The Climate change atlas for Africa of tree species prioritized for forest landscape restoration in Ethiopia

A description of methods used to develop the atlas

Roeland Kindt
Abrham Abiyu
Peter Borchardt
Ian K Dawson
Sebsebe Demissew
Ramni Jamnadass
Jens-Peter B Lillesø
Søren Moestrup
Fabio Pedercini
Jan J Wieringa
Wubalem Tadesse
Lars Graudal

WORKING PAPER 17

Working Paper 17

The Climate change atlas for Africa of tree species prioritized for forest landscape restoration in Ethiopia

A description of methods used to develop the atlas

Roeland Kindt
Abrham Abiyu
Peter Borchardt
Ian K Dawson
Sebsebe Demissew
Ramni Jamnadass
Jens-Peter B Lillesø
Søren Moestrup
Fabio Pedercini
Jan J Wieringa
Wubalem Tadesse
Lars Graudal

CIFOR-ICRAF

Working Paper 17

DOI: 10.17528/cifor-icraf/008977

Kindt R, Abiyu A, Borchardt P, Dawson IK, Demissew S, Graudal L, Jamnadass R, Lillesø J-PB, Moestrup S, Pedercini F, Wieringa JJ, Wubalem T. 2023. *The Climate change atlas for Africa of tree species prioritized for forest landscape restoration in Ethiopia: A description of methods used to develop the atlas*. Working Paper No. 17. Bogor, Indonesia; and Nairobi, Kenya: Center for International Forestry Research and World Agroforestry (CIFOR-ICRAF).

CIFOR
Jl. CIFOR, Situ Gede
Bogor Barat 16115
Indonesia
T +62 (251) 8622622
F +62 (251) 8622100
E cifor@cifor-icraf.org

ICRAF
United Nations Avenue, Gigiri
PO Box 30677, Nairobi, 00100
Kenya
T +254 (20) 7224000
F +254 (20) 7224001
E ICRAF@cifor-icraf.org

cifor-icraf.org

We would like to thank all donors who supported this work through their contributions to the CGIAR Trust Fund. A list of donors is available at: http://www.cgiar.org/about-us/our-funders/

Contents

About the authors **v**
Acknowledgements **vi**
Acronyms **vii**

1 Introduction **1**

2 Selection of species **3**

3 Predictor variables **8**

4 Future climates **12**

5 Compilation of occurrence observations **13**

6 Spatial thinning of occurrence observations **15**

7 Environmental thinning of occurrence observations **20**

8 Compilation of background observations **22**

9 Spatial folding of occurrence and background observations **26**

10 Calibration of species distribution models and generation of suitability maps **31**

11 Discrimination of areas with novel environmental conditions **38**

12 Generation of habitat change maps **39**

13 Convex hulls for an *a posteriori* distance constraining method **40**

14 Base map and map annotations used in the atlas **42**

15 Inset maps showing country distribution of Plants of the World Online **43**

References **44**
Appendices **49**
1 Synonyms and authorities for species names 49
2 Data sets accessed from GBIF 56
3 Data sets accessed from BIEN 58

List of box, tables and figures

Box

1 The online Climate change atlas for Africa of tree species prioritized for forest landscape restoration in Ethiopia. 1

Tables

1 153 tree species selected as initial candidates for species distribution modelling. 4
2 Final subset of predictor variables selected for species suitability modelling in our analysis. 9
3 Datasets of occurrence observations used in our analysis. 13
4 Results from the spatial thinning of occurrence observations. 15
5 Minimum number of occurrence observations recommended for species distribution modelling, according to selected references, ordered by number of recommended observations. 20
6 Target number of environmentally thinned occurrence observations. 21
7 Mountains peaks added to the set of background locations. 23
8 Target number of background observations 24
9 Number of occurrence and background locations used for spatial folding. 27
10 Argument settings for the BiodiversityR::ensemble.calibrate.weights function. 31
11 Evaluation statistics for the selected ensemble models used to generate species distribution maps. 33
12 Rules to classify habitat change. 39

Figures

1 Scatterplot matrix showing correlations among AriditySeason and selected predictor variables from WorldClim. 10
2 Scatterplot matrix showing correlations among AriditySeason and selected predictor variables from envirem. 11
3 Scatterplot matrix showing correlations among AriditySeason and selected predictor variables from ENVIREM and SoilGrids250. 11
4 Candidate background locations. 22
5 For each species, background locations (small, green symbols) used for modelling were restricted to a 500 km buffer around occurrence locations (large, red symbols). 24
6 Occurrence and background locations for each species were assigned to four spatial folds with a size of 500 km. 26
7 Convex hull for a hypothetical tree species with a relatively narrow range across Africa. 40
8 Convex hull for a hypothetical tree species with a wide range across Africa. 41

About the authors

Roeland Kindt, Abrham Abiyu, Ian Dawson, Lars Graudal, Ramni Jamnadass, Jens-Peter Barnekow Lillesø, Søren Moestrup and Fabio Pedercini work for World Agroforestry (ICRAF), part of the CIFOR-ICRAF partnership. Lars Graudal, Jens-Peter Barnekow Lillesø, Søren Moestrup and Fabio Pedercini also work for the University of Copenhagen, Copenhagen, Denmark, and Ian Dawson for Scotland's Rural College (SRUC), Edinburgh, Scotland. Peter Borchardt is the founder of ARBONETH – Networking for Trees in Ethiopia and manages Plant-for-Ethiopia. Sebsebe Demissew works for Addis Ababa University and Jan Wieringa at the Naturalis Biodiversity Center in Leiden. Wubalem Tadesse works for Ethiopian Forestry Development. The authors are interested in research and practical capacity building for forest landscape restoration, agroforestry and broader tree planting, and in the conservation of natural habitats.

Acknowledgements

The authors gratefully acknowledge the support of Norway's International Climate and Forest Initiative in the development of the **Climate change atlas for Africa of tree species prioritized for forest landscape restoration in Ethiopia**, through funding of the Provision of Adequate Tree Seed Portfolio in Ethiopia (PATSPO) project (http://www.worldagroforestry.org/project/provision-adequate-tree-seed-portfolio-ethiopia). We also gratefully acknowledge the support of CGIAR funding partners through the CGIAR Research Program on Forests, Trees and Agroforestry (2011–2021). In addition, our thanks to David Henry for editorial support.

Acronyms

AUC	Area Under the receiver-operator Curve
BIEN	Botanical Information and Ecology Network
BIOCLIM	Bioclimate analysis and prediction system
BMCP	Buffered Minimum Convex Polygon
BSO	Breeding Seedling (or Seed) Orchard
CMIP5	Coupled Model Intercomparison Project for the Fifth Assessment Report of the Intergovernmental Panel on Climate Change
CMIP6	Coupled Model Intercomparison Project for the Sixth Assessment Report of the Intergovernmental Panel on Climate Change
ENVIREM	Environmental Rasters for Ecological Modelling
FRC	Forestry Research Centre
GBIF	Global Biodiversity Information Facility
GCM	Global Climate Model (or General Circulation Model)
ICRAF	World Agroforestry
MESS	Multivariate Environmental Similarity Surfaces
PATSPO	Provision of Adequate Tree Seed Portfolio in Ethiopia project
PET	Potential Evapotranspiration
POWO	Plants of the World Online
RAINBIO	African RAINforest community dynamics: implications for tropical BIOdiversity conservation and climate change mitigation
RCP	Representative Concentration Pathway
SDM	Species Distribution Modelling (or Model)
SEDI	Symmetric Extremal Dependence Index
SoW-FGR	The State of the World's Forest Genetic Resources report
SSP	Shared Socioeconomic Pathway
TSS	True Skill Statistic
VIF	Variance Inflation Factor
WFO	World Flora Online

1 Introduction

This working paper describes the methods used to develop the online **Climate change atlas for Africa of tree species prioritized for forest landscape restoration in Ethiopia (**http://atlas.worldagroforestry.org/). The atlas shows the baseline and 2050s habitat distributions across Africa for 127 tree species. The purpose of the atlas is to indicate how alterations in environmental conditions caused by climate change will likely affect the locations where particular tree species can grow in Africa. This is important for planning current and future tree-planting activities, including tree-planting-based forest landscape restoration actions. The atlas will help ensure that the right species are chosen for planting in particular locations, and is an important part of the process of operationalizing **Climate Appropriate Portfolios of Tree Diversity** (Kindt et al. 2023). The atlas is part of a larger set of tools developed by CIFOR-ICRAF for tree species selection for planting purposes that can be found on the **Global Tree Knowledge Platform** (https://www.worldagroforestry.org/tree-knowledge). Further background on the atlas and why it is important is provided in Box 1.

Box 1. The online Climate change atlas for Africa of tree species prioritized for forest landscape restoration in Ethiopia.

The tree species that have been mapped in the online atlas are priorities identified through the Provision of Adequate Tree Seed Portfolio in Ethiopia (PATSPO) project (https://www.worldagroforestry.org/project/provision-adequate-tree-seed-portfolio-ethiopia). This project, now in its second phase, is developing tree seed supply capacity in Ethiopia to help reach the country's large forest landscape restoration target of 15 million hectares. The PATSPO project is describing existing tree seed sources and is planting breeding seedling (or seed) orchards (BSOs) for tree improvement; these BSOs further act as high-quality adapted seed sources. Mapping where tree species can grow under future climate helps PATSPO to plan for sustainable, appropriate tree seed supply.

Although the PATSPO project focuses on Ethiopia, the online atlas covers species distributions for the whole of Africa. Principally, this is to anticipate situations where suitable habitat shifts across national boundaries. Such shifts will occur when novel future climatic conditions for Ethiopia are already experienced in other African countries under their baseline climate. Another reason to model at the continental scale is to increase the number, and to reduce the bias, in the occurrence observations used for the model calibrations (see also Luedeling et al. 2014; Meyer et al. 2016). Finally, by scaling out to the whole of Africa, the atlas can be used by researchers, restoration planners and tree planters in other African nations, for those tree species that are common priorities with Ethiopia.

The modelling of contemporary and likely future tree species distributions, as carried out for the online atlas, can be used in three ways to narrow down what tree species to plant: first, by taking account of contemporary climate only; second, by considering future climate only; or, third, by considering both situations. In the last case, priority is given to the tree species that are predicted to be present in the future and that are present currently. This last option is an attractive one for both maximizing the probability of initial tree-planting success (establishment) and the likelihood of obtaining products and services from planted trees when these will only be fully realized decades later (e.g., when the product is timber or the service is carbon sequestration). For further information on these points, please refer to Kindt et al. (2023).

In this working paper, we describe the methods behind the creation of the online atlas. These methods, and most of the occurrence observations behind our maps, are available publicly. By sharing our methods, we hope they can be used more widely for mapping tree species distributions in current and predicted future climates. This would apply for mapping other tree species in Africa and for undertaking mapping on other continents. In this working paper, we do not discuss the interpretation of our maps – this will be covered in other, forthcoming publications. Readers of this working paper should also note that its purpose is not to provide an introduction to species distribution modelling methods. For readers who are not familiar with the basic methods for creating habitat suitability maps from species occurrence data and environmental raster data, we suggest they read the references we provide in our description of steps in the subsequent sections of this paper.

Overall, our modelling relies heavily on scripts run in the *R* software package.[1] Guisan et al. (2017) and Hijmans and Elith (2016–2021; https://rspatial.org/raster/sdm/index.html) specifically address the use of *R* for species distribution modelling. Another example *R* script for species distribution modelling, which showcases many of the same methods used for creating the current atlas, is available from https://rpubs.com/Roeland-KINDT/854918. The following tutorial shows how to use the graphical user interface of BiodiversityR for species distribution modelling: https://www.researchgate.net/publication/301515736_Ensemble_suitability_modelling_with_the_new_GUI_interface_of_BiodiversityR. Good starting points for an overall understanding of species distribution modelling are Guisan and Thuiller (2005), Guisan et al. (2017), Booth (2018) and Kindt (2018b). Note also that the following video is a recording of a seminar about our atlas: https://www.youtube.com/watch?v=csKvEeHl3jA.

In the following sections of this working paper, we discuss, step-by-step, our methods for atlas development. The different steps proceed from the collection of information on environmental predictor variables and species' occurrences for prioritized species, through data processing and model calibration, to the generation of the final maps. We also summarize the visualization of outputs in the online atlas.

1 We ran scripts in *R* version 3.6.1 (R Core Team 2019) for procedures described in Sections 3 to 9 of this working paper; in *R* version 3.6.0 for model calibrations and the generation of suitability maps in Section 10; and in *R* version 4.0.2 (R Core Team 2020) for procedures in Sections 11 to 15, and for the creation of the maps shown in the atlas (Section 16).

2 Selection of species

In this section, we explain how we came up with an initial list of tree species for modelling species distributions. We started with 153 species at this stage, a number later reduced to 127 species, as will be explained in subsequent sections.

An initial selection of priority tree species for the PATSPO project was undertaken in 2017. A 'Top 96' list of species was compiled first (Kindt 2018a). This included 25 priority tree species[2] identified in the Ethiopian Country Report for the State of the World's Forest Genetic Resources report (SoW-FGR; Institute of Biodiversity Conservation 2012); and other tree species in the SoW-FGR that were listed as important for solid wood production, for energy, for non-wood products, for agroforestry systems, for environmental services, and that have social values (see Table 4.1 in Kindt 2018a). Also included in Kindt's 'Top-96' list are tree species mentioned in the SoW-FGR for which genetic variability has been assessed; for which there are genetic or seed improvement programs; that are target species for *in situ* conservation; that have seed production areas; and for which seed are distributed by Ethiopia's Forestry Research Centre (FRC). The 'Top 96' of Kindt (2018a) further included tree species recorded on the seed price lists of the national and subnational (regional) tree seed centers in Ethiopia; and species imported by the High Value Tree Crops project.

Species were classified as native or exotic to Ethiopia based on information available from the SoW-FGR and the Useful Trees and Shrubs of Ethiopia publication (Bekele-Tesemma et al. 2007). For species that were not described in these sources, the Plants of the World Online portal (POWO; http://powo.science.kew.org/; see also Section 15) was consulted to identify their origin (accessed 22 November 2017).

From a 'long list' of 240 candidate species for species distribution modelling prepared at the same time as the 'Top 96' list (Kindt 2018a: Appendix II therein), the 'Top 96' list was expanded to 153 species. This was done by adding 57 further species that were native to Ethiopia and that were also included either in the Agroforestree Database (Orwa et al. 2009) or the University of Copenhagen Seed Leaflets series (from 1983 ongoing). Inclusion of the species in the Agroforestree Database or Seed Leaflets series was used as a proxy for the general usefulness of the trees in agroforestry and forestry. Table 1 lists all 153 species taken forward at this initial stage for distribution modelling.

2 Two species identified among the priority 27 in the Ethiopian Country Report for the SoW-FGR (Table 4 therein List of priority forest tree and shrub species), *Acacia drepanolobium* and *Prosopis juliflora*, are considered to be invasive species, and were not included in our modelling.

Table 1. 153 tree species[3] selected as initial candidates for species distribution modelling. The 'Criterion' column indicates how each species was selected (T25: among the 'Top 25' species; T96: otherwise among the 'Top 96' species; A: native species listed in the Agroforestree Database; L: native species listed in the Seed Leaflets series). Origin distinguishes between native (N) and exotic (E) to Ethiopia. The remaining columns document whether the species is listed in the Ecocrop database (E), the Selection of Forages for the Tropics (F), the Global Species Matrix (G), the Tropical Forestry Handbook (H), the Food Composition database (U) and the Wood Database (W).[4]

Species	Criterion	Origin	E	F	G	H	U	W
Acacia abyssinica	T96	N	x	-	-	-	-	-
Acacia decurrens	T96	E	x	-	-	x	-	-
Acacia lahai	A	N	-	-	-	-	-	-
Acacia melanoxylon	T96	E	x	-	-	x	-	x
Acacia nilotica	T96	N	x	x	x	x	-	-
Acacia polyacantha	T96	N	x	-	-	-	-	-
Acacia saligna	T96	E	x	-	x	x	-	-
Acacia senegal	T25	N	x	-	x	x	-	-
Acacia seyal	T96	N	x	-	x	-	-	x
Acacia sieberiana	A	N	x	-	-	-	-	-
Acacia tortilis	T96	N	x	-	x	x	-	-
Adansonia digitata	T25	N	x	-	x	-	-	-
Afrocarpus falcatus	T25	N	x	-	-	-	-	-
Albizia grandibracteata	T96	N	-	-	-	-	-	-
Albizia gummifera	T96	N	-	-	-	-	-	-
Albizia lebbeck	T96	E	x	x	x	x	-	x
Albizia schimperiana	T96	N	-	-	-	-	-	-
Annona senegalensis	A	N	x	-	-	-	-	-
Anogeissus leiocarpa	L	N	-	-	-	-	-	-
Antiaris toxicaria	A	N	-	-	-	-	-	-
Azadirachta indica	T96	E	x	-	x	x	-	-
Balanites aegyptiaca	T96	N	x	-	x	-	-	-
Bauhinia thonningii	T96	N	x	-	x	-	-	-
Berchemia discolor	A	N	x	-	-	-	-	-
Borassus aethiopum	AL	N	x	-	x	-	-	-
Boswellia microphylla	T96	N	-	-	-	-	-	-
Boswellia neglecta	T96	N	-	-	-	-	-	-
Boswellia ogadensis	T96	N	-	-	-	-	-	-
Boswellia papyrifera	T25	N	-	-	-	-	-	-
Boswellia pirottae	T96	N	-	-	-	-	-	-
Boswellia rivae	T96	N	-	-	-	-	-	-
Bridelia micrantha	AL	N	x	-	-	-	-	-
Cajanus cajan	T96	E	x	x	x	-	x	-

continued on next page

3 Species names in the atlas are current names standardized with World Flora Online (May 2019 version, WFO 2021; http://www.worldfloraonline.org/) via the WorldFlora package (Kindt 2020). Naming authorities are provided in Table A1.1 (Appendix 1). Synonyms are available from Tables A1.2 and A1.3.

4 Information from these (and other) databases was 'mined' recently to generate a prioritized list of 100 tree species for planting in the tropics and subtropics (Kindt et al. 2021). Details about the databases are provided in this publication.

Table 1. Continued

Species	Criterion	Origin	E	F	G	H	U	W
Calliandra calothyrsus	T96	E	x	x	x	x	-	-
Callistemon citrinus	T96	E	-	-	-	-	-	-
Calotropis procera	A	N	x	-	x	-	-	-
Capparis tomentosa	A	N	x	-	-	-	-	-
Carica papaya	T96	E	x	-	-	-	x	-
Casuarina cunninghamiana	T96	E	x	-	-	-	-	-
Casuarina equisetifolia	T96	E	x	-	x	x	-	-
Catha edulis	T25	N	x	-	-	-	-	-
Ceiba pentandra [5]	AL	E	x	-	x	x	-	-
Celtis africana	T96	N	-	-	-	-	-	-
Citrus sinensis	T96	E	x	-	-	-	x	-
Coffea arabica	T25	N	x	-	-	-	-	-
Combretum aculeatum	AL	N	x	-	-	-	-	-
Combretum collinum	A	N	-	-	-	-	-	-
Combretum molle	T96	N	x	-	-	-	-	-
Commiphora africana	T96	N	x	-	-	-	-	-
Commiphora guidottii	T96	N	-	-	-	-	-	-
Commiphora myrrha	T25	N	-	-	-	-	-	-
Cordeauxia edulis	T25	N	x	-	x	-	-	-
Cordia africana	T25	N	x	-	-	-	-	-
Corymbia citriodora	T96	E	-	-	-	x	-	x
Croton macrostachyus	T96	N	-	-	-	-	-	-
Cupressus lusitanica	T25	E	x	-	-	x	-	x
Cupressus sempervirens	T96	E	x	-	-	-	-	x
Cytisus proliferus	T96	E	x	-	x	-	-	-
Dalbergia melanoxylon	AL	N	x	-	-	-	-	x
Delonix regia	T96	E	x	-	-	x	-	-
Dichrostachys cinerea	A	N	x	-	-	-	-	-
Diospyros mespiliformis	A	N	x	-	-	-	-	-
Dobera glabra	A	N	x	-	-	-	-	-
Dodonaea viscosa	T96	N	x	-	-	-	-	-
Dombeya torrida	A	N	-	-	-	-	-	-
Dovyalis abyssinica	T96	N	-	-	-	-	-	-
Dovyalis caffra	T96	E	x	-	-	-	-	-
Ekebergia capensis	T96	N	x	-	-	-	-	-
Entada abyssinica	T96	N	x	-	-	-	-	-
Erythrina abyssinica	T96	N	x	-	-	-	-	-
Erythrina brucei	T96	N	-	-	-	-	-	-
Eucalyptus camaldulensis	T25	E	x	-	x	x	-	x

continued on next page

5 The exotic species *Ceiba pentandra* was included among the 153 candidate species as it had been identified as native to Ethiopia by Bekele-Tesemma et al. (2007), whereas information from Plants of the World Online (POWO) – compiled later and taken as a more authoritative source – indicated it to be exotic to the country.

