

Book Review

Stochastic Modelling for Systems Biology.

Darren J. Wilkinson

Mathematical and Computational Biology,
Chapman & Hall/CRC, London, UK;
ISBN: 1-58488-540-8; Hardcover; 254 pp.;
2006; £79.95.

‘Stochastic Modelling for Systems Biology’ was designed to fill an important gap in the educational materials available for students learning about modelling methods for biological systems. Specifically, while stochastic models are emerging as perhaps the preferred method for modelling cellular and subcellular biochemistry in research practice, they remain unfamiliar to most of those who are not specialists in the field. The underlying mathematical and computational methods are well described in the literature of other fields, but the translation to biological practice is largely documented only in the current scientific literature. There are few teaching materials available for these models, particularly for beginning students in biological modelling who lack the background to follow the current scientific literature or the dense mathematical treatments available in texts from other fields. The material in this book arose out of a class the author teaches on stochastic systems biology to master’s students in bioinformatics. The text therefore takes a practice-oriented approach to the material, assuming a limited background, focusing on practical considerations in model design and implementation, and making extensive use of example systems.

The text provides a solid overview of the basics of stochastic kinetic modelling for the model developer. Chapter 1 introduces the topic by covering some basic concepts and applications of modelling for biology. Chapter 2 describes some representations of biochemical models that are used throughout the rest of the text. Chapters 3 through 5 then provide background material helpful in following the later sections. Chapter 3 covers some basic probability theory and a few important probability distributions,

Chapter 4 techniques for sampling from probability distributions in general, and Chapter 5 concepts in Markov models including extensions to continuous time and space. Chapters 6 and 8 provide the bulk of the material specific to stochastic models in biology. Chapter 6 covers the basic theory, models, and methods behind standard Gillespie simulations of reaction chemistry. Chapter 8 provides a more thorough consideration of algorithmic issues in stochastic chemical modelling, including a survey of the leading methods for exact and approximate stochastic kinetic models. In between, Chapter 7 provides four extended examples: dimerization systems, Michaelis–Menten enzyme kinetics, a generic autoregulatory gene network, and the *lac* operon. Chapters 9 and 10 discuss techniques for model fitting, providing a useful if not exhaustive overview of some common methods.

This is indeed a timely addition to the literature and nicely fills the gap Wilkinson identifies in the available teaching materials for biological modelling. The practical, hands-on approach Wilkinson takes will not satisfy all readers. Treatment of background theory is sparse and even non-specialists may find that they need to delve into the suggestions for further reading. But this approach makes the book more useful and accessible to the beginner than a denser but deeper text would be. The book is filled with useful practical advice on the gaps between theoretical models and realistic systems and data sets, as well as techniques for bridging these gaps in practice. There are many pointers to other texts and primary literature that should meet the needs of those requiring greater theoretical depth than this text provides. The text also has a companion website on which the author intends to keep an up-to-date directory of literature and tools for systems biology modelling. Wilkinson’s practice-oriented approach is also reflected in the several extended examples presented in the text, which are likely to greatly help the beginner. Formal specifications for these examples are provided in the Systems Biology Markup Language (SBML), which should make it

easy for readers to get hands-on practice with the models.

There are some specific topics for which broader coverage would nonetheless have been desirable. Stochastic differential equations (SDEs) receive a limited, almost parenthetical treatment in the context of continuous state space Markov models. They are an important alternative to Gillespie-style models for stochastic simulation, though, and warrant a more thorough treatment of algorithms and examples in a text on stochastic models in systems biology. The discussion of parameter fitting is likewise briefer than might have been desirable. It covers several of the standard techniques used by statisticians—e.g. Gibbs sampling, Metropolis–Hastings and other MCMC methods—but omits consideration of the many continuous optimization methods that are also quite important to parameter tuning in practice.

There are also some minor design decisions with which one could argue. For instance, the text presents code examples in R, a standard programming language for statisticians but an obscure choice for those from other backgrounds. The author reasonably defends this decision by noting that he must pick some language and that R is public domain and relatively readable. A more widely used language or even a generic pseudo-code might nonetheless have been more accessible. The discussion of many currently popular tools and standards, while a nice complement to the practice-oriented approach of the text, may also date it quickly.

This book would be most useful for its originally intended purpose: an introductory graduate-level

course on stochastic models for biology. The text is, by itself, a bit sparse for a full-semester graduate course. But it would lend itself well to a project-based class or to supplementation with current literature, much of which Wilkinson references in his suggestions for further reading. Some of the material, particularly the treatment of parameter inference methods, seems likely to demand a firmer background in probability and statistics than the text itself provides. An instructor might therefore do better to require a full introductory class on probability or statistics as a prerequisite and omit Wilkinson's background chapters. The text could also be useful as an independent study aid or reference for scientists with at least basic understanding of reaction systems to introduce themselves to the fundamentals of stochastic kinetic modelling. And although the text does not delve deeply enough for it to be sufficient for modelling specialists, the field is sufficiently young that even the expert is likely to learn something new from browsing it. Aside from a few minor criticisms, this is an excellent text that is likely to find an enthusiastic audience among instructors of systems biology and biological modelling.

Russell Schwartz
Department of Biological Sciences,
Carnegie Mellon University,
4400 Fifth Avenue,
Pittsburgh, PA 15213, USA
E-mail: russells@andrew.cmu.edu