

TR-QC-05-2014

Review of time-based techniques for modelling space

Revision: 1.0; April 8, 2014

Author(s): Vashti Galpin (UEDIN), Cheng Feng (UEDIN), Jane Hillston (UEDIN), Mieke Massink (CNR), Mirco Tribastone (SOTON), Max Tschaikowski (SOTON)

Publication date: 8 April 2014

Funding Scheme: Small or medium scale focused research project (STREP)

Topic: ICT-2011 9.10: FET-Proactive 'Fundamentals of Collective Adaptive Systems' (FOCAS)

Project number: 600708

Coordinator: Jane Hillston (UEDIN)

e-mail: Jane.Hillston@ed.ac.uk

Fax: +44 131 651 1426

Part. no.	Participant organisation name	Acronym	Country
1 (Coord.)	University of Edinburgh	UEDIN	UK
2	Consiglio Nazionale delle Ricerche – Istituto di Scienza e Tecnologie della Informazione "A. Faedo"	CNR	Italy
3	Ludwig-Maximilians-Universität München	LMU	Germany
4	Ecole Polytechnique Fédérale de Lausanne	EPFL	Switzerland
5	IMT Lucca	IMT	Italy
6	University of Southampton	SOTON	UK
7	Institut National de Recherche en Informatique et en Automatique	INRIA	France

Abstract

This document provides details about different mathematical techniques for modelling space and time. It first describes a classification of space modelling techniques, covering two time-based approaches and one space-only approach, that of topological and metric spaces. It then goes into detail about discrete space approaches providing basic definitions and considering how such techniques have been used in the ecology, biology, epidemiological and networking literature. Following this, continuous space techniques are considered.

Contents

1	Introduction	3
2	A classification of space and movement modelling techniques	3
2.1	Notation	4
2.2	Non-spatial modelling dimensions	4
2.3	Time-based non-spatial modelling techniques	5
2.4	Introducing space	6
2.5	Spatial modelling techniques	8
2.6	Movement	11
2.6.1	Discrete space	11
2.6.2	Continuous space, no aggregation	12
2.6.3	Continuous space, aggregation	12
2.7	Conclusion	12
3	Discrete space modelling techniques	12
3.1	Spatial homogeneity	13
3.2	Spatial regularity	14
3.3	Neighbours and neighbourhoods	15
3.4	Boundary conditions	15
3.5	Discrete space without state-based aggregation	16
3.6	Discrete space with state-based aggregation	16
3.7	Examples of existing use	18
3.7.1	Ecology	19
3.7.2	Biology	21
3.7.3	Epidemiology	22
3.7.4	Networking	24
3.7.5	Other	25
4	Continuous space modelling techniques	26
4.1	Continuous space without state-based aggregation	26
4.2	Continuous space with state-based aggregation	26
4.3	Examples of existing use	27
4.3.1	Ecology	27
4.3.2	Biology	28
4.3.3	Epidemiology	29
4.3.4	Networking	30
5	Conclusion	30
	References	31

1 Introduction

The goal of this document is to provide further details of the various time-based spatial modelling techniques that are referenced in Deliverable 2.1 [GBC⁺14]. For this reason, the next section on classification is almost identical to that of the deliverable. The sections on discrete space techniques and continuous space techniques follow the same structure as in D2.1 (except that they do not have sections about the case studies). Topological space is not time-based, so is not covered here but additional details over and above those given in Deliverable 2.1 can be found in QUANTICOL technical report TR-QC-01-2014 [CLM14].

This reports focusses on the following details.

- General definition of the approaches that fall into the categories defined by considering aggregation versus none, discrete space versus continuous space, and where appropriate, subcategories of discrete space.
- Description of the approaches found across various disciplines including ecology, biology, epidemiology and others, both generally and in terms of specific examples of use.

The report does not attempt to fully cover the following issues, some of which are for further work in Task 2.1.

- Classification of examples in terms of the types of homogeneity introduced in this document: location, transfer, (spatial) parameter and spatial homogeneity.
- Identification of multiscale aspects of existing uses including whether in patch models, rates are faster with patches than those between patches or *vice versa*.
- Assumptions in the modelling techniques that make them suitable for a particular application, although some assumptions are mentioned.
- Costs of analysis as this topic will be considered in future documents of Task 2.1.

The goal of this report is to capture a flavour of the existing literature to do with spatial modelling, so that the process of reading and summarising the literature does not need to be repeated later in the project.

2 A classification of space and movement modelling techniques

This section provides a framework in which to understand the various choices that can be made in terms of spatial modelling techniques. There exists a rich literature about these techniques and it is necessary to understand this before making decisions relating to the spatial aspects of the languages and their semantics in the QUANTICOL project.

We are interested in modelling the collective behaviour of many individuals using a stochastic approach (based on Markov chains that are discrete time or continuous time) and we use mean-field techniques¹ to obtain a fluid or deterministic approximation of their behaviour [Kur81]. We start from information about how each individual changes state over time, and we are interested in how this affects the overall behaviour of the system over time, either obtaining an exact stochastic analysis of the system, or a deterministic approximation of its behaviour.

Our specific modelling goals are determined by the case studies in the QUANTICOL project which are smart transport and smart grids. Choice of modelling technique should be informed by the sort of questions for which answers are required. The requirements section in Deliverable 5.1 [TCG⁺14] considers some of these questions and others will be revealed as the project continues and properties for verification by spatial model-checking are considered. Furthermore, existing models of case studies are discussed in Deliverable 2.1 [GBC⁺14] and the goal of each model is identified.

¹See the tutorial [BHL13] and the references therein for more details of the mean-field approach.

2.1 Notation

To set the scene, some notation will be useful. For illustrative purposes in this document, we consider two *populations* P_A and P_B . At each point in time, each individual in P_A is in exactly one of the *states* A_1, \dots, A_n and each individual in P_B is in exactly one of the states B_1, \dots, B_m . We can consider the populations in terms of individuals (we may need a naming convention to be able to refer to each individual) or in aggregation by counting the number of individuals in each state. This is called a *state-based aggregation* and it provides a view of the number of individuals in a population in a particular state, for each population and each state allowed for that population. Thus, let $N_{A_i}(t)$ refer to the number of individuals in population P_A that are in state A_i at time t and let $N_{B_j}(t)$ refer to the number of individuals in population P_B that are in state B_j at time t . These will be called *subpopulations*. The total number of individuals in the two populations at time t can be expressed as $N_A(t) = \sum_{i=1}^n N_{A_i}(t)$ and $N_B(t) = \sum_{j=1}^m N_{B_j}(t)$ respectively. Clearly, these counts are in \mathbb{N} (which includes zero). Furthermore, if no births or deaths are assumed, and an individual must be in one of the available states², then $N_A(t_1) = N_A(t_2)$ for all times t_1 and t_2 and the size of P_A is a constant N_A ; similarly the size of P_B is N_B .

To complete the notation required for populations, we use $X_{A_i}(t) \in \mathbb{R}_{\geq 0}$ to represent a non-negative real-valued description of the population P_A which in certain approaches is an approximation to $N_{A_i}(t)$ and similarly for $X_{B_i}(t) \in \mathbb{R}_{\geq 0}$ for P_B .

2.2 Non-spatial modelling dimensions

Before space is considered, there are already a number of choices that lead to different approaches to modelling dynamic systems in a quantified manner. We now consider the dimensions and the choices on each dimension as informed by the prior research of members of the project into dynamic modelling of systems [BHL13]. For example, the time dimension considers how time is treated in different types of Markov chains. There are other aspects of time such as non-determinism and causality, but these are not a strong focus of our general modelling approach, and so are not included in the classification.

Time: Time is usually non-negative, strictly increasing and infinite, and can either be a non-negative real or integer. In some models, a finite end-point may be used to delimit the period of interest.

discrete: In the context of this research, discrete time is used in those modelling approaches where choices are probabilistic. At each clock tick (which could be associated with an integer if useful for the specific model), each individual chooses probabilistically its next state. For example, discrete time Markov chains (DTMCs) use this approach [KS76, Nor98].

continuous: Here, time is continuous and this is captured by the fact that actions such as changing state have a duration associated with them. In the case of continuous time Markov chains (CTMCs), stochasticity is introduced by having random durations that are drawn from exponential distributions [Nor98].

State: States can be viewed as capturing a quality or attribute of an individual. As described above, an individual is assumed to be in a single state at each point in time³.

discrete: Usually when the states associated with an individual are discrete, there are a finite number of them. However, in the case of an attribute like *year-of-birth*, there may be a countably infinite number of values.

²In some models, births and deaths can be included for a fixed size population by introducing a “dead” state. However, this requires that there is a finite maximum population size.

³An individual could have more than one attribute, and then the individual’s state is multidimensional with a value for each attribute. In this case, the individual’s state can be seen as a tuple of values.

TIME	discrete			
AGGR	none		state	
STATE	discrete	continuous	discrete	continuous
	DTMC (discrete time Markov chain) [Nor98]	LMP (labelled Markov process) [Pan09]	population DTMC [BHLM13]	difference equations, ordinary differential equations (ODEs) [BHLM13, MNS11]

TIME	continuous			
AGGR	none		state	
STATE	discrete	continuous	discrete	continuous
	CTMC (continuous time Markov chain) [Nor98]	CTMP (continuous time Markov process) [DP03]	population CTMC [BHLM13, Kur81]	population ODEs [BHLM13, Kur81]

Figure 1: Classification of mathematical models in terms of time, aggregation and state

continuous: A continuous-valued state can be interpreted as measurement of some quantity associated with the individual. An example of this would be *temperature* or *height*.

Aggregation: As discussed previously, individuals can be considered separately, or the focus can be on the number of individuals in each state. This is more relevant to discrete state approaches than continuous state. In the continuous case, aggregation can be described by a function, or discretisation can be applied to obtain frequency data.

none: Behaviour of each individual is considered separately.

state-based: The behaviour of groups of individuals is considered by counting the number of individuals in each state over time (giving a non-negative integer value), or by having a non-negative real-valued approximation to this number. This approach appears under a number of different names in the literature including *population-based*, *state frequency data*, *numerical vector form*, and *counting abstraction*. The term *occupancy measure* is used when counts are normalised by the population size.

These possibilities can be expressed in a table, which can then be populated with mathematical modelling techniques from the literature. Figure 1 illustrates this and describes the modelling techniques that fit each combination of each element of each dimension.

2.3 Time-based non-spatial modelling techniques

An important aspect of our prior research is the application of the mean-field technique where the analysis of a population CTMC or DTMC can be approximated by an analysis using ordinary differential equations (ODEs) [Kur81, BHLM13]. As the number of states of a Markov chain increases (the “state-space explosion” problem), the analysis of the Markov chain becomes intractable. Modelling a large number of individuals can lead to a very large Markov chain. This can be mitigated by using a population Markov chain where behaviour is considered at a population level rather than at an individual level. The choice of a population Markov chain means we are interested in how many individuals

are in each state i , given by N_{A_i} , and the states in the Markov chain have the form $(N_{A_1}, \dots, N_{A_n})$. However, for large systems this may still not be sufficient to obtain reasonable analysis times, and an approximation using ODEs obtained from the population Markov chain can be used. This gives a system of ODEs for the variables $(X_{A_1}, \dots, X_{A_n})$. This population Markov chain considers non-negative integer-valued population counts whereas the ODEs take a fluid approach and population quantities are non-negative real values. Considering the modelling techniques in Figure 1 for both discrete time and continuous time, the Markov chain obtained by considering many individuals (in the first column) can be transformed to a smaller Markov chain (in the third column) which can then be approximated by ODEs (in the fourth column).

This last transformation uses the mean-field approximation technique which comes originally from physics, where it refers to the approach where the movement of an individual particle is considered in the field generated by other particles rather than trying to solve the more complex problem of many particles interacting [CL07, MP12]. In modelling of systems, it has come to mean an approach where it is assumed that the number of individuals in a stochastic system becomes very large so that the population-level behaviour of the system can be expressed as ODEs which provide an “average” behaviour of the system. Results such as those proved by Kurtz [Kur81] demonstrate that under certain conditions, convergence occurs, namely as the number of individuals tends to infinity, the difference between the stochastic trajectories of the subpopulation sizes and the deterministic trajectories of the subpopulation sizes tends to zero. Practically, in many cases, good approximations using the ODE approach over the stochastic approach can be achieved at relatively low numbers of individuals [TGH12] and there are error bounds on the approximations [DN08].

Markov processes (in the second column) do not fit into this work flow and seem different from the other modelling techniques, as they are characterised by a continuous state space (which can also be interpreted as any continuous aspect of a model, including space).

2.4 Introducing space

Space can be considered in different ways.

continuous: Here, space is represented by real values in the case of one-dimensional space, pairs of real values in the two-dimensional case and triples of real values in the three-dimensional case. It is always (uncountably) infinite but may be bounded in extent. Continuous space used in this way can be seen as an exact representation of actual physical space.

discrete: Approaches that use discrete space assume a number (usually finite) of distinct locations where connectivity between locations is described by an adjacency relation. At each location, there can be multiple individuals, although in some cases, such as cellular automata [Ila01], this may be restricted to a single individual. A location may be an abstraction or aggregation of actual space.

topological: This approach to space considers the relationships between points in space and contains no notion of time. It can be applied to both discrete and continuous space. Topological spaces consider space in an abstract manner using open sets from which concepts of continuity, adjacency and neighbourhoods are defined. Metric spaces have a notion of distance which is used to define these concepts.

As before, we introduce notation for the remainder of this document. For the purposes of this section, we only consider 2-dimensional space when we consider continuous space, either $\mathbb{R} \times \mathbb{R}$ or a bounded contiguous subset of $\mathbb{R} \times \mathbb{R}$. Since this provides an uncountably infinite set of points, the most straightforward way to refer to each point is by its coordinates (x, y) .

In the case of discrete space, we assume a finite (or at most countably infinite) set of points \mathcal{L} with some naming convention. There can be a relationship between points in 2-dimensional continuous

TIME	continuous			
AGGR	none		state and/or space	
STATE	discrete	continuous	discrete	continuous

SPACE				
discrete				
general	multiple individuals CTMC	TDSHA, PDMP [BP10, Dav93]	patch CTMC [CLBR09]	patch ODEs [CLBR09]
regular	interacting particle system (IPS) [DL94b]		foraging model [MSH05]	foraging model [MSH05]
homo- geneous			bike sharing [FG12]	bike sharing [FG12]
continuous	molecular dynamics [CPB08] agents	CTMP [DP03]	spatio-temporal point processes [SBG02]	partial differential equations (PDES) [HLBV94]

Figure 2: Classification of mathematical models in terms of time, aggregation, state and space

space and locations in discrete space. If there is a partition of continuous space into regions then each location can represent a region in continuous space.

Discrete space also requires notions of adjacency and neighbourhood. In the general case, the set of locations \mathcal{L} can be taken as the vertices of a graph, and the connections between locations (the adjacency relation) can be defined as edges in that graph. Then the edges of the graph $E_{\mathcal{L}}$ are drawn from the subsets of size two of the location set $\mathcal{P}_2(\mathcal{L})$, so $E_{\mathcal{L}} \subseteq \mathcal{P}_2(\mathcal{L})$. Each edge has the form $\{l_1, l_2\}$, and edges of the form $\{l, l\}$ are permitted. For reasons we describe in Section 3, we have chosen to use an undirected graph which is to be understood as allowing movement or interaction in at least one direction between the two locations.

