Quantitative Modelling Approaches

Chapter · January 2018

DOI: 10.1016/B978-0-12-809633-8.20454-8

4 authors:

Filippo Castiglione
Italian National Research Council
200 PUBLICATIONS **3,038** CITATIONS
SEE PROFILE

Emiliano Mancini
Hasselt University
23 PUBLICATIONS **94** CITATIONS
SEE PROFILE

Marco Pedicini
Università Degli Studi Roma Tre
55 PUBLICATIONS **397** CITATIONS
SEE PROFILE

Abdul Salam Jarrah
American University of Sharjah
83 PUBLICATIONS **1,140** CITATIONS
SEE PROFILE

20454 Quantitative Modelling Approaches

Filippo Castiglione, Institute for Applied Computing - National Research Council of Italy, Rome, Italy

Emiliano Mancini, Institute for Advanced Studies - University of Amsterdam, Amsterdam, Netherlands

Marco Pedicini, Department of Mathematics and Physics - Roma Tre University, Rome, Italy

Abdul Salam Jarrah, Department of Mathematics and Statistics - American University of Sharjah, Sharjah, UAE

Glossary

Complexity science A field of study aiming at defining methods and tools to assess and predict the dynamics of a complex system using a holistic approach.
Complex system A hierarchical self-organized system with a large number of components characterized by non-linear connections. In such systems order is an emergent property of its internal microscopic dynamics.
Computation All steps necessary for a digital computer to solve a problem. A computation transforms the input, i.e., the definition of a problem in mathematical terms, into outputs, i.e., the solution of the problem in terms of values of some of the problem's variables.
Dynamics Area of classical mechanics concerned with the study of forces and torques and their effect on motion. More in general, the dynamics of a system describes the trajectory of the constituting variables in a proper mathematical space.
Mathematical modelling Description of a natural systems by means of the mathematical language that is then suitable for transformations and other mathematical treatments to find non-trivial relations among system elements.
Networks Graphical description of the relations among entities that is then viable to mathematical treatments such as those of graph theory.
Simulation computer recreation of the dynamics of natural or artificial systems.
Stochasticity The presence of something that is not deterministically determined. In mathematics, it is used to indicate the use of a random element in the description of a system.

Introduction

The recent wide availability of experimental high-throughput omics data (genomics, proteomics, metabolomics, etc.) is fostering the development of theoretical biology. This data is analysed by established bioinformatics tools that have been developed in the past decades as well as by novel approaches devised at a great pace by the community. However, the results of such analysis are in some cases inconclusive and further investigation is necessary to make sense of it. This is when mathematical, statistical and computational approaches such as modelling and simulation come into play. This article briefly describes few but well-representative methods that have been used to this purpose in the last couple of decades.

Quantitative modelling of biological phenomena comprise all methods borrowed by other quantitative or *exact* disciplines such as physics, mathematics, computer science or engineering, that are nowadays used to tame the biological complexity. Given the enormous diversity found in biological systems, it is in general difficult to obtain quantitative estimation of parameters of interest and often the results are dependent on the modelling assumptions. Therefore, the methods range between the two extremes of being completely qualitative to perfectly quantitative. Qualitative methods provide an approximate explanation of what is going on in the studied system without being able to estimate values precisely or give clear indications about how to influence the system (**Saadatpour and Albert, 2016**). On the other hand quantitative methods provide answers to specific questions with numbers, that is, with something that is readily useful in real practice without the need of further (complex) transformations or assumptions (**Mogilner *et al*., 2006**).

The usefulness of computational models in systems biology has been pointed out in many recent literature review articles (Bartocci and Lió, 2016; Fisher and Henzinger, 2007; Ji *et al*., 2017; Machado *et al*., 2011; Materi and Wishart, 2007). This article, however, aims at providing an informal and not exhaustive introduction to the main computational modelling approaches adopted in systems biology. Besides the classical differential equation approach, this paper presents other computational dynamic approaches such as Petri nets, Boolean networks, cellular automata, agent based models, and the statistical methods of Bayesian networks and formal systems such as process algebras.

Background/Fundamentals

Mathematical modelling has a long tradition and is central to most disciplines of science and engineering. It characterizes the scientific effort in describing nature in quantitative terms. Newton's work which was published in the 17th century, is probably the most striking example of application of the concepts of calculus to describe a physical system (i.e., the motion of planets in terms of differential equations). A model is, today as then, a mathematical description of the elements of a process and the ways in which they influence each other. The mathematical description, sometimes amenable of symbolic transformations, leads to a formulation of the "solution" which reveals aspects of the structure and dynamics of the system that are not plainly recognisable at a first sight. Mathematical (now computational alike) modelling is an iterative process which goes from the real-life phenomenon to the virtual mathematical object called the "model" that produces "results" in terms of how it behaves in hypothetical scenarios. To be scientifically sound, these results should be compared to actual data that were produced by the real-life phenomenon and were not used to build or infer the model yet This comparison, which is called model validation, usually identifies shortcomings or problems in the model, advocating some model adjustment, hence the iterative scheme described in **Fig. 1**.

This practice is not different when the system under consideration comprise life. On one side, we have a biological system we wish to understand and on the other side we have a model corresponding to our understanding of the real system that allows us to perform

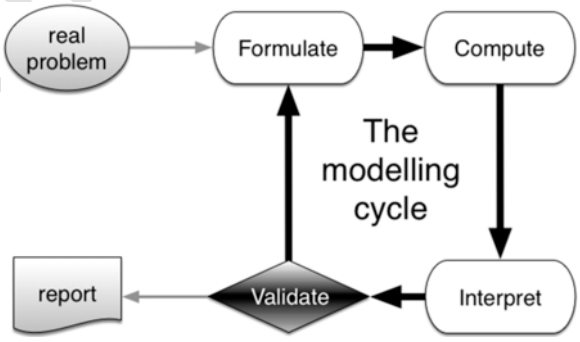

Fig. 1 Mathematical or computational modelling is an iterative process which involves the translation of the most significant characteristics of the real problem into a mathematical formalism amenable of computation (or analytical treatment) to produce results that are then interpreted to be translated back to reality. Those need then to be compared to the real system to be validated so to show the correctness and even usefulness of the model. If this is not the case, the model is improved and the whole cycle is repeated.

analysis and testing of hypothetical scenarios. This is for instance the case of a model describing the pathogenesis of a disease permitting to test different therapeutic interventions thus enabling the search for the most successful treatment options.