Table 1. Continued

Species	Criterion	Origin	E	F	G	H	U	W
Eucalyptus globulus	T25	E	x	-	x	x	-	-
Eucalyptus grandis	T96	E	x	-	x	x	-	x
Eucalyptus saligna	T96	E	x	-	-	x	-	-
Eucalyptus viminalis	T96	E	x	-	-	x	-	-
Euphorbia tirucalli	A	N	x	-	x	x	-	-
Faidherbia albida	T25	N	x	-	x	x	-	-
Ficus carica	T96	E	x	-	-	-	x	-
Ficus sur	T96	N	-	-	-	-	-	-
Ficus sycomorus	T96	N	x	-	-	-	-	-
Flacourtia indica	A	N	x	-	-	-	-	-
Flueggea virosa	A	N	-	-	-	-	-	-
Garcinia livingstonei	A	N	-	-	-	-	-	-
Gardenia volkensii	L	N	-	-	-	-	-	-
Grevillea robusta	T25	E	x	-	-	x	-	x
Grewia damine	A	N	x	-	-	-	-	-
Grewia villosa	A	N	-	-	-	-	-	-
Hagenia abyssinica	T25	N	x	-	-	-	-	-
Hyphaene thebaica	A	N	x	-	x	-	-	-
Ilex mitis	A	N	-	-	-	-	-	x
Jacaranda mimosifolia	T96	E	x	-	-	x	-	-
Jatropha curcas	T96	E	x	-	x	-	-	-
Juniperus procera	T25	N	x	-	-	-	-	x
Kigelia africana	AL	N	x	-	-	-	-	-
Lawsonia inermis	A	N	x	-	x	-	-	-
Leucaena leucocephala	T96	E	x	x	x	x	-	-
Maerua aethiopica	T96	N	-	-	-	-	-	-
Malus domestica	T96	E	x	-	-	-	x	x
Mangifera indica	T96	E	x	-	-	-	x	x
Markhamia lutea	AL	N	x	-	-	-	-	-
Melia azedarach	T96	E	x	-	x	x	-	x
Milicia excelsa	AL	N	x	-	-	x	-	x
Millettia ferruginea	T96	N	-	-	-	-	-	-
Moringa oleifera	AL	N	x	-	x	-	x	-
Moringa stenopetala	T25	N	x	-	x	-	-	-
Nuxia congesta	A	N	-	-	-	-	-	-
Olea capensis	A	N	x	-	-	-	-	x
Olea europaea	T96	N	x	-	x	-	x	x
Oxytenanthera abyssinica	T25	N	-	-	x	-	-	-
Parkinsonia aculeata	T96	E	x	-	x	x	-	-
Persea americana	T96	E	x	-	x	-	x	-
Phoenix reclinata	T96	N	x	-	-	-	-	-
Pinus patula	T96	E	x	-	-	x	-	x

continued on next page

Table 1. Continued

Species	Criterion	Origin	E	F	G	H	U	W
Polyscias fulva	A	N	x	-	-	-	-	-
Pouteria adolfi-friedericii	T25	N	-	-	-	-	-	x
Prunus africana	T25	N	x	-	-	-	-	-
Pterolobium stellatum	T96	N	-	-	-	-	-	-
Rhamnus prinoides	T25	N	x	-	-	-	-	-
Saba comorensis	A	N	-	-	-	-	-	-
Salvadora persica	A	N	x	-	x	-	-	-
Sarcocephalus latifolius	A	N	-	-	-	-	-	-
Schefflera abyssinica	T96	N	-	-	-	-	-	-
Schinus molle	T96	E	x	-	-	x	-	-
Sclerocarya birrea	AL	N	x	-	x	-	-	-
Searsia natalensis	A	N	-	-	-	-	-	-
Securidaca longipedunculata	A	N	x	-	-	-	-	-
Senna didymobotrya	A	N	-	-	-	-	-	-
Sesbania bispinosa	T96	E	x	-	-	x	-	-
Sesbania sesban	T96	N	x	x	x	-	-	-
Shirakiopsis elliptica	A	N	x	-	-	-	-	-
Spathodea campanulata	T96	N	x	-	-	x	-	-
Steganotaenia araliacea	A	N	-	-	-	-	-	-
Stereospermum kunthianum	A	N	x	-	-	-	-	-
Strychnos henningsii	A	N	-	-	-	-	-	-
Strychnos innocua	A	N	x	-	-	-	-	-
Strychnos spinosa	A	N	x	-	-	-	-	-
Syzygium guineense	A	N	x	-	-	-	-	-
Tamarindus indica	T25	N	x	-	x	x	x	x
Tamarix aphylla	A	N	x	-	-	x	-	-
Terminalia brownii	T96	N	x	-	-	-	-	-
Trichilia emetica	AL	N	-	-	x	-	-	-
Vangueria madagascariensis	A	N	x	-	-	-	-	-
Vepris nobilis	A	N	-	-	-	-	-	-
Vernonia amygdalina	A	N	x	-	-	-	-	-
Vitellaria paradoxa	T25	N	x	-	x	-	-	-
Vitex doniana	A	N	x	-	-	-	-	-
Warburgia ugandensis	T96	N	x	-	-	-	-	-
Ximenia americana	AL	N	x	-	-	-	-	-
Yushania alpina	T25	N	-	-	-	-	-	-
Ziziphus jujuba [6]	T25	N	-	-	-	-	x	-
Ziziphus mucronata	A	N	x	-	-	-	-	-
Ziziphus spina-christi	T96	N	-	-	-	x	-	-

6 In the more recent (January 2023) version of World Flora Online, *Ziziphus mauritiana* is no longer listed as a synonym of *Ziziphus jujuba*. *Z. mauritiana* was listed in the original 'Top 25' species, whereas the synonym of *Z. jujuba* is used in the atlas.

3 Predictor variables

In this section, we explain how we came up with the list of bioclimatic, soil and topographic variables used to model the distributions of prioritized tree species (see also https://rspatial.org/raster/sdm/4_sdm_envdata.html).

Nineteen bioclimatic candidate predictor variables for generating species distributions were downloaded for the historical baseline climate of 1970 to 2000 from WorldClim 2.1 (Fick and Hijmans 2017; https://www.worldclim.org/; accessed April 2020). Downloads were at resolutions of 30 arc-seconds (~1 km) and 2.5 arc-minutes (150 arc-seconds, ~5 km). The higher resolution values were used for model calibrations and the lower resolution values were used to generate the actual maps for the baseline climate (a similar approach was used by Hannah et al. 2020 for a pantropical study). These and other geospatial raster layers were accessed in *R* via the *raster* package (versions 2.8-19 and 3.4-5; Hijmans 2020).

The 19 bioclimatic candidate predictor variables from WorldClim 2.1, as above, were expanded with 16 further bioclimatic predictors variables available from *envirem* (version 2.2; Title and Bemmels 2018) using the ***envirem::generateRasters*** function. Input rasters for this function, declared via the function of ***envirem::assignNames***, included monthly precipitation and minimum, maximum and mean monthly temperatures, downloaded from WorldClim 2.1 at resolutions of 30 and 150 arc-seconds (see previous paragraph for reasoning). As *envirem* calculations further required information on extraterrestrial solar radiation, raster layers with this information were calculated for each year from 1970 to 2000 separately via the ***envirem::ETsolradRasters*** function, and then averaged.

Included also as a separate candidate bioclimatic predictor variable was the Moisture Index, calculated by dividing annual precipitation (bioclimatic variable *BIO12* from WorldClim) by the annual potential evapotranspiration (PET) (bioclimatic variable *annualPET* from *envirem*).

Also added as a candidate bioclimatic predictor variable was *AriditySeason*, which is the balance between precipitation and PET for the dry season with the largest (most negative) such balance, calculated by the function of ***BiodiversityR::ensemble.PET.season*** (version 2.12-2; Kindt and Coe 2005). Among their environmental predictor variables, Hannah et al. (2020) used the similar variable *accumulated aridity index*, defined by the longest period where monthly PET was larger than the monthly precipitation. Another related variable to *AriditySeason* is the *maximum climatological water deficit*, as used by Chave et al. (2014; see also Do et al. 2021) to estimate the aboveground biomass of tropical trees, but *AriditySeason* also considers the occurrence of more than one rainy season in a particular location.

From the ENVIREM website (https://envirem.github.io/; accessed September 2016; note the use of capitals to differentiate the website from the *envirem* package mentioned above[7]), the topographic wetness index (variable *topoWet*) and topographic roughness index (variable *tri*) were also downloaded at resolutions of 30 and 150 arc-seconds. As above, the two different resolutions were required for model calibrations and actual map projections, respectively.

Soil measurements selected as candidate predictor variables were average bulk density (fine earth fraction in cg cm^{-3}), clay content (particles < 0.002 mm in the fine earth fraction in g kg^{-1}), silt content (particles ≥ 0.002 mm and ≤ 0.05 mm in the fine earth fraction in g kg^{-1}) and soil pH in H_2O (x 10). These measurements were obtained from SoilGrids250 (Hengl et al. 2017; https://www.isric.org/explore/

7 We use the same notations of 'envirem' and 'ENVIREM' in the remaining text to differentiate between the package and the website.

soilgrids; May 2020 release version). Values were taken for each variable from soil depths of 5 to 15 cm, 15 to 30 cm, 30 to 60 cm and 60 to 100 cm, before averaging for the variable for all soil depths. These were the same averaged soil variables used by Hannah et al. (2020), with the exception of depth to bedrock, which they also used.[8] Another difference in our analysis compared with Hannah et al. was that we used a higher resolution of 250 m for model calibrations. Soil variables, only at the higher resolution of 2.5 arc-minutes, were downloaded (August 2020) as raster layers by adapting an *R* script available from https://git.wur.nl/isric/soilgrids/soilgrids.notebooks/-/blob/master/markdown/wcs_from_R.md. This script results in averaged soil data at the selected resolution. After creating subsets of spatially thinned occurrence observations (see Sections 6 and 7) for each tree species, soil information was extracted as comma-separated data files from the highest resolution of 250 m of SoilGrids250 via Soilgrids REST API (https://rest.isric.org/soilgrids/v2.0/docs). This was done separately for each observation. The same method was used to extract soil information for background locations (see Section 8). A particular reason for us to include soil variables in our analysis was to model tree species that are edaphic specialists (Corlett and Tomlinson 2020, Hannah et al. 2020).

Once sets of candidate variables had been extracted, a Variance Inflation Factor (VIF; Fox and Monette 1992) analysis was carried out via function ***BiodiversityR::ensemble.VIF.dataframe*** to select a subset of lesser-correlated predictor variables for actual modelling, setting argument *VIF.max* to 5. First, a ***data.frame*** was created that contained all the information for the full set of background locations from the highest resolution data sets (250 m for soil variables and 30 arc-seconds for the other variables). After excluding the records with missing data, the data.frame contained 9,898 records.[9] Initially, we had intended to keep all of the variables of *ariditySeason*, *BIO6* (minimum temperature of the coldest month[10]), Moisture Index and *growingDegDays5* in our final subset of chosen variables for modelling, based in part on our reading of Booth (2016). However, within our final subset, we only retained *AriditySeason*, as several of the above variables had a final VIF > 10.

Table 2. Final subset of predictor variables selected for species suitability modelling in our analysis.
VIF = Variance Inflation Factor.

Predictor variable	VIF	VIF range	Source	Comment
AriditySeason	9.309	9.14 - 9.96	BiodiversityR	
BIO18	4.144	3.99 - 4.20	WorldClim	Precipitation of the warmest quarter
PETWettestQuarter	3.379	3.41 - 3.68	envirem	
topoWet	3.041	3.56 - 3.77	ENVIREM	
PETColdestQuarter	2.688	2.75 - 2.89	envirem	
tri	2.595	2.99 - 3.19	ENVIREM	
BIO15	2.515	2.22 - 2.33	WorldClim	Precipitation seasonality
BIO2	2.500	2.50 - 2.67	WorldClim	Mean diurnal range
PETDriestQuarter	2.478	2.49 - 2.73	envirem	
bdod	2.165	1.48 - 1.54	SoilGrids250	
BIO14	1.970	1.94 - 2.03	WorldClim	Precipitation of driest month
clay	1.750	1.84 - 1.89	SoilGrids250	
BIO19	1.658	1.61 - 1.67	WorldClim	Precipitation of driest quarter
silt	1.613	1.75 - 1.82	SoilGrids250	
monthCountByTemp10	1.552	1.36 - 1.45	envirem	

8 Depth to bedrock was excluded, as R^2 values for 10-fold cross-validations were below 55% (Hengl et al. 2017: Table 1). For our retained variables, R^2 values ranged from 72.6% to 83.4%.

9 10,000 random locations were selected across Africa with data on bioclimatic conditions (see Section 8). Missing data were a result of missing values for soil variables.

10 We tried both *BIO5* (maximum temperature of the warmest month) and *BIO6* as alternatives in VIF analyses. For both, VIF values were above 10.

Our final predictor subset consisted of 15 variables, all with VIF values below 5, except for *AriditySeason* with a VIF below 10 (Table 2). Settling on these VIF limits was consistent with previous studies. Ranjitkar et al. (2014a), for example, used a VIF threshold of 5 to select predictor variables for suitability modelling, as recommended by Rogerson (2000). Naimi et al. (2013), Ranjitkar et al. (2014b), de Sousa et al. (2019) and Ramirez-Villegas et al. (2020) in their analyses used a threshold of 10 for predictor variable selection. For our chosen subset of variables, high pairwise correlations (of magnitude ≥ 0.8; the limit set by Ranjitkar et al. 2014a) were observed only between *AriditySeason* and *BIO18*, and between *topoWet* and *tri* (Figures 1 to 3).

VIF was calculated for the highest resolution data available. The VIF range was obtained from 10 repetitions of the analysis, with default settings for the ***BiodiversityR::ensemble.VIF*** function using the 2.5 arc-minutes raster layers as predictors to create baseline maps.

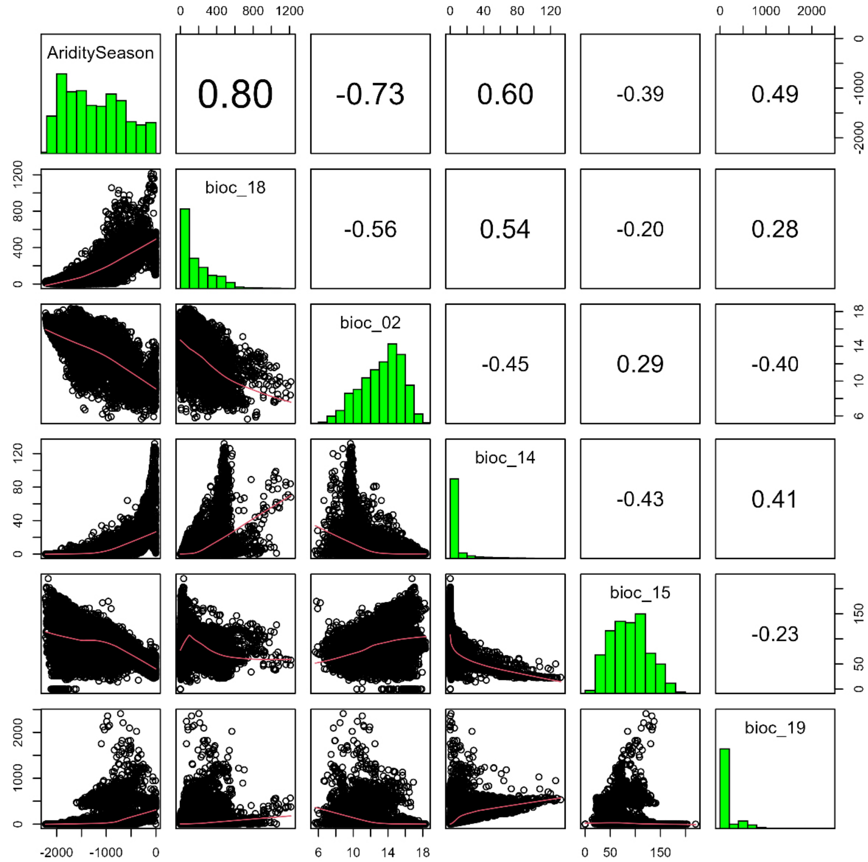

Figure 1. Scatterplot matrix showing correlations among *AriditySeason* and selected predictor variables from WorldClim. The graph was created with function ***BiodiversityR::ensemble.pairs*** using default settings (1,000 randomly selected points) for the baseline raster layers at 2.5 minutes resolution.

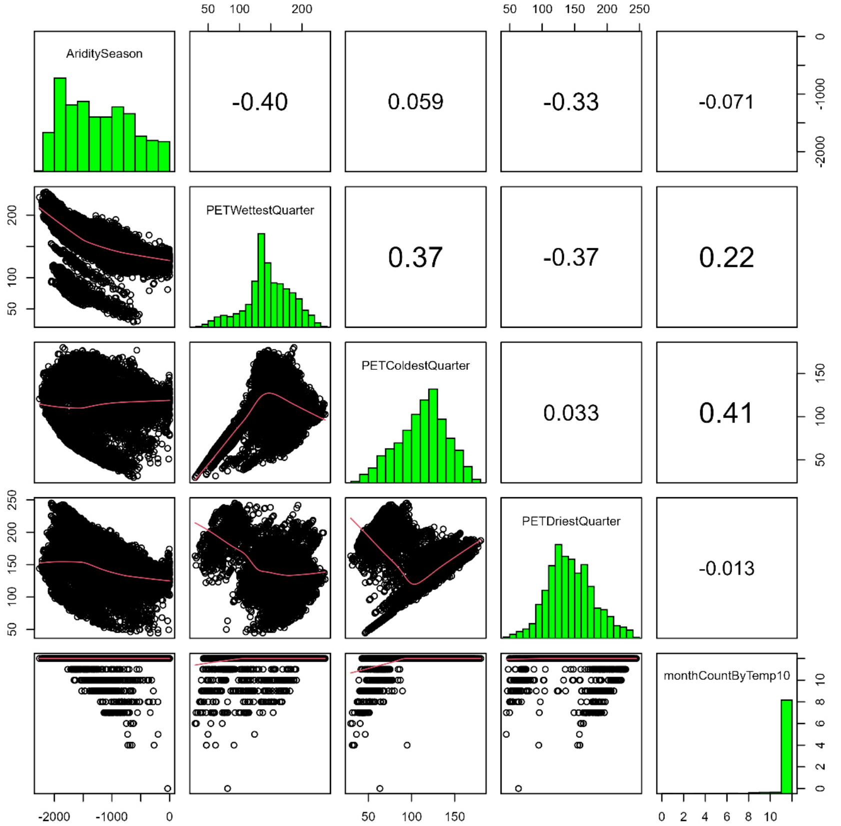

Figure 2. Scatterplot matrix showing correlations among *AriditySeason* and selected predictor variables from envirem. The graph was created with function ***BiodiversityR::ensemble.pairs*** with default settings (1,000 randomly selected points) using the baseline raster layers at 2.5 minutes resolution.

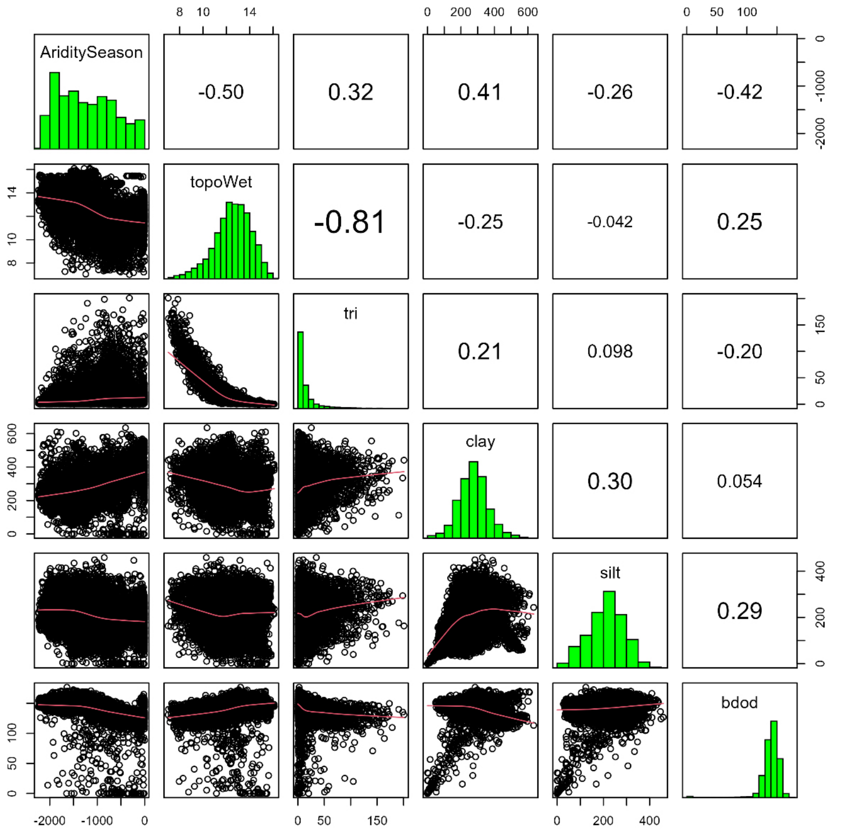

Figure 3. Scatterplot matrix showing correlations among *AriditySeason* and selected predictor variables from ENVIREM and SoilGrids250. The graph was created with function ***BiodiversityR::ensemble.pairs*** with default settings (1,000 randomly selected points) using the baseline raster layers at 2.5 minutes resolution.

4 Future climates

In this section, we provide information about the future climates that we used for species distribution modelling for prioritized tree species in the atlas.

Future climates in the atlas correspond to projections for the middle of the 21st century (2050s, 2041–2060) under a low-emissions scenario (Shared Socioeconomic Pathway [SSP] 1-2.6) and a high-emissions scenario (SSP 3-7.0) from CMIP6. SSP 1-2.6 is the CMIP6 equivalent of the CMIP5 low-emissions scenario of RCP2.6. SSP 3-7.0 is a middle-of-the-road high-emissions scenario of CMIP6, between worst case and optimistic outcomes when the world fails to enact any climate policies (for further information, see: https://www.carbonbrief.org/cmip6-the-next-generation-of-climate-models-explained).

For both of these scenarios, bioclimatic and monthly climatic data were downloaded for nine Global Climate Models (GCMs, or General Circulation Models) available from WorldClim 2.1. These are: BCC-CSM2-MR, CNRM-CM6-1, CNRM-ESM2-1, CanESM5, GFDL-ESM4, IPSL-CM6A-LR, MIROC-ES2L, MIROC6 and MRI-ESM2-0. The resolution of the raster layers we used was 2.5 arc-minutes. This was the highest resolution available for future climates from WorldClim 2.1, when downloading these raster layers in 2020.

Values for the expanded set of bioclimatic variables (*AriditySeason*, Moisture Index and variables generated via *envirem*) were calculated for each GCM and each emission scenario with similar methods to those used for the baseline climate layers. As these calculations required details on extraterrestrial solar radiation, relevant raster layers were first created for each year from 2041 to 2060 via the function ***envirem::ETsolradRasters***, and these were then averaged.

5 Compilation of occurrence observations

In this section, we explain how we collected occurrence data for prioritized tree species. Species occurrence observations in geographic space are the basis for modelling individual species distributions.