By adding restrictions to the general case, subclasses of discrete space can be obtained. For example, *regular discrete space* modelling techniques assume that there is a *regularity* in defining the neighbours of a location but not necessarily in parameters [DL94b, OS71]. In contrast, *homogeneous space* techniques take a different approach and assume *full connectivity* between all regions (giving a complete graph) and equality of parameters between locations and at locations [Che81]. Both of these issues will be discussed in more detail and formality in the section on discrete space (Section 3).

We now address an issue of terminology. The term *map* will be used to represent a two-dimensional continuous representation of space that is to scale, meaning that the ratio between distances is preserved. On the other hand, the terms *location graph* or simply *graph* will be used to denote a representation that is not to scale but indicates the connections between locations. In between these two, is the category of *topological map* which is

“a type of diagram that has been simplified so that only vital information remains and unnecessary detail has been removed. These maps lack scale, and distance and direction are subject to change and variation, but the relationship between points is maintained. A good example is the tube map of the London Underground.” [Wik13b]

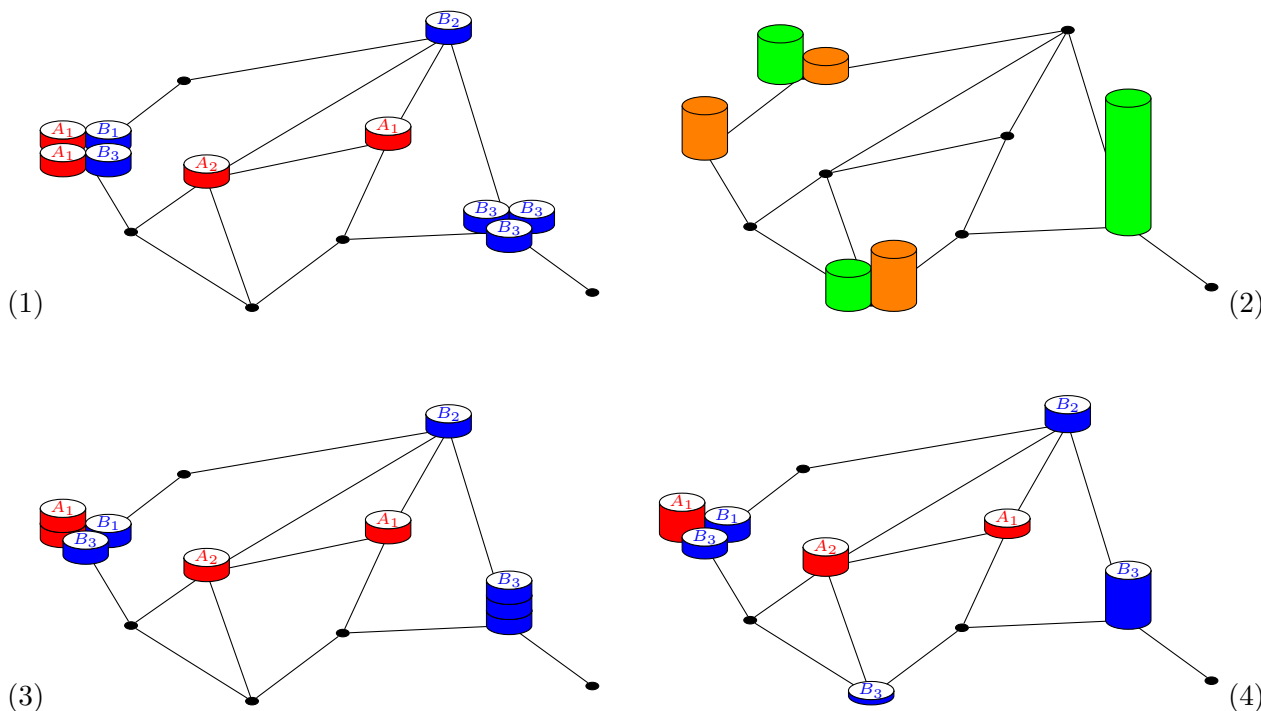


Figure 3: Discrete space: (1) no aggregation, discrete state; (2) no aggregation, continuous state; (3) aggregation of state, possible aggregation of space, discrete state (4) aggregation of state, possible aggregation of space, continuous state

The distinction between a topological map and a graph is that although both represent relationships between points, the topological map is a continuous deformation of the original map, and it also is a two-dimensional representation of a graph. A graph is more abstract and has no specific two-dimensional representation. Note that planarity (the lack of overlapping edges when embedded in the plane) is not required for topological maps or graphs. Any topological map can be abstracted to a graph of locations, hence defining discrete space.

A table has been constructed to identify mathematical models for the different combinations of time, aggregation, state and space (see Figure 2). Here, we have chosen to focus on continuous time models; however there are discrete time models of various approaches, for example, some variants of interacting particle systems (IPs) use probabilities [DL94b].

All the models appearing in the table consider changing behaviour over time. Characteristics of space may or may not change as time passes⁴. In the case where there are no changes, space can be considered independently of time and represented as a topological space. When there are changes in the characteristics of space over time, the characteristics of space at a specific point in time can be considered topologically.

The next section considers each entry of the table in Figure 2 and illustrates the ideas using a consistent diagrammatic framework.

2.5 Spatial modelling techniques

The techniques described in this section are mainly continuous time, although some have discrete time analogues, as mentioned above.

⁴An example of this would be off-peak road closures for the painting of road markings.

Discrete space, no aggregation, discrete state: The techniques in this category consider space to consist of a (finite) number of locations that have connections between them. The most straightforward way is to consider these models as graphs with the locations as nodes and the links as edges. This type of model is illustrated in Figure 3(1). Here, and throughout this subsection, we assume individuals are from the two populations already defined. The first, P_A consists of red and white tokens, and has states A_1 and A_2 , using our previously introduced notation. The second, P_B consists of blue and white tokens with states B_1 , B_2 and B_3 . The current state of an individual is indicated on the top of the token. Note that the four diagrams in Figure 3 represent four single points in time and do not show change over time (and similarly for subsequent diagrams). For two-dimensional and three-dimensional space, the best visualisation method for change over time is video. For one-dimensional space, a graph with two axes can be used.

Regular space models in this category have a regular pattern of locations [DL94a, DL94b]. For example, the locations could be laid out in the rectangular grid, or a hexagonal tiling. The locations that represent space can be placed at the nodes of the regular graphs or in the spaces (faces) created by the regular graph as shown in Figure 4(1). Some models only allow one individual in each location, such as interacting particle systems (IPSS) [DL94b] and cellular automata (CA) [Ila01], but others may allow multiple individuals. There is no aggregation of individuals. In Section 3, regular space will be formally defined.

Discrete space, no aggregation, continuous state: These techniques differ from those above in the fact that the state is continuous. This is indicated by a solid token where the height indicates the value of a single continuous state. This is an inherently continuous value rather than the notion of approximation by continuous values described earlier in this section. This could be viewed as a measurement such as strength of radio signal or length of battery life. In Figure 3(2), there is an assumption of at most one individual per node, and two values associated with that individual. Different colours have been used in the diagram to make it clear that the values are continuous.

The major difference between this category and the previous one where regular space is considered is the fact that instead of having discrete states, one or more non-negative real values are associated with each individual as shown in Figure 4(1).

Discrete space, aggregation, discrete state: These techniques differ from those in the first category above in the fact that there is aggregation [MP12]. This means that instead of each individual being treated separately, individuals in the same state are considered as populations. This is illustrated in Figure 3(3) by the fact that individual tokens are grouped into stacks at nodes in the network. Figure 4(3) shows the regular space case where populations are aggregated at each location [EE04].

Discrete space, aggregation, continuous state: Here each region or point is associated with approximations to the discrete population approach mentioned in the previous item [MP12], as shown in Figure 3(4). At each node, for each state in each population, there is a real number that approximates the number of individuals in that state. This is illustrated in the graph by a column with a real-valued height for each state in each population. Note that in Figure 3(4), the lowest node has a non-zero value for blue tokens in state B_3 although there were none in the CTMC model in Figure 3(3), illustrating that approximation can occur. The case of regular space [LD96] is illustrated in Figure 4(4).

Continuous space, no aggregation, discrete state: These are approaches where each individual's location and state are modelled separately from those of other individuals. An example of this type of model is where the movement and interaction of each molecule is modelled individually in molecular dynamics [BU10]. Agent-based models take a similar approach. Figure 5(1) illustrates this. The continuous space is indicated by a bounded area and each individual is shown at its own location. These models are typically computationally expensive to simulate.

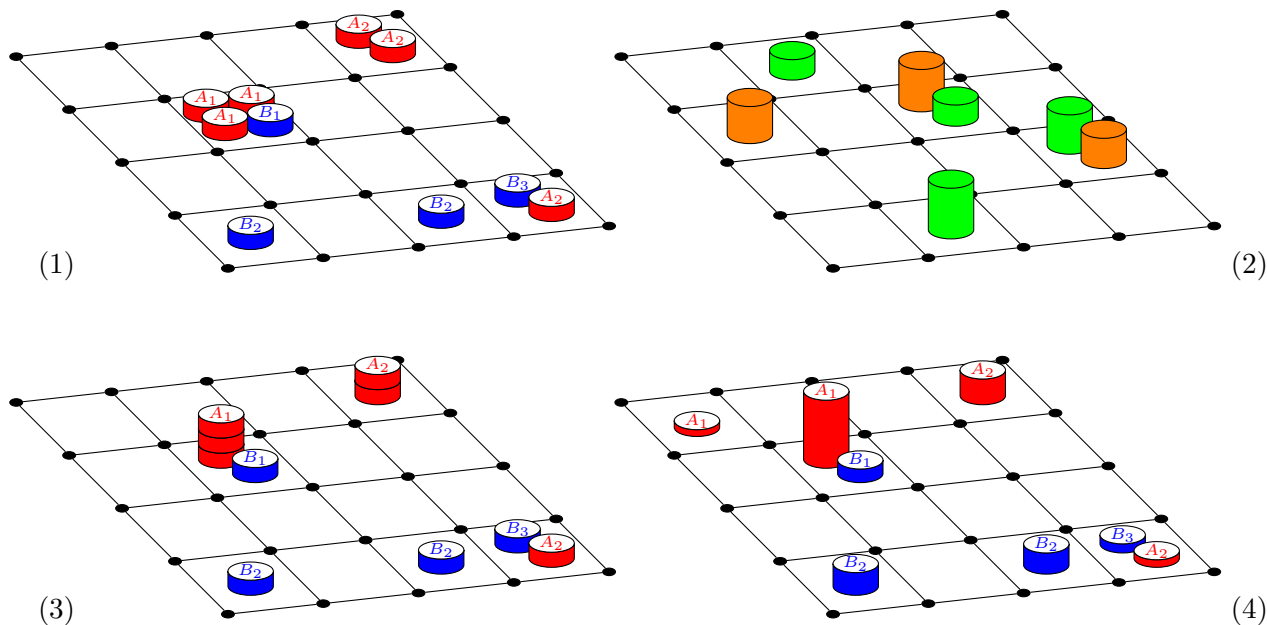


Figure 4: Regular discrete space: (1) no aggregation, discrete state; (2) no aggregation, continuous state; (3) aggregation of state, possible aggregation of space, discrete state (4) aggregation of state, possible aggregation of space, continuous state

Continuous space, no aggregation, continuous state: In contrast with the previous category, the state is now continuous rather than discrete [DP03]. Since there is no aggregation, this approach models individuals rather than populations. The continuous space is indicated by a bounded area and each individual is shown at its own location. The continuous state is indicated by the varying heights of the tokens, and in Figure 5(2), it is assumed that there is only one (non-spatial) measurement per individual, although two different qualities may be measured.

Continuous space, aggregation, discrete state: In these techniques, each point in space can be filled by one or more individuals [SBG02]. Hence for each point in space, it is possible to aggregate the number of individuals in each state. Figure 5(c) shows a fairly sparse number of individuals but much denser arrangements are also possible.

Continuous space, aggregation, continuous state: At each point in space, there is a real value describing an approximation to the number of individuals at that point [OL01, CPB08]. In the case of two-dimensional space, the population of each state can be represented in three-dimensions by surfaces. Figure 5(d) illustrates a surface describing the number of individuals at each point for state A_1 . In contrast to Figure 5(c), this figure illustrates a very dense situation.

As is the case with techniques that do not include space, presented in Figure 1, the techniques using continuous state without aggregation (the second column of models in Figure 2) seem distinctly different to the other approaches. The techniques that can be applied to models without space described early in this document (approximation by ODEs of a population DTMC or CTMC) can be applied to discrete space since the Markov chain involved is a population Markov chain that takes location into account. Furthermore, taking the hydrodynamic limit of IPS models provides PDEs [DMP91]. This subsection has not considered movement which will be examined now.

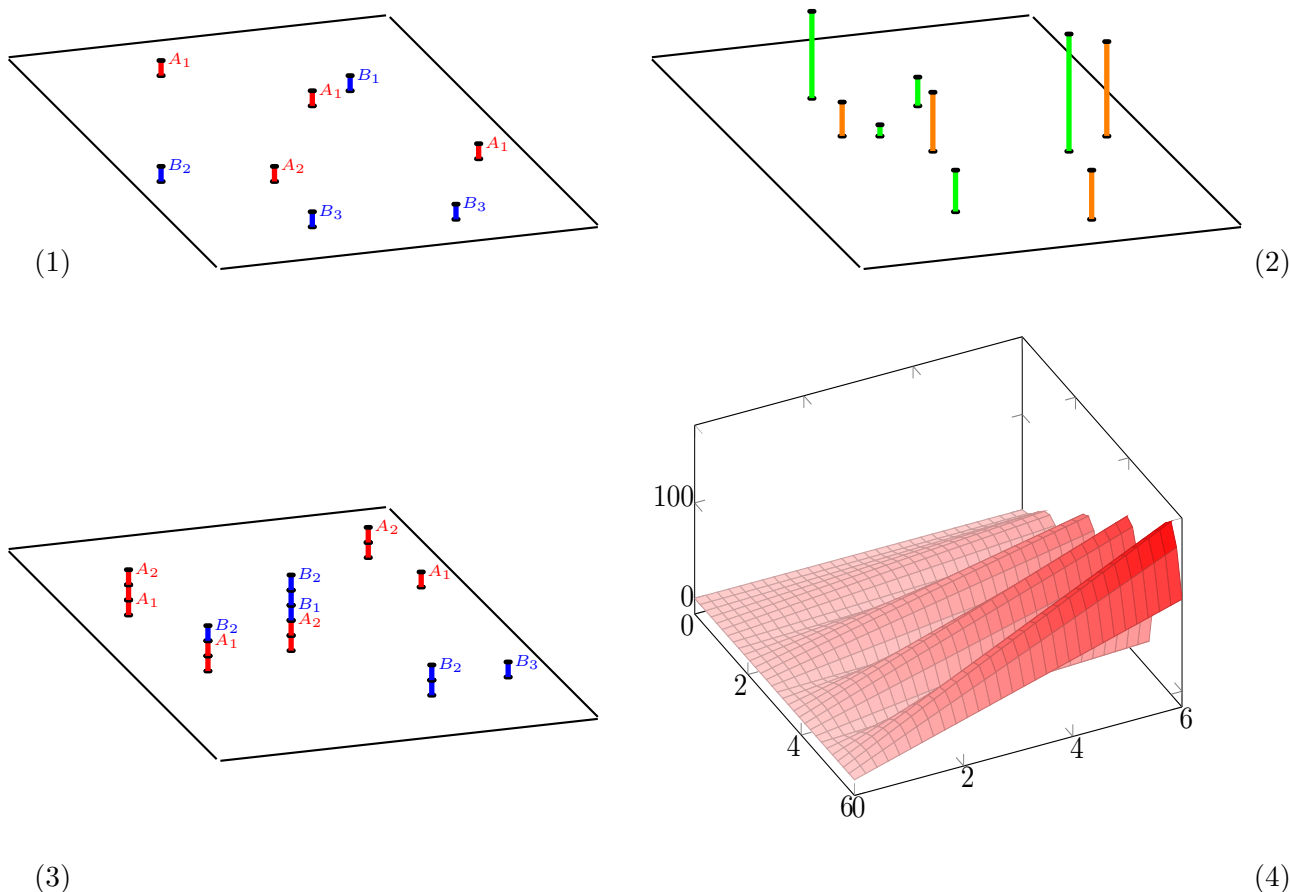


Figure 5: Continuous space: (1) no aggregation, discrete state; (2) no aggregation, continuous state; (3) aggregation, discrete state; (4) aggregation, continuous state.