Models can be broadly classified according to their structure. A first important distinction to be made relates to whether or not time is one of the variables of the model. A *dynamic* model accounts for time-dependent changes in the state of the system whereas *static* models only consider the system at equilibrium. Static models are often used to describe the invariable relations among object such as, for instance, proteins, metabolites or genes. A subfield of systems biology called network biology deals with such descriptions in terms of graphs or networks. Dynamical models allow for "forecasting" hypothetical scenarios and are quite valuable, for instance, in finding optimal clinical interventions in modelling of disease progress.

Another distinction concerns the involvement of the concept of randomness among the elements of the model. A model is *deterministic* if the current state of the model is uniquely determined by the model parameters and its previous states. That is, given an initial state, a deterministic model always follows the same "trajectory". In contrast, the trajectories of a *stochastic* model may vary depending on the randomness present, in other words, by the variance of the stochastic variables that are therefore not described by unique values, but rather by probability distributions.

A third important classification criterion is the structure of the space from which the variables of the model take on their values. A variable is called *discrete* if it takes on values from a countable set (finite such as the set {0,1} or infinite such as the set of integers), otherwise, it is called *continuous*., that is the case when its values can be arbitrarily close real numbers. If all variables are discrete then the model is said to be discrete. On the other hand, if all variables stand for continuous values then the model is continuous. Finally, hybrid models contain both continuous and discrete components. For instance a model of bacterial growth could represent with $n(t,x)$ the number of bacteria at time t in a certain volume at position x with time and space continuous but state n discrete.

In the following sections, we briefly illustrate different mathematical or computational modelling paradigms classifiable according to the mentioned criteria above. We leave to the reader the leisure of recognising each of them. The paragraphs also report examples of application of the methodologies in the biomedical field with the hope to convey the message of the convenience of quantitative reasoning in these fields. References of relevant works are provided and further references for interested readers are provided at the end of the article.

Differential Equations

Ordinary Differential Equations

To our knowledge one of the earliest work in the literature that used differential equations in biology was the influential and controversial paper by Daniel Bernoulli which was published in 1760. Using a simple model of differential equations, Bernoulli was able to show that the long-term benefits of inoculation overweight the risk of immediate death as it may lead to increasing the life expectancy . For more information about the original Bernoulli's model see Chapter 4 of **Bacaër (2011)**, and, for recent work on Bernoulli's model, see **Dietz and Heesterbeek (2002)**.

For a given biological system of interacting species (substances, chemicals, …) where their concentrations are functions of only one variable (time, age, or cell size,…), an ordinary differential equation (ODE) model of the system is basically a collection of equations, one for each species, describing the rate of change in the concentration of the species, with respect to the variable, in terms of the concentrations of other species in the system, see **Fig. 2**, which presents the well-known Lotka-Volterra or predator-prey model of two interacting quantities x (prey) and y (predator) with the initial concentrations being x_0 and y_0, respectively. This model has four parameters, namely A, B, C, and D. For more information about predator-prey models, see Chapter 13 of **Bacaër (2011)**. A similar but more specific example of ODE system applied to clinic is those in Dell'Acqua and Castiglione (2009) and Jarrah *et al*. (2014) about modelling muscular dystrophy.

Values of the parameters of a model are determined based on the available data and using any of several parameter estimation methods such as Genetic Algorithm, Particle Swarm, Simulated Annealing, Steepest Descent, to name few (**Moles *et al*., 2003**) and multiple shooting method for stochastic systems (**Peifer and Timmer, 2007**). For biochemical networks, the parameters could be assumed to be of one of two forms depending on the underlying kinetics: mass action or Hill function which includes Michaelis-Menten as a special case. While Mass Action (polynomial form) is appropriate when the reaction rate among species is proportional to the probability of the collision of the particles of the species, the Hill function in general is used in enzyme kinetics as well as protein's activation or inhibition.

Regardless of the optimization method, parameter estimation is probably the hardest step in the ODE approach for modelling biological systems (**Chou and Voit, 2009**) mainly due to the large number of species being modelled, their nonlinear interactions, and the fact that available quantitative data are limited. Yet the ODE approach is probably the most popular among all, mainly due to its solid and old mathematical foundation as well as its huge success in modelling a variety of biological systems at all levels; such as the famous population dynamic model predator-prey, Hodgkin–Huxley model of excitable neuron (**Hodgkin and Huxley, 1952**), and gene regulatory network of yeast cell cycle (**Tyson and Novák, 2015**) to name a few.

There are many different variations of the ODE approach, in particular, stochastic ODE (**Bachar *et al*., 2013**), delay ODE (**Parmar *et al*., 2015**). Also the S-systems (**Savageau, 1988**) is a special class of ODE models that assumes a specific structure of the system.

There are several general ODE solvers as well as specialized standalone software that can be used for parameter estimations and solving ODE. The standalone software CellDesigner (**Matsuoka *et al*., 2014**) and COPASI (**Hoops *et al*., 2006**) can be used for the modelling and simulation of biochemical networks and both have most of the mentioned methods and features above.