We combined occurrence data for our prioritized species from eight sources (Table 3). Three of the datasets used (AERTS, DEMISSEW and BORCHARDT) only documented occurrences in Ethiopia, but we included these given our particular focus on that country. (Note that spatial and environmental thinning procedures described in the following sections reduce potential bias towards Ethiopia.) As the AERTS, DEMISSEW and BORCHARDT datasets refer only to Ethiopia, we did not use the data cleaning protocols described in the next paragraph. (For the same reason, we also did not use such protocols for the Burkina Faso TERRIBLE dataset.)

For the datasets of GBIF and NATURALIS, we used data cleaning protocols available via the *CoordinateCleaner* package (version 1.0-7; Zizka et al. 2019). We deemed these cleaning procedures unnecessary for pan-African datasets where the procedures for excluding erroneous locations were clearly documented (for BIEN, version 1.2.4, see Maitner et al. 2017; for RAINBIO, see Dauby et al. 2016). We used current names for location determinations and known synonyms (see Tables A1.2 and A1.3, Appendix 1).

Table 3. Datasets of occurrence observations used in our analysis. The 'Records' column indicates the number of references in total to our initially chosen 153 tree species.

Dataset	*Records*	Compilation
GBIF https://www.gbif.org/	*257,988*	Data were downloaded on 1 October 2018 from GBIF via the function of ***dismo::gbif*** (version 1.1-4; Hijmans et al. 2017). Records without longitude and latitude data were removed. Records for synonym names that were not accepted synonyms according to *ThePlantList* (consulted in November 2018) were removed. Observations where the basis of the record was "FOSSIL_SPECIMEN" or "UNKNOWN" were removed. Records flagged by the ***CoordinateCleaner::Clean Coordinates*** function as records with potentially erroneous geolocation data were removed, including records where the country designation did not correspond to the country boundaries from ***rnaturalearth::ne_countries(scale = 10)***. See Appendix 2 for identities of the occurrence datasets.
BIEN https://bien.nceas.ucsb.edu/bien/	*74,213*	Data were downloaded on 18^{n} July 2020 via the function of ***BIEN::BIEN_occurrence_species*** with default settings of the function. See Appendix 3 for details on the custodians of the data.
RAINBIO https://gdauby.github.io/rainbio/	*11,991*	Data were compiled from the RAINBIO mega database, sourcing data both from a file with native species and a file with exotic species.
NATURALIS https://bioportal.naturalis.nl	*9,587*	Data were compiled by co-author Jan Wieringa from herbarium records available in the Naturalis Herbarium. Records without longitude and latitude data were removed. Records flagged by the ***CoordinateCleaner::CleanCoor dinates*** function as records with potentially erroneous geolocation data were removed, including records where the country designation did not correspond to the country boundaries from ***rnaturalearth::ne_countries(scale = 10)***.

continued on next page

Table 3. Continued

Dataset	*Records*	Compilation
TERRIBLE	*7,823*	Data were originally compiled by lead author Roeland Kindt for species suitability modelling for Burkina Faso, available from a map of tree and shrub distributions in the country prepared by Terrible (1975), as documented in Gaisberger et al. (2016).
AERTS	*1,911*	Data were compiled from supplementary materials available from a study on church forests in Ethiopia by Aerts et al. (2016).
DEMISSEW	*597*	Data were compiled by co-author Sebsebe Demissew from information available in the Ethiopian herbarium.
BORCHARDT	*249*	Data were compiled by co-author Peter Borchardt from information on mother trees selected as individual seed sources for tree planting and restoration projects in which he has been involved.

6 Spatial thinning of occurrence observations

In this section, we explain how we spatially thin occurrence data to reduce sampling biases that can otherwise occur in generating species distribution maps.

To reduce possible sampling biases with occurrence data, we applied spatial thinning using the functions ***BiodiversityR::ensemble.spatialThin*** and ***BiodiversityR::ensemble.spatialThin.quant***, based on a similar algorithm to ***spThin::thin*** (Aiello-Lammens 2015). This procedure thins out species records using a random systematic approach to record removal until all paired occurrences are above a minimum distance threshold.

First, we rounded all occurrence coordinates to four decimal places and removed duplicate records for each species (Table 4). Second, occurrence data were limited to Africa, the region covered by the atlas. Then, for species where the number of retained records in Africa was above 50, the argument of ***thin.km*** in the ***BiodiversityR::ensemble.spatialThin*** and ***BiodiversityR::ensemble.spatialThin.quant*** functions was set at 10 km. This minimum distance has been widely used in species distribution modelling studies (Aiello-Lammens 2015, Title and Bemmels (2018), Castellanos et al. 2019, van Zonneveld et al. 2020). The criterion of 10 km was applied to 121 tree species from our prioritized species list (see 'km' column in Table 4); only two of these species[11] then failed additional criteria for species distribution modelling (as explained in Section 10).

For 32 tree species on our initial prioritized list for which the number of retained records in Africa was 50 or lower (Table 4: 'Africa' column), the argument of ***thin.km*** was set at a less stringent 2 km for spatial filtering with ***BiodiversityR::ensemble.spatialThin***. This meant that one occurrence observation was retained per 30 arc-seconds grid cell, a similar procedure for limiting occurrence records to unique grid cells to that used for model calibrations by a number of other authors (e.g., de Sousa et al. 2019, Thuiller et al. 2019, Fremout et al. 2020, Hannah et al. 2020). Lowering the distance criterion from 10 km to 2 km captured another eight species[12] that could be taken forward for distribution modelling (meaning 121 - 2 + 8 = 127 species in total).

Table 4. Results from the spatial thinning of occurrence observations. 0.0001: Number of records retained after rounding longitude and latitude to four decimals and removing duplicate records; Africa: Number of records retained in Africa; km: setting of argument thin.km for function ***BiodiversityR::ensemble.spatialThin*** and ***BiodiversityR::ensemble.spatialThin.quant***; Thinned: Number of records retained after thinning, used to sort species in the table; Percentage: % retained from the '0.0001' records; SDM: whether a species distribution model was fitted ultimately (see text and Table 9).

Species	0.0001	Africa	km	Thinned	Percentage	SDM
Combretum molle	1405	1403	10	910	64.9	YES
Dichrostachys cinerea	1436	1377	10	896	65.1	YES
Combretum collinum	1853	1852	10	837	45.2	YES
Vitellaria paradoxa	10176	10176	10	774	7.6	YES
Tamarindus indica	2746	1752	10	743	42.4	YES
Syzygium guineense	1046	1033	10	702	68.0	YES
Annona senegalensis	1232	1227	10	697	56.8	YES

continued on next page

11 These were *Albizia grandibracteata* and *Millettia ferruginea*: see Table 4.

12 These were *Acacia decurrens*, *Boswellia neglecta*, *Boswellia papyrifera*, *Cupressus sempervirens*, *Dobera glabra*, *Pouteria adolfi-friedericii*, *Warburgia ugandensis* and *Yushania alpina*; see Table 4.

Table 4. Continued

Species	0.0001	Africa	km	Thinned	Percentage	SDM
Anogeissus leiocarpa	4337	4331	10	689	15.9	YES
Flueggea virosa	2973	1291	10	687	53.2	YES
Acacia senegal	899	882	10	673	76.3	YES
Bauhinia thonningii	1315	1313	10	660	50.3	YES
Acacia nilotica	904	791	10	655	82.8	YES
Acacia sieberiana	1285	1281	10	655	51.1	YES
Strychnos spinosa	1073	1067	10	648	60.7	YES
Ficus sur	1326	1321	10	636	48.1	YES
Leucaena leucocephala	5499	1625	10	629	38.7	YES
Azadirachta indica	2425	2019	10	620	30.7	YES
Diospyros mespiliformis	1384	1381	10	606	43.9	YES
Sclerocarya birrea	1010	1004	10	606	60.4	YES
Adansonia digitata	1168	1096	10	589	53.7	YES
Ziziphus mucronata	675	669	10	564	84.3	YES
Ximenia americana	1890	801	10	561	70.0	YES
Acacia polyacantha	807	805	10	533	66.2	YES
Vitex doniana	1124	1118	10	533	47.7	YES
Grevillea robusta	2918	1461	10	503	34.4	YES
Calotropis procera	2237	1232	10	495	40.2	YES
Euphorbia tirucalli	1326	1191	10	492	41.3	YES
Phoenix reclinata	765	714	10	491	68.8	YES
Melia azedarach	4874	993	10	472	47.5	YES
Delonix regia	1820	1028	10	471	45.8	YES
Jacaranda mimosifolia	5079	1004	10	471	46.9	YES
Ekebergia capensis	652	650	10	470	72.3	YES
Ceiba pentandra	3220	2378	10	462	19.4	YES
Kigelia africana	741	690	10	461	66.8	YES
Balanites aegyptiaca	841	772	10	440	57.0	YES
Olea europaea	18801	536	10	436	81.3	YES
Sarcocephalus latifolius	913	907	10	432	47.6	YES
Bridelia micrantha	600	594	10	428	72.1	YES
Milicia excelsa	1366	1364	10	414	30.4	YES
Senna didymobotrya	1156	978	10	405	41.4	YES
Dodonaea viscosa	24913	552	10	397	71.9	YES
Entada abyssinica	784	783	10	396	50.6	YES
Acacia tortilis	547	437	10	384	87.9	YES
Trichilia emetica	637	634	10	381	60.1	YES
Acacia seyal	490	487	10	375	77.0	YES
Commiphora africana	445	442	10	374	84.6	YES
Ficus sycomorus	614	497	10	368	74.0	YES
Borassus aethiopum	1421	1418	10	364	25.7	YES
Sesbania sesban	668	482	10	363	75.3	YES
Celtis africana	462	453	10	359	79.2	YES

continued on next page

Table 4. Continued

Species	0.0001	Africa	km	Thinned	Percentage	SDM
Stereospermum kunthianum	803	803	10	344	42.8	YES
Albizia lebbeck	1334	742	10	326	43.9	YES
Ilex mitis	519	514	10	326	63.4	YES
Flacourtia indica	586	466	10	325	69.7	YES
Parkinsonia aculeata	2699	582	10	300	51.5	YES
Jatropha curcas	1300	450	10	292	64.9	YES
Cajanus cajan	1212	356	10	288	80.9	YES
Rhamnus prinoides	336	334	10	279	83.5	YES
Vernonia amygdalina	354	338	10	276	81.7	YES
Securidaca longipedunculata	334	333	10	265	79.6	YES
Capparis tomentosa	321	318	10	263	82.7	YES
Combretum aculeatum	314	312	10	255	81.7	YES
Nuxia congesta	330	325	10	252	77.5	YES
Albizia gummifera	358	354	10	251	70.9	YES
Faidherbia albida	371	290	10	251	86.6	YES
Salvadora persica	369	291	10	251	86.3	YES
Antiaris toxicaria	1987	1718	10	237	13.8	YES
Dovyalis caffra	665	619	10	226	36.5	YES
Saba comorensis	319	317	10	221	69.7	YES
Dalbergia melanoxylon	278	273	10	220	80.6	YES
Olea capensis	315	307	10	219	71.3	YES
Shirakiopsis elliptica	285	285	10	205	71.9	YES
Prunus africana	279	277	10	204	73.6	YES
Strychnos innocua	394	393	10	202	51.4	YES
Ziziphus jujuba	628	234	10	193	82.5	YES
Searsia natalensis	285	277	10	191	69.0	YES
Spathodea campanulata	869	261	10	190	72.8	YES
Steganotaenia araliacea	244	244	10	185	75.8	YES
Garcinia livingstonei	226	222	10	179	80.6	YES
Mangifera indica	1835	754	10	179	23.7	YES
Erythrina abyssinica	218	215	10	178	82.8	YES
Grewia villosa	205	179	10	160	89.4	YES
Berchemia discolor	190	187	10	158	84.5	YES
Acacia saligna	3237	270	10	151	55.9	YES
Moringa oleifera	531	238	10	148	62.2	YES
Cordia africana	191	182	10	147	80.8	YES
Oxytenanthera abyssinica	242	242	10	139	57.4	YES
Lawsonia inermis	413	165	10	137	83.0	YES
Juniperus procera	201	171	10	132	77.2	YES
Sesbania bispinosa	183	137	10	132	96.4	YES
Croton macrostachyus	177	177	10	131	74.0	YES
Vepris nobilis	175	171	10	120	70.2	YES
Afrocarpus falcatus	224	174	10	117	67.2	YES

continued on next page

Table 4. Continued

Species	0.0001	Africa	km	Thinned	Percentage	SDM
Gardenia volkensii	130	124	10	110	88.7	YES
Markhamia lutea	144	138	10	106	76.8	YES
Acacia abyssinica	133	132	10	105	79.5	YES
Hyphaene thebaica	354	342	10	104	30.4	YES
Pterolobium stellatum	127	125	10	103	82.4	YES
Polyscias fulva	136	134	10	102	76.1	YES
Strychnos henningsii	131	128	10	100	78.1	YES
Vangueria madagascariensis	172	141	10	100	70.9	YES
Pinus patula	794	223	10	92	41.3	YES
Catha edulis	110	98	10	86	87.8	YES
Acacia melanoxylon	22812	131	10	85	64.9	YES
Carica papaya	1470	141	10	81	57.4	YES
Grewia damine	90	87	10	79	90.8	YES
Eucalyptus camaldulensis	25511	241	10	73	30.3	YES
Dombeya torrida	102	96	10	71	74.0	YES
Acacia lahai	101	101	10	66	65.3	YES
Ziziphus spina-christi	835	76	10	63	82.9	YES
Dovyalis abyssinica	88	81	10	61	75.3	YES
Albizia schimperiana	69	69	10	59	85.5	YES
Terminalia brownii	74	74	10	59	79.7	YES
Hagenia abyssinica	91	81	10	57	70.4	YES
Albizia grandibracteata	82	82	10	53	64.6	NO
Coffea arabica	1106	77	10	53	68.8	YES
Schinus molle	4438	56	10	52	92.9	YES
Persea americana	1793	72	10	50	69.4	YES
Boswellia neglecta	50	50	2	49	98.0	YES
Pouteria adolfi-friedericii	50	49	2	49	100.0	YES
Millettia ferruginea	58	58	10	46	79.3	NO
Schefflera abyssinica	56	56	10	44	78.6	YES
Boswellia papyrifera	45	45	2	43	95.6	YES
Casuarina equisetifolia	1372	55	10	39	70.9	YES
Warburgia ugandensis	35	35	2	34	97.1	YES
Erythrina brucei	33	33	2	33	100.0	NO
Yushania alpina	35	35	2	33	94.3	YES
Acacia decurrens	2233	28	2	28	100.0	YES
Dobera glabra	25	21	2	21	100.0	YES
Commiphora myrrha	28	20	2	20	100.0	NO
Cupressus sempervirens	836	20	2	20	100.0	YES
Tamarix aphylla	581	23	2	19	82.6	NO
Boswellia rivae	19	19	2	18	94.7	NO
Casuarina cunninghamiana	2969	15	2	15	100.0	NO
Cupressus lusitanica	1184	14	2	14	100.0	NO
Eucalyptus globulus	5096	12	2	12	100.0	NO

continued on next page

Table 4. Continued

Species	0.0001	Africa	km	Thinned	Percentage	SDM
Ficus carica	397	12	2	11	91.7	NO
Callistemon citrinus	1694	10	2	10	100.0	NO
Maerua aethiopica	11	11	2	10	90.9	NO
Eucalyptus grandis	1270	9	2	9	100.0	NO
Citrus sinensis	343	8	2	8	100.0	NO
Commiphora guidottii	9	9	2	8	88.9	NO
Boswellia microphylla	7	7	2	7	100.0	NO
Corymbia citriodora	1121	7	2	7	100.0	NO
Boswellia pirottae	6	6	2	6	100.0	NO
Calliandra calothyrsus	357	5	2	5	100.0	NO
Moringa stenopetala	9	5	2	5	100.0	NO
Boswellia ogadensis	4	4	2	4	100.0	NO
Eucalyptus saligna	2907	4	2	4	100.0	NO
Eucalyptus viminalis	12937	4	2	4	100.0	NO
Cordeauxia edulis	5	3	2	3	100.0	NO
Cytisus proliferus	735	3	2	3	100.0	NO
Malus domestica	18931	0	2	0		NO

7 Environmental thinning of occurrence observations

In this section, we explain how we environmentally thin occurrence data to reduce sampling biases that can otherwise occur in generating species distribution maps.

Table 5. Minimum number of occurrence observations recommended for species distribution modelling, according to selected references, ordered by number of recommended observations.

Reference	Records	Comments
Santini et al. (2021)	200–500	These authors suggest using a large sample but also stress that no magic number exists, given: uncertainties about the true ecology of a species and its link to predictor variables; whether the species is in equilibrium with the environment; and whether presence points are biased
Feeley and Silman (2011)	200	
Varela et al. (2014)	50	For non-filtered and biased data
Wisz et al. (2008)	30	
van Proosdij et al. (2015)	25	For widespread species in Africa
Rivers et al. (2011)	15	Number of herbarium records required for IUCN Red List calculations
van Proosdij et al. (2015)	14	For narrow-ranged species in Africa
Varela et al. (2014)	5	For environmentally-filtered data. This number may increase for species with more complex niches, but also in these cases smaller environmentally-filtered datasets are expected to outperform larger biased datasets

According to different authors, the minimum number of occurrence records required for species distribution modelling ranges from 5 to > 200 observations (Table 5). Selecting occurrence observations less biased in environmental space can increase the performance of species suitability models (Varela et al. 2014, Castellanos et al. 2019). We therefore applied the ***BiodiversityR::ensemble.environmentalThin*** function for environmental filtering of the occurrence observations of some species.

We used the following rules to calculate the number of occurrence observations to retain for each of our tree species (this number is used for argument ***thin.n*** in the function):

- Remove at least one third of observations closest in environmental space for species with at least 75 occurrences after spatial thinning.
- Maximally retain 200 observations after environmental thinning.
- Minimally retain 50 observations after environmental thinning for species with 50 to 74 observations.
- No environmental thinning when the initial number of spatial occurrence records is 50 or below.

The targets of retained occurrence observations are summarized in Table 6. Our overall aim was to retain a high number of occurrence observations to calibrate species distribution models (see also Castellanos et al. 2019: Figure 3 therein).

For a subset of species, the environmental thinning process resulted in a dataset with fewer records than the target number. This was due to the second algorithm applied in the ***BiodiversityR::ensemble.environmentalThin*** function, where the random selection process attempts to create smaller subsets with the same minimum environmental distance.

Table 6. Target number of environmentally thinned occurrence observations. Input represents the number of occurrence observations used to calculate the target.

Input	Target	Input	Target	Input	Target
5	5	105	70	205	136
10	10	110	73	210	140
15	15	115	76	215	143
20	20	120	80	220	146
25	25	125	83	225	150
30	30	130	86	230	153
35	35	135	90	235	156
40	40	140	93	240	160
45	45	145	96	245	163
50	50	150	100	250	166
55	50	155	103	255	170
60	50	160	106	260	173
65	50	165	110	265	176
70	50	170	113	270	180
75	50	175	116	275	183
80	53	180	120	280	186
85	56	185	123	285	190
90	60	190	126	290	193
95	63	195	130	295	196
100	66	200	133	≥ 300	200

Environmental thinning failed for three species: *Albizia grandibracteata*, *Coffea arabica* and *Schinus molle*. For these species, the occurrences of spatially-thinned observations alone were retained for modelling (see Table 4).

8 Compilation of background observations

In this section, we explain how we selected background (pseudo-absence) locations for species distribution modelling.

Across the domain covered by the atlas, we randomly selected 10,000 background (pseudo-absence) locations (see https://rspatial.org/raster/sdm/3_sdm_absence-background.html) that had non-missing values for bioclimatic predictor variables[13] at the resolution of 30 arc-seconds. This was done via the ***dismo::randomPoints*** function and resulted in the locations shown in Figure 4. For each of the background locations, soil data were obtained via Soilgrids REST API (see Section 3). We excluded from our initial 10,000 background locations those with missing soil data, resulting in a final tally of 9,898 random background locations with complete details for predictor variables.

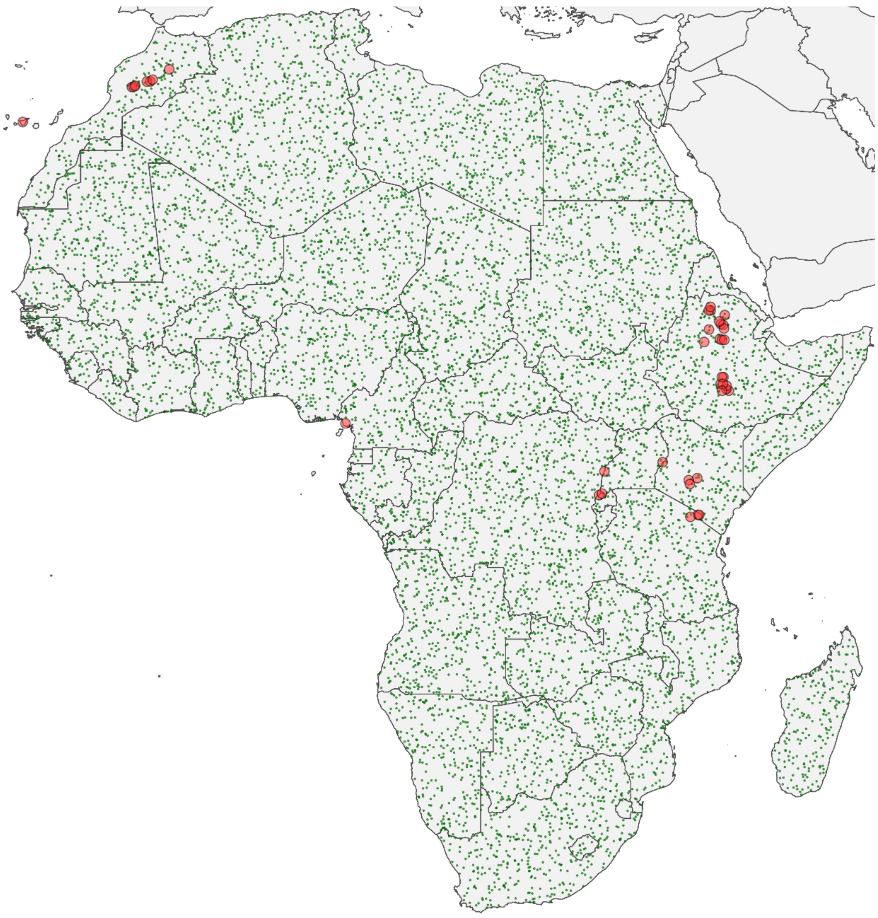

Figure 4. Candidate background locations. Small, green symbols depict 9,898 randomly selected locations across Africa. Large, red symbols depict selected mountain peaks (see explanation in text).

