



# High-resolution topographic variables accurately predict the distribution of rare plant species for conservation area selection in a narrow-endemism hotspot in New Caledonia

Guillaume Lannuzel<sup>1</sup> · Joan Balmot<sup>1,2</sup> · Nicolas Dubos<sup>3</sup> · Martin Thibault<sup>4</sup> · Bruno Fogliani<sup>1</sup>

Received: 12 January 2020 / Revised: 10 January 2021 / Accepted: 20 January 2021 /  
Published online: 22 February 2021  
© The Author(s) 2021

## Abstract

Species distribution models (SDMs) represent a widely acknowledged tool to identify priority areas on the basis of occurrence data and environmental factors. However, high levels of topographical and climatic micro-variation are a hindrance to reliably modelling the distribution of narrow-endemic species when based on classic occurrence and climate datasets. Here, we used high-resolution environmental variables and occurrence data obtained from dedicated field studies to produce accurate SDMs at a local scale. We modelled the potential current distribution of 23 of the 25 rarest species from Mount Kaala, a hotspot of narrow-endemism in New Caledonia, using occurrence data from two recent sampling campaigns, and eight high-resolution (10 m and 30 m) environmental predictors in a Species Distribution Modelling framework. After a first sampling operation, we surveyed six additional areas containing, overall, 13 of the 20 species modelled at this stage, to validate our projections where the highest species richness levels were predicted. The ability of our method to define conservation areas was largely validated with an average 84% of predicted species found in the validation areas, and additional data collected enabling us to model three more species. We therefore identified the areas of highest conservation value for the whole of Mount Kaala. Our results support the ability of SDMs based on presence-only data such as MaxEnt to predict areas of high conservation value using fine-resolution environmental layers and field-collected occurrence data in the context of small and heterogeneous systems such as tropical islands.

**Keywords** Digital elevation models · Field validation · Islands · MaxEnt · Narrow-endemism

---

Communicated by Daniel Sanchez Mata.

---

✉ Guillaume Lannuzel  
guillaume.lannuzel@gmail.com

Extended author information available on the last page of the article

## Introduction

More than half of plant extinction events over the last 500 years occurred in oceanic islands, which are home to more than 30% of critically endangered plants (Tershy et al. 2015). In this regard, islands are priority territories for the development of plant conservation programs (McGeoch et al. 2016). In an era of global biodiversity crisis (Ceballos et al. 2015), the world's biodiversity hotspots have been defined according both to the level of richness and endemism of their flora and the threats they are facing, nearly half of them being island territories (Myers et al. 2000). As stressed by Cañadas et al. (2014), some hotspots are composed with micro- or nano-hotspots due to high levels of narrow-endemism and should be the focus of conservation actions. The presence of such narrow-hotspots was highlighted in New Caledonia by Wulff et al. (2013) and was recently identified as in need of urgent conservation planning (Gâteblé et al. 2018). In these areas of high conservation value, management planning often suffers from a lack of high-resolution background data on precipitation and temperature regimes, resulting in a poor understanding of biogeographical patterns and species ecology (Heywood 2011). Knowledge about the ecology of rare species is even more limited in the case of narrow endemic species (NES) occurring in inaccessible locations (Kier et al. 2009). Pressures from global changes are particularly concerning in islands (Weigelt et al. 2016) and NES are particularly threatened (Wulff et al. 2013; Cañadas et al. 2014). There is a need to characterise the ecological niche of rare insular plants and their potential response to environmental changes.

Species distribution models (SDMs) are now widely used to produce transferable projections of the potential distribution of known species, which helps identify important areas for conservation among other goals (Guisan et al. 2013; Leroy et al. 2014). In this regard, Heywood (2011) proposed the use of this tool to guide the protection of island flora in data-poor territories. However, modelling the distribution of rare plant species is challenging in several aspects. First, the resolution of the most commonly used bioclimatic data (Booth et al. 2014) such as the WorldClim Database (Fick and Hijmans 2017) or Chelsea climate data (Karger et al. 2017) are too coarse to properly represent highly heterogeneous habitats. The most up to date WorldClim dataset is available at a 1 km resolution (WorldClim2—Fick and Hijmans 2017), while it is known that rare plant distributions can be locally driven by topo-climatic variations at a 10–100 m scale (Franklin et al. 2013). The resolution of these global datasets is therefore insufficient to properly model species distributions and guide decision-making tools at the local scale (Kier et al. 2009). Secondly, occurrence data of rare or narrow-ranging plant species are difficult to obtain from existing databases. Generalist datasets and herbarium collections are the most common sources of such data (Loiselle et al. 2003; Gallardo and Aldridge 2013). However, they often lack occurrence data on rare species, and herbarium collections may be geographically biased (Ter Steege et al. 2011). Ad hoc collection of input presence data may thus form a better basis for high resolution modelling of rare plant species in narrow ranges. Finally, the selection of absence data is a key point in SDM projections, hence calling for a precise description of the survey effort underlying each occurrence dataset (Phillips 2008). Poorly defined absence data may indeed result in overpredictions in surveyed areas and produce “precise answers to the wrong questions” (Yackulic et al. 2013).

While SDM studies have advanced toward more efficient algorithms at large scales, there have been relatively few applications to the projection of NES distributions. Recent studies are answering some of these challenges by integrating microvariation predictors into their projections (Tomlinson et al. 2019). However, high-resolution background data

are not always available. When such background data are lacking, a combination of dedicated field collection of presence data and post hoc field validation of projections should produce reliable guidelines for in situ conservation purposes.

The New Caledonia archipelago is one of the world's biodiversity hotspots of the highest conservation value (Myers et al. 2000; Wulff et al. 2013). Almost 1% of the world's flora is endemic to this southwest Pacific archipelago (Munzinger et al. 2020). In this region, about 20% of the indigenous flora is considered as narrow-endemic, i.e., restricted to one, two or three locations (Wulff et al. 2013). This high endemism is the consequence of a unique geological history, resulting in a high topographic and climatic heterogeneity. Ultramafic soils are very frequent in the south of Grande Terre (the main island) and sparsely distributed northwards along the west coast in isolated and steep mountains (Bonvallot et al. 2012). In terms of plant ecology, these mountains are considered as ultramafic islands within a non-ultramafic matrix (Isnard et al. 2016). The original composition of ultramafic soil, particularly the abundant presence of nickel ore, offers substantial economic value to the mining industry that is widely developed in New Caledonia.

Mining represents a major threat to biodiversity through vegetation removal, topographic alteration, and persisting pollution (Lefcort et al. 2010; Palmer et al. 2010; Duran et al. 2013). In New Caledonia, the number of mining tenements issued rose by about a hundred between 1997 and 2017, resulting in more than 45% of all ultramafic areas currently being under mining concessions (Ibanez et al. 2019). Furthermore, no protected areas have been implemented on ultramafic substrates in the northern half of the main island (Ibanez et al. 2019), which emphasises the need for urgent conservation programs to identify and protect its endemic-rich flora. In 2019, the conservation status of one third of the New Caledonian flora was assessed by the local Red List Authority (RLA-NC). Out of a thousand species, 43% were classified as threatened, mainly by open-cast mining, bush-fires or invasive alien species, and 11% were considered "data-deficient" (RLA-NC 2019). This lack of information needs to be addressed urgently for the design of efficient management strategies (Collen et al. 2008; IUCN 2012; IUCN/SSC 2013).

In this study, we coupled empirical field data collection and species distribution modeling to produce high-resolution "conservation priority" maps and promote the conservation of narrow-endemic taxa in one of New-Caledonia's hotspots of narrow-endemism threatened by mining activity. We predicted the distribution of 23 species among the rarest taxa of Mt. Kaala, using presence data from two recent sampling campaigns, and eight available habitat variables (10 m and 30 m resolution). To evaluate the ability of our projections to define areas for conservation, we carried out six additional surveys in areas where the highest species richness was predicted. Our results are discussed in light of the local and regional strategies for plant conservation. The study provides an integrative framework that should be considered by environment managers in hotspots of narrow-endemism.

## Methods

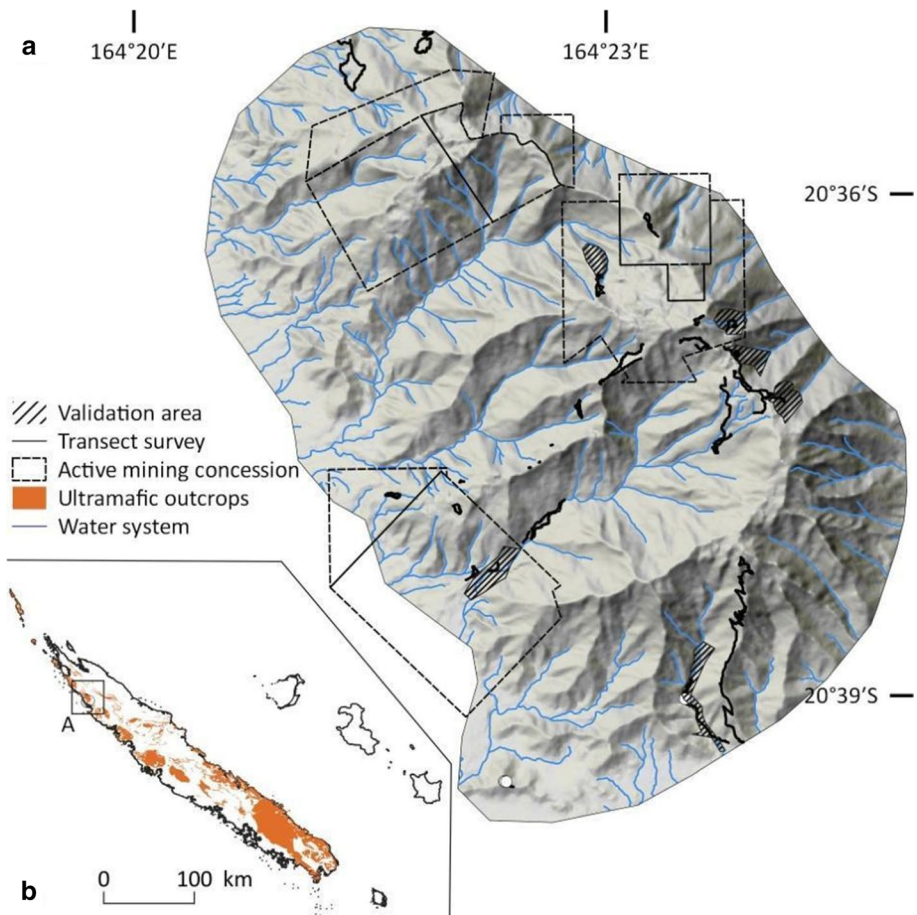
### Study site

The New Caledonia archipelago is located in the Coral Sea, 1500 km east of Australia and 2000 km north-northwest of New Zealand. It encompasses the main island of Grande Terre, the Loyalty Islands, and a number of smaller islands and islets. Grande Terre is 400 km long and 50–70 km wide.

Our study site is defined by the geological entity known as Mount Kaala, located in northwest Grande Terre (164°23'E, 20°36'S), with an altitude of 1070 m (Fig. 1). The site is dominated by ultramafic substrates with a serpentinite outcrop at the bottom, and is isolated in a large volcano-sedimentary coastal plain. The vegetation of Mount Kaala is a mosaic of varyingly degraded shrublands called “maquis miniers” (ca. 61% of the study area based on Barrière et al. 2007), dense humid forest (ca. 19%) and degraded secondary herbaceous patches (ca. 16%) on the slopes. The summit plateau, exploited for ore extraction, is composed of bare soils (ca. 4% of the study area) and secondary herbaceous patches.

