and Social Simulation*. 2009;**12**(4):9
- [71] Kavak H et al. Big data, agents, and machine learning: Towards a data-driven agent-based modeling approach. In: *Proceedings of the Annual Simulation Symposium*. Society for Computer Simulation International; 2018
- [72] Zhong J et al. Learning behavior patterns from video for agent-based crowd modeling and simulation. *Autonomous Agents and Multi-Agent Systems*. 2016;**30**(5):990-1019
- [73] Zhang H et al. Data-driven agent-based modeling, with application

to rooftop solar adoption. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems. Istanbul, Turkey: International Foundation for Autonomous Agents and Multiagent Systems; 2015. pp. 513-521

[74] Pruyt E et al. From data-poor to data-rich: System dynamics in the era of big data. In: 32nd International Conference of the System Dynamics Society, Delft, The Netherlands, 20-24 July 2014; Authors Version. The System Dynamics Society; 2014

[75] Michopoulos J et al. Agent-based simulation of data-driven fire propagation dynamics. In: International Conference on Computational Science. Springer; 2004

[76] Goodchild MF. Geographic information systems and science: Today and tomorrow. *Annals of GIS*. 2009;15(1):3-9

[77] Nadi S, Delavar MR. Spatio-temporal modeling of dynamic phenomena in GIS. In: ScanGIS, Espoo, Finland; 2003. pp. 215-225

[78] Wang F, Zhou C, Nie Y. Event processing in sensor streams. In: Managing and Mining Sensor Data. Boston MA, USA: Springer; 2013. pp. 77-102

[79] Song J et al. Application of dynamic data driven application system in environmental science. *Environmental Reviews*. 2014;22(3):287-297

[80] Nittel S et al. Shared ride trip planning with geosensor networks. In: Societies and Cities in the Age of Instant Access. Dordrecht, Netherlands: Springer; 2007. pp. 179-194

[81] Sagl G, Resch B, Blaschke T. Contextual sensing: Integrating contextual information with human and technical geo-sensor

information for smart cities. *Sensors*. 2015;15(7):17013-17035

[82] Hernández-Pliego J et al. Combined use of tri-axial accelerometers and GPS reveals the flexible foraging strategy of a bird in relation to weather conditions. *PLoS One*. 2017;12(6):e0177892

[83] Júnior JF et al. Driver behavior profiling: An investigation with different smartphone sensors and machine learning. *PLoS One*. 2017;12(4):e0174959

[84] Sahli N, Jabeur N, Badra M. Agent-based framework for sensor-to-sensor personalization. *Procedia Computer Science*. 2013;19:197-205

[85] Dagdeviren O et al. A survey of agent technologies for wireless sensor networks. *IETE Technical Review*. 2011;28(2):168-184

[86] Wang Z. Integrating spatio-temporal data into agent-based simulation for emergency navigation support. *GIST Rapport*, No. 58. 2012

[87] Kennedy C et al. AIMSS: An architecture for data driven simulations in the social sciences. In: International Conference on Computational Science. Berlin Heidelberg, Germany: Springer; 2007

[88] Ranjit S et al. Agent-based modeling of taxi behavior simulation with probe vehicle data. *ISPRS International Journal of Geo-Information*. 2018;7(5):177

[89] Bonnell TR et al. Emergent group level navigation: An agent-based evaluation of movement patterns in a folivorous primate. *PLoS One*. 2013;8(10):e78264

[90] Scheutz M, Mayer T. Combining agent-based modeling with big data methods to support architectural and urban design. In: Understanding

Complex Urban Systems. Cham, Switzerland: Springer; 2016. pp. 15-31

[91] Hu X. Dynamic data driven simulation. SCS M&S Magazine. 2011;5:16-22

[92] Rodríguez R, Cortés A, Margalef T. Injecting dynamic real-time data into a DDDAS for forest fire behavior prediction. In: International Conference on Computational Science. Springer; 2009

[93] Allen G. Building a dynamic data driven application system for hurricane forecasting. In: International Conference on Computational Science. Springer; 2007

[94] Darema F. Dynamic Data Driven Applications Systems: A New Paradigm for Application Simulations and Measurements. Berlin, Heidelberg: Springer Berlin Heidelberg; 2004

