



HAL
open science

Pre-assessments of plant conservation status in islands: the case of French Overseas Territories

Simon Véron, Arthur Bernard, Elise Lebreton, Carlos Rodrigues-Vaz, M Durand, Lilian Procopio, Mike Héliion, Marc Gayot, Guillaume Viscardi, Gary A Krupnick, et al.

► To cite this version:

Simon Véron, Arthur Bernard, Elise Lebreton, Carlos Rodrigues-Vaz, M Durand, et al.. Pre-assessments of plant conservation status in islands: the case of French Overseas Territories. *Biodiversity and Conservation*, 2023, 32 (4), pp.1165-1187. 10.1007/s10531-023-02544-8 . mnhn-03952485

HAL Id: mnhn-03952485

<https://mnhn.hal.science/mnhn-03952485v1>

Submitted on 14 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Pre-assessments of plant conservation status in islands: 2 the case of French Overseas Territories

3

4 **Authors:** Véron S.^{*a}, Bernard A.^a, Lebreton E.^{a,b}, Rodrigues-Vaz C.^{a,c}, Durand M.^a, Procopio L.^d, Héliou
5 M.^d, Gayot M.^e, Viscardi G.^f, Krupnick G.A.^g, Carrington C.M.S.^h, Boulet V.^{ij}, Mallet B.ⁱ, Dimassi A.^k,
6 Pailler T.^l, Hivert J.ⁱ, Lebouvier M.^m, Agnola P.ⁿ, Bruy D.^{o,p}, Gateblé G.^q, Lannuzel G.^r, Meyer S.^r,
7 Gargominy O.^s, Gigot G.^s, Invernón V.^a, Leblond S.^{a,s}, Pignal M.^a, Terçerie S.^s, Muller S.^a, Rouhan G.^a

8 *Corresponding author¹: simon.veron@mnhn.fr - +33 1 47 07 78 58

9 ^a Institut de Systématique, Évolution, Biodiversité (ISYEB), Muséum National d'Histoire Naturelle,
10 Sorbonne Université, École Pratique des Hautes Études, CNRS, Université des Antilles, 57 Rue Cuvier,
11 CP 39, 75005, Paris, France

12 ^b Institut de Botanique B22, Campus du Sart Tilman - Quartier VALLEE 1 Chemin de la Vallée, 44000
13 Liège, Belgique

14 ^c Diversité, adaptation, développement des plantes (DIADE), Institut de Recherche pour le
15 Développement (IRD), Université de Montpellier, 34394 Montpellier, France

16 ^d Gwada Botanica Chemin La Chaise - 97 115 Sainte-Rose, Guadeloupe

17 ^e Arboretum de Montebello, Office Nationale des Forêts. 97170 Petit Bourg, Guadeloupe

18 ^f Conservatoire Botanique National de Martinique, Espace Camille Darsières – BP 4033 - 97254 Fort
19 de France CEDEX, Martinique

20 ^g Department of Botany, National Museum of Natural History, Smithsonian Institution, PO Box 37012,
21 MRC-166, Washington DC 20013-7012, USA.

22

23 ^h Dept. of Biological & Chemical Sciences, The University of the West Indies, Cave Hill Campus,
24 BB11000, Barbados

25 ⁱ Conservatoire Botanique National de Mascarin, 2 rue du Père Georges - Les Colimaçons, 97436,
26 Saint-Leu, La Réunion, France

27 ^j Université de Bretagne Occidentale, 29200, Brest, France

28 ^k Conservatoire Botanique National de Mascarin - Antenne de Mayotte, 1 ruelle Chamodeau, Route
29 Nationale 2 Coconi, 97670 Ouangani, Mayotte

¹ Present address : Comité français de l'UICN- Union internationale pour la conservation de la nature
259-261 rue de Paris - 93100 MONTREUIL – simon.veron@uicn.fr

30 ^l UMR C53 Cirad / université de La Réunion. Peuplements Végétaux et Bioagresseurs en Milieu
31 Tropical 117 rue du Général Ailleret, 97430 Le Tampon, Réunion – France

32 ^m CNRS, EcoBio (Ecosystèmes, biodiversité, évolution) - UMR 6553, Université Rennes 1, 35000,
33 Rennes, France

34 ⁿ Terres Australes et Antarctiques Françaises, Direction de l'environnement, 1 Rue Gabriel Dejean,
35 97410, Saint-Pierre, la Réunion
36