### Studied species

The species list we used in this study was determined in a previous study (Table 1, Lanuzel, unpublished.), and is based on a comprehensive screening of all herbarium data



**Fig. 1** Detailed map of Mount Kaala with survey transects, validation areas and active mining areas (a) and its location in New Caledonia (b)

**Table 1** Studied species, with local IUCN and narrow-endemism statuses

Family	Taxa	IUCN	NES*	Occ	Family	Taxa	IUCN	NES*	Occ
Apocynaceae	<i>Marsdenia kaalaensis</i>	CR	x	8	Phyllanthaceae	<i>Phyllanthus pterocladus</i>	EN		67
Araceae	<i>Burretio Kentia hapala</i>	EN		13	Primulaceae	<i>Myrsine novocaledonica</i> subsp. <i>kaalaensis</i>	DD	x	24
Araliaceae	<i>Polycias</i> sp. nov.	CR	x	5	Rubiaceae	<i>Gynochthodes truncata</i>	EN		68
Asparagaceae	<i>Arthropodium</i> sp. nov.	CR	x	19	Rubiaceae	<i>Psychotria</i> sp. nov.	NE	x	6
Asteraceae	<i>Pytinicarpa</i> sp. nov.	CR	x	39	Rubiaceae	<i>Thiollierea kaalaensis</i>	EN	x	2
Capparaceae	<i>Capparis parvifolia</i>	CR	x	1	Rutaceae	<i>Comptonella fruticosa</i>	EN		20
Euphorbiaceae	<i>Cleidion velutinum</i>	EN		61	Rutaceae	<i>Zieria chevalieri</i>	EN	x	259
Myrtaceae	<i>Tristaniopsis minutiflora</i>	CR	x	90	Salicaceae	<i>Casearia kaalaensis</i>	EN		65
Olacaceae	<i>Olax hypoleuca</i> var. <i>microphylla</i>	EN		7	Salicaceae	<i>Xylosma kaalaense</i>	EN		76
Phyllanthaceae	<i>Phyllanthus deciduiramus</i>	CR	x	25	Sapotaceae	<i>Planchonella minutiflora</i>	CR		5
Phyllanthaceae	<i>Phyllanthus favieri</i> var. <i>kaalaensis</i>	CR	x	9	Sapotaceae	<i>Planchonella pinifolia</i>	EN		4
Phyllanthaceae	<i>Phyllanthus koumacensis</i> var. <i>brevitepalus</i>	EN		6	Thymelaeaceae	<i>Deltaria brachyblastophora</i>	EN		59
Phyllanthaceae	<i>Phyllanthus pitifer</i> var. <i>pitifer</i>	EN		55					

CR critically endangered, EN endangered, DD data deficient, NE non evaluated

\*NES: Narrow-endemic species i.e. the species is only known from Mount Kaala, corresponding to the 1 locality narrow-endemic species conception of Wulff et al. (2013). Occ = final occurrence number for the taxon

available worldwide (mainly Paris—P, Nouméa—NOU, Missouri Botanical Garden—MO) for the Mount Kaala area. We filtered the resulting species list using the IUCN criteria and methodology (IUCN 2012). This process highlighted 23 taxa assessed as either Critically Endangered or Endangered in 2017 by the local Red List Authority (RLA-NC). We also kept two non-evaluated taxa because we assumed their extinction risk was very high. We identified four specimens collected in the early stages of the study as new species and the publication of their formal description is pending. We assume the species list we used is up to date for this area. However, given that new species are continuously described in New Caledonia (Gâteblé et al. 2018), our study provides recommendations for conservation on the basis of current knowledge.

## Environmental variables

We used six environmental predictors in this study, all based on topography and remote sensing. Because of the high resolution needed, the use of global bioclimatic data was not appropriate. Recently, Tomlinson et al. (2019) showed the relevance of using edaphic proxies versus downscaled climatic variables in the case of narrow-range endemic species. As a consequence of the lack of reliable edaphic predictors in New Caledonia, we used proxies representing the local topo-climatic variations (Pouteau et al. 2012), as well as a geological layer, as proxies for edaphic conditions. We computed four abiotic layers all derived from a Digital Elevation Model (DEM) freely available at a 10 m resolution (accuracy  $\pm 2$  m) (Gouvernement de la Nouvelle-Calédonie 2017), following the TopoTo Raster methods (Juffroy 2012). The first proxy is the DEM per se, because elevation is linked to temperature at the local scale (Bonvallot et al. 2012). Slope steepness (hereafter referred to as “slope”) affects the velocity of both surface and subsurface flow, and hence soil water content, erosion potential and soil formation (Pouteau et al. 2012; Wilson and Gallant 2000). Slope exposure (hereafter “exposure”) was used as a proxy of both insolation and windwardness, which can influence the drying rate of vegetation (Wilson and Gallant 2000). We used the topographic wetness index (TWI) as a proxy for soil moisture, and therefore water availability for plants (Western et al. 1999; Radula et al. 2018). In New Caledonia, TWI was recently proven to affect species assemblages at local scales (Blanchard et al. 2019). Slope and exposure were calculated using the QGIS 2.18.7 “terrain analysis” tool. TWI was computed with the SAGA GIS dedicated algorithm. It is based on slope and catchment area, the latter also computed using SAGA GIS. We also used a geological polygon layer (1/50,000 scale) with substrate type categories (Maurizot et al. 2005) rasterized at a 10 m resolution.

The only biotic predictor used was the Normalised Difference Vegetation Index (NDVI), a continuous index representing vegetation productivity (Pettorelli et al. 2005). Recent studies showed a correlation between tree species richness and single NDVI scene taken in the wet season in New Caledonia (Pouteau et al. 2018). We therefore interpreted it as a proxy for habitat suitability because many taxa are restricted to specific vegetation types (Isnard et al. 2016). NDVI was calculated from 30 m-resolution Landsat 8 images from 24th March 2017 (USGS 2017), corresponding to one of the only scenes without clouds in the study area during the wet season. We accounted for collinearity between environmental predictors using Pearson correlation coefficients (Le Lay et al. 2010; Gogol-Prokurat 2011) with a threshold set at 0.70 (see Appendix Tables 4). Correlation indices of the predictors ranged between  $-0.48$  (Soil-DEM) and  $0.39$  (Slope-DEM).

## Sampling effort

We acquired a first dataset during the five field studies in 2017 and 2018. We sought the most cost-effective method for data collection considering the steepness of the mountain and limited accessibility. Field surveys were designed to equally sample all altitudinal (DEM—100 m classes) and vegetation classes (NDVI—0.1 index value classes). We used strip transects, which is recommended in the case of highly clustered multi-species assemblages (Melville and Welsh 2001; Ogutu et al. 2006). Because strip transect accuracy depends on width, we assumed that detection was possible within a 5 m wide strip on either side of the observer. We chose a 5 m limit as the maximum distance at which we confidently detected a shrub species in dense shrublands, and small herbaceous species in herbaceous maquis. Each GPS mark was then located at the centre of a 10 m wide square-cell within which all species of interest were recorded.

A total of twenty-one strip transects were carried out on Mount Kaala (Fig. 1a). Transect locations were recorded using a GPS device (Garmin GPSMap64; spatial accuracy  $\pm 3$  m). All data were transmitted to the local Red List Authority and are available on the [www.endemia.nc](http://www.endemia.nc) website.

## Modelling

To broaden the appeal of our approach to the widest user community, we chose a modelling algorithm that is reproducible without needing high computational power and which can be undertaken by any stakeholder with moderate background knowledge of SDMs. We used MaxEnt (v.3.3.3; Phillips et al. 2006), a user-friendly and well documented software. This algorithm is widely acknowledged to be efficient for rare species and performs well with small datasets (Phillips et al. 2004, 2006; Pearson et al. 2007; Wisz et al. 2008). The convergence threshold was set at 0.00001 with a maximum number of iterations of 500. We chose the logistic output format with suitability values ranging from 0 (unsuitable) to 1 (optimal; Phillips and Dudík 2008). We ran five cross-validation replicates for each species. When the number of occurrences was lower than five, the software automatically reduced this number. For each replicate, it randomly splits the dataset into two parts: 80% for training, 20% for testing. The remaining parameters were kept as default settings.

## Sampling bias correction

The sampling methodology inevitably produced spatial and ecological auto-correlation of occurrences, which could represent a major bias (Kramer-Schadt et al. 2013). One expected consequence of this bias is to make predictions biased toward over-sampled areas (Phillips et al. 2009). As stated by Phillips et al. (2009), presence-only modelling is fairly robust as long as background points have the same sample selection bias as the presence points (Yackulic et al. 2013). However, this background selection method was hard to implement where the sampling effort is not precisely reported (Phillips et al. 2009; Kramer-Schadt et al. 2013; Fourcade et al. 2014). In our case, the dedicated sampling inherently provided sampled and unsampled areas. Moreover, this method enabled the use of all presence locations, as opposed to split or spatial filtering methods (Fourcade et al. 2014). It is therefore preferable for rare species with fewer occurrence

data. We added a 5 m buffer to our GPS tracks to generate 10-m wide polygons corresponding to our sampling design. The resulting polygon layer was rasterised to obtain a 10-m resolution raster layer with a value of 1 in sampled pixels, and 0 in unsampled pixels. This layer was implemented in Maxent as a “Bias file”. The software then selected background points within the surveyed area, while results were projected on the overall study area.

## Model performance

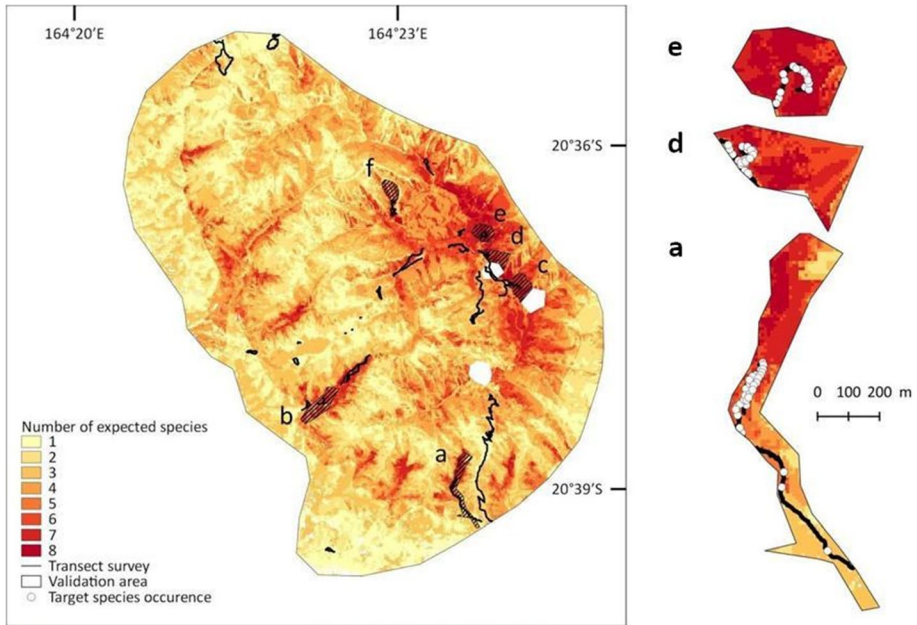
Selection of appropriate validation metrics such as AUC and TSS is a matter of debate in presence-pseudo-absence modelling, especially for rare species when prevalence is low (Leroy et al. 2018). To reduce misleading results, Leroy et al. (2018) advocated for the use of similarity tests in such cases. To evaluate model performance independently from prevalence, we used the continuous Boyce index ( $B_{\text{cont}(0,1)}$ ; Hirzel et al. 2006) which avoids arbitrary selection of class cutting thresholds by using a sliding window, and consequently might be more efficient for small sample tests. This indicator was computed for each replicate using the “ecospat” package (Di Cola et al. 2017) in R (R Core Team 2018). The Boyce index varies from  $-1$  to  $1$ . A value of  $1$  corresponds to a perfect match between observed and predicted occurrences, while values under  $0$  are no better than random. Every acceptable prediction ( $B_{\text{cont}(0,1)} > 0.5$ ) layer was transformed into a binary presence/absence map using the maximum sensitivity/specificity threshold (Jiménez-Valverde and Lobo 2007; Cao et al. 2013). Considering the debate on validation metrics for SDM, we also computed the true skill statistics (TSS), because many authors consider it as prevalence-independent, but did not refer to it for further analyses. Model performance is considered as poor if TSS value is lower than  $0.4$ , moderate between  $0.4$  and  $0.6$ , and good above  $0.6$  (Beauregard and de Blois 2014). We used the maximum sensitivity/specificity threshold to calculate TSS (Allouche et al. 2006; Williams et al. 2009). We also provide AUC (that is automatically calculated by MaxEnt) in Appendix Table 5. We present average predicted values for each taxon.

