

Introduction

Here, we present the main features of the book: its aims, structure, content, terminology, readership, supporting material, expected pre-requisites, and how it differs from other books already available. The structure of the book follows the main modeling steps, so we recommend reading the sections about the book's structure and content before reading the other parts.

Cambridge University Press
978-0-521-76513-8 — Habitat Suitability and Distribution Models
Antoine Guisan , Wilfried Thuiller , Niklaus E. Zimmermann
Excerpt
[More Information](#)

1 · *General Content of the Book*

1.1 What Is This Book About?

This book aims to present a class of models used to quantify taxa–environment relationships (i.e. taxa’s habitat suitability or ecological niche), and how these can be used to predict the geographical distribution of taxa. This book is also about understanding the drivers behind those distributions. This book mainly focuses on modeling the distribution of single species, yet the concept of a species itself is increasingly being called into question (especially in some groups). It is therefore important to note that these approaches can also be used to model other biological entities, for instance, genes, haplotypes, or clades, within or across species. Communities, ecosystems, or biomes can also be modeled using these approaches, directly as fixed entities or by assembling individual species predictions. Further examples of modeling these new entities are rapidly being added to the literature. This is why we prefer to use the term “habitat suitability modeling” (see below), as it does not solely refer to modeling species, but more generally modeling suitable conditions for any given biological entity. For the sake of simplicity, throughout this book we will remain at species level for all examples and explanations, but most of the techniques, metrics, and approaches presented can be directly applied to almost any other modeling entity. We also mainly discuss distributions in terms of simple occurrences or presence–absence of species, as these are the most commonly available data, but much of the reasoning can be extended to use abundances when available, provided the appropriate statistical methods are also used.

1.2 How Is the Book Structured?

The book is made of seven main parts, including this introduction. The first five each address one or several of the modeling steps described in Chapter 2 (see also Guisan and Zimmermann, 2000; Austin, 2002),

4 · General Content of the Book

Part VI introduces the tools and datasets used in this book, with two developed examples, and Part VII provides some conclusions and perspectives. The seven main parts are as follows:

- Part I (Chapters 2–5) presents the theoretical framework on which these models are built, i.e. what drives species distributions and how to use it when building models.
- Part II (Chapters 6–8) presents the data necessary to build HSMs, i.e. what we would need theoretically, what is available or what should be sampled, on what scale, how to prepare or sample them, and what associated problems might occur.
- Part III (Chapters 9–14) presents a representative sample of the statistical techniques commonly used to model habitat suitability, how they work, and how to parameterize them in R.
- Part IV (Chapters 15–16) presents ways of evaluating the models fitted in Part III, presenting different approaches to keep some data “independent”, and the metrics used to compare predictions with observations.
- Part V (Chapter 17) presents ways of deriving spatial predictions and projections in time and space from the fitted models, with the associated uncertainties and additional assumptions specific to model transferability.
- Part VI (Chapters 18–19) presents the data and tools used in this book, as well as two developed case studies, all freely available and explained. Chapter 18 presents an overview of all data and tools used throughout the book, but because many of the model-fitting examples are developed with the `biomod2` R package. Chapter 19 develops two illustrative case studies specifically with `biomod2`, and in particular ensemble models using a variety of techniques.
- Part VII (Chapter 20) provides a short conclusion and some perspectives. HSM research has become a very active field, with fast internal dynamics, and there is still much to be done to make it a more mature field. We suggest some directions that we think are interesting and worth pursuing in this regard.

1.3 Why Write a Textbook with R Examples?

The book is conceived both as a textbook to teach the main modeling steps, but also as a practical guide to run these models in R.¹ The choice of R is deliberate here, as it is freeware available on all platforms, and is now by far

¹ www.r-project.org

the most commonly used statistical software for building HSMs, and more generally, for running ecological analyses of niches and distributions. The book is richly illustrated with examples and guidelines for building HSMs in R from sampling design and data construction to model building, testing, projections and interpretation. We have attempted to build all examples from publicly available information, which can be partly retrieved directly *from within* R (official packages), or at least from internet downloads.