2.6 Movement

In all of the models described in the previous section, there may be interaction between individuals (even if this interaction is expressed at the population level). Opportunity for interaction is often related to colocation or proximity (which requires some notion of neighbourhood or distance). Many models capture movement of individuals explicitly and then use colocation or proximity to determine the possibility of interaction, although there are some models that only use proximity without movement such as IPSs and CA. In these two modelling techniques, space is regular and discrete and at most one individual is present at each location.

Because of the importance of movement in the modelling of smart transport, we must consider the choices that can be made, and they are now discussed for the two different types of time-based spatial modelling techniques, described in Section 2.4. However, in the case of continuous space, this discussion is split into modelling techniques where there is aggregation and those where there is none.

2.6.1 Discrete space

Assuming an undirected graph of locations, the presence of an edge between two locations describes the fact that movement or interaction along that edge is possible in at least one direction. The absence of an edge can be interpreted as meaning that movement and interaction can never take place, in either direction. As we will see in Section 3, parameters associated with an edge express (possibly in a time-varying manner) the propensity for movement or interaction in either direction. If it is zero at a particular time for a particular direction, it means that no active interaction or movement can take place at that time point. Hence, the graph of locations provides a skeleton for describing what

movement or interaction is possible.

The adjacencies created in a location graph of regular space can capture where movement or interaction may occur (possibly with some weighting to capture likelihood) or what the neighbourhood of a location is.

2.6.2 Continuous space, no aggregation

In the case of continuous space where individuals are not aggregated, there are many different models of movement through two-dimensional space, such as models of animal movement and models for ad hoc and opportunistic networks [CBD02]. These are often random and capture the probability of movement in a particular direction at a certain speed. An additional concern is to determine what happens at the boundary of the space. This concern can be avoided by assuming the space is the surface of a torus and hence has no boundaries – this is more common than assuming the surface of a sphere. There are also models to describe the movement of a related group of individuals through the space [CBD02]. Connectivity models on the other hand, describe interaction (for example, contact duration and time between contacts) rather than location [KP07, CFB09, CMRM07]. Interaction can be interpreted as dynamic graphs with the individuals as the nodes. Connectivity models are more abstract than movement models.

2.6.3 Continuous space, aggregation

Here, movement is expressed through the form of the PDE. Diffusion-reaction PDEs are used since they can express movement as diffusion and interaction as reactions [CPB08, HLBV94, OS71, Tur53]. The diffusion terms can capture drift which is caused by obstacles or external stimuli such as wind, the likelihood of continuing in the same direction, the effect of the density of other individuals, and the impact of environmental characteristics. The reaction term describes interactions between individuals.

2.7 Conclusion

Figure 2 provides a classification of different spatial models taking into account aspects of time, aggregation and state. Movement models can be associated with particular aspects of space. This section has outlined some basic ideas, and the next three sections consider three different approaches to space. Two are time-based, namely discrete space and continuous space, and one is static, topological space. Each section will provide further details about the different approaches to modelling space, including a discussion of existing models and the areas in which they have been applied, and how these models are related to the smart transport case studies considered in the project.

3 Discrete space modelling techniques

We will focus here on the continuous time models, with pointers to the discrete time models where appropriate. As mentioned in the previous section, we assume a set of locations \mathcal{L} and an undirected graph over locations $(\mathcal{L}, E_{\mathcal{L}})$ with $E_{\mathcal{L}} \subseteq \mathcal{P}_2(\mathcal{L})$. Each edge has the form $\{l_1, l_2\}$, and loops such as $\{l, l\}$ are allowed. In some cases, a hypergraph may be useful and here, each edge is a subset of \mathcal{L} rather than subset of size two. For a particular application, there may be a reason to restrict edges to subsets of fixed size. We do not consider hypergraphs further in this document.

Locations in discrete space models can have two main sources, either they are essentially locations on a map, such as bike stations or bus stops, or alternatively each location represents a region on a two-dimensional map, and space is aggregated. The edges of the graph can be determined by various factors. Adjacency of regions is an obvious choice, but there may be other context-specific elements such as presence of connections between regions such as railway lines or similar.

If each location is used to represent a distinct region of continuous 2-dimensional space, a function $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathcal{L}$ can be defined. Since this is a function, it is the case that $f^{-1}(l_1) \cap f^{-1}(l_2) = \emptyset$ for locations $l_1 \neq l_2$ and hence regions are disjoint. Depending on the context, the union of $f^{-1}(l)$ for all l may be the whole continuous space under consideration so f defines a partition of the space, or this union may be a subset of the continuous space, only representing regions of interest. In some cases, where not all of continuous space is of interest, a single location can be used to represent the uninteresting regions, thus ensuring a partition.

A modelling technique with discrete space will have parameters that depend on locations, or links between locations. We can consider two groups of parameters; those that are associated with locations, namely with vertices of the graph and those that are associated with interaction or movement, namely the edges of the graph, and we define two functions to describe these parameter sets as follows

- $\lambda(l)$ for $l \in \mathcal{L}$, and
- $\eta(l_1, l_2)$ and $\eta(l_2, l_1)$ for $\{l_1, l_2\} \in E_{\mathcal{L}}$.

The range of these functions will remain abstract for the purposes of this discussion. Note that although the edges of the graph are not directed, the function η is sensitive to direction. We have chosen to use undirected graphs to give a basic skeleton to possible movement and interaction between locations, and to use the parameters to capture the directionality of that movement or interaction. Movement is obviously directional. Interaction can be undirected when considering an abstract view of effect or communication. Alternatively, it can be directed if one party is the sender and the other the recipient. This could be synchronous when the sending and receipt happen simultaneously, or asynchronous when the receipt happens later. Our choice of an undirected graph allows these details to be expressed in parameters. This separation of concerns is also useful when different subpopulations have different forms of interaction. In the rest of this document, the term *transfer* will be used to refer to both movement and interaction.

A topic whose exploration is beyond the scope of the current document but should be mentioned is that of how to divide a map in regions. A simple approach is to base it on a tiling of the plane using triangles, quadrilaterals or hexagons. More complex approaches involve taking local information into account and creating irregular patches. This is a topic for further research within QUANTICOL.

The discrete space approach as described above is very general as it allows arbitrary graphs over locations, as well as heterogeneity for parameters. In the literature there are modelling techniques that are defined for specific graph subclasses and we now identify two important spatial subclasses: spatially homogeneous and regular.

3.1 Spatial homogeneity

To be able to discuss formally aspects of discrete space, we develop the following definitions, leading to a definition of spatial homogeneity (a term which is used in the literature but not formally defined), by considering the location-related parameters. A spatial model is

- *location homogeneous* if $\lambda(l_i) = \lambda(l_j)$ for all locations $l_i, l_j \in \mathcal{L}$.
- *transfer homogeneous* if $\eta(l_i, l_j) = \eta(l_j, l_i) = \eta(l_{i'}, l_{j'}) = \eta(l_{j'}, l_{i'})$ for all edges $\{l_i, l_j\}, \{l_{i'}, l_{j'}\} \in E_{\mathcal{L}}$.
- *(spatially) parameter homogeneous* if it is both location and transfer homogeneous.
- *spatially homogeneous*⁵ if it is parameter homogeneous, and its location graph is complete⁶. Regular connections between locations which do not give total connectivity are discussed in the next section on regular space.

⁵This is a different notion to the graph theoretic definition of homogeneous graph which is a condition on isomorphic subgraphs [Gar76] and to homogeneous Markov graphs which are a specific class of random graphs [FS86].

⁶A complete undirected graph has an edge $\{l, l'\}$ between each pair of vertices l and l' .

Models with spatial homogeneity have a symmetry that can allow for analyses that are not possible for more complex models. Examples are the bike sharing system considered in [FG12] where the metrics of interest are the number of empty and full bike stations.

Spatial inhomogeneity can be introduced in two ways: the first involves connectivity where equal accessibility is no longer assumed, and the second where all locations are still accessible from all locations, but parameters vary between locations. These are not necessarily distinct concepts. Consider the case where there is a parameter $\rho_{i,j} \in \eta(l_i, l_j)$ which describes the rate of movement from location i to location j . If $\rho_{i,j}$ is the same for all i and j and no other parameters vary by location then the model is spatially homogeneous. However, if $\rho_{i,j}$ can vary and possibly be zero then not only does a specific parameter vary by location but additionally, equal accessibility no longer holds (either because on average it takes longer depending on the rate, or if the rate is zero there is no accessibility). However, if $\rho_{i,j}$ is constant for all i and j but other parameters vary by locations, then the model is spatially inhomogeneous.

3.2 Spatial regularity

The category of regular discrete space covers those spatially inhomogeneous models where the organisation of space is regular (rather than an arbitrary graph where each vertex may have an arbitrary number of edges) but parameters can vary for each location in space

In contrast to spatial homogeneity, regularity of space is more difficult to define formally when starting from a graph, although it is very straightforward to identify visually [OS71]. There are three possible approaches to describing regularity in two dimensions.

1. A lattice, grid or mesh graph is defined as “a graph whose drawing, embedded in some Euclidean space \mathbb{R}^n , forms a regular tiling” [Wik13a]. Here, we focus on two-dimensional space and hence we are considering planar graphs. Since a regular tiling is created using regular polygons, there are only three regular tilings of the plane: equilateral triangles, squares and regular hexagons. The three lattice graphs we obtain from these tilings are infinite, and for modelling we may only wish to consider finite graphs. This can be achieved by taking a connected induced subgraph⁷ of a lattice graph.
2. Another approach to specifying graphs of regular space is to specify how many edges each face⁸ of the graph has and what the degree of each vertice is. A regular location graph with triangular faces is a planar graph in which each face has three edges and each vertex has degree six, a regular location graph with quadrilateral faces is planar, has faces with four edges and with vertices of degree four, and a regular location graph with hexagonal faces is planar, has faces with six edges and vertices with degree three.
3. Finally, a graph with regular structure can be constructed by identifying points in $\mathbb{Z} \times \mathbb{Z}$ or $\mathbb{R} \times \mathbb{R}$, and adding links. For example, the graph with a hexagonal structure can be constructed as follows.

Assume the initial coordinates are (x, y) then let

$$\begin{array}{lll} p_1 = (x, y) & p_2 = (x + 1, y + 1) & p_3 = (x + 2, y + 1) \\ p_4 = (x + 3, y) & p_5 = (x + 2, y - 1) & p_6 = (x + 1, y - 1) \end{array}$$

be vertices, and add the following edge set: $\{(p_1, p_2), (p_2, p_3), (p_3, p_4), (p_4, p_5), (p_5, p_6), (p_6, p_1)\}$. This creates an initial hexagon and other hexagons can be placed at appropriate points so that the space is tiled.

⁷An induced subgraph is obtained by taking a subset of the vertices of a graph and keeping all of the edges that have both endpoints in the subset.

⁸In a planar graph, a face is a region bounded by edges.

There are other divisions of two-dimensional space that can be viewed as regular such as that provided by a dartboard but we will not attempt that level of generality for discrete space beyond saying that regular space should have the property that at each location (except possibly at boundary locations) there is a uniform way to determine the neighbours. However, we exclude from this definition n -hop neighbours in an arbitrary graph (see definition of n -hop in the next subsection).

One-dimensional regular space can be represented simply as an undirected path. We do not tackle the definition of three-dimensional regular space.

Note that the complete graph requirement for spatial homogeneity means that a location graph cannot be both regular and homogeneous⁹. However, a spatially regular location graph can be parameter homogeneous. We will use the term *irregular space* whenever the usage of space is not regular or homogeneous. To distinguish irregularity from regularity, we will consider the pattern of connectivity between neighbours and this is discussed further in the next subsection.

3.3 Neighbours and neighbourhoods

In an undirected graph of locations representing discrete space, the links between locations are used to define neighbours. Given a location l , its *immediate neighbours* are those vertices l' such that $\{l, l'\}$ is an edge in the graph. Its *n -hop neighbours* are those that can be reached through a path in the location graph of at most n steps (but excluding the location l itself). In the case of a regular grid graph, the immediate neighbours (west, north, east and south) are referred to as the Von Neumann neighbourhood. The larger neighbourhood that includes the northwest, northeast, southeast and southwest points as well as the immediate neighbourhood is known as the Moore neighbourhood. Both types of neighbourhoods can be extended to n -hop neighbours and also applied to hexagonal and triangular regular location graphs.

This is a purely spatial approach to defining neighbourhoods. However, in some cases, it can be the entity or process itself that defines its neighbourhood depending on its capability. Other approaches use a (perception) function that determines the neighbours of an individual by specifying the other individuals with which it can interact.

To distinguish irregular location graphs from regular location graphs, one can say that if it is possible to define the 1-hop neighbours of a location in a regular fashion except at boundaries, then this will be considered a regular discrete space model even if the parameters can vary by location. An example of this is given in the approach taken to modelling fluid limits for stochastic mobile networks [TT13]).

3.4 Boundary conditions

An issue for discrete space (and continuous space) is determining what happens at the boundaries of the space. One approach is to ensure there are none by working with infinite structures such as infinite graphs, or alternatively boundaryless structures such as tori. A rectangular region can be transformed into a torus by joining the top and bottom edges (to form a cylinder) and then joining the left and right ends (by curving the tube). Other approaches work with boundaries and either choose to keep individuals inside the region (by reflection or other techniques) or to treat boundary locations as sources and/or sinks.

Furthermore, when solving ODEs, there may be boundary conditions that constrain the ODEs by specifying the values that must occur in boundary regions [AMR95]. The conditions could constrain the value of the solution (Dirichlet, first-type) or the value of the derivative (Neumann, second-type). Cauchy boundary conditions provide a curve or surface that constrains both solution and derivative.

⁹The complete graph with three vertices is regular and can be spatially homogeneous. However, since we are considering tilings/graphs with multiple faces, this graph is not included in our definitions.

3.5 Discrete space without state-based aggregation

We now consider the different modelling techniques that have been applied to discrete space starting with those that do not involve aggregation. When there is no aggregation and state is discrete, the technique models individuals and can be seen as an agent-based system where space is discrete. Hence each individual has some state and is located at exactly one location. Multiple individuals can be located at a single location. The movement of individuals between locations can be determined by a rate in the case of continuous time. To describe these in their most general form, we assume that each individual I (where I is a unique name for the individual) has associated time-based information:

- $\text{loc}(I, t) \in \mathcal{L}$ which is its location at time t
- $\text{state}(I, t) = A_i$ which is its state¹⁰ at time t

Moreover, there are rules that describe how an individual can change location or change state. Since this is a continuous time model, these rules may specify rate constants (each rate defining an exponential distribution) to describe how long it takes for the changes to occur, or the rates may be functional (but still defining an exponential distribution) that take into account the presence of others at that location, the characteristics of the location or even the current time (thus introducing time inhomogeneity). The behaviour of the agents in this modelling technique is thus described as they individually change state and/or location. Assuming a fixed population size, we can model this system as a CTMC, where each state in the CTMC is a tuple consisting of information about each individual in the system. If we assume N individuals then a state has the following form

$$((\text{loc}(I_1, t), \text{state}(I_1, t)), \dots, (\text{loc}(I_N, t), \text{state}(I_N, t)))$$

There are $(L \times n)^N$ states in this Markov chain if there are L locations and n states and if it is possible for all individuals to be in all possible combinations of location and state.