## Field validation

We produced a general map by stacking every specific binary map (Fig. 2, and Appendix Fig. 4). We obtained a raster layer giving the expected number of species present within each pixel. Six areas labelled from “a” to “f” (Fig. 2) were chosen for survey within the areas where pixel values fell within the 4th upper quartile (Rhoden et al. 2017). These areas were surveyed, during a sampling trip on 6th and 7th February 2018. We then computed a post-hoc test, i.e. the Positive Predictive Power index (PPP; Fois et al. 2015) to assess the efficiency of our SDMs. PPP represents the ratio between true positives and the sum of predicted positives. Here, we considered a true positive when a predicted species was actually observed in the area considered and a false positive when a species was predicted but not observed.

All data collected during this second field campaign were added to the total dataset to produce the final prediction maps.  $B_{\text{cont}(0,1)}$  were re-calculated accordingly.





**Fig. 2** Stack map of every acceptable binary occurrence prediction of 20 rare plant taxa on Mt Kaala (New Caledonia) obtained during the 1st field study, which we used to define validation areas (“a”–“f”). The location of transect surveys are also indicated. On the right, examples of validation areas with transect surveys and studied species’ occurrences

### Plant conservation priority scoring

A lot of literature is available to determine important areas for conservation (Blasi et al. 2011; Balaguru et al. 2006; Darbyshire et al. 2017). We adapted the different criteria to match available data and specific local characteristics, such as narrow-endemism. Darbyshire et al. (2017) identified three important criteria while considering the establishment of important plant areas (IPAs): threatened species, botanical richness, and threatened habitat. The threatened species criterion is relatively straightforward if IUCN assessments are locally available, which was the case here. Other criteria are harder to implement in a data-poor context. When no exhaustive inventories are available, botanical richness can be assessed by the density of important/valuable species (Darbyshire et al. 2017). We used the predicted presence of species in our list, weighted by the IUCN status, to inform both these criteria. The IUCN (2016) recommended considering the irreplaceability of a species or population. In New Caledonia, NES that are restricted to one locality are irreplaceable, because, if they disappear from Mount Kaala in our example, they go extinct. We thus considered every predicted presence of such species as of irreplaceable value in our index. Threatened habitats are not known clearly in New Caledonia. In the absence of such a reference source, scores of the “threatened habitat” criterion were computed using two components. The first was the hierarchy of the given vegetation type within the post-fire ecological succession based on McCoy et al. (1999), from bare soil (lower level) to dense humid forest (upper level). The second was an area index representing the relative scarcity of each vegetation type on the mountain. The vegetation map used at this step was

**Table 2** Criteria, modality and scores used for the identification of priority areas for the conservation of rare plant species of Mt Kaala (New Caledonia)

Layer	Criteria	Modality	Score
A	IUCN status	EN	1
		CR	2
B	Narrow-endemism	No	0
		Yes	1
C	Vegetation	Anthropic area	0
		Herbaceous maquis	$1 \times A_{\text{herb}}$
		Open shrubland	$2 \times A_{\text{op-sh}}$
		Dense shrubland	$3 \times A_{\text{de-sh}}$
		Para-forest	$4 \times A_{\text{pa-fo}}$
Forest	$5 \times A_{\text{fo}}$		

EN endangered, CR critically endangered, A vegetation relative rarity based on the area occupied by one vegetation type

produced by Barrière et al. (2007). This results in the following formula where  $i$  is the vegetation type:

$$C_i = 1 - (Area_i / Area_{total})$$

We then proposed the list of criteria shown in Table 2.

The final layer was calculated to determine an index of conservation priority  $I$ , computed with the following formula:

$$I = \frac{C_i}{C_{max}} + \sum_{j=1}^{jmax} \frac{A_j}{A_{max}} + \sum_{j=1}^{jmax} \frac{B_j}{B_{max}}$$

where A, B, and C are the priority conservation criteria mentioned in Table 2 and  $j$  is the taxon considered. Criteria A and B were calculated for each taxon, while criterion C was calculated once for the whole study area. The  $I$  index therefore varies between 0 and 3, where 0 corresponds to an area with the lowest conservation priority, and 3 represents the highest level of priority.

## Results

### Pre-validation modelling

At the end of the first data collection phase, the mean number of occurrences was  $33 \pm 11$  SE per species. Three species were detected only once, and a maximum of 259 occurrences was reached for *Zieria chevalieri*. All elevation and NDVI classes were included in the surveys. Relative to the whole massif area, transects represented about 0.5% of Mount Kaala. A total of 1610, pixels ( $10 \times 10$  m) were sampled, for a total of 800 “presence” pixels. For elevation classes,  $0.68 \pm 0.19\%$  of areas were surveyed with an over-representation of the highest classes (700 m and higher). For NDVI,  $0.36 \pm 0.07\%$  of areas were surveyed with a slight over-representation of the 0.1–0.2 classes. An under-representation on higher classes (0.4 and more) corresponded to dense humid forests.

Species with one or two occurrences (four species) were discarded for the modelling procedure. Distributions of twenty-one species were projected, among which 17 had a  $B_{\text{cont}(0,1)}$  ranging from 0.58 to 1 and were thus considered reliable. The four models with a  $B_{\text{cont}(0,1)} < 0.5$  were not considered for the validation step.

Binary maps highlighted different distributions. Species' projected niches were either restricted to low elevation (e.g. *Casearia kaalaensis*, *Cleidion velutinum*, *Deltaria brachyblastophora*), high elevation (e.g. *Comptonella fruticosa*, *Psychotria* sp. nov., *Pytinicarpa* sp. nov.), or more broadly distributed (e.g. *Gynochthodes truncata*, *Xylosma kaalaense*). Niche overlap values ranged from 0 to 8 species predicted per pixel.

## Field validation

We conducted surveys in six areas projected to host more than five species. We selected two areas at low altitude, one in the middle altitudinal range, and three at high altitude. Mean PPP was  $0.80 \pm 0.02$  SE, with no PPP under 0.7 (Table 3). Surveyed areas contained 15 of 17 predicted taxa, and no unpredicted species were observed.

Surveys represented 5% of each selected area on average, and 460 additional surveyed pixels (46,000 m<sup>2</sup>—28% of the initial dataset). Additional occurrence data were then added to the occurrence datasets. This allowed modelling of the distribution of one more species, and improved the validation metric of four others. Final results included 22 prediction maps, with  $B_{\text{cont}(0,1)}$  values ranging from 0.51 to 1. Each predictor's contribution varied among species. DEM was the predominant predictor with 51% contribution on average, but some species' projections were mostly driven by NDVI (e.g. *Gynochthodes truncata*), exposure (e.g. *Marsdenia kaalaensis*) or slope (e.g. *Myrsine novocaledoniae* subsp. *kaalaensis*). Soil contribution represented 7% on average, but emerged as a strong predictor for some species (e.g. *Olex hypoleuca*—28%, *Gynochthodes truncata*—26%). TWI was the poorest predictor with an average 2% contribution, and a maximum contribution of 6%. Detailed statistics and results are given in Appendix Table 5. Detailed response curves are also given in Appendix Fig. 5. For most species, the marginal response curves of each predictor showed no significant differences with the response obtained with models created using only the corresponding variable. In a few specific cases only (*C. kaalaensis*, *C. velutinum*, *P. deciduiramus*, *Psychotria* sp. nov., *T. minutiflora*), exposition and slope show different curve shapes. The final specific maps, once stacked, showed a spatial niche overlap (spatial congruence) of 71% with at least two species predicted per pixel. The number of predicted species per pixel ranged from 0 to 12 (mean = 2.77). The 4th upper quartile of the total number of species per pixel (i.e., more than 9 species predicted) represented 0.8% of the study area.

**Table 3** Predictive power index (PPP) calculated for each area used for field validation

Area	Predicted species	Observed predicted species	Observed unpredicted species	PPP
a	6	6	1	1
b	6	5	0	0.83
c	7	5	0	0.71
d	8	7	0	0.88
e	8	7	0	0.88
f	7	5	1	0.71

## Priority areas for rare plant conservation

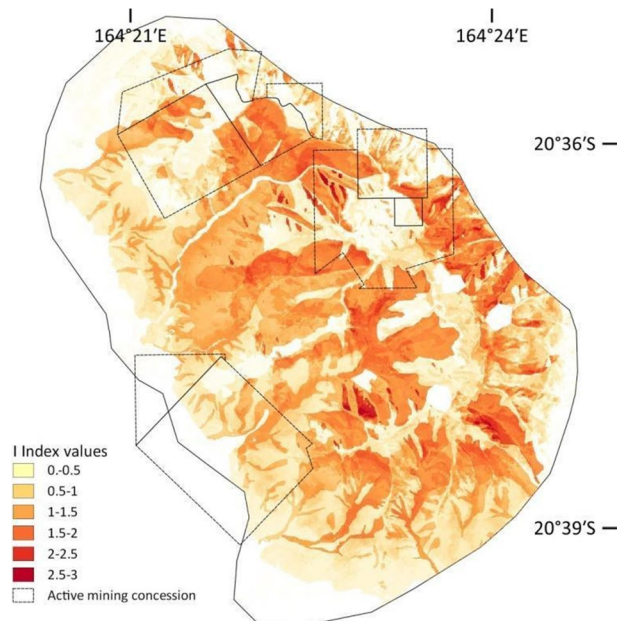
The I index varied within the study area and was rather low on average (median  $I=0.60$ ). Pixels with  $I > 1.5$  represented 6.3% of the total study area and were mostly located on the slopes, at intermediate altitudinal range inside and outside active mining concessions (Fig. 3). On the western slopes, it represented some areas at the top end of the three main valleys, where vegetation is less impacted. On the southern and eastern slopes, valleys are less deep, and areas of high I index were revealed in every little thalweg protected from mining erosion and bushfires.

## Discussion

In a context of highly heterogeneous habitats, our results showed that the occurrence of narrow-endemic species can be accurately predicted using topographic, geological and vegetation predictors. The Boyce validation metric seemed efficient in assessing the quality of predictions, even with only four presence data (*P. pinifolia*,  $B_{\text{cont}(0,1)}=0.97$ , verified during validation). Our field validation surveys confirmed 80% of projected occurrences, which supports the reliability of our models (Jiménez-Valverde 2012; Somodi et al. 2017; Leroy et al. 2018). Hence, we were able to use the projected maps to develop an evidence-based spatially explicit index for local environment managers.

Data sampling was designed to offset the lack of presence data in international databases. However, the production of presence-only data sets depends on three parameters: (i) sampling probability, (ii) occurrence probability, and (iii) detection probability. Together, they produce a bias often occurring in presence-only modelling studies (Yackulic et al. 2013). Our bias correction method maximised sampling-probability using transect tracks as bias files, without removing any occurrence data, which is a key point when modelling

**Fig. 3** Distribution of priority areas for conservation (computed with the I index) at Mt Kaala (New Caledonia). High I values represent areas with a high number of rare plant species, weighted according to IUCN status, narrow-endemism status and habitat importance. Active mining concessions are circled in black



the distribution of rare species (Fourcade et al. 2014). This method, here applied to transect sampling, also addressed the detection probability issue by using the 5 m-buffer added to the tracks. Finally, we measured a first likelihood of occurrence, and used it for further field surveys which enabled us to add new occurrence data for the final models. This was only possible because we started from dedicated surveys, allowing a high confidence in the definition of sampled/unsampled pixel maps. Controlling for these three parameters is obviously more difficult on SDMs at large scales when the sampling effort underlying occurrence data is often unknown (Phillips et al. 2009; Kramer-Schadt et al. 2013; Fourcade et al. 2014).

