

## Predictive mapping of tree species assemblages in an African montane rainforest

Biotropica

Babaasa, Dennis; Finn, John T.; Schweik, Charles M.; Fuller, Todd K.; Sheil, Douglas  
<https://doi.org/10.1111/btp.13302>

This publication is made publicly available in the institutional repository of Wageningen University and Research, under the terms of article 25fa of the Dutch Copyright Act, also known as the Amendment Taverne.

Article 25fa states that the author of a short scientific work funded either wholly or partially by Dutch public funds is entitled to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed using the principles as determined in the Association of Universities in the Netherlands (VSNU) 'Article 25fa implementation' project. According to these principles research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and / or copyright owner(s) of this work. Any use of the publication or parts of it other than authorised under article 25fa of the Dutch Copyright act is prohibited. Wageningen University & Research and the author(s) of this publication shall not be held responsible or liable for any damages resulting from your (re)use of this publication.

For questions regarding the public availability of this publication please contact  
[openaccess.library@wur.nl](mailto:openaccess.library@wur.nl)

# Predictive mapping of tree species assemblages in an African montane rainforest

Dennis Babaasa<sup>1,2</sup>  | John T. Finn<sup>2</sup> | Charles M. Schweik<sup>2</sup>  | Todd K. Fuller<sup>2</sup>  |  
 Douglas Sheil<sup>3</sup> 

<sup>1</sup>Institute of Tropical Forest Conservation, Mbarara University of Science and Technology, Kabale, Uganda

<sup>2</sup>Department of Environmental Conservation, University of Massachusetts, Amherst, Massachusetts, USA

<sup>3</sup>Forest Ecology and Forest Management Group, Wageningen University and Research, Wageningen, The Netherlands

## Correspondence

Dennis Babaasa, Institute of Tropical Forest Conservation, Mbarara University of Science and Technology, Kabale, Uganda.

Email: [dbabaasa@must.ac.ug](mailto:dbabaasa@must.ac.ug)

## Funding information

International Foundation for Science; Institute of Tropical Forest Conservation, Mbarara University of Science and Technology; British Ecological Society; University of Massachusetts Amherst; Mohammed bin Zayed Species Conservation Fund; Wildlife Conservation Society

Associate Editor: Jennifer Powers

Handling Editor: Laura Schneider

## Abstract

Conservation of mountain ecosystems can benefit from knowledge of habitats and their distribution patterns. This benefit is particularly true for diverse ecosystems with high conservation values such as the “Afromontane” rainforests. We mapped the vegetation of one such forest: the rugged Bwindi Impenetrable Forest, Uganda—a World Heritage Site known for its many restricted-range plants and animal taxa including several iconic species. Given variation in elevation, terrain and human impacts across Bwindi, we hypothesized that these factors influence the composition and distribution of tree species. To test this, detailed surveys were carried out using stratified random sampling. We established 289 georeferenced sample sites (each with 15 trees  $\geq 20$  cm dbh) ranging from 1320 to 2467 m a.s.l. and measured 4335 trees comprising 89 species that occurred in four or more sample sites. These data were analyzed against 21 digitally mapped biophysical variables using various analytical techniques including nonmetric multidimensional scaling (NMDS) and random forests. We identified six tree species assemblages with distinct compositions. Among the biophysical variables, elevation had the strongest correlation with the ordination ( $r^2=0.5$ ;  $p<0.001$ ). The “out-of-bag” (OOB) estimate of the error rate for the best final model was 50.7% meaning that nearly half of the variation was accounted for using a limited set of variables. We demonstrate that it is possible to predict the spatial pattern of such a forest based on sampling across a highly complex landscape. Such methods offer accurate mapping of composition that can guide conservation.

## KEY WORDS

Bwindi impenetrable Forest, elevation gradient, human disturbance, random forests, vegetation mapping

## 1 | INTRODUCTION

Mountain ecosystems often possess high conservation value due to their diverse and often geographically restricted biota. This diversity and rarity arises both from the sharp habitat gradients and numerous habitats often found within a compact region, and the role

of such highlands as refugia when climate varies. Understanding these systems, their biota and their needs and dependencies is challenging (Perrigo et al., 2020; Trew & Maclean, 2021). Mountain landscapes create complex interacting environmental gradients often covering broad ranges in temperature, moisture, illumination, nutrients, disturbance processes and histories (Ghazoul

& Sheil, 2010; Richter, 2008). Many environmental factors vary with elevation, leading to some characteristic zonation patterns (Hamilton, 1975; Hedberg, 1951; Schmitt et al., 2010). Plant communities are further structured at a local scale by topography, reflecting differences in soil depth, structure, moisture, nutrients, exposure and aspect contributing to local variations in suitability for different plants to establish and persist (Eilu et al., 2004; Ghazoul & Sheil, 2010; Lovett et al., 2001). Tropical species are known to be sensitive to disturbance and shifts in environmental conditions (Sheil, 2016; Ssali et al., 2017). For example, disturbance (the availability of open areas where trees can establish) have been shown to relax the constraints imposed by competition and extend effective elevation ranges of some species that are not constrained by short-range dispersal, particularly those in secondary forest, to warmer and cooler climates (Muñoz Mazón et al., 2019). Thus, montane vegetation spatial patterns are a result of complex interactions between the terrain, local conditions and histories (Hoersch et al., 2002), as well as interactions among species such as competition (Pfeffer et al., 2003). This makes the understanding of relationships between physical features of the environment and vegetation challenging.

Montane forest is the rarest vegetation type on the African continent (Linder, 2014; White, 1983), and also among the least studied (Desalegn & Beierkuhnlein, 2010; Sainge et al., 2019). This biologically rich and unique "Afromontane" ecosystem is increasingly threatened by deforestation, degradation and defaunation (Cordeiro et al., 2007; Plumptre et al., 2007), while global climate change has added new threats (Ayebare et al., 2013 and Ayebare et al., 2018; Wright et al., 2022). Deforestation is threatening the high aboveground carbon stock of montane forests (Cuni-Sánchez et al., 2021). This calls for prioritization of conservation actions to identify and protect vulnerable species and habitats. Mapping of forest vegetation can help to assess, plan, and guide conservation management in addressing these challenges by clarifying what taxa typically occur where and how these distributions are determined (Brinkmann et al., 2011; Fjeldsa, 2007). For example, variation in plant species composition along an elevation gradient is highly correlated with dietary variability (Elgart, 2010; Ganas et al., 2004) and variation in the female population genetic structure (Guschanski et al., 2008) of mountain gorillas (*Gorilla beringei beringei*).

Predictive vegetation mapping is defined as predicting the vegetation composition across a landscape from mapped environmental variables (Franklin, 1995). It is normally done using a combination of field data with digital maps of topography, as well as climatic limiting factors, facilitated by flexible, algorithmic modeling approaches and is driven by the need to map vegetation over large areas for conservation planning (Evans & Cushman, 2009; Hoersch et al., 2002; Pfeffer et al., 2003). Availability of biophysical predictor variables like elevation, temperature and precipitation mapped at relevant spatial scales and advances in species distribution modeling (Elith & Leathwick, 2009; Franklin, 2010; Hastie et al., 2009), provide opportunities to quantitatively analyze, predict and map the flora of biodiversity hotspots, if based on field data, or at least are validated

or informed by field data. This increases our understanding of the vegetation composition and plant community patterns in relation to environmental factors. Various approaches are available for predictive mapping of species distributions (Elith et al., 2006; Hijmans & Elith, 2016). We highlight three general approaches ranked according to their "function complexity" (Pearce & Ferrier, 2000): first, are the earliest and simplest methods that include Bioclim, Domain and Mahalanobis. These methods do not consistently perform well when compared with other approaches (Elith et al., 2006). Second, are the generalized regression models—the Generalized Linear Models (GLM; McCullagh & Nelder, 1989) and Generalized Additive Models (GAM; Hastie & Tibshirani, 1986). Though more complex than the first group, the generalized regression models provide mixed results because the nature of the relationships with the predictor variables may vary across the range of each species (Franklin, 1995; Moore et al., 1991). Third, are the nonparametric machine learning techniques that can reveal very complex patterns between species and physical features of the environment and can yield good predictive models when data are sufficient (Breiman, 2001; Prasad et al., 2006; Cutler et al., 2007). The earliest approaches were the Artificial Neural Networks (ANN) and Classification and Regression Trees (CART; Breiman et al., 1984). Later models include MaxEnt (Maximum Entropy; Phillips et al., 2006), Random Forests (RF) (Breiman, 2001), Boosted Regression Trees (Elith et al., 2009), Multivariate Adaptive Regression Splines (MARS; Friedman, 1991) and Support Vector Machines (SVM; Guo et al., 2005). Comparisons among such options found that RF perform best in mapping current and future species distributions (Prasad et al., 2006).