13 Soil raster data were only obtained for the resolution of 2.5 arc-minutes, as described in Section 3.

Table 7. Mountains peaks added to the set of background locations[14].

Mountain peak	Height (in m)	Mountain peak	Height (in m)
Kibo (Uhuru Pk)	5895	Mount Cameroon	4040
Mount Kenya (Batian)	5199	Weshema / Wasema?	4030
Mawenzi (Hans Meyer Pk)	5148	Oldoinyo Lesatima	4001
Ngaliema / Mt Stanley	5109	Jebel n'Tarourt / Tifnout / Iferouane	3996
Mount Meru (Socialist Pk)	4566	Muggia	3950
Ras Dashen	4550	Dubbai	3941
Karisimbi	4507	Taska n'Zat	3912
Bwahit	4437	Mount Kinangop	3902
Tullu Demtu	4377	Cimbia	3900
Mount Elgon (Wagagai)	4321	Ieciuol ?	3840
Amba Farit	4270	Kawa / Caua / Lajo	3830
Abune Yosef / Guliba Amba	4260	Jbel Tignousti	3819
Bada	4195	Filfo / Encuolo	3805
Kaka / Kecha / Chiqe	4193	Kosso Amba	3805
Jbel Toubkal	4167	Baylamtu / Gavsigivla	3777
Muhavura	4127	Ouaougoulzat	3763
Guna	4120	Somkaru	3760
Choqa / Choke / Birhan	4100	Abieri	3750
Chilalo	4071	Arin Ayachi	3747
Ighil Mgoun	4068	Teide	3718

To the 9,898 randomly selected background locations, we added 40 locations representing the highest mountain peaks in Africa that were otherwise easy to miss in background location selection. This was important because some of our initial model runs predicted some (Afromontane) tree species to be suitable for mountain peaks. An initial list of 98 mountain peaks with location details was downloaded from Wikipedia (https://en.wikipedia.org/wiki/List_of_highest_mountain_peaks_of_Africa; accessed 14 July 2020), and the peaks were ordered by altitude. We then created a new peak list, eliminating peaks lower down on the ordered original list if they were located less than 10 km from a mountain peak higher on the list. Pairwise geographical distances between mountain peaks were calculated via the function of ***geosphere::distGeo*** (version 1-5-10; Hijmans 2019). Ultimately, we decided to retain the 40 highest mountain peaks (Table 7).

Then, for each tree species, a separate set of background locations was selected as a subset from the full set of background locations constituting the initial random locations and the locations of mountain peaks. In a first step, for each species a subset of background locations was created by selecting only those locations within a 500 km buffer (created via function ***dismo::circles***) of occurrence observations (Figure 5). This buffer width had been used earlier by Hannah et al. (2020) (see also mentions of 500 km as a potential migration or dispersal distance by Lazarus and McGill 2014, Hoenner et al. 2018 and Iverson et al. 2019).

14 The question marks associated with mountain peak names are as in the initial Wikipedia names list

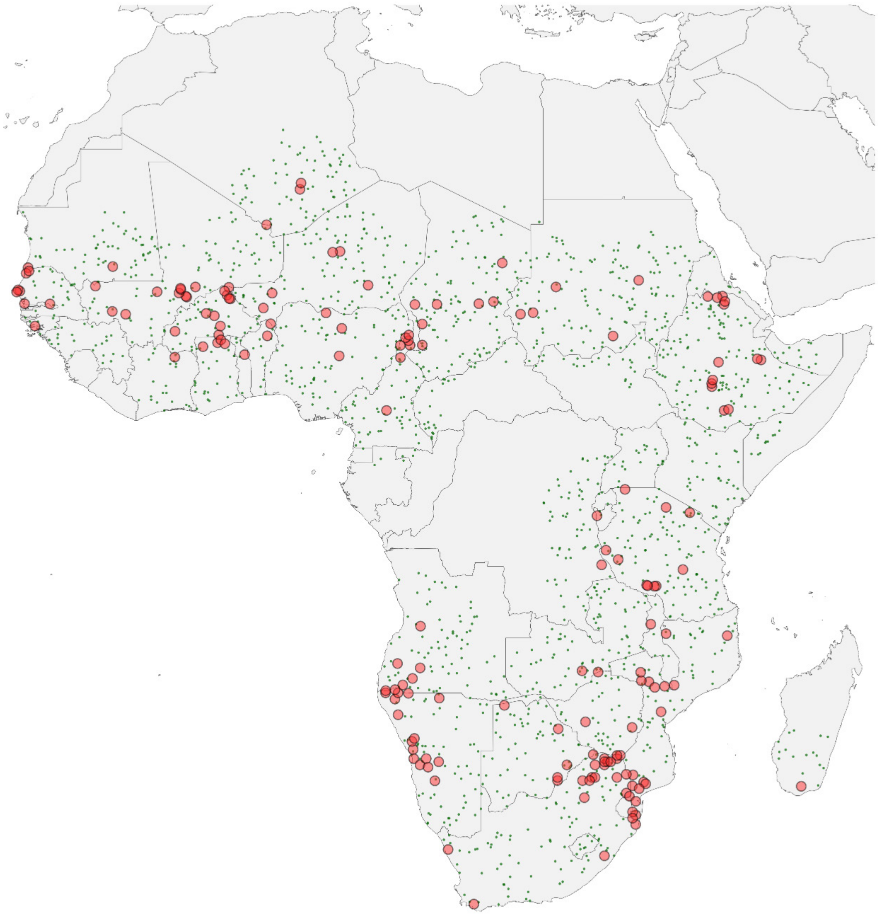

Figure 5. For each species, background locations (small, green symbols) used for modelling were restricted to a 500 km buffer around occurrence locations (large, red symbols). The map shown here depicts occurrence and background locations for *Faidherbia albida*.

Table 8. Target number of background observations

Occurrence	Background	Occurrence	Background	Occurrence	Background
5	400	75	750	145	1450
10	400	80	800	150	1500
15	400	85	850	155	1550
20	400	90	900	160	1600
25	400	95	950	165	1650
30	400	100	1000	170	1700
35	400	105	1050	175	1750
40	400	110	1100	180	1800
45	450	115	1150	185	1850
50	500	120	1200	190	1900
55	550	125	1250	195	1950
60	600	130	1300	200	2000
65	650	135	1350	500	5000
70	700	140	1400	1000	10000

In a second step, the set of random background locations within the buffer was randomly subsetted via the ***base::sample*** function, using the following rules (see also Table 8):

- The target number for the subset of background locations should be 10 times the number of occurrence observations when these observations are 40 or more.
- When the number of occurrence observations is less than 40, the target should still be 400 background observations.
- If fewer background locations are available than the target, then all available background locations should be retained.

For species with 50 or more occurrence observations, our algorithm for subsetting applied the same rules as Khoury et al. (2019). However, unlike in our case, Khoury et al. (2019) used background locations that were 100 times the number of occurrence locations for species with fewer than 50 occurrences.[15]

Our approach took account of the observations of the simulation study of Grimmet et al. (2020: their Figure 4) who showed that a prevalence of 0.1 (the ratio of occurrence to background locations) is a good compromise for the performance of different algorithms of species distribution modelling. The same authors found that using a prevalence of 0.05 also resulted in acceptable, but slightly lower, model performances among different algorithms.[16] As in our case the lowest number of occurrence observations was 20 for the species retained for species suitability modelling (for *Cupressus sempervirens*, see Table 9 in next section), we expect our choices on setting background locations (at 400 for species with fewer than 40 presence observations, corresponding to the lowest prevalence of 0.05 = 20 / 400 for *Cupressus sempervirens*), to be appropriate.

15 The Khoury et al. (2019) algorithm creates the anomaly that the number of background locations does not always decrease for decreasing numbers of occurrence observations. For example, the algorithm selects 500 background observations for 50 species observations and 3,000 background observations for 30 species observations.

16 As locations of mountain peaks within the 500 km buffer were added, prevalence values were slightly below 0.1 for most species.

9 Spatial folding of occurrence and background observations

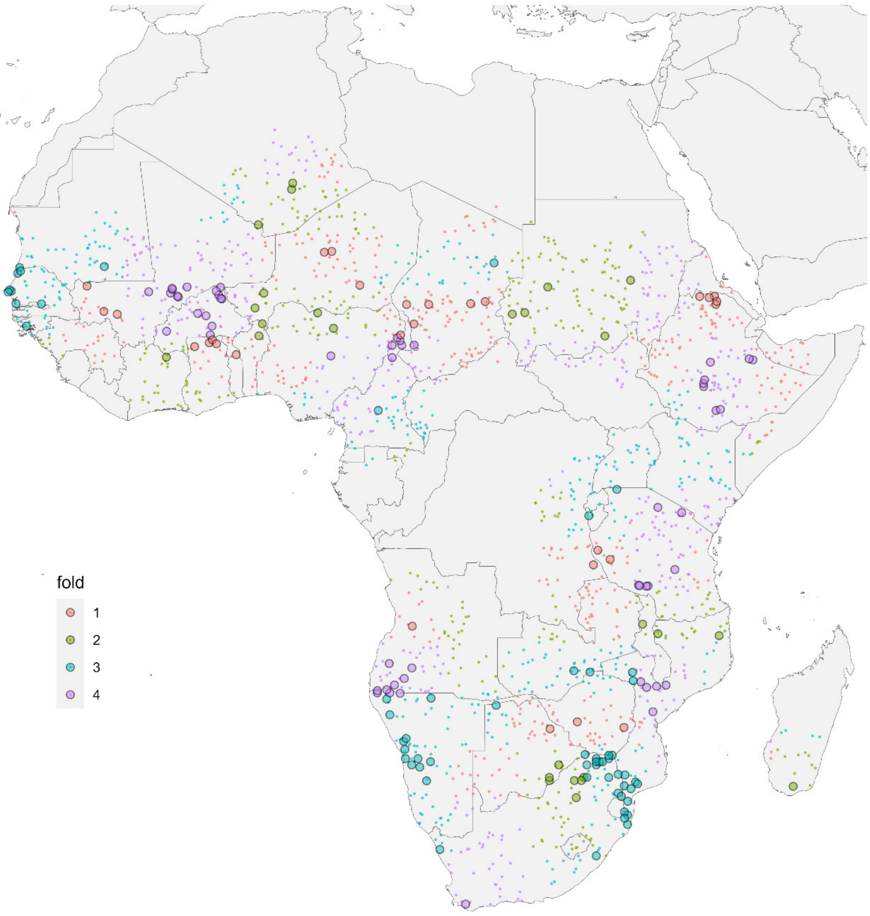

Figure 6. Occurrence and background locations for each species were assigned to four spatial folds with a size of 500 km. Larger symbols show occurrence observations, smaller symbols background locations, with the colour representing the fold. This figure shows the same locations for *Faidherbia albida* as for Figure 5.

In this section, we explain how we applied a spatial folding scheme to group presence and background locations for subsequent cross-validations in the generation of species distribution maps.

Occurrence and background locations were assigned to four spatial folds via the ***BiodiversityR::ensemble.spatialBlock*** function that internally calls the ***blockCV::spatialBlock*** function (Valavi et al. 2018). In fourfold model cross-validation, locations from each single fold are used for the evaluation of models calibrated with locations from the other three folds (see Box 1 in Valavi et al. 2018). As argued by Valavi et al., spatial folding methods are preferred to conventional random techniques of cross-validation since the latter can lead to underestimation of prediction error.

We set the size of the folds to 500 km (argument ***theRange*** = 500000), the same size as the magnitude of the circular buffer used to select background locations (see Section 8). In a first run for each species, we set the minimum number of locations in each fold (argument ***numLimit***) to 20. For species where this target could not be achieved, we reduced the minimum number of locations to 10. In a final run for species where the target of 10 could not be reached, we set the minimum target number to 5. As we judged 5 locations to be a minimum to evaluate model calibrations in cross-validation tests, species where this target could not be met were excluded from model calibrations (26 species in Table 9 with minimum occurrences in a fold < 5). As a consequence, the final atlas we generated contains maps for 127 species, corresponding to the species in Table 9 with minimum occurrences in a fold ≥ 5.

Prior to applying the spatial folding function, occurrence and background locations were transformed to the equal-area Mollweide projection (https://spatialreference.org/ref/esri/53009/), as spatial folding requires equal-area coordinate reference systems (Valavi et al. 2018). After spatial folding, the locations were transformed back to their latitudes and longitudes (https://spatialreference.org/ref/epsg/4326/).

Table 9. Number of occurrence and background locations used for spatial folding. Minimum values are the minima among the four folds. Column SDM indicates whether a species distribution model was calibrated for a species, based on the minimum number of locations per fold to be 5 or larger**.** Species are ordered by 'Occurrence', 'Minimum'.

Species	Occurrence		Background		SDM
	All	Minimum	All	Minimum	
Ficus sycomorus	198	45	2011	496	YES
Kigelia africana	200	44	2031	468	YES
Jatropha curcas	191	43	1941	413	YES
Adansonia digitata	200	41	2022	478	YES
Vitex doniana	195	41	1981	476	YES
Acacia nilotica	200	40	2031	449	YES
Phoenix reclinata	196	40	1991	458	YES
Calotropis procera	200	38	2037	457	YES
Ekebergia capensis	200	38	2031	388	YES
Combretum collinum	197	38	2001	445	YES
Entada abyssinica	193	38	1961	389	YES
Balanites aegyptiaca	198	37	2011	434	YES
Borassus aethiopum	196	37	1983	479	YES
Parkinsonia aculeata	200	36	2037	451	YES
Ficus sur	199	36	2021	445	YES
Albizia gummifera	165	36	1681	357	YES
Flueggea virosa	199	35	2031	451	YES
Syzygium guineense	199	35	2031	465	YES
Strychnos spinosa	198	35	2017	481	YES
Acacia sieberiana	194	35	1970	450	YES
Ilex mitis	200	34	2031	418	YES
Combretum molle	198	34	2011	449	YES
Annona senegalensis	194	34	1969	456	YES
Ximenia americana	193	34	1970	465	YES
Milicia excelsa	192	34	1931	416	YES
Dodonaea viscosa	200	33	2031	383	YES
Celtis africana	199	33	2021	453	YES
Sesbania sesban	195	33	1981	446	YES
Cajanus cajan	192	33	1951	436	YES
Acacia tortilis	200	32	2038	462	YES

continued on next page

Table 9. Continued

Species	Occurrence		Background		SDM
Ziziphus mucronata	200	32	2031	384	YES
Senna didymobotrya	198	32	2010	430	YES
Diospyros mespiliformis	196	32	1991	446	YES
Bridelia micrantha	192	32	1951	440	YES
Vernonia amygdalina	184	32	1871	424	YES
Albizia lebbeck	200	31	2031	431	YES
Ceiba pentandra	200	31	2023	400	YES
Sarcocephalus latifolius	199	31	2020	444	YES
Sclerocarya birrea	196	30	1990	429	YES
Leucaena leucocephala	192	30	1951	405	YES
Salvadora persica	167	30	1700	388	YES
Commiphora africana	200	29	2030	463	YES
Tamarindus indica	191	29	1941	383	YES
Shirakiopsis elliptica	136	29	1391	307	YES
Olea europaea	200	28	2037	484	YES
Bauhinia thonningii	194	28	1971	465	YES
Grevillea robusta	192	28	1950	399	YES
Combretum aculeatum	170	28	1730	408	YES
Nuxia congesta	168	28	1711	358	YES
Olea capensis	146	28	1491	306	YES
Saba comorensis	143	28	1461	285	YES
Azadirachta indica	200	27	2031	443	YES
Delonix regia	200	27	2031	479	YES
Melia azedarach	200	27	2037	464	YES
Dichrostachys cinerea	198	27	2011	474	YES
Capparis tomentosa	168	27	1711	362	YES
Acacia seyal	195	26	1981	479	YES
Securidaca longipedunculata	173	26	1761	409	YES
Faidherbia albida	167	26	1701	363	YES
Dalbergia melanoxylon	144	26	1473	332	YES
Jacaranda mimosifolia	200	25	2030	437	YES
Stereospermum kunthianum	200	25	2031	376	YES
Trichilia emetica	197	25	1997	376	YES
Euphorbia tirucalli	190	25	1946	430	YES
Rhamnus prinoides	186	25	1891	372	YES
Steganotaenia araliacea	122	25	1250	243	YES
Acacia senegal	200	24	2030	395	YES
Acacia polyacantha	192	24	1951	396	YES
Dovyalis caffra	146	24	1490	273	YES
Flacourtia indica	200	23	2031	460	YES
Vitellaria paradoxa	200	23	2005	442	YES
Grewia villosa	103	23	1058	240	YES
Strychnos innocua	132	22	1349	295	YES

continued on next page

Table 9. Continued

Species	Occurrence		Background		SDM
Erythrina abyssinica	115	22	1181	270	YES
Anogeissus leiocarpa	197	21	1991	451	YES
Prunus africana	136	21	1391	294	YES
Searsia natalensis	127	21	1301	293	YES
Ziziphus jujuba	127	21	1299	274	YES
Garcinia livingstonei	118	21	1209	247	YES
Berchemia discolor	102	21	1047	192	YES
Antiaris toxicaria	158	20	1599	379	YES
Spathodea campanulata	126	20	1282	284	YES
Mangifera indica	115	20	1177	261	YES
Acacia saligna	100	20	1028	211	YES
Moringa oleifera	98	18	1008	207	YES
Afrocarpus falcatus	77	16	797	161	YES
Cordia africana	94	15	971	195	YES
Lawsonia inermis	90	14	945	201	YES
Juniperus procera	86	14	887	208	YES
Croton macrostachyus	86	13	891	181	YES
Gardenia volkensii	71	13	724	141	YES
Hyphaene thebaica	68	13	707	152	YES
Oxytenanthera abyssinica	92	12	944	229	YES
Acacia abyssinica	70	12	730	155	YES
Vangueria madagascariensis	66	12	690	154	YES
Coffea arabica	53	12	558	126	YES
Sesbania bispinosa	88	11	886	189	YES
Vepris nobilis	79	11	820	190	YES
Markhamia lutea	69	11	701	154	YES
Polyscias fulva	68	11	710	153	YES
Strychnos henningsii	64	11	647	124	YES
Pterolobium stellatum	63	11	660	151	YES
Pinus patula	59	11	612	144	YES
Catha edulis	57	11	600	110	YES
Carica papaya	54	11	562	112	YES
Persea americana	50	11	524	119	YES
Terminalia brownii	50	11	530	106	YES
Ziziphus spina-christi	50	11	520	107	YES
Acacia melanoxylon	54	10	574	103	YES
Schinus molle	52	10	553	122	YES
Grewia damine	50	10	530	127	YES
Hagenia abyssinica	49	9	520	104	YES
Eucalyptus camaldulensis	48	8	481	91	YES
Boswellia neglecta	49	7	517	94	YES
Pouteria adolfi-friedericii	49	7	520	112	YES
Acacia lahai	48	7	507	103	YES

continued on next page

Table 9. Continued

Species	Occurrence		Background		SDM
Casuarina equisetifolia	37	7	421	86	YES
Warburgia ugandensis	34	7	420	83	YES
Yushania alpina	33	7	431	89	YES
Dovyalis abyssinica	49	6	520	89	YES
Acacia decurrens	28	6	403	76	YES
Albizia schimperiana	49	5	520	77	YES
Dombeya torrida	49	5	520	97	YES
Schefflera abyssinica	44	5	471	81	YES
Boswellia papyrifera	43	5	452	84	YES
Dobera glabra	21	5	427	79	YES
Cupressus sempervirens	20	5	407	66	YES
Boswellia rivae	18	2	418	73	NO
Casuarina cunninghamiana	15	2	422	76	NO
Erythrina brucei	33	1	420	49	NO
Cupressus lusitanica	14	1	420	72	NO
Eucalyptus globulus	11	1	424	54	NO
Ficus carica	11	1	427	58	NO
Callistemon citrinus	10	1	410	59	NO
Eucalyptus grandis	9	1	403	74	NO
Corymbia citriodora	7	1	403	73	NO
Albizia grandibracteata	53	0	557	111	NO
Millettia ferruginea	46	0	480	53	NO
Commiphora myrrha	20	0	420	48	NO
Tamarix aphylla	19	0	424	89	NO
Maerua aethiopica	10	0	427	95	NO
Citrus sinensis	8	0	425	64	NO
Commiphora guidottii	8	0	408	81	NO
Boswellia microphylla	7	0	411	72	NO
Boswellia pirottae	6	0	420	76	NO
Calliandra calothyrsus	5	0	411	86	NO
Moringa stenopetala	5	0	409	67	NO
Boswellia ogadensis	4	0	268	25	NO
Eucalyptus saligna	4	0	407	75	NO
Eucalyptus viminalis	4	0	400	59	NO
Cordeauxia edulis	3	0	203	29	NO
Cytisus proliferus	3	0	391	71	NO
Malus domestica	0				NO

10 Calibration of species distribution models and generation of suitability maps

In this section, we explain how we calibrated the species distribution models and generated the suitability maps.

Species distribution models were calibrated via the functions ***BiodiversityR::ensemble.calibrate.weights*** and ***BiodiversityR::ensemble.calibrate.models*** via procedures of ensemble suitability modelling described by Kindt (2018b). Similar procedures were used in other species distribution studies with *BiodiversityR*, such as those by Ranjitkar et al. (2014a,b), de Sousa et al. (2019), Fremout et al. (2020) and van Zonneveld et al. (2020). The ensemble procedures of *BiodiversityR* calculate habitat suitability as a weighted average of predictions from different algorithms,[17] an approach that may significantly increase model performance (Marmion et al. 2009, Hao et al. 2019).

Table 10 provides critical argument settings for the ***BiodiversityR::ensemble.calibrate.weights*** function that performed the fourfold cross-validation tests (see previous section and Box 1 in Valavi et al. 2018) to calculate the weights[18] for each of the considered algorithms in contributing to the ensemble suitability.

