Manuscript of article in press in the *Encyclopedia of Ecology, Second edition*
Brian Fath, Editor-in-Chief

Numerical Ecology

Pierre Legendre

Université de Montréal, Montréal, QC, Canada

Author contact information

Département de sciences biologiques, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Québec, Canada H3C 3J7

E-mail: Pierre.Legendre@umontreal.ca

Abstract

Numerical ecology is the field of quantitative ecology devoted to the numerical analysis of data, mostly multivariate, with emphasis on community composition data. It is a sub-discipline of ecology, not of statistics or other mathematical discipline. Many of the methods used in numerical ecology have been developed by ecologists, specialists of classification methods, geneticists and other researchers who were facing questions about multivariate data in their fields of study. The field of numerical ecology results from the work of a large number of dedicated scientists who experimented with numerical methods of analysis before and in the computer era. They developed a broad palette of methods of analysis to answer specific ecological questions. The article lists some of the pioneer researchers, as well as the scientists who wrote the first textbooks describing how numerical and statistical analysis of ecological data could and should be done to answer ecological questions and test hypotheses. The field is also indebted to the scientists who developed the many computer packages now available for numerical analysis of ecological data.

Keywords: Beta diversity; Canonical analysis; Clustering; Community composition data; Dissimilarity; Fourth-corner problem; Multivariate data; Ordination; Paleoecological analysis; Quantitative ecology; R software; Space-time analysis; Spatial analysis; Spatial eigenfunction analysis; Temporal analysis

Introduction

Numerical ecology is the field of quantitative ecology devoted to the numerical analysis of [mostly multivariate, but also time series] ecological data, with emphasis on community composition data. Community ecologists, whose data are multivariate by nature (many species, several environmental variables), are the primary users of these methods. Hence, population dynamics, single-species distribution models or the analysis of single species spatial patterns, which are powerful applications of mathematical ecology, are not considered parts of numerical ecology *sensu stricto*.

Numerical ecology is a sub-discipline of ecology, not of statistics or other mathematical discipline. In numerical ecology, the analysis starts with consideration of an ecological question and the data available to answer it. Numerical analysis methods are chosen to answer the question at hand and test ecological hypotheses about the data. When tests of statistical significance are in order, ecologists use permutation tests in most cases. These tests are applicable to non-normal univariate or multivariate data, in particular multivariate community composition data.

Many of the methods used in numerical ecology have been developed by ecologists, specialists of classification methods, geneticists and other researchers who were facing questions about multivariate data in their fields of study. Their training often combined statistics and some field of ecology or biology, in different proportions. So, historically, it is the people who needed to analyze intricate data sets to address high-level scientific questions in their fields of research, and had statistical or numerical training, who often developed statistical or numerical methods of data analysis. The first statisticians, people like Galton, Pearson and Spearman, who created the bases of modern-day statistics, had not been trained as statisticians either: that field did not exist before their work.

The field was reviewed and synthesized by Legendre and Legendre in five editions of a book published in French ("*Écologie numérique*") and in English ("*Numerical Ecology*") from 1979 to 2012. Because of the successive editions of this successful book over more than three decades, people often associate the field to the names of these authors. This short article will show how these books were part of a trend in the ecological literature that started before the 1960's and involved many researchers.

When Legendre and Legendre published the first editions of their book, they called it *"Écologie numérique"* and "*Numerical Ecology*" to emphasize the lineage with the field of *numerical taxonomy*, founded in 1963 by microbiologist Peter H. A. Sneath and population geneticist Robert R. Sokal. Numerical taxonomy aimed at testing hypotheses about biological systematics, population biology, and phenetic, phyletic and phylogenetic relationships, using multivariate data analysis. The approach included explicit steps to create dendrograms and cladograms using numerical methods, instead of the subjective syntheses of data that were generally favored until then. Likewise, numerical ecology includes steps to test ecological hypotheses using data and explicit methods of numerical analysis.

A brief history

Pioneer researchers

Numerical ecology developed thanks to the work of a numerous researchers. Until about 1970, community ecology had been mostly a descriptive science, although some ecologists had ventured into mathematical analyses. Pioneer researchers who developed key concepts and numerical methods of great importance for multivariate data analysis include the following well-known scientists:

• Vegetation scientist Paul Jaccard, working in the Alps, developed the first similarity coefficient used to analyze vegetation survey data (Jaccard 1900). His coefficient is still in wide use nowadays in all fields where scientists analyze multivariate presence-absence observational data.