Many machine-learning based vegetation mapping studies have focused on temperate forests, but rarely have they been applied to explore the complex, highly mixed and species-rich rainforests in the tropics (Lin et al., 2020). Yet, accurate maps of vegetation can aid management and contribute to various research as well as identifying restricted and vulnerable communities and habitats to support priority conservation planning (Brinkmann et al., 2011; Sainge et al., 2019). While numerous studies have focused on large geographical areas, fewer studies have identified and mapped vegetation communities along environmental gradients in smaller landscapes or mountains (Latt & Parker, 2022). However, local studies contribute to prioritization of areas in need of urgent conservation action particularly in biodiversity hotspots at high risk of habitat degradation and loss (Fjeldsa, 2007; Seddon et al., 2020).

Bwindi Impenetrable Forest in SW Uganda (henceforth "Bwindi") is well suited, as a challenging case, for evaluating the performance of predictive mapping. It is one of the few forests in all of Africa where lowland and montane forests are in a continuum (Hamilton, 1974 and Hamilton 1975). The forest has a richer tree diversity compared to other forests in the ecoregion, attributed to high rainfall and soil characteristics (Eilu et al., 2004). The terrain of Bwindi is extremely rugged with high topographic diversity (Howard, 1991; Leggat & Osmaston, 1961). This limits conventional vegetation mapping approaches such as aerial photo interpretation or data derived from satellite images as their spatial resolution data is often insufficient

for mapping vegetation in specific areas such as gullies and steep slopes where topographic shadows mask the vegetation (Pfeffer et al., 2003).

The forest also lies in an area with one of the highest rural human population densities in Africa (Cordeiro et al., 2007). Because of this, little natural forest persists outside the boundaries of the park (Twongyirwe et al., 2011). Heavy human disturbance of the past within the forest greatly modified the vegetation structure and ecology (Babaasa et al., 2004; Sheil, 2012; Ssali et al., 2017). Before it became a park, the forest was subjected to a broad gradient of human disturbance with pit-sawing and poaching being the most widely distributed in the forest (Howard, 1991; McNeilage et al., 1998), while mining (gold and tungsten) was concentrated in specific locations (Butynski, 1984; Harcourt, 1981). Human-induced wildfires and harvesting of non-timber products impacted the forest periphery (Butynski, 1984; Cunningham, 1996), while numerous human trails and two public roads traversed the forest (Butynski, 1984; Harcourt, 1981). Human disturbance across the forest varied greatly in intensity, distribution and time periods (McNeilage et al., 2001). Ongoing disturbance processes include localized landslides, forest elephants (*Loxodonta cyclotis*), occasional wildfires following long periods of drought, harvesting of poles and wire snaring (Babaasa, 2000; Hickey et al., 2019; Olupot et al., 2009), and a public road cutting through the narrow forest corridor and high elevation zone (Barr et al., 2015). This complexity of factors contributes to making Bwindi a complex site for investigations of plant communities.

Numerous studies have demonstrated variations in rainforest composition with elevation (Eilu et al., 2004; Hamilton, 1975; Lieberman et al., 1996; Lovett et al., 2001), topography (Werner & Hormeier, 2015; Jucker et al., 2018) and with human disturbance (Chapman & Chapman, 1997; Plumptre, 1996; Ssali et al., 2017). Given the wide variation in elevation, topography and past and current human disturbances across Bwindi Forest and evidence from elsewhere, we hypothesized that elevation, topography, and human disturbance, either acting in isolation or in combination, would play major roles in determining the composition and distribution of tree species and communities. We carried out extensive ground surveys followed by multivariate statistical techniques to evaluate a variety of environmental variables as potential predictive attributes to test this hypothesis.

## 2 | METHODS

### 2.1 | Study area

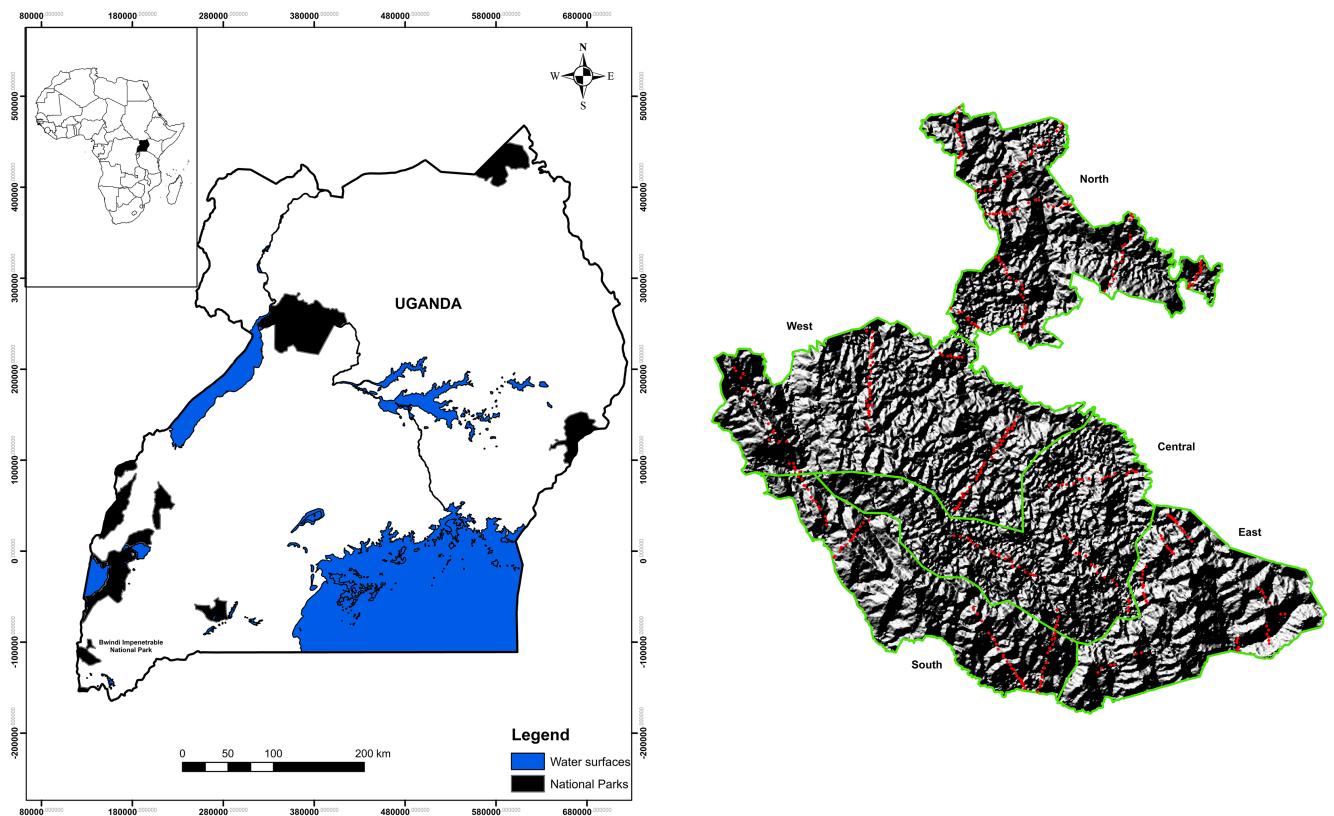
Bwindi Forest is located in SW Uganda (Figure 1), in the Kigezi Highlands at the eastern edge of the Albertine Rift (latitude 0°53'-0°8'S and longitude 29°35'- 29°50' E). Covering an area of 331 km<sup>2</sup>, Bwindi lies at the North West end of the Kigezi Highlands which are associated with the up-warping and faulting during the formation of the Albertine Rift and are underlain by Precambrian shale, phyllite, quartz, quartzite, schist and granite of the Karagwe-Ankolean System

(Leggat & Osmaston, 1961). The Government of Uganda (1967) classified the soils of Bwindi as belonging to the "non-differentiated humic ferrallitic" types and having been derived from the foregoing geological formation, and they are composed of mainly tropical red earths with an overlying layer of brown to black spongy humus. The erosive action of the numerous rivers in their youth stage within the Highlands has caused the topography of the park to be extremely rugged with narrow, steep-sided valleys and deep gullies that run in various directions, bound by emergent hill crests lying between 1190 m in the northwest and 2607 m a.s.l. in the southeast (Howard, 1991). The forest has been classified as a 'moist lower montane forest' (Hamilton, 1994), ranging from near the upper boundary of lowland forest to montane forest (Hamilton, 1975).

Information on Bwindi's vegetation cover is limited. A forest type map prepared by Cahusac (1961) for managing timber harvesting operations is now outdated given the management history of Bwindi that include more than four decades (from 1947 to 1991) of intense and forest-wide timber harvesting (Howard, 1991; McNeilage et al., 1998), and whose long-term impact was the creation of a diverse patch mosaic of vegetation types differing in age, hence succession, and broken canopy cover (Babaasa et al., 2004; Sheil, 2012). Incessant, heavy human and natural disturbances have prevented reclosure of the forest canopy (Ssali et al., 2017). Previous descriptions of the tree flora of Bwindi were based on general tree species inventory and limited to a small area of high elevation in the southeast (Hamilton, 1969; Howard, 1991; Leggat & Osmaston, 1961) or low elevation in the north (Eilu et al., 2004). Later work on trees (Davenport et al., 1996), though covering more representative areas of the forest, recorded only the species encountered but did not georeference the sample sites, preventing spatial analyses and extrapolation. Until now, it had remained unclear how forest-wide floristic patterns were spatially structured in relation to the environment.