37 ^o AMAP, Université Montpellier, IRD, CIRAD, CNRS, INRAE, Montpellier, France

38 ^p AMAP, IRD, Herbar de Nouvelle-Calédonie, Nouméa, Nouvelle-Calédonie

39 ^q Institut Agronomique néo-Calédonien, Equipe ARBOREAL, BP 711, 98810 Mont-Dore, Nouvelle-
40 Calédonie

41 ^r Endemia, New Caledonia Plants Red List Authority, BP 4682, 98847, Nouméa, Nouvelle-Calédonie

42 ^s Unité d'appui et recherche PatriNat (OFB-CNRS-MNHN), 36 rue Geoffroy Saint-Hilaire, Maison
43 Buffon - CP 41, Cedex 05, 75231, Paris, France

44 **Acknowledgments:** We thank the Office Français de la Biodiversité (French Agency of Biodiversity)
45 and the Ministère des Outre-mer (French Ministry of Overseas Territories) for funding and supporting
46 this project. G. Gateblé, D. Bruy, G. Lannuzel and S. Meyer are grateful to “CNRT Nickel et son
47 Environnement, Nouméa, Nouvelle-Calédonie” for funding the ERMINES project through Grant CSF
48 n°4PS2017-CNRT.IAC/ERMINE. The authors thank the French Southern and Antarctic Territories
49 (TAAF) for their support for the inventory programs of the flora and vegetation of the Scattered
50 Islands. We are grateful to RECOLNAT national Research Infrastructure (ANR-11-INBS-0004) and
51 especially all contributors to the “Herbonautes” citizen science program for databasing more than
52 16,000 Paris herbarium specimens. We would like to thanks the INRAE Guadeloupe for allowing
53 access to Fournet’s herbarium data. Last but not least, we thank all contributors for regional datasets
54 that were used to conduct this study.

55

56

57

58

59

60

61

62 **Abstract**

63 Assessment methods have been developed to estimate a preliminary conservation status for
64 species and subsequently to facilitate the building of Red Lists. Such pre-assessment methods
65 could be particularly useful in the French Overseas Territories (FOTs) where Red Lists tend
66 to be out-dated or absent and where a high number of endemic species face detrimental
67 anthropogenic pressures. We first aimed to conduct a preliminary assessment (hereafter, pre-
68 assessment) of the conservation status of endemic plants from Guadeloupe, Martinique,
69 Réunion, Mayotte, French sub-Antarctic islands, New Caledonia, and Scattered Islands. We
70 then compared the various methods used in conducting the pre-assessment and discussed
71 ways to adapt these methods to small territories. We compiled occurrence data of endemic
72 species identified thanks to a previous taxonomic work and pre-assessed their conservation
73 status under Red List criteria A and B and the use of a Random Forest algorithm. We then
74 measured the accuracy, specificity, and sensitivity of each method based on existing Red
75 Lists. The Random Forest algorithm and a method based on range-size performed best at
76 correctly attributing conservation status. Using these pre-assessment methods, we estimated
77 that up to 60% of the endemic flora of the FOTs is potentially threatened. Range restriction
78 but also anthropogenic pressures were key factors that explained these risks. Pre-assessment
79 methods are useful tools to get a first measure of species conservation status. These methods
80 should be adapted to the territories considered and their conservation issues in order to reach a
81 good performance.

82 **Keywords:** Red List, pre-assessments, islands, tracheophytes, French overseas territories,
83 endemics

84

85 **1. Introduction**

86 Plants are the dominant kingdom in terms of Earth biomass, representing around 82% (c.a.
87 450Gt), and represent around 370 000 species (Bar-On et al. 2018; Roskov et al. 2019,
88 Freiberg et al. 2020). Plants are essential to the functioning of ecosystems and provide
89 invaluable services to humanity. Yet, almost 22% of vascular plants may be threatened with
90 extinction, although only 14% of them have had their conservation status assessed at the
91 global scale (Brummitt et al. 2015; IUCN Red List version 2022-2). The conservation status
92 of species is essential for the prioritization of conservation actions and one of the tools
93 designed for this purpose is the Red List of threatened species of the International Union for
94 Conservation of Nature (IUCN) (Mace et al. 2008). The IUCN Red List is the world's most
95 commonly used system for gauging the extinction risk faced by species. It is based on criteria

96 and data to assess for each species: i. a population decline (criterion A); ii. a reduced range
97 (criterion B); iii. a small population and decline (criterion C); iv. a very small population
98 (criterion D); and/or v. a quantitative analysis of the risk of extinction (criterion E) (IUCN
99 2012). These criteria are used to assign a threat category to each species. The Red List allows
100 to consistently document the conservation status of species, to identify areas in need of special
101 protection, to support conservation actions and policies, and to monitor the changing state of
102 biodiversity on Earth (Rodrigues et al. 2006).