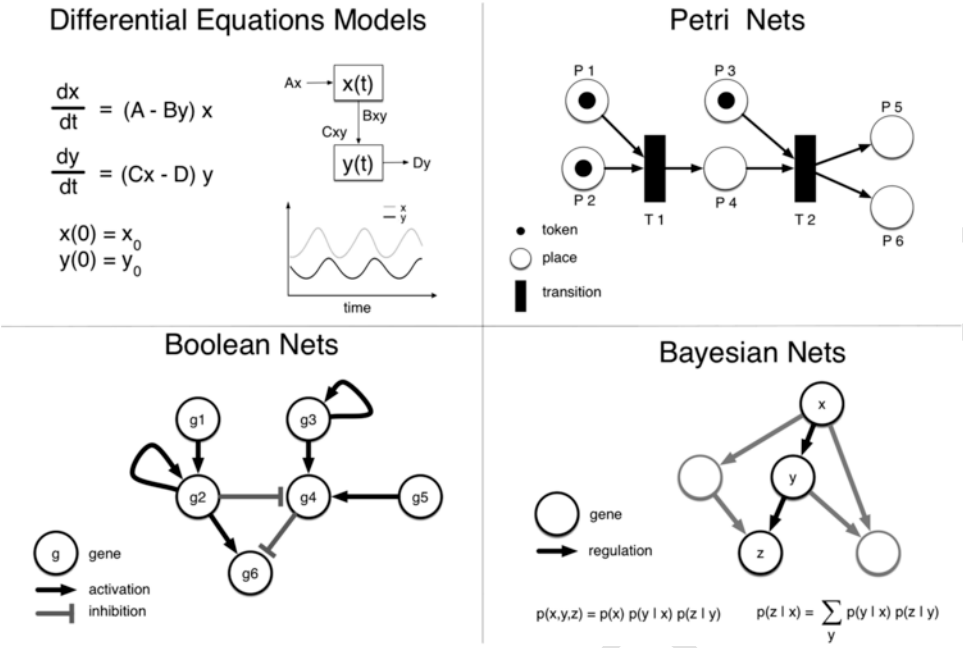

Fig. 2 Differential equations models exemplified in a predator-prey relation. The Petri-nets modelling paradigm is very generic and apt to model concurrency not just in biological systems. It consists in tokens moving among places through transitions epitomising conditions. Boolean networks have been conceived to describe gene regulation but are general enough to be applied to other field of science well. Bayesian networks are statistical methods which allow to estimate the strength (probabilistically speaking) of a relation among two objects such as genes, give a set of data.

One of the main limitations of the ODE approach is the fact that it assumes that the species are well-mixed, and hence it cannot be used to properly model a system that is space-depended. This can be easily resolved when using partial differential equations.

Partial Differential Equations

This approach enables modelling interacting species as functions of more than one variable. In particular, a partial differential equation (PDE) represents the change in the concentration of a substance with respect to not only time, as most ODE models, but also to other variables such as place (spatial coordinates), age and size of a cell.

Alan Turing developed a simple PDE model of two diffusible and interacting chemicals, and hence the name reaction-diffusion model, and he used the model to explain the pattern formation in the developing animal embryo (**Turing, 1952**). Since then, the reaction-diffusion model has been the basis of most PDE methods for the PDE modelling and simulation of complex biological patterns (**Kondo and Miura, 2010**). For a comprehensive treatment of the subject, see **Murray (2003)**.

Modelling methods using the PDE approach have been developed to study various kinds of biological systems at all levels; such as ion singling (Ramay et al, 2010)(, chemotaxis-driven migration of T-cells toward infection sites (**Vroomans *et al*., 2012**), and climate change (**Goosse *et al*., 2010**) as well as weather prediction (**Steppeler *et al*., 2003**).

Similar to the ODE approach, solving PDE equations is not easy and is usually done using numerical techniques (**Tadmor, 2012**). Furthermore, methods for the parameter estimations of PDE models have recently been developed using different approaches such as least squares method (**Muller and Timmer, 2002**) and Bayesian methods (**Xun *et al*., 2013**).

Difference Equations

In *Liber Abaci*, which was published in 1202, Fibonacci investigated how fast rabbits could breed. Assuming ideal environment, he developed a model for the number of rabbits using a difference equation and produced the famous Fibonacci's sequence: 1, 1, 2, 3, 5, 8, 13, 21, …. It is worth mentioning that Fibonacci's sequence has appeared in other areas within as well as outside biology. Fast forward to the 20th century, difference equations models of various biological systems are ubiquitous, in particular, the difference equations approach provides a discrete version of population biology in general (**Cull *et al*., 2005**). Also, finite difference equations are used to develop numerical methods for solving ODE and PDE systems (**Sewell, 2006**). Furthermore, difference equations approach provide a generalization as well as a mathematical framework for cellular automata (**Itoh and Chua, 2009**) and other finite discrete methods. See the classical textbook (**Elaydi, 2005**) for more information on the theory of difference equations, and **Allman and Rhodes (2004)** and

Cull *et al.* (2005) for examples of how the difference equations approach is used to model several biological systems such logistic growth and phylogenetic trees. Difference equations are best for studying discrete-time, population dynamics with stochastic features (**Witten and de la Torre, 1984**).

Petri Nets

Petri nets (PN) are directed bipartite graphs with nodes belonging to one of the two sets called states (also called places) and transitions (see **Fig. 2**). They were conceived by Carl Adam Petri in 1962 to describe chemical reactions (**Petri and Reisig, 2008**). Indeed, the states indicate substances and the transitions the reactions. Nodes identifying states are filled with tokens to indicate that the relative substance is present (alternatively, that the substance is present in sufficient concentration to allow the reaction to take place). When a transition has more than one incoming arcs (called *take* arcs) linked to corresponding states, those states all have to contain sufficient tokens to allow the transition to fire, meaning that the reaction takes place and the token is moved ahead in the states linked downward the exit arcs (called *give* arcs) (see **Fig. 2**). In this manner, the states code for conditions that need to be satisfied for a transition (or more generically an action) to take place. The flexibility of this modelling formalism allow its use to describe concurrent process in distributed systems. It has indeed been extensively used in computer science to describe distributed processing. In systems biology, Petri nets and its variants, developed to remove constraints such as the use of a discrete number of tokens (i.e., the coloured PN), or the use of deterministic transitions (i.e., the stochastic PN), have been employed to model biochemical networks (Chaouiya, 2007; Goss and Peccoud, 1998; Srivastava *et al*., 2001) such as metabolic processes (**Simao *et al*., 2005**) and cell signalling (**Li *et al*., 2007**). The description of a biological process such as those at the molecular level just mentioned, is quite straightforward with PN. Moreover, given the intrinsic account for the temporal scale, such formalism is quite appealing for biologist that can analyse, mainly in a qualitative fashion but in principle also in a quantitative manner, complex biological phenomena by answering what-if questions such as "what happen if the production of a molecular compound is inhibited?"

Boolean Networks

Boolean networks (BN) have been introduced by Kauffman in 1969 to model gene regulation (Kauffman, 1969, 1993). They represent genes as nodes of a graph with two types of arcs, activation or inhibition, denoting respectively an activation of, say, gene $g2$ subsequent to the activation of gene $g1$, and inhibition, indicating the capacity of an express gene to block the transcription of another gene (see **Fig. 2**). The model describes the temporal evolution, carried out in discrete steps, of the active/inactive states of the set of genes. At each time step the state of each gene is determined by a Boolean logical rule which is a function of the state of the genes which impinge upon it (i.e., its regulators). For instance, for the network in **Fig. 2**, the Boolean formula for node g4 could be $g4(t+1)=g3(t)\ AND\ g5(t)\ AND\ NOT\ g2(t)$.