• In 1954, the vegetation ecologist David Goodall was the first to use factor analysis in community ecology. Goodall proposed the term "ordination" to designate this type of analysis, a term now widely used in textbooks and publications in community ecology and many other fields (Goodall, 1954). At that before-computer time, several other ecologists had experimented with numerical methods to address ecological questions.

• Robert R. Sokal (State University of New York, Stony Brook, USA) developed *numerical taxonomy* with Peter H. A. Sneath (their foundation textbooks were published in 1963 and 1973; see Bibliography) and promoted the use of multivariate data analysis in biology and ecology. These two researchers proposed several methodological developments, including similarity coefficients and clustering methods, and experimented with the use of computers.

• John C. Gower spent his career (1955 to present) developing numerical methods of analysis for numerical taxonomy, numerical ecology and agricultural experimentation. He was also a pioneer in the use of computers at the Rothamsted Experimental Station in England. The author of this article had the privilege to work with him on the properties of dissimilarity coefficients (Gower and Legendre 1986).

• Robert H. Whittaker (Cornell University, USA) proposed the five-kingdom taxonomic classification of the world's biota into the Animalia, Plantae, Fungi, Protista, and Monera (Whittaker 1969) and developed the key ecological concepts of alpha, beta and gamma diversity (Whittaker 1972). He hired young collaborators who wrote and distributed important computer software for community ecology.

• Cajo J. F. ter Braak (Wageningen University, The Netherlands) developed canonical correspondence analysis and many other key methods related to canonical ordination. He also wrote the Canoco program (the first version was developed in 1985), which was the first generally available software for community ordination, simple and canonical. Successive versions were released to researchers from 1988 (version 2.1) and in the following years up to now. The history of the Canoco software is recounted in ter Braak (1988) and in section 1.2 of the successive versions of the Canoco manual, e.g. ter Braak and Šmilauer (2002).

Numerical ecology is the result of many years of collaborative work among many dedicated researchers in the fields of numerical classification and quantitative ecology, too many to be listed here. Many of them are cited in the *References* sections of the "*Numerical Ecology*" and "*Numerical Ecology with R*" books. These collaborations are illustrated in Figure 1.

Numerical ecology has been able to make great progress in the computer age thanks to the dedication of many developers of statistical packages, especially in the R language, who wrote software designed to analyze ecological data. Cited here, in alphabetic order, are some of the packages, available on the Comprehensive R Archive Network (CRAN) Web site, that have been developed by and for ecologists: ade4, adespatial, BiodiversityR, cocorresp, codep, ecodist, FactoMineR, FD, labdsv, lmodel2, mvpart, pastecs, picante, princurve, rioja, vegan, vegclust. That list is not exhaustive. Many other packages and functions are available on researchers personal Web pages or in appendices of published papers describing statistical methods for ecological analysis.

Textbooks

From 1969 to 1979, the contributions of the previous decades were synthesized in four textbooks that marked the foundation of the field of numerical ecology:

• Statistical ecologist Evelyn Christine Pielou, Professor of mathematical biology at Queens' University (Kingston, Canada), formally introduced the field in 1969 by publishing a textbook entitled "*An Introduction to Mathematical ecology*". The Preface opened with the following sentences:

"The fact that ecology is essentially a mathematical subject is becoming ever more widely accepted. Ecologists everywhere are attempting to formulate and solve their problems by mathematical reasoning, using whatever mathematical knowledge they have acquired, usually in undergraduate courses or private study. The purpose of this book is to serve as a text for these students and to demonstrate the wide array of ecological problems that invite continued investigation." (Pielou 1969, p. v.)

• László Orlóci, University of Western Ontario, London, Canada, published in 1975 "*Multivariate Analysis in Vegetation Research*" with a clear orientation towards community ecology. The main chapters described ways of computing resemblance functions as well as methods of ordination and classification.

• Roger Green, who was Orlóci's colleague at the University of Western Ontario, London, Canada, published in 1979 "*Sampling Design and Statistical Methods for Environmental Biologists*" oriented towards animal ecology and sampling designs. The book is a comprehensive guide to the principles of sampling design and methods of statistical analysis. It reviews the principles of inference, sampling and statistical design, and hypothesis formulation, with reference to ecological data.

