

## Book Review

EPB: Urban Analytics and City Science

2020, Vol. 47(7) 1309–1310

© The Author(s) 2020

Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)[journals.sagepub.com/home/epb](https://journals.sagepub.com/home/epb)

Robin Lovelace, Jakub Nowosad and Jannes Muenchow. *Geocomputation with R*. Chapman and Hall/CRC Press: UK, 2019, 335 pp. ISBN: 9781138304512, £66.99 (hbk), ISBN: 9780203730058, £60.29 (eBook)

Resources and code: <https://geocompr.robinlovelace.net>

**Reviewed by:** Chris Brunson, National Centre for Geocomputation, Maynooth University, Ireland

DOI: 10.1177/2399808320945456

This book sets out to explain the key ideas in geocomputation – more specifically manipulating, visualising, modelling and analysing geographical data. A further aim is to achieve all of this using only open source software. Developments of recent years have made this an achievable goal, and this book provides a good starting point for readers wishing to do this. The authors already have a strong reputation for their work with R, in particular their promotion of open and reproducible research methods in the areas of human geography and planning. The book is divided into three parts. The first two parts follow a natural sequence from the basics of how geographic data are represented and handled in R. With timely emphasis on the more recent ‘sf’ (simple features) package and data model, in the first part, to more advanced modelling and other techniques using R and related software (such as QGIS) in the second. The third part showcases a number of ‘real-world’ applications demonstrating how the ideas of the first to parts may be put into practice.

The first part (*Foundations*) outlines the ‘nuts and bolts’ of R as a tool for manipulating geographical data – and highlights a number of features of the *simple features* (sf) approach to manipulating geographical data. One area that is focussed on and plays an important role in sf’s approach is the use of *units* – for example specifying areas as square meters or square feet. Keeping track of these attributes is important to ensure like-for-like operations are carried out between variables storing measurement data – in a similar way to the importance of keeping a check on the map projection used for geographical objects. This is a relatively new feature and is nicely explained here. Another highlight of this part of the book in chapter 7 is the discussion of the use of web data resources within R. As in much of the book, the technical issues are well outlined and thoroughly backed up with R examples, followed by exercises at the end of each chapter.

The second part of the book (*Extensions*) turns to more advanced topics. Interestingly, explaining the concepts of algorithms is saved until this section, with operations in R covered up to this point generally achieved by typing in individual commands at the R console. The notions of loops of operations and conditional statements are introduced here.

Perhaps as someone familiar with R, and before that with writing algorithms more generally, I found this unusual – but that said, far from unreasonable. This part of the book also goes on to outline linking R with GIS software, using R itself to create maps, as well as statistical modelling. The final part (*Applications*) considers real-world geographical and planning problems tackled with R. Rather than a general overview with a brief discussion covering a large list of topics, the section selects three major areas (reflecting the authors' expertise) and focuses on these. This works well, and, although the subject areas might not be exactly those of the reader, the detailed coverage of R operations afforded by this approach should transfer reasonably well to similar problems in cognate topics.

In terms of style, the book is very much 'hands-on' and provides some very useful information on more practical issues. For example, although R is syntactically equivalent across operating systems, some packages may require different installation procedures. These are a very welcome inclusion here, as a book of this kind is often used in front of a computer trying out the code examples. Making an effort to ensure the machine is 'up and running' for a wide range of users is crucial. An interesting point is that recent years have seen a (re)turn to script and code-based approaches for geographical data processing, with the adoption of R and Python (and a number of other programming languages to some extent) as popular tools. Clearly this book reflects this, although interestingly the fundamentals of algorithm design (functions, loops and so on) are not formally introduced until almost the end of the second part of the book. This is perhaps a consequence of the fact that R is an interactive command line tool and a great deal can be achieved by entering functions from existing packages directly into the console, particularly when working with 'tidy' data manipulation tools such as those provided by 'dplyr' and companion packages. Although the focus is on R, as the book title suggests, in terms of script-based vs. GUI-based approaches, the tone is fairly agnostic; Chapter 9 considers bridges to GIS software, including QGIS, SAGA and GRASS, all of which offer a more GUI-based approach. A discussion of the pros and cons of CLI vs. GUI is included here, although some of the points made reflect the current situation rather than fundamental distinctions. For example, provision of support for some geodatabase management system may not currently be available in R, but there is no reason why this could not be implemented in the future. However, perhaps for a book that is practical in tone, this is not a major issue.

The focus of the book is perhaps not strongly on data analysis – statistical modelling is mainly covered in a single chapter – but other texts cover this in more detail. The strength of this book is therefore on the computational aspects (for example producing R-based interactive web sites using shiny) and as a comprehensive overview of the kinds of data one is likely to work with (for example taking care to ensure raster and vector data are both well represented here). The final part showcasing real-world applications is also strong. Another highlight of the book are the exercises at the end of each chapter. Some are coding exercises, essential on a 'practice makes perfect' principle, while others focus on understanding the underlying ideas, and as such enable readers to innovate as well as merely acquiring a fixed set of skills. Finally, the book comes with excellent resources, including an electronic version of the book itself (as a *bookdown* output) together with all the datasets and useful links (<https://geocompr.robinlovelace.net>). The emphasis of this book is perhaps most strongly on data manipulation, management and visualisation. If these are your main needs, it provides a comprehensive and readable companion.