Time Series Analysis: Forecasting and Control,5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc...

Article *in* Journal of Time Series Analysis · March 2016

DOI: 10.1111/jtsa.12194

CITATIONS
77

READS
60,608

1 author:

Granville Tunnicliffe Wilson
Lancaster University
65 PUBLICATIONS **1,668** CITATIONS

SEE PROFILE

JOURNAL OF TIME SERIES ANALYSIS
J. Time. Ser. Anal. (2016)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1111/jtsa.12194

BOOK REVIEW

TIME SERIES ANALYSIS: FORECASTING AND CONTROL, 5TH EDITION, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1

Draft chapters of the first edition of this book were exchanged across the Atlantic by George Box and Gwilym Jenkins just 50 years ago. Published in 1970 by Holden-Day, it soon became, and still remains, a classic text in its field. The second edition in 1976 included a section with exercises to enhance the use of the book as a course text; otherwise, the additions and amendments were modest. Gwilym Jenkins died shortly before his 50th birthday in 1982, following a long illness, and the third edition, published by Prentice Hall, did not appear until 1994, with Greg Reinsel as a new co-author. By that time the subject had expanded considerably and much new material was added to the book. The fourth edition, published by Wiley in 2009, included two new chapters. In his preface to that edition, George Box credited much of the new material in that and the previous edition, to Greg Reinsel, who died suddenly and unexpectedly at the age of 56 years, a few years before that edition was printed. Greta Ljung was asked by George Box to help in the preparation of this fifth edition, published last year by Wiley. The material has been in part re-organized, expanded and updated, the exposition enhanced by revising and modifying text, and the presentation of tables and graphs made more effective. The exercises, previously in their own section, now appear at the end of each chapter, with their number increased. Sadly, after a long and exceptionally distinguished life, George Box died before this edition appeared. In her preface, Greta Ljung dedicates this book to his memory and that of his previous distinguished co-authors.

I can recommend this book, in the first place, for the content of the first edition. Although enhanced in subsequent editions, the importance of this original material is profound, and many of those coming new to the analysis of time series would benefit from its instruction. A high standard of modelling of the structure of time series is paramount for valid applications, whether in forecasting or in control. Box and Jenkins achieve this through their scheme and philosophy for building ARMA models of the stochastic structure of time series and transfer function models for the dynamic structure of dependence between time series. Their particular formulation by which a dependent variable is represented as the sum of two components, the dynamic dependence on an explanatory variable and the remaining, possibly seasonal, stochastic variation, has proved of enormous value over subsequent years. Examples from my own experience are as varied as modelling the dependence of daily gas demand on temperature, the weekly abundance of moths on rainfall in the UK and the weekly sales of consumer products on advertising promotions. Box–Jenkins modelling, as it is now widely known, enables valid inference of the magnitude and timing of the dependence between time series, in contrast to the misleading inference that might be obtained from simple regression. The value of the models is often more in the wider sense of prediction than forecasting and control, informing policy and planning decisions as, for example, on the investment in product promotions.

The original work was motivated by the modelling of manufacturing processes suffering from substantial stochastic disturbances that affected the efficiency of their control and quality of output. The original book therefore devoted two chapters to explaining how the models, once constructed, could be used to improve process control. This material has its foundations in the area of statistical process control rather than the engineering methods of automatic process control but serves to link these areas. To emphasize the developments in industrial quality

improvement, the material on control was completely rewritten as a single chapter in the third and subsequent editions of the book. This section is unfamiliar to many who might claim to have an understanding of Box–Jenkins methodology, but is of value to those working in the economically very important area of process improvement through experimentation and model building.