Model evaluation statistics were calculated via ***BiodiversityR::ensemble.evaluate*** and included the Area Under the receiver-operator Curve (AUC, e.g. Hijmans 2012, Castellanos et al. 2019, Grimmet et al. 2020), the Symmetric Extremal Dependence Index (SEDI, Wunderlich et al. 2019) and the True Skill Statistic[19] (TSS, Allouche et al. 2006). Although the use of AUC has been criticized (e.g., by Jimenez-Valverde 2011), it provides a valid measure of relative model performance for the same species and study area (Wisz et al. 2008), and therefore it is also valid for comparing the performance of a consensus model and the individual performance of contributing algorithms (Kindt 2018b).

Table 10. Argument settings for the *BiodiversityR::ensemble.calibrate.weights* function.

Argument	Setting	Comment
ENSEMBLE.tune	TRUE	Determines the optimal combination of options from ***ENSEMBLE.exponent*** and ***ENSEMBLE.min***
ENSEMBLE.min	c(0.55, 0.60, 0.65)	Minimum AUC for an algorithm to be included in the ensemble. See Kindt (2018b) for details.
ENSEMBLE.exponent	c(1, 2, 3)	Exponent applied to algorithm AUC values to convert these into weights. See Kindt (2018b) for details.
MAXENT	0	Do not fit a maximum entropy model via ***dismo::maxent***. Not fitted as maximum entropy model obtained via MAXNET.
MAXNET	1	Fit a maximum entropy model via ***maxnet::maxnet*** (Philips et al. 2017)

continued on next page

17 These methods have also been described as 'consensus methods', where a relevant combination of several model outputs results in a prediction that has higher accuracy than those of individual model outputs; similar consensus methods are used in meteorology, climatology and economics (Marmion 2009). For an online example, check https://rspatial.org/raster/sdm/6_sdm_methods.html#combining-model-predictions.

18 See Kindt (2018b) for how weights are calculated from AUC values. Weights can be zero so that results from algorithms with weight zero are not included in the calculations. When only one weight is larger than zero, the ensemble predictions are equal to those of the selected (best) model. When it is not known *a priori* which algorithm has the best predictions for a particular species or study region, selection of the single best algorithm still is a method of ensemble suitability modelling.

19 In this working paper, we do not report results for the TSS, as Wunderlich et al. (2019) demonstrated that SEDI has superior qualities as an evaluation metric.

Table 10. Continued

Argument	Setting	Comment
MAXLIKE	1	Fit a maximum likelihood model via ***maxlike::maxlike*** (see Kindt 2018b for citation)
GBM	1	Fit a boosted regression trees model via ***gbm::gbm*** (see Kindt 2018b for citation)
GBMSTEP	1	Fit a stepwise boosted regression trees model via ***dismo::gbm.step*** (see Kindt 2018b for citation)
RF	1	Fit a random forest model via ***randomForest::randomForest*** (see Kindt 2018b for citation)
CF	0	Do not fit a random forest model via ***party::cforest***. Not fitted as calibration and projections consume a lot of time.
GLM	1	Fit a generalized linear model via ***stats::glm*** (see Kindt 2018b for citation)
GLMSTEP	1	Fit a stepwise generalized linear model via ***MASS::stepAIC*** (see Kindt 2018b for citation)
GAM	1	Fit a generalized additive model via ***gam::gam*** (see Kindt 2018b for citation)
GAMSTEP	1	Fit a stepwise generalized additive model via ***gam::step.Gam*** (see Kindt 2018b for citation)
MGCV	1	Fit a generalized additive model via ***mgcv::gam*** (see Kindt 2018b for citation)
MGCFIX	0	Do not fit a generalized additive model with fixed d.f. regression splines
EARTH	1	Fit a multivariate adaptive regression spline model via ***earth::earth*** (see Kindt 2018b for citation)
RPART	1	Fit a recursive partitioning and regression tree model via ***rpart::rpart*** (see Kindt 2018b for citation)
NNET	1	Fit an artificial neural network via ***nnet::nnet*** (see Kindt 2018b for citation)
FDA	1	Fit a flexible discriminant analysis model via ***mda::fda*** (see Kindt 2018b for citation)
SVM	1	Fit a support vector machine model via ***kernlab::ksvm*** (see Kindt 2018b for citation)
SVME	1	Fit a support vector machine model via ***stats::glm*** (see Kindt 2018b for citation)
GLMNET	0	Do not fit a generalized linear model with lasso or elasticnet regularization. Not fitted as calibration and projections consume a lot of time.
BIOCLIM.0	0	Do not fit the original BIOCLIM algorithm. Not fitted as the alternative implementation was done via BIOCLIM.
BIOCLIM	1	Fit a BIOCLIM model via ***dismo::bioclim*** (see Kindt 2018b for citation)
DOMAIN	0	Do not fit a model via the DOMAIN algorithm. Not fitted as calibration and projections consume a lot of time, and as this algorithm typically does not perform as well as other algorithms (Wisz et al. 2008b)
MAHAL	0	Do not fit models via the Mahalanobis algorithm. Not fitted as calibrations and projections consume a lot of time.
MAHAL01	0	Do not fit models via the Mahalanobis algorithm. Not fitted as calibrations and projections consume a lot of time.
PROBIT	TRUE	Transform suitability predictions from each algorithm to probabilities via a probit transformation. See Kindt (2018b) for details.

As some machine-learning algorithms use randomization approaches, we carried out model calibration procedures 5 times[20] for each species. Based on their average AUC in the fourfold cross-validations (AUC-mean), from the 5 calibrated ensemble models for each species we selected the ensemble model with the highest AUC-mean (see Table 11) to generate habitat suitability maps for the baseline and future climates (GCMs × scenarios) via ***BiodiversityR::ensemble.raster***.

Model performance statistics shown in Table 11 correspond to the selected ensemble model. In the majority of cases (102 species when comparing AUC-mean statistics), the ensemble model outperformed individual algorithms. In cases where an individual algorithm had higher AUC-mean than the ensemble model, GLMSTEP ranked first for 7 species, and MAXNET and GAMSTEP ranked first for 6 species each. There were only 4 cases where the ensemble model ranked third and only 2 cases where it ranked fourth, but for all these cases the final model had an AUC-final value above 90%. Only for *Catha edulis* was the AUC-mean difference between the best ranking algorithm and the ensemble model larger than 2%.

Elsewhere, prediction accuracies with AUC-mean of more than 90% are classed as excellent,[21] from 90% to 80% as good, from 80% to 70% as fair, from 70% to 60% as poor,[22] and below 60% as a fail. On this basis, in the current analysis no species failed for cross-validation, and the model was considered to be poor for 5 species only: *Grewia damine*, *Ziziphus mucronata*, *Adansonia digitata*, *Combretum aculeatum* and *Hyphaene thebaica* (in descending order of AUC-mean value; Table 11). Taking the same classification thresholds, the minimum AUC in a fold (AUC-min) was classified as a fail for 3 species, *Ziziphus mucronata*, *Commiphora africana* and *Garcinia livingstonei*; and as poor for 39 species, ranging from 69.9% for *Anogeissus leiocarpa* to 60.7% for *Entada abyssinica*. The prediction accuracies for the final models that were fitted with the full set of occurrence and background locations were all classed as good or excellent, with *Commiphora africana* ranked lowest with a value of 82% for AUC-final.

Table 11. Evaluation statistics for the selected ensemble models used to generate species distribution maps. AUC-mean: the mean AUC percentage value over the fourfold cross-validations, used to order entries in the table; AUC-min: the minimum AUC over the fourfold cross-validations; Rank: the rank of the ensemble model, based on AUC-mean, when compared with individual algorithms; AUC-final: the AUC percentage value for the model calibrated with all observations; SEDI-final: the SEDI percentage value for the model calibrated with all observations.

Species	AUC-mean	AUC-min	Rank	AUC-final	SEDI-final
Yushania alpina	95.901%	85.768%	1	99.543%	97.954%
Pouteria adolfi-friedericii	93.913%	86.391%	4	99.059%	97.640%
Ilex mitis	92.857%	90.509%	3	96.573%	91.718%
Dombeya torrida	92.510%	90.309%	1	98.528%	97.192%
Schefflera abyssinica	92.095%	83.940%	1	98.113%	95.741%
Juniperus procera	92.081%	86.903%	1	96.259%	94.616%
Hagenia abyssinica	91.999%	87.574%	1	99.031%	95.956%
Rhamnus prinoides	91.331%	86.473%	1	95.058%	91.356%
Prunus africana	91.082%	89.102%	1	94.260%	92.855%
Acacia saligna	90.125%	88.671%	1	97.313%	93.584%
Acacia decurrens	89.988%	81.699%	1	99.369%	97.531%

continued on next page

20 Including the fourfold cross-validation runs, this means that for each species 20 cross-validation models were calibrated for each algorithm. These were followed by 5 final model calibrations ('ensemble models') with the full set of occurrence and background observations. From these final 5 ensemble models, we selected the ensemble model with the highest AUC-mean.

21 These thresholds are taken from an Operating Manual for BIOMOD from 2009. For SEDI, we used the same thresholds.

22 Hijmans (2012) mentions that an AUC threshold of 0.7 is often used to identify "good" models.

Table 11. Continued

Species	AUC-mean	AUC-min	Rank	AUC-final	SEDI-final
Acacia melanoxylon	89.608%	80.874%	2	97.677%	93.828%
Nuxia congesta	89.510%	86.560%	1	94.491%	87.248%
Dodonaea viscosa	89.246%	86.128%	1	94.131%	86.681%
Dovyalis abyssinica	89.181%	85.569%	1	96.762%	95.035%
Olea europaea	88.828%	80.846%	1	93.414%	92.094%
Olea capensis	88.219%	85.051%	2	96.167%	87.815%
Coffea arabica	87.919%	83.816%	1	96.789%	90.656%
Polyscias fulva	87.536%	84.789%	1	98.293%	94.082%
Acacia abyssinica	87.210%	82.698%	2	94.683%	92.284%
Albizia gummifera	86.385%	85.421%	1	91.953%	84.699%
Warburgia ugandensis	85.322%	78.795%	1	98.866%	95.991%
Afrocarpus falcatus	85.313%	78.093%	1	95.511%	90.277%
Acacia lahai	85.110%	81.884%	2	96.857%	94.026%
Celtis africana	85.104%	80.770%	1	90.857%	86.438%
Boswellia papyrifera	84.604%	73.038%	1	97.235%	92.681%
Albizia schimperiana	84.591%	81.515%	1	97.233%	95.041%
Shirakiopsis elliptica	84.376%	82.647%	1	91.114%	87.644%
Senna didymobotrya	84.366%	82.084%	1	91.846%	86.626%
Cordia africana	84.346%	78.087%	2	92.232%	86.974%
Ekebergia capensis	84.342%	78.910%	1	89.230%	86.537%
Markhamia lutea	84.068%	81.110%	1	95.326%	92.678%
Casuarina equisetifolia	83.606%	77.713%	1	97.991%	94.368%
Catha edulis	83.596%	76.925%	4	95.474%	90.279%
Croton macrostachyus	83.521%	78.485%	3	95.406%	92.372%
Terminalia brownii	83.453%	71.437%	1	97.332%	93.331%
Pterolobium stellatum	83.296%	79.335%	1	94.382%	93.471%
Vitellaria paradoxa	82.736%	77.058%	1	92.716%	89.087%
Ceiba pentandra	82.426%	75.212%	1	89.771%	81.954%
Searsia natalensis	82.326%	75.446%	1	91.664%	88.272%
Parkinsonia aculeata	82.301%	74.034%	1	91.122%	86.063%
Cupressus sempervirens	82.225%	62.692%	2	96.878%	95.378%
Grevillea robusta	82.222%	71.482%	1	92.301%	84.533%
Dobera glabra	82.001%	73.400%	1	97.502%	93.795%
Ficus sur	81.928%	77.965%	1	88.543%	83.704%
Schinus molle	81.788%	72.854%	1	96.346%	90.864%
Bridelia micrantha	81.575%	75.626%	1	88.709%	83.693%
Ziziphus spina-christi	81.575%	77.995%	1	94.019%	88.497%
Azadirachta indica	81.356%	78.672%	2	90.753%	77.444%
Milicia excelsa	81.268%	77.780%	2	90.106%	86.705%
Spathodea campanulata	81.263%	76.517%	1	89.937%	84.240%
Combretum molle	81.069%	76.017%	1	87.948%	86.203%
Vepris nobilis	81.062%	70.567%	1	92.877%	88.928%
Sarcocephalus latifolius	80.981%	75.278%	1	88.552%	87.735%

continued on next page

Table 11. Continued

Species	AUC-mean	AUC-min	Rank	AUC-final	SEDI-final
Salvadora persica	80.726%	75.034%	1	90.242%	80.832%
Phoenix reclinata	80.726%	75.641%	1	88.569%	84.701%
Euphorbia tirucalli	80.634%	76.096%	3	90.420%	88.065%
Syzygium guineense	80.375%	79.141%	1	89.865%	85.842%
Strychnos henningsii	80.155%	75.509%	2	96.839%	83.111%
Delonix regia	79.994%	76.179%	2	90.768%	84.178%
Dovyalis caffra	79.993%	74.989%	3	92.752%	86.494%
Borassus aethiopum	79.810%	74.470%	2	93.462%	81.520%
Gardenia volkensii	79.805%	66.162%	1	93.794%	85.139%
Sclerocarya birrea	79.586%	75.107%	2	87.626%	78.063%
Melia azedarach	79.203%	74.282%	1	88.811%	81.344%
Jacaranda mimosifolia	79.123%	72.312%	1	91.934%	86.541%
Calotropis procera	78.355%	66.753%	1	90.908%	81.285%
Erythrina abyssinica	78.274%	73.474%	1	89.941%	86.579%
Annona senegalensis	77.637%	65.726%	1	89.218%	77.782%
Capparis tomentosa	77.585%	66.237%	1	87.770%	82.118%
Vangueria madagascariensis	77.429%	71.948%	2	95.362%	88.838%
Pinus patula	77.319%	65.759%	2	93.411%	85.310%
Strychnos innocua	77.300%	72.289%	1	89.961%	80.308%
Dichrostachys cinerea	77.277%	72.002%	1	89.001%	78.842%
Tamarindus indica	77.271%	71.814%	1	88.624%	82.023%
Saba comorensis	77.266%	75.021%	1	88.766%	87.223%
Antiaris toxicaria	77.051%	74.604%	1	94.009%	87.527%
Strychnos spinosa	76.647%	72.860%	2	88.385%	83.960%
Cajanus cajan	76.361%	68.623%	1	84.837%	77.031%
Kigelia africana	76.019%	73.205%	1	85.195%	75.626%
Diospyros mespiliformis	76.008%	70.081%	1	87.415%	85.057%
Oxytenanthera abyssinica	75.914%	74.371%	1	88.853%	82.636%
Eucalyptus camaldulensis	75.879%	73.138%	1	94.932%	86.851%
Ziziphus jujuba	75.743%	71.667%	1	88.270%	77.257%
Acacia nilotica	75.738%	73.213%	1	89.651%	78.765%
Persea americana	75.722%	66.590%	2	92.958%	91.337%
Steganotaenia araliacea	75.625%	71.318%	1	88.455%	82.837%
Anogeissus leiocarpa	75.513%	69.920%	1	89.418%	85.274%
Acacia seyal	75.450%	72.258%	1	91.279%	75.608%
Berchemia discolor	75.377%	61.864%	1	96.005%	89.038%
Trichilia emetica	74.960%	73.501%	2	89.915%	83.142%
Vitex doniana	74.877%	69.149%	1	84.575%	83.169%
Securidaca longipedunculata	74.807%	68.074%	1	87.983%	88.187%
Boswellia neglecta	74.612%	65.565%	1	90.424%	85.909%
Flueggea virosa	74.566%	67.150%	1	86.108%	76.209%
Leucaena leucocephala	74.412%	69.705%	1	88.702%	81.157%
Balanites aegyptiaca	74.391%	72.645%	1	87.256%	77.289%

continued on next page

Table 11. Continued

Species	AUC-mean	AUC-min	Rank	AUC-final	SEDI-final
Vernonia amygdalina	74.314%	67.829%	1	85.068%	83.952%
Flacourtia indica	74.178%	65.577%	1	87.194%	81.934%
Acacia polyacantha	74.073%	70.104%	1	83.869%	84.585%
Ficus sycomorus	73.573%	67.038%	1	84.840%	81.286%
Carica papaya	73.567%	65.761%	1	92.372%	86.681%
Sesbania sesban	73.491%	61.065%	1	84.178%	70.209%
Bauhinia thonningii	73.454%	72.154%	1	85.054%	81.677%
Dalbergia melanoxylon	73.308%	65.872%	1	88.100%	79.298%
Sesbania bispinosa	72.868%	66.814%	1	88.325%	78.164%
Lawsonia inermis	72.671%	70.150%	1	89.504%	83.966%
Combretum collinum	72.422%	66.586%	2	83.553%	84.150%
Entada abyssinica	72.398%	60.702%	1	87.685%	82.912%
Faidherbia albida	72.201%	64.193%	1	86.193%	81.070%
Acacia tortilis	72.001%	63.421%	1	85.184%	76.900%
Grewia villosa	71.922%	67.260%	1	88.550%	79.386%
Acacia sieberiana	71.715%	68.201%	1	91.491%	79.332%
Acacia senegal	71.674%	66.404%	1	82.811%	81.442%
Ximenia americana	71.606%	64.783%	1	87.034%	79.593%
Stereospermum kunthianum	71.486%	61.479%	1	88.607%	77.052%
Albizia lebbeck	70.977%	66.627%	1	87.601%	81.424%
Garcinia livingstonei	70.827%	57.391%	1	86.723%	80.516%
Jatropha curcas	70.718%	65.446%	1	88.781%	76.836%
Moringa oleifera	70.686%	61.770%	2	86.808%	77.122%
Mangifera indica	70.423%	62.064%	1	92.769%	79.963%
Commiphora africana	70.205%	58.099%	1	82.058%	79.256%
Grewia damine	69.729%	61.908%	1	90.374%	83.454%
Ziziphus mucronata	69.461%	58.630%	1	88.712%	79.126%
Adansonia digitata	69.411%	66.058%	1	90.830%	75.315%
Combretum aculeatum	68.815%	67.122%	1	92.950%	66.390%
Hyphaene thebaica	67.989%	62.291%	1	84.893%	77.200%

The ***BiodiversityR::ensemble.raster*** function generates three types of habitat suitability maps (Kindt 2018b) that are all included in the atlas:

- **Predicted presence maps** depict areas where a species is predicted to be suitable (present) or not suitable (absent) as predicted by a particular ensemble model.
- **Predicted suitability maps** depict the probability that a species is suitable across the mapped area as predicted by a particular ensemble model.
- **Count suitability maps** show how many of the algorithms that are used by a particular ensemble model predict that a species is suitable (consensus map).

Predicted presence maps are derived from predicted suitability maps by applying a suitability threshold value that discriminates areas where a species is predicted to be present (above the threshold) or absent (below the threshold). Various approaches exist to calculate the threshold; based on these we applied a threshold that maximizes the sum of sensitivity and specificity, as recommended by Liu et al. (2013, 2016). This threshold has been widely used by a range of authors in species suitability investigations (e.g., de Sousa et al. 2019, Grimmet et al. 2020, Ramirez-Villegas et al. 2020, Sillero et al. 2021) and is the default argument setting for functions ***BiodiversityR::ensemble.calibrate.weights*** and ***BiodiversityR::ensemble.calibrate.models***.

For the predicted suitability maps, the species absence-presence threshold was used in combination with maximum suitability values to classify species suitability maps to four quartile ranges above the threshold using function ***raster::quantile*** with argument ***probs=c(0, 0.25, 0.5, 0.75, 1.0)***. Below the threshold and with minimum suitability values, we calculated a 90% percentile value to show some of the areas (those of highest suitability) where the species was predicted not to be suitable.

The ***BiodiversityR::ensemble.raster*** function also generates count suitability maps that show the number of algorithms that predict species presence (Kindt 2018b). These types of maps can be used to investigate consensus among the different algorithms used to calculate habitat suitability by the ensemble model.

Water bodies were masked out in the generation of maps by applying the inland and ocean water layer of Lamarche et al. (2017). We reprojected ('warped') this layer to the resolution of our predictor variables and the suitability maps in QGIS, using a quantile method whereby a raster was classed as a water body when at least 25% was covered by water.

11 Discrimination of areas with novel environmental conditions

In this section, we explain how we separately mapped areas with novel environmental conditions. An argument can be made that species should be predicted not to be suitable (should be predicted absent) in areas with novel environmental conditions (conditions outside the observed environmental range of the species).

Identifying areas with novel conditions is conceptually related to the multivariate environmental similarity surface (MESS) methodology developed by Elith et al. (2010), but for the atlas we use a binary classification of novel versus not-novel environments, where these had negative MESS values and values within the observed range, respectively.

The observed environmental ranges of each tree species for our 43 environmental variables were obtained via the ***BiodiversityR::ensemble.novel.object*** function. In two separate exercises, one using our predictor variables and one using the full set of 43 environmental variables, we created maps for each species that showed areas where there are novel conditions[23] within the predicted habitat.

The maps for novel conditions allow discrimination between novel conditions for variables not used for calibrating the model ('Extrapolated #1'); and novel conditions for variables that were used for calibrating the model ('Extrapolated #2').

23 The phenomenon whereby the ensemble suitability model predicts that a species is suitable for novel environmental conditions can be described as 'extrapolation', the opposite of 'interpolation' (e.g., https://www.dictionary.com/e/interpolation-vs-extrapolation/).

12 Generation of habitat change maps

In this section, we explain how we generated habitat change maps in the online atlas.

Habitat change maps were created by comparing the predicted presence of a species in baseline climate with the predicted presence in future climates, projected for each of the nine GCMs. Separate maps were created for our two chosen climate change scenarios, and for when areas with novel conditions were excluded (masked out, see Section 11) or included.

Classifications of habitat change were based on the likelihood scale developed for the 5th Assessment Report of the IPCC by Mastrandea et al. (2011). With this scale, 66% to 90% probability is classified as 'likely' and 90% to 100% probability as 'very likely' (Table 12). The same scale was used to investigate habitat change in a climate change atlas for Central America that some of the current authors contributed to (de Sousa et al. 2019).

Table 12. Rules to classify habitat change.