1.4 What Is This Book Not About?

This book is not meant to be an exhaustive review of all the recent developments in the field of habitat suitability modeling *sensu lato* (see terminology issues below), which is best found in review papers or other work (e.g. Franklin, 2010a). It is not intended to provide detailed explanations of standard statistical methods (best found in statistical textbooks), or of standard geographic analyses (best found in geography and geographic information systems textbooks), nor does it provide a detailed review of fundamental knowledge in ecology and biogeography (best found in the relevant textbooks).

1.5 Why Was This Book Needed?

The twenty-first century is experiencing a major biodiversity crisis and our planet now faces the risk of a sixth – human-driven – major species extinction. Major threats currently include pollution, habitat destruction, modification and fragmentation, whereas future threats also include biological invasions and climate change. Such threats to species, biodiversity, and ecosystems are usually estimated using multi-scale assessments based on International Union for Conservation of Nature (IUCN) geographic criteria (e.g. geographic extent and/or area of occupancy). Therefore, sound approaches and methods are needed to forecast the future distribution of life on Earth. At the same time, science is progressing fast. The last century was paralleled by major technological advances, especially in bioinformatics and biomathematics. Not only have we improved our knowledge and understanding of the living world, and particularly what drives species distributions, but we have also improved our ability to model and predict it, with applications in evolutionary biology, biogeography and conservation biology. As a result, interest in predictive HSMs of organisms has grown dramatically during the twenty-first century. The reasons for this interest are twofold. First, the last two decades witnessed an exponential increase

6 · General Content of the Book

in computing power, with the advent of geographic information systems and remote-sensing technologies leading to the development of high-resolution environmental datasets, which together gave greater freedom to analyse, model, and predict a large number of taxa over various spatial scales (from the very local to global). Second, boosted by these numerical developments and with no other approaches applicable to large number of species in many different regions or over large geographic extents, HSMs have emerged as a vital tool in applied ecological and environmental sciences. They have proved particularly useful for evaluating the potential impact of global anthropogenic environmental changes on biodiversity and ecosystems (e.g. climate change, biological invasions, habitat destruction; (Johnson and Gillingham, 2005; Rodríguez et al., 2007; Franklin, 2010b; Schwartz, 2012; Guisan et al., 2013; Tulloch et al., 2016), as demonstrated by their key role in the first assessments by the Intergovernmental Panel on Biodiversity and Ecosystem Services (IPBES²). The strong societal need and political pressure to address these environmental problems have consolidated the role of HSMs in conservation sciences, but also require the formalization of the modeling methods used for predictions. Consequently, HSMs have also been increasingly commonly used in fundamental ecological and evolutionary sciences, to improve our understanding of species' ecological niches or to test biogeographic hypotheses (see more examples in Scott et al., 2002; Guisan and Thuiller, 2005; Franklin, 2010a; Peterson et al., 2011). This book can thus be seen as an additional contribution to establish modeling standards in basic and applied HSM research.

1.6 Who Is This Book For?

We have written this book for advanced students for use as textbook in university classes, at the level of third-year bachelor, masters, and doctoral studies, or for any scientist who wants to familiarize themselves with the principles and methods of habitat suitability modeling and species distribution predictions.