The choice of environmental predictors is a challenging issue in fine-scale SDMs (Nezer et al. 2016). Hijmans et al. (2005) showed the inaccuracy of global climate products such as Wordclim in the case of remote oceanic islands (including New Caledonia), because of the scarcity of local meteorological station networks (Turner et al. 2003; Pouteau et al. 2015). These products are thus unlikely to produce reliable predictions of very rare plants' suitable habitat, often influenced by 10–100 m scale variations (Franklin et al. 2013). The proxies we used here, mostly derived from the DEM layer, are the most relevant and locally available at high resolution for plant ecology (Pouteau et al. 2012; Wilson et al. 2013). However, three predictors were computed from the DEM which increases the risk of multicollinearity issues. We tested this issue using the Spearman correlation index, but also analysed the response curves of each predictor and species (Appendix Fig. 5) to confirm the absence of multicollinearity. For each species and predictors, two curves are computed. The first one represents the response of the model to the predictor while all other predictors are set at their mean value in the sample. The second represents the response to the predictor in a model built only with this predictor. The comparison between both curves shows whether a predictor is independent (both curves are similar) or correlated (both curves show different shapes—Elith et al. 2011; Phillips 2017). Almost all results show independence between predictors, except for some species, for which exposition and slope had slightly different shapes. In these cases, the influence of a predictor on a species' distribution must be read on the response curve of the predictor alone as it is more accurate if predictors are correlated, even slightly. The level of collinearity between variables was relatively low (maximum Spearman index =  $-0.48$  and  $0.39$ ) and was low across species. Therefore, we believe potential collinearity between our environmental predictors may occur at the margin of species distributions, but may not alter our conclusions.

DEM was the most influential predictor overall. However, the contribution of each predictor varied with species. *Gynochthodes truncata* for example, one of the most commonly recorded target species, was found all along the altitudinal gradient, and had a distribution influenced mostly by NDVI, while *Phyllanthus pterocladus* was equally influenced by DEM (34%) and NDVI (29%). The potential distribution of most NES restricted to Mount Kaala was influenced primarily by DEM and high altitude. This could be explained by a stronger island-like isolation of high-altitude areas on ultramafic mountains like Mount Kaala in New Caledonia. Further studies on other ultramafic mountains would give more

insight on narrow-endemism patterns, and contribute to understanding the drivers of such restricted distributions in the New Caledonian flora.

DEM per se and exposure are considered as climate proxies for temperature, and windwardness/insolation respectively. Slope, along with TWI and soil, influence soil formation and water content. In New Caledonia, L'Huillier et al. (2010) showed a causal relationship between elevation, slope, and soil characteristics on ultramafic substrates. Brown hypermagnesian soils such as colluvial formations (Isnard et al. 2016) cover lower altitudes, whereas ferralsols with high rock content influenced by slope and proximity to the ridges (Isnard et al. 2016) occur at middle and higher altitudes. The combination of soil substrate, elevation, and slope can be considered as reliable proxies of soil characteristic diversity. Our models were based on edaphic and climate proxies, which are known to be more relevant than strictly climatic ones for distinguishing subtle variations in plant distributions (Illan et al. 2010; Beaugard and de Blois 2014; Meineri and Hylander 2017).

Because of the uniqueness of New Caledonian flora, with high narrow-endemism rates and topo-climatic heterogeneity (Jaffré 1993; Isnard et al. 2016), we adapted already-existing criteria for the selection of conservation areas (Balaguru et al. 2006; Blasi et al. 2011; IUCN 2016; Darbyshire et al. 2017). The criteria we used considered both habitat suitability and vegetation conservation value. Narrow-endemism represents a major conservation stake on mined ultramafic mountains (Wulff et al. 2013). Considering the potential distribution of NES in our irreplaceability index highlighted new areas that would not have emerged from vegetation and species density-based indices. Our scoring therefore accounts for both the local flora specificity and the lack of comprehensive survey data. In the same way, since the country still has no red list of ecosystems, we evaluated the relative rarity of each vegetation type and its position in the vegetation succession in order to assess conservation values. We believe that this index constitutes a strong basis for further implementation of conservation programs in Mount Kaala. Our framework could even be replicated in similar ultramafic mountain contexts, provided that the reliability of irreplaceability indices could be controlled. One may be surprised by the relatively low I index value of low altitude areas, despite the proven high presence of IUCN threatened species. This is likely to be a result of the absence of NES at low altitude, and the relative rarity of forests on these areas mainly covered by shrublands on brown hypermagnesian soils. Consequently, the lower value attributed to early-stage vegetation compared to nearly climax forests tends to decrease the global conservation value of low altitude slopes. However, recent works emphasised the sensitivity of these rich vegetation types that are highly threatened by bushfires (Ibanez et al. 2019). Further developments of vegetation conservation value indices may give more weight to these habitats harbouring high levels of IUCN threatened taxa. Here, we consider the higher I index values in higher elevations relevant with regard to the threats due to mining activity.

Ibanez et al. (2019), following previous works (Jaffré et al. 2010; Wulff et al. 2013), also stressed the need to consider the high narrow-endemic species diversity on the island. The local network of protected areas, composed of few relatively large areas (Ibanez et al. 2019), does not sit well with the distribution of highly restricted taxa. A network of micro-scale conservation areas, such as the ones identified in this study, could enhance

the conservation status of flora across the territory overall. Our results will be transmitted to mining operators on Mount Kaala to promote the protection of areas of high conservation value, and to local environmental authorities as an operational tool for conservation planning.

The high PPP values in every checked area corroborated the statistical results and enabled us to find additional presence data points, consolidating the dataset for final analyses and unravelling additional occurrences of rare species. MaxEnt-guided surveys confirmed the relevance of our models for onsite conservation planning. However, the extent of these surveys was restricted due to time/funding limitations, and we focused on areas with the highest density of rare species. In so doing, we validated the areas of greater conservation interest on the basis of NES diversity with a cost-effective approach. This method cannot therefore be considered as a statistical validation of models *stricto sensu* (Rhoden et al. 2017; Hertzog et al. 2014), as we did not validate the predicted areas of NES absence on Mount Kaala. A further improvement could be a validation sampling protocol in which every quartile, including predicted absences, of projected species density is equally surveyed (Rhoden et al. 2017; Hertzog et al. 2014). On the other hand, a confirmation of “low conservation” value areas may legitimate further degradations and may not appear to be an efficient resource investment to funding institutions. We validated entire areas containing several species, but didn’t check the pixel per pixel validity of the model. Such validation would certainly lead to poorer PPP values. Another solution would have been the use of coarser resolution (e.g. 30 m). However, such resolution would have required additional sampling efforts to balance the lower detection probability if we aimed at being exhaustive in bigger pixels, significantly reducing the cost-effectiveness of our approach. We thus consider our predictions as a fair trade-off between limitations and conservation objectives, i.e. spatial prioritisation rather than species-specific distribution modelling. This prioritisation relied on fewer environmental predictors than in other areas with more fine-grained data available and must be used with caution. The partial field validation process appeared here as an acceptable counterpart to the lack of ecological data.

Island floras are characterised by high levels of endemism and sensitivity often associated with greater knowledge gaps (Caujapé-Castells et al. 2010), resulting in a lack of reliable data for local-scale biodiversity management (Cayuela et al. 2009; Collen et al. 2008). Plus, in rugged terrain areas (Whittaker and Fernandez-Palacios 2007), plant distributions are driven by very fine scale topo-climatic variations. Hence, biodiversity conservation needs to be assessed at a scale that matches these variations. Tomlinson et al. (2019) showed the relevance of coupled edaphic and topographic predictors in modelling short-range endemic species distributions through high resolution SDMs. In their case, the effect of topography was supposed to come from the correlation between soil characteristics and elevation, with certain soil types occurring only at high altitude. As shown by L’Huillier et al. (2010), New Caledonian ultramafic edaphic diversity is partly linked to elevation, slope, and position on the slope as well, justifying our choice of predictors. A detailed mapping of fine soil characteristics would be of great help to improve our understanding of the discrete drivers of plant species’ distribution in New Caledonia, but the cost–benefit

ratio of such data acquisition would need to be assessed. Considering the local context, and the urgent need for conservation planning, we assumed that the risk of using limited environmental proxies was lower than the benefits from the acquisition of such predictions of conservation value. Our maps may be used as guidance for the determination of new occurrences for the considered rare species, and our index as a spatial guide for the prioritisation of management actions.

The global conservation effort suffers from a lack of funding, as has been highlighted for tropical islands (Caujapé-Castells et al. 2010). One aim here was to propose a reproducible framework while trying to use the best practices in biodiversity assessment using SDMs (Araujo et al. 2019). We used only freely available data covering all of New Caledonia, and the computational power required was that of a common office computer. Hence, we believe that our study addresses the difficulties encountered by biodiversity managers in areas with high topo-climatic diversity and low data availability such as tropical islands (Koch et al. 2017; El-Gabbas et al. 2020), and would be easily transferable to decision-makers (Guisan et al. 2013). In addition, both dedicated surveys and field validation of our results gave weight to our model projections despite the constraints encountered. Both the distribution maps and the spatially explicit index we produced represent an opportunity for local stakeholders to develop their conservation plans. Our study promotes further collaboration efforts between researchers, managers and landowners to incorporate our results into existing management strategies and reproduce similar assessments in other rich and understudied areas. Finally, we encourage the use of user-friendly modelling tools in urgent conservation contexts to stimulate further initiatives and help rare plants step into the limelight.

## Appendix

See Tables 4 and 5, Figs. 4 and 5.

**Table 4** Correlation matrix based on Pearson index between all environmental variables considered

	Expo	DEM	NDVI	Slope	Soil	TWI
Expo	1.00	0.09	0.01	0.13	−0.09	−0.05
DEM		1.00	−0.07	0.39	−0.48	−0.02
NDVI			1.00	0.11	−0.02	0.01
Slope				1.00	−0.35	−0.02
Soil					1.00	0.00
TWI						1

*Expo* exposition, *Irada* solar irradiation, *DEM* digital elevation model, *NDVI* Normalised Difference Vegetation Index, *TWI* topographic wetness index