103 Assessing the conservation status of a species is a time-consuming and costly process that
104 requires extensive studies on the taxonomy and chorology of species, monitoring the changes
105 of populations and threats in a territory, mobilizing and training a large number of experts,
106 conducting dedicated workshops, processing data and publishing results, among other tasks
107 (UICN France 2018). It therefore proves impractical to assess and regularly update the status
108 of every plant species on the planet, especially as new species are described every day (on
109 average 2 400 annually in the last 10 years [IPNI 2021]). In addition, existing plant
110 assessments have not accumulated on the Red List in a systematic way due to different types
111 of biases. For example, more attention is paid to plant species expected to be threatened, so
112 that species that may be secure are under-evaluated (Bachman et al. 2019). To overcome these
113 difficulties, methods and tools for preliminary conservation assessments (hereafter, pre-
114 assessments) have been developed to support and speed up the evaluation of the conservation
115 status of species while following standards close to those used in the Red List framework (e.g.
116 Stévant et al. 2019; Bachman et al. 2020). Although an assessment following Red List
117 standards is necessary to know the genuine conservation status of a species as defined by the
118 IUCN, such pre-assessment methods have been used for comprehensive analyses of potential
119 extinction risks of the flora of Puerto Rico (Miller et al. 2012), Hawaii (Krupnick et al. 2009),
120 the Lesser Antilles (Carrington et al. 2017), and Greece (Kougioumoutzis et al. 2021).

121 Automatic pre-assessment methods have therefore a high potential to assist Red List
122 assessments (Cazalis et al. 2022).

123 In this study we focus on selected French overseas territories (FOTs) and aim to pre-assess the
124 conservation status of their endemic flora. According to existing regional Red Lists (New
125 Caledonia, French Polynesia, Guadeloupe, Martinique, Réunion, Mayotte), more than half of
126 the seed plants endemic to the FOTs could be threatened with extinction, of which nearly 1/3
127 have been assessed as critically endangered (Véron et al. 2021). This figure remains high in
128 ferns and lycophytes since 37% of endemic species are threatened with extinction (Véron et
129 al. 2021). However, there is still a large number of species whose conservation status is
130 unknown or out-of-date and no longer corresponds to the current conservation status of the
131 taxon. The stakes are high since nearly 4 000 species of tracheophytes are endemic to the
132 FOTs (Véron et al. 2021), of which 10 are part of biodiversity hotspots (Myers et al. 2000). In
133 these territories, plant taxa face a vast array of threats, such as deforestation, invasive species,
134 changes in land use, fires, tourism, mining activities or climate change (e.g. UICN France et
135 al. 2014, Meyer et al. 2021). With the exception of French Guiana and Adélie Land, FOTs are
136 islands, and species found in insular territories are likely more vulnerable to anthropogenic
137 pressures than continental species (Russell and Kueffer 2019). Thus, the proportion of
138 threatened endemic plants is higher in the FOTs than in metropolitan France and even higher
139 than on a global scale (Nic Lughadha et al. 2020; Véron et al. 2021).

140 The application of pre-assessment methods to the endemic flora of the FOTs is an important
141 contribution to conservation actions in these territories. These pre-assessments compile
142 baseline data for future actions and these new data could support future Red Lists. The
143 numerous botanical field collecting campaigns carried out for several centuries in the overseas
144 territories, the study of herbaria specimens and their digitization, the existence of national and

145 regional taxonomic repositories, recent inventories and participatory sciences are all assets
146 allowing for the analysis of these pre-evaluations. The detailed knowledge of the flora of
147 certain, but not all, territories and the existence of recent Red Lists (e.g. New Caledonia)
148 make it possible to test the performance of these methods and their bias in order to improve
149 upon them. Knowledge of island characteristics, biogeography, environmental pressures,
150 collecting efforts, conservation issues and available floristic data in the overseas territories
151 means that factors that have not yet been taken into account can be tested to analyze the
152 performance of pre-assessments methods globally. The objectives of this study are therefore
153 to:

