Nonlinear Time Series: Nonparametric and Parametric Methods.

Jianqing FAN and Qiwei YAO. New York: Springer-Verlag, 2003. ISBN 0-387-95170-9. xix + 551 pp. $79.95.

Although *Nonlinear Time Series* is the only part of the title to appear on the spine of this new book by Fan and Yao, the word "nonparametric" in the subtitle really deserves top billing. There are hints here and there that the authors follow the viewpoint emphasized by Tong (1990), that there is a true underlying nonlinear dynamic law generating the time series data. Nonparametric methods can help uncover this dynamic law. But the book also works from a purely pragmatic standpoint; nonparametric estimation methods are useful in constructing effective forecasting algorithms, whatever the true dynamic law. The standard time series modeling paradigm for building a forecast algorithm reduces an observed series to approximate stationarity through removal of trend and seasonal components, identifies the covariance structure and possible parametric models through the estimated autocovariance function of the resulting series, selects among competing estimated models through order-selection criteria (such as the Akaike information criterion) and residual diagnostics, and then computes optimal predictors for future values of the series. Autoregressive moving average (ARMA) models form a broad class of *linear* models that can approximate quite general autocovariance structures arbitrarily closely. Dependence outside the second-order moment structure, on the other hand, may require nonlinear models.

Nonparametric methods have a long history in time series analysis and appear throughout the standard modeling paradigm, particularly in estimation of trend and seasonal components for nonstationary time series and in estimation of spectral density functions and marginal probability densities for stationary time series. The authors hope to extend the use of nonparametric tools for identifying and estimating nonlinear time series models. These models may have flexible nonparametric specifications, or the nonparametric analysis may suggest parametric nonlinear models. Nonparametric methods are also useful in constructing predictors for nonlinear processes.

The authors provide considerable background material on both time series and nonparametrics. Introductory chapters on characteristics of time series (Chap. 2) and ARMA modeling and forecasting (Chap. 3) borrow heavily from texts by Brockwell and Davis (1991, 2002) and associated ITSM software. Subsequent chapters review important classes of parametric nonlinear time series models [threshold, generalized autoregressive conditional heteroscedastic (GARCH), and bilinear; Chap 4] and introduce nonparametric methods through density estimation (Chap. 5) and spectral density estimation (Chap. 7). The subjects of main interest—those that really set this book apart—are concentrated in later chapters: smoothing with dependent data (Chap. 6), nonparametric time series models (Chap. 8), model validation (Chap. 9), and nonlinear prediction (Chap. 10).

The broad range of topics covered in this book makes for a large and awkward load. It is like coming home from the grocery store and trying to get all of the bags into the house in one trip; losing a few things on the way up the steps, crushing a few more while pushing through the door, and cracking one or two eggs when dropping the bags on the counter. Everything in the bags must be examined carefully for scratches, bruises, and breaks, and some items are lost altogether.

This book has scratches scattered throughout, in the form of abundant errors and inconsistencies in the technical typesetting. There are quite a few bruises as well: incorrect figure references, typos in formulas, garbled phrases, and terms used before they are defined.

Some breaks are noticeable, especially the proof of Theorem 7.4. The maximum periodogram ordinate for an iid non-Gaussian sequence, suitably normalized, does indeed converge in distribution to the standard Gumbel distribution. The authors' argument makes this result appear trivial, but the key approximation they use is incorrect. (See Davis and Mikosch 1999 for a valid proof, which relies on a Gaussian approximation technique for sums of independent random vectors.) Elsewhere, there are cracks in the exposition, with statements that are not quite right, like the claim on page 420 that point transformations of weakly stationary series are weakly stationary.

After taking stock of the damage, we might ask whether anything is missing. One omission is suggested by the authors' comments on page 16 that "the validity of a parametric model for a large real data set over a long time span is always questionable," and that this, among other factors, has "led to a rapid development of computationally intensive methodologies... that are designed to identify complicated data structures by exploring local lower-dimensional structures." These comments seem to ignore the possibility of parametric *hierarchical* models, which often take the form of parameter-driven generalized state-space models in the time series context. Such models can capture a variety of nonstationary and nonlinear behaviors (e.g., Kitagawa 1987; Harvey 1989; Durbin and Koopman 2001). The hierarchical model specifies dynamics of observations given time-dependent "local parameters" (or states) and dynamics of local parameters given time-invariant "global parameters" (or hyperparameters). Such hierarchical models can often successfully describe real datasets over long time spans by allowing the local parameters to change smoothly over time, suggesting that this parametric methodology has some relationship with nonparametric methods. Indeed, certain smoothing splines can be computed using the Kalman recursions, because they are the optimal fixed-interval smoothers for an integrated random walk plus noise, a simple hierarchical time series (see, e.g., Durbin and Koopman 2001 and references therein). Some mention of this relationship, and perhaps some discussion of the authors' perspective on the use of nonparametric methods in the identification of hierarchical models, would have been nice to see.