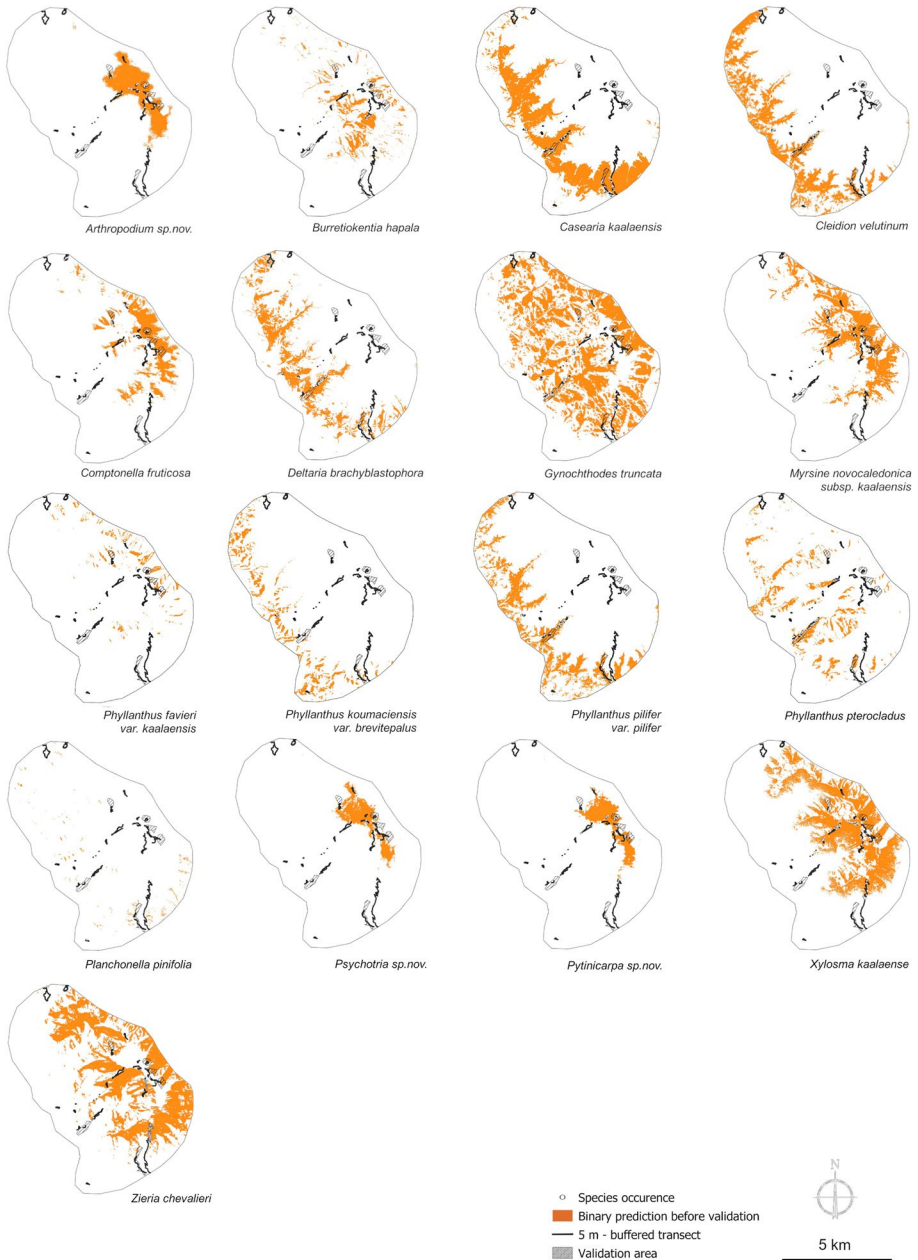
**Table 5** Occurrence number, validation metrics, and percent contribution (in %) of each environmental predictors for every modeled species on the Mount Kaala (New Caledonia)

Family	Taxon	Pre-validation				Post-validation							Vd area						
		Nb occurrences	Background point	Bcont(0,1)	TSS	AUC	Nb occurrences	Background point	Bcont(0,1)	TSS	AUC	DEM		Slope	Irradiation	NDVI	Exposition	Soil	TW1
Apocynaceae	<i>Marsdenia kaalaensis</i>	1	-	-	-	-	8	2005	1	0.9	0.95	23	9	0	4	64	0	0	(X)
Araliaceae	<i>Polyscias sp.nov. (Lamunzel 95)</i>	1	-	-	-	-	5	2005	1	0.59	0.94	23	17	0	0	60	0	0	(X)
Araceae	<i>Burretiokentia hapala</i>	13	1610	0.95	0.92	0.99	13	2005	0.95	0.92	0.99	28	11	19	25	17	1	0	
Asparagaceae	<i>Aristolochium sp.nov. (Karmadi 268)</i>	13	1610	0.94	0.76	0.92	19	2005	0.88	0.62	0.86	61	16	0	10	11	0	2	X
Asteraceae	<i>Pycnanthus sp.nov. (Lamunzel 134)</i>	19	1610	1	0.73	0.92	39	2005	0.89	0.89	0.91	61	3	4	1	27	0	4	X
Capparidaceae	<i>Capparis parvifolia</i>	1	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	
Euphorbiaceae	<i>Cleidion velutinum</i>	51	1610	0.79	0.81	0.93	61	2005	0.95	0.87	0.94	73	2	1	18	1	3	3	X
Myrtaceae	<i>Tristanopsis minimaiflora</i>	90	1610	0.8	0.87	0.99	90	2005	0.8	0.87	0.99	42	32	22	4	0	0	0	
Myrtaceae	<i>Olea hypoleuca</i> var. <i>microphylla</i>	4	1610	0.8	-0.08	0.99	7	2005	1	0.62	0.85	58	0	25	2	0	15	0	(X)
Phyllanthaceae	<i>Phyllanthus decalvatus</i>	21	1610	0.73	0.94	0.98	25	2005	0.9	0.74	0.96	57	10	9	14	3	6	0	X
Phyllanthaceae	<i>Phyllanthus fontieri</i> var. <i>kaalaensis</i>	9	1610	1	0.33	0.99	9	2005	1	0.33	0.99	0	9	35	0	53	0	3	
Phyllanthaceae	<i>Phyllanthus koumacensis</i> var. <i>brevipetalus</i>	6	1610	1	0.48	0.98	6	2005	1	0.48	0.98	41	7	3	14	17	17	0	
Phyllanthaceae	<i>Phyllanthus pilifer</i> var. <i>pilifer</i>	47	1610	0.93	0.68	0.93	55	2005	0.93	0.66	0.93	68	4	0	13	5	8	1	X
Phyllanthaceae	<i>Phyllanthus pierocladus</i>	54	1610	0.94	0.63	0.85	67	2005	0.96	0.68	0.85	32	6	5	29	9	18	1	X
Primulaceae	<i>Myrsine novaealedonica</i> subsp. <i>kaalaensis</i>	12	1610	0.94	0.52	0.75	24	2005	0.92	0.49	0.89	34	38	0	4	20	3	2	X
Rubiaceae	<i>Gynochthodes truncata</i>	51	1610	0.77	0.53	0.69	68	2005	0.7	0.36	0.75	20	4	6	37	5	24	5	X
Rubiaceae	<i>Psychotria</i> sp. nov. (Lamunzel 91)	6	1610	1	0.7	0.96	6	2005	1	0.7	0.96	60	1	38	0	0	0	1	
Rubiaceae	<i>Thiolliera kaalaensis</i>	2	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	
Rutaceae	<i>Comptonella fruticosa</i>	10	1610	1	0.74	0.89	20	2005	0.93	0.83	0.92	35	3	0	19	35	5	3	X
Rutaceae	<i>Zieria chevalieri</i>	259	1610	0.8	0.65	0.65	259	2005	0.8	0.65	0.89	47	1	41	5	4	1	2	-
Salicaceae	<i>Casuaria kaalaensis</i>	47	1610	0.95	0.41	0.86	65	2005	0.85	0.71	0.89	80	5	1	1	13	0	X	
Salicaceae	<i>Xylocarpus kaalaense</i>	64	1610	0.88	0.61	0.83	76	2005	0.89	0.27	0.77	67	23	5	2	0	2	1	X
Sapotaceae	<i>Planchonella minimaiflora</i>	4	1610	-0.46	0.75	0.96	5	2005	1	-0.03	0.68	77	3	4	0	1	15	0	(X)
Sapotaceae	<i>Planchonella pinifolia</i>	4	1610	1	0.99	0.99	4	2005	1	0.99	0.99	31	12	0	0	44	8	6	X
Thymelaeaceae	<i>Delaria brachyblastophora</i>	47	1610	1	0.42	0.91	59	2005	0.97	0.58	0.89	77	7	0	5	3	8	1	X

Cells in grey are poor Bcont(0,1), TSS and AUC values

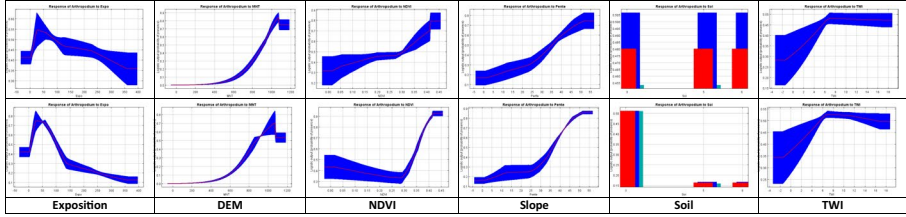
Occ nb = occurrence number, Bg nb = background point number, Bcont(0,1) = Boyce continuous index (Hirzel et al. 2006), DEM = digital elevation model, NDVI = Normalised Difference Vegetation Index, TW1 = Topographic Wetness Index, Vd Area: in validation areas, “X” = species predicted and observed, “(X)” = species not predicted but observed, “-” = species predicted but not observed

**Fig. 5** MaxEnt outputs for each taxon. Marginal response curves above (above), showing how the logistic prediction changes as each environmental variable is varied, keeping all other environmental variables at their average sample value. And responses curves representing a model created using only the corresponding variable (below)

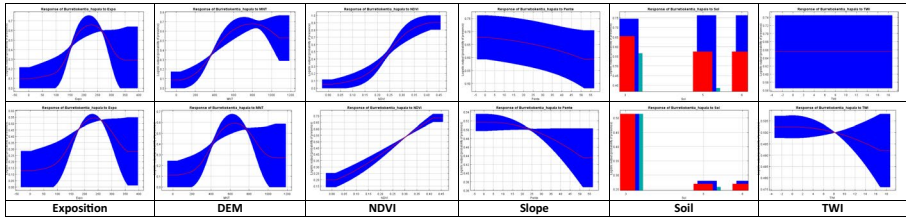


**Fig. 4** Specific maps showing binary predictions obtained before the validation campaign, and all (pre and post validation) occurrences obtained for each species

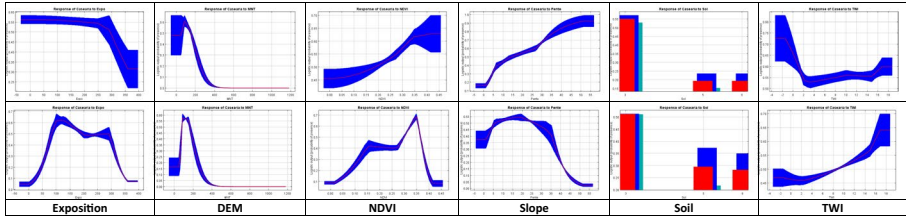
*Arthropodium sp.nov.* (Karnadi 268)