• The first French and English editions of the Legendre and Legendre (1979) numerical ecology textbook (1979 and 1983) provided a differently oriented synthesis of statistical methods aimed at all fields of ecology. The authors presented the mathematical bases of the methods of data analysis, and illustrated these methods with easy-to-compute numerical examples and real-data ecological applications drawn from the published literature.

Contribution of the Legendre brothers

In May 1975, a dozen or so ecologists, mostly marine, sat during three days in a classroom on the second floor of a historical building of the *Station marine de Villefranche-sur-Mer* (Université Paris 6, France), a few meters away from the Mediterranean shore, to discuss developments concerning a new trend in the ecological literature: the statistical analysis of multivariate ecological data. The meeting was called "Séminaire de mathématiques appliquées à l'océanographie biologiques" and had a marine ecology orientation.

Because they had both worked in data analysis, Louis Legendre (oceanographer, Université Laval, Canada) and Pierre Legendre (community ecologist, Université du Québec à Montréal, Canada) had been independently invited to participate in the seminar, where they contributed several presentations. On the evening of the closing day of the meeting, sitting at the terrace of a restaurant with view on the harbor, Louis and Pierre Legendre wrote, on a paper place mat, a list of subjects, which was to become the table of contents of a book about the new subdiscipline of ecology that had been discussed during the seminar. They published the first edition of the book in 1979, in French, under the title "*Écologie numérique*".

History of publication of the Numerical Ecology *textbook*

• The first edition of "*Écologie numérique*" (in French) was published in 1979 by Masson, in Paris, and Presses de l'Université du Québec in Québec City (two volumes, 473 pages in total).

• The work was translated into English, under the supervision of the two authors, and published in 1983 by Elsevier Scientific Co. in Amsterdam under the title "*Numerical Ecology*" (435 pages).

• A second French edition, revised and augmented, was published in 1984 by the two original publishing houses (2 volumes, 618 pages in total).

• During the 1980's, community ecologists started to study species-environment relationships thanks to the computer package Canoco made available by Cajo ter Braak. During the 1990's, they became aware of the importance of spatial structures to understand the spatial variation of community composition (Levin 1992, Legendre 1993). The second English edition of "*Numerical Ecology*" was published in 1998 by Elsevier (868 pages). It mostly focused on modeling the multivariate structure of community composition data. It incorporated a whole chapter on canonical ordination and one on spatial analysis, and described the partitioning of the variation of community composition data into spatial end environmental components, a method, now very popular, that had been proposed by Borcard et al. (1992).

• The years 2000 were marked by the development of multiscale variation partitioning, a development initiated by Borcard and Legendre (2002), and the progress of methods for the analysis of beta diversity. The third English edition of "*Numerical Ecology*", published in 2012 (1006 pages), featured a new chapter on multiscale spatial eigenfunction analysis, as well as substantial additions to most other chapters.

In the meantime, a companion book, "*Numerical Ecology with R*", had been written by Daniel Borcard, François Gillet and Pierre Legendre and published by Springer Science in 2011 in the *Use R!* book series. The book contained detailed accounts of the computation of the numerical ecology methods of analysis using R packages. It was based on the major developments of R packages for ecologists since the year 2000, produced by various groups and their collaborators around the world, including the packages vegan (2001) and ade4 (2002). The list of R packages used in the various chapters occupies several pages at the end of the book. A second edition of the R book will be published by Springer in early 2018. The data sets used in the book and the scripts of all analyses are freely available on a Web page cited at the end of this article. The R book was translated to Chinese by Jiangshan Lai (Institute of Botany, Chinese Academy of Science) and published in 2014 by Higher Education Press (Beijing).

In February 2018, the various editions of the "*Numerical ecology*" textbook had been cited more than 18,000 times in the scientific literature and the "*Numerical Ecology with R*" book more than 1600 times.

Important papers across the years

Users of numerical methods and graduate students often wonder where the basic ideas of the methods we are routinely using come from and how they were developed. Here is a selection of papers that have changed the way ecologists analyze multivariate data during the past 50 years and the teaching of numerical ecology to graduate students in universities. The following list is by no means exhaustive.

The years 1960 and 1970 — Development of redundancy analysis (RDA); Rao (1964) called the method "principal components of instrumental variables"; Wollenberg (1977) called it "redundancy analysis". Principal coordinate analysis: Gower (1966). The concepts of alpha-beta-gamma diversity: Whittaker (1972). Time-constrained clustering: Gordon and Birks (1972, 1974).