The state of the genes can be updated all at once (synchronous update) or one at a time with a specified or unspecified order (asynchronous update). The two evolutions are not equivalent with the latter being much more complex to compute than the former. The space of the possible trajectories of the set of gene-expression grows exponentially with the number of genes thus its exploration to determine the dynamical properties of the BN becomes quickly unfeasible. On the other hand, the analysis of network properties such as its robustness or the steady-states of its dynamics which identify specific biological functions , are exactly the reward of using such models (**Li *et al*., 2004**).

Boolean networks are generally inferred from time-series of gene expression data (**D'haeseleer *et al*., 2000**) or are constructed by expert manual curation from literature data (**Pedicini *et al*., 2010**). Boolean networks have also been employed to model signalling pathways (**Saez-Rodriguez *et al*., 2007**). For these biological applications the limitation of the two-values logic has been removed in the multi-valued extension (**Schaub *et al*., 2007**). Moreover, to account for the presence of noise and uncertainty, the stochastic extensions of probabilistic Boolean networks were introduced (Shmulevich *et al*., 2002).

Bayesian Networks

A Bayesian network is a statistical model describing a set of variables and their conditional dependencies via a directed acyclic graph (see **Fig. 2**). The nodes of the network represent random variables whereas the edges stand for dependencies between couple of variables. The value of each variable is determined by a probabilistic function of the variables associated to the nodes than impinges upon it.

Bayesian networks were introduced by the work of Pearl (**Pearl, 1988**) as generic probabilistic methods to infer relationships among variables given a set of data. In fact, based on the formula from Bayes, there exist learning methods to infer the probability parameters also in case of incomplete data. Data coming from gene regulation activity provide a natural example of applicability of Bayesian networks. In this case the nodes-variables are the expression levels of the genes and the probabilistic relationships identified by the edges are estimated by means of the inference algorithms (**Friedman, 2004**). Data from signalling networks is amenable to this representation and analysis (**Sachs *et al*., 2005**). The disadvantage due to the inability to model feedback loops present in biological networks has been removed in an extension called dynamic Bayesian networks (Husmeier, 2003; Kim *et al*., 2003; Zou and Conzen, 2005).

Microscopic Ab-initio Models

Models in this category are usually discrete (both in time and space) dynamic models that aim to simulate/address biological problems for which spatial properties strongly affects the resulting dynamics. Discrete models are better suited to capture the discrete nature of individual cells than continuous models, especially when dealing with complex systems involving many variables such as sets of cells in different stages of their life cycle or distinguishable by the affinity of their receptors for different antigens. Cell-based models such as cellular automata, cellular Potts and agent-based models are able to represent the behaviour and interactions of each individual cell in a biological system. In these models, the macroscopic dynamics of a system is an emergent property of the microscopic interactions and for this reason they are considered bottom-up modelling approaches.

The Ising model (1925) is the simplest theoretical model of ferromagnetism used to study ferromagnetic phase transitions and is considered by many the precursor of cellular automata and agent based models. In the Ising model each magnetic dipole moment of atomic spins is represented as a discrete variable characterized by one of two states (+1 or −1) and arranged on a lattice. The interactions between spins are local and each spin is allowed to change synchronously based on the spin state of its neighbours.

Cellular Automata

Cellular automata (1940s) are a general computational model characterized by discrete space, state, and time (Ilachinski, 2001; Toffoli and Margolus, 1987; Wolfram, 2002) inspired in the first half of the 20th century by the work of John von Neumann (**von Neumann, 1966**) and Stanislaw Ulam (**Ulam, 1952**). A cellular automata model is a discrete dynamic system defined by a grid (a finite metric space), a finite set of possible cells states, a function describing the neighbourhood of a cell (the set of cells influencing the one under consideration) and a transition function determining the state of a cell in the next time step depending on the state of the cells in its neighbourhood (see **Fig. 3**).

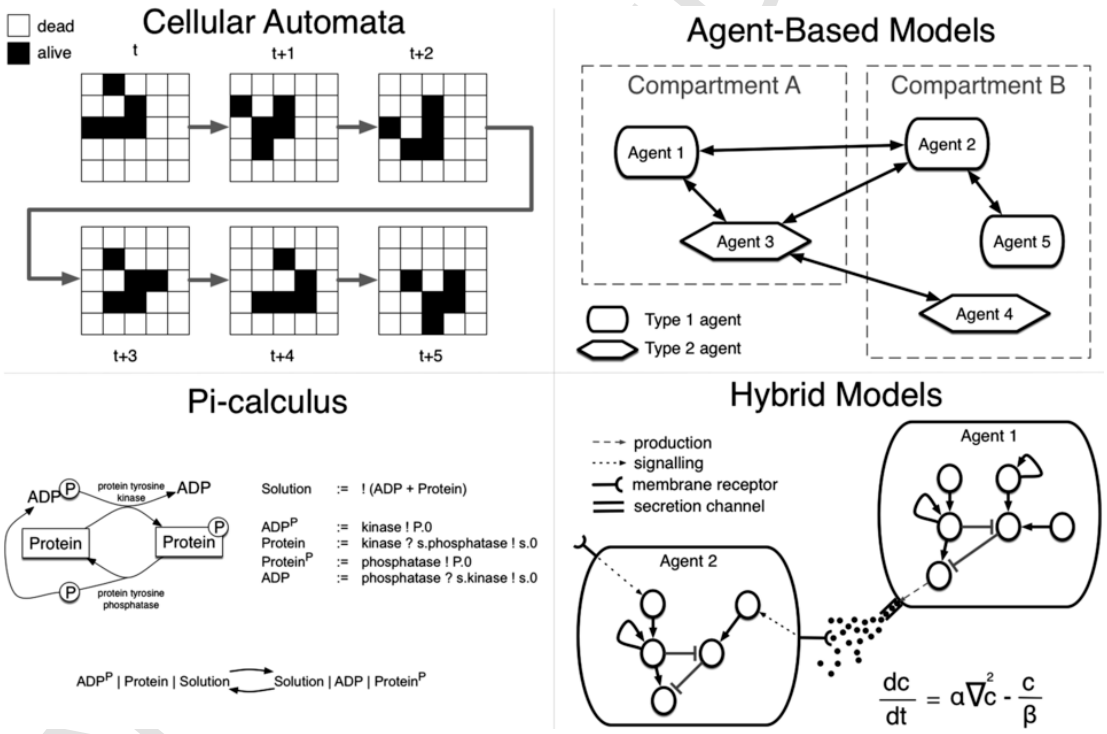

Fig. 3 Cellular automata and agent-based models represent entities individually. These evolve according to defined rules whose complexity range from rather simple (often Boolean) functions in cellular automata to sophisticated algorithms for the agents' behaviour in agent-based models. Furthermore, as an example, one can construct complex hybrid models embedding Boolean networks in agents representing biological cells to set up the rule for its differentiation. Pi-calculus is a different approach used to formally describe a complex system such as a chemical reactor and to conduct (in the case of stochastic calculi, e.g., stochastic k-calculus) simulations much like in the previously described methodologies.