Simulation suits this type of model, and techniques for simulating systems where behaviour is based on functional exponential rates are well understood [Gil77]. Typically in the case of general discrete space, movement is assumed to be single-hop so that movement is only possible to an immediate neighbour. In regular space, movement is often possible to an n -hop neighbourhood, and in spatially homogeneous models, movement is possible to any other location as a result of the completeness of the location graph.

Next, considering discrete space modelling techniques without aggregation where the state is continuous, instead of having a rule describing what the next state is, there needs to be a rule describing how this continuous value changes over time. A good candidate for this type of rule is an ODE. These techniques are hybrid in that they exhibit both continuous and stochastic behaviour. Additionally, they may also have instantaneous behaviour. Transition-driven stochastic hybrid automata (TDSHAs) [BP10] and piecewise deterministic Markov chains (PDMPs) [Dav93] are suitable modelling techniques.

3.6 Discrete space with state-based aggregation

In the two approaches that are described next, aggregation of state occurs. It is assumed that we have many individuals to whom the same set of rules apply with the same parameters, and we choose to view them as a population, P_A and to reason about them as a population. To extend the notation introduced earlier, assume we have a fixed number of locations, l_1, \dots, l_L . We can now consider the counts of subpopulations at each location. So for P_A , we have a value $N_{A_i}^{(k)}$ which is the number of

¹⁰If the population P_A has multiple attributes $A^{[1]}, \dots, A^{[p]}$, then $\text{state}(I, t) = (A_{i_1}^{[1]}, \dots, A_{i_p}^{[p]})$ represents a tuple of states.

individuals at location k in state i . Additionally

$$N_{A_i} = \sum_{k=1}^L N_{A_i}^{(k)} \quad \text{and} \quad N_A^{(k)} = \sum_{i=1}^n N_{A_i}^{(k)} \quad \text{and} \quad N_A = \sum_{i=1}^n N_{A_i} = \sum_{k=1}^L N_A^{(k)}$$

We can create a continuous time Markov chain smaller than that of the previous section consisting of at most $(N_A + 1)^{L \times n}$ states where each state has the form

$$(N_{A_1}^{(1)}, \dots, N_{A_n}^{(1)}, \dots, N_{A_1}^{(k)}, \dots, N_{A_n}^{(k)}, \dots, N_{A_1}^{(L)}, \dots, N_{A_n}^{(L)})$$

This provides a discrete aggregated representation of space where for each location, we know how many individuals are in each state without knowing exactly which individual at that location is in which state. As in the case without space, this is a population CTMC. This model is amenable to state-based analysis techniques (assuming a small enough state space) and stochastic simulation. There appears to be no real difference to the size of the Markov chain when using regular space in this technique. In the case of spatial homogeneity, the fact that parameters are identical may make the model amenable to an analytic approach, rather than requiring simulation [FG12].

In the continuous state variant of this technique, the notation $X_{A_i}^{(k)}$ is used for the real value that describes the quantity of individuals in state i at location k . Since this can be a non-integer value, it is an approximation to the actual count $N_{A_i}^{(k)}$. Since the subpopulation sizes are treated as continuous values, a standard modelling technique is to express the change in this quantity in terms of an ODE.

$$\frac{dX_{A_i}^{(k)}}{dt} = F_{i,k}((X_{A_1}^{(1)}, \dots, X_{A_n}^{(1)}, \dots, X_{A_1}^{(k)}, \dots, X_{A_n}^{(k)}, \dots, X_{A_1}^{(L)}, \dots, X_{A_n}^{(L)}), t)$$

This is a population ODE because it tracks the changes in subpopulation sizes over time. There are $L \times n$ variables in total; one for each combination of state and location. The inclusion of t as an argument to $F_{i,j}$ indicates that it can be a time-inhomogeneous ODE. This ODE often has the following form

$$\begin{aligned} \frac{dX_{A_i}^{(k)}}{dt} = & f_{i,k}(X_{A_1}^{(k)}, \dots, X_{A_n}^{(k)}) + \\ & \sum_{j=1, j \neq k}^L (g_{i,k,j}(X_{A_1}^{(k)}, \dots, X_{A_n}^{(k)}, X_{A_1}^{(j)}, \dots, X_{A_n}^{(j)}) - h_{i,k,j}(X_{A_1}^{(k)}, \dots, X_{A_n}^{(k)}, X_{A_1}^{(j)}, \dots, X_{A_n}^{(j)})) \end{aligned}$$

where $f_{i,k}$ captures the local behaviour which only depends on the subpopulation sizes locally, $g_{i,k,j}$ describes the inflow of population from location j to location k , $h_{i,k,j}$ describes the outflow of population from location k to location j , and these flows depend only on the subpopulation sizes in location k and location j . This is a time-homogeneous ODE since change over time is only dependent on subpopulation sizes (that are dependent on time) rather than on time directly.

For both the general and regular space cases and assuming only movement/interaction between 1-hop neighbours, then a term $X_{A_i}^{(j)}$ should only appear in the right hand side of the ODE if $\{l_k, l_j\}$ is an edge in the location graph. A spatially homogeneous model would require terms from all locations.

The modelling techniques in the categories of discrete state (with or without aggregation) and continuous state with aggregation (these are the categories in the first, third and fourth columns for general discrete space in Figure 2) are very similar to those in the same categories (and the same columns) in Figure 1. The discrete state approaches without aggregation (first column) are characterised by CTMCs, discrete state approaches with aggregation (third column) by population CTMCs which abstract from individuals, and continuous state approaches (fourth column) by ODEs. This similarity is not surprising, as in the general case of discrete space, location is essentially another attribute and hence the same techniques apply. Similarly to before, the technique for continuous

state modelling without aggregation (second column) is out of step with the other techniques since it requires a hybrid approach when discrete space is introduced.

Introducing locations while keeping the population size fixed can result in a decrease in accuracy of the approximations as now the number of variables (and ODEs) has increased by a factor of L and hence each ODE refers to a smaller number of individuals, thus leading to population fragmentation and less justification for applying mean-field techniques at a location.

3.7 Examples of existing use

This section reviews the literature of modelling with discrete space techniques, considering different approaches in different disciplines. It is selective and illustrative rather than exhaustive.

Before going further, we diverge into a discussion of terminology. As mentioned above, terminology is not consistent between different disciplines or even within disciplines. Examples of this confusion include “mean-field” and “patch model”, and before the use of discrete space modelling is discussed further, this issue is tackled.

The term mean-field has been used differently in ecology. Morozov and Poggiale [MP12] identify five different uses, and relate them to spatial modelling in ecology. They express a preference for the last two in the following list:

- (A) An ODE model with the assumption of well-mixed populations that allows the inference that there is equal probability of interaction between individuals that are close to each other regardless of where these individuals are located. This assumption is one of implicit spatial homogeneity. Morozov and Poggiale link these to “small world” approaches where individuals have equal chance of interacting regardless of distance.
- (B) A non-spatial ODE model that has terms added to represent explicit space on which an approximation is done by ignoring the spatial terms.
- (C) An ODE model that has terms that implicitly capture spatial heterogeneity and hence differs from a non-spatial ODE model.
- (D) The reduction of a spatially explicit model to one that is not spatially explicit but contains spatial moments.
- (E) A model that expresses spatially explicit individual-based models as spatially explicit population-based PDEs or integro-differential equations.

In this document, the term “mean-field” is not used with any implication of spatial aspects and hence matches interpretation (A). It is used here for those approaches based on differential equations that take a fluid approximation of populations whether there is discrete space in the model or not. Hence the approach described in this section for approximating a population CTMC by population ODEs is a mean-field approach.

Another confusing term is “patch models”. In Durrett and Levin’s 1994 paper [DL94a] comparing different approaches to space and showing that different outcomes are achieved¹¹, they describe a patch model (as defined by Chesson [Che81]) as one in which all locations are connected to all other locations. In later papers, the terminology refers to discrete space models with arbitrary graphs, and this is the sense in which we will use “patch”. Hence, both the discrete and continuous population models defined over discrete space in this section are patch models.

¹¹Durrett and Levin use the term “mean-field” in the sense of (A) in [DL94a].

3.7.1 Ecology

Space plays an crucial role in ecological models but the goals of these models are often qualitative as the focus is on generic mechanisms. Ecologists are interested in qualities of the whole space such as whether species persist or can co-exist, as well as dynamic patterns such as stationarity, oscillatory behaviour, chaos or multistability [MP12]. Hence, there is a focus on global behaviour.

Berec [Ber02] provides a classification of spatial models where he considers the time dimension (discrete or continuous), the space dimension (discrete or continuous) and the population dimension (discrete or continuous). This last dimension differs from our classification, as it does not separate out aggregation in the population dimension. He identifies the following models for discrete space.

discrete population, continuous time: These models are interacting particle systems where behaviour is determined by a set of rules. They are discussed in more detail below. Also falling into this category are stochastic cellular automata [OS97, LS11], contact processes [Har74, BG90] and binary-valued Markov random graphs considered on countable space [Ish81] which include the Ising model [Bru67]. These Markov random fields are a Gibbs state with a nearest neighbour potential [Ish81].

continuous population, continuous time: Reaction-dispersal networks describe change over time by a system of ODEs over species in locations. These are called metapopulation models [Lev69] or patch models [Lev74, LP74, Lev76, Tak96], and they are the same as ODE patch models in our terminology. The ODEs used in these models often focus on the probability of each patch being in one of a number of states (such as uninhabitable, habitable but unoccupied, occupied) [XFAS06]. The goals of the model are to be able to describe the overall proportion of occupied patches and to identify equilibria. Colonisation rates reflect distances between patches, so these models are not transfer homogeneous. Wu and Loucks [WL95] provide arguments for hierarchical patch models where different spatial and time scales are taken into account.

discrete population, discrete time: These are individual-based models where the behaviour of an individual is given by a list of rules which are applied at each time tick. Probabilistic cellular automata fall into this class as do interacting particle systems which use probabilities rather than rates [DL94a]. The order in which rules are applied is significant.

continuous population, discrete time: Coupled-map lattices are a type of discrete time equation and are defined by systems of difference equations [HCM91, Kan98]. They use regular space, and allow continuous population sizes. Examples of their use include plant growth [HMW96] and host-parasite interaction [CHM92].

Berec considers how the discrete population models can become mean-field models in the sense of (C) defined above and he assumes homogeneity of spatial parameters.

Durrett and Levin [DL94a] compare (i) a non-spatial mean-field model, (ii) a PDE model that uses continuous space, (iii) a spatially homogeneous model by Chesson [Che81] and (iv) a grid based IPS model. The first two models are simulated numerically and the last two are simulated stochastically. The authors identify three scenarios.

1. Two species enhance the presence of each other, and all models agree at the global level.
2. Two species compete for a single resource, and the two spatial models [(ii) and (iv)] differ from the two non-spatial models [(i) and (iii)] in global outcome.
3. One species always does better than the other but which will go extinct if the other dies out, and the two deterministic models [(i) and (ii)] differ from the two stochastic models [(iii) and (iv)] in global behaviour.

This exercise illustrates that choice of model is important and can affect the global results of the model. Later work considers scenarios with more than two species and shows differences between spatial and non-spatial treatment [DL98].

Durrett and Levin [DL94b] have also surveyed IPSs which have a discrete view of space, and an assumption of at most a single individual at each site. These models use regular space, and are homogeneous in terms of their spatial parameters. Time in this approach can be discrete or continuous, and measures of interest include global characteristics such as extinction or the proportion of occupied sites. The models are simulated and one-dimensional model behaviour can be expressed in two dimensions. For models with two-dimensions, snapshots of behaviour are given but video is probably the best way to display these simulations.

Morozov and Poggiale [MP12] start from a discrete space, discrete aggregation model (a population CTMC) and consider the various ways to obtain mean-field models. The first four are essentially moment closure models that provide mean-field models of global population size of species by introducing spatial terms into the ODE so they fall into category (4). The last takes a different approach.

spatial moment closure: This approach is based on expressing models as ODEs over the moments of the model. Given a patch model, ODEs of interest describe how the means of some quantities vary. Each patch will vary in these quantities, and it is the global variation over time that is of interest. The ODEs for these means can be expressed in terms of higher moments, for example the covariance of these quantities which itself can be expressed in terms of even higher moments, leading to an infinite regress. There are various ways to truncate this sequence and the general approach is to assume that the higher order moments do not have a significant impact on the model. Because the covariance captures spatial variation, it must be retained otherwise the ODE would have no spatial component but third order and higher moments can be approximated. In the case of the mean of the product of three variables $E[xyz]$, it can be approximated by $E[xy]E[z]$, $E[xz]E[y]$ or $E[yz]E[x]$ or by using a distribution-based approximation [MSH05]. An example of moment closure is given by a regular discrete space model of foraging [MSH05]. In the initial model, the foragers can move to any location but they are more likely to move to a location with high grass. A refinement of this is to only allow them to move to a location with higher grass that is nearby. These are presented both as discrete population (aggregation) models and continuous population models. The goal is to derive ODEs for average height of the grass over all locations and the average density of animals over all locations. This is achieved by using moments including the variance of the grass height and the covariance of the grass height and forager density. The closure is obtained by first approximating expectations of the form $E[x_i y_j z_k]$ where each variable is associated with a patch and captures height or density with global expectations for height and density, and then means of products with products of means. Single patch third-order terms are replaced by log normal approximations. This model introduces spatial heterogeneity by allowing movement to be dependent on characteristics of the destination location. More realistic movement is then captured by allowing foragers to only assess the characteristics of nearby locations. Different approaches to spatial moment closure using different distributions for approximation have been compared [KCMG05].

pairwise approximation: This modelling technique is expressed in terms of ODEs which describe changes in the probabilities of certain pairs (adjacent locations) in the model [WKB07, MOSS92]. From these ODEs, the proportion of locations in a particular state can be determined. The ODEs include conditional probabilities over triples of sites and these conditional probabilities are approximated by simpler conditional probabilities. Ellner [Ell01] considers a multiscale pair approximation technique.

modified mean-field: In this technique, higher order moments are approximated by functions of first order moments, usually of the form $c.E[x]^a.E[y]^b$ where x and y are variables and a , b and c are constants [PRL11].

scale transition theory: Choice of spatial scale can affect modelling outcomes, and this theory explains the effect of spatial aggregation and allows one to achieve correct dynamics at different scales. It is a form of moment closure based on Taylor expansion of moments [Che12].

aggregation: This technique takes advantage of the structure of the model. If the proportion of a species in a patch can be expressed in terms of global variables then a mean-field model can be constructed. This technique can be applied when dispersal rates between patches are much faster than demographic rates within patches. First, the proportion of species across patches can be determined after which the demographic changes can be investigated [APS12, ACVP00]. This model can perform much better than the spatial moment technique when there is substantially demographic variation across patches but it does require differences in rates. This has similarities with the Quasi-Steady State assumption from biological modelling.

Reducing the complexity of mean-field models is possible in certain circumstances. When considering stability of spatially homogeneous solutions, Jansen and Lloyd [JL00] show that analysis of a k -species, n -patch system can be reduced to that of n uncoupled k -dimensional single-patch systems.