### 2.2 | Study design

To account for the different environmental conditions, we employed a stratified random sampling approach. The park was divided into five strata based on geological formations visible on the Digital Elevation Model (DEM) of Bwindi and the starting points on the boundary of the DEM in each stratum were selected randomly with the random point function within ArcGIS (version 10.5; ESRI, Redlands, CA, USA). Line transects were drawn on the DEM in each stratum from the random starting points to traverse the topographic positions of the ridges (Figure 1; Table 1). The number and length of the transects selected varied with area, accessibility and shape of the strata. We then superimposed the transect drawings on high-resolution (0.5 m) true color, digital aerial photographs of Bwindi. The aerial photos were visually interpreted along the transects by drawing polygons around areas perceived to be of uniform tree community structure based on differences in tone and texture. This allowed the sample sites to be placed in what we perceived to be distinct tree communities.



**FIGURE 1** Location of Bwindi Impenetrable National Park, Uganda, and of tree sampling transects and sample sites (red dots) among the strata superimposed on a Digital Elevation Model of the study area.

Stratum	Transect elevation range (m.a.s.l.)	No. of sample points				
		Valley	Hillside	Ridge	Gully	Total
Central	2014-2237	8	31	16	1	56
East	2020-2467	1	59	2	1	63
South	1479-2342	3	50	11	2	66
North	1320-1732	3	48	6	1	58
West	1479-2124	2	31	10	3	46
All	1320-2467	17	219	45	8	289

**TABLE 1** Sampling design for trees ( $\geq 20$  cm dbh) for floristic gradient modeling in Bwindi Impenetrable National Park, Uganda.

### 2.3 | Tree species sampling

We carried a printed copy of the digitized polygons, overlaid with a coordinate grid, and used a hand-held Global Positioning System (GPS) device to locate the digitized polygons in the field. A single random point within each digitized polygon was selected for tree species sampling. At each random sample point, we used the point-to-tree distance technique or plotless sampling method to sample trees (Hall, 1991; Klein & Vilcko, 2006; Sheil et al., 2003). This technique involved selecting the nearest 15 trees ( $\geq 20$  cm dbh) around the random center-point. The selected trees were identified to species level and we measured the diameter at breast height (dbh) of each individual. We named the tree species following nomenclature used in Kalema and

Hamilton (2020). The distance from the sample site center-point to the 15th farthest tree was measured and regarded as the sample site radius. This procedure is suitable for rapid and robust assessments of vegetation where tree density varies, such as in patchy and disturbed tropical forest (Klein & Vilcko, 2006; Sheil et al., 2003). At each center-point, eight environmental attributes were recorded: aspect—as the compass direction facing down slope; and steepness of the slope using a clinometer. Untransformed aspect and slope are poor for quantitative analysis, so slope was transformed to a more suitable index by taking the sine of the slope in degrees; aspect was also transformed into a suitable index by taking the negative cosine of the angle in degrees minus 35 (McCune & Grace, 2002). Four physiographic positions of valley, hillside, ridge tops and gully were simply

recorded as "1" if the sample site was in that physiographical class and "0" otherwise. The final recorded site characteristics consisted of spatial variables—the Universal Transverse Mercator (UTM) coordinates of easting and northing (datum WGS 84) using a hand-held GPS unit, standardized to zero mean and unit variance. All the data were collected at 289 sample sites spread across the forest (Figure 1).

## 2.4 | Biophysical predictor variables

We acquired 20 digitally mapped biophysical variables summarized in Table 2 and detailed in Appendix S1. These were: the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) 30×30m Version 2 (<http://gdem.ersdac.jspacesystems.or.jp/>), and nineteen 1-km-scale (30sec) bioclimatic variables, that include 11 temperature and eight precipitation variables from WorldClim version 2 (<http://worldclim.org/version2.1>; 1970–2000; Fick & Hijmans, 2017). We clipped the ASTER GDEM and the 19 bioclimatic rasters to a window covering Bwindi Forest only. Lastly, the human disturbance across the park was based on the combined relative encounter of human activity signs that were likely to have an impact on vegetation – wood cutting, bee hives, old pitsaw sites, disused mineral extraction pits, snares, and fireplaces (McNeilage et al., 1998). We used the geo-references of all the 289 sample sites to extract site values from the predictor rasters. We constructed an environmental matrix of extracted variable values together with site measurements (aspect, slope, topographic position and x, y point coordinates) and tested for pairwise collinearity and one of any pair of highly correlated variables (Pearson  $r \geq 0.75$ ) discarded (Table S1).

TABLE 2 A summary of potential explanatory variable names and descriptions for floristic gradient modeling in Bwindi Impenetrable National Park, Uganda.

Biophysical variable name	# of variables	Reference
Temperature	1	Fick and Hijmans (2017)
Precipitation	8	Fick and Hijmans (2017)
30m Digital Elevation Model	1	NASA & METI (2011)
Human disturbance	1	McNeilage et al., (1998)
Steepness of the slope	1	Field measurement
Aspect	1	Field measurement
Ridge tops	1	Field measurement
Valley	1	Field measurement
Gully	1	Field measurement
Hillside	1	Field measurement
Latitude	1	Field measurement
Longitude	1	Field measurement

## 2.5 | Data analyses

Based on field data, a sample sites-versus-species matrix (using basal area values of each tree species relative to the area of the sample site [ $m^2 ha^{-1}$ ]) was created. We only considered tree species occurring in more than four sample sites, resulting in a 289 sample site by 89 tree species matrix. The data were natural log-transformed following the generalized procedure (McCune & Grace, 2002) to minimize the influence of large trees. We subjected the data matrix to various multivariate techniques to identify associations and spatial patterns among the tree species using R software (version 4.2.1; R Core Team, 2022). To describe the different tree species assemblages within our study area, we clustered the sample sites with similar tree species that tend to occur together using a polythetic, agglomerative hierarchical clustering (HC) with the flexible beta linkage method ( $\beta = -0.25$ ; Lance & Williams, 1966; Legendre & Legendre, 1998) with the agnes function and Bray-Curtis as a distance measure. The clustering results were portrayed by a dendrogram or clustering tree. We determined the optimal number of clusters using Indicator Species Analysis (ISA) procedure, described below. Because we had numerous sample sites ( $n = 289$ ), we simplified the presentation using composite sample sites from the original sample sites by computing the centroid of each of the cluster as the mean of the basal area/ha of each tree species per cluster.

We used Multi-Response Permutation Procedures (MRPP) and a Mantel's test to test for differences in composition between the clusters. We also utilized ISA, a method that combines frequency and mean abundance tables, to identify the characteristic tree species for each cluster. Statistical significance of the indicator tree species was determined by random permutations for each species at  $p < .05$  significance level. Indicator values vary from 0 (no indication) to 1 (perfect indication).

A Kruskal's nonmetric multidimensional scaling (NMDS) based on Bray-Curtis coefficient was used to evaluate the ecological tendencies reflected in the cluster dendrogram and the relationship between sample sites, clusters and environmental variables. The sample sites were ordinated, then overlaid with the cluster centroids from cluster analysis with surrounding confidence ellipses at  $\pm 2$  SD from the mean (enclosing approximately 95% of sample sites within each cluster). Lastly, 12 least correlated site environmental variables (Table 2) were fitted onto the ordinations using 1000 permutations.

We applied RF techniques to predict and map the spatial distribution of the clusters (Evans & Cushman, 2009; Lin et al., 2020; Prasad et al., 2006). RF is a data mining technique that should produce accurate predictions without overfitting the data (Breiman, 2001; Breiman & Cutler, 2003). In R software, it is implemented in the function 'randomForest' in a package with same name (Liaw & Wiener, 2002). In RF, bootstrap samples are drawn to construct multiple trees; each tree is grown with a randomized subset of predictors, hence the name 'random' forests. A large number of trees (500 to 2000) are grown, hence a 'forest' of trees. The number of predictors used to find the best split at each node is a random subset

of the total. The trees are grown to maximum size without pruning, and aggregation is by averaging the trees. Out-of-bag (OOB) samples are used to calculate an unbiased error rate and variable importance, eliminating the need for a test set or cross-validation. Because a large number of trees are grown, there is limited generalization error (that is, the true error of the population as opposed to the training error only), which means that no overfitting is possible, a useful feature for prediction. By growing each tree to maximum size without pruning and selecting only the best split among a random subset at each node, RF maintains some ability for prediction (Breiman, 2001). Random predictor selection diminishes correlation among unpruned trees and reduces bias; by taking an ensemble of unpruned trees, variance is also moderated. RF provides several metrics that aid in interpretation. The importance of each predictor variable is evaluated based on how much worse the prediction would be if the data for that predictor were permuted randomly. Since our response variable, the clusters, was a factor (categorical), we performed a classification procedure in the analysis. Six least correlated variables (Pearson's  $r < 0.75$ ) that are digitally mapped for the whole park: isothermality, temperature seasonality, minimum temperature of the coldest month, annual precipitation, mean temperature of the coldest quarter and elevation were tested as predictors for cluster distribution.