- 154 1) Pre-assess the conservation status of the endemic flora of selected FOTs, in
155 particular for Not Evaluated species;
- 156 2) Compare the performance of four assessment methods to correctly predict IUCN
157 Red List conservation status while considering factors absent from previous
158 studies;

159

160 **2. Method**

161 2.1. Compiling and cleaning data

162 The first step of this study was to compile occurrence data of strict endemic and regional
163 endemic plant species (see definition in Supplementary Information 1) of Réunion, Mayotte,
164 Guadeloupe, Martinique, Scattered Islands, French sub-Antarctic Islands and New Caledonia.
165 We focused on these FOTs due to conservation issues for endemic plants and data availability
166 (Supplementary Information 1). To select endemic species we used the French national
167 taxonomic repository TAXREF (Gargominy et al. 2021) and a recent checklist of the FOTs

168 endemic plants (Véron et al. 2021). Only species records within the geographic boundary of
169 the FOTs were used in the analysis. The main data sources included local inventory datasets
170 (e.g. <https://mascarine.cbnm.org/>; Carrington et al. 2018), the National inventory of natural
171 heritage – the official data source of French National Red Lists - (INPN; MNHN and OFB
172 2021), the Global Biodiversity Information Facility (GBIF), the collections of the Paris
173 herbarium (P) along with the ones of the FOTs herbaria (e.g., GUAD, MTK, REU, MAO,
174 NOU) and other international herbaria (e.g., K, MO, G) (all data sources are in Table 1). In
175 total we compiled occurrence records for 3 545 taxa, i.e. 89% of endemic tracheophytes in the
176 FOTs (Véron et al. 2021). Some studies proposed to use only data freely available on the web
177 (GBIF, BIEN etc.), yet this would lead to inaccurate results for the FOTs because they are
178 incomplete and/or not precise enough (all specimen data are not in GBIF, many sensitive data
179 have had their coordinates approximated etc.).

180 Within these compiled datasets (one for each territory), a comparison was made between the
181 names of the species and the French national taxonomic repository TAXREF (Gargominy et
182 al. 2021). The spelling of the names of the taxa present in these datasets was compared to the
183 names present in TAXREF and corrected when necessary. Only taxa considered endemic or
184 subendemic in TAXREF and in the Endemia repository for New-Caledonia were kept. The
185 names in the datasets that are considered synonyms in TAXREF were changed to their
186 accepted names. We kept occurrences at the species and infra-specific levels because: *i.* data
187 often allow the identification of taxa below the species level, *ii.* Red Lists can accommodate
188 infra-specific taxa, *iii.* many taxa are endemic at the infra-specific but not the specific level
189 (see Supplementary Information 2 for results of a complementary analysis at the specific
190 level). It was also necessary to filter the data by removing duplicate occurrences, and by
191 removing or correcting erroneous coordinates (especially occurrences in the sea or at the
192 center of the territory).

193 Many occurrences had locality information but were not georeferenced. However, pre-
194 assessments of IUCN status mostly require spatial coordinates for each collected specimen.
195 When necessary, we therefore created locality indexes (i.e. we attributed spatial coordinates to
196 locality names found on herbarium labels) mainly with the help of botanical experts of the
197 FOTs who have a comprehensive knowledge of these islands and their botanical collections.
198 This made it possible to *post-facto* assign longitude, latitude, and uncertainties on the
199 coordinates for places in Martinique, Guadeloupe, Réunion, New Caledonia, and the French
200 sub-Antarctic Islands. Walker et al. (2022) showed that minimal data-cleaning performed well
201 in the context of automated assessments and that excluding uncertain data would considerably
202 reduce the coverage of automatic assessments. Thus, we preferred to keep most of uncertain
203 data, although we excluded the ones that mentioned a locality that was too vague (e.g.
204 “Guadeloupe”). Moreover, we estimated whether coordinate uncertainty was higher or lower
205 than 2 km, which is the size of the grid cell recommended to estimate species Area of
206 Occupancy (IUCN 2012). We then ran an analysis on the influence of uncertain coordinates
207 on our results (see section 2.3 and Supplementary Information 2). However, we acknowledge
208 that coordinates and uncertainties estimations were often subjective. In total we compiled ca.
209 255 000 georeferenced occurrences of FOTs endemic taxa. Regarding the Scattered Islands
210 we directly extracted range distribution data from Hivert et al. (2018) who estimated species
211 conservation status based on criteria that were adapted from the Red List criteria due to the
212 very small size of the territory and conservation management purposes.