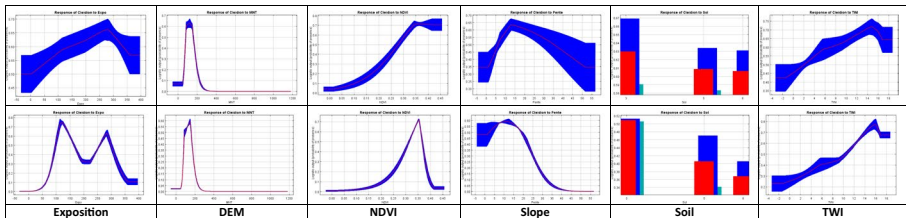
*Burretiokentia hapala*



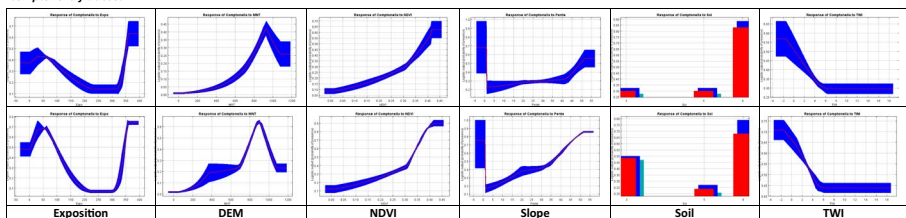
*Casearia kaalaensis*



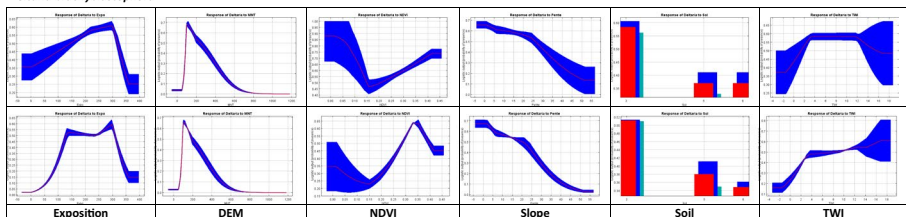
*Cleidion velutinum*



*Comptonella fruticosa*



*Deltaria brachyblastophora*



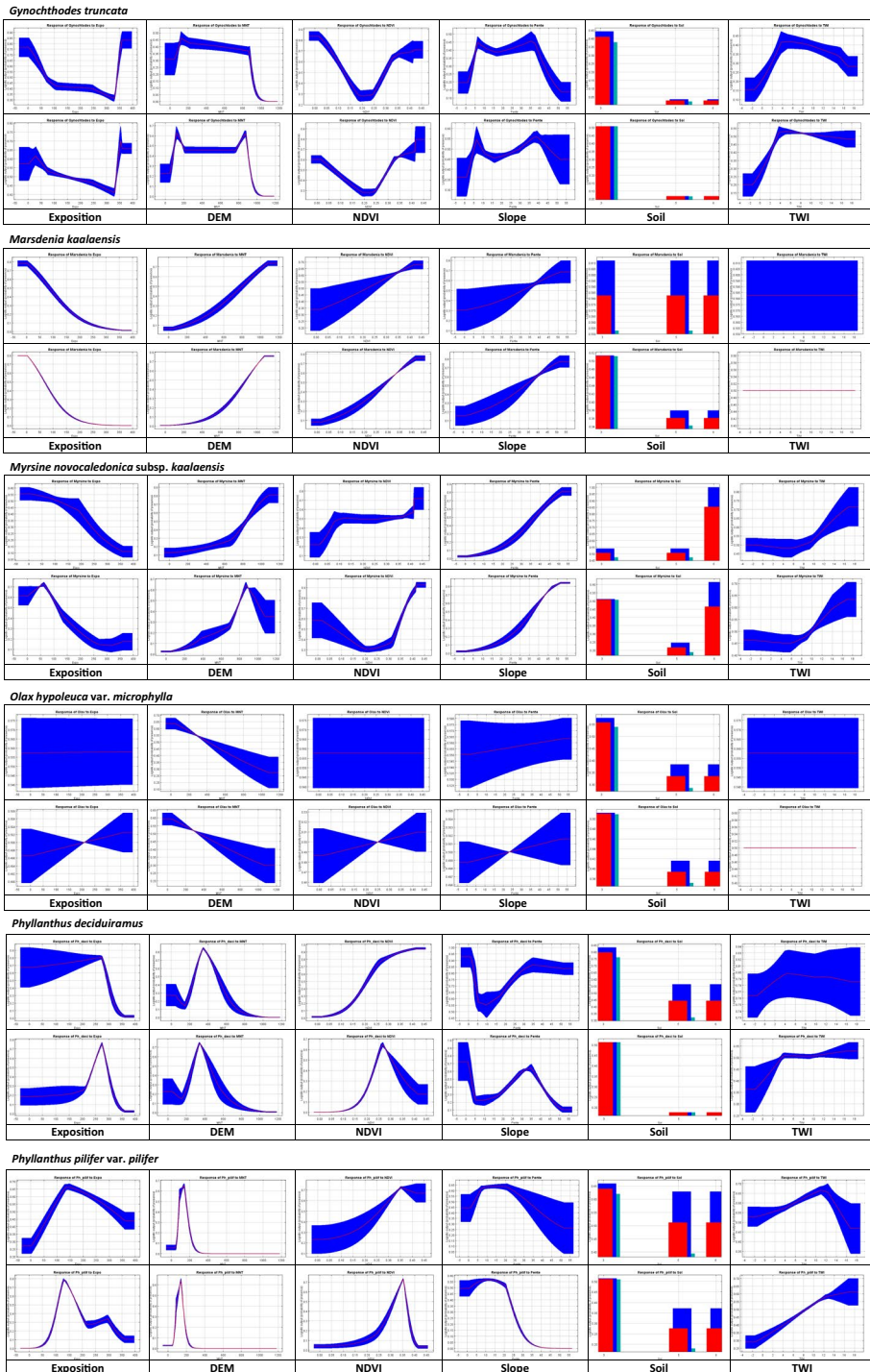
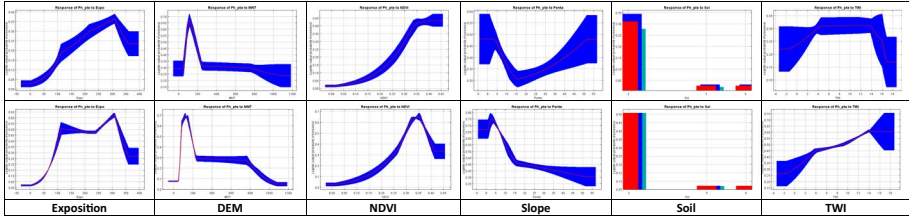
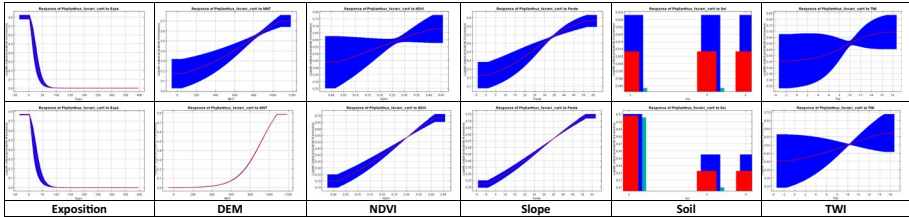


Fig. 5 (continued)

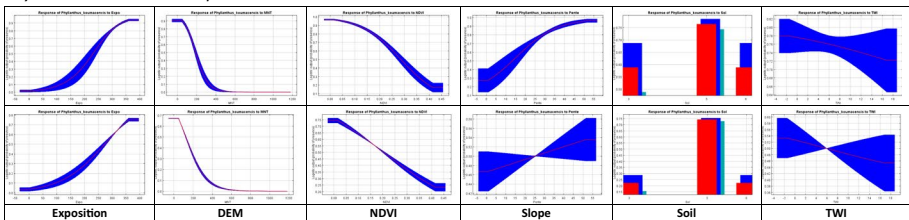
*Phyllanthus pterocladus*



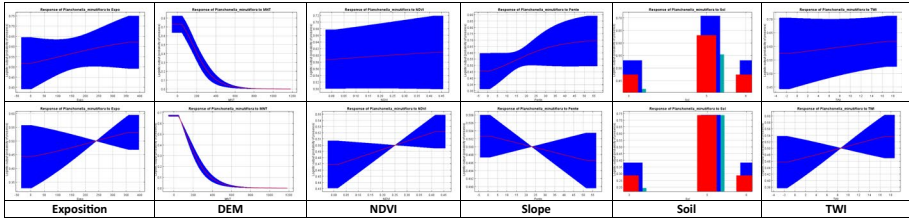
*Phyllanthus javieri* var. *kaalaensis*



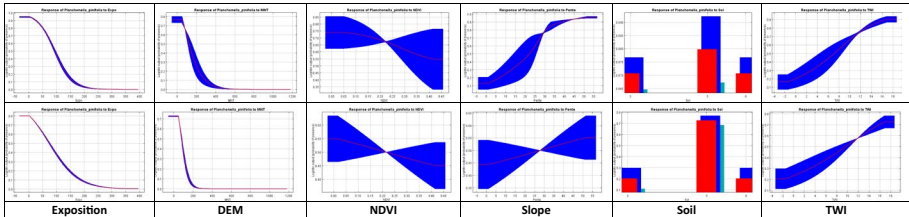
*Phyllanthus koumaccensis* var. *brevitapalus*



*Planchonella minutiflora*



*Planchonella pinifolia*



*Polyscias* sp. nov. (Lannuzel 95)

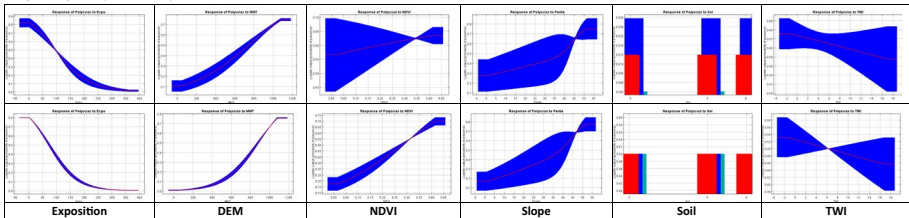


Fig. 5 (continued)

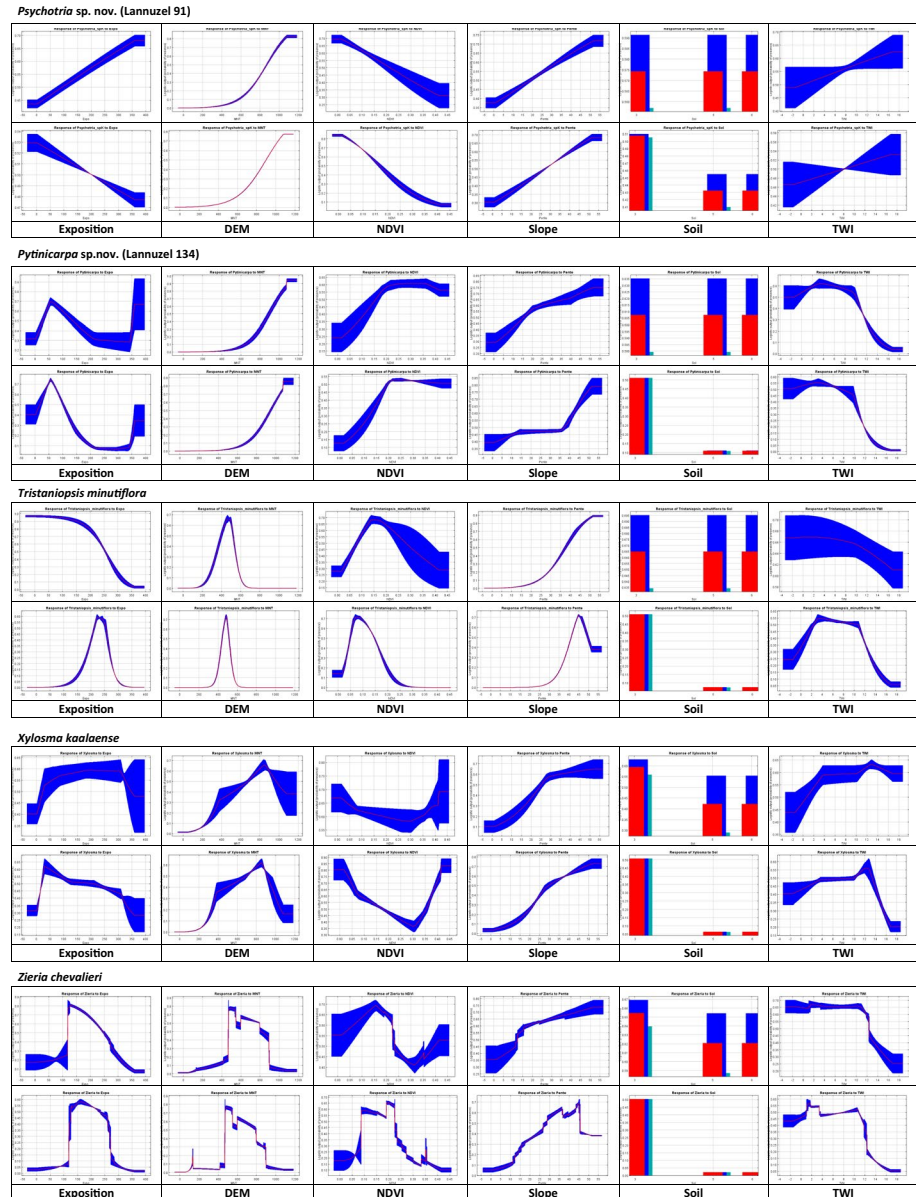


Fig. 5 (continued)

**Acknowledgements** We thank Boris Leroy for useful discussions and wise advice. The authors are grateful to every consulted herbarium for giving access to their data. Roy Benyon is also acknowledged for his precious advice on the English translation. The project was funded by the Société des Mines de la Tontouta, which operates a mine on the top of the mountain, in order to improve conservation efforts within their operational area, and in the surroundings. We thank the Institut Agronomique néo-Calédonien (IAC) for funding JB, MT and BF, and providing field logistical support. We thank the Société Le Nickel (SLN) for

giving access to the southern parts of the mountain, and the Kaala-Gomen municipal authorities for giving us access to the rest of the mountain. We are grateful to François Munoz, Jérôme Munzinger and Sean Tomlinson for their useful comments and additions to the manuscript.