Classical cellular automata are synchronous (all cells update their state simultaneously thus the grid state is independent from the order of update of the grid), local, and homogeneous. Many modifications of the classical model exist: asynchronous automata, heterogeneous automata, non-uniform cellular automata, stochastic automata and many others such as movable cellular automata (**Psakhie *et al*., 1995**), Brownian cellular automata (**Lee and Peper, 2010**).

Cellular automata have been extensively used as a modelling paradigm in biology (**Ermentrout and Edelstein-Keshet, 1993**). Some applications include the study of idiotypic B cell proliferation in the human immune system (**De Boer *et al*., 1992**) or the dynamics of infectious diseases (**Zorzenon dos Santos and Coutinho, 2001**).

Cellular Potts models, also known as Glazier Graner-Hogeweg models (Graner and Glazier, 1992; Glazier and Graner, 1993), are an evolution of cellular automata in which each cell occupies more than one lattice site or pixel. For this reason, each cell has a shape which varies over time as a function of a Hamiltonian representing the adhesion energy of the cell and other mechanical constraints such as surface-area-to-volume ratio. One of the most known framework for cellular Potts model is CompuCell3D (Chaturvedi *et al*., 2005; Popławski *et al*., 2008) which has been used in several different applications, from anatomical and pathological conditions of tissues and organs to models of single-species and multi-species biofilm growth. In this hybrid framework, the mechanical properties of the cell membrane can be integrated with continuous models of reaction-diffusion dynamics and models of intracellular signaling dynamics achieving a complete and realistic multi-scale model of a complex biological system.

Agent-Based Models

Agent-based models (ABM) are an evolution of the simpler cellular automata model in which each agent is an autonomous decision-making entity interacting on the basis of a given set of rules which are not necessarily identical (Bonabeau, 2002; Gilbert, 2008). Strictly speaking, ABMs are not discrete in time and space as they are exemplary of hybrid systems where discrete time variables are combined with continuous state-space ones. Nevertheless, in their applications to biology and medicine and with the purpose to speed up execution on a computer, agents are usually placed on a structure (a two dimensional or three-dimensional lattice representing real-world spatial environment such as a lymph node or a complex network representing social contacts or sexual interactions) that defines the couples or groups of agents that are going to interact at a given time. Agent-based models have been used to study many different problems. For instance, they have been used to simulate HIV or gonorrhoea epidemics where agents represent individuals part of a sexual contact network (**Mei *et al*., 2011**) or to describe the population dynamics of ecosystems in which each agent is an animal or a plant. They have also being used to model the intracellular dynamics phenomenain which agents are viral particles moving within the cytoplasm of an infected cell. Finally, they have been extensively used to describe the dynamics of the immune system. In this case the agents represent all the entities involved in the immune response to a pathogen such as immune cells, epithelial cells, bacteria, viral particles etc (**Castiglione and Celada, 2015**).

Several frameworks can be used to implement agent based models: NETLOGO, FLAME, MASON, JASON and REPAST are some of the most versatile and supported frameworks.

Process Calculi

The design of a computational implementation of any biological system starts from the knowledge of the living system under consideration and by abstracting its main characteristics. This task is fulfilled by means of some formal method such as those describing the syntax of a computer language. The most striking feature of biological phenomena that is difficult to capture in symbolic formal methods is concurrency. Formal systems, therefore, need to be able to fully denote such biological complexity or otherwise turn out inaccurate and eventually useless.

Process calculi or process Algebras (PA) are formal systems devised to easily express parallel computations and their properties (like Church's lambda-calculus introduced to express Turing computable arithmetic functions). The most successful system was Milner's pi-calculus introduced to express interactions between concurrent processes (**Milner, 1989**). PAs share some similarity with formalism used in the description of chemical reactions (see **Fig. 3** for the example of the phosphorylation/dephosphorylation reaction expressed in pi-calculus). This was precisely established with the introduction of the Chemical Abstract Machine or CHAM (**Berry and Boudol, 1992**), where the states of the parallel machine are chemical solutions and molecules can interact according to reaction rules.

Biological phenomena involve millions of moving particles in any single cell while millions of cells form tissues which transport substances, transform energy use signals, etc. Notably, biological systems process information necessary to the organization of life by the same mechanisms they use to processes any other element, like the nutrients. Process algebras are considered suitable formalisms to model these kinds of systems because the transmitted information is processed *concurrently* and *asynchronously*. Moreover, part of essential transactions which happen in biological systems are inherently computational and therefore they can be expressed with formal languages used to describe concurrent processes.

While representing biological systems with PAs, several researchers realised that there was a missing dimension: quantitative aspects were neglected in the description of the interaction rules. The stochastic pi-calculus (**Priami, 1995**) was an enriched Milner's formalism with a quantitative context semantics giving access to the world of simulations of biochemical processes (**Regev *et al*., 2001**; **Priami *et al*., 2001**). Further developments were obtained by deriving new formal systems by adding missing features to PAs: bio-ambients (**Cardelli and Gordon, 2000**) and brane-calculi (**Cardelli, 2005**) allowing compartments, the kappa-calculus (**Danos and Laneve, 2004**) with the specific purpose of representing protein interactions.

The prominent aspect to remark regarding the convenience for using this approach in describing biological processes is that it enables, for instance, the syntactic and semantic comparison (hence prompt translation of the properties of the first to the second system) between two concurrent systems once formalised with a PA or to build incrementally complex process from a multitude of elementary processes being either independent or interrelated.