[95] Fujimoto R et al. A Dynamic Data Driven Application System for Vehicle Tracking. Procedia Computer Science; 2014;29(1):1203-1215

[96] Castrillón M et al. Forecasting and visualization of wildfires in a 3D geographical information system. Computers & Geosciences. 2011;37(3):390-396

[97] Heppenstall A, Malleson N, Crooks A. "Space, the final frontier": How good are agent-based models at simulating individuals and space in cities? Systems. 2016;4(1):9

[98] Rai S, Hu X. Behavior pattern detection for data assimilation in agent-based simulation of smart environments. In: Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013 IEEE/WIC/ACM International Joint Conferences on. IEEE; 2013

[99] Xue H, Gu F, Hu X. Data assimilation using sequential Monte Carlo methods in wildfire spread

simulation. ACM Transactions on Modeling and Computer Simulation (TOMACS). 2012;22(4):23

[100] Ward JA, Evans AJ, Malleson NS. Dynamic calibration of agent-based models using data assimilation. Royal Society Open Science. 2016;3(4):150703

[101] Rai S, Hu X. Data assimilation with sensor-informed resampling for building occupancy simulation. In: Proceedings of the 2017 Winter Simulation Conference. Las Vegas, Nevada: IEEE Press; 2017. pp. 1-12

[102] Nguyen HM et al. An alternative approach to avoid overfitting for surrogate models. In: Proceedings of the Winter Simulation Conference. Winter Simulation Conference; 2011

[103] Kashif A et al. Agent based Framework to Simulate Inhabitants' Behaviour in Domestic Settings for Energy Management. In: Proceedings of 3rd International Conference of Agents and Artificial Intelligence. Rome, Italy; 2011. pp. 190-199

[104] Calvez B, Hutzler G. Automatic tuning of agent-based models using genetic algorithms. In: International Workshop on Multi-Agent Systems and Agent-Based Simulation. Springer; 2005

[105] Oloo F, Wallentin G. An adaptive agent-based model of homing pigeons: A genetic algorithm approach. ISPRS International Journal of Geo-Information. 2017;6(1):27

[106] Chu X, Buyya R. Service oriented sensor web. In: Sensor networks and Configuration. Berlin, Heidelberg, Germany: Springer; 2007. pp. 51-74

[107] Tan X et al. Agent-as-a-service-based geospatial service aggregation in the cloud: A case study of flood response. Environmental Modelling & Software. 2016;84:210-225

- [108] Tang W et al. Agent-based modeling within a cyberinfrastructure environment: A service-oriented computing approach. *International Journal of Geographical Information Science*. 2011;25(9):1323-1346
- [109] Vinyals M, Rodriguez-Aguilar JA, Cerquides J. A survey on sensor networks from a multiagent perspective. *The Computer Journal*. 2011;54(3):455-470
- [110] Fachada N et al. Parallelization strategies for spatial agent-based models. *International Journal of Parallel Programming*. 2017;45(3):449-481
- [111] Parry HR, Bithell M. Large scale agent-based modelling: A review and guidelines for model scaling. In: *Agent-based Models of Geographical Systems*. Springer; 2012. pp. 271-308
- [112] Nichols JA, Hallam TG, Dimitrov DT. Parallel simulation of ecological structured communities: Computational needs, hardware capabilities, and nonlinear applications. *Nonlinear Analysis: Theory, Methods & Applications*. 2008;69(3):832-842
- [113] Tang W, Wang S. HPABM: A hierarchical parallel simulation framework for spatially-explicit agent-based models. *Transactions in GIS*. 2009;13(3):315-333
- [114] Li Z et al. 4D-SAS: A distributed dynamic-data driven simulation and analysis system for massive spatial agent-based modeling. *ISPRS International Journal of Geo-Information*. 2016;5(4):42
- [115] Shook E, Wang S, Tang W. A communication-aware framework for parallel spatially explicit agent-based models. *International Journal of Geographical Information Science*. 2013;27(11):2160-2181
- [116] Gong Z et al. Parallel agent-based simulation of individual-level spatial interactions within a multicore computing environment. *International Journal of Geographical Information Science*. 2013;27(6):1152-1170
- [117] Pereira GM. Dynamic data driven multi-agent simulation. In *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT 2007*. 2008
- [118] Xiaoyingjie X, Zhanghao Z, Lisong L. Dynamic data driven multi-agent simulation in maritime traffic. In: *Proceedings—2009 International Conference on Computer and Automation Engineering, ICCAE 2009*. 2009
- [119] Hernandez S et al. From on-body sensors to in-body data for health monitoring and medical robotics: A survey. In: *2014 Global Information Infrastructure and Networking Symposium (GIIS)*. 2014
- [120] Darema F. DDDAS, a key driver for large-scale-big-data and large-scale-big-computing. *Procedia Computer Science*. 2015;51:2463
- [121] Van Berkel S et al. Automatic discovery of algorithms for multi-agent systems. In: *GECCO'12—Proceedings of the 14th International Conference on Genetic and Evolutionary Computation Companion*. 2012
- [122] Sottolare RA. Using learner data to influence performance during adaptive tutoring experiences. In: *International Conference on Augmented Cognition*. Springer; 2014
- [123] Shin S et al. Design and analysis of cost-efficient sensor deployment for tracking small UAS with agent-based modeling. *Sensors*. 2016;16(4):575
- [124] Yamins D, Nagpal R. Automated global-to-local programming in 1-d spatial multi-agent systems. In: *Proceedings of the 7th International*