**Author contributions** GL, JB and BF contributed to the study design and conception. Material preparation and data collection were performed by GL and JB. Data analysis and preparation of the first draft of the manuscript were performed by GL, ND and MT. All authors contributed to reviewing and editing the final manuscript. The entire study was supervised by BF.

**Funding** The Société des Mines de la Tontouta (SMT) entirely funded the corresponding author's salary. It gave access to the mining site and surroundings. The SMT has no other role, neither in the study design, nor in the results and interpretations. The Institut Agronomique néo-Calédonien (IAC) supported wages for the rest of the local team, field logistics and other costs, and it participated actively in the study design, results and interpretations through the involvement of its team.

**Data availability** All species studied here are considered as threatened. Therefore, precise locations were not transmitted with the study. However, all data were transmitted to the local red list authority held by the NGO Endemia ([www.endemia.nc](http://www.endemia.nc)) and are available on the website, at a lower resolution. References for all source codes used in statistical analysis were duly acknowledged in the manuscript; they are also available on demand.

## Compliance with ethical standards

**Conflict of interest** The authors declare they have no financial or non-financial conflict of interest to disclose.

**Ethical approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** All authors consent for the publication of this study.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* 43(6):1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Araujo MB, Anderson RP, Barbosa AM, Beale CM, Dormann CF, Early R, Garcia RA, Guisan A, Maiorano L, Naimi B, O'hara RB, Zimmermann NE, Rahbek C (2019) Standards for distribution models in biodiversity assessments. *Sci Adv* 5:eaat4558
- Balaguru B, John Britto SJS, Nagamurugan N, Natarajan D, Soosairaj S (2006) Identifying conservation priority zones for effective management of tropical forests in eastern Ghats of India. *Biodivers Conserv* 15(4):1529–1543. <https://doi.org/10.1007/s10531-004-6678-1>
- Beauregard F, de Blois S (2014) Beyond a climate-centric view of plant distribution: edaphic variables add value to distribution models. *PLoS One* 9(3):e92642. <https://doi.org/10.1371/journal.pone.0092642>
- Barrière R, Dagostini G, Rigault F, Munzinger J (2007) Caractérisation et cartographie de la végétation des milieux miniers de la côte Nord Ouest—Technical report. IRD/DDEE Province Nord, Nouméa

- Blanchard G, Munoz F, Ibanez T, Hequet V, Vandrot J, Girardi J, Birnbaum P (2019) Regional rainfall and local topography jointly drive tree community assembly in lowland tropical forests of New Caledonia. *J Veg Sci* 30:345–356. <https://doi.org/10.1111/jvs.12781>
- Blasi C, Marignani M, Copiz R, Fipaldini M, Bonacquisti S, Del Vico E et al (2011) Important plant areas in Italy: from data to mapping. *Biol Conserv* 144(1):220–26. <https://doi.org/10.1016/j.biocon.2010.08.019>
- Bonvallot J, Gay J-C, Habert E (2012) Atlas de la Nouvelle-Calédonie. IRD, Nouméa
- Booth TH, Nix HA, Brusby JR, Hutchinson MF (2014) BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Divers Distrib* 20:1–9. <https://doi.org/10.1111/ddi.12144>
- Cañadas EM, Fenu G, Peñas J, Lorite J, Mattana E, Bacchetta G (2014) Hotspots within hotspots: endemic plant richness, environmental drivers, and implications for conservation. *Biol Conserv* 170:282–291. <https://doi.org/10.1016/j.biocon.2013.12.007>
- Cao Y, DeWalt RE, Robinson JL, Tweddle T, Hinz L, Pressino M (2013) Using Maxent to model the historic distributions of stonefly species in Illinois streams: the effects of regularization and threshold selections. *Ecol Model* 259:30–39. <https://doi.org/10.1016/j.ecolmodel.2013.03.012>
- Caujapé-Castells J, Tye A, Crawford DJ, Santos-Guerra A, Sakai A, Beaver K et al (2010) Conservation of oceanic island floras: present and future global challenges. *Perspect Plant Ecol Evol Syst* 12:107–129. <https://doi.org/10.1016/j.ppees.2009.10.001>
- Cayuela L, Golicher DJ, Newton AC, Kolb M, de Albuquerque FS, Arets EJM et al (2009) Species distribution modelling in the tropics: problems, potentialities, and the role of biological data for effective species conservation. *Trop Conserv Sci* 2(3):319–352. <https://doi.org/10.1177/194008290900200304>
- Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM (2015) Accelerated modern human-induced species losses: entering the sixth mass extinction. *Sci Adv* 1(5):e1400253. <https://doi.org/10.1126/sciadv.1400253>
- Collen B, Ram M, Zamin T, McRae L (2008) The tropical biodiversity data gap: addressing disparity in global monitoring. *Trop Conserv Sci* 1(2):75–88. <https://doi.org/10.1177/194008290800100202>
- Darbyshire I, Anderson S, Asatryan A, Byfield A, Cheek A, Clubbe C et al (2017) Important plant areas: a revised selection criteria for a global approach to plant conservation. *Biodivers Conserv* 26:1767. <https://doi.org/10.1007/s10531-017-1336-6>
- Di Cola V, Groenmann O, Petitpierre B, Breiner FT, D’Amen M, Randin C, Engler R, Pottier J, Pio D, Dubuis A, Pelissier L (2017) ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography* 40:774–787. <https://doi.org/10.1111/ecog.02671>
- Duran AP, Rauch J, Gaston KJ (2013) Global spatial coincidence between protected areas and metal mining activities. *Biol Conserv* 160:272–278. <https://doi.org/10.1016/j.biocon.2013.02.003>
- El-Gabbas A, Gilbert F, Dormann CF (2020) Spatial conservation prioritisation in data-poor countries: a quantitative sensitivity analysis using multiple taxa. *BMC Ecol* 20:35
- Elith J, Phillips SJ, Hastie T, Dudik M, En Chee Y, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Divers Distrib* 17:43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- Fick SE, Hijmans RJ (2017) Worldclim2: new 1 km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302–4315. <https://doi.org/10.1002/joc.5086>
- Fois M, Fenu G, Lombrana AC, Cogoni D, Bacchetta G (2015) A practical method to speed up the discovery of unknown populations using Species Distribution Models. *J Nat Conserv* 24:42–48. <https://doi.org/10.1016/j.jnc.2015.02.001>
- Fourcade Y, Engler JO, Rödder D, Secondi J (2014) Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PLoS One* 9(5):e97122. <https://doi.org/10.1371/journal.pone.0097122>
- Franklin J, Davis FW, Ikegami M, Syphard A, Flints LE, Hannah L (2013) Modelling plant species distributions under future climates: how fine scale do climate projections need to be? *Glob Chang Biol* 19:473–483. <https://doi.org/10.1111/gcb.12051>
- Gallardo B, Aldridge DC (2013) Evaluating the combined threat of climate change and biological invasions on endangered species. *Biol Conserv* 160:225–233. <https://doi.org/10.1016/j.biocon.2013.02.001>
- Gâteblé G, Barrabé L, McPherson G, Munzinger N, Swenson U (2018) One new endemic plant species on average per month in New Caledonia, including eight more species from Ile Art (Belep Islands), a major micro-hotspot in need of protection. *Aust Syst Bot* 31:448–480. <https://doi.org/10.1071/SB18016>
- Gogol-Prokurat M (2011) Predicting habitat suitability for rare plants at local spatial scales using a species distribution model. *Ecol Appl* 21(1):33–47. <https://doi.org/10.1890/09-1190.1>
- Gouvernement de la Nouvelle-Calédonie (2017) Modèle Numérique de Terrain (MNT). Géorep-Portail de l’information géographique de la Nouvelle-Calédonie. [www.georep.nc](http://www.georep.nc)





- Guisan A, Tingley R, Baumgartner JB, Naujokaitis-Lewis I, Sutcliffe PR, Tulloch AIT et al (2013) Predicting species distributions for conservation decisions. *Ecol Lett* 16:1424–35. <https://doi.org/10.1111/ele.12189>
- Hertzog LR, Besnard A, Jay-Robert P (2014) Field validation shows bias-corrected pseudo-absence selection is the best method for predictive species–distribution modelling. *Divers Distrib* 20:1403–1413
- Heywood VH (2011) The hazardous future of island floras. In: Bramwell D, Caujapé-Castells J (eds) *The biology of island floras*. Cambridge University Press, Cambridge, p 522
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surface for global land areas. *Int J Climatol* 25:1965–1978. <https://doi.org/10.1002/joc.1276>
- Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat models to predicts species presences. *Ecol Model* 199:142–152. <https://doi.org/10.1016/j.ecolmodel.2006.05.017>
- Ibanez T, Birnbaum P, Gâteblé G, Hequet V, Isnard S, Munzinger J et al (2019) Twenty years after Jaffré et al. (1998), is the system of protected areas now adequate in New Caledonia? *Biodivers Conserv* 28(1):245–254. <https://doi.org/10.1007/s10531-018-1659-y>
- Illan JG, Gutiérrez D, Wilson RJ (2010) The contributions of topoclimate and land cover to species distributions and abundance: fine-resolution tests for a mountain butterfly fauna. *Glob Ecol Biogeogr* 19:159–173. <https://doi.org/10.1111/j.1466-8238.2009.00507.x>
- Isnard S, L’Huillier L, Rigault F, Jaffré T (2016) How did the ultramafic soils shape the flora of the New Caledonian hotspot? *Plant Soil* 403:53–76. <https://doi.org/10.1007/s11104-016-2910-5>
- IUCN (2012) *Catégories et critères de la liste rouge de l’UICN*. 3.1. Gland: UICN
- IUCN (2016) « A global standard for the identification of key biodiversity areas ». 1. Gland, Switzerland: IUCN
- IUCN/SSC (2013) « Guidelines for reintroductions and others conservation translocations ». 1. Gland, Switzerland: IUCN Species Survival Commission
- Jaffré T (1993) Relation between ecological diversity and floristic diversity in New Caledonia. *Biodivers Lett* 1:82–91. <https://doi.org/10.2307/2999751>
- Jaffré T, Munzinger J, Lowry PP (2010) Threats to the conifer species found on New Caledonian’s ultramafic massifs and proposal for urgently needed measures to improve their protection. *Biodivers Conserv* 19:1485–1502. <https://doi.org/10.1007/s10531-010-9780-6>
- Jiménez-Valverde A (2012) Insights into the area under the receiver operating characteristic curve (AUC) as a discriminating measure in species distribution modeling. *Glob Ecol Biogeogr* 21:498–507. <https://doi.org/10.1111/j.1466-8238.2011.00683.x>
- Jiménez-Valverde A, Lobo JM (2007) Threshold criteria for conversion of probability of species presence to either–or presence–absence. *Acta Oecol* 31:361–369. <https://doi.org/10.1016/j.actao.2007.02.001>
- Juffroy F (2012) Génération du Modèle numérique de terrain au pas de 10 m sur la Nouvelle-Calédonie. DTSI, Nouméa, p 15
- Karger DN, Conrad O, Böhrer J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M (2017) Climatologies at high resolution for the earth’s land surface areas. *Sci Data* 4:170122. <https://doi.org/10.1038/sdata.2017.122>
- Kier G, Kreft H, Lee TM, Jetz W, Ibisch PL, Nowicki J et al (2009) A global assessment of endemism and species richness across island and mainland regions. *PNAS* 106(23):9322–27. <https://doi.org/10.1073/pnas.0810306106>
- Koch R, Almeida-Cortez J, Kleinschmitt B (2017) Revealing areas of high nature conservation importance in a seasonally dry tropical forest in Brazil: combination of modelled plant diversity hot spots and threat patterns. *J Nat Conserv* 35:24–39
- Kramer-Schadt S, Niedballa J, Pilgrim JD, Schröder B, Lindenborn J, Reinfelder V et al (2013) The importance of correcting for sampling bias in MaxEnt species distribution models. *Divers Distrib* 19(11):1366–79. <https://doi.org/10.1111/ddi.12096>
- L’Huillier L, Jaffré T, Wulff A (2010) Mines et environnement en Nouvelle-Calédonie: les milieux sur substrats ultramafiques et leur restauration. IAC Edition, p 412
- Le Lay G, Engler R, Franc E, Guisan A (2010) Prospective sampling based on model ensembles improves the detection of rare species. *Ecography* 33(6):1015–1027. <https://doi.org/10.1111/j.1600-0587.2010.06338.x>
- Lefcort H, Vancura J, Lider EL (2010) 75 years after mining ends stream insect diversity is still affected by heavy metals. *Ecotoxicology* 19(8):1416–1425. <https://doi.org/10.1007/s10646-010-0526-8>
- Leroy B, Bellard C, Dubos N et al (2014) Forecasted climate and land use changes, and protected areas: the contrasting case of spiders. *Divers Distrib* 20:686–697. <https://doi.org/10.1111/ddi.12191>