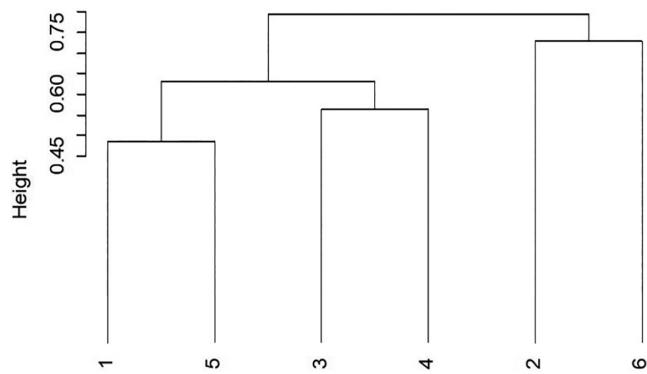
### 3 | RESULTS

We recorded 4335 individual trees ( $\geq 20$  cm dbh) comprising 121 species in 51 families from 289 sample sites. The sample sites spanned an elevation of 1320 to 2467 m.a.s.l. broadly representing Bwindi's elevation range and topographic variation. Just 89 tree species from 47 families occurred in four or more sample sites and were included in the subsequent analyses. The richest two sample sites included 12 species while the poorest sites were 11 each with four species. The commonest tree was *Strombosia scheffleri* that occurred in 130 (45%) sample sites, while the least common species were nine (*Anthocleista vogellii*, *Antiaris toxicaria*, *Casearia battiscombei*, *Celtis durandii*, *Dichaetanthera corymbosa*, *Ficus sur*, *Hanno longipes*, *Memecylon myrianthum* and *Pauridiantha callicarpoides*) each occurring in only four (1.4%) sample sites.

The cluster dendrogram (Figure 2) grouped the sample sites into six clusters. This cluster-stage yielded the lowest average summed  $p$ -values and the highest number of significant indicator species based on ISA (Figure 3). The six groups were further grouped in three pairs, with clusters 1 and 5 being most similar in composition, while clusters 2 and 6 were more distinct.

The MRPP test showed that the clusters were significantly different from random association (observed  $\delta = 0.79$ , expected  $\delta = 0.89$ ,  $p < .001$ ). The Mantel test ( $r = 0.32$ ,  $p < .001$ ) also indicated that the six clusters differed significantly in composition.

In total, 63 indicator tree species proved significant for the six clusters (summarized in Table 3 and detailed in Table S2). All the species, with the exception of *Gambeya albida* in cluster 4, had indicator



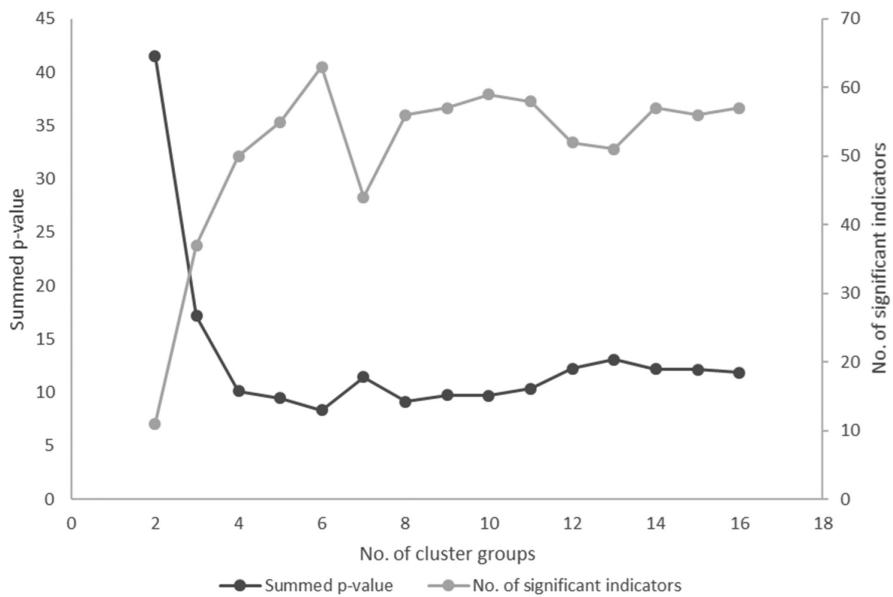
**FIGURE 2** Polythetic, agglomerative, hierarchical clustering dendrogram depicting the relationships between the composite clusters. Numbers correspond to the clusters in Table 3 and 4, Figure 4; Appendix S2 and S3.

values of less than 50% implying that they also occurred, at lower abundance and frequency, in other clusters. The clusters had numerous indicator species with clusters 2, 5 and 6 having the largest number (14, 14 and 18, respectively), while clusters 1, 3 and 4 had the fewest (6, 7 and 4, respectively). Twenty-four species were pioneers (early successional) and mid-successional species (non-pioneer light demanders, NPLDs), spread across the clusters.

NMDS resulted in a 3-axis optimal solution and a good fit with a clear positive relationship between observed community dissimilarity and ordination distances. The NMDS ordination diagram produced a dense cluster of sample site points and HC analysis confidence ellipses that greatly overlapped on all the three axes (Figure S1). The sample site points that were positioned where the ellipses intersected did not belong exclusively to one cluster but had compositions indicative of two or more clusters. The overall patterns show a complex range of variables that show consistent and nonrandom relationships to species composition and distribution ( $p < .001$ ). The strongest correlations were observed for elevation and longitude, but ridge-top, precipitation, temperature and human disturbance factors also showed a nonrandom role (Figure S2). Since elevation and longitude arrows were close and pointed in the same direction, it means they are positively correlated implying that clusters 2, 3, 4 and 6 positioned along the direction of the arrows are depicted to be at high elevation to the east of the park.

The RF predictive map for the six clusters (Figure 4) and description of the clusters (Table 3 and Table S2), revealed distinct vegetation geographically arranged along the north-west and south-east latitudinal and longitudinal axes of Bwindi. Cluster 1 (white) and 5 (light green) were largely in areas north and west of Bwindi, while cluster 2 (pink) and 6 (dark green) were primarily in the extreme southeasterly areas. Clusters 3 (yellow orange) and 4 (yellow green) were in the middle of the forest. However, in some areas, the clusters formed mosaics, with patches of cluster 6 within clusters 2 and 5, while those of cluster 2 were within cluster 5, and those of cluster 1 within clusters 2 and 5. We found no evidence of any of our six clusters being more or less associated with forest edges or large gaps located a distance from the forest edges.

**FIGURE 3** Changes in summed p-values and number of indicator tree species with  $p \leq .05$  from randomization tests across 2–16 clusters.



**TABLE 3** Indicator tree species for the clusters in Bwindi Impenetrable National Park, Uganda.

Cluster number	# Sample sites	Elevation range; mean elevation (m.a.s.l.); and location of cluster	# Indicator tree species	# Pioneer and NPLD species
1	58	1372–2350; 1956 Mid-elevation transitional forest in the center and west of the park	6	3
2	35	1778–2407; 2195 Submontane forest on ridge-tops in the centre-south of the park	14	7
3	45	1479–2217; 2023 Mid-elevation forest in the center-south of the park	7	2
4	33	1773–2311; 2188 Sub-montane forest in the center-south of the park	4	2
5	75	1320–2131; 1696 Low-elevation forest in the north, forest corridor and west of the park	14	5
6	43	1416–2467; 2203 High elevation montane forest to the southeast of the park	18	5

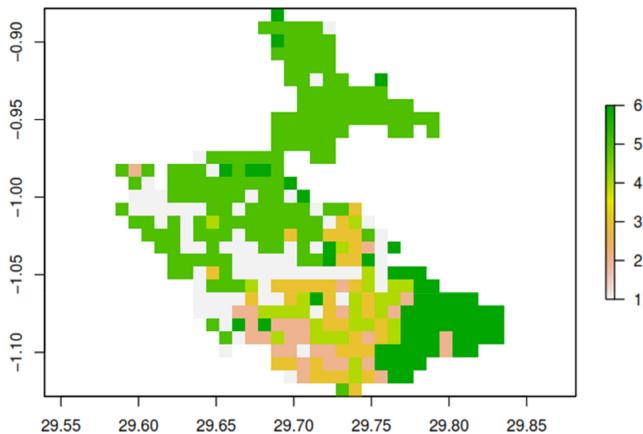
Estimated importance of the variables in predicting cluster composition and spatial distribution estimates revealed that elevation had the greatest estimated importance in terms of both accuracy and Gini score (Figure 5). Only two other variables had importance of nearly similar magnitude for predictive accuracy: precipitation of driest month and annual precipitation. The OOB estimate of the error rate was 50.7% (Table 4).