Closing Remarks

Computational methods are used to make sense of the vast amount of data produced by experimental biology high throughput technologies and, importantly, to "fill the dots" where data is incomplete in revealing the underlying structure or in describing the temporal evolution of the system under study. Mathematical or computational modelling methods are applied and even developed anew to analyse such data, to dig information out of it, to discover structures, to generalize conclusions, to provide scenarios to choose from, etc. The accuracy of these computational approaches ranges from low (that is when we call the approach qualitative) to high (i.e., that is quantitative) depending on the data availability.

The methodologies range from classical differential equations to more unconventional agent-based simulation, passing through process algebra and statistical approaches. These models have shown to be promising if not even concretely useful in some cases, revealing biological aspects of the phenomena under consideration which were impossible to see just by inspecting the data. Vice versa, biology is providing a number of challenging problems to appraise the methods themselves, encouraging computational scientists to extend and improve existing methodologies (**Machado *et al*., 2011**).

One of the greatest challenge in this respect is the construction of multi-scale models embracing the wide range of time and length scales of biological phenomena. Whereas few modelling techniques though theoretical able to handle different scales (e.g., process algebras or Petri nets), their practical use is impaired by the combinatorial explosion of variants eventually leading to computational intractable problems. In this case the obvious trick is to combine two or (rarely) more approaches that have their strength in different time/length scale into a unified hybrid multi-scale system (Fisher and Henzinger, 2007; Bortolussi and Policriti, 2008; Fromentin *et al*., 2010). This is for instance the case of agent-based models of cellular systems where each agent/cell is equipped with a system of ordinary differential equations or a Boolean network to describe the regulatory network among genes driving its behaviour (see **Fig. 3**).

Computational biology is a flourishing field of applied mathematics and computer science. Only in the past two decades the collective effort of computational scientists combined with the expertise of field biologist have produced a number of open-software that offer the model techniques described in this brief article to analyse experimental data and/or to simulate hypothetical scenarios. Review papers such as **Bartocci and Lió (2016)** offer a nice view of what is available at the time of that publication. There is no reason to believe that such vigorous effort in the systems biology community will grow smaller, but rather just the opposite, new tools will be available soon, either as improvements of existing methodologies or able to exploit and combine different formalisms and modelling paradigms to provide powerful platforms for the analysis of the ever increasing experimental and diagnostic data.

Acknowledgement

Abdul Jarrah was supported by the National Science Foundation under Grant No. DMS-1440140 in the Fall 2018 while the author was in residence at the Mathematical Sciences Research Institute in Berkeley, California, USA. Filippo Castiglione acknowledge partial support from the Institute for Advanced Studies of the University of Amsterdam. Emiliano Mancini acknowledges partial support from the SimCity project of the Dutch NWO, eScience agency under contract C.2324.0293.

Uncited references

(Needham *et al*., 2007; Ramay *et al*., 2010).

References

Allman, E., Rhodes, J., 2004. Mathematical Models in Biology: An Introduction. Cambridge University Press, New York.
Bacaër, N., 2011. A Short History of Mathematical Population Dynamics. Springer-Verlag, London.
Bachar, M., Batzel, J., Ditlevsen, S., 2013. Stochastic Biomathematical Models: With Applications to Neuronal Modeling. Springer, Heidelberg.
Bartocci, E., Lió, P., 2016. Computational modeling, formal analysis, and tools for systems biology. PLOS Computational Biology 12 (1), e1004591. https://doi.org/10.1371/journal.pcbi.1004591.
Berry, G., Boudol, G., 1992. The chemical abstract machine. Theoretical Computer Science 96, 217–248.
Bonabeau, E., 2002. Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences of the United States of America 99, 7280–7287.
Bortolussi, L., Policriti, A., 2008. Hybrid systems and biology. In: Bernardo, Marco, Degano, Pierpaolo, Zavattaro, Gianluigi (Eds.), Formal Methods for Computational Systems Biology. vol. 5016 of LNCS. Springer, pp. 424–448.
Cardelli, L., Gordon, A.D., 2000. Mobile ambients. Theoretical Computer Science 240 (1), 177–213.