Despite these problems, this book has much that is interesting and useful. The discussions of ergodicity in Section 2.1.4 and of mixing conditions in Section 2.6 are handy. The presentation of ARCH and GARCH models in Section 4.2 is a concise introduction to this vast literature from a statistics standpoint, and Chapter 5 gives a nice overview of nonparametric density estimation, with particular emphasis on results and references for density estimation with dependent data. In fact, most chapters end with extensive bibliographical notes. These will certainly be valuable resources for researchers, particularly in the later chapters that describe evolving areas.

These later chapters, notably Chapters 6 and 8–10, constitute the book's main contribution—topics not found in typical time series or nonparametrics texts. Chapter 6 covers smoothing with dependent data, in both the time domain (a standard topic in traditional time series analysis) and the state domain (not so standard). Chapter 8, on nonparametric time series models, includes functional coefficient autoregressions, additive autoregressions, and index models, among others. Model validation, in Chapter 9, focuses on generalized likelihood ratios for testing against nonparametric alternatives (for which the nonparametric maximum likelihood estimator may not exist or may be too constrained to be of use). Chapter 10 covers nonlinear prediction, including point predictors, minimum-length prediction intervals, and predictive distributions.

Many of the interesting ideas presented in these chapters are highlighted through examples worked out in detail. These include both classic datasets (Canadian lynx and Wolf's sunspots, naturally), and new examples. Chapter 8 provides some informative examples of financial applications (a technical trading rule applied to pound/dollar exchange rates, and a value-at-risk analysis for the Standard and Poor's 500 index).

The material presented in these chapters includes techniques that anyone with a solid background in time series analysis could appreciate and could implement immediately with standard software. Even so, as the authors point out, there are plenty of open questions about these techniques. This makes the material in these chapters appealing to practitioners and researchers alike, although the practitioner would need to pick through lots of technical detail to extract the useful applied bits. *Nonlinear Time Series: Nonparametric and Parametric Methods* is best suited as a stimulating research monograph. It is not a textbook (in particular, it has no exercises), but it has sufficient breadth that it could serve as the focus of a graduate reading course or as a source of supplemental teaching materials for an advanced time series class.

Jay BREIDT
Colorado State University

REFERENCES

Brockwell, P. J., and Davis, R. A. (1991), *Time Series: Theory and Methods* (2nd ed.), New York: Springer-Verlag.

——— (2002), *Introduction to Time Series and Forecasting* (2nd ed.), New York: Springer-Verlag.

Davis, R. A., and Mikosch, T. (1999), "The Maximum of the Periodogram of a Non-Gaussian Sequence," *The Annals of Probability*, 27, 522–536.

Durbin, J., and Koopman, S. J. (2001), *Time Series Analysis by State Space Methods*, Oxford, U.K.: Oxford University Press.

Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, U.K.: Cambridge University Press.

Kitagawa, G. (1987), “Non-Gaussian State-Space Modeling of Nonstationary Time Series” (with discussion), *Journal of the American Statistical Association*, 82, 1032–1063.
Tong, H. (1990), *Non-Linear Time Series: A Dynamical Systems Approach*, Oxford, U.K.: Oxford University Press.

Statistical Inference and Simulation for Spatial Point Processes.

Jesper MØLLER and Rasmus P. WAAGEPETERSEN. Boca Raton: Chapman & Hall/CRC, 2004. ISBN 1-5848-8265-4. xv + 237 pp. $69.95.

This book is an extremely well-written summary of important topics in the analysis of spatial point processes. The text is an agreeable blend of technical and heuristic approaches, containing a thorough presentation of spatial point process models and a detailed survey of methods for their simulation and inference. The authors do an excellent job focusing on those theoretical concepts and methods that are most important in applied research. Although other good books on spatial point processes are available, this is the first text to tackle difficult issues of simulation and simulation-based inference for such processes, including methods based on Markov chain Monte Carlo and related techniques. As the authors correctly note, as computer power and speed increase, and because analytic expressions for expectations of statistics for complex point process models are often unavailable, simulation-based approaches for spatial point processes should become increasingly important and widespread.

Readers may find the text moderately difficult to read. The level of exposition is about halfway between Brian Ripley’s (1981) brilliantly simplistic *Spatial Statistics* and the far more theoretical text by Daley and Vere-Jones (1988), which deals little with spatial point processes but is widely (and correctly) considered the indispensable book on point processes in general. While Møller and Waagepetersen’s book focuses on important practical topics, such as simulation and inference, it is less a manual for applied statistics than a description of key concepts and mathematical results justifying important techniques used in applied research. Considering that the text states most results in their full mathematical precision and includes proofs of key theorems, it is remarkably easy to follow.

The book has appeal far beyond the realm indicated by its title. Following a brief introduction featuring some examples of spatial point process datasets, the book provides a terrific summary of a wide variety of spatial point process models, and in fact hardly even mentions simulation and related topics until the seventh of its eleven chapters. Basic methods for the description, estimation, and display of key features of point processes, which are subjects described in great length in other texts such as those by Ripley (1981), Diggle (1983), and Cressie (1993), are all packed into Chapter 4 under the heading “Summary Statistics,” which makes this 27-page chapter an appealing dense summary of such resources. The later chapters delve meticulously into simulation procedures for various models, even including detailed algorithms for simulation methods and their use in likelihood inference.