1980–1989 — Spatially-constrained clustering: Lefkovitch (1978); Legendre and Legendre (1984b). Metric and Euclidean properties of dissimilarity coefficients: Gower and Legendre (1986). Canonical correspondence analysis (CCA): ter Braak (1986, 1987a,b). Spatial analysis as a tool for community ecologists: Legendre and Fortin (1989).

1990–1999 — The method of variation partitioning: Borcard et al. (1992). Spatial autocorrelation, a new paradigm for ecology: Levin (1992), Legendre (1993). Co-inertia analysis: Doledec and Chessel (1994). Indicator species analysis: Dufrêne and Legendre (1997). RLQ analysis: Dolédec et al. (1996). Fourth-corner analysis: Legendre et al. (1997), Dray and Legendre (2008), Dray et al. (2014). Distance-based redundancy analysis (dbRDA): Legendre and Anderson (1999).

2000–2009 — Transformations for community composition data used in linear ordination, Legendre and Gallagher (2001), leading to transformation-based PCA (tbPCA) and transformation-based RDA (tbRDA). Spatial eigenfunction analysis – Moran's eigenvector maps (MEM): Borcard and Legendre (2002), Dray et al. (2006); asymmetric eigenvector maps (AEM): Blanchet et al. (2008). Concordance analysis of species associations: Legendre (2005). The rationale for estimation of beta diversity by the variance of the community composition data table, Var($\mathbf{Y}$): Legendre et al. (2005). Improving indicator species analysis: De Cáceres and Legendre (2009), De Cáceres et al. (2010).

2010 to present — Should the Mantel test be used in spatial analysis? Legendre and Fortin (2010), Legendre et al. (2015). Testing the space-time interaction in community surveys: Legendre et al. (2010). Test of significance of the canonical axes in RDA: Legendre et al. (2011). Partitioning beta diversity: Legendre and De Cáceres (2012), Legendre (2014). Temporal and space-time analysis of beta diversity: Legendre and Gauthier (2014). Study of temporal beta diversity: Legendre and Salvat (2015). Multiscale codependence analysis (MCA), which quantifies the joint spatial distribution of a pair of variables at different spatial scales (Guénard et al. 2010), was generalized to handle multivariate response data (Guénard and Legendre 2018).

Workshops

On 3–11 June 1986, Pierre and Louis Legendre, assisted by Marie-Josée Fortin (now Professor at University of Toronto), organized a NATO Advanced Study Workshop on Numerical Ecology at the *Station biologique de Roscoff* in France. Methods of data analysis were presented by statisticians and methodologists, followed by discussions of their application to ecological problems by working groups of ecologists. A book of Proceedings was published after the workshop (Legendre and Legendre 1987).

On 26–28 May 2008, a workshop entitled Spatial Ecological Data Analysis with R (SEDAR) was held at *Université Claude Bernard* in Lyon. It had been organized by Stéphane Dray to coordinate efforts among researchers developing the spatial analysis of ecological data and make plans for the future. One of the results of this workshop was a new R package, adespatial, for spatial and time-series analysis of community data. Written under the direction of Stéphane Dray, adespatial appeared on CRAN on 6 June 2016. New functions are still being added to this package.

On 6–7 October 2016, a workshop organized by Pedro Peres-Neto (Concordia University) and Marie-Josée Fortin (University of Toronto) was convened at Concordia University in Montreal. Twenty participants discussed future developments of the field. Following the meeting, one of the participants, Prof. Vladimir Makarenkov, computed a network describing the scientific collaborations that produced the wealth of references to numerical methods included in the 2012 edition of the "*Numerical Ecology*" book. Although the list of references at the end of a textbook is admittedly biased in favor of its authors, this network (Figure 1) illustrates the fact that the development of data analysis methods for ecologists is the result of a broad and fruitful collaboration among many scientists.

Developments in progress

Community ecology

One of the primary concerns of community ecology nowadays is to test hypotheses about the processes that generate and maintain biodiversity in ecosystems, in particular beta diversity (i.e. the spatial variation in community composition among sites) in a region, through neutral processes, abiotic environmental species filters and biotic interaction filters. Numerical ecology develops and provides the statistical methods to test such ecological hypotheses.

Methodological developments on which researchers are presently working include the following topics:

• Beta diversity analysis of spatially distributed genetic, molecular and trait data. This work extends the concept of beta diversity, which was originally defined as the spatial variation of community composition data, to other types of biodiversity data.