However, it is the seasonal ARIMA model and model building that became the most well-known methodology of Box and Jenkins, and is the topic of Chapter 9 in this, as in the original, edition. New ideas in the stochastic formulation of seasonal models appeared in the aftermath of the first edition and are now usefully described in this chapter of the present edition, although the reader is referred to other texts, which give fuller accounts. One major new formulation was the structural component seasonal time series model and its state space formulation. The state space representation is first introduced for (non-seasonal) ARIMA models in Chapter 5 as a means of constructing forecasts. The Kalman filter is presented there as the technique for estimating the states and constructing the forecasts conditional on past values. The estimation of smoothed state variables conditional upon all observed series values is also explained and derived. The use of state space methods for calculation of the likelihood for ARIMA model estimation is described in Chapter 7. This material, which first appeared in the third edition, is a valuable addition to the original. Chapter 9 also describes the closely related unobserved components model and how the components can be estimated by more classical Wiener filtering methods, as well as the Kalman filter. These methods have been widely applied to seasonal adjustment, but only very brief mention is made in Chapter 9 of this topic. The seasonal models have been very widely applied to extend the X-11 seasonal adjustment methods of the USA Census Bureau and its subsequent versions, but this receives no mention. There have also been developments, over the years, of automatic methods of identifying seasonal models, as well as different statistical strategies for aid in reliable model identification. Some reference to, or summary of, these strategies would help less experienced practitioners make good or better use of these models.

The third edition also saw the introduction of a chapter on intervention analysis, outliers and missing values, which remains as Chapter 13. The methodology for this, and the topic of regression models with time series (ARIMA) error terms presented towards the end of Chapter 9, is implicit in, and was facilitated by, the development of transfer function model building and estimation. The importance of these topics justifies, however, their separate treatment and adds further to my recommendation of this book. I have recently read an article on forecasting call centre volumes by an experienced professional statistician who uses multiple regression methods with dummy variables and lagged values of the volumes series, described as an autoregressive model. Such methods are still widespread when best time series practice, established for decades and widely available using high-quality licenced or free software, is to use regression with an autoregressive (or ARIMA) model error term. Chapter 13 alone justifies the purchase of this book by those faced with such forecasting tasks, or training others how to solve them.

Chapters 10 and 12, first introduced in the fourth edition, are now 10 and 14 and have both been expanded in this edition. The first of these is an overview of topics in univariate time series that, again, emerged in the years following the first edition. These topics are tests for unit roots in ARMA models, conditional heteroscedastic models, nonlinear time series models and long-memory time series processes. The strong foundation of ARIMA model building presented in earlier chapters is a valuable prerequisite for studying these topics. This chapter gives a useful grounding in the ideas and methodology of these topics from which the reader can progress to other texts, which give a much fuller treatment. In contrast, Chapter 14, written originally by Reinsel, gives a very thorough introduction to multivariate or vector ARMA models and their extension to integrated multivariate processes, although Reinsel's own book on this subject should be read for an in-depth treatment. The identifiability of the models, their canonical forms and identification methods using canonical correlation are described and illustrated using a bivariate example. Vector autoregressive models are the models of choice for much econometric time series analysis. The treatment, in a short section, of integrated forms of these models and inference on co-integration is necessarily limited yet still useful as an introduction.

Spectral properties of time series models are only briefly mentioned, in Chapters 2 and 3 and the appendices of Chapters 12 and 14. Model order identification methods, which were only just becoming widely used when the second edition appeared, are now extensively referenced. The exponentially weighted moving average form of forecasts is mentioned as a property of the IMA(1,1) model, but there is no description of the many extensions

DOI: 10.1111/jtsa.12194

of such exponential smoothing forecasting methods to seasonal time series that are still being actively used and developed. Software to implement Box–Jenkins methods, originally written by their respective research groups, was soon promoted professionally and did much to disseminate the methods and their application before the major statistical packages incorporated them. A feature of this edition is the introduction of the freely available **R** programming language as a tool for implementing much of the modelling methodology described in the book; it has also been used to enhance and update the numerical illustrations. In places, **R** code is listed to enable the reader to reproduce graphical and numerical results and, more generally, will be useful for teachers and practitioners of the methods.

The book is nicely structured and presented, well referenced and indexed. The material contributions of Greta Ljung to its content are substantial, although dispersed throughout the book rather than focused on a single chapter or two. The care and thought she has taken with the overall compilation of the material have produced a volume that remains at the forefront of its subject and should be read by all who wish to attain a high standard in the practice of time series analysis and forecasting.

GRANVILLE TUNNICLIFFE WILSON
Dept. of Mathematics and Statistics
University of Lancaster, UK

J. Time. Ser. Anal. (2016)
DOI: 10.1111/jtsa.12194