1.7 Where Can I Find Supporting Material?

To facilitate the learning process and transfer of technology to real case problems, supporting material (scripts, data, and manuals) is available on the companion website to the book.³

² www.ipbes.org

³ www.unil.ch/hsdm

1.8 What Are Readers Assumed to Know Already?

We assume that readers have basic understanding of ecological theory, biogeography, and macroecology, in biostatistics – such as multivariate statistics, inference tests, bootstrapping, etc. – and in GIS and spatial analyses – such as geodata handling, spatial interpolation, and spatial autocorrelation. We also assume that readers have a grasp of the R environment.⁴

1.9 How Does This Book Differ From Previous Ones?

HSMs have become very popular and have undergone tremendous development in the past 10 to 15 years, yet there are still very few books on the subject. Our book is an ideal complement to the two most recent works published: Franklin (2010a) and Peterson et al. (2011). Franklin (2010a) is an excellent monograph on these models, but it does not include the codes explaining how to practically run the different analyses. Peterson et al. (2011) is a multi-authored book on species' ecological niches and geographic distributions that provides a series of chapters on specific topics, with some detailed examples, but no code for running the analyses. Other books are more miscellaneous, by having each chapter contributed by different authors, and a large number of authors overall (e.g. Corsi et al., 2000; Scott et al., 2002).

Our book has three unique features, namely:

1. It is a textbook which can be used to teach classes in universities, technical high schools or other courses. It follows a simple sequential structure, introducing the theoretical concepts, data preparation, model building, model evaluation, and spatial predictions, which thus follows the logical successive modeling steps.
2. It is unique in providing fully developed examples and practical case studies that can be run in the R language, using the most advanced tools available in R to model habitat suitability and niches, and to predict species distributions.
3. It comes with a companion website,⁵ where additional resources can be found, regularly updated, and discussed in an online forum.

⁴ www.r-project.org

⁵ www.unil.ch/hsdm

8 · General Content of the Book

It is therefore a more practical book, which – we hope! – will allow interested students and scientists to quickly get off to a good start in the field.

1.10 What Terminology Is Used in This Book?

Many different names and acronyms are used in the literature for the same class of models (see also Glossary): habitat suitability models (HSM), the term mainly used throughout this book, but also ecological niche models (ENM), species distribution models (SDMs), habitat distribution models (HDM), climate-envelope models (CEM), resource selection functions (RSF), and many other more minor variants. It appears that all the models in this class can be used to investigate both species niches (i.e. niche modeling, NM) and species distributions (i.e. distribution modeling, DM).

Although we recognize that all terms can be used depending on the study context, we decided to use the HSM terminology in the title and throughout the book because:

- it best reflects the basis of what all these models do: quantifying the species–environment (habitat) relationship;
- not all applications of these models need to predict the geographic distribution of the modeled entity;
- it is more generally applicable to entities other than species (e.g. genes or communities), to which the niche concept may not apply;
- it is still applicable to cases when a niche is only partially captured in a model, i.e. when an envelope of suitable habitats is modeled, but possibly not the full species niche, as typically observed in studies confined to geographic extents smaller than the full distribution range of the modeled species.

There are also issues of scale and hierarchy in the use of these different terms. As we will see in Chapters 2 and 3, the environmental niche is the envelope of all suitable habitats for a species across its whole range. All the conditions the species can withstand therefore needs to be encompassed to capture the whole species niche in a model.

The main terms used throughout the different parts are defined in the glossary at the end of the book.

PART I · Overview, Principles, Theory, and Assumptions Behind Habitat Suitability Modeling

In this first part of the book, we begin by briefly presenting the general procedure (i.e. the series of methodological steps) used to build and apply HSMs (Chapter 2). We next summarize our ecological and evolutionary understanding of the factors driving species distributions and related biogeographical theory (Chapter 3). It is by no means our intention to present an exhaustive review of all existing theories, which can best be found in textbooks (Lomolino et al., 2010; Smith and Smith, 2015), but rather to focus on the most useful concepts for HSMs. Readers familiar with the theory behind species' niches and geographic distributions may prefer to start directly with Chapter 4, where we explain the main principles of habitat suitability modeling, how predictions for individual species can be assembled to predict communities, and what the main applications of these models are. We finally present the main working assumptions that are made when fitting such models (Chapter 5; see also Part V for assumptions specifically related to projections in time and space).