## 4 | DISCUSSION

Our study shows that the methods we used can be useful for ecological understanding of mountain habitats. We demonstrate that the RF model, supported by multivariate statistical techniques, was effective in delineating tree communities and predictively mapping them in response to complex topographic gradients,

compounded by vegetation that is highly mixed, species-rich with a heavy and long disturbance history. For the final result to be optimal, all our data were used in training the model instead of forfeiting them for an independent validation. Thus, we chose to use the OOB estimate of the error rate provided in RF. Our OOB error statistic was moderately high. Brinkmann et al. (2011) and Lin et al. (2020) attribute inaccuracies of predicted maps to mapped predictor variables being too coarse to provide exact variable estimates for each sample site, especially, in hilly and topographically diverse areas or noninclusion of causal factors (Barbet-Massin & Jetz, 2014; Bedia et al., 2013). But more important could be the observation of Volkov et al. (2003) that the distribution of plant abundances in natural communities at local scales is (or can be considered in effect to be) largely random with species being ecologically equivalent, or structured more by dispersal than by differences in abiotic conditions. Pfeffer et al. (2003) suggest that most

plant species are tolerant to a wide range of varying conditions, and therefore, there can never be a perfect species-environment correlation. Moreover, species data tend to be redundant and noisy (Kent, 2012; Ter Braak, 1995) making it difficult to completely explain how a rich community of vegetation is determined by a limited set of interacting factors (Pfeffer et al., 2003). Nonetheless,

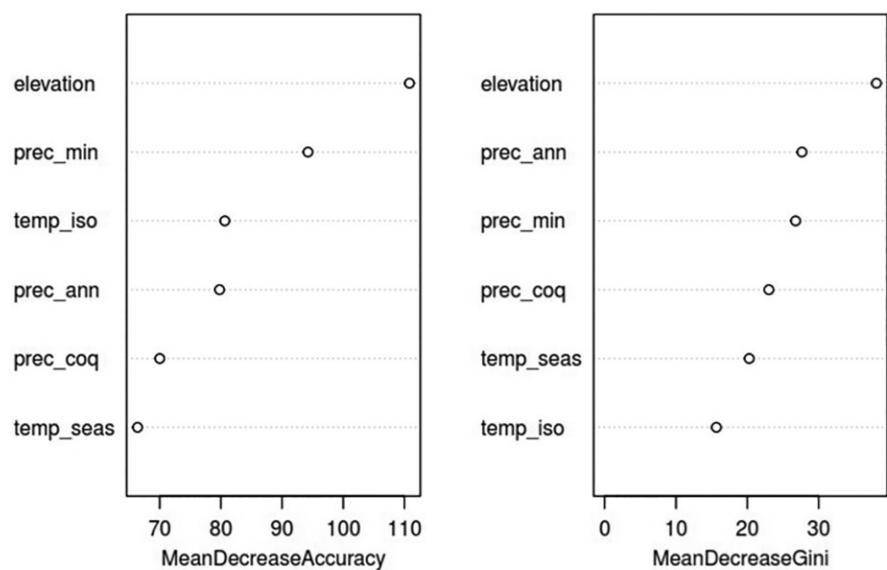


**FIGURE 4** Predictive map of tree species assemblage spatial patterns in Bwindi Impenetrable National Park, Uganda. Numbers in the legend correspond to the clusters in [Tables 3 and 4](#), [Figure 2](#); [Appendix S2](#) and [S3](#).

the most influential variables indicated by the RF model are those expected to influence large scale patterns of montane vegetation (Hedberg, 1951; Lieberman et al., 1996; Eilu et al., 2004; Schmitt et al., 2010; Sainge et al., 2019; Lin et al., 2020). This, together with the fact that at least half of the tree community variation was accurately predicted using a limited set of predictor variables, permits us to predictively map the vegetation at unsampled locations with reasonable confidence.

Our approach to vegetation mapping in mountains provides an alternative to the sometimes limited approaches provided by remote sensing (Pfeffer et al., 2003). Remote sensing is poor at mapping plant communities that merge along elevational gradients (Singh et al., 2001; Townsend & Walsh, 2001). In addition, remote detection of specific plant species is most likely to be viable if the target species possess unique growth forms, phenology or other readily detectable characteristic (He et al., 2015). Vegetation types with the same overall physiognomy or plant form like trees, but with varied in floristic composition are difficult to differentiate using both spectroscopy and remote sensing, resulting in misinterpretation and misclassification of remote sensing images (Townsend & Walsh, 2001).

Several other factors that can influence vegetation composition of communities were not considered in our model. Inclusion of geology, edaphic conditions, hydrological regime, interspecific



**FIGURE 5** Variable importance plot generated by RF algorithm showing the ranked importance of the model's predictive factors, measured using the mean decrease accuracy and mean decrease in gini index. The labels indicate: prec\_min = precipitation of the driest month; prec\_ann = annual precipitation; prec\_coq = precipitation of the coldest quarter; temp\_isol = isothermality; temp\_seas = precipitation seasonality.

Cluster	1	2	3	4	5	6	Class error %
1	14	6	9	7	16	5	75.4
2	3	15	6	4	2	5	57.1
3	5	8	20	5	6	0	54.4
4	3	8	5	16	1	0	51.5
5	10	1	5	2	50	4	30.5
6	2	4	1	2	8	24	41.5
OOB estimate of error rate							50.7

**TABLE 4** Confusion matrix of OOB estimate of error rate for the predictive tree assemblage spatial distribution pattern map of Bwindi Impenetrable National Park, Uganda. The cluster number corresponds to those in [Table 3](#), [Figures 2 and 4](#).

interactions, dispersal ability, and biotic or abiotic interactions (Godsoe & Harmon, 2012) might enable more accurate description of the vegetation pattern if suitable data were available.

A combined HC analysis and ISA identified six tree communities across Bwindi each with their own indicator species. The numerous indicator species for each tree community, without a single species being dominant, is indicative of a large proportion of the forest being secondary vegetation communities. In many tropical forests, single-species dominance is a characteristic of mature (climax) communities (Eggeling, 1947; Connell & Lowman, 1989; Sheil & Burslem, 2003). Furthermore, close to 40% of the indicator species, spread across the tree communities, were pioneers such as *Macaranga barteri* at low elevation and *Polyscias fulva* and *Neoboutonia macrocalyx* at high elevation and mid-successional or non-pioneer light demanders (NPLDs) such as *Gambeya albida*, *Newtonia buchananii*, and *Entandrophragma excelsum*. This is further evidence that considerable areas of the forest are in early succession stages. Human disturbances of the past, particularly pitsawing, were spread throughout the forest (Howard, 1991; McNeilage et al., 1998). However, the disturbances varied in intensity and were concentrated in different areas at different time periods (McNeilage et al., 2001). This led the forest vegetation to be structurally and compositionally complex, with a mosaic of patches of disturbed and regenerating forest, differing in size and succession stage, being superimposed on differences caused by topography and possibly soil characteristics. It could be a very long time before the forest regenerates to the primary (climax) stage. Aside from the past disturbances, there are additional reasons why much of Bwindi's vegetation remains in early successional stages. Ssali et al., (2017) found that forest regrowth in the many gaps dominated by bracken fern scattered across the forest is impeded in multiple ways including repeated damage and high seed predation. The prevalence of pioneer species and NPLDs we found across the forest likely reflect similar influences (Chapman & Chapman, 1997). The observed species have longer seed dormancy, greater growth potential and are less constrained by dispersal and seed predation (Muñoz Mazón et al., 2019; Sheil, 2016). Late successional species may take long periods to reestablish due to of the scarcity of remaining seed sources and limited dispersal (Ssali et al., 2017).

Ordination analysis demonstrated that the tree communities were mainly arranged along two collinear gradients—elevation and sample site's position in the forest landscape (i.e., longitude). Bwindi is elongated from north-west to south-east, exhibiting a gradual rise in mean elevation in this direction. Although elevation had the strongest relationship with tree composition, additional variability was explained by topographic position and climate (temperature and precipitation). Topography has multiple potential mechanisms through which to impact vegetation: for example, the soils of slopes and lower valleys are generally moister, better structured, and richer in plant nutrients than the more drought-prone and better illuminated ridge tops (Ghazoul & Sheil, 2010; Jucker et al., 2018). For example, before it was sought out and harvested in accessible areas, the slow-growing and

shade-tolerant *Podocarpus latifolius* was common on Bwindi's hill-ridges >2000 m.a.s.l. (Hamilton, 1994). Our results, as expected, show that temperature tends to decline as elevation increases (Lieberman et al., 1996) with tolerance of low-temperatures often offset by lower competitive ability under warmer conditions (Sheil, 2016; Slik et al., 2009). Our RF model indicated that annual precipitation and precipitation of driest month also contribute to differences in tree composition. While many variables are clearly related to tree community patterns these are often correlated and cannot be unambiguously assigned to distinct mechanisms so their individual contributions could not be analytically separated.