As can be seen, there are multiple approaches to mean-field models for patch-based models in ecology. Most consider global outcomes and qualitative assessment with few focussing on local details.

3.7.2 Biology

Bittig and Urmacher [BU10] identify five distinct methods for spatial modelling in cell biology that offer different granularities in their approximation of physical reality. Two of these are continuous space approaches and will be discussed in the next section.

compartments: In this model, space is divided into compartments which may be contained in other compartments, giving nesting of space, although adjacency is also used. Compartments are necessary to represent physical reality by ensuring that the correct concentrations are used in rate equations when concentrations may differ in different parts of the cell. Within compartments, a non-spatial approach is used implying an assumption of well-mixedness and standard algorithms such as Gillespie [Gil77] and extensions [Gil07] are used for stochastic simulation of population CTMCs. Transfer rates are defined between compartments and in some cases, the membranes between compartments. This technique is an irregular discrete space patch-based population CTMC approach or ODE approach, depending on whether the modelling is stochastic or deterministic.

discrete space, lattice: Typically, the size of the lattice is chosen to reflect average molecule distance and diameter, and it is assumed that at most one molecule can occupy a lattice block, requiring a way in which to deal with collisions. This is a similar approach to that taken by cellular automata: discrete regular space with single occupancy of locations [Ila01]. This approach has also been used for models of tissue.

discrete space, subvolumes: Like the previous point, a lattice is used, however, each block in the lattice can contain multiple molecules and molecules can diffuse to other blocks. If individual molecules are modelled then this is a discrete regular space approach without state-based aggregation. If aggregation is used, then it is a regular patch-based population model. The Next Subvolume Method [EE04, WHXY11] is used for simulation of these models, using a Standard Gillespie approach for reaction in a subvolume and diffusion to an adjacent subvolume, and Gibson and Brucks Next Reaction Method [GB00] to determine in which subvolume the next reaction takes place.

The subvolume approaches are based on a master equation named the Reaction Diffusion Master Equation (RDME) [GMWM76, Isa08] which describes how the probability of being in each state

changes over time, where a state is a vector representing the quantities of species in each voxel. A voxel is a small cube of 3-dimensional space and each voxel is assumed to be small enough to allow concentrations of molecules to be homogeneous throughout the voxel. The relationship between RDME and PDE approaches have been investigated [Isa09, II09]. For an overview of techniques to model diffusion, both stochastically and continuously, see [ECM07]. Choice of compartment size is important [EC09]. The importance of stochastic modelling over mean-field approaches is demonstrated in the successful modelling of *E. coli* phenotypes [FE06].

Marion *et al* [MMRL02] consider a simple biochemical reaction system and how to represent it when there are heterogeneous instances of the system over a rectangular grid of locations. This is essentially a patch model, and although there is a difference between locations, there is no notion of neighbourhood and the model assumes that it is possible for a biochemical species to migrate from any location to any other. A deterministic model of a single location is developed and extended to a (non-spatial) stochastic model. A spatial stochastic model by adding migration actions is derived and an “analytic approximation” is derived that provides equations to approximate first and second moments. Moment closure is done by using the log-normal distribution and the authors show that it gives closer results to the full stochastic model than the mean-field approximation. The measures of interest are mean, variance and covariance across all locations, so are global in nature.

The cellular Potts model [GG92, CHK⁺05] is based on the Ising model from physics [Bru67] and is a regular discrete space model similar to a cellular automaton. The lattice can be two- or three-dimensional and the rules for updating include a rule to determine which points in the lattice are upgraded and a potential energy function that determines what the updates should be. This approach has been applied to embryo development [CHK⁺05]. In biological modelling with this technique, each cell can be represented by multiple lattice points so it can be viewed as discretisation of actual space.

Othmer and Stevens [OS97] investigate movement of myxobacteria on a two-dimensional grid (using stochastic cellular automata) and the plane (using PDEs). They assume that a non-diffusive modulator chemical can be present at each cell or point, and they are interested in the three distinct dynamics that can occur in the PDE system dependent on this modulator. Later research [Ste00b, Ste00a] considers these models further, and use a novel approach to the limiting procedure since the equations do not meet the ellipticity condition used by Oelschläger [Oel89].

Pattern formation is also important in biology. Turing’s paper gave an initial insight into this process [Tur53] and later work has considered this further [OS71, LS85] both in terms of regular discrete space models that describe cells in a matrix and partial differential equations. Investigation of the stability of patterns is an important aspect of this research.

Spatial modelling in biology covers a range of topics from molecular interaction to movement of bacteria, and a range of techniques are used depending on the requirements of the model.

3.7.3 Epidemiology

Riley [Ril07] identifies four distinct approaches to disease spread modelling.

Patch-based transmission: Individuals within a patch have the same risk of infection which is determined by the current rate of infection over all patches, and some distance-based likelihood of infection from other patches.

Distance transmission: Here transmission is pairwise and the risk is determined by distance between pairs.

Multigroup transmission: This is a finer grained model that considers households and other groups with group membership ensuring a high risk of infection.

Network transmission: This is more fine-grained than the previous approach, and considers individuals and their interactions.

Although some of these approaches mention distance, this does not necessarily imply continuous space, although continuous space may be used in these models. Dorjee *et al* [DPR⁺13] classify modelling approaches to the spread of influenza taking into account those models that use spatially explicit networks or spatially explicit compartments.

Patch-based or *metapopulation* models¹² are used extensively in modelling of epidemics both generally [AP02, AvdD03, vdD08, ADH⁺05, Mol77] and for various diseases: rabies [Mur03, Chapter 13], measles [BG95, Gre92], cholera [BCG⁺09], influenza between trading posts in Canada [vdD08] and bovine tuberculosis in possums [vdD08]. These models often focus on the calculation of the basic reproduction number, R_0 , which describes the average number of infecteds each infected generates. If this number is less than 1, then a disease will die out, and if it is greater than 1 the disease will spread to the whole population. Furthermore, for certain models it can be used to determine the proportion of the population that should be vaccinated to control the spread of the disease which is $1 - 1/R_0$. R_0 is also used in ecology to describe reproductive success. Many of these models assume that the population is static in the sense that they do not change patches, and that the disease spreads within patches and between patches.

To understand the development of patch models in epidemiology, note that the standard SIR (susceptible-infected-recovered) model does not consider space and must be extended. In the extension without movement between patches, each patch i has values S_i , I_i and R_i described by differential equations where the rate of infection in each patch can be calculated from a weighted sum of the number of infecteds across all patches [vdD08]. For the case where the populations of each patch are the same size, and infection from outside patches is less than that from the patch to itself but identical across all outside patches, an analytic solution can be determined. To include movement, birth and death, first an ODE model is constructed without disease. If there are n patches, then for each patch, there are n variables to keep track of the individuals in that patch with respect to where they were born. Individuals can move from where they are born to any other patch and then back to their native patch. They are not able to move between foreign patches. Then disease is added to the model so that there are groups with different disease statuses in each patch. If the graph determined by the travel rates (an edge is present between patches if the travel rate is non-zero) is strongly connected, then equilibrium in one patch (disease-free equilibrium or epidemic equilibrium) implies all other patches are at the same equilibrium. Other models of movement can also be used.

The patch model developed by [ADH⁺05] allows for multiple species. The model has a basic model of mobility between patches for each species and a more complex model of disease transmission within and between species. A number of results relating to disease-free equilibrium are proved. Since this model considers discrete space and multiple populations with multiple states, the details are worth investigating further as part of the QUANTICOL project.

Arrigoni and Pugliese [AP02] investigate patch models in the limit where both the number of patches and the number of individuals go to infinity. This provides an countably infinite set of PDEs, with each PDE describing how the number of patches with n infecteds varies over time. This is an example of how a patch model that uses discrete space can be made more fluid without moving away from a model of discrete space.

Levin and Durrett [LD96] consider the linkages between ecological and epidemiological modelling. They consider a patch-model where the local dynamics are dependent on the average number of susceptibles, infecteds and recovered across the whole population. This leads to ODEs for the averages which include terms of high moments leading to the closure problem as mentioned in the earlier section on ecology. Taking a different approach by considering individuals on a grid, the change in the number of infected sites can be expressed with an ODE that depends on the probability of two adjacent states being in a particular configuration which is also not a closed system. Assuming

¹²The basic epidemiological model is called the compartment model [Bra08a] and this consists of a single population where each individual has a state related to whether they have the disease, can become infected with the disease or are immune. It has no spatial aspects and it should not be confused with the compartment models in biology which are patch-based models.

adjacent sites are independent, using a form of pair approximation leads to inaccuracies, hence ODEs for the probability of two adjacent states being in a particular configuration are considered. If this is closed by an approximation to these rates of change, a better estimate of the probability of infection is obtained. This approach is considered in the case of general contact processes where neighbourhoods are defined more generally, and a size-based approximation approach is used where certain terms are excluded from the second-order ODE because they are very small, resulting in a linear approach. The authors terms this the *dyad heuristic*. They also consider an interacting particle system with multiple individuals at each location and derive ODEs describing changes in the total population.

Other techniques can be used in epidemiology. For example, movement of individuals can be considered in space by regular space two-dimensional models where there are individuals at various points on a grid, both in the case of disease spread and effects of vaccination [RA97]. A certain population density is required for the disease to infect the whole population.

A network model of an epidemic can be constructed where the network represents contacts between people [Bra08b]. This model technique can be used when modelling the initial stages of an epidemic because using an assumption of homogeneous mixing may overestimate the size of the epidemic. Although this is not meant as an inherently spatial model, it does capture space. The model consists of a graph where nodes are individuals and edges represent contact between individuals (in either direction). Two parameters are required for the calculations: p_k , the fraction of nodes with k edges for each $k > 1$ and T , the probability of transmission. Generating functions are used to reason about the spread of the epidemic. R_0 , the basic reproduction number, is calculated and if this is less than one, the epidemic will die out (as is standard). If R_0 is greater than 1 then using fixed points, it is possible to reason about whether there will be a minor outbreak, in the sense that the number of infecteds will grow but then decrease and the disease will die out, or whether there will be a full-blown epidemic. Various networks are considered including networks where the values of p_k are given by the Poisson distribution, networks with superspreaders, small world networks and scale free networks.

The cholera model of [BCG⁺09] is interesting as it adds in a river modelling dynamic by describing the flow of disease between links by considering a water-borne method of transition. This is appropriate for cholera as it is transmitted by water contamination. In this research, different network topologies are considered including regular one-dimensional grids, regular two-dimensional grids, Peano networks (which are structured in a similar fashion to the water courses in a watershed) and optimal channel networks (which scale like real world networks). They investigate this model as a discrete space model over these networks and also as PDEs. Other research into cholera modelling [GMB⁺12, MBR⁺12] describes a model of spatial epidemics, illustrated with real examples of cholera epidemics in Haiti and in KwaZulu (South Africa), where space is modelled as two networks of nodes, with the same vertices but different edges, representing waterways and human mobility. The networks are reconstructed from GIS data, and provide a detailed but tractable model of geographical features, such as water basins, and of human mobility (using the so-called gravity model to model flows of people between locations). The model is given by a set of ODEs, duplicated for each node, and thus can be seen as a mean-field model. Gatto *et al* [GMB⁺12] characterise the reproduction number R_0 in terms of the eigenvalues of a matrix that takes the network structure into account, and provides a characterisation of epidemic outbreaks at each node by the corresponding eigenvector.

3.7.4 Networking

Computer networks, in particular ad hoc networks and mobile networks, often require spatial modelling for evaluation. For example, computer and mobile phone virus spread modelling involves spatial aspects and much of this research draws on epidemiological approaches [CGK03, HMCV09, KW91, KW93, PSV02, SPW02, WGH09, ZGT02]. Many of these approaches are focussed on the overall level of infection rather than an understanding of local details. Routing protocols may have spatial aspects that can be discrete or continuous and an overview can be found in [ZAS12].

Chaintreau *et al* [CLBR09] develop a mean-field model of movement and data ageing using real data collected from cabs in the San Francisco Bay area. The region is divided into 15 regular patches in a grid with a sixteenth patch representing the rest of the world (although only four patches connect directly to the sixteenth patch due to the geography). They first develop an ODE model of movement between patches which describes the number of cabs in each patch over time. This is later parameterised by the real data. The system that models how data ages is then described as a set of PDEs with one PDE for each patch. It takes into account movement between patches, the ageing of data, and how each cab can obtain younger data either from a base station or from other cabs. Rates to describe the occurrence of opportunistic contacts between cabs in the same patch, and in neighbouring patches, are again derived from the real data. Although the model is expressed as a PDE (due to two variables, time and age), the spatial aspect is treated discretely. The PDEs have a unique solution defined as a non-linear ODE problem. Analytic results are provided for the single-patch case, and the multi-patch case is approximated by considering low age and high age asymptotics. The paper provides some ideas for parameterisation when detailed GPS data is available.

A mean-field model is also constructed of peer-to-peer networking in the case of mobile networks and has similarities to gossip patch model [ZIM11]. ZebraNet, a delay-tolerant network involving wildlife has been modelled where each animal and the data collection vehicle are treated individually, giving a continuous space model [Fen12]. Recently, this has been transformed into a mean-field patch model where waterholes are used to divide the space into Voronoi cells which are then the patches of the model [Fen14].

3.7.5 Other

Propagation of forest fires is investigated in the context of Multi-class Multi-type Markovian Agent Model (M²MAM) [CGB⁺10]. The approach models individual agents in discrete space which may be regular or not and from this, a patch ODE model is derived. Forest fires have also been modelled using stochastic cellular automata as part of a climate model [LS11].

Irregular discrete space models have been used for a number of applications in several domains by members of the project. Among these domains are emergency egress, swarm robotics and crowd movement. The applications use discrete locations that are not situated in a regular grid, but that model relevant areas in space through which system components (people or robots) move.

The robotics case study consists of a swarm of robots that have to collectively identify a shortest path [MBL⁺13]. The division of a path in a number of separate sections considered as discrete locations provides a way to approximate the actual distribution of the traversal time by real robots.

The modelling of emergency egress from a multi-story building [MLB⁺12] involves a typical three story building with a number of building elements such as rooms, corridors and stairwells, doors and several exits, and a population of evacuees was modelled. The model included the commonly used human factors characteristics of average speed of walking, size of doors and corridors, and the characteristic number of people that can pass through areas such as stairwells that are used in the field of emergency egress modelling.

Other research concerns the understanding of the aspects of emergence, in the mathematical sense, of spontaneous drinking parties that is a common phenomenon in cities in the south of Spain [RG03]. In particular, the introduction of small variations that break symmetry, both in space and in the degree of connectivity between locations but also in the behaviour of the individuals can lead to new behaviour [BLM13].

Bike sharing systems have been modelled with homogeneous discrete space using a population CTMC approach with an associated mean-field model [FG12]. When space is not homogeneous, a clustering approach has been used to group similar locations together [FGM12].

4 Continuous space modelling techniques

Continuous space is more straightforward to define than discrete space. In this section, we will focus on two-dimensional space; however, both one- and three-dimensional space may be useful for modelling transport. Continuous space can either be the Euclidean plane extending infinitely in all directions or it can be a bounded connected (contiguous) subset of this plane. Points in the plane can be referred to by their coordinates $(x, y) \in \mathbb{R} \times \mathbb{R}$. As with discrete space, we can consider two cases.

4.1 Continuous space without state-based aggregation

In these models, we consider identifiable individuals. If I is an individual, then it has associated information, similar to the discrete state case.