Joint Conference on Autonomous Agents and Multiagent Systems. Vol. 2. International Foundation for Autonomous Agents and Multiagent Systems; 2008

[125] Keller N, Hu X. Data driven simulation modeling for mobile agent-based systems. In: Proceedings of the 2016 Spring Simulation Multiconference—TMS/DEVS Symposium on Theory of Modeling and Simulation, TMS/DEVS 2016. 2016

[126] Grimm V, Railsback SF. Agent-based models in ecology: Patterns and alternative theories of adaptive behaviour. In: Agent-based Computational Modelling. Heidelberg, Germany: Springer; 2006. pp. 139-152

[127] Haghighi M, Cliff D. Sensomax: An agent-based middleware for decentralized dynamic data-gathering in wireless sensor networks. In: 2013 International Conference on Collaboration Technologies and Systems (CTS). 2013

[128] Boulis A, Han C-C, Srivastava MB. Design and implementation of a framework for efficient and programmable sensor networks. In: Proceedings of the 1st International Conference on Mobile Systems, Applications and Services. ACM; 2003

[129] Aiello F et al. Maps: A mobile agent platform for wsns based on java sun spots. Proceedings of ATSM. 2009

[130] Garriga J et al. Expectation-maximization binary clustering for behavioural annotation. PLoS One. 2016;**11**(3):e0151984

[131] Necula E. Analyzing traffic patterns on street segments based on GPS data using R. Transportation Research Procedia. 2015;**10**:276-285

[132] Folino G, Forestiero A, Spezzano G. An adaptive flocking algorithm for performing approximate clustering. Information Sciences. 2009;**179**(18):3059-3078

[133] Jakob M et al. Agents vs. pirates: Multi-agent simulation and optimization to fight maritime piracy. In: 11th International Conference on Autonomous Agents and Multiagent Systems 2012, AAMAS 2012: Innovative Applications Track. 2012

[134] Oremland M, Laubenbacher R. Optimization of agent-based models: Scaling methods and heuristic algorithms. Journal of Artificial Societies and Social Simulation. 2014;**17**(2):6

[135] Heppenstall AJ, Evans AJ, Birkin MH. Genetic algorithm optimisation of an agent-based model for simulating a retail market. Environment and Planning B: Planning and Design. 2007;**34**(6):1051-1070

[136] Zambonelli F et al. Developing pervasive multi-agent systems with nature-inspired coordination. Pervasive and Mobile Computing. 2015;**17**:236-252

[137] Shibuya K. A framework of multi-agent-based modeling, simulation, and computational assistance in an ubiquitous environment. Simulation. 2004;**80**(7-8):367-380

[138] Taillandier P et al. GAMA: A simulation platform that integrates geographical information data, agent-based modeling and multi-scale control. In: International Conference on Principles and Practice of Multi-Agent Systems. Springer; 2010