Whilst the influence of elevation on tree community composition and distribution was systematically identified and statistically demonstrated, the tree communities in Bwindi do not form distinct zones along the elevation gradient. This accords with results from other montane forests such as Udzungwe Mountains National Park, Tanzania (Lovett et al., 2006). There are multiple possible reasons for lack of elevation zoning. We highlight two: Bwindi is one of the few forests in Africa where lowland and highland forests are in a continuum (Hamilton, 1974 and Hamilton 1975). In a rather limited area, there is intermingling of lowland tree species that reach their upper boundary and highland tree species at their lower boundary. Our results showed an overlap of the cluster confidence ellipses on all the 3-dimensions of the ordination diagram and numerous sample sites and species positioned where the ellipses intersect indicating that they can reasonably belong to more than one community (Kent, 2012; Townsend & Walsh, 2001). In addition, lack of perfect indicator species (having indicator value  $\leq 0.5$ ) implies that numerous tree species are spread widely but are most frequent and abundant within the tree community they indicate, presumably representing their ecological optimum (Austin, 2013; Mueller-Dombois & Ellenberg, 2002). The continuous nature of change of plant species in moist forest communities is related to a continuous change in environmental variables (i.e., precipitation and temperature) along the elevation gradient (Lovett & Lindberg, 1993). Secondly, the intense forest-wide human disturbance of the past that involved mainly large-scale species-selective cutting and removal of large hardwoods could have relaxed the constraints imposed by competition, making it possible for some tree species to colonize the disturbed areas resulting in extension of their elevation range limits (Sheil, 2016; Muñoz Mazon et al., 2019). In addition to natural disturbances such as drought and occasional landslides, human disturbances such as clearing, timber cutting, fires and many others have occurred at different times in different parts of the forest, leading to spatial patterns that obscure what would be expected in a system lacking human impacts (Chapman & Chapman, 2004).

## 5 | CONCLUSIONS

Prediction of vegetation across an extensive mountain landscape using mapped environmental variables offers a potential in addition

to remote sensing. Remote sensing is challenging in species rich mountainous regions like Bwindi with considerable topographic complexity and a multifaceted history of natural and human impacts. Efforts should be geared towards accessing and refining predictor variables at the appropriate scales so as to further improve such mapping.

We see that Bwindi is a mosaic of patches of different successional stages resulting from a complex history. Recovery to late successional forest is slow, likely due to seed limitation and continued disturbances. Small island forests such as Bwindi, surrounded by dense human populations, are under immense pressure and their long-term viability and survival remain in doubt. Mapping tree species composition and patterns helps understand the extent, variation and ecological impacts of past and ongoing disturbance and other factors. Such information can assist protected area management to direct conservation efforts so as to avoid or reverse further degradation of the fragile ecosystem and monitor restoration programmes. Also, the information can be used to model wildlife abundance and distribution, estimation of forest on-ground carbon stocks and to examine the consequences of various climate change scenarios. Predicting changes in composition is instrumental to inform adaptive management strategies and conserving biodiversity.

## AUTHOR CONTRIBUTIONS

DB conceived the idea, designed the study and collected field data; DB, JF, and CS collaborated on the data analyses and manuscript writing; TK and DS reviewed and edited the manuscript; all authors discussed and approved the manuscript.

## ACKNOWLEDGMENTS

This study was supported by the Institute of Tropical Forest of Mbarara University of Science and Technology, Wildlife Conservation Society, the International Foundation for Science, the Mohammed bin Zayed Species Conservation Fund, the British Ecological Society, and the University of Massachusetts, Amherst. We thank two anonymous reviewers for their useful comments on a previous version of this manuscript, and the Subject Editor Laura Schneider for the insightful suggestions which allowed us to improve the present paper.

## CONFLICT OF INTEREST STATEMENT

The authors have no relevant financial or non-financial interests to disclose.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the Dryad Digital Repository: <https://doi.org/10.5061/dryad.rv15dv4fp> (Babaasa et al., 2024).

## ORCID

Dennis Babaasa  <https://orcid.org/0000-0002-4855-4534>

Charles M. Schweik  <https://orcid.org/0000-0002-0199-3740>

Todd K. Fuller  <https://orcid.org/0000-0002-4805-5460>

Douglas Sheil  <https://orcid.org/0000-0002-1166-6591>

## REFERENCES

Austin, M. P. (2013). Vegetation and environment: Discontinuities and continuities. In E. van der Maarel & J. Franklin (Eds.), *Vegetation ecology* (2nd ed., pp. 71–106). Wiley-Blackwell.

Ayebare, S., Plumptre, A. J., Kujirakwinja, D., & Segan, D. (2018). Conservation of the endemic species of the Albertine rift under future climate change. *Biological Conservation*, 220, 67–75. <https://doi.org/10.1016/j.biocon.2018.02.001>

Ayebare, S., Ponce-Reyes, R., Segan, D. B., Watson, J. E. M., Possingham, H. P., Seimon, A., & Plumptre, A. J. (2013). *Identifying climate resilient corridors for conservation in the Albertine rift*. Unpublished report by the Wildlife Conservation Society to MacArthur Foundation.

Babaasa, D. (2000). Habitat selection by elephants in Bwindi impenetrable National Park, Uganda. *African Journal of Ecology*, 38, 116–122. <http://ir.must.ac.ug/xmlui/handle/123456789/1683>

Babaasa, D., Eilu, G., Kasangaki, A., Bitariho, R., & McNeilage, A. (2004). Gap characteristics and regeneration in Bwindi impenetrable National Park, Uganda. *African Journal of Ecology*, 42, 217–224. <https://doi.org/10.1111/j.1365-2028.2004.00519.x>

Babaasa, D., Finn, J. T., Schweik, C. M., Fuller, T. K., & Sheil, D. (2024). Predictive mapping of tree species assemblages in an African montane rainforest. *Biotropica*, 00, 1–13. <https://doi.org/10.1111/btp.13302>

Barbet-Massin, M., & Jetz, W. (2014). A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. *Diversity and Distributions*, 20, 1285–1295. <https://doi.org/10.1111/ddi.12229>

Barr, R., Arrea, I.B., Asuma, S., Masozera, A.B., & Gray, M. (2015). *Pave the impenetrable? An economic analysis of potential Ikumba-Ruhija road alternatives in and around Uganda's Bwindi impenetrable National Park*. Conservation Development Strategy. Conservation Strategy Fund Technical Series No. 35. Sebastopol, CA.

Bedia, J., Herrera, S., & Gutierrez, J. M. (2013). Dangers of using global bioclimatic datasets for ecological niche modeling. Limitations for future climate projections. *Global and Planetary Change*, 107, 1–12. <https://doi.org/10.1016/j.gloplacha.2013.04.005>

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>

Breiman, L., & Cutler, A. (2003). Setting up, using, and understanding Random Forests. <http://www.stat.berkeley.edu/users/breiman/rf.html>

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. Chapman & Hall/CRC.

Brinkmann, K., Patzelt, A., Schlecht, E., & Buerkert, A. (2011). Use of environmental predictors for vegetation mapping in semi-arid mountain rangelands and the determination of conservation hotspots. *Applied Vegetation Science*, 14, 17–30. <https://doi.org/10.1111/j.1654-109X.2010.01097.x>

Butynski, T. M. (1984). *Ecological survey of the impenetrable Forest (Bwindi) forest Uganda, and recommendations for its conservation and management*. Unpublished report to the Uganda Government.

Chahusac, A. B. (1961). Working Plan for the Impenetrable CFR. In G. J. Leggat & H. A. Osmaston (Eds.), *Impenetrable C.F.R. Forest types (interpretation report 6). Reference map No. K1/45 appendix G* (pp. 1961–1971). Uganda Forest Department.

Chapman, C. A., & Chapman, L. J. (2004). Unfavourable successional pathways and conservation value of logged tropical forest. *Biodiversity and Conservation*, 13, 2089–2105. <https://doi.org/10.1023/B:BIOT.0000040002.54280.41>

Chapman, C. J., & Chapman, L. J. (1997). Forest regeneration in logged and unlogged forests of Kibale National Park, Uganda.

*Biotropica*, 29, 396–412. <https://doi.org/10.1111/J.1744-7429.1997.TB00035.X>

Connell, J. H., & Lowman, M. D. (1989). Low-diversity tropical rain forests: Some possible mechanisms for their existence. *The American Naturalist*, 134(1), 88–119. <https://doi.org/10.1086/284967>

Cordeiro, N. J., Burgess, N. D., Dovie, D. B., Kaplin, B. A., Plumptre, A. J., & Marrs, R. (2007). Conservation in areas of high population density in sub-Saharan Africa, editorial. *Biological Conservation*, 134, 155–163. <https://doi.org/10.1016/j.biocon.2006.08.023>

Cuni-Sanchez, A., Sullivan, M. J. P., Platts, P. J., Lewis, S. L., Marchant, R., Imani, G., Hubau, W., Abiem, I., Adhikari, H., Albrecht, T., Altman, J., Amani, C., Aneseyee, A. B., Avitabile, V., Banin, L., Batumike, R., Bauters, M., Beeckman, H., Begne, S. K., ... Zibera, E. (2021). High aboveground carbon stock of African tropical montane forests. *Nature*, 596, 536–542. <https://doi.org/10.1038/s41586-021-03728-4>

Cunningham, A. B. (1996). *People, park and plant use. Recommendations for multiple-use zones and development alternatives around Bwindi impenetrable National Park, Uganda*. People and Plants working paper 4. UNESCO.

Cutler, D. R., Edwards, T. C., Jr., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. (2007). Random forests for classification in ecology. *Ecology*, 88, 2783–2792. <https://doi.org/10.1890/07-0539.1>

Davenport, T., Howard, P., & Matthews, R. (1996). *Bwindi impenetrable national park biodiversity report*. Forest Department.