- $\text{loc}(I, t) \in \mathbb{R} \times \mathbb{R}$ which is its location at time t , and
- $\text{state}(I, t) = A_i$ which is its state¹³ at time t .

There are rules which describe how the individual changes state that may take into account the individual's current location, and rules that describe an individual's movement through space which may take into account the individual's state. In the networking literature, there are a number of movement models and these will be discussed later in this section. As with discrete space, the rates for state change are exponential and can be functional. Unlike with discrete space, it is not useful to construct a Markov chain whose states are obtained from the locations and states of each individual.

In the case that the state is continuous, then

- $\text{state}(I, t) = Y$ which is a continuous variable representing its state at time t .

As with the discrete space case, some way is required that describes the change of state over time, and an ODE can be used for this. Some models require both discrete and continuous non-aggregated states and this requires a hybrid solution.

A different approach to modelling continuous state with continuous time is that of continuous time Markov processes (CTMP) [DP03]. A CTMP is a tuple (S, Σ, R, L) where (S, Σ) forms a specific type of topological manifold and $R : S \times \Sigma \rightarrow \mathbb{R}_{\geq 0}$ is a rate function which is measurable in its first coordinate and a measure on its second coordinate. L is a state labelling function. Applying this in the context of space, the manifold is $(\mathbb{R} \times \mathbb{R}, \Sigma)$ where Σ consists of the open sets of $\mathbb{R} \times \mathbb{R}$, hence defining a σ -algebra. A notion of path through this space can be defined describing the behaviour of an individual. Furthermore, if there are additional continuous quantities associated with the individual then additional dimensions of \mathbb{R} can be used.

4.2 Continuous space with state-based aggregation

When individuals are aggregated, there is no need to keep track of them individually and *densities* become more important. In spatio-temporal point processes¹⁴, each point in space (x, y) has an associated integral count for a state in a population at a specific point in time t . We can denote this as $N_{A_i}((x, y), t)$ and its behaviour is described by a function $\lambda((x, y), t)$. In general, λ can depend on all preceding events, but in the case of a Poisson process, it only depends on (x, y) and t [SBG02]. If λ is a constant, then there is no spatial heterogeneity. If the equation defining λ includes comparison with other points, then either clustering or inhibitory behaviour can be defined. If time and space are independent then λ can be defined by $\lambda((x, y), t) = \lambda_1(x, y)\lambda_2(t)$.

¹³As with discrete space, if the population P_A has multiple attributes $A^{[1]}, \dots, A^{[p]}$, then $\text{state}(I, t) = (A_{i_1}^{[1]}, \dots, A_{i_p}^{[p]})$ representing a tuple of states.

¹⁴In contrast to spatio-temporal point processes, spatial point processes describe distributions in space, and do not include a notion of change over time [BBS07] and hence are unsuitable for our purposes.

For continuous aggregation of populations, we now consider the classical model of movement in continuous space, partial differential equations. For populations described by $X_{A_i}((x, y), t)$, the general form is

$$F_i(x, y, t, X_{A_1}, \dots, X_{A_n}, \frac{\partial X_{A_i}}{\partial x}, \frac{\partial X_{A_i}}{\partial y}, \frac{\partial X_{A_i}}{\partial t}, \frac{\partial^2 X_{A_i}}{\partial x^2}, \frac{\partial^2 X_{A_i}}{\partial xy}, \frac{\partial^2 X_{A_i}}{\partial y^2}) = 0$$

if we assume that we are interested in second order partial derivatives over space only for the population $X_{A_i}((x, y), t)$. Note that writing the PDE in this form simply allows it to be described as a function over all the derivatives of interest rather than as a single partial derivative being equal to a function of other derivatives. There are various techniques for solving PDES, many of which involve discretising the area into a mesh [SSML03]

4.3 Examples of existing use

PDEs of reaction-diffusion type are very well understood in many disciplines, such as ecology [OL01], biology [Mur02], and chemistry [Van07]. It is beyond the scope of this document to provide a general overview of the literature. Instead, we focus on a few articles to illustrate the techniques used.

4.3.1 Ecology

Spatio-temporal point processes have been used to model plant growth and dispersal [BP97] and other applications [Dig05]. The local density of a point is defined in terms of a competition kernel that determines the level of competition in terms of distance from the point. This leads to two integro-differential moment equations: one for average density that includes the competition kernel and the spatial autocovariance that given a distance measures the association for that distance, namely the level of association between two points. The other moment equation is for the covariance of the density, and it is expressed in terms of higher moments, as described before. This equation is approximated by assuming that the third central moment is negligible and it is shown to be reasonably accurate for moderate neighbourhood sizes. Later work introduces a dispersal kernel that describes where offspring grow [BP99]. Markov random graphs on continuous space over continuous time can also be considered as spatio-temporal point processes [Ish81].

Holmes *et al* [HLBV94] review the use of PDEs in ecological applications. In the least spatially heterogeneous case, the PDE expressing Brownian (random) motion where movement rate is independent of time and space, is defined as follows

$$\frac{\partial X_{A_i}((x, y), t)}{\partial t} = D \left(\frac{\partial^2 X_{A_i}}{\partial x^2} + \frac{\partial^2 X_{A_i}}{\partial y^2} \right) = D \Delta X_{A_i}$$

where D is called the *diffusion constant* and

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

is called the *Laplacian*. This only captures movement and if interaction is to be included an additional term is needed, giving the following equation.

$$\frac{\partial X_{A_i}}{\partial t} = D \Delta X_{A_i} + f_i(X_{A_1}, \dots, X_{A_n})$$

This function captures the interaction of the various populations at a specific point and time. However, these PDEs do not take into account any spatial heterogeneity and also have a positive probability that an individual can move an arbitrarily long distance in an arbitrarily short time. The PDE above can be modified to include terms that represent drift due to wind or water currents.

$$\frac{\partial X_{A_i}}{\partial t} = D \Delta X_{A_i} - w_x \frac{\partial X_{A_i}}{\partial x} - w_y \frac{\partial X_{A_i}}{\partial y} + f_i(X_{A_1}, \dots, X_{A_n})$$

where w_x and w_y are drift velocities. Moving from Brownian motion to a more directed form of movement, the following PDE captures correlated motion. The PDE is also called the *telegraph* equation.

$$\frac{\partial X_{A_i}}{\partial t} = \frac{s^2}{2\lambda} \left(\frac{\partial^2 X_{A_i}}{\partial x^2} + \frac{\partial^2 X_{A_i}}{\partial y^2} \right) - \frac{1}{2\lambda} \frac{\partial^2 X_{A_i}}{\partial t^2} + f_i(X_{A_1}, \dots, X_{A_n})$$

where s is the speed of the individuals, and $1/2\lambda$ measures the correlation between directions of travel. To capture the effect of other individuals or of the environment, a function D that depends on density of the species or some feature of the space can be introduced, leading to the following PDE.

$$\frac{\partial X_{A_i}}{\partial t} = \frac{\partial}{\partial x} \left(D(X_{A_i}, (x, y)) \frac{\partial X_{A_i}}{\partial x} \right) + \frac{\partial}{\partial y} \left(D(X_{A_i}, (x, y)) \frac{\partial X_{A_i}}{\partial y} \right) + f_i(X_{A_1}, \dots, X_{A_n})$$

In this equation, the function D only depends on X_{A_i} and (x, y) but a function that also depends on X_{A_1}, \dots, X_{A_n} could be used, thus including effects from other subpopulations. However, the more complex the PDE, the more time is required to solve it numerically.

In another PDE model of migration, an expression is derived for the spatial autocorrelation and the scaling of this autocorrelation is investigated for strong and weak density regulation [ELS02].

4.3.2 Biology

Bittig and Uhrmacher [BU10] describe two continuous space approaches for cellular modelling.

particle space: In the case of *molecular* dynamics each particle is modelled individually and movement is done by rules, leading to some form of random walk. These models can be made more efficient by assuming that each particle is only affected by nearby events, so that the effect of other events can be excluded from consideration. Green's Function Reaction-Diffusion algorithm can be used if reactions consisting of multiple molecules are decomposed in bimolecular reactions, allowing for exact reactions.

gradients/PDEs: These have been discussed above. Often in biomolecular modelling, only simple diffusion is required. As discussed below the link between these models and those based on discrete regular space is an area of ongoing research.

Fange *et al* [FBSE10] describe three different techniques for spatially heterogeneous stochastic kinetics as *microscopic* when each individual particle is considered in terms of its position (continuous-space), as *mesoscopic* when the Reaction Diffusion Master Equation (RDME) is used (discrete space) and as *macroscopic* when PDEs are used. The mesoscopic approach is known to give unrealistic results for very fine discretisation and the authors extend the RDME to resolve this problem by introducing microscopic modelling within voxels. The RDME corresponds to the forward equations of a CTMC that models a network of biochemical reactions occurring at discrete sites, and molecular transitions across sites. The relationship with the PDE model [Isa09] can be established by means of a procedure that closes the equations for the average populations obtained by summing across the forward equations and approximates the expected value of a function of random variables with the function of their expectations, rather than a more rigorous limiting procedure in the sense of Kurtz [Kur81].

PDEs can also be obtained by taking the hydrodynamic limit of IPSs, namely as the number of particles tends to infinity [DMP91, DMFL86, DN94]. Chemical reactions have also been modelled as PDEs in a series of papers [AT80, Kot86, Blo91, Blo93, Blo94, Blo96, HO00]. However, in their

stochastic models, local reaction rates model only birth and death, and there are no interactions between multiple types of agents.

Random walks (RW) on regular space models involving individuals have been surveyed in [CPB08]. Starting either from regular discrete space, or continuous space, this movement is described by PDEs in the limit. Restrictions on the walk give rise to different variants including biased RW, RW with waiting times, RW with barriers, correlated RW, the telegraph equation, reinforced random walks and combinations of these. Some of these RW are no longer Markovian, for example, the correlated RW captures the tendency to move in the same direction requiring knowledge of what that direction is. The mean squared displacement (MSD) of each type of walk is used to determine whether a RW is realistic or can allow unrealistically large change in location in a small time interval. Othmer *et al* [ODA88] derive PDEs for walks consisting of pauses and jumps, as well as those consisting of alternation of direction choice and walks in the chosen direction. The former give rise to the diffusion PDE and the latter to the telegraph equation.

Lévy walks are based on a power-law distribution rather than an exponential distribution and are characterised by short distances and occasional long distances and have been used to investigate the movement of mussels [dJWH⁺11]. Other authors [JMP12] have argued that composite Brownian motion, expressed as the sum of multiple exponential distributions (a hyperexponential distribution) is a better description of this movement.

In an investigation of the movement of bacteria, random walks on stochastic cellular automata have been investigated in the limit giving rise to PDEs [OH02]. This work investigates when the dynamics provide aggregating, blowup or collapsing behaviour.

Pattern formation is important in biology, as mentioned above, and can be considered in regular discrete space or continuous space. An important PDE in this context is the Swift-Hohenberg equation

$$\frac{\partial u}{\partial t} = ru - (1 + \nabla^2)^2 u + N(u)$$

where $u(x, y, t)$ is a scalar function defined on the plane, r is a bifurcation parameter, and $N(u)$ is a smooth nonlinear function [SH77].

4.3.3 Epidemiology

Spatial point processes have been used to model the spread of foot and mouth disease [Dig06]. In terms of PDEs, Kendall [Ken65] proposed the first spatial epidemic model based on the Kermack-McKendrick nonspatial compartment model. Andersson and Djehiche [AD95] investigate a similar model and demonstrate that it is the limit of a one dimensional regular space model where there are populations at each position in regular space. The Kendall model can be extended to the Diekmann-Thieme model where traits of individuals affect both their susceptibility to infection and their infectiveness to other individuals [Rua07]. These equations contain a kernel $K(x, y, t - s)$ that describes interactions at point x at time t that are influenced by interactions at point y at s seconds before.

Wu [Wu08] uses Taylor series expansions and partial derivatives (as two different approaches) to illustrate how equations are derived and a diffusion coefficient is introduced to give the final form of the PDE which captures both change over time and change over space. This formalisation (which is Fick's law of diffusion) can be interpreted as individuals moving from areas of high infection to areas of low infection, although the author warns that outbreaks of disease can change movement behaviour which could make this model inappropriate.

Another aspect that is of interest in spatial disease modelling is generation of waves or fronts as disease moves through the population which is distributed in space. Murray [Mur03, Chapter 13] develops a one-dimensional continuous space PDE model of rabies outbreaks amongst foxes and derives solutions for the wavefront speed, and width of breaks that would halt the disease spread. He also describes a two-dimensional continuous space model that takes into account local fox densities and carrying capacities of an outbreak of rabies in the south of England. A similar two-dimensional

continuous space model is extended to consider both natural immunity of foxes to rabies and size of breaks to curb the spread of the disease.

The cholera model of [BCG⁺09] is also defined in one- and two-dimensional space as PDEs using diffusion terms that take into account the concentration of cholera in the water. General PDE models are considered in [Wu08], and also in [MZ08, MM91] where directed movement such as moving away from areas of high infection and/or overcrowding is modelled and in [LZ09] where latency of disease and mobility of infected individuals is modelled.

4.3.4 Networking

There is a substantial amount of work on mobility models, both at the analytical level and experimentally through traces. We briefly review the literature; for a more extensive discussion, we refer the reader to various surveys [CBD02, MM09, Bet01]. Group mobility has also been studied [HGPC99]. The importance of mobility models where average speed does not decay is highlighted in [YLN03] where a technique to transform models to be sound is presented.

Owing to its analytical tractability, the random walk (RW) model has been extensively studied in networking research. Unbiased RWs are proposed in [AHL96] and [ALLC00] to study movements across cellular networks, and to study routing protocols [IM06, GV06] and performance characteristics in ad hoc networks [GMPS04]. Direction based on an individual's state is a feature of a discrete time Markovian model developed for the comparison of update strategies in cellular networks [BNKS94]. In all cases, mobility is not coupled with a model of interaction between nodes. RWs are also used in [DT05] as the basic mobility model by which the authors arrive at a reaction-diffusion type equation for information propagation in ad hoc networks; their analysis is carried out at the deterministic/macroscale level without considering a limiting regime of a counterpart stochastic/microscopic process. Instead, the PDEs used in [GL07] are interpreted as the deterministic limit of the empirical measure of node concentrations, by appealing to the strong law of large numbers.

Connectivity models provide an abstraction of mobility models in that they provide information about intercontact time [KP07, KR12, CFB09, CMRM07]. It is often assumed that these intercontact times are exponentially distributed; however analysis of experimental data has shown that they are better approximated by a power law distribution [CHC⁺07].

Stochastic geometry is applied to wireless networks [BB09a] and is used to evaluate the space and time performance of MAC (Medium Access Control) networks and mobile ad hoc networks [BB09b]. Epidemiological approaches have been taken to routing in networks and evaluation is often done assuming a nonspatial model. Klein *et al* develop a PDE to describe routing behaviour and travelling waves that occur [KHM10].

5 Conclusion

This document has described in some detail various mathematical techniques for discrete space and continuous space modelling. As Deliverable 2.1 [GBC⁺14] concludes, patch models are the most relevant approaches for the smart transport case studies in the QUANTICOL project. As the literature review has demonstrated, there are many techniques that can be used when dealing with mean-field models that involve space, and these will be investigated further, building from this technical report. This investigation will take two forms: development of new techniques relevant to the QUANTICOL case studies, and modification of existing techniques for determination of local rather than global measures.