Cardelli, L., 2005. Brane Calculi. In: In: Danos, V., Schachter, V. (Eds.), Computational Methods in Systems Biology (CMSB 2004). Lecute Notes in Computer Science, 3082. Springer, Berlin, Heidelberg, pp. 257–278.
Castiglione, F., Celada, F., 2015. Immune System Modeling and Simulation. CRC Press, Boca Raton, FL, USA.
Chaouiya, C., 2007. Petri net modelling of biological networks. Briefings in Bioinformatics 8, 210–219.
Chaturvedi, R., Huang, C., Kazmierczak, B., et al., 2005. On multiscale approaches to three-dimensional modeling of morphogenesis. Journal of the Royal Society Interface 2, 237–253.
Chou, C., Voit, E.O., 2009. Recent developments in parameter estimation and structure identification of biochemical and genomic systems. Mathematical Biosciences 219, 57–83.
Cull, P., Flahive, M., Robson, R., 2005. Difference Equations: From Rabbits to Chaos. Springer, New York.
Danos, V., Laneve, C., 2004. Formal molecular biology. Theoretical Computer Science 325 (1), 69–110.
De Boer, R.J., Segel, L.A., Perelson, A.S., 1992. Pattern formation in one- and two-dimensional shape-space models of the immune system. Journal of Theoretical Biology 155 (3), 295–333.
Dell'Acqua, G., Castiglione, F., 2009. Stability and phase transitions in a mathematical model of Duchenne muscular dystrophy. Journal of Theoretical Biology 260 (2), 283–289. https://doi.org/10.1016/j.jtbi.2009.05.037.
Dietz, K., Heesterbeek, J.A.P., 2002. Daniel Bernoulli's epidemiological model revisited. Mathematical Biosciences 180, 1–21.
D'haeseleer, P., Liang, S., Somogyi, R., 2000. Genetic network inference: From co- expression clustering to reverse engineering. Bioinformatics 16 (8), 707–726.
Elaydi, S., 2005. An Introduction to Difference Equations. Springer, New York.
Ermentrout, G.B., Edelstein-Keshet, L., 1993. Cellular automata approaches to biological modeling. Journal of Theoretical Biology 160 (1), 97–133.
Fisher, J., Henzinger, T.A., 2007. Executable cell biology. Nature Biotechnology 25 (11), 1239–1249. https://doi.org/10.1038/nbt1356.
Friedman, N., 2004. Inferring cellular networks using probabilistic graphical models. Science 303 (5659), 799–805.
Fromentin, J., Eveillard, D., Roux, O., 2010. Hybrid modeling of biological networks: Mixing temporal and qualitative biological properties. BMC Systems Biology 4, 79. https://doi.org/10.1186/1752-0509-4-79 (PMID: 20525331).
Gilbert, N., 2008. Agent-Based Models. Sage Publications, Los Angeles.
Glazier, J.A., Graner, F., 1993. Simulation of the differential adhesion driven rearrangement of biological cells. Physical Review E 47, 2128–2154.
Goosse, H., Barriat, P.Y., Lefebvre, W., Loutre, M.F., Zunz, V., 2010. Chapter 3: Modelling the Climate System of Introduction to Climate Dynamics and Climate Modeling. Online textbook available at http://www.climate.be/textbook.
Goss, P.J., Peccoud, J., 1998. Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets. Proceedings of the National Academy of Sciences of the United States of America 95, 6750–6755.
Graner, F., Glazier, J.A., 1992. Simulation of biological cell sorting using a two-dimensional extended Potts model. Physical Review Letters 69 (13), 785–790.
Hodgkin, A.L., Huxley, A.F., 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. The Journal of Physiology 117, 500–544.
Hoops, S., Sahle, S., Gauges, R., et al., 2006. COPASI: A complex pathway sImulator. Bioinformatics 22, 3067–3074.
Husmeier, D., 2003. Sensitivity and specificity of inferring genetic regulatory interactions from microarray experiments with dynamic Bayesian networks. Bioinformatics 19 (17), 2271–2282.
Ilachinski, A., 2001. Cellular Automata: A Discrete Universe. World Scientific Publishing Company, River Edge, NJ, USA.
Itoh, M., Chua, L., 2009. Difference equations for cellular automata. International Journal of Bifurcation and Chaos 19, 805–830.
Jarrah, A.S., Castiglione, F., Evans, N.P., Grange, R.W., Laubenbacher, R., 2014. A mathematical model of skeletal muscle disease and immune response in the mdx mouse. BioMed Research International 871810. https://doi.org/10.1155/2014/871810.
Ji, Z., Yan, K., Li, W., Hu, H., Zhu, X., 2017. Mathematical and computational modeling in complex biological systems. BioMed Research International 5958321. https://doi.org/10.1155/2017/5958321.
Kauffman, S., 1969. Metabolic stability and epigenesis in randomly constructed genetic nets. Journal of Theoretical Biology 22, 437–467.
Kauffman, S.A., 1993. The Origins of Order: Self-organization and Selection in Evolution. Oxford University Press, New York.
Kim, S., Imoto, S., Miyano, S., 2003. Inferring gene networks from time series microarray data using dynamic Bayesian networks. Briefings in Bioinformatics 4 (3), 228–235.
Kondo, S., Miura, T., 2010. Reaction-diffusion model as a framework for understanding biological pattern formation. Science 329, 1616–1620.
Lee, J., Peper, F., 2010. Efficient computation in Brownian cellular automata. In: In: Peper, F., Umeo, H., Matsui, N., Isokawa, T. (Eds.), Natural Computing. Proceedings in Information and Communications Technology, 2. Springer, Tokyo.
Li, C., Ge, Q.W., Nakata, M., Matsuno, H., Miyano, S., 2007. Modelling and simulation of signal transductions in an apoptosis pathway by using timed Petri nets. Journal of Biosciences 32, 113–127.
Li, F., Long, T., Lu, Y., Ouyang, Q., Tang, C., 2004. The yeast cell-cycle network is robustly designed. Proceedings of the National Academy of Sciences of the United States of America 101 (14), 4781–4786.
Machado, D., Costa, R.S., Rocha, M., et al., 2011. Modeling formalisms in systems biology. AMB Express 1, 45. https://doi.org/10.1186/2191-0855-1-45.
Materi, W., Wishart, D.S., 2007. Computational systems biology in drug discovery and development: Methods and applications. Drug Discovery Today 12 (7–8), 295–303. https://doi.org/10.1016/j.drudis.2007.02.013.
Matsuoka, Y., Funahashi, A., Ghosh, S., Kitano, H., 2014. Modeling and simulation using CellDesigner. In: In: Miyamoto-Sato, E., Ohashi, H., Sasaki, H., Nishikawa, J., Yanagawa, H. (Eds.), Transcription Factor Regulatory Networks. Methods in Molecular Biology (Methods and Protocols). vol. 1164. Humana Press, New York.
Mei, S., Quax, R., van de Vijver, D., Zhu, Y., Sloot, P.M.A., 2011. Increasing risk behaviour can outweigh the benefits of antiretroviral drug treatment on the HIV incidence among men-having-sex-with-men in Amsterdam. BMC Infectious Diseases 11, 118.
Milner, R., 1989. Communication and Concurrency. Prentice Hall, Inc., Upper Saddle River, NJ, USA (ISBN:0-13-115007-3).
Mogilner, A., Wollman, R., Marshall, W.F., 2006. Quantitative modeling in cell biology: What is it good for?. Developmental Cell 11, 279–287.
Moles, C.G., Mendes, P., Banga, J.R., 2003. Parameter estimation in biochemical pathways: A comparison of global optimization methods. Genome Research 13, 2467–2474.
Muller, T., Timmer, J., 2002. Fitting parameters in partial differential equations from partially observed noisy data. Physical Review 171, 1–7.
Murray, J., 2003. Mathematical Biology. Springer-Verlag, Berlin.
.
Parmar, K., Blyuss, K.B., Kyrychko, Y.N., Hogan, S.J., 2015. Time-delayed models of gene regulatory networks. Computational and Mathematical Methods in Medicine 2015, 16. https://doi.org/10.1155/2015/347273 (Article ID 347273).
Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers Inc, San Francisco.
Pedicini, M., Barrenas, F., Clancy, et al., 2010. Combining network modeling and gene expression microarray analysis to explore the dynamics of Th1 and Th2 cell regulation. PLOS Computational Biology 6 (12), e1001032. https://doi.org/10.1371/journal.pcbi.1001032.

