Journal of Statistical Software

February 2016, Volume 69, Book Review 6. *doi: 10.18637/jss.v069.b06*

Reviewer: Dipankar Bandyopadhyay
Virginia Commonwealth University

Sampling Spatial Units for Agricultural Surveys

Roberto Benedetti, Federica Piersimoni, Paolo Postiglione
Springer-Verlag, Berlin, 2015.
ISBN 978-3-662-46007-8. 325 pp. USD 179.00.
http://www.springer.com/9783662460078

Currently, there are a number of nice textbooks available on spatial statistics and survey sampling (Bivand, Pebesma, and Gómez-Rubio 2008; Lumley 2011; Plant 2012; Banerjee, Carlin, and Gelfand 2014), with an emphasis on exploring the subject matter via the R software. I take great interest in using parts of these books while teaching a PhD level course in spatial statistics. In particular, if the application area is on ecology and agriculture, I would most certainly rely on Prof. Plant's book (Plant 2012). However, keeping up with the recent momentum on 'big data', there has been a dearth of a nice book that summarizes the tools necessary for handing spatial referencing while analyzing datasets from complex agricultural surveys, or in more general terms, from national level surveys. I personally thank Professors Benedetti, Piersimoni and Postiglione for filling out this void with this commendable work. This book is a meticulously organized treatise of applying spatial data methods to sample surveys (primarily in agriculture), with the computational engine powered by the R software.

Little over 300 pages, the book is divided into 12 chapters. Chapter 1, the introductory one, is all about definitions and statistical concepts pertaining to spatial data methods and survey sampling. After introducing the concepts of finite population sampling mainly from the design-based perspective, the three broad kinds of spatial datasets (viz. geostatistical, lattice and point pattern) are discussed with nice graphical illustrations. Chapter 2 is mainly a historical anecdote on agricultural surveys, with various examples that relied on spatial reference frames. Chapter 3 introduces the basics of geographical information systems and the spatial connection, while Chapter 4 presents various aspects of analyzing remotely sensed datasets, including dimension reduction, information extraction, unsupervised and supervised classification, etc. Moving on, Chapter 5 explores choice of the statistical unit for data collection. In particular, I enjoyed reading Section 5.3 which discusses the advantages and disadvantages of using the spatial frames over the list frames. The chapter on sampling designs (Chapter 6) is pretty standard, and can be found in any standard book on sampling theory. I would consider Chapters 7 and 8 (on spatial sampling designs and related sample size estimations, respectively) as the crux of this book. After introducing the necessary details corresponding to each section, R codes are presented for demonstration. Chapter 9 discusses

a variety of steps that constitute survey data collection and processing. Here, the authors justifiably state that this topic is usually neglected in any standard book on sampling theory. Chapter 10 considers advanced topics on complex surveys, such as efficiency enhancement of the estimation process via introduction of auxiliary information, related calibration approach, non-response adjustments, etc, all adjusted to the spatial framework. No book on survey sampling is complete without a chapter on small area estimation, and Chapter 11 explores that, once again in conjunction to the spatial angle. Finally, as a continuation of the debate between design-based/model-based approaches, Chapter 12 presents (spatial) model-based survey methods.

On the overall, the presentation mode is lucid with excellent sequencing of concepts. I enjoyed going through most chapters, some with considerable details. The level of mathematical rigor is acceptable to someone with a good background on calculus, statistical inference of the Cassella-Berger (Casella and Berger 2002) level, matrix algebra and some knowledge of complex surveys. A noteworthy feature of this book is the availability of well-documented R codes in almost every chapter, whenever necessary, for practical illustration of the methodology described. This feature can assist a user to quickly apply the methods, or modify, without combing through the theoretical details. Arguably, in recent times, the R language has emerged as the language of choice for statistical programming, and the authors considered this with seriousness while developing the content. Another interesting feature is the list of references at the end of each chapter pertaining to the topics discussed, allowing the reader easy access without scrolling to the end of the book.

For a one-semester special course on this topic, this book may not serve as the sole book that I can rely upon. It does not include well-thought-out problems at the end of each chapter. Also, the treatment is entirely classical and design-based, and would not cater to someone looking for Bayesian methods. It is mainly an intermediate-level reference book for graduate students and (agricultural) researchers to get introduced to the nuances of spatial statistics in survey sampling, and quickly move to hands-on computing. Yet, without any suitable competitor, it is the current book of choice. If you are enamoured with the versatility of R, I highly recommend buying it.

References

Banerjee S, Carlin BP, Gelfand AE (2014). *Hierarchical Modeling and Analysis for Spatial Data*. CRC Press.

Bivand RS, Pebesma EJ, Gómez-Rubio V (2008). *Applied Spatial Data Analysis with R*. Springer-Verlag.

Casella G, Berger RL (2002). *Statistical Inference*, volume 2. Duxbury, Pacific Grove.

Lumley T (2011). *Complex Surveys: A Guide to Analysis Using R*. John Wiley & Sons.

Plant RE (2012). *Spatial Data Analysis in Ecology and Agriculture Using R*. CRC Press.

Reviewer:

Dipankar Bandyopadhyay
Department of Biostatistics
School of Medicine, Virginia Commonwealth University
Richmond, VA 23298-0032, United States of America
E-mail: dbandyop@vcu.edu
URL: http://people.vcu.edu/~dbandyop/

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
February 2016, Volume 69, Book Review 6 *Published:* 2016-02-01
doi:10.18637/jss.v069.b06