References

- [ACVP00] P. Auger, S. Charles, M. Viala, and J.-C. Poggiale. Aggregation and emergence in ecological modelling: integration of ecological levels. *Ecological Modelling*, 127(1):11–20, 2000.
- [AD95] H. Andersson and B. Djehiche. Limit theorems for multitype epidemics. *Stochastic processes and their applications*, 56(1):57–75, 1995.
- [ADH⁺05] J. Arino, J.R. Davis, D. Hartley, R. Jordan, J.M. Miller, and P. van den Driessche. A multi-species epidemic model with spatial dynamics. *Mathematical Medicine and Biology*, 22(2):129–142, 2005.
- [AHL96] I.F. Akyildiz, J.S.M. Ho, and Y.-B. Lin. Movement-based location update and selective paging for PCS networks. *IEEE/ACM Transactions on Networking*, 4(4):629–638, 1996.
- [ALLC00] I.F. Akyildiz, Y.-B. Lin, W.-R. Lai, and R.-J. Chen. A new random walk model for PCS networks. *IEEE Journal on Selected Areas in Communications*, 18(7):1254–1260, 2000.
- [AMR95] U.M. Ascher, R.M.M. Mattheij, and R.D. Russell. *Numerical solution of boundary value problems for ordinary differential equations*. SIAM, 1995.
- [AP02] F. Arrigoni and A. Pugliese. Limits of a multi-patch SIS epidemic model. *Journal of Mathematical Biology*, 45(5):419–440, November 2002.
- [APS12] P. Auger, J.C. Poggiale, and E. Sánchez. A review on spatial aggregation methods involving several time scales. *Ecological Complexity*, 10:12–25, June 2012.
- [AT80] L. Arnold and M. Theodosopulu. Deterministic limit of the stochastic model of chemical reactions with diffusion. *Advances in Applied Probability*, 12(2):367–379, 1980.
- [AvdD03] J. Arino and P. van den Driessche. A multi-city epidemic model. *Mathematical Population Studies*, 10(3):175–193, 2003.
- [BB09a] F. Baccelli and B. Błaszczyszyn. *Stochastic Geometry and Wireless Networks: Volume I: Theory*. NOW Publishers, 2009.
- [BB09b] F. Baccelli and B. Błaszczyszyn. *Stochastic Geometry and Wireless Networks: Volume II: Application*. NOW Publishers, 2009.
- [BBS07] A. Baddeley, I. Bárány, and R. Schneider. Spatial point processes and their applications. *Stochastic Geometry: Lectures given at the CIME Summer School 2004*, pages 1–75, 2007.
- [BCG⁺09] E. Bertuzzo, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo. On spatially explicit models of cholera epidemics. *Journal of The Royal Society Interface*, 7(43):321–333, 2009.
- [Ber02] L. Berec. Techniques of spatially explicit individual-based models: construction, simulation, and mean-field analysis. *Ecological Modelling*, 150(1):55–81, 2002.
- [Bet01] C. Bettstetter. Smooth is better than sharp: a random mobility model for simulation of wireless networks. In *4th ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM 2001)*, pages 19–27. ACM, 2001.
- [BG90] C. Bezuidenhout and G. Grimmett. The critical contact process dies out. *The Annals of Probability*, 18(4):1462–1482, 1990.
- [BG95] B. Bolker and B. Grenfell. Space, persistence and dynamics of measles epidemics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 348(1325):309–320, 1995.
- [BHLM13] L. Bortolussi, J. Hillston, D. Latella, and M. Massink. Continuous approximation of collective systems behaviour: a tutorial. *Performance Evaluation*, 70(5):317–349, 2013.
- [BLM13] L. Bortolussi, D. Latella, and M. Massink. Stochastic process algebra and stability analysis of collective systems. In *15th International Conference on Coordination Models and Languages (COORDINATION 2013)*, LNCS 7890, pages 1–15. Springer, 2013.
- [Blo91] D. Blount. Comparison of stochastic and deterministic models of a linear chemical reaction with diffusion. *The Annals of Probability*, 19(4):1440–1462, 1991.

- [Blo93] D. Blount. Limit theorems for a sequence of nonlinear reaction-diffusion systems. *Stochastic Processes and their Applications*, 45(2):193–207, 1993.
- [Blo94] D. Blount. Density-dependent limits for a nonlinear reaction-diffusion model. *The Annals of Probability*, 22(4):2040–2070, 1994.
- [Blo96] D. Blount. Diffusion limits for a nonlinear density dependent space-time population model. *The Annals of Probability*, 24(2):639–659, 1996.
- [BNKS94] A. Bar-Noy, I. Kessler, and M. Sidi. Mobile users: To update or not to update? In *INFOCOM 1994*, pages 570–576, 1994.
- [BP97] B. Bolker and S.W. Pacala. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theoretical Population Biology*, 52(3):179–197, 1997.
- [BP99] B.M. Bolker and S.W. Pacala. Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. *The American Naturalist*, 153(6):575–602, 1999.
- [BP10] L. Bortolussi and A. Policriti. Hybrid dynamics of stochastic programs. *Theoretical Computer Science*, 411:2052–2077, 2010.
- [Bra08a] F. Brauer. Compartmental models in epidemiology. In L.J.S. Allen, F. Brauer, P. van den Driessche, and J. Wu, editors, *Mathematical Epidemiology*, pages 19–80. Springer, 2008.
- [Bra08b] F. Brauer. An introduction to networks in epidemic modeling. In L.J.S. Allen, F. Brauer, P. van den Driessche, and J. Wu, editors, *Mathematical Epidemiology*, pages 133–146. Springer, 2008.
- [Bru67] S.G. Brush. History of the Lenz-Ising model. *Reviews of Modern Physics*, 39(4):883, 1967.
- [BU10] A.T. Bittig and A.M. Uhrmacher. Spatial modeling in cell biology at multiple levels. In *Winter Simulation Conference (WSC 2010)*, pages 608–619. IEEE, 2010.
- [CBD02] T. Camp, J. Boleng, and V. Davies. A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing*, 2(5):483–502, 2002.
- [CFB09] H. Conceição, M. Ferreira, and J. Barros. A cautionary view of mobility and connectivity modeling in vehicular ad-hoc networks. In *IEEE 69th Vehicular Technology Conference (VTC 2009)*, pages 1–5. IEEE, 2009.
- [CGB⁺10] D. Cerotti, M. Gribaudo, A. Bobbio, C.T. Calafate, and P. Manzoni. A Markovian agent model for fire propagation in outdoor environments. In *Seventh European Performance Engineering Workshop (EPEW 2010)*, LNCS 6342, pages 131–146. Springer, 2010.
- [CGK03] Z. Chen, L. Gao, and K. Kwiaty. Modeling the spread of active worms. In *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications (INFOCOM 2003)*, volume 3, pages 1890–1900. IEEE, 2003.
- [CHC⁺07] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of human mobility on opportunistic forwarding algorithms. *IEEE Transactions on Mobile Computing*, 6(6):606–620, June 2007.
- [Che81] P.L. Chesson. Models for spatially distributed populations: the effect of within-patch variability. *Theoretical Population Biology*, 19(3):288–325, 1981.
- [Che12] P. Chesson. Scale transition theory: its aims, motivations and predictions. *Ecological Complexity*, 10:52–68, 2012.
- [CHK⁺05] R. Chaturvedi, C. Huang, B. Kazmierczak, T. Schneider, J.A. Izaguirre, T. Glimm, H.G.E. Hentschel, J.A. Glazier, S.A. Newman, and M.S. Alber. On multiscale approaches to three-dimensional modelling of morphogenesis. *Journal of the Royal Society Interface*, 2(3):237–253, 2005.
- [CHM92] H.N. Comins, M.P. Hassell, and R.M. May. The spatial dynamics of host-parasitoid systems. *Journal of Animal Ecology*, 61(3):735–748, 1992.
- [CL07] P.M. Chaikin and T.C. Lubensky. *Principles of Condensed Matter Physics*. Cambridge University Press, 4th edition, 2007.

- [CLBR09] A. Chaintreau, J.-Y. Le Boudec, and N. Ristanovic. The age of gossip: spatial mean field regime. In *Eleventh International Joint Conference on Measurement and Modeling of Computer Systems (SIGMETRICS/Performance 2009)*, pages 109–120. ACM, 2009.
- [CLM14] V. Ciancia, D. Latella, and M. Massink. Logics of space and time. Technical Report TR-QC-01-2014, QUANTICOL, 2014.
- [CMRM07] R. Calegari, M. Musolesi, F. Raimondi, and C. Mascolo. CTG: a connectivity trace generator for testing the performance of opportunistic mobile systems. In *6th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering*, pages 415–424. ACM, 2007.
- [CPB08] E.A. Codling, M.J. Plank, and S. Benhamou. Random walk models in biology. *Journal of the Royal Society Interface*, 5(25):813–834, 2008.
- [Dav93] M.H.A. Davis. *Markov Models and Optimization*. Chapman & Hall, 1993.
- [Dig05] P.J. Diggle. Spatio-temporal point processes: methods and applications. Working Paper 78, Department of Biostatistics, Johns Hopkins University, 2005. <http://biostats.bepress.com/jhubiostat/paper78>.
- [Dig06] P.J. Diggle. Spatio-temporal point processes, partial likelihood, foot and mouth disease. *Statistical Methods in Medical Research*, 15(4):325–336, August 2006.
- [dJWH⁺11] M. de Jager, F.J. Weissing, P.M.J. Herman, B.A. Nolet, and J. van de Koppel. Levy walks evolve through interaction between movement and environmental complexity. *Science*, 332(6037):1551–1553, June 2011.
- [DL94a] R. Durrett and S. Levin. The importance of being discrete (and spatial). *Theoretical Population Biology*, 46(3):363–394, 1994.
- [DL94b] R. Durrett and S.A. Levin. Stochastic spatial models: a user’s guide to ecological applications. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 343(1305):329–350, 1994.
- [DL98] R. Durrett and S. Levin. Spatial aspects of interspecific competition. *Theoretical Population Biology*, 53(1):30–43, 1998.
- [DMFL86] A. De Masi, P. A. Ferrari, and J. L. Lebowitz. Reaction-diffusion equations for interacting particle systems. *Journal of Statistical Physics*, 44(3-4):589–644, 1986.
- [DMP91] A. De Masi and E. Presutti. *Mathematical Methods for Hydrodynamic Limits*. Lecture Notes in Mathematics. Springer-Verlag, 1991.
- [DN94] R. Durrett and C. Neuhauser. Particle systems and reaction-diffusion equations. *The Annals of Probability*, 22(1):289–333, 1994.
- [DN08] R.W.R. Darling and J.R. Norris. Differential equation approximations for Markov chains. *Probability Surveys*, 5:37–79, 2008.
- [DP03] J. Desharnais and P. Panangaden. Continuous stochastic logic characterizes bisimulation of continuous-time markov processes. *Journal of Logic and Algebraic Programming*, 56(1-2):99–115, 2003.
- [DPR⁺13] S. Dorjee, Z. Poljak, C.W. Revie, J. Bridgland, B. McNab, E. Leger, and J. Sanchez. A review of simulation modelling approaches used for the spread of zoonotic influenza viruses in animal and human populations. *Zoonoses and Public Health*, 60(6):383–411, 2013.
- [DT05] M. Durvy and P. Thiran. Reaction-diffusion based transmission patterns for ad hoc networks. In *INFOCOM 2005*, volume 3, pages 2195–2205, 2005.
- [EC09] R. Erban and S.J. Chapman. Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions. *Physical Biology*, 6(4):046001, 2009.
- [ECM07] R. Erban, J. Chapman, and P. Maini. A practical guide to stochastic simulations of reaction-diffusion processes. *arXiv preprint arXiv:0704.1908*, 2007.
- [EE04] J. Elf and M. Ehrenberg. Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *Systems Biology*, 1(2):230–236, 2004.

- [Ell01] S.P. Ellner. Pair approximation for lattice models with multiple interaction scales. *Journal of Theoretical Biology*, 210(4):435–447, June 2001.
- [ELS02] S. Engen, R. Lande, and B.-E. Sæther. Migration and spatiotemporal variation in population dynamics in a heterogeneous environment. *Ecology*, 83(2):570–579, 2002.
- [FBSE10] D. Fange, O.G. Berg, P. Sjöberg, and J. Elf. Stochastic reaction-diffusion kinetics in the microscopic limit. *Proceedings of the National Academy of Sciences*, 107(46):19820–19825, 2010.
- [FE06] D. Fange and J. Elf. Noise-induced min phenotypes in *e. coli*. *PLoS Computational Biology*, 2(6):e80, 2006.
- [Fen12] C. Feng. *Modelling opportunistic networks with HYPE*. MSc dissertation, School of Informatics, University of Edinburgh, 2012.
- [Fen14] C. Feng. Patch-based hybrid modelling of spatially distributed systems by using stochastic HYPE – ZebraNet as an example. In *Twelfth International Workshop on Quantitative Aspects of Programming Languages and Systems (QAPL 2014)*, 2014.
- [FG12] C. Fricker and N. Gast. Incentives and regulations in bike-sharing systems with stations of finite capacity. arXiv preprint arXiv:1201.1178, 2012.
- [FGM12] C. Fricker, N. Gast, and H. Mohamed. Mean field analysis for inhomogeneous bike sharing systems. *DMTCS Proceedings*, 01:365–376, 2012.
- [FS86] O. Frank and D. Strauss. Markov graphs. *Journal of the American Statistical Association*, 81(395):832–842, 1986.
- [Gar76] A. Gardiner. Homogeneous graphs. *Journal of Combinatorial Theory, Series B*, 20(1):94–102, 1976.
- [GB00] M.A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *Journal of Physical Chemistry A*, 104(9):1876–1889, 2000.
- [GBC⁺14] V. Galpin, L. Bortolussi, V. Ciancia, A. Clark, R. De Nicola, C. Feng, S. Gilmore, N. Gast, J. Hillston, A. Lluch-Lafuente, M. Loreti, M. Massink, L. Nenzi, D. Reijbergen, V. Senni, F. Tiezzi, M. Tribastone, and M. Tschaikowski. A preliminary investigation of capturing spatial information for CAS. QUANTICOL Deliverable 2.1, March 2014.
- [GG92] F. Graner and J.A. Glazier. Simulation of biological cell sorting using a two-dimensional extended Potts model. *Physical Review Letters*, 69(13):2013, 1992.
- [Gil77] D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
- [Gil07] D.T. Gillespie. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, 58:35–55, 2007.
- [GL07] M. Garetto and E. Leonardi. Analysis of random mobility models with partial differential equations. *IEEE Transactions on Mobile Computing*, 6(11):1204–1217, 2007.
- [GMB⁺12] M. Gatto, L. Mari, E. Bertuzzo, R. Casagrandi, L. Righetto, I. Rodriguez-Iturbe, and A. Rinaldo. Generalized reproduction numbers and the prediction of patterns in waterborne disease. *Proceedings of the National Academy of Sciences*, 109(48):19703–19708, 2012.
- [GMPS04] A.El. Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay trade-off in wireless networks. In *INFOCOM 2004*, 2004.
- [GMWM76] C.W. Gardiner, K.J. McNeil, D.F. Walls, and I.S. Matheson. Correlations in stochastic theories of chemical reactions. *Journal of Statistical Physics*, 14(4):307–331, 1976.
- [Gre92] B.T. Grenfell. Chance and chaos in measles dynamics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 54(2):383–398, 1992.
- [GV06] M. Grossglauser and M. Vetterli. Locating mobile nodes with EASE: learning efficient routes from encounter histories alone. *IEEE/ACM Transactions on Networks*, 14(3):457–469, 2006.
- [Har74] T.E. Harris. Contact interactions on a lattice. *The Annals of Probability*, 2(6):969–988, 1974.