213 2.2.Preliminary Conservation Assessments (objective 1)

214 Pre-assessments are conducted with available data and aim to overcome the difficulty (or
215 impossibility) to exactly reproduce the IUCN approach and standards automatically. To allow
216 approaching the potential threats status of species, these methods sometimes deviate from

217 IUCN criteria and sub-criteria, but instead use variables related to extinction risks which
218 make them highly performant (e.g. Stévant et al. 2019, see also Discussion). Here, we used
219 four successful methods developed in the literature to assign pre-assessment conservation
220 statuses to endemic tracheophytes of the FOTs (Fig. 1). We adapted these methods to each
221 territory according to its geographic characteristics and data availability (table 1).

222 i) Criterion B, considering jointly criteria B1a, B2a, B1b, B2b

223 • Criteria B1a and B2a

224 This method calculates the Area Of Occupancy (AOO), Extent Of Occurrence (EOO) and the
225 number of locations for each species based on georeferenced data (IUCN 2012). The EOO
226 provides information on the overall geographical distribution while the AOO provides
227 information on the area occupied by a species (see detailed definition in IUCN 2012). A
228 location is "a particular area from the ecological and geographical point of view in which a
229 single threatening event can rapidly affect all individuals of the taxon present". AOO, EOO
230 and number of locations allows pre-assessment using criteria B1a and B2a according to
231 Dauby et al. (2017). We used the R package "ConR" and the default options of the
232 "IUCN.eval" function, although sometimes adapted to the geographical characteristics of the
233 territory (Table 1). Thus, the AOO was calculated using a 2km*2km grid whose origin varies
234 over a given number of iterations and we retained the minimum value of the AOO over all
235 iterations (see other methods in Moat et al. (2018)). For a trade-off between efficiency and
236 speed, we chose to vary the grid origin 20 times. For the identification of locations, and due to
237 the difficulty of complying with the definition stated above using automatic methods, a
238 location was defined as a 10km*10km or 5km*5km grid cell encompassing the largest
239 possible number of occurrences (Table 1). Nevertheless, we performed a sensitivity analysis
240 to estimate the influence of grid size on our results (Supplementary Information 2). We did

241 not assess fragmentation as no methods exist to automatically do so. The thresholds of EOO,
242 AOO and number of locations used to pre-assess the category of a taxon were generally those
243 defined by the IUCN (IUCN 2012), except for very small islands. Indeed, IUCN thresholds to
244 assess extinction risk assessments do not always comply with management objectives in very
245 small territories and it is therefore not recommended to use the IUCN standards there. New
246 thresholds were therefore defined for the Scattered islands, Crozet, Saint-Paul and Amsterdam
247 (islands of the French sub-Antarctic islands) at both local/single island and regional scales
248 following Hivert et al. (2018) (Supplementary Information 3). These adaptations do not allow
249 measuring a genuine IUCN conservation status but they are useful for conservation
250 management issues in small territories.

251 • Criteria B1b and B2b

252 This method is adapted from Carrington et al. (2017) and aims to pre-assess a continuing
253 population decline, i.e. criteria B1b and B2b. It is used to estimate a decline in one of the
254 following elements relative to a reference year:

- 255 - the AOO of a taxon estimated using the ConR package (Dauby et al. 2017);
- 256 - habitat quality by assessing the proportion of occurrence of a taxon outside any
257 protected areas (as defined in protectedplanet.net and assuming that protected
258 areas prevent population declines). A potential decline is estimated for this
259 criterion when half of the occurrences are outside the protected areas;
- 260 - the number of locations (Dauby et al. 2017);
- 261 - the proportion of yearly observations for a taxon. We chose to estimate the
262 changes of the proportion of individuals rather than the number of individuals in
263 order to roughly take into account sampling differences between years.

264

265 However, we acknowledge the difficulty of estimating a continuous decline using automatic
266 methods, the often 'opportunistic' data available, and the differences in sampling in time and
267 space (see below how this limitation was taken into account in the analysis of method
268 performance). We chose 2010 as the reference year assuming that this conforms to the IUCN
269 recommendations to estimate a continuous decline (IUCN 2012) and that a 10 year period
270 should ideally be used to estimate the state of a species (UICN France 2018). Yet, as the
271 IUCN guidelines do not recommend a specific threshold to estimate a continuous decline, we
272 also conducted additional analyses by choosing the reference year as the median year of
273 collection for each territory (Supplementary Information 2). Finally, the reference year does
274 not influence the binary classification of species as “threatened” or “non-threatened” (see
275 below).