• Beta diversity analyses across temporal and space-time surveys. Comparison of two and multiple surveys across time. Identification of the processes that cause changes in community composition for species, genetic, molecular, and trait data.

• New advances in spatial modeling by spatial eigenfunction analysis. Translation of landscape resistance networks into spatial eigenfunctions.

• Paleoecological analysis: new advances in time-constrained clustering and other methods for modeling abrupt and changes in multi-species paleoecological data series.

• Three-table analysis. The basic methods, called RLQ analysis (Dolédec et al. 1996) and fourth-corner analysis (Legendre et al. 1997), test hypotheses of relationships between species traits and environmental characteristics of the sites mediated by the observed site-by-species data matrix. The two methods were unified by Dray and Legendre (2008) and Dray et al. (2014). Future developments, recently published or under discussion, involve other characteristics of the species, e.g. their phylogeny, and other characteristics of the sites, e.g. their spatial structure. Examples of such extensions are given in Dray et al. (2014). Mathematical extensions of the method are also considered.

• Analysis of multi-species community data using multivariate generalized linear mixed models (GLMM) or a latent variable model (LVM) that combines GLM with Markov Chain Monte Carlo (MCMC) methods (Warton et al. 2015), as in the boral R package (Hui 2017).

• Analysis of ecological networks.

The methods developed for community ecology can be transferred to other fields where empirical research is also concerned with frequency data, namely: gene frequencies, molecular data (including those used in microbiology nowadays), as well as biological and behavioral trait analysis.

Software development

In most cases nowadays, new methods of analysis are published accompanied by software. For decades, Fortran, then C dominated the programming environment. An important program, which implemented a variety of simple and canonical ordination methods and allowed researchers to implement them, is Canoco; version 2.1 became available in 1988 (ter Braak 1988).

R is a free software environment for statistical computing and graphics distributed on the CRAN Web site. The first stable version, R 1.0, appeared on CRAN on 29 February 2000. The R language is in fashion at the moment, and it is likely to be around for quite some time, given that the research community has produced thousands of packages, each containing from a few to hundreds of functions for data analysis. More than 12,000 packages are presently distributed on the CRAN site, in addition to the many other packages and functions available on individual researchers' Web pages or found in appendices of methodological papers. R may eventually be replaced by other software development environments, or complemented by other more specialized programming and computing environments. The future will tell, but for sure, ecologists will keep computing.

Conclusion

As pressing new ecological questions emerge in the world, ecologists and methodologists will keep developing methods of data analysis to answer these questions using multivariate data and enrich the methodological framework of numerical ecology.

Acknowledgments

Many thanks to Louis Legendre (Université Paris 6, France) for comments on a first draft of this article, and to Vladimir Makarenkov (Université du Québec à Montréal, Canada) who computed the network of scientific collaborations shown in Figure 1.

References

Blanchet, F. G., P. Legendre and D. Borcard. (2008). Modelling directional spatial processes in ecological data. *Ecological Modelling* **215**, 325–336.

Borcard, D., Gillet, F. and Legendre, P. (2011). *Numerical ecology with R*. Use R! series, New York: Springer Science.

Borcard, D., Gillet, F. and Legendre, P. (2014). *Numerical ecology with R, Chinese edition* (translation: J. Lai, Institute of Botany, Chinese Academy of Sciences). Beijing: Higher Education Press.

Borcard, D. and Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling* **153**, 51–68.

Borcard, D., Legendre, P. and Drapeau, P. (1992). Partialling out the spatial component of ecological variation. *Ecology* **73**, 1045–1055.

De Cáceres, M., and P. Legendre. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology* **90**, 3566–3574.

De Cáceres, M., P. Legendre and M. Moretti. (2010). Improving indicator species analysis by combining groups of sites. *Oikos* **119**, 1674–1684.

Dolédec, S. and D. Chessel. (1994). Co-inertia analysis: an alternative method for studying speciesenvironment relationships. *Freshwater Biology* **31**, 277–294.

Dolédec, S., Chessel, D., ter Braak, C. J. F. and Champely, S. (1996). Matching species traits to environmental variables: a new three-table ordination method. *Environmental and Ecological Statistics* **3**, 143–166.