Baseline	GCMs	Habitat change	Comment
Suitable	0	Lost	Very likely as > 90% of GCMs predict habitat loss
Suitable	1 - 3	Lost	Likely as ≥ 66% of GCMs predict habitat loss
Suitable	4 - 5	Uncertain	
Suitable	6 - 9	Kept	Likely as ≥ 66% of GCMs predict habitat kept
Not suitable	6 - 9	New	Likely as ≥ 66% of GCMs predict new habitat

13 Convex hulls for an *a posteriori* distance constraining method

In this section, we explain how we generated *a posteriori* distance constraining hulls. For each species, the hulls are used to show areas in the modelled distribution (outside the hull) where the conditions for the species are predicted to be suitable, but that are distant from known presence observations of the species. Areas outside hulls are expected not to be reachable by natural dispersal processes and to be outside the natural range of a species by consequence.

Maps in the online atlas include a convex hull that was created via the ***BiodiversityR::ensemble.chull.create*** function. All spatially thinned occurrence observations for a species were used to create this hull (Figures 7 and 8). The argument setting of ***buffer.maxmins = TRUE*** selected the Buffered Minimum Convex Polygon (BMCP) as an *a posteriori* distance constraining method, as described by Mendes et al. (2020).

Mendes et al. (2020) investigated various methods of adding distance constraints to reduce overprediction in species distribution modelling. These methods include dispersal constraints that essentially exclude areas unlikely to have been colonized by a species. The BMCP algorithm was among the *a posteriori* methods that reduced overprediction without incurring high omission errors. For each species, the buffer width for BMCP corresponds to the maximum calculated from the distances to the nearest neighbour for each occurrence location.

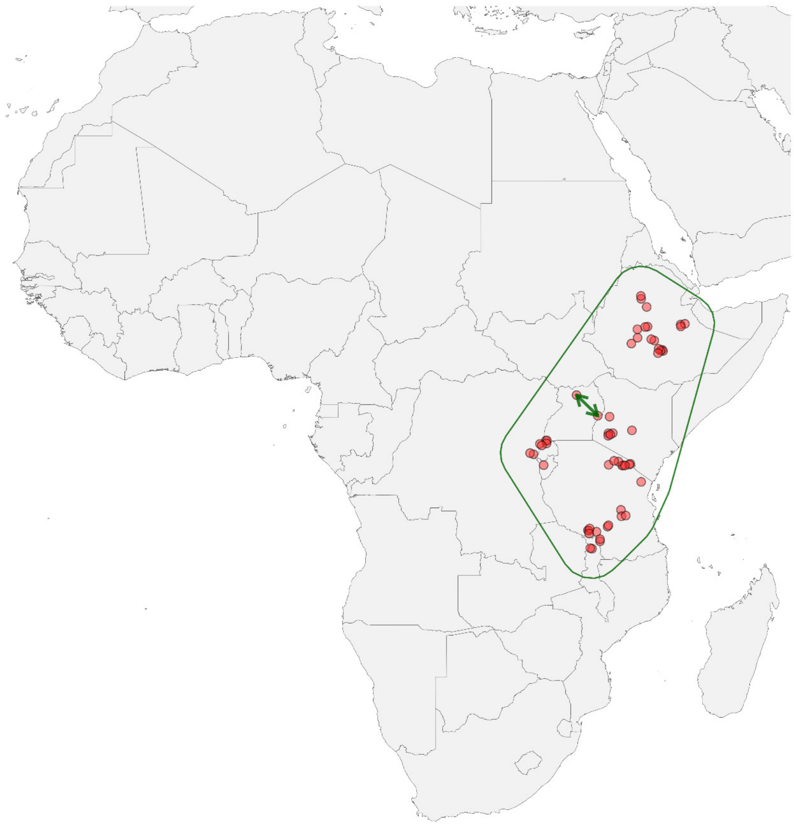

Figure 7. Convex hull for a hypothetical tree species with a relatively narrow range across Africa. The arrowed line shows the buffer width distance and the two locations used to calculate this distance. Red circles show the full set of spatially thinned occurrence observations for the species.

Convex hulls were added to our maps as a visual aid for users to identify areas where it is unlikely that species occur naturally. Note that we have not used convex hulls as masks for the area calculations given in the atlas. The area statistics shown, calculated after reprojecting the rasters to the equal-area Mollweide projection (https://spatialreference.org/ref/esri/53009/), include areas outside the hulls.

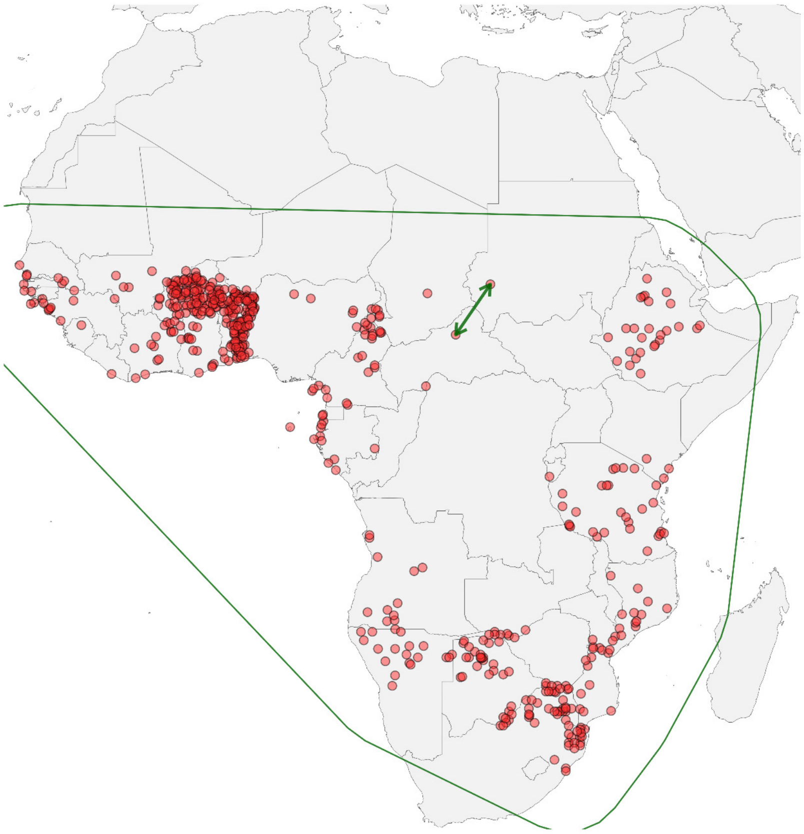

Figure 8. Convex hull for a hypothetical tree species with a wide range across Africa. The arrowed line shows the buffer width distance and the two locations used to calculate this distance. Red circles show the full set of spatially thinned occurrence observations for the species.

14 Base map and map annotations used in the atlas

In this section, we explain how the main atlas maps were created.

The base (continental) map used throughout the atlas was obtained from https://maps.wikimedia.org via the *basemapR* package (version 0.1.0; Bailey 2020; accessed 23 February 2021).

North arrows and the scale bar were added using ***ggspatial::annotation_north_arrow*** and ***ggspatial::annotation_scale*** functions, respectively (*ggspatial* version 1.1.5; Dunnington 2021), to maps created via the *ggplot2* package (version 3.3.3; Wickham 2016).

For the species presence maps in the baseline climate, spatially thinned occurrence observations from only the RAINBIO database (Dauby et al. 2016), one of the databases used to compile occurrence observations (see Section 5), were added. Spatial thinning was done via ***BiodiversityR::ensemble.spatial.thin*** for occurrences closer than 100 km. This was carried out for graphical reasons, to avoid the overlap of occurrence symbols and increase the visibility of the habitat suitability layer.

Our reason to select RAINBIO occurrences only for visual purposes was based on the numerous quality checks, including manual checks by African flora experts, undertaken while georeferencing these occurrences. The inclusion of occurrences from RAINBIO in the maps therefore provides users with a reliable check for the modelled distributions. However, as RAINBIO is geographically focused on mainland Africa, and especially to areas south of the Sahel and north of southern Africa, the distribution of occurrences from RAINBIO should not be used to inspect the reliability of the models outside the area that is thereby defined.[24]

The spatially thinned RAINBIO locations served as inputs for the creation of a concave hull (different from the convex *a posteriori* distance constraining hulls described in Section 13) via the ***ggforce::geom_mark_hull*** function (version 0.3.2; Pedersen 2020). Default parameter settings were used, except for ***concavity = 1.5***. These concave hulls were added to maps to assist the user in locating the spatially thinned observations from the RAINBIO database.

24 Inset maps in the atlas from *Plants of the World Online* help counter this problem of limited geographic range (see Section 15 of this working paper).

15 Inset maps showing country distribution of Plants of the World Online

In this section, we explain how the inset atlas maps were created.

For each of the tree species mapped in the atlas, we compiled country distributions from Plants of the World Online (POWO; http://powo.science.kew.org/; accessed 4 September 2021) using the classifications of 'Native', 'Introduced' and 'Doubtful'. Countries that were not listed by POWO for a species were classified as 'Absent'.

An inset map was created with the Natural Earth (NE) 1:10 million Admin0 shapefile (version 4.1.0, downloaded May 2018 via https://www.naturalearthdata.com/downloads/10m-cultural-vectors/10m-admin-0-details/). Insets were only shown for baseline climate maps, to avoid any misunderstanding that POWO would provide future climate projections.

To match country names from POWO with country names in the NE shapefile, a lookup table was created with past and current names where these were different for countries (e.g., Democratic Republic of the Congo versus Zaïre; and Eswatini versus Swaziland). Whereas POWO mapped the Cape Provinces, Free State, KwaZulu-Natal and Northern Provinces separately for South Africa, the inset maps show information aggregated for South Africa. Other areas that were mapped separately by POWO were Cabinda (mapped in the inset map as Angola), the Gulf of Guinea Islands (mapped in the inset map as Equatorial Guinea) and the Caprivi Strip (mapped in the inset map as Namibia). As POWO does not separately list the new countries of South Sudan and Somaliland, the data for these nations were contained within the POWO data for Sudan and Somalia, respectively.

Similar to the visualization of full convex hulls (see Section 13), and of the occurrence locations as well as the concave hulls derived from RAINBIO (see Section 14), the inset maps can aid users to visually check the performance of suitability models.

References

Aerts R, Van Overtveld K, November E, Wassie A, Abiyu A, Demissew S, Daye DD, Giday K, Haile M, TewoldeBerhan S, et al. 2016. Conservation of the Ethiopian church forests: Threats, opportunities and implications for their management. *Science of the Total Environment* 551–552:404–414. https://doi.org/10.1016/j.scitotenv.2016.02.034

Aiello-Lammens ME, Boria RA, Radosavljevic A, Vilela B, Anderson RP. 2015. spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* 38:541–545. https://doi.org/10.1111/ecog.01132

Allouche O, Tsoar A, Kadmon R. 2006. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (Tss). *Journal of Applied Ecology* 43(6):1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Bailey C. 2020. basemapR: Base maps for ggplot2. R package version 0.1.0. https://github.com/Chrisjb/basemapR

Bekele-Tesemma A. 2007. Useful trees and shrubs of Ethiopia: Identification, propagation and management in 17 agroclimatic zones. Nairobi, Kenya: World Agroforestry. Accessed in 2018 via http://www.worldagroforestry.org/usefultrees

Booth TH. 2018. Species distribution modelling tools and databases to assist managing forests under climate change. *Forest Ecology and Management* 430:196–203. https://doi.org/10.1016/j.foreco.2018.08.019

Booth TH. 2016. Estimating potential range and hence climatic adaptability in selected tree species. *Forest Ecology and Management* 366:175–183. https://doi.org/10.1016/j.foreco.2016.02.009

Booth TH, Nix HA, Busby JR, Hutchinson MF. 2014. BIOCLIM: The first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Diversity and Distributions* 20(1):1–9. https://doi.org/10.1111/ddi.12144

Castellanos AA, Huntley JW, Voelker G, Lawing AM. 2019. Environmental filtering improves ecological niche models across multiple scales. *Methods in Ecology and Evolution* 10(4):481–492. https://doi.org/10.1111/2041-210X.13142

Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC, et al. 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology* 20(10):3177–3190. https://doi.org/10.1111/gcb.12629

Corlett RT and Tomlinson KW. 2020. Climate change and edaphic specialists: Irresistible force meets immovable object? *Trends in Ecology & Evolution*, 35(4):367–376. https://doi.org/10.1016/j.tree.2019.12.007

Dauby G, Zaiss R, Blach-Overgaard A, Catarino L, Damen T, Deblauwe V, Dessein S, Dransfield J, Droissart V, Duarte MC, et al. 2016. RAINBIO: A mega-database of tropical African vascular plants distributions. *PhytoKeys* 74:1–18. https://doi.org/10.3897/phytokeys.74.9723

de Sousa K, van Zonneveld M, Holmgren M, Kindt R, Ordoñez JC. 2019. The future of coffee and cocoa agroforestry in a warmer Mesoamerica. *Scientific Reports* 9(1):8828. https://doi.org/10.1038/s41598-019-45491-7

Do HTT, Zimmer HC, Vanclay JK, Grant JC, Trinh BN, Nguyen HH, Nichols JD. 2021. Site form classification—A practical tool for guiding site-specific tropical forest landscape restoration and management. *Forestry: An International Journal of Forest Research* 95(2):261–273. https://doi.org/10.1093/forestry/cpab046

Dunnington D. 2021. ggspatial: Spatial data framework for ggplot2. R package version 1.1.5. https://CRAN.R-project.org/package=ggspatial

Elith J, Kearney M, Phillips S. 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution* 1(4):330–342. https://doi.org/10.1111/j.2041-210X.2010.00036.x

Feeley KJ and Silman MR. 2011. Keep collecting: Accurate species distribution modelling requires more collections than previously thought. *Diversity and Distributions* 17(6):1132–1140. https://doi.org/10.1111/j.1472-4642.2011.00813.x

Fick SE and Hijmans RJ. 2017. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37(12):4302–4315. https://doi.org/10.1002/joc.5086

Fox J and Monette G. 1992. Generalized collinearity diagnostics. *Journal of the American Statistical Association* 87(417):178–183. https://doi.org/10.1080/01621459.1992.10475190

Fremout T, Thomas E, Gaisberger H, van Meerbeek K, Muenchow J, Briers S, Gutierrez-Miranda CE, Marcelo-Peña JL, Kindt R, Atkinson R, et al. 2020. Mapping tree species vulnerability to multiple threats as a guide to restoration and conservation of tropical dry forests. *Global Change Biology* 26(6):3552–3568. https://doi.org/10.1111/gcb.15028

Gaisberger H, Kindt R, Loo J, Schmidt M, Bognounou F, Da SS, Diallo OB, Ganaba S, Gnoumou A, Lompo D, et al. 2018. Spatially explicit multi-threat assessment of food tree species in Burkina Faso: A fine-scale approach. *PLOS ONE* 13(1):e0190760. https://doi.org/10.1371/journal.pone.0190760

Govaerts R, Nic Lughadha E, Black N, Turner R, Paton A. 2021. The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity. *Scientific Data* 8(1):215. https://doi.org/10.1038/s41597-021-00997-6

Grimmett L, Whitsed R, Horta A. 2020. Presence-only species distribution models are sensitive to sample prevalence: Evaluating models using spatial prediction stability and accuracy metrics. *Ecological Modelling* 431:109194. https://doi.org/10.1016/j.ecolmodel.2020.109194

Hannah L, Roehrdanz PR, Marquet PA, Enquist BJ, Midgley G, Foden W, Lovett JC, Corlett RT, Corcoran D, Butchart SHM, et al. 2020. 30% land conservation and climate action reduces tropical extinction risk by more than 50%. *Ecography* 43(7):943–953. https://doi.org/10.1111/ecog.05166

Hao T, Elith J, Guillera-Arroita G, Lahoz-Monfort JJ. 2019. A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Diversity and Distributions* 25(5):839–852. https://doi.org/10.1111/ddi.12892

Hengl T, Mendes de Jesus J, Heuvelink GBM, Ruiperez Gonzalez M, Kilibarda M, Blagotić A, Shangguan W, Wright MN, Geng X, Bauer-Marschallinger B, et al. 2017. SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE* 12(2):e0169748. https://doi.org/10.1371/journal.pone.0169748

Hijmans RJ. 2020. raster: Geographic Data Analysis and Modeling. R package version 3.4-5. https://CRAN.R-project.org/package=raster

Hijmans RJ. 2019. geosphere: Spherical Trigonometry. R package version 1.5-10. https://CRAN.R-project.org/package=geosphere

Hijmans RJ. 2012. Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null model. *Ecology* 93(3):679–688. https://doi.org/10.1890/11-0826.1

Hijmans RJ, Phillips S, Leathwick J, Elith J. 2017. dismo: Species Distribution Modeling. R package version 1.1-4. https://CRAN.R-project.org/package=dismo

Hoenner X, Huveneers C, Steckenreuter A, Simpfendorfer C, Tattersall K, Jaine F, Atkins N, Babcock R, Brodie S, Burgess J, et al. 2018. Australia's continental-scale acoustic tracking database and its automated quality control process. *Scientific Data* 5(1):170206. https://doi.org/10.1038/sdata.2017.206

Institute of Biodiversity Conservation. 2012. The State of Forest Genetic Resources of Ethiopia. Country Report submitted to FAO on the State of Forest Genetic Resources of Ethiopia. Addis Ababa, Ethiopia: Institute of Biodiversity Conservation. https://www.fao.org/3/i3825e/i3825e23.pdf

Iverson LR, Prasad AM, Peters MP, Matthews SN. 2019. Facilitating adaptive forest management under climate change: A spatially specific synthesis of 125 species for habitat changes and assisted migration over the eastern United States. *Forests* 10(11):989. https://doi.org/10.3390/f10110989

Jiménez-Valverde A. 2012. Insights into the area under the receiver operating characteristic curve (Auc) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography* 21(4):498–507. https://doi.org/10.1111/j.1466-8238.2011.00683.x

Khoury CK, Amariles D, Soto JS, Diaz MV, Sotelo S, Sosa CC, Ramírez-Villegas J, Achicanoy HA, Velásquez-Tibatá J, Guarino L. 2019. Comprehensiveness of conservation of useful wild plants: An operational indicator for biodiversity and sustainable development targets. *Ecological Indicators* 98:420–429. https://doi.org/10.1016/j.ecolind.2018.11.016

Kindt R. 2020. WorldFlora: An R package for exact and fuzzy matching of plant names against the World Flora Online taxonomic backbone data. *Applications in Plant Sciences* 8(9). https://doi.org/10.1002/aps3.11388

Kindt R. 2018a. Provision of Adequate Tree Seed Portfolios (PATSPO): Preparation for species distribution modelling. Nairobi, Kenya: World Agroforestry. https://www.worldagroforestry.org/output/preparation-species-distribution-modelling

Kindt R. 2018b. Ensemble species distribution modelling with transformed suitability values. *Environmental Modelling & Software* 100:136–145. https://doi.org/10.1016/j.envsoft.2017.11.009

Kindt R and Coe R. 2005. Tree diversity analysis: A manual and software for common statistical methods for ecological and biodiversity studies. Nairobi, Kenya: World Agroforestry. https://www.worldagroforestry.org/publication/tree-diversity-analysis-manual-and-software-common-statistical-methods-ecological-and

Kindt R, Dawson IK, Lillesø J-PB, Muchugi A, Pedercini F, Roshetko JM, van Noordwijk M, Graudal L, Jamnadass R. 2021. The one hundred tree species prioritized for planting in the tropics and subtropics as indicated by database mining. Nairobi, Kenya: World Agroforestry. https://worldagroforestry.org/publication/one-hundred-tree-species-prioritized-planting-tropics-and-subtropics-indicated-database

Kindt R, Graudal L, Jamnadass R, Pedercini F, McMullin S, Hendre PS, Carsan S, Moestrup S, Abiyu A, Lillesø J-PB., et al. 2023. Operationalizing climate appropriate portfolios of tree diversity. CIFOR-ICRAF Infobrief No. 383. https://DOI: 10.17528/cifor-icraf/008850

Lamarche C, Santoro M, Bontemps S, d'Andrimont R, Radoux J, Giustarini L, Brockmann C, Wevers J, Defourny P, Arino O. 2017. Compilation and validation of SAR and optical data products for a complete and global map of inland/ocean water tailored to the climate modeling community. *Remote Sensing* 9(1):36. https://doi.org/10.3390/rs9010036

Lazarus ED and McGill BJ. 2014. Pushing the pace of tree species migration. *PLoS ONE* 9(8):e105380. https://doi.org/10.1371/journal.pone.0105380

Liu C, Newell G, White M. 2016. On the selection of thresholds for predicting species occurrence with presence-only data. *Ecology and Evolution* 6(1):337–348. https://doi.org/10.1002/ece3.1878

Liu C, White M, Newell G. 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography* 40(4):778–789. https://doi.org/10.1111/jbi.12058

Luedeling E, Kindt R, Huth NI, Koenig K. 2014. Agroforestry systems in a changing climate – Challenges in projecting future performance. *Current Opinion in Environmental Sustainability* 6:1–7. https://doi.org/10.1016/j.cosust.2013.07.013

Maitner BS, Boyle B, Casler N, Condit R, Donoghue II J, Durán SM, Guaderrama D, Hinchliff CE, Jørgensen PM, Kraft NJB, et al. 2018. The BIEN R package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods in Ecology and Evolution* 9(2):373–379. https://doi.org/10.1111/2041-210X.12861

Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W. 2009. Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions* 15(1):59–69. https://doi.org/10.1111/j.1472-4642.2008.00491.x

Mastrandrea MD, Mach KJ, Plattner G-K, Edenhofer O, Stocker TF, Field CB, Ebi KL, Matschoss PR. 2011. The IPCC AR5 guidance note on consistent treatment of uncertainties: A common approach across the working groups. *Climatic Change* 108:675–691. https://doi.org/10.1007/s10584-011-0178-6

Mendes P, Velazco SJE, de Andrade AFA, De Marco P. 2020. Dealing with overprediction in species distribution models: How adding distance constraints can improve model accuracy. *Ecological Modelling* 431:109180. https://doi.org/10.1016/j.ecolmodel.2020.109180

Meyer C, Weigelt P, Kreft H. 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecology Letters* 19(8):992–1006. https://doi.org/10.1111/ele.12624

Naimi B, Hamm NAS, Groen TA, Skidmore AK, Toxopeus AG. 2014. Where is positional uncertainty a problem for species distribution modelling? *Ecography* 37(2):191–203. https://doi.org/10.1111/j.1600-0587.2013.00205.x

Nix HA. 1986. A biogeographic analysis of Australian elapid snakes. *In* Longmore R. ed. *Atlas of elapid snakes of Australia. Australian Flora and Fauna Series 7.* Canberra: Australian Government Publishing Service 4–15

Orwa C, Mutua A, Kindt R, Jamnadass R, Simons A. 2009. *Agroforestree Database: A tree reference and selection guide. Version 4.* Nairobi, Kenya: World Agroforestry. https://www.worldagroforestry.org/output/agroforestree-database

Pedersen TL. 2020. ggforce: Accelerating 'ggplot2'. R package version 0.3.2. https://CRAN.R-project.org/package=ggforce

Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME. 2017. Opening the black box: An open-source release of Maxent. *Ecography* 40(7):887–893. https://doi.org/10.1111/ecog.03049

Proosdij ASJ, Sosef MSM, Wieringa JJ, Raes N. 2016. Minimum required number of specimen records to develop accurate species distribution models. *Ecography* 39(6):542–552. https://doi.org/10.1111/ecog.01509

R Core Team. 2019. R: A language and environment for statistical computing. Version 3.6.1. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/

R Core Team. 2020. R: A language and environment for statistical computing. Version 4.0.2. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Ramirez-Villegas J, Khoury CK, Achicanoy HA, Mendez AC, Diaz MV, Sosa CC, Debouck DG, Kehel Z, Guarino L. 2020. A gap analysis modelling framework to prioritize collecting for ex situ conservation of crop landraces. *Diversity and Distributions* 26(6):730–742. https://doi.org/10.1111/ddi.13046

Ranjitkar S, Kindt R, Sujakhu NM, Hart R, Guo W, Yang X, Shrestha KK, Xu J, Luedeling E. 2014. Separation of the bioclimatic spaces of Himalayan tree rhododendron species predicted by ensemble suitability models. *Global Ecology and Conservation* 1:2–12. https://doi.org/10.1016/j.gecco.2014.07.001

Ranjitkar S, Xu J, Shrestha KK, Kindt R. 2014. Ensemble forecast of climate suitability for the Trans-Himalayan Nyctaginaceae species. *Ecological Modelling* 282:18–24. https://doi.org/10.1016/j.ecolmodel.2014.03.003

Rivers MC, Taylor L, Brummitt NA, Meagher TR, Roberts DL, Lughadha EN. 2011. How many herbarium specimens are needed to detect threatened species? *Biological Conservation* 144(10):2541–2547. https://doi.org/10.1016/j.biocon.2011.07.014

Rogerson PA. 2014. *Statistical methods for geography: A student's guide*. London: SAGE Publications.