276 The joint analysis of the pre-assessments of sub-criteria B1a, B2a, B1b, B2b, i.e. based on
277 AOO, EOO, number of locations and population decline, makes it possible to estimate
278 criterion B following the IUCN analysis grid (see IUCN 2012) and attribute a status Critically
279 Endangered (CR), Endangered (EN), Vulnerable (VU) or Near-Threatened or Least-Concern
280 (NT or LC) to species. The expertise of experienced botanists is indispensable to estimate
281 sub-criterion c) [extreme fluctuations] and, to our knowledge, it is not assessable using
282 automatic methods, so it was not considered in this study.

283 ii) Criterion A

284 The purpose of this method is to estimate the decline of a population based on the proportion
285 of the AOO of each species that intersects a degraded area (Stévant et al. 2019). Indeed, the
286 decline of a population can be inferred from habitat losses according to the IUCN (IUCN
287 2012). We identified degraded areas based on land cover data (e.g. Bossard et al. 2000),
288 evolution of land cover (Bossard et al. 2000) and on pressure maps specific to each territory
289 (e.g. loss of forest cover, fire frequency, invasive species) (Table 1). A raster map combining

290 all these degraded areas was built for each territory and was used to calculate a potential
291 decline in the AOO as defined by Stévant et al. (2019). Following Stévant et al. (2019)
292 method, a species is pre-assessed CR, EN or VU if more than 80%, 70% or 50% of its AOO
293 intersects degraded areas, respectively.

294 iii) Criteria A, B

295 This third methodological approach attributes a conservation status corresponding to the most
296 threatened status between criterion A and criterion B. Used jointly, criteria A and B are
297 similar to the PACA method of Stévant et al. (2019), with the difference that the potential
298 decline of a population (criteria B1b and B2b) is estimated differently: by the presence in a
299 degraded habitat according to Stévant et al. (2019), by the comparison to a reference year in
300 our study. Yet, the former option was tested in Supplementary Information 2. In addition,
301 when we did not observe any decline for a taxon but the ConR algorithm (criterion B1a and
302 B2a) classified the species in a threat category based on the EOO, the AOO and the number of
303 locations, we followed the method of Stévant et al. (2019) and defined the categories "likely
304 rare" (CR or EN according to ConR but without observed decline) and "potentially rare" (VU
305 according to ConR but without observed decline). The classification "likely rare" is analogous
306 to the VU category under criterion D2 (i.e. restricted area of occupancy or number of
307 locations with a plausible future threat that could drive the taxon to CR or EX in a very short
308 time). Conversely, a taxon pre-assessed LC or NT based on its AOO or EOO but facing a
309 decline was classified as "LC or NT, potential decline".

310

311 iv) Random Forest (Breiman 2001) – no defined criteria

312 Random Forest is a machine learning method particularly suitable for pre-assessments of
313 extinction risks (Bland et al. 2015; Pelletier et al. 2018; Walker et al. 2022). Here we used the

314 "Random Forest" algorithm to classify endemic plants in a conservation status category, the
315 predictors being the variables used for criteria B (AOO, EOO, number of locations, decline in
316 one of the variables indicated in point *i.*, proportion of occurrence within a protected area) and
317 A (proportion of the AOO degraded for each pressure taken individually and for all pressure
318 layers combined). We also estimated the importance of these variables on IUCN Red List
319 statuses by calculating the Importance of Gini according to the Random Forest algorithm
320 (Supplementary Information 4; Breiman 2001). The model parameters for each territory were
321 adjusted to achieve the best possible performance. The model was trained and evaluated
322 across the entire dataset of taxa having an IUCN status (excluding Data Deficient [DD] and
323 Not Evaluated [NE] taxa), as recommended by Walker et al. (2022) whose results "favor
324 using all assessed species" for training, "even when well-designed sub-samples are available".
325 We then used the classification trees generated in order to estimate a pre-status for DD and
326 NE taxa.

327 We also explored another machine learning method, the neural network approach from (Zizka
328 et al. 2020), using similar predictors (Supplementary Information 2).