Dray, S., Choler, P., Dolédec, S., Peres-Neto, P. R., Thuillier, W., Pavoine, S. and ter Braak, C. J. F. (2014). Combining the fourth-corner and the RLQ methods for assessing trait responses to environmental variation. *Ecology* **95**, 14–21.

Dray, S. and Legendre, P. (2008). Testing the species traits-environment relationships: the fourth-corner problem revisited. *Ecology* **89**, 3400–3412.

Dray, S., P. Legendre and P. R. Peres-Neto. (2006). Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* **196**, 483–493.

Dufrêne, M. and P. Legendre. (1997). Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs* **67**, 345–366.

Goodall, D. W. (1954). Objective methods for the classification of vegetation. III. An essay in the use of factor analysis. *Australian Journal of Botany* **2**, 304–324.

Gordon, A. D. and H. J. B. Birks. (1972). Numerical methods in Quaternary palaeoecology. I. Zonation of pollen diagrams. *New Phytologist* **71**, 961–979.

Gordon, A. D. and H. J. B. Birks. (1974). Numerical methods in Quaternary palaeoecology. II. Comparison of pollen diagrams. *New Phytologist* **73**, 221–249.

Gower, J. C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**, 325–338.

Gower, J. C. and Legendre, P. (1986). Metric and Euclidean properties of dissimilarity coefficients. *Journal of Classification* **3**, 5–48.

Green, R. H. (1979). *Sampling design and statistical methods for environmental biologists*. New York: John Wiley & Sons.

Guénard, G. and Legendre, P. (2018). Bringing multivariate support to multiscale codependence analysis: assessing the drivers of community structure across spatial scales. *Methods in Ecology and Evolution* **9**, 292–304.

Guénard, G., Legendre, P., Boisclair, D. and Bilodeau, M. (2010). Multiscale codependence analysis: an integrated approach to analyze relationships across scales. *Ecology* **91**, 2952–2964.

Hui, F. K. C. (2017). *boral: Bayesian ordination and regression analysis.* R package version 1.3.1. https://CRAN.R-project.org/package=boral

Jaccard, P. (1900). Contribution au problème de l'immigration post-glaciaire de la flore alpine. *Bulletin de la Société Vaudoise des Sciences Naturelles* **36**, 87–130.

Lefkovitch, L. P. (1978). Cluster generation and grouping using mathematical programming. *Mathematical Biosciences* **41**, 91–110.

Legendre, L. and Legendre, P. (1979). *Écologie numérique. Tome 1: Le traitement multiple des données écologiques. Tome 2: La structure des données écologiques.* Paris: Masson and Québec: Presses de l'Université du Québec.

Legendre, L. and Legendre, P. (1983). *Numerical ecology*. Developments in Environmental Modelling, Vol. 3. Amsterdam: Elsevier Scientific Publ. Co.

Legendre, L. and Legendre, P. (1984a). *Écologie numérique, deuxième édition. Tome 1: Le traitement multiple des données écologiques. Tome 2: La structure des données écologiques.* Paris: Masson and Québec: Presses de l'Université du Québec.

Legendre, P. (1993). Spatial autocorrelation: trouble or new paradigm? *Ecology* **74**, 1659–1673.

Legendre, P. (2005). Species associations: the Kendall coefficient of concordance revisited. *Journal of Agricultural Biological and Environmental Statistics* **10**, 226–245.

Legendre, P. (2014). Interpreting the replacement and richness difference components of beta diversity. *Global Ecology and Biogeography* **23**, 1324–1334.

Legendre, P. and M. J. Anderson. (1999). Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs* **69**, 1–24.

Legendre, P., D. Borcard and P. R. Peres-Neto. (2005). Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs* **75**, 435–450.

Legendre, P. and M. De Cáceres. (2013). Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters* **16**, 951–963.

Legendre, P., M. De Cáceres, and D. Borcard. (2010). Community surveys through space and time: testing the space-time interaction in the absence of replication. *Ecology* **91**, 262–272.

Legendre, P. and M.-J. Fortin. (1989). Spatial pattern and ecological analysis. *Vegetatio* **80**, 107–138.

Legendre, P. and M.-J. Fortin. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources* **10**, 831–844.

Legendre, P., M.-J. Fortin and D. Borcard. (2015). Should the Mantel test be used in spatial analysis? *Methods in Ecology and Evolution* **6**, 1239–1247.

Legendre, P. and E. D. Gallagher. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**, 271–280.

