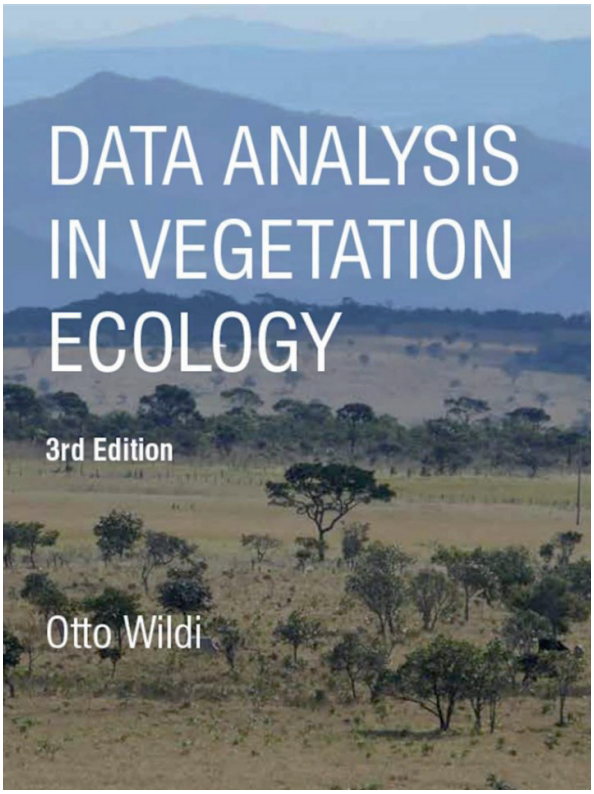
Community Ecology 18(3): 319-320, 2017
1585-8553 © Akadémiai Kiadó, Budapest
DOI: 10.1556/168.2017.18.3.11

Book review

Otto Wildi. 2017. **Data Analysis in Vegetation Ecology.** 3rd ed. CABI, Wallingford, UK. xxi + 333 pages. ISBN: 978-1-78639-422-4. Price: €55.00. Also available as e-pdf and e-pub.



This is the third edition of a book, earlier versions of which were handled by a different publisher. The author, Otto Wildi from the Swiss Federal Research Institute (Birmensdorf), has a life-long experience in the field of data analysis in vegetation science. He has published quite many research papers, books and book chapters and his own multivariate data analysis software, which guarantee by themselves the high quality of the material presented. In addition to conventional multivariate analytical methods, largely belonging to the area of cluster analysis and ordination, the book has much more to tell us, by discussing novel analytical approaches that have received increasing attention in the past decades. In other words, the contents have been updated to suit the needs of practically any scientist working in the field of vegetation ecology.

The book is divided into fourteen chapters plus a bibliography, two appendices and a subject index. The Introduction is unusually short, giving only a brief explanation of fundamental paradigms that may rule vegetation analysis. Of these, I must emphasize "Similarity theory" originally developed by L. Orlóci and E. Feoli which provides a conceptual underpinning of any approach concerned with the joint analysis of biological and environmental data. It is important that the author correctly separates the real world from the *data world*, the "image", which can only be some simplified representation of reality. Then, the whole book is concerned with procedures to construct, manipulate, evaluate and interpret various forms of this image.

Pattern recognition and the sampling procedures appropriate to this purpose come next. At this point, the reviewer misses some more details of sampling theory because this is the first, and perhaps the most crucial step that takes us from the real world to the image – thus influencing all other subsequent steps of the study. At the end, we find examples from the R package, a practice followed by the author throughout the book.

Data types and their transformation are the subject matter of the next part of the book. The discussion of scale types is most welcome here, because this is an issue frequently overlooked or downweighed in importance in similar books and research articles. Without knowledge of the possibilities, data analysis may produce mathematically incorrect and ecologically uninformative results – for which we have published examples from noted journals.

The central issue of the next chapter is resemblance in multivariate space. The reader is acquainted with the most frequently applied measures of distance, dissimilarity and correlation. One of the few truly embarrassing errors of the book appears here: function 4.4 is not named on p. 40 (it is the normalized Canberra metric), and in its formula on the next page division is by p rather than by pp.

Classification of vegetation data is described in Chapter 5. Here, the different colors are especially useful in improving the interpretability of figures – this a definite novelty of the book, in related works published to date the color option was rarely used. It is a pity, however, that the author (like many of his predecessors) has chosen the wrong term for incremental sum of squares clustering on p. 62 – it is obvious even from the figure caption that it is not variance that is being minimized by this algorithm.

The most commonly used multivariate analysis procedures belong to the domain of ordination: the "art" of efficient reduction of dimensionality and gradient recovery. Chapter 6 begins with an introduction to the most fundamental ordination method, namely principal component analysis, and continues with the description of almost all currently applied techniques. "Almost all", because one would expect the appearance of constrained ordination methods, such as redun-

dancy analysis, in this part of the book. The discussion of these most popular techniques, however, is deferred somewhat unexpectedly to section 7.4 in the next chapter, which starts with pattern analysis. In other words, ranking methods, the Mantel test and correlograms are, for example, inserted in between – which does not seem entirely logical. Then, the chapter concludes with tabular rearrangement techniques, another important toolkit in vegetation science.

Chapter 8 turns to the most fashionable subject of trait-based analysis in community ecology. This is especially important at large scale, when species based analyses are meaningless or even impossible. Matrix operations are richly illustrated here, although some of them would have been more didactical to occur earlier, at the beginning of the ordination chapter.

The following chapters are devoted to topics that are usually missing from standard vegetation ecology books. These include static predictive modeling in Chapter 9 which gives the opportunity to describe generalized models and regression trees. Changes in time may be evaluated by methods introduced next, with many examples on time series data. The author devotes quite few pages to Markov chains, another topic infrequently occurring in ecology textbooks. The same is true for the next chapter, which discusses dynamic modeling. In this, the author reminds us of classical models of population change, and does not forget about space processes either. The subject of classification is revisited in Chapter 12 in which important details, such as the revision of vegetation classifications and the relevance of nomenclature.

A big advantage of the book is that all methods have been demonstrated using the same program package (however less friendly it seems first) and a relatively small collection of examples. This is the main reason that I can warmly recommend it to everyone who wants to learn more on this complicated subject matter. Further advantage is that the data sets described in the book are freely available online, and the same is true for the R program package. The publisher contributed greatly to the high quality of its output by adding colors to practically all illustrations. I am absolutely happy to own a copy – and will sit down soon to learn R.

János Podani
Eötvös University,
Budapest, Hungary