- Leroy B, Delsol R, Huguény B, Meynard CN, Barhoumi C, Barbet-Massin M, Bellard C (2018) Without quality presence–absence data, discrimination metrics such as TSS can be misleading measures of model performance. *J Biogeogr* 45(9):1994–2002. <https://doi.org/10.1111/jbi.13402>
- Loiselle BA, Howell CA, Graham CH, Goerck JM, Brooks T, Smith KG, Williams PH (2003) Avoiding pitfalls of using species distribution models in conservation planning. *Conserv Biol* 17(6):1591–1600. <https://doi.org/10.1111/j.1523-1739.2003.00233.x>
- Maurizot P, Schmitt C, Vendé-Leclerc M (2005) Harmonisation de la couverture cartographique géologique numérique de la Nouvelle-Calédonie, phase 4. BRGM, Nouméa, p 22
- McCoy S, Jaffré T, Rigault F, Ash JE (1999) Fire and succession in the ultramafic maquis of New Caledonia. *J Biogeogr* 26(3):579–594. <https://doi.org/10.1046/j.1365-2699.1999.00309.x>
- McGeoch MA, Genovesi P, Bellingham PJ, Costello MJ, McGrannachan C, Sheppard A (2016) Prioritizing species, pathways, and sites to achieve conservation targets for biological invasion. *Biol Invasions* 18(2):299–314. <https://doi.org/10.1007/s10530-015-1013-1>
- Meineri E, Hylander K (2017) Fine-grain, large domain climate models based on climate station and comprehensive topographic information improve microrefugia detection. *Ecography* 40:1003–1013. <https://doi.org/10.1111/ecog.02494>
- Melville GJ, Welsh AH (2001) Line transect sampling in small regions. *Biometrics* 57(4):1130–1137. <https://doi.org/10.1111/j.0006-341X.2001.01130.x>
- Munzinger J, Morat P, Jaffré T, Gâteblé G, Pillon Y, Rouhan G, Bruy D, Veillon J-M, Chalopin M (2020) [continuously updated] FLORICAL: checklist of the vascular indigenous flora of New Caledonia. <http://publish.plantnet-project.org/project/florical>
- Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403(6):853–858. <https://doi.org/10.1038/35002501>
- Nezer O, Bar-David S, Gueta T, Carmel Y (2016) High-resolution species-distribution model based on systematic sampling and indirect observations. *Biodivers Conserv* 26:421. <https://doi.org/10.1007/s10531-016-1251-2>
- Ogutu JO, Bholá N, Piepho HP, Reid R (2006) Efficiency of strip- and line-transect surveys of African savanna mammals. *J Zool*. <https://doi.org/10.1111/j.1469-7998.2006.00055.x>
- Palmer MA, Bernhardt ES, Schlesinger WH, Eshleman KN, Foufoula-Georgiou E, Hendryx MS et al (2010) Mountaintop mining consequences. *Science* 327(5962):148–149. <https://doi.org/10.1126/science.1180543>
- Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J Biogeogr* 34:102–117. <https://doi.org/10.1111/j.1365-2699.2006.01594.x>
- Pettorelli N, Vik JO, Mysterud A, Gaillard J-M, Tucker CJ, Stenseth NC (2005) Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol Evol* 20(9):503–509. <https://doi.org/10.1016/j.tree.2005.05.011>
- Phillips SJ (2008) Transferability, sample selection bias and background data in presence only modelling: a response to Peterson et al. (2007). *Ecography* 31:272–278. <https://doi.org/10.1111/j.0906-7590.2008.5378.x>
- Phillips SJ (2017) A brief tutorial on MaxEnt. [http://biodiversityinformatics.amnh.org/open\\_source/maxent/](http://biodiversityinformatics.amnh.org/open_source/maxent/)
- Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31:161–175. <https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- Phillips SJ, Dudík M, Schapire RE (2004) A maximum entropy approach to species distribution modeling. In: Proceedings of the twenty-first international conference on machine learning, p 83. ACM. <http://dl.acm.org/citation.cfm?id=1015412>
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Model* 190(3–4):231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: implication for background and pseudo-absence data. *Ecol Appl* 19:181–197. <https://doi.org/10.1890/07-2153.1>
- Pouteau R, Meyer J-Y, Taputuarai R, Stoll B (2012) Support vector machines to map rare and endangered native plants in Pacific islands forests. *Ecol Inform* 9:37–46. <https://doi.org/10.1016/j.ecoinf.2012.03.003>
- Pouteau R, Bayle E, Blanchard E, Birnbaum P, Cassan J-J, Hequet V, Ibanez T, Vandrot H (2015) Accounting for the indirect area effect in stacked species distribution models to map species richness in a montane biodiversity hotspot. *Divers Distrib* 21:1329–1338. <https://doi.org/10.1111/ddi.12374>

- Pouteau R, Gillespie TW, Birnbaum P (2018) Predicting tropical tree species richness from normalized difference vegetation index time series: the devil is perhaps not in the detail. *Remote Sens* 10:698. <https://doi.org/10.3390/rs10050698>
- R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>
- Radula MW, Szymura TH, Szymura M (2018) Topographic wetness index explains soil moisture better than bioindication with Ellenberg's indicator values. *Ecol Indic* 85:172–179. <https://doi.org/10.1016/j.ecolind.2017.10.011>
- Rhoden CM, Peterman WE, Taylor CA (2017) Maxent-directed field surveys identify new populations of narrowly endemic habitat specialists. *PeerJ* 5:e3632. <https://doi.org/10.7717/peerj.3632>
- RLA-NC (2019) La Liste rouge de la flore menacée de Nouvelle-Calédonie-synthèse Mars 2019, p 4, Nouméa. <http://endemia.nc/files/201903-plaquette-RLA-Flore-web-v4.pdf>
- Somodi I, Lepesi N, Botta-Dukat Z (2017) Prevalence dependence in model goodness measures with special emphasis on true skill statistics. *Ecol Evol* 7:863–872. <https://doi.org/10.1002/ece3.2654>
- Tershy BR, Shen K-W, Newton KM, Holmes ND, Croll DA (2015) The importance of islands for the protection of biological and linguistic diversity. *Bioscience* 65(6):592–597. <https://doi.org/10.1093/biosci/biv031>
- Ter Steege H, Haripersaud PP, Banki OS, Schieving F (2011) A model of botanical collectors' behaviour in the field: never the same species twice. *Am J Bot* 98(1):31–37. <https://doi.org/10.3732/ajb.1000215>
- Tomlinson S, Lewandowski W, Elliott CP, Miller BP, Turner SR (2019) High-resolution distribution modeling of a threatened short-range endemic plant informed by edaphic factors. *Ecol Evol* 10:763–777. <https://doi.org/10.1002/ece3.5933>
- Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M (2003) Remote sensing for biodiversity science and conservation. *Trends Ecol Evol* 18:306–314. [https://doi.org/10.1016/S0169-5347\(03\)00070-3](https://doi.org/10.1016/S0169-5347(03)00070-3)
- USGS (2017) « Landsat 8 March ». Courtesy of the U.S. Geological Survey
- Weigelt P, Steinbauer MJ, Cabral JS, Kreft H (2016) Late Quaternary climate change shapes island biodiversity. *Nature* 532(7597):99. <https://doi.org/10.1038/nature17443>
- Western AW, Grayson RB, Blöschl G, Willgoose GR, McMahon TA (1999) Observed spatial organization of soil moisture and its relation to terrain indices. *Water Resour Res* 35(3):797–810. <https://doi.org/10.1029/1998WR900065>
- Whittaker RJ, Fernandez-Palacios JM (2007) Island biogeography: ecology, evolution, and conservation, 2nd edn. Oxford University Press, Oxford, p 402
- Williams JN, Seo C, Thorne J, Nelson JK, Erwin S, O'Brien JM, Schwartz MW (2009) Using species distribution models to predict new occurrences for rare plants. *Divers Distrib* 15:565–576. <https://doi.org/10.1111/j.1472-4642.2009.00567.x>
- Wilson JP, Gallant JC (2000) Terrain analysis: principles and applications. Wiley, New York, p 479
- Wilson JW, Sexton JO, Jobe RT, Haddad NM (2013) The relative contribution of terrain, land cover, and vegetation structure indices to species distribution models. *Biol Conserv* 164:170–176. <https://doi.org/10.1016/j.biocon.2013.04.021>
- Wisn MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, NCEAS Predicting Species Distributions Working Groups (2008) Effects of sample size on the performance of species distribution models. *Divers Distrib* 14(5):763–773. <https://doi.org/10.1111/j.1472-4642.2008.00482.x>
- Wulff A, Hollingsworth PM, Ahrends A, Jaffré T, Veillon J-M, L'Huillier L, Fogliani B (2013) Conservation priorities in a biodiversity hotspot: analysis of narrow endemic plant species in New Caledonia. *PLoS One* 8(9):e73371. <https://doi.org/10.1371/journal.pone.0073371>
- Yackulic CB, Chandler R, Zipkin EF, Royle JA, Nichols JD, Campbell Grant EH, Veran S (2013) Presence-only modelling using MAXENT: when can we trust the inferences? *Methods Ecol Evol* 4:236–243. <https://doi.org/10.1111/2041-210x.12004>

## Authors and Affiliations

**Guillaume Lannuzel<sup>1</sup>**  · **Joan Balmot<sup>1,2</sup>** · **Nicolas Dubos<sup>3</sup>** · **Martin Thibault<sup>4</sup>** · **Bruno Fogliani<sup>1</sup>** 

<sup>1</sup> Institut Agronomique néo-Calédonien (IAC), Equipe ARBOREAL (Agriculture Biologique et Valorisation), BP 73, 98890 Païta, New Caledonia

<sup>2</sup> Département de Biologie et Ecologie, Université de Montpellier, Montpellier, France

<sup>3</sup> Centre d'Ecologie et des Sciences de la Conservation (CESCO UMR 7204), Sorbonne Universités, MNHN, Paris, France

<sup>4</sup> UMR250 ENTROPIE, Institut de Recherche pour le Développement, Nouméa, New Caledonia