Desalegn, W., & Beierkuhnlein, C. (2010). Plant species and growth forms richness along altitudinal gradients in the southwest Ethiopian highlands. *Journal of Vegetation Science*, 21, 617–626. <https://doi.org/10.1111/j.1654-1103.2010.01177.x>

Eggeling, W. J. (1947). Observations on the ecology of the Budongo rain forest, Uganda. *Journal of Ecology*, 34, 20–87. <https://doi.org/10.2307/2256760>

Eilu, G., Hafashimana, D. L., & Kasenene, J. M. (2004). Density and species diversity of trees in four tropical forests of the Albertine rift, western Uganda. *Diversity and Distributions*, 10, 303–312. <https://doi.org/10.1111/j.1366-9516.2004.00089.x>

Elgart, A. A. (2010). Are the gorillas in Bwindi impenetrable National Park “true” mountain gorillas? *American Journal of Physical Anthropology*, 141, 561–570. <https://doi.org/10.1002/ajpa.21172>

Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J., Peterson, A. T., ... Zimmerman, N. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>

Elith, J., Leathwick, J. R., & Hastie, T. (2009). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77, 802–881. <https://doi.org/10.1111/j.1365-2656.2008.01390.x>

Evans, J. S., & Cushman, S. A. (2009). Gradient modeling of conifer species using random forests. *Landscape Ecology*, 24, 673–683. <https://doi.org/10.1007/s10980-009-9341-0>

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302–4315. <https://doi.org/10.1002/joc.5086>

Fjeldsa, J. (2007). How broad-scale studies of patterns and processes can serve to guide conservation planning in Africa. *Conservation Biology*, 21, 659–667. <https://doi.org/10.1111/j.1523-1739.2007.00706.x>

Franklin, J. (1995). Predictive vegetation mapping: Geographic modeling of biospatial patterns in relation to environmental gradients.

Progress in Physical Geography: Earth and Environment, 19, 474–499. <https://doi.org/10.1177/03091339501900403>

Franklin, J. (2010). *Mapping species distributions: Spatial inference and prediction*. Cambridge.

Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics*, 19, 1–141. <https://doi.org/10.1214/aos/1176347963>

Ganas, J., Robbins, M. M., Nkurunungi, J. B., Kaplin, B. A., & McNeilage, A. (2004). Dietary variability of mountain gorillas in Bwindi impenetrable National Park, Uganda. *International Journal of Primatology*, 25(5), 1043–1072.

Ghazoul, J., & Sheil, D. (2010). *Tropical rain forest ecology, diversity and conservation*. Oxford University Press.

Godsoe, W., & Harmon, L. J. (2012). How do species interactions affect species distribution models? *Ecography*, 35, 811–820. <https://doi.org/10.1111/j.1600-0587.2011.07103.x>

Government of Uganda. (1967). *Atlas of Uganda* (2nd ed.). Lands and Survey Department.

Guo, Q., Kelly, M., & Graham, C. (2005). Support vector machines for predicting distribution of sudden oak death in California. *Ecological Modeling*, 182, 75–90. <https://doi.org/10.1016/j.ecolmodel.2004.07.012>

Guschanski, K., Caillaud, D., Robbins, M., & Vigilant, L. (2008). Females shape the genetic structure of a gorilla population. *Current Biology*, 18(22), 809–1814. <https://doi.org/10.1016/j.cub.2008.10.031>

Hall, J. (1991). Multiple-nearest-tree sampling in an ecological survey of Afromontane catchments forests. *Forest Ecology and Management*, 42, 245–266. [https://doi.org/10.1016/0378-1127\(91\)90028-T](https://doi.org/10.1016/0378-1127(91)90028-T)

Hamilton, A. C. (1969). The vegetation of south-west Kigezi. *Uganda Journal*, 33, 175–199.

Hamilton, A. C. (1974). Distribution patterns of forest trees in Uganda and their historical significance. *Vegetatio*, 29, 29–32. <https://www.jstor.org/stable/20036795>

Hamilton, A. C. (1975). A quantitative analysis of altitudinal zonation in Uganda forests. *Vegetatio*, 30, 99–106. <https://doi.org/10.1007/BF02389611>

Hamilton, A. C. (1994). Tropical rain forest ecosystems. In H. Leith & M. J. Werger (Eds.), *Ecosystems of the world*, 14B (pp. 155–182). Elsevier.

Harcourt, A. (1981). Can Uganda's gorillas survive? A survey of the Bwindi Forest reserve. *Biological Conservation*, 19, 269–282. [https://doi.org/10.1016/0006-3207\(81\)90003-3](https://doi.org/10.1016/0006-3207(81)90003-3)

Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical Science*, 1(3), 297–318.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

He, K. S., Bethany, A., Cord, A. F., Rocchini, D., Tuanmu, M., Schmidlein, S., Turner, W., Wegmann, M., & Pettorelli, N. (2015). Will remote sensing shape the next generation of species distribution models? *Remote Sensing in Ecology and Conservation*, 1, 4–18. <https://doi.org/10.1002/rse2.7>

Hedberg, O. (1951). Vegetation belts of the east African mountains. *Svensk Botanisk Tidskrift*, 45, 140–202.

Hickey, J. R., Uzabaho, E., Akantorana, M., Arinaitwe, J., Bakebwa, I., Bitariho, R., Eckardt, W., Gilardi, K. V., Katutu, J., Kayijamahe, C., Kierepka, E. M., Mugabukomeye, B., Musema, A., Mutabaazi, H., Robbins, M. M., Sacks, B. N., & Zikusoka, G. K. (2019). *Bwindi Sarambwe 2018. Surveys: Monitoring mountain gorillas, other select mammals, and human activities*. GVTC.

Hijmans, R., & Elith, J. (2016). Species distribution modelling with R. <https://rspatial.org/raster/sdm>

Hoersch, B., Braun, G., & Schmidt, U. (2002). Relation between landform and vegetation in alpine regions of Wallis, Switzerland: A multi-scale remote sensing and GIS approach. *Computers, Environment and Urban Systems*, 26, 113–139. [https://doi.org/10.1016/S0198-9715\(01\)00039-4](https://doi.org/10.1016/S0198-9715(01)00039-4)

Howard, P. C. (1991). *Nature conservation in Uganda's tropical forest reserves*. IUCN.

Jucker, T., Bongalov, B., Burslem, D. F. R. P., Nilus, R., Dalponte, M., Lewis, S. M., Phillips, O. L., Qie, L., & Coomes, D. A. (2018). Topography shapes the structure, composition and function of tropical forest landscapes. *Ecology Letters*, 21, 989–1000. <https://doi.org/10.1111/ele.12964>

Kalema, J., & Hamilton, A. (2020). *Field guide to the forest trees of Uganda*. CABI Oxfordshire.

Kent, M. (2012). *Vegetation description and data analysis: A practical approach* (2nd ed.). Wiley-Blackwell.

Klein, C., & Vilcko, F. (2006). A new empirical approach for estimation in k-tree sampling. *Forest Ecology and Management*, 237, 522–533. <https://doi.org/10.1016/j.foreco.2006.09.072>

Lance, G. N., & Williams, W. T. (1966). A generalised sorting strategy for computer classification. *Nature*, 211, 218. <https://doi.org/10.1038/212218a0>

Latt, M. M., & Parker, B. B. (2022). Tree species composition and forest community types along environmental gradients in Htamanthi wildlife Santuary, Myanmar: Implications for action prioritization in conservation. *Plants*, 11, 2180. <https://doi.org/10.3390/plants11162180>

Legendre, P., & Legendre, L. (1998). *Numerical Ecology* (2nd ed.). Elsevier Sciences BV. Amsterdam.

Leggat, G. J., & Osmaston, H. A. (1961). *Working plan for the impenetrable central Forest reserve, Kigezi District, western Uganda*, 1961–71. Uganda Forest Department.

Liaw, A., & Wiener, M. (2002). Classification and regression by random forests. *R News*, 2(3), 18–22.

Lieberman, D., Lieberman, M., Peralta, R., & Hartshorn, G. S. (1996). Tropical forest structure and composition on a large-scale altitudinal gradient in Costa Rica. *Journal of Ecology*, 84, 137–152. <https://doi.org/10.2307/2261350>

Lin, H.-Y., Li, C.-F., Chen, T.-Y., Hsieh, C.-F., Wang, G., Wang, T., & Hu, J.-M. (2020). Climate-based approach for modeling the distribution of montane forest vegetation in Taiwan. *Applied Vegetation Science*, 23, 239–253. <https://doi.org/10.1111/avsc.12485>

Linder, H. P. (2014). The evolution of African plant diversity. *Frontiers in Ecology and Evolution*, 2, 1–14. <https://doi.org/10.3389/fevo.2014.00038>

Lovett, G. M., & Lindberg, S. E. (1993). Atmospheric deposition and canopy interactions of nitrogen in forests. *Canadian Journal of Forest Research*, 23, 1603–1616. <https://doi.org/10.1139/x93-200>

Lovett, J. C., Clarke, G. P., Moore, R., & Morrey, G. H. (2001). Elevational distribution of restricted range forest tree taxa in eastern Tanzania. *Biodiversity and Conservation*, 10, 541–550. <https://doi.org/10.1023/A:1016610526242>

Lovett, J. C., Marshall, A. R., & Carr, J. (2006). Changes in tropical forest vegetation along an altitudinal gradient in the Udzungwa Mountains National Park, Tanzania. *African Journal of Ecology*, 44, 478–490. <https://doi.org/10.1111/j.1365-2028.2006.00660.x>

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). Chapman &.