Santini L, Benítez-López A, Maiorano L, Čengić M, Huijbregts MAJ. 2021. Assessing the reliability of species distribution projections in climate change research. *Diversity and Distributions* 27(6):1035–1050. https://doi.org/10.1111/ddi.13252

Seed Leaflets series. 1983–2001. Short descriptions of a number of tropical tree species with particular emphasis on seed issues e.g., harvest, treatment, storage and sowing of seeds. Copenhagen: University of Copenhagen. https://ign.ku.dk/english/publications/publications/previous-publications-series/seed-leaflets/

Sillero N, Arenas-Castro S, Enriquez-Urzelai U, Vale CG, Sousa-Guedes D, Martínez-Freiría F, Real R, Barbosa AM. 2021. Want to model a species niche? A step-by-step guideline on correlative ecological niche modelling. *Ecological Modelling* 456:109671. https://doi.org/10.1016/j.ecolmodel.2021.109671

Terrible, M. 1975. *Atlas de la Haute-Volta. Arbres et arbustes. Carte de répartition et de présence de 65 espèces*. Paris: Institut Géographique National.

Thuiller W, Guéguen M, Renaud J, Karger DN, Zimmermann NE. 2019. Uncertainty in ensembles of global biodiversity scenarios. *Nature Communications* 10(1):1446. https://doi.org/10.1038/s41467-019-09519-w

Title PO and Bemmels JB. 2018. ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography* 41(2):291–307. https://doi.org/10.1111/ecog.02880

Valavi R, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G. 2019. BLOCK CV: An R package for generating spatially or environmentally separated folds for *k*-fold cross-validation of species distribution models. *Methods in Ecology and Evolution* 10(2):225–232. https://doi.org/10.1111/2041-210X.13107 van Zonneveld M, Kindt R, Solberg SØ, N'Danikou S, Dawson IK. 2021. Diversity and conservation of traditional African vegetables: Priorities for action. *Diversity and Distributions* 27(2):216–232. https://doi.org/10.1111/ddi.13188

Varela S, Anderson RP, García-Valdés R, Fernández-González F. 2014. Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography* 37:1084–1091. https://doi.org/10.1111/j.1600-0587.2013.00441.x

Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. 2009. New York: Springer Verlag. https://doi.org/10.1007/978-0-387-98141-3

Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, NCEAS Predicting Species Distributions Working Group. 2008. Effects of sample size on the performance of species distribution models. *Diversity and Distributions* 14(5):763–773. https://doi.org/10.1111/j.1472-4642.2008.00482.x

Wunderlich RF, Lin Y-P, Anthony J, Petway JR. 2019. Two alternative evaluation metrics to replace the true skill statistic in the assessment of species distribution models. *Nature Conservation* 35:97–116. https://doi.org/10.3897/natureconservation.35.33918

Zizka A, Silvestro D, Andermann T, Azevedo J, Duarte Ritter C, Edler D, Farooq H, Herdean A, Ariza M, Scharn R, et al. 2019. CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution* 10(5):744–751. https://doi.org/10.1111/2041-210X.13152

Appendices

Appendix 1. Synonyms and authorities for species names

Current names and synonyms were checked via World Flora Online (WFO; May 2019 version of the taxonomic backbone; http://www.worldfloraonline.org/) via the *WorldFlora* package (version 1.9; Kindt 2020). We also checked for current names in the World Checklist of Vascular Plants (WCVP; version 6 of September 2021; https://wcvp.science.kew.org/; Govaerts et al. 2021) via a modified script from https://rpubs.com/Roeland-KINDT/812716 for using ***WorldFlora::WFO.match***.

Where the same current name was retrieved for the WFO and the WCVP, the identification number for the record in WCVP was included in Table A1.1. Table A1.2 lists the species where the WCVP retrieved an alternative current name.[25]

When compiling occurrence data (Section 5), searches included synonyms listed in Table A1.3.

As World Flora Online has been updated regularly, we list changes in accepted names in Table A1.4 using the most recent downloadable version of the WFO taxonomic backbone. In the more recent version of WFO, *Ziziphus mauritiana* (wfo-0000430322), a species selected as a 'Top 25' one (Table 1), a is no longer treated as a synonym for *Ziziphus jujuba*.

Table A1.1. Authorship and ID for current species names in World Flora Online (May 2019) obtained via the *WorldFlora* package (Kindt 2020). The WCVP.ID shows the ID from the World Checklist of Vascular Plants if the same current name was retrieved as for WFO.

Species	Authorship	WFO.ID	WCVP.ID
Acacia abyssinica	Benth.	wfo-0000187479	
Acacia decurrens	Willd.	wfo-0000192434	470138-1
Acacia lahai	Benth.	wfo-0000201927	
Acacia melanoxylon	R.Br.	wfo-0000204086	470873-1
Acacia nilotica	(L.) Delile	wfo-0000205536	
Acacia polyacantha	Willd.	wfo-0000209605	
Acacia saligna	(Labill.) Wendl.	wfo-0000210801	471383-1
Acacia senegal	(L.) Willd.	wfo-0000210855	
Acacia seyal	Delile	wfo-0000210994	
Acacia sieberiana	DC.	wfo-0000211037	
Acacia tortilis	(Forssk.) Hayne	wfo-0000211235	
Adansonia digitata	L.	wfo-0000519672	558628-1
Afrocarpus falcatus	(Thunb.) C.N.Page	wfo-0000522640	946473-1
Albizia grandibracteata	Taub.	wfo-0000183441	473256-1
Albizia gummifera	(J.F.Gmel.) C.A.Sm.	wfo-0000183535	473259-1
Albizia lebbeck	(L.) Benth.	wfo-0000184271	99109-3
Albizia schimperiana	Oliv.	wfo-0000186271	473389-1

continued on next page

25 Note that these included spelling variants.

Table A1.1. Continued

Species	Authorship	WFO.ID	WCVP.ID
Annona senegalensis	Pers.	wfo-0000537928	72309-1
Anogeissus leiocarpa	(DC.) Guill. & Perr.	wfo-0000538097	
Antiaris toxicaria	Lesch.	wfo-0000538857	850341-1
Azadirachta indica	A.Juss.	wfo-0000557668	1213180-2
Balanites aegyptiaca	(L.) Delile	wfo-0000313273	813589-1
Bauhinia thonningii	Schum.	wfo-0000170425	
Berchemia discolor	(Klotzsch) Hemsl.	wfo-0000564133	716679-1
Borassus aethiopum	Mart.	wfo-0000350303	664869-1
Boswellia microphylla	Chiov.	wfo-0000569709	127052-1
Boswellia neglecta	S.Moore	wfo-0000569712	127056-1
Boswellia ogadensis	Vollesen	wfo-0000569717	905757-1
Boswellia papyrifera	(Caill. ex Delile) Hochst.	wfo-0000569719	127060-1
Boswellia pirottae	Chiov.	wfo-0000569720	127061-1
Boswellia rivae	Engl.	wfo-0000569722	127063-1
Bridelia micrantha	(Hochst.) Baill.	wfo-0000421441	340183-1
Cajanus cajan	(L.) Millsp.	wfo-0000179103	1152177-2
Calliandra calothyrsus	Meisn.	wfo-0001050431	
Callistemon citrinus	(Curtis) Skeels	wfo-0000775642	
Calotropis procera	(Aiton) Dryand.	wfo-0000581500	1004515-2
Capparis tomentosa	Lam.	wfo-0000585223	146824-1
Carica papaya	L.	wfo-0000588009	30011248-2
Casuarina cunninghamiana	Miq.	wfo-0000590647	159845-1
Casuarina equisetifolia	L.	wfo-0000590663	159856-1
Catha edulis	(Vahl) Endl.	wfo-0000590815	941530-1
Ceiba pentandra	(L.) Gaertn.	wfo-0000592594	1166232-2
Celtis africana	Burm.f.	wfo-0000593393	850978-1
Citrus sinensis	(L.) Osbeck	wfo-0001249323	
Coffea arabica	L.	wfo-0000910097	747038-1
Combretum aculeatum	Vent.	wfo-0000616040	169878-1
Combretum collinum	Fresen.	wfo-0000616192	170004-1
Combretum molle	R.Br. ex G.Don	wfo-0000616553	170290-1
Commiphora africana	(A.Rich.) Endl.	wfo-0000617158	127576-1
Commiphora guidottii	Chiov. ex Guid.	wfo-0000617297	127676-1
Commiphora myrrha	(Nees) Engl.	wfo-0000617380	127741-1
Cordeauxia edulis	Hemsl.	wfo-0000165271	487135-1
Cordia africana	Lam.	wfo-0000620224	113939-1
Corymbia citriodora	(Hook.) K.D.Hill & L.A.S.Johnson	wfo-0000925431	986336-1
Croton macrostachyus	Hochst. ex Delile	wfo-0000931591	342917-1
Cupressus lusitanica	Mill.	wfo-0000630722	
Cupressus sempervirens	L.	wfo-0000630789	261974-1
Cytisus proliferus	L.f.	wfo-0000185459	
Dalbergia melanoxylon	Guill. & Perr.	wfo-0000172325	490328-1

continued on next page

Table A1.1. Continued

Species	Authorship	WFO.ID	WCVP.ID
Delonix regia	(Hook.) Raf.	wfo-0000166389	491231-1
Dichrostachys cinerea	(L.) Wight & Arn.	wfo-0000176871	492423-1
Diospyros mespiliformis	Hochst. ex A.DC.	wfo-0000649333	322702-1
Dobera glabra	(Forssk.) Juss. ex Poir.	wfo-0000652723	779320-1
Dodonaea viscosa	(L.) Jacq.	wfo-0000653170	30058367-2
Dombeya torrida	(J.F.Gmel.) Bamps	wfo-0000654003	823248-1
Dovyalis abyssinica	(A.Rich.) Warb.	wfo-0000925138	111558-1
Dovyalis caffra	(Hook.f. & Harv.) Sim	wfo-0000925143	111560-1
Ekebergia capensis	Sparrm.	wfo-0000663623	578362-1
Entada abyssinica	A.Rich.	wfo-0000205748	493817-1
Erythrina abyssinica	DC.	wfo-0000180423	494336-1
Erythrina brucei	Schweinf.	wfo-0000180564	494368-1
Eucalyptus camaldulensis	Dehnh.	wfo-0000954597	592777-1
Eucalyptus globulus	Labill.	wfo-0000954998	592965-1
Eucalyptus grandis	W.Hill	wfo-0000955035	592976-1
Eucalyptus saligna	Sm.	wfo-0000955842	593334-1
Eucalyptus viminalis	Labill.	wfo-0000956115	593454-1
Euphorbia tirucalli	L.	wfo-0000965116	348517-1
Faidherbia albida	(Delile) A.Chev.	wfo-0000186081	494764-1
Ficus carica	L.	wfo-0000687690	852556-1
Ficus sur	Forssk.	wfo-0000690530	853792-1
Ficus sycomorus	L.	wfo-0000690537	853797-1
Flacourtia indica	(Burm.f.) Merr.	wfo-0000925655	365348-1
Flueggea virosa	(Roxb. ex Willd.) Royle	wfo-0000967255	1013601-1
Garcinia livingstonei	T.Anderson	wfo-0000694422	428049-1
Gardenia volkensii	K.Schum.	wfo-0000971256	751323-1
Grevillea robusta	A.Cunn. ex R.Br.	wfo-0000709544	50798-3
Grewia damine	Gaertn.	wfo-0000709875	
Grewia villosa	Willd.	wfo-0000710393	834635-1
Hagenia abyssinica	(Bruce ex Steud.) J.F.Gmel.	wfo-0000994920	725448-1
Hyphaene thebaica	(L.) Mart.	wfo-0000216304	667540-1
Ilex mitis	(L.) Radlk.	wfo-0000729632	83531-1
Jacaranda mimosifolia	D.Don	wfo-0000778761	130936-2
Jatropha curcas	L.	wfo-0000219580	131462-2
Juniperus procera	Hochst. ex Endl.	wfo-0000355729	262311-1
Kigelia africana	(Lam.) Benth.	wfo-0000778884	109874-1
Lawsonia inermis	L.	wfo-0000366658	553638-1
Leucaena leucocephala	(Lam.) de Wit	wfo-0000164084	138955-2
Maerua aethiopica	(Fenzl) Oliv.	wfo-0001290548	147641-1
Malus domestica	Borkh.	wfo-0001008355	726282-1
Mangifera indica	L.	wfo-0000371248	69913-1
Markhamia lutea	(Benth.) K.Schum.	wfo-0000779039	110020-1
Melia azedarach	L.	wfo-0000450150	578949-1

continued on next page

Table A1.1. Continued

Species	Authorship	WFO.ID	WCVP.ID
Milicia excelsa	(Welw.) C.C.Berg	wfo-0000447908	910900-1
Millettia ferruginea	(Hochst.) Baker	wfo-0000199980	507366-1
Moringa oleifera	Lam.	wfo-0001085051	584736-1
Moringa stenopetala	(Baker f.) Cufod.	wfo-0001085058	584747-1
Nuxia congesta	R.Br. ex Fresen.	wfo-0000797418	546816-1
Olea capensis	L.	wfo-0000817299	610645-1
Olea europaea	L.	wfo-0000817273	610675-1
Oxytenanthera abyssinica	(A.Rich.) Munro	wfo-0000882687	410276-1
Parkinsonia aculeata	L.	wfo-0000170206	512242-1
Persea americana	Mill.	wfo-0000465160	325643-2
Phoenix reclinata	Jacq.	wfo-0000269796	668943-1
Pinus patula	Schiede ex Schltdl. & Cham.	wfo-0000481882	314961-2
Polyscias fulva	(Hiern) Harms	wfo-0000280060	91769-1
Pouteria adolfi-friedericii	(Engl.) A.Meeuse	wfo-0000281508	
Prunus africana	(Hook.f.) Kalkman	wfo-0000995790	729417-1
Pterolobium stellatum	(Forssk.) Brenan	wfo-0000170624	516643-1
Rhamnus prinoides	L'Hér.	wfo-0000460040	718580-1
Saba comorensis	(Bojer ex A.DC.) Pichon	wfo-0000299250	81757-1
Salvadora persica	L.	wfo-0000492914	779348-1
Sarcocephalus latifolius	(Sm.) E.A.Bruce	wfo-0000303384	
Schefflera abyssinica	(Hochst. ex A.Rich.) Harms	wfo-0000305601	
Schinus molle	L.	wfo-0000435157	71044-1
Sclerocarya birrea	(A.Rich.) Hochst.	wfo-0000434908	71162-1
Searsia natalensis	(Bernh. ex C.Krauss) F.A.Barkley	wfo-0000434889	71180-1
Securidaca longipedunculata	Fresen.	wfo-0000503535	
Senna didymobotrya	(Fresen.) H.S.Irwin & Barneby	wfo-0000163726	234467-2
Sesbania bispinosa	(Jacq.) W.Wight	wfo-0000186833	518441-1
Sesbania sesban	(L.) Merr.	wfo-0000178461	518533-1
Shirakiopsis elliptica	(Hochst.) Esser	wfo-0000309756	1014108-1
Spathodea campanulata	P.Beauv.	wfo-0000779723	110661-1
Steganotaenia araliacea	Hochst.	wfo-0000431247	849204-1
Stereospermum kunthianum	Cham.	wfo-0000779673	110792-1
Strychnos henningsii	Gilg	wfo-0000502962	547225-1
Strychnos innocua	Delile	wfo-0000502968	547244-1
Strychnos spinosa	Lam.	wfo-0000502889	547485-1
Syzygium guineense	(Willd.) DC.	wfo-0000318724	601750-1
Tamarindus indica	L.	wfo-0000170926	520167-1
Tamarix aphylla	(L.) H.Karst.	wfo-0000458771	828051-1
Terminalia brownii	Fresen.	wfo-0001296425	171010-1
Trichilia emetica	Vahl	wfo-0000455454	579419-1

continued on next page

Table A1.1. Continued

Species	Authorship	WFO.ID	WCVP.ID
Vangueria madagascariensis	J.F.Gmel.	wfo-0000331269	769766-1
Vepris nobilis	(Delile) Mziray	wfo-0000420153	969503-1
Vernonia amygdalina	Delile	wfo-0000072744	
Vitellaria paradoxa	C.F.Gaertn.	wfo-0000332885	790034-1
Vitex doniana	Sweet	wfo-0000333061	865694-1
Warburgia ugandensis	Sprague	wfo-0000427581	146038-1
Ximenia americana	L.	wfo-0000428247	316341-2
Yushania alpina	(K.Schum.) W.C.Lin	wfo-0000907601	
Ziziphus jujuba	Mill.	wfo-0000430303	719213-1
Ziziphus mucronata	Willd.	wfo-0000430319	719359-1
Ziziphus spina-christi	(L.) Desf.	wfo-0001131308	719427-1

Table A1.2. Synonyms (current names in the World Checklist of Vascular Plants (WCVP) with the naming authority from the same database) and ID for species where World Flora Online and the World Checklist on Vascular Plants disagreed on the current name.

Species	Synonym	Authority	WCVP.ID
Acacia abyssinica	*Vachellia abyssinica*	(Hochst. ex Benth.) Kyal. & Boatwr.	77131675-1
Acacia lahai	*Vachellia lahai*	(Steud. & Hochst. ex Benth.) Kyal. & Boatwr.	77131736-1
Acacia nilotica	*Vachellia nilotica*	(L.) P.J.H.Hurter & Mabb.	77089275-1
Acacia polyacantha	*Senegalia polyacantha*	(Willd.) Seigler & Ebinger	60451312-2
Acacia senegal	*Senegalia senegal*	(L.) Britton	518304-1
Acacia seyal	*Vachellia seyal*	(Delile) P.J.H.Hurter	77089276-1
Acacia sieberiana	*Vachellia sieberiana*	(DC.) Kyal. & Boatwr.	77131781-1
Acacia tortilis	*Vachellia tortilis*	(Forssk.) Galasso & Banfi	77087190-1
Anogeissus leiocarpa	*Terminalia leiocarpa*	(DC.) Baill.	171202-1
Bauhinia thonningii	*Piliostigma thonningii*	(Schumach.) Milne-Redh.	514346-1
Calliandra calothyrsus	*Calliandra houstoniana* var. *calothyrsus*	(Meisn.) Barneby	1010291-1
Callistemon citrinus	*Melaleuca citrina*	(Curtis) Dum.Cours.	77108602-1
Citrus sinensis	*Citrus ×aurantium*	L.	59600-2
Cupressus lusitanica	*Hesperocyparis lusitanica*	(Mill.) Bartel	60451554-2
Cytisus proliferus	*Chamaecytisus albus*	(Hacq.) Rothm.	485939-1
Grewia damine	*Grewia tiliifolia*	Vahl	834597-1
Pouteria adolfi-friedericii	*Pouteria adolfi-friederici*	(Engl.) A.Meeuse	788785-1
Sarcocephalus latifolius	*Nauclea latifolia*	Sm.	757144-1
Schefflera abyssinica	*Astropanax abyssinicus*	(Hochst. ex A.Rich.) Seem.	89896-1
Securidaca longipedunculata	*Securidaca longepedunculata*	Fresen.	692714-1
Vernonia amygdalina	*Gymnanthemum amygdalinum*	(Delile) Sch.Bip.	210886-1
Yushania alpina	*Oldeania alpina*	(K.Schum.) Stapleton	77131105-1

Table A1.3. Confirmed synonym names of the candidate species for species distribution modelling (Table 1). Synonyms were confirmed with World Flora Online where there are no entries in the Comment column.