Legendre, P., Galzin, R. and Harmelin-Vivien, M. L. (1997). Relating behavior to habitat: solutions to the fourth-corner problem. *Ecology* **78**, 547–562.

Legendre, P. and O. Gauthier. (2014). Statistical methods for temporal and space-time analysis of community composition data. *Proceedings of the Royal Society B* **281**, 20132728.

Legendre, P. and Legendre, L. (eds.) (1987). *Developments in numerical ecology*. NATO ASI series, Vol. G-14. Berlin: Springer-Verlag.

Legendre, P. and Legendre, L. (1998). *Numerical ecology, 2nd English edition*. Developments in Environmental Modelling, Vol. 20. Amsterdam: Elsevier Science BV.

Legendre, P. and Legendre, L. (2012). *Numerical ecology, 3rd English edition*. Developments in Environmental Modelling, Vol. 24. Amsterdam: Elsevier Science BV.

Legendre, P. and V. Legendre. (1984b). Postglacial dispersal of freshwater fishes in the Québec peninsula. *Canadian Journal of Fisheries and Aquatic Sciences* **41**, 1781–1802.

Legendre, P., J. Oksanen and C. J. F. ter Braak. (2011). Testing the significance of canonical axes in redundancy analysis. *Methods in Ecology and Evolution* **2**, 269–277.

Legendre, P. and B. Salvat. (2015). Thirty-year recovery of mollusc communities after nuclear experimentations on Fangataufa atoll (Tuamotu, French Polynesia). *Proceedings of the Royal Society B* **282**, 20150750.

Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology* **73**, 1943–1967.

Orlóci, L. (1975). *Multivariate analysis in vegetation research*. The Hague: Dr. W. Junk B. V.

Pielou, E. C. (1969). *An introduction to mathematical ecology*. New York: John Wiley & Sons.

Rao, C. R. (1964). The use and interpretation of principal component analysis in applied research. *Sankhyā – The Indian Journal of Statistics Series A* **26**, 329–358.

Sneath, P. H. A. and Sokal, R. R. (1973). *Numerical taxonomy – The principles and practice of numerical classification*. San Francisco: W. H. Freeman.

Sokal, R. R. and Sneath, P. H. A. (1963). *Principles of numerical taxonomy*. San Francisco: W. H. Freeman.

ter Braak, C. J. F. (1986). Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* **67**, 1167–1179.

ter Braak, C. J. F. (1987a). The analysis of vegetation-environment relationships by canonical correspondence analysis. *Vegetatio* **69**, 69–77.

ter Braak, C. J. F. (1987b). Ordination. In Jongman, R. H. G., C. J. F. ter Braak & O. F. R. van Tongeren (eds.) *Data analysis in community and landscape ecology*, pp. 91–173. Wageningen, The Netherlands: Pudoc. Reissued in 1995 by Cambridge, England: Cambridge University Press.

ter Braak, C. J. F. (1988). CANOCO – an extension of DECORANA to analyze species-environment relationships. *Vegetatio* **75**, 159–160.

ter Braak, C. J. F. and Šmilauer, P. (2002). *CANOCO reference manual and CanoDraw for Windows user's guide – Software for canonical community ordination (version 4.5)*. Ithaca: Microcomputer Power.

van den Wollenberg, A. L. (1977). Redundancy analysis. An alternative for canonical correlation analysis. *Psychometrika* **42**, 207–219.

Warton, D. I., Blanchet, F. G., O'Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C. and Hui, F. K. C. 2015. So many variables: joint modeling in community ecology. *Trends in Ecology and Evolution* **30**, 766–779.

Whittaker, R. H. (1969). New concepts of kingdoms or organisms. *Science* **163**, 150–160.

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon* **21**, 213–251.

Relevant Websites

Numerical ecology, n.d., www.numericalecology.com — Legendre Numerical ecology page.

Numerical Ecology with R (NEwR) books, n.d., http://adn.biol.umontreal.ca/_numericalecology/numecolR/ — NEwR book.

Numerical taxonomy, n.d., https://en.wikipedia.org/wiki/Numerical_taxonomy — Numerical taxonomy.

Paul Jaccard, n.d., https://en.wikipedia.org/wiki/Paul_Jaccard — Paul Jaccard.

David Goodall, n.d., https://en.wikipedia.org/wiki/David_W._Goodall — David Goodall.

John C. Gower, n.d., http://onlinelibrary.wiley.com/doi/10.1111/insr.12094/pdf— A conversation with John C. Gower.