McCune, B., & Grace, J. B. (2002). *Analysis of ecological communities*. MJM Software Design.

McNeilage, A., Plumptre, A., Brock-Doyle, A., & Vedder, A. (1998). Bwindi impenetrable National Park, Uganda gorilla and large mammal census. In *WCS working paper No. 14*. Wildlife Conservation Society. Wildlife Conservation Society.

McNeilage, A., Plumptre, A. J., Brock-Doyle, A., & Vedder, A. (2001). Bwindi impenetrable National Park, Uganda: Gorilla census, 1997. *Oryx*, 35, 39–47. <https://doi.org/10.1046/j.1365-3008.2001.00154.x>

Moore, D. E., Lees, B. G., & Davey, S. M. (1991). A new method for predicting vegetation distribution using decision tree analysis in a geographic information system. *Journal of Environmental Management*, 15, 59–71. <https://doi.org/10.1007/BF02393838>

Mueller-Dombois, D., & Ellenberg, H. (2002). *Aims and methods of vegetation ecology*. The Blackburn Press.

Muñoz Mazón, M., Klanderud, K., Finegan, B., Veintimilla, D., Bermeo, D., Murrieta, E., Delgado, D., & Sheil, D. (2019). Disturbance and the elevation ranges of woody plant species in the mountains of Costa Rica. *Ecology and Evolution*, 9, 14330–14340. <https://doi.org/10.1002/ece3.5870>

NASA & METI. (2011). *ASTER global digital elevation model (GDEM)*. <http://gdem.ersdac.jspacesystems.or.jp/>

Olupot, W., Barigiyira, R., & Chapman, C. A. (2009). The status of anthropogenic threat at the people-park interface of Bwindi impenetrable National Park, Uganda. *Environmental Conservation*, 36, 41–50. <https://doi.org/10.1017/S0376892909005347>

Pearce, A., & Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133, 225–245. [https://doi.org/10.1016/S0304-3800\(00\)00322-7](https://doi.org/10.1016/S0304-3800(00)00322-7)

Perrigo, A., Hoorn, C., & Antonelli, A. (2020). Why mountains matter for biodiversity. *Journal of Biogeography*, 47(2), 315–325.

Pfeffer, K., Pebesma, E. J., & Burrough, P. A. (2003). Mapping alpine vegetation observations and topographic attributes. *Landscape Ecology*, 18, 759–776. <https://doi.org/10.1023/B:LAND.0000014471.78787.d0>

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259.

Plumptre, A. J. (1996). Changes following sixty years of selective timber harvesting in the Budongo Forest reserve, Uganda. *Forest Ecology and Management*, 89, 101–113. [https://doi.org/10.1016/S0378-1279\(96\)03854-6](https://doi.org/10.1016/S0378-1279(96)03854-6)

Plumptre, A. J., Davenport, T. R. B., Behangana, M., Kityo, R., Eilu, G., Ssegawa, P., Ewango, C., Meirte, D., Kahindo, C., Herremans, M., Peterhans, J. K., Pilgrim, J. D., Wilson, M., Languy, M., & Moyer, D. (2007). The biodiversity of the Albertine rift. *Biological Conservation*, 134, 178–194. <https://doi.org/10.1016/j.biocon.2006.08.021>

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems*, 9, 181–199. <https://doi.org/10.1007/s10021-005-0054-1>

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing URL <https://www.R-project.org/>

Richter, M. (2008). Tropical mountain forests: Distribution and general features. *Biodiversity and Ecology Series*, 2, 7–24.

Sainge, M. N., Lyonga, N. M., Mbatchou, G. P. T., Kenfack, D., Nchu, F., & Peterson, A. T. (2019). Vegetation, floristic composition and structure of a tropical montane forest in Cameroon. *Bothalia*, 49(1), 2270. <https://doi.org/10.4102/abc.v49i1.2270>

Schmitt, C. B., Denich, M., Demissew, S., Friis, I., & Boechmer, H. J. (2010). Floristic diversity in fragmented Afromontane rainforests: Altitudinal variation and conservation importance. *Applied Vegetation Science*, 13, 291–304. <https://doi.org/10.1111/j.1654-109X.2009.01067.x>

Seddon, N., Chausson, A., Berry, P., Girardin, C. A., Smith, A., & Turner, B. (2020). Understanding the value and limits of nature-based solutions to climate change and other global challenges. *Philosophical Transactions of the Royal Society B*, 375, 20190120. <https://doi.org/10.1098/rstb.2019.0120>

Sheil, D. (2012). The impenetrable challenge of an overwhelming understorey. *Bulletin of British Ecological Society*, 43, 45–47.

Sheil, D. (2016). Disturbance and distributions: Avoiding exclusion in a warming world. *Ecology and Society*, 21(1), 10. <https://doi.org/10.5751/ES-07920>

Sheil, D., & Burslem, D. F. (2003). Disturbing hypotheses in tropical forests. *Trends in Ecology & Evolution*, 18(1), 18–26. [https://doi.org/10.1016/S0169-5347\(02\)00005-8](https://doi.org/10.1016/S0169-5347(02)00005-8)

Sheil, D., Ducey, M. J., Sidiyasa, K., & Samsoedin, I. (2003). A new type of sample unit for the efficient assessment of diverse tree communities in complex forest landscapes. *Journal of Tropical Forest Science*, 15, 117–135. <https://www.jstor.org/stable/23616330>

Singh, T. P., Singh, S., Roy, P. S., & Rao, B. S. P. (2001). Vegetation mapping and characterization in west Siang District of Arunachal Pradesh, India – A satellite remote sensing-based approach. *Current Science*, 83(10), 1221–1230.

Slik, J. F., Raes, N., Aiba, S. I., Brearley, F. Q., Cannon, C. H., Meijaard, E., Nagamasu, H., Nilus, R., Paoli, G., Poulsen, A. D., & Sheil, D. (2009). Environmental correlates for tropical tree diversity and distribution patterns in Borneo. *Diversity and Distributions*, 15, 523–532. <https://doi.org/10.1111/j.1472-4642.2009.00557.x>

Ssali, F., Moe, S. R., & Sheil, D. (2017). A first look at the impediments to forest recovery in bracken-dominated clearings in the African highlands. *Forest Ecology and Management*, 402, 166–176. <https://doi.org/10.1016/j.foreco.2017.07.050>

ter Braak, C. J. F. (1995). Ordination. In R. H. G. Jongman, C. J. F. ter Braak, & O. F. R. van Tongeren (Eds.), *Data analysis in community and landscape ecology* (pp. 91–173). Cambridge University Press.

Townsend, P., & Walsh, S. (2001). Remote sensing of forested wetlands: Application of multitemporal and multispectral satellite imagery to determine plant community composition and structure in southeastern USA. *Plant Ecology*, 157, 129–149. <https://doi.org/10.1023/A:1013999513172>

Trew, B. T., & Maclean, I. M. (2021). Vulnerability of global biodiversity hotspots to climate change. *Global Ecology and Biogeography*, 30(4), 768–783. <https://doi.org/10.1111/geb.13272>

Twongyirwe, R., Majaliwa, J. G. M., Ebanyat, P., Tenywa, M. M., Sheil, D., Heist, M. V., Oluka, M., & Kumar, L. (2011). Dynamics of forest cover conversion in and around Bwindi impenetrable forest, southwestern Uganda. *Journal of Applied Sciences and Environmental Management*, 15, 189–195. <https://doi.org/10.4314/jasem.v15i1.68439>

Volkov, I., Banavar, J. R., Hubbell, S. P., & Maritan, A. (2003). Neutral theory and relative species abundance in ecology. *Nature*, 424, 1035–1037. <https://doi.org/10.1038/nature01883>

Werner, F. A., & Homeier, J. (2015). Is tropical montane forest heterogeneity promoted by a resource-driven feedback cycle? Evidence from nutrient relations, herbivory and litter decomposition along a topographical gradient. *Functional Ecology*, 29, 430–440. <https://doi.org/10.1111/1365-2435.12351>

White, F. (1983). *The vegetation of Africa: A descriptive memoir to accompany the Unesco/AETFAT/UNSO vegetation map of Africa*. Unesco.

Wright, E., Eckardt, W., Refisch, J., Bitariho, R., Grueter, C. C., Ganas-Swaray, J., Stoinski, T. S., & Robbins, M. M. (2022). Higher maximum temperature increases the frequency of water drinking in mountain gorillas (*gorilla beringei beringei*). *Frontiers of Conservation Science*, 3, 738820. <https://doi.org/10.3389/fcosc2022.738820>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Babaasa, D., Finn, J. T., Schweik, C. M., Fuller, T. K., & Sheil, D. (2024). Predictive mapping of tree species assemblages in an African montane rainforest. *Biotropica*, 56, e13302. <https://doi.org/10.1111/btp.13302>