

# Tree diversity analysis

A manual and software for common statistical  
methods for ecological and biodiversity studies

**Roeland Kindt and Richard Coe**

World Agroforestry Centre, Nairobi, Kenya



World Agroforestry Centre  
TRANSFORMING LIVES AND LANDSCAPES



World Agroforestry Centre  
TRANSFORMING LIVES AND LANDSCAPES

**Suggested citation:** Kindt R and Coe R. 2005. *Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies*. Nairobi: World Agroforestry Centre (ICRAF).

Published by the World Agroforestry Centre  
United Nations Avenue  
PO Box 30677, GPO 00100  
Nairobi, Kenya  
Tel: +254(0)20 7224000, via USA +1 650 833 6645  
Fax: +254(0)20 7224001, via USA +1 650 833 6646  
Email: [icraf@cgiar.org](mailto:icraf@cgiar.org)  
Internet: [www.worldagroforestry.org](http://www.worldagroforestry.org)

© World Agroforestry Centre 2005  
ISBN: 92 9059 179 X

Design and Layout: K. Vanhoutte  
Printed in Kenya

This publication may be quoted or reproduced without charge, provided the source is acknowledged. Permission for resale or other commercial purposes may be granted under select circumstances by the Head of the Training Unit of the World Agroforestry Centre. Proceeds of the sale will be used for printing the next edition of this book.

# Contents

Acknowledgements	iv
Introduction	v
Overview of methods described in this manual	vi
<b>Chapter 1</b> Sampling	1
<b>Chapter 2</b> Data preparation	19
<b>Chapter 3</b> Doing biodiversity analysis with Biodiversity.R	31
<b>Chapter 4</b> Analysis of species richness	39
<b>Chapter 5</b> Analysis of diversity	55
<b>Chapter 6</b> Analysis of counts of trees	71
<b>Chapter 7</b> Analysis of presence or absence of species	103
<b>Chapter 8</b> Analysis of differences in species composition	123
<b>Chapter 9</b> Analysis of ecological distance by clustering	139
<b>Chapter 10</b> Analysis of ecological distance by ordination	153

# Acknowledgements

We warmly thank all that provided inputs that lead to improvement of this manual. We especially appreciate the comments received during training sessions with draft versions of this manual and the accompanying software in Kenya, Uganda and Mali. We are equally grateful to the thoughtful reviews by Dr Simoneta Negrete-Yankelevich (Instituto de Ecología, Mexico) and Dr Robert Burn (Reading University, UK) of the draft version of this manual, and to Hillary Kipruto for help in editing of this manual.

We highly appreciate the support of the Programme for Cooperation with International Institutes (SII), Education and Development Division of the Netherlands' Ministry of Foreign Affairs, and VVOB (The Flemish Association for Development Assistance and Technical Cooperation, Flanders, Belgium) for funding

the development for this manual. We also thank VVOB for seconding Roeland Kindt to the World Agroforestry Centre (ICRAF).

This tree diversity analysis manual was inspired by research, development and extension activities that were initiated by ICRAF on tree and landscape diversification. We want to acknowledge the various donor agencies that have funded these activities, especially VVOB, DFID, USAID and EU.

We are grateful for the developers of the R Software for providing a free and powerful statistical package that allowed development of Biodiversity.R. We also want to give special thanks to Jari Oksanen for developing the vegan package and John Fox for developing the Rcmdr package, which are key packages that are used by Biodiversity.R.

# Introduction

This manual was prepared during training events held in East- and West-Africa on the analysis of tree diversity data. These training events targeted data analysis of tree diversity data that were collected by scientists of the World Agroforestry Centre (ICRAF) and collaborating institutions. Typically, data were collected on the tree species composition of quadrats or farms. At the same time, explanatory variables such as land use and household characteristics were collected. Various hypotheses on the influence of explanatory variables on tree diversity can be tested with such datasets. Although the manual was developed during research on tree diversity on farms in Africa, the statistical methods can be used for a wider range of organisms, for different hierarchical levels of biodiversity, and for a wider range of environments.

These materials were compiled as a second-generation development of the Biodiversity Analysis Package, a CD-ROM compiled by Roeland Kindt with resources and guidelines for the analysis of ecological and biodiversity information. Whereas the Biodiversity Analysis Package provided a range of tools for different types of analysis, this manual is accompanied by a new tool (Biodiversity.R) that offers a single software environment for all the analyses that are described in this manual. This does not mean that Biodiversity.R is the only recommended package for a particular type of analysis, but it offers the advantage for training purposes that users only need to be introduced to one software package for statistically sound analysis of biodiversity data.

It is never possible to produce a guide to all the methods that will be needed for analysis of biodiversity data. Data analysis questions are continually advancing, requiring ever changing data collection and analysis methods. This manual focuses on the analysis of species survey data. We describe a number of methods that can be used to analyse hypotheses that are frequently important in biodiversity research. These are not the only methods that can be used to analyse these hypotheses, and other methods will be needed when the focus of the biodiversity research is different.

Effective data analysis requires imagination and creativity. However, it also requires familiarity with basic concepts, and an ability to use a set of standard tools. This manual aims to provide that. It also points the user to other resources that develop ideas further.

Effective data analysis also requires a sound and up to date understanding of the science behind the investigation. Data analysis requires clear objectives and hypotheses to investigate. These have to be based on, and push forward, current understanding. We have not attempted to link the methods described here to the rapidly changing science of biodiversity and community ecology. Data analysis does not end with production of statistical results. Those results have to be interpreted in the light of other information about the problem. We can not, therefore, discuss fully the interpretation of the statistical results, or the further statistical analyses they may lead to.

# Overview of methods described in this manual

On the following page, a general diagram is provided that describes the data analysis questions that you can ask when analysing biodiversity based on the methodologies that are provided in this manual. Each question is discussed in further detail in the respective chapter. The arrows indicate the types of information that are used in each method. All information is derived from either the species data or the environmental data of the sites. Chapter 2 describes the species and environmental data matrices in greater detail.

Some methods only use information on species. These methods are depicted on the left-hand side of the diagram. They are based on biodiversity statistics that can be used to compare the levels of biodiversity between sites, or to analyse how similar sites are in species composition.

The other methods use information on both species and the environmental variables of the sites. These methods are shown on the right-hand side of the diagram. These methods provide insight into the influence of environmental

variables on biodiversity. The analysis methods can reveal how much of the pattern in species diversity can be explained by the influence of the environmental variables. Knowing how much of a pattern is explained will especially be useful if the research was conducted to arrive at options for better management of biodiversity. Note that in this context, ‘environmental variables’ can include characteristics of the social and economic environment, not only the biophysical environment.

You may have noticed that Chapter 3 did not feature in the diagram. The reason is that this chapter describes how the Biodiversity.R software can be installed and used to conduct all the analyses described in the manual, whereas you may choose to conduct the analysis with different software. For this reason, the commands and menu options for doing the analysis in Biodiversity.R are separated from the descriptions of the methods, and placed at the end of each chapter.