329

330 Pre-assessments were performed with the R software (R Core Team 2020) and scripts are
331 available online at https://github.com/arthur-b-1303/Projet_Fentom

332 **Table 1:** Data used for each territory. Criterion B parameters vary between territories in order to reach the best performance possible for this
 333 pre-assessment method. CLC=Corinne Land Cover (Bossard et al. 2000)

	Main sources of species records	Number of occurrence records compiled	Median age of records	Pre-assessed endemic species	Anthropogenic pressures (criterion A)	Criterion B parameters
New Caledonia	Endemia and RLA Flore NC 2019; ERMINES project (for ultramafic substrates); Herbaria P and NOU; Other international herbaria; INPN; GBIF	126 329 records including 121 657 georeferenced	1978	2 591 (out of 2 758 endemic species)	Urban and agricultural environments (Gouvernement de la Nouvelle-Calédonie 2014); Loss of forest cover (Hansen et al. 2013); Fires (OEIL 2021); Active mines (Gouvernement de la Nouvelle-Calédonie 2014)	AOO, EOO according to recommendations of (IUCN 2012) Number of locations = 10km*10km square grid
Réunion	Mascarine; Herbarium REU; Herbarium P; INPN; GBIF	102 406 records including 102 208 georeferenced	2010	413 (out of 489 endemic species)	Urban and agricultural environments (Bossard et al. 2000); Land use change 2012-2018 (Bossard et al. 2000); loss of forest cover (Hansen et al. 2013); invasive species (Fenouillas et al. 2021)	AOO, EOO (IUCN 2012) Number of locations = 10km*10km square grid
Mayotte	Mascarine; Herbarium MAO; Herbarium P; INPN; GBIF	7 178 records including 7	2012	117 (out of 124 endemic)	Urban and agricultural	AOO [EOO not considered] (IUCN

		088 georeferenced		species)	environments (Bossard et al. 2000); Land use change 2012-2018 (Bossard et al. 2000); loss of forest cover (Hansen et al. 2013)	2012) Number of locations = 5km*5km square grid
Martinique	CBN Martinique; Herbarium MTK; Herbarium P; Dataset in Carrington et al. (2017); INPN; GBIF	4 514 records including 3640 georeferenced	1945	183 (out of 221 endemic species)	Urban and agricultural environments (Bossard et al. 2000); Land use change 2012-2018 (Bossard et al. 2000); loss of forest cover (Hansen et al. 2013)	AOO, EOO (IUCN 2012) Number of locations = 10km*10km square grid
Guadeloupe	Gwada Botanica (https://www.gwadabotanica.fr/); Herbarium GUAD; Herbarium P; Dataset in Carrington et al. (2017); INPN; GBIF	8 551 records including 7 985 georeferenced	1981	187 (out of 206 endemic species)	Urban and agricultural environments (Bossard et al. 2000); Land use change 2012-2018 (Bossard et al. 2000); loss of forest cover (Hansen et al. 2013)	AOO, EOO (IUCN 2012) Number of locations = 5km*5km square grid
French sub- Antarctic islands (Kerguelen, Crozet, Saint- Paul,	National Nature Reserve of the French sub-Antarctic Islands; Project IPEV 136 Subanteco; GBIF; Herbarium P.	12 652 records of which 12 372 geolocated	2017	16 (out of 15 endemic species + 2 patrimonial species)	Not assessed	<u>Kerguelen</u> : AOO, EOO (IUCN 2012) Number of locations = 10km*10km square

Amsterdam)						<p>grid</p> <p><u>Saint Paul, Amsterdam, Crozet</u> AOO, 100m*100m square grid, Number of locations = 100m*100m square grid.</p> <p>CR = $\leq 0.5\%$ of grid cells; EN = $\leq 3.5\%$; VU = $\leq 7.5\%$</p>
Scattered Islands (Europa, Glorioso, Juan de Nova, Tromelin)	Conservatoire Botanique National Mascarin (Hivert et al. 2018)	Hivert et al. (2018)	Hivert et al. (2018)	38 (out of 38 endemic species)	Not assessed	<p>AOO, 100m*100m square grid. CR = $\leq 0,35$ to $0,96\%$ of grid cells; EN = $\leq 2,88$ to $3,5\%$; VU = $\leq 6,73$ to $7,48\%$.</p> <p>Besides , different thresholds are used to estimate extinction risks at local and regional scales (see Hivert et al. 2018)</p> <p>Number of locations = number of islands</p>