Peifer, M., Timmer, J., 2007. Deterministic inference for stochastic systems using multiple shooting and a linear noise approximation for the transition probabilities. IET System Biology 1, 78–88.
Petri, C.A., Reisig, W., 2008. Petri net. Scholarpedia 3 (4), 6477. Available at: http://www.scholarpedia.org/article/Petri_net.
Popławski, N.J., Shirinifard, A., Swat, M., Glazier, J.A., 2008. Simulation of single-species bacterial-biofilm growth using the Glazier-Graner-Hogeweg model and the CompuCell3D modeling environment. Mathematical Biosciences and Engineering 5 (2), 355.
Priami, C., 1995. Stochastic pi-calculus. Computer Journal 38 (7), 578–589.
Priami, C., Regev, A., Shapiro, E., Silverman, W., 2001. Application of a stochastic name-passing calculus to representation and simulation of molecular processes. Information Processing Letters 80 (1), 25–31.
Psakhie, S.G., Horie, Y., Korostelev, S.Y., et al., 1995. Method of movable cellular automata as a tool for simulation within the framework of mesomechanics. Russian Physics Journal 38 (11), 1157–1168.
Ramay, H., Jafri, M.S., Lederer, W.J., Sobie, E.A., 2010. Predicting local SR Ca^{2+} dynamics during Ca^{2+} wave propagation in ventricular myocytes. Biophysical Journal 98, 2515–2523.
Regev, A., Silverman, W., Shapiro, E., 2001. Representation and simulation of biochemical processes using the π-calculus process algebra. In: In: Altman, R.B., Dunker, A.K., Hunter, L., Klein, T.E. (Eds.), Pacific Symposium on Biocomputing, 6. World Scientific Press, Singapore, pp. 459–470.
Saadatpour, A., Albert, R., 2016. A comparative study of qualitative and quantitative dynamic models of biological regulatory networks. EPJ Nonlinear Biomedical Physics 4, 5. https://doi.org/10.1140/epjnbp/s40366-016-0031-y.
Saez-Rodriguez, J., Simeoni, L., Lindquist, J., et al., 2007. A logical model provides insights into T cell receptor signaling. PLOS Computational Biology 3 (8), e163.
Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D., Nolan, G., 2005. Causal protein-signaling networks derived from multi-parameter single-cell data. Science 308 (5721), 523–529.
Savageau, M.A., 1988. Introduction to S-systems and the underlying power-law formalism. Mathematical and Computer Modelling 11, 546–551.
Schaub, M.A., Henzinger, T.A., Fisher, J., 2007. Qualitative networks: A symbolic approach to analyze biological signaling networks. BMC Systems Biology 1, 4.
Sewell, G., 2006. The Numerical Solution of Ordinary and Partial Differential Equations. John Wiley & Sons, Chicester, United Kingdom.
Simao, E., Remy, E., Thieffry, D., Chaouiya, C., 2005. Qualitative modelling of regulated metabolic pathways: Application to the tryptophan biosynthesis in E. coli. Bioinformatics 21 (2), 190–196.
Srivastava, R., Peterson, M.S., Bentley, W.E., 2001. Stochastic kinetic analysis of the Escherichia coli stress circuit using sigma(32)-targeted antisense. Biotechnology and Bioengineering 75, 120–129.
Steppeler, J., Hess, R., Schattler, U., Bonaventura, L., 2003. Review of numerical methods for nonhydrostatic weather prediction models. Meteorology and Atmospheric Physics 82 (1–4), 287–301.
Tadmor, E., 2012. A review of numerical methods for nonlinear partial differential equations. Bulletin of the American Mathematical Society 49, 507–554.
Toffoli, T., Margolus, N., 1987. Cellular Automata Machines: A New Environment for Modeling. MIT Press, Cambridge, MA, USA.
Turing, A., 1952. The chemical basis of morphogenesis. Philosophical Transactions of the Royal Society of London Series B Biological Sciences 237, 37–72.
Tyson, J., Novák, B., 2015. Models in biology: Lessons from modeling regulation of the eukaryotic cell cycle. BMC Biology 13, 46https://doi.org/10.1186/s12915-015-0158-9.
Ulam, S., 1952. Random processes and transformations. In: Proceedings of the International Congress of Mathematicians, pp. 264–275. Rhode Island: American Mathematical Society.
von Neumann, J., 1966. The Theory of Self-reproducing Automata. University of Illinois Press, Urbana, IL.
Vroomans, R.M.A., Marée, A.F.M., de Boer, R.J., Beltman, J.B., 2012. Chemotactic migration of T cells toward dendritic cells promotes the detection of rare antigens. PLOS Computational Biology 8, e1002763.
Witten, M., de la Torre, D., 1984. Biological populations obeying difference equations: The effects of stochastic perturbation. Journal of Theoretical Biology 111, 493–507.
Wolfram, S., 2002. A New Kind of Science. Wolfram Media, Inc, Champaign, IL.
Xun, X., Cao, J., Mallick, B., Carroll, R.J., Maity, A., 2013. Parameter estimation of partial differential equation models. Journal of the American Statistical Association 108 (503)https://doi.org/10.1080/01621459.2013.794730.
Zorzenon dos Santos, R.M., Coutinho, S., 2001. Dynamics of HIV infection: A cellular automata approach. Physical Review Letters 87, 168102.
Zou, M., Conzen, S., 2005. A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. Bioinformatics 21 (1), 71–79.

Further Reading

Castiglione, F., Celada, F., 2015. Immune System Modeling and Simulation. CRC Press, Boca Raton.
Ciocchetta, F., Hillston, J., 2009. Bio-PEPA: A framework for the modelling and analysis of biological systems. Theoretical Computer Science 410 (33–34), 3065–3084.
Deutsch, A., Dormann, S., 2005. Cellular Automaton Modeling of Biological Pattern Formation: Characterization, Applications, and Analysis. Birkhäuser, Boston.
Elaydi, S., 2005. An Introduction to Difference Equations. Springer, New York.
Ellner, S.P., Guckenheimer, J., 2006. Dynamic Models in Biology. Princeton University Press, New Jersey.
Gilbert, N., 2008. Agent-Based Models. Sage Publications, Los Angeles.
Ilachinski, A., 2001. Cellular Automata: A Discrete Universe. World Scientific Publishing Company, River Edge, NJ, USA.
Kauffman, S.A., 1995. At Home in the Universe: The Search for Laws of Self-organization and Complexity. Oxford University Press, Oxford.
Laneve, C., Tarissan, F., 2007. Simple calculus for proteins and cells. Electronic Notes in Theoretical Computer Science 171, 139–154.
Milner, R., 1999. Communicating and Mobile Systems: The π-Calculus. Cambridge University Press, Cambridge.
Murray, J., 2003. Mathematical Biology. Springer-Verlag, Berlin.