Robert H. Whittaker, n.d., https://en.wikipedia.org/wiki/Robert_Whittaker — Robert H. Whittaker.

========

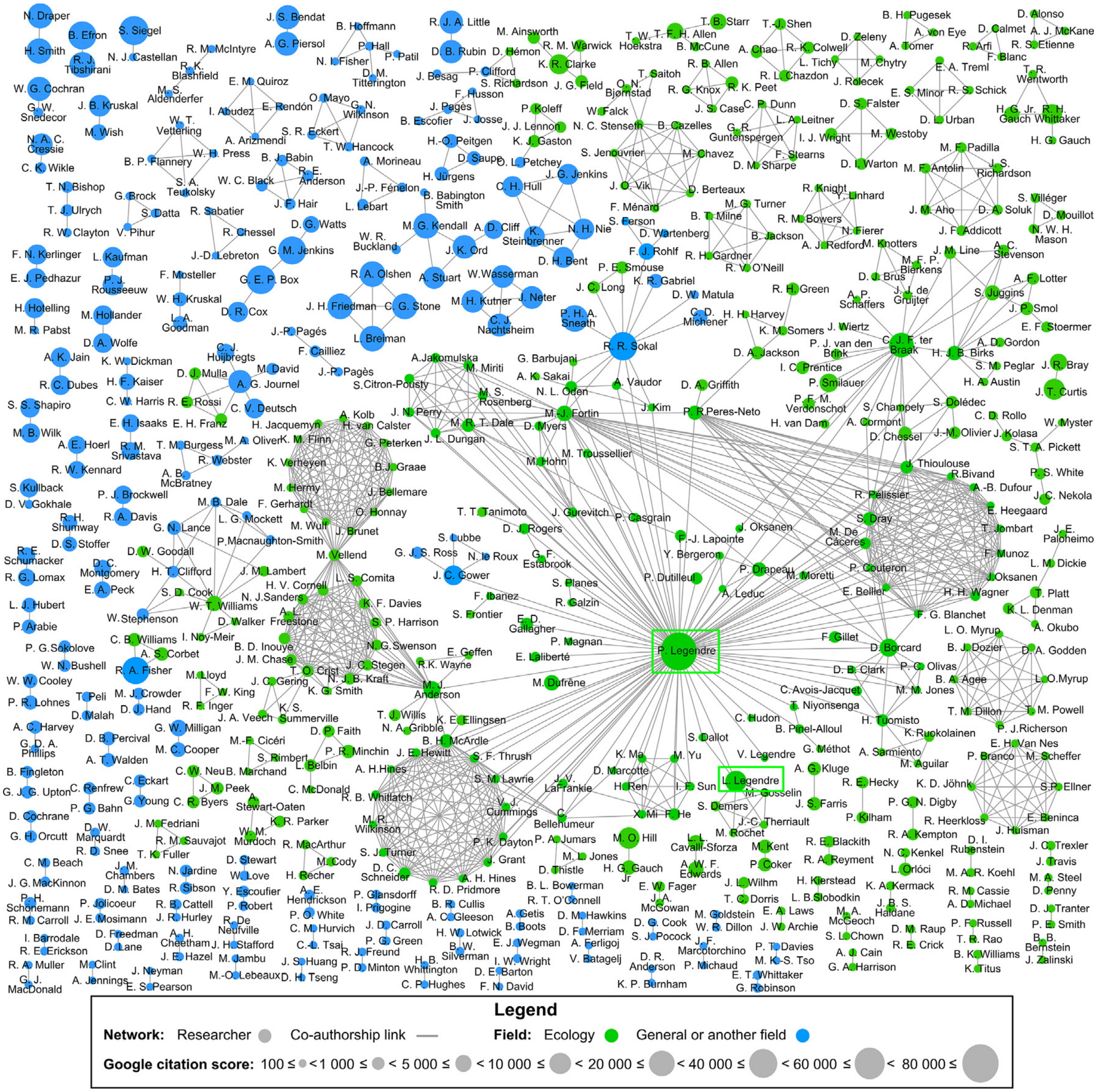

Figure 1. This figure describes the network of collaborators who produced the references in the 2012 edition of the *Numerical Ecology* book. Single-author references were excluded. Network computed and kindly provided for use in this article by Prof. V. Makarenkov, Department of Computer Sciences, Université du Québec à Montréal.

========