Species	Synonym	Comment
Acacia abyssinica	*Vachellia abyssinica*	see Table A1.2
Acacia lahai	*Vachellia lahai*	see Table A1.2
Acacia polyacantha	*Senegalia polyacantha*	see Table A1.2
Acacia senegal	*Senegalia senegal*	
Acacia seyal	*Vachellia seyal*	see Table A1.2
Acacia sieberiana	*Vachellia sieberiana*	see Table A1.2
Afrocarpus falcatus	*Podocarpus falcatus*	
Anogeissus leiocarpa	*Terminalia leiocarpa*	
Bauhinia thonningii	*Piliostigma thonningii*	
Calliandra calothyrsus	*Calliandra houstoniana*	see Table A1.2
Citrus sinensis	*Citrus aurantium*	see Table A1.2
Combretum molle	*Combretum rochetanum*	
Corymbia citriodora	*Eucalyptus citriodora*	
Cupressus sempervirens	*Cupressus pyramidalis*	
Cytisus proliferus	*Chamaecytisus palmensis*	
Cytisus proliferus	*Chamaecytisus proliferus*	
Dodonaea viscosa	*Dodonaea angustifolia*	*Dodonaea viscosa* subsp. *angustifolia*
Dombeya torrida	*Dombeya schimperiana*	
Faidherbia albida	*Acacia albida*	
Grewia damine	*Grewia bicolor*	
Prunus africana	*Pygeum africanum*	
Sarcocephalus latifolius	*Nauclea latifolia*	see Table A1.2
Searsia natalensis	*Rhus natalensis*	
Sesbania bispinosa	*Sesbania aculeata*	
Shirakiopsis elliptica	*Sapium ellipticum*	
Spathodea campanulata	*Spathodea nilotica*	
Vepris nobilis	*Teclea nobilis*	
Vernonia amygdalina	*Gymnanthemum amygdalinum*	see Table A1.2
Yushania alpina	*Oldeania alpina*	see Table A1.2
Yushania alpina	*Arundinaria alpina*	
Yushania alpina	*Sinarundinaria alpina*	
Ziziphus jujuba	*Ziziphus mauritiana*	

Table A1.4. Authorship and ID for current species names in World Flora Online (January 2023) via the *WorldFlora* package (Kindt 2020). Only species where the name has changed from Table A1.1 have been included.

Species	Type	taxonID	Current name	Current Authorship
Acacia abyssinica	new	wfo-0001336820	*Vachellia abyssinica*	(Hochst. ex Benth.) Kyal. & Boatwr.
Acacia lahai	new	wfo-0001336856	*Vachellia lahai*	(Steud. & Hochst. ex Benth.) Kyal. & Boatwr.
Acacia nilotica	new	wfo-0001284776	*Vachellia nilotica*	(L.) P.J.H.Hurter & Mabb.
Acacia polyacantha	new	wfo-0000744649	*Senegalia polyacantha*	(Willd.) Seigler & Ebinger
Acacia senegal	new	wfo-0001281302	*Senegalia senegal*	(L.) Britton

continued on next page

Table A1.4. Continued

Acacia seyal	new	wfo-0001284777	*Vachellia seyal*	(Delile) P.J.H.Hurter
Acacia sieberiana	new	wfo-0000201188	*Acacia hamiltoniana*	Maiden
Acacia tortilis	new	wfo-0001285358	*Vachellia tortilis*	(Forssk.) Galasso & Banfi
Anogeissus leiocarpa	new	wfo-0000408839	*Terminalia leiocarpa*	Baill.
Bauhinia thonningii	new	wfo-0000170413	*Piliostigma thonningii*	(Schumach.) Milne-Redh.
Calliandra calothyrsus	new	wfo-0000199357	*Calliandra houstoniana var. calothyrsus*	(Meisn.) Barneby
Callistemon citrinus	new	wfo-0000239474	*Melaleuca citrina*	(Curtis) Dum.Cours.
Citrus sinensis	new	wfo-0000607909	*Citrus × aurantium*	L.
Cytisus proliferus	new	wfo-0001057343	*Chamaecytisus prolifer*	(L.f.) Link
Sarcocephalus latifolius	new	wfo-0000249572	*Nauclea latifolia*	Sm.
Schefflera abyssinica	new	wfo-0000294220	*Astropanax abyssinicum*	(Hochst. ex A.Rich.) Seem.
Securidaca longipedunculata	spelling	wfo-0000503535	*Securidaca longepedunculata*	Fresen.
Vernonia amygdalina	new	wfo-0000096111	*Gymnanthemum amygdalinum*	(Delile) Sch.Bip.
Yushania alpina	new	wfo-0001336675	*Oldeania alpina*	(K.Schum.) Stapleton

Appendix 2. Data sets accessed from GBIF

Data accessed from GBIF on 1 October 2018 (Section 5) belonged to the following occurrence datasets:

https://doi.org/10.15468/1ojlip
https://doi.org/10.15468/22kwre
https://doi.org/10.15468/2a9ebc
https://doi.org/10.15468/2wbxxh
https://doi.org/10.15468/3gmsj9
https://doi.org/10.15468/3v0exk
https://doi.org/10.15468/4i21ik
https://doi.org/10.15468/4owz9w
https://doi.org/10.15468/4vlfas
https://doi.org/10.15468/50rhvf
https://doi.org/10.15468/5cdfgv
https://doi.org/10.15468/5sl7sh
https://doi.org/10.15468/5wurxy
https://doi.org/10.15468/5yvniw
https://doi.org/10.15468/6habyb
https://doi.org/10.15468/6xu3nq
https://doi.org/10.15468/7c2j4n
https://doi.org/10.15468/7gudyo
https://doi.org/10.15468/7pjxmn
https://doi.org/10.15468/7uigwo
https://doi.org/10.15468/81w5bx
https://doi.org/10.15468/8tx9fr
https://doi.org/10.15468/8wushq
https://doi.org/10.15468/8wyijj
https://doi.org/10.15468/92bnx4
https://doi.org/10.15468/9fvpm9
https://doi.org/10.15468/9tfduu
https://doi.org/10.15468/9wycdn
https://doi.org/10.15468/ab3s5x
https://doi.org/10.15468/abe1lg
https://doi.org/10.15468/axtkuz
https://doi.org/10.15468/aywock
https://doi.org/10.15468/bbtur8
https://doi.org/10.15468/bdv8ee
https://doi.org/10.15468/bhvwem
https://doi.org/10.15468/blra4p
https://doi.org/10.15468/bmbo9q
https://doi.org/10.15468/bpx7wi
https://doi.org/10.15468/bvg1jz
https://doi.org/10.15468/bvrtsv
https://doi.org/10.15468/bznvdf
https://doi.org/10.15468/c4w4co
https://doi.org/10.15468/ckbuue
https://doi.org/10.15468/cmnjt1
https://doi.org/10.15468/cmrlox
https://doi.org/10.15468/cpnhcc
https://doi.org/10.15468/cribcp
https://doi.org/10.15468/d4ssas
https://doi.org/10.15468/dayejw
https://doi.org/10.15468/djzgie
https://doi.org/10.15468/dlwwhz
https://doi.org/10.15468/dq9umt
https://doi.org/10.15468/e4y4nr
https://doi.org/10.15468/e8rhqm
https://doi.org/10.15468/eabysc
https://doi.org/10.15468/encqea
https://doi.org/10.15468/enkiul
https://doi.org/10.15468/exh7vo
https://doi.org/10.15468/exwp2m
https://doi.org/10.15468/fbbbfl
https://doi.org/10.15468/g8pbz5
https://doi.org/10.15468/gcoa2d
https://doi.org/10.15468/gcvvtr
https://doi.org/10.15468/geylm0
https://doi.org/10.15468/gm1qxi
https://doi.org/10.15468/h1txwb
https://doi.org/10.15468/h3mptq
https://doi.org/10.15468/h85hfz
https://doi.org/10.15468/hja69f
https://doi.org/10.15468/hnhrg3
https://doi.org/10.15468/hnlnjq
https://doi.org/10.15468/hnvndq
https://doi.org/10.15468/hseoaq
https://doi.org/10.15468/hy2rea
https://doi.org/10.15468/iinlqm
https://doi.org/10.15468/iokuu0
https://doi.org/10.15468/iozuua
https://doi.org/10.15468/irkrku
https://doi.org/10.15468/izzxhf
https://doi.org/10.15468/jbp1u9
https://doi.org/10.15468/jbqjzg
https://doi.org/10.15468/jcpwq5
https://doi.org/10.15468/jgauer
https://doi.org/10.15468/jgnvoj
https://doi.org/10.15468/jgqm7v
https://doi.org/10.15468/jhysv3
https://doi.org/10.15468/jmlkgk
https://doi.org/10.15468/jyrthk
https://doi.org/10.15468/kpwam2
https://doi.org/10.15468/ku21hg

https://doi.org/10.15468/l5pasu
https://doi.org/10.15468/lbjfvo
https://doi.org/10.15468/lcbrct
https://doi.org/10.15468/lhlwfx
https://doi.org/10.15468/lmf6x5
https://doi.org/10.15468/looi3y
https://doi.org/10.15468/ly60bx
https://doi.org/10.15468/lzrybq
https://doi.org/10.15468/m0cqdu
https://doi.org/10.15468/m2s30i
https://doi.org/10.15468/mboy4p
https://doi.org/10.15468/mci9vv
https://doi.org/10.15468/mdnmzb
https://doi.org/10.15468/mi3taw
https://doi.org/10.15468/mnys1f
https://doi.org/10.15468/mpx4qk
https://doi.org/10.15468/mqurnm
https://doi.org/10.15468/msqavy
https://doi.org/10.15468/mz2hp2
https://doi.org/10.15468/nc6rxy
https://doi.org/10.15468/nheq3e
https://doi.org/10.15468/nud2pn
https://doi.org/10.15468/nxnqzf
https://doi.org/10.15468/nyhbez
https://doi.org/10.15468/oblwrk
https://doi.org/10.15468/oeqwrv
https://doi.org/10.15468/ohy7xv
https://doi.org/10.15468/oi1ego
https://doi.org/10.15468/oqniiy
https://doi.org/10.15468/owja31
https://doi.org/10.15468/p4cwmb
https://doi.org/10.15468/p4hm6o
https://doi.org/10.15468/pff0t6
https://doi.org/10.15468/pgrqej
https://doi.org/10.15468/pkgevu
https://doi.org/10.15468/pucjqn
https://doi.org/10.15468/pzlxvg
https://doi.org/10.15468/qbb7sg
https://doi.org/10.15468/qvbvdp
https://doi.org/10.15468/r9azth
https://doi.org/10.15468/ra9vp0
https://doi.org/10.15468/rckmn2
https://doi.org/10.15468/rmiqhw
https://doi.org/10.15468/rqn53h
https://doi.org/10.15468/rvjdu1
https://doi.org/10.15468/rydcn2
https://doi.org/10.15468/s1f4aw
https://doi.org/10.15468/s3o29s
https://doi.org/10.15468/s5auru
https://doi.org/10.15468/s6ctus
https://doi.org/10.15468/szhrlq
https://doi.org/10.15468/t5azeg
https://doi.org/10.15468/tixwop
https://doi.org/10.15468/tjq4xs
https://doi.org/10.15468/tnj8wm
https://doi.org/10.15468/tobbxc
https://doi.org/10.15468/tsrjm0
https://doi.org/10.15468/uc1apo
https://doi.org/10.15468/ufmslw
https://doi.org/10.15468/ugkanx
https://doi.org/10.15468/uhrfuc
https://doi.org/10.15468/ulk1iz
https://doi.org/10.15468/uul0mj
https://doi.org/10.15468/uzfatz
https://doi.org/10.15468/vlr656
https://doi.org/10.15468/vnatbk
https://doi.org/10.15468/vp9lmx
https://doi.org/10.15468/w35jmd
https://doi.org/10.15468/wbw5v0
https://doi.org/10.15468/we6uaw
https://doi.org/10.15468/wh7guo
https://doi.org/10.15468/wnual9
https://doi.org/10.15468/wojfbp
https://doi.org/10.15468/wrthhx
https://doi.org/10.15468/xmucap
https://doi.org/10.15468/xsp4ka
https://doi.org/10.15468/xwvi5c
https://doi.org/10.15468/y8jm57
https://doi.org/10.15468/ydhpki
https://doi.org/10.15468/yn20mq
https://doi.org/10.15468/yo3mmu
https://doi.org/10.15468/ypcgtg
https://doi.org/10.15468/ypoair
https://doi.org/10.15468/yr534j
https://doi.org/10.15468/yucybm
https://doi.org/10.15468/z1qxgu
https://doi.org/10.15468/zmcpyo
https://doi.org/10.15468/zplh9m
https://doi.org/10.18165/qssyvr
https://doi.org/10.3897/bdj.4.e8286
https://doi.org/10.5519/0002965

Appendix 3. Data sets accessed from BIEN

Data accessed from BIEN on 1 October 2018 (Section 5) were attributed to the following custodians, compiled via the field of 'dataowner' of the downloaded datasets. Abbreviations for some of the herbaria correspond to those listed in https://bien.nceas.ucsb.edu/bien/data-contributors/herbaria/ .

A, AAU, ABH, AD, AK, AMAZ, ARAN, AS, AUT023, AUT024, AUT052, AZE009, AZE015, B, BA, BAA, BAB, Badru Mugerwa, BAF, BC, BCN, BDBCV, BEREA, BFL, BG, BIGA, BIGU, BIO, BIO-UNIPI, BM, BMO, BOLV, BPBM, BR, Brad Boyle, BRH, BRI, BRLU, C, CAH, CANB, CAS, CAS-BOT-BC, CAY, CBM, CDA, CDBI, CDMB, Cenargen, Centre National de la Recherche Scientifique et Technologique / Institut de l'environnement et de recherches agricoles, Centre National de la Recherche Scientifique et Technologique / Institut de lenvironnement et de recherches agricoles, Centre National des Semences Forestières, CETI, CHAPA, CHEP, chilesp, CHR, CHSC, CIB, CIB-UV, CIBYC-UAEM, CICY, Ciência e Tecnologia (IFAM), CIHS-UAC, CIIDIR-IPN, CJBN, CLARK-A, CNARP, CNS, CNS-UT, COA, COFC, COI, COL, Comissão Executiva do Plano da Lavoura Cacaueira (CEPLAC), CONC, CPR, CR, CRSN_LWIRO, CSUSB, CTES, CU, CUZ, DAKAR, DAV, David Kenfack, DBF-NHMD, DBG, DICTUS-USON, DNA, Douglas Sheil, DSM, DUKE, E, EA, EAP, EB-BUAP, ECOSUR, EFG, Eileen Larney, Emanuel Martin, EMAU, EMMA, Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), EMY, ENAG, ENCB-IPN, ENT, ESALQ, ESP003, ESP007, ESP046, ESP089, ESP119, ESP197, F, FA-UAS, FAPESP, FB-UMSNH, FC-UNAM, FCB-UANL, FCF-UANL, FCN-UAQ, FCO, FCQ, FESI-UNAM, FFPRI, FHO, FI, FLAS, FMNH, FR, Francesco Rovero, FT, FTG, Fundação Universidade Federal de Mato Grosso do Sul (UFMS), FundaciÃ³n Puerto Rastrojo Â- ColÃ´mbia, Fundación GAIA, FURB, FUVATES, FZ-UACH, G, GA, GABON, GB, George Chuyong, GH, GI, GLM, GMBA, GMNHJ, GRBGT, GUAY, GZU, HAL, Harvard University, HBG, HCM, HCSM, HEM, Herb. Hinton, Herbario Amazónico Colombiano, Herbario de la Universidad Industrial de Santander, Herbario Universidad de Antioquia, Herbarium togoense, HGM, HIB, HN, HNB, HNC, HNMN, HO, HOXA, HSB, HSC, HSS, HU, HUA, HUAL, HULE, HUSA, HUT, HYO, IA, IAC, IADIZA, IANIGLA, IAP, IAvH, IAVH, IBE, IBK, IBOT SAS, IBSC, IBt, IBT, IBUNAM, ICESI, ICN, IE-UNAM, IEA, IEA-UAT, IF, IGB, IGL-UNAM, IHNE, IICT, IIZD-UASLP, ILCA, IMA, IMECBIO-UDG, INB, INBio, INCIVA, INECOL, INIFAP-CECOY, INIREB, INM, INPA, Institut de Recherche Agronomique de Guinée (IRAG), Instituto Amazónico de Investigaciones Científicas SINCHI, Instituto de Botânica, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt (IAvH), Instituto Federal de Educação, Instituto Nacional de Pesquisas da Amazônia (INPA), Instituto Plantarum de Estudos da Flora Ltda (HPL), IPA, IPT, IRENAT-CP, IRVC, ISA, ISKW, ITIC, James S. MIller, JAUM, JBAG, JBBJCM, JBGP, JBRJ, JBS, JCT, Jean Claude Razafimahaimodison, JEO, JEPS, JSCM, JUA, JYV, K, KAW, KE, Keith Pohs, KMN, KOM, KPM, KTU, KU, KUN, KURA, L, LA, LAGU, LBG, LBV, LD, LE, LEB, LEGON-GC, LG, LIL, LISC, LMA, LMU, LOJA, LP, LPB, LSU, LTB, LUKI_INERA, LWI, M, MA, MAK, MAL, Mar-Elise Hill, Matteo Detto, Mauricio Bonifacino, MB, MBK, MBM, MBML, MCNAM, MEDEL, MEL, MELU, MEXU, Meyner Nusalawo, MGC, MHES, MHNG, MHU, MICH, Miriam van Heist, MISS, MKD001, MNHM, MNHN, MNHNL, MO, MOL, Moscow State University, MPN, MPU, MSC, MUB, MUHNAC, Museo de La Salle - Universidad de La Salle, Museo Nacional de Costa Rica (MNCR), Museu de Ciências Naturais - Fundação Zoobotânica do Rio Grande do Sul (MCN-FZBRS), MY, NAS, nbf, ND, NE, NH, NHMM, NHMUK, NHT, NMNH-SI, NMNL, NMNS, NO DISPONIBLE, NOU, NSW, NSW Office of Environment and Heritage, NU, NY, NYBG, NZFRI, O, OBI, Oliver Phillips, OSA, OTS, P, PAMP, Patricia Alvarez-Loayza, Patrick Boundja, Patrick Jansen, PDA, PE, PERTH, PNFM, PRE, PRU, PTHM, PUCRS, PUJ, PY, QCA, QCNE, Richard Condit, Rob Hunt, Robert Peet, RPSC, RSA, S, SALA, SANBI, SANT, Sarah Yoga Bengbate, SAV, SBBG, SD, SDNHM, SDSU, SEINET, SEL, SERBO, SERG, SEV, SFV, SI, SJSU, SMNH, SMNK, SMU, SNSB-M, SP, SRGH, STU, Susan Letcher, SUVA, SW, TAES, TAI, TAIF, TALL, TAM, TAN, TEF, TEFH, TFD, Tim Killeen, TKPM, TLMF, TOYA, TRH, TROM, TUB, U, UA, UAAAN, UACh, UADY, UAZ, UC, UCALDAS, UCD, UCO, UConn, UCR, UCS, UDEA, UDENAR, UDFJC, UEFS, UEL, UEM, UEPA, UESB, UESC, UFBA, UFC, UFERSA, UFES, UFMA, UFPB, UFPE, UFPI, UFPR, UFRB, UFRN, UFRPE, UFS, UFSC, UFSJ, UJAT, UJLOG, ULB, ULM, ULS, UM, UMO, UNAL, UNAN-LEON, UnB, UNEMAT, UNESC, UNESP, UNEX, Uniamazonia, UNICACH, UNICAMP, UNICAP, UNICORDOBA, UNISANTA, UNITINS, UNIVASF, Universidad Católica de Oriente, Universidad de Antioquia, Universidad Industrial de Santander, Universidad Tecnológica del Chocó, Universidade Estadual de Londrina, Universidade Estadual Paulista, Universidade Federal da Bahia (UFBA), Universidade Federal de Mato Grosso

(UFMT), Universidade Federal de Minas Gerais (UFMG), Universidade Federal de Sergipe, Universidade Federal do Ceará, Universidade Federal do Pará (UFPA), Universidade Federal do Paraná, Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Universidade Federal Rural de Pernambuco, Universidade Regional de Blumenau, Universidade Tecnológica Federal do Paraná (UTFPR), Université de Montréal Biodiversity Centre, Université de Strasbourg, University of Alberta Museums, University of British Columbia, UNM, UoB-IB, UPN, UPNA, UPS, UQUINDÍO, US, USCG, USF, USM, USMS, USP, USZ, UTEP, UTFPR, UvA-IBED, UVAL, VAL, VIT, W, WAG, WELT, WII, WIS, WOLL, WTU, WU, XAL, YUG027, YUG047

DOI: 10.17528/cifor-icraf/008977

CIFOR-ICRAF *Working Papers* contain preliminary or advanced research results on important tropical forest issues that need to be published in a timely manner to inform and promote discussion. This content has been internally reviewed but has not undergone external peer review.

This working paper describes the methods used to develop the online **Climate change atlas for Africa of tree species prioritized for forest landscape restoration in Ethiopia**. The purpose of the atlas, available at http://atlas.worldagroforestry.org/, is to indicate how climate change is likely to affect the locations where particular tree species can grow in Africa. The atlas shows the baseline and 2050s habitat distributions across Africa for 127 tree species. Methods behind the creation of the atlas described in this working paper include: the selection of tree species; the processing and selection of predictor variables; the selection of future climates; the compilation of occurrence observations, and their spatial and environmental thinning; the compilation of background observations; the spatial folding of occurrence and background observations; the calibration of species distribution models and the generation of suitability maps; the discrimination of areas with novel environmental conditions; the generation of habitat change maps; and the creation of convex hulls for an *a posteriori* distance constraining method. This working paper is not a beginner's guide to species distribution modelling; however, for users who also require an initial introduction, we provide references to appropriate resources.

UNIVERSITY OF COPENHAGEN

NICFI Norway's International Climate and Forest Initiative

PATSPO II Provision of adequate tree seed portfolios

SRUC

cifor-icraf.org

cifor.org | worldagroforestry.org

CIFOR-ICRAF

The Center for International Forestry Research (CIFOR) and World Agroforestry (ICRAF) envision a more equitable world where trees in all landscapes, from drylands to the humid tropics, enhance the environment and well-being for all. CIFOR-ICRAF are CGIAR Research Centers.

CGIAR