The authors have a very impressive knack for explaining complicated topics very clearly, and readers unfamiliar with the subject matter will benefit greatly from their expositions, some of which are quite innovative. For instance, most other authors start by describing a point process heuristically as a random countable collection of points in some space, but then proceed to define it instead as a random measure or stochastic process, so that the subject is embedded in a more general and more theoretically developed research area. Instead, Møller and Waagepetersen continue throughout to define a point process simply as a set, that is, a countable collection of points. This makes things considerably easier for the reader less familiar with measure theory, and it is remarkable how little difficulty the authors have in explaining rather sophisticated concepts or techniques using this definition. For instance, superpositions of point processes are expressed as unions of random sets, and the nearest-neighbor function in Chapter 4 is defined using set differences. The more experienced reader may find these changes rather unconventional, but they do not cause any real problems and do often seem to simplify the exposition.

The book’s main weakness is in its use of examples. The authors draw the inexperienced reader in with several rather nicely graphically depicted examples of spatial point process datasets, which illustrate the scope of the methods in the text. However, these examples are less than stimulating, and it is unclear what the questions of primary interest relating to these datasets are. Why, for instance, should the reader be interested in the locations of 1,382 weed plants in a Danish barley field? Perhaps in an effort not to distract attention away from the book’s primary focus, the examples are not very thoroughly explained. The reader is provided little information on design and sampling issues, available covariates, and background scientific knowledge, all of which would be very important for actual applications. These omissions are very understandable; such information might seem tangent to the book’s main topic, and is not provided by other spatial point process books either. Unfortunately, however, the models and summary statistics applied to these datasets fail to inform the reader of much of profound interest, and so the reader may be left with the impression that the main purpose of such datasets is to facilitate the understanding and appreciation of spatial point process models, rather than the other way around.

The authors claim in the Preface (p. xiii) that the text is intended to be accessible to “senior undergraduate students and Ph.D. students in statistics, experienced statisticians, and applied probabilists.” The book’s mathematical rigor (as well as its lack of homework exercises) make it too difficult to use for an undergraduate course, but it could certainly be used for graduate students, especially if supplemented with a project involving some of the concepts in the text.

Statistical Inference and Simulation for Spatial Point Processes will no doubt prove an outstanding resource for researchers and students interested in spatial point processes. Its excellent survey of the vast array of models is reason enough to own it. As computer technology and speed advance and simulation continues to play an ever-increasing role in statistical inference, the authors’ clear, detailed, and comprehensive survey of simulation methods for spatial point processes will become increasingly important.

Frederic P. SCHOENBERG
University of California Los Angeles

REFERENCES

Cressie, N. A. C. (1993), *Statistics for Spatial Data* (2nd ed.), New York: Wiley.
Daley, D. J., and Vere-Jones, D. (1988), *An Introduction to the Theory of Point Processes*, New York: Springer-Verlag.
Diggle, P. J. (1983), *Statistical Analysis of Spatial Point Patterns*, London: Academic Press.
Ripley, B. D. (1981), *Spatial Statistics*, New York: Wiley.

Distribution Theory of Runs and Patterns and Its Applications: A Finite Markov Chain Imbedding Approach.

James C. FU and W. Y. LOU. River Edge, NJ: World Scientific Publishing, 2003. ISBN 981-02-4587-4. x + 162 pp. $38.00.

This book’s title gives an accurate impression of its content. The basic model is a sequence $X_1, X_2, \ldots$ of independent random variables having a common distribution on a finite set of outcomes. A *simple pattern* is a chosen finite string of outcomes, and a *compound pattern* is a finite set of simple patterns. The problem is to calculate the distribution of the number of times that the pattern is observed in the finite sequence $X_1, X_2, \ldots, X_n$, using either “overlap counting” or “nonoverlap counting.” For instance, in coin tossing, a simple pattern of interest might be *HH*. In the finite sequence *HHHH*, the chosen pattern occurs twice with nonoverlap counting and three times with overlap counting.

It might seem amazing that anything new can be said about a setting and problems that could have been formulated in the seventeenth century. As the book’s subtitle reveals, the authors’ aim is to present the method of “finite Markov chain imbedding,” and although this can be seen as a reformulation of a known technique, it does provide a helpful framework for handling the more intricate run and pattern distributions.

Let N_n be the number of occurrences of the chosen pattern, using one or other method of counting, within the finite sequence $X_1, X_2, \ldots, X_n$. The random variable N_n is called *finite Markov chain imbeddable* if

$$P(N_n = x) = P(Y_n \in C_x)$$

for all values x taken by N_n. Here $Y_0, Y_1, \ldots, Y_n$ is a Markov chain on some finite state space, started with some particular initial distribution, and the sets C_x form a partition of the state space.

To see how this can be done in a simple case, let $X_0, X_1, \ldots$ be independent coin tosses, with values H and T, and let the pattern be $\{HH, TT\}$. On setting

$$I_t = \begin{cases} 1 & \text{if } X_t = X_{t-1} \\ 0 & \text{if not,} \end{cases}$$