- [HCM91] M.P. Hassell, H.N. Comins, and R.M. May. Spatial structure and chaos in insect population dynamics. *Nature*, 353(6341):255–258, 1991.
- [HGPC99] X. Hong, M. Gerla, G. Pei, and C.-C. Chiang. A group mobility model for ad hoc wireless networks. In *2nd ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pages 53–60. ACM, 1999.
- [HLBV94] E.E. Holmes, M.A. Lewis, J.E. Banks, and R.R. Veit. Partial differential equations in ecology: spatial interactions and population dynamics. *Ecology*, 75(1):17–29, 1994.
- [HMCV09] H. Hu, S. Myers, V. Colizza, and A. Vespignani. WiFi networks and malware epidemiology. *Proceedings of the National Academy of Sciences*, 106(5):1318–1323, January 2009.
- [HMW96] R.J. Hendry, J.M. McGlade, and J. Weiner. A coupled map lattice model of the growth of plant monocultures. *Ecological Modelling*, 84(1):81–90, 1996.
- [HO00] T. Hillen and H.G. Othmer. The diffusion limit of transport equations derived from velocity-jump processes. *SIAM Journal on Applied Mathematics*, 61(3):751775, 2000.
- [II09] S.A. Isaacson and D. Isaacson. Reaction-diffusion master equation, diffusion-limited reactions, and singular potentials. *Physical Review E*, 80(6):066106, 2009.
- [Ila01] A. Ilachinski. *Cellular Automata: A Discrete Universe*. World Scientific, 2001.
- [IM06] S. Ioannidis and P. Marbach. A Brownian motion model for last encounter routing. In *INFOCOM 2006*, 2006.
- [Isa08] S.A. Isaacson. Relationship between the reaction-diffusion master equation and particle tracking models. *Journal of Physics A: Mathematical and Theoretical*, 41(6):065003, 2008.
- [Isa09] S.A. Isaacson. The reaction-diffusion master equation as an asymptotic approximation of diffusion to a small target. *SIAM Journal on Applied Mathematics*, 70(1):77–111, January 2009.
- [Ish81] V. Isham. An introduction to spatial point processes and Markov random fields. *International Statistical Review/Revue Internationale de Statistique*, 41(9):21–43, 1981.
- [JL00] V.A.A. Jansen and A.L. Lloyd. Local stability analysis of spatially homogeneous solutions of multi-patch systems. *Journal of Mathematical Biology*, 41(3):232–252, September 2000.
- [JMP12] V.A.A. Jansen, A. Mashanova, and S. Petrovskii. Comment on "Levy walks evolve through interaction between movement and environmental complexity". *Science*, 335(6071):918–918, 2012.
- [Kan98] K. Kaneko. Diversity, stability, and metadynamics: Remarks from coupled map studies. In J. Bascompte and R.V. Solé, editors, *Modeling Spatiotemporal Dynamics in Ecology*, pages 27–45. Springer-Verlag, 1998.
- [KCMG05] I. Krishnarajah, A. Cook, G. Marion, and G. Gibson. Novel moment closure approximations in stochastic epidemics. *Bulletin of Mathematical Biology*, 67(4):855–873, 2005.
- [Ken65] D.G. Kendall. Mathematical models of the spread of infection. In *Mathematics and Computer Science in Biology and Medicine*, pages 213–225. Medical Research Council London, 1965.
- [KHM10] D.J. Klein, J. Hespanha, and U. Madhow. A reaction-diffusion model for epidemic routing in sparsely connected MANETs. In *INFOCOM 2010*, pages 1–9. IEEE, 2010.
- [Kot86] P. Kotelenez. Law of large numbers and central limit theorem for linear chemical reactions with diffusion. *The Annals of Probability*, 14(1):173–193, 1986.
- [KP07] T. Kathiravelu and A. Pears. Reproducing opportunistic connectivity traces using connectivity models. In *2007 ACM CoNEXT conference*, page 34. ACM, 2007.
- [KR12] T. Kathiravelu and N. Ranasinghe. Self similarity and predictability of contact opportunities in opportunistic networks. In *18th IEEE International Conference on Networks (ICON 2012)*, pages 459–464. IEEE, 2012.
- [KS76] J.G. Kemeny and J.L. Snell. *Finite Markov Chains*. Springer, 1976.
- [Kur81] T.G. Kurtz. *Approximation of Population Processes*. SIAM, 1981.

- [KW91] J.O. Kephart and S.R. White. Directed-graph epidemiological models of computer viruses. In *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, pages 343–359. IEEE, 1991.
- [KW93] J.O. Kephart and S.R. White. Measuring and modeling computer virus prevalence. In *IEEE Computer Society Symposium on Research in Security and Privacy*, page 215. IEEE, 1993.
- [LD96] S.A. Levin and R. Durrett. From individuals to epidemics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1347):1615–1621, 1996.
- [Lev69] R. Levins. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the ESA*, 15(3):237–240, 1969.
- [Lev74] S.A. Levin. Dispersion and population interactions. *American Naturalist*, 108(960):207–228, 1974.
- [Lev76] S.A. Levin. Population dynamic models in heterogeneous environments. *Annual Review of Ecology and Systematics*, 7(1):287–310, 1976.
- [LP74] S.A. Levin and R.T. Paine. Disturbance, patch formation, and community structure. *Proceedings of the National Academy of Sciences*, 71(7):2744–2747, 1974.
- [LS85] S.A. Levin and L.A. Segel. Pattern generation in space and aspect. *SIAM Review*, 27(1):45–67, 1985.
- [LS11] K. Lichtenegger and W. Schappacher. A carbon-cycle-based stochastic cellular automata climate model. *International Journal of Modern Physics C*, 22(06):607–621, 2011.
- [LZ09] J. Li and X. Zou. Modeling spatial spread of infectious diseases with a fixed latent period in a spatially continuous domain. *Bulletin of Mathematical Biology*, 71(8):2048–2079, 2009.
- [MBL⁺13] M. Massink, M. Brambilla, D. Latella, M. Dorigo, and M. Birattari. On the use of Bio-PEPA for modelling and analysing collective behaviors in swarm intelligence. *Swarm Intelligence*, 7(2-3):201–228, 2013.
- [MBR⁺12] L. Mari, E. Bertuzzo, L. Righetto, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo. Modelling cholera epidemics: the role of waterways, human mobility and sanitation. *Journal of The Royal Society Interface*, 9(67):376–388, 2012.
- [MLB⁺12] M. Massink, D. Latella, A. Bracciali, M.D. Harrison, and J. Hillston. Scalable context-dependent analysis of emergency egress models. *Formal Aspects of Computing*, 24(2):267–302, 2012.
- [MM91] D.B. Meade and F.A. Milner. An SIR model for epidemics with diffusion to avoid infection and overcrowding. In *13th IMACS World Congress on Computation and Applied Mathematics*, volume 3, pages 1444–1445, 1991.
- [MM09] M. Musolesi and C. Mascolo. Mobility models for systems evaluation. In B. Garbinato, H. Miranda, and L. Rodrigues, editors, *Middleware for Network Eccentric and Mobile Applications*, pages 43–62. Springer, 2009.
- [MMRL02] G. Marion, X. Mao, E. Renshaw, and J. Liu. Spatial heterogeneity and the stability of reaction states in autocatalysis. *Physical Review E*, 66(5):051915, November 2002.
- [MNS11] C. McCaig, R. Norman, and C. Shankland. From individuals to populations: a mean field semantics for process algebra. *Theoretical Computer Science*, 412(17):1557–1580, 2011.
- [Mol77] D. Mollison. Spatial contact models for ecological and epidemic spread. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(3):283–326, 1977.
- [MOSS92] H. Matsuda, N. Ogita, A. Sasaki, and K. Satō. Statistical mechanics of population: The lattice Lotka-Volterra model. *Progress of Theoretical Physics*, 88(6):1035–1049, 1992.
- [MP12] A. Morozov and J.-C. Poggiale. From spatially explicit ecological models to mean-field dynamics: the state of the art and perspectives. *Ecological Complexity*, 10:1–11, 2012.
- [MSH05] G. Marion, D.L. Swain, and M.R. Hutchings. Understanding foraging behaviour in spatially heterogeneous environments. *Journal of Theoretical Biology*, 232(1):127–142, 2005.
- [Mur02] J.D. Murray. *Mathematical Biology I: An Introduction*. Springer, 3rd edition, 2002.

- [Mur03] J.D. Murray. *Mathematical Biology II: Spatial Models and Biomedical Applications*. Springer, 2003.
- [MZ08] F.A. Milner and R. Zhao. SIR model with directed spatial diffusion. *Mathematical Population Studies*, 15(3):160–181, 2008.
- [Nor98] J.R. Norris. *Markov Chains*. Cambridge University Press, 1998.
- [ODA88] H.G. Othmer, S.R. Dunbar, and W. Alt. Models of dispersal in biological systems. *Journal of Mathematical Biology*, 26(3):263–298, 1988.
- [Oel89] K. Oelschläger. On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes. *Probability Theory and Related Fields*, 82(4):565–586, 1989.
- [OH02] H.G. Othmer and T. Hillen. The diffusion limit of transport equations II: chemotaxis equations. *SIAM Journal on Applied Mathematics*, 62(4):1222–1250, 2002.
- [OL01] A. Okubo and S. A. Levin. *Diffusion and Ecological Problems: Modern Perspectives*. Springer, 2001.
- [OS71] H.G. Othmer and L.E. Scriven. Instability and dynamic pattern in cellular networks. *Journal of Theoretical Biology*, 32(3):507–537, 1971.
- [OS97] H.G. Othmer and A. Stevens. Aggregation, blowup, and collapse: The ABC’s of taxis in reinforced random walks. *SIAM Journal on Applied Mathematics*, 57(4):1044–1081, 1997.
- [Pan09] P. Panangaden. *Labelled Markov Processes*. Imperial College Press, 2009.
- [PRL11] M. Pascual, M. Roy, and K. Laneri. Simple models for complex systems: exploiting the relationship between local and global densities. *Theoretical Ecology*, 4(2):211–222, 2011.
- [PSV02] R. Pastor-Satorras and A. Vespignani. Immunization of complex networks. *Physical Review E*, 65(3):036104, 2002.
- [RA97] C.J. Rhodes and R.M. Anderson. Epidemic thresholds and vaccination in a lattice model of disease spread. *Theoretical Population Biology*, 52(2):101–118, 1997.
- [RG03] J. E. Rowe and R. Gomez. El Botellón: Modeling the movement of crowds in a city. *Complex Systems*, 14:363–370, 2003.
- [Ril07] S. Riley. Large-scale spatial-transmission models of infectious disease. *Science*, 316(5829):1298–1301, 2007.
- [Rua07] S. Ruan. Spatial-temporal dynamics in nonlocal epidemiological models. In *Mathematics for Life Science and Medicine*, pages 97–122. Springer, 2007.
- [SBG02] F.P. Schoenberg, D.R. Brillinger, and P. Guttorp. Point processes, spatial-temporal. In *Encyclopedia of Environmetrics*, volume 3, pages 1573–1577. Wiley, 2002.
- [SH77] J. Swift and P.C. Hohenberg. Hydrodynamic fluctuations at the convective instability. *Physical Review A*, 15(1):319, 1977.
- [SPW02] S. Staniford, V. Paxson, and N. Weaver. How to own the internet in your spare time. In *USENIX Security Symposium*, pages 149–167, 2002.
- [SSML03] B.M. Slepchenko, J.C. Schaff, I. Macara, and L.M. Loew. Quantitative cell biology with the virtual cell. *Trends in Cell Biology*, 13(11):570–576, 2003.
- [Ste00a] A. Stevens. The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems. *SIAM Journal on Applied Mathematics*, 61(1):183–212, 2000.
- [Ste00b] A. Stevens. A stochastic cellular automaton modeling gliding and aggregation of myxobacteria. *SIAM Journal on Applied Mathematics*, 61(1):172–182, 2000.
- [Tak96] Y. Takeuchi. *Global Dynamical Properties of Lotka-Volterra Systems*. World Scientific, 1996.
- [TCG+14] M. Tribastone, A. Clark, N. Gast, S. Gilmore, and D. Reijnders. Data validation and requirements for case studies. QUANTICOL Deliverable 5.1, March 2014.

- [TGH12] M. Tribastone, S. Gilmore, and J. Hillston. Scalable differential analysis of process algebra models. *IEEE Transactions on Software Engineering*, 38(1):205–219, 2012.
- [TT13] M. Tschaikowski and M. Tribastone. Spatial fluid limits for stochastic mobile networks. *arXiv preprint arXiv:1307.4566*, 2013.
- [Tur53] A.M. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society (Part B)*, 237:37–72, 1953.
- [Van07] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, 3rd edition, 2007.
- [vdD08] P. van den Driessche. Spatial structure: Patch models. In L.J.S. Allen, F. Brauer, P. van den Driessche, and J. Wu, editors, *Mathematical Epidemiology*, pages 179–190. Springer, 2008.
- [WGHB09] P. Wang, M.C. Gonzalez, C.A. Hidalgo, and A.-L. Barabasi. Understanding the spreading patterns of mobile phone viruses. *Science*, 324(5930):1071–1076, April 2009.
- [WHXY11] B. Wang, B. Hou, F. Xing, and Y. Yao. Abstract Next Subvolume Method: A logical process-based approach for spatial stochastic simulation of chemical reactions. *Computational Biology and Chemistry*, 35(3):193–198, 2011.
- [Wik13a] Wikipedia. Lattice graph, 2013. Last accessed: 4 Feb 2013, http://en.wikipedia.org/wiki/Lattice_graph.
- [Wik13b] Wikipedia. Topological map, 2013. Last accessed: 4 Feb 2013, http://en.wikipedia.org/wiki/Topological_map.
- [WKB07] S.D. Webb, M.J. Keeling, and M. Boots. Host-parasite interactions between the local and the mean-field: how and when does spatial population structure matter? *Journal of Theoretical Biology*, 249(1):140–152, 2007.
- [WL95] J. Wu and O.L. Loucks. From balance of nature to hierarchical patch dynamics: a paradigm shift in ecology. *The Quarterly Review of Biology*, 70(4):439–466, 1995.
- [Wu08] J. Wu. Spatial structure: partial differential equations models. In L.J.S. Allen, F. Brauer, P. van den Driessche, and J. Wu, editors, *Mathematical Epidemiology*, pages 191–204. Springer, 2008.
- [XFAS06] D. Xu, Z. Feng, L.J.S. Allen, and R.K. Swihart. A spatially structured metapopulation model with patch dynamics. *Journal of Theoretical Biology*, 239(4):469–481, 2006.
- [YLN03] J. Yoon, M. Liu, and B. Noble. Sound mobility models. In *9th Annual International Conference on Mobile Computing and Networking*, pages 205–216. ACM, 2003.
- [ZAS12] A.M. Zungeru, L.-M. Ang, and K. P. Seng. Classical and swarm intelligence based routing protocols for wireless sensor networks: A survey and comparison. *Journal of Network and Computer Applications*, 35(5):1508–1536, September 2012.
- [ZGT02] C.C. Zou, W. Gong, and D. Towsley. Code red worm propagation modeling and analysis. In *9th ACM Conference on Computer and Communications Security*, pages 138–147. ACM, 2002.
- [ZIM11] X. Zhou, S. Ioannidis, and L. Massoulié. On the stability and optimality of universal swarms. *ACM SIGMETRICS Performance Evaluation Review*, 39(1):301–312, 2011.