334 * Occurrence records include preserved specimens, living specimens, observations, photographs, and material samples

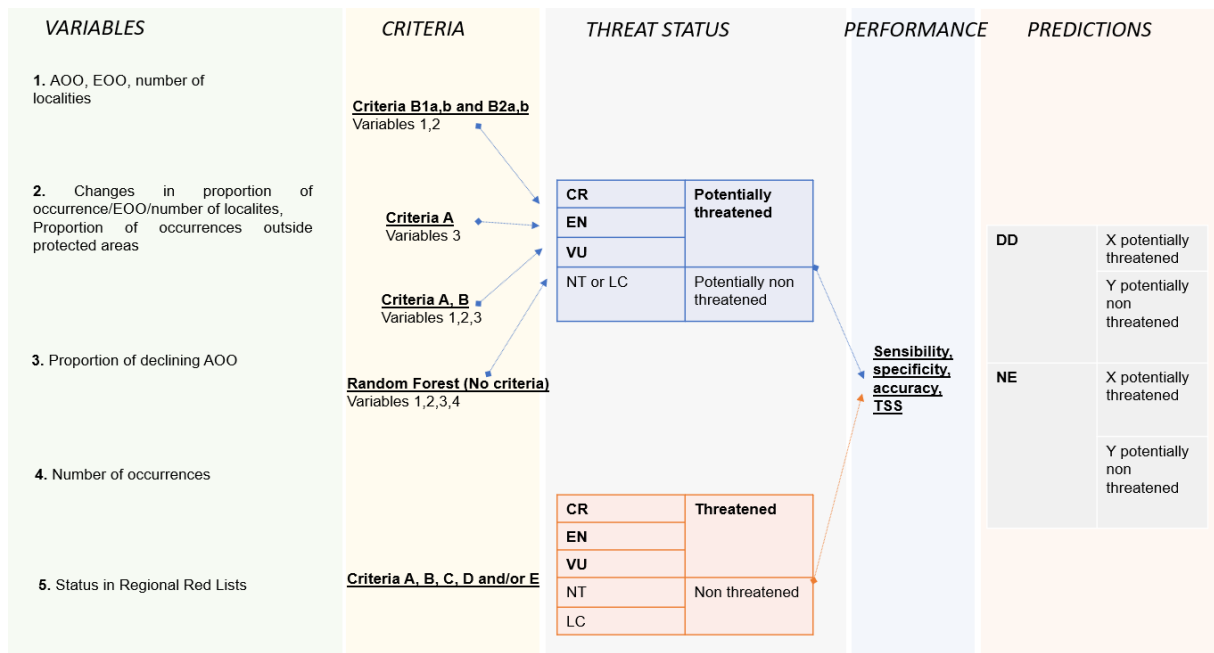
335

336 2.3. Method performance and potential changes in conservation status (objective 2)

337 Following the pre-assessment of conservation statuses, we aimed to test the effectiveness and
338 limitations of each method (Fig. 1).

339 For each method we compared existing IUCN statuses to the pre-assessed ones. We assessed
340 each status independently (Supplementary Information 5 and 6) and then classified the species
341 as potentially threatened (CR, EN, VU) or non-threatened (LC or NT). In addition, we
342 adopted a precautionary approach and considered 'likely rare' and 'potentially rare' species as
343 threatened because of the similarity of these categories to the VU status under criterion D2.
344 We then analyzed the proportion of species that 1) remained threatened after pre-assessments
345 (sensitivity) and 2) remained non-threatened (specificity) following pre-assessments.
346 Therefore, DD and NE species were excluded from the calculation of sensitivity and
347 specificity indices. A high sensitivity indicated that there was a high proportion of species
348 classified as threatened by the IUCN that were pre-assessed potentially threatened. A high
349 specificity method indicated that there was a high proportion of species classified as non-
350 threatened (LC or NT) by the IUCN that were pre-assessed non-threatened.

351 We estimated the accuracy of the pre-assessments, i.e. the proportion of species correctly
352 assessed, and the True Skill Statistic (TSS) which is equal to "sensitivity + specificity - 1".
353 TSS scores range from 1 to -1, with 1 indicating a perfect pre-assessment model, while a
354 value of 0 or lower indicates that pre-assessments are no better than if status had been pre-
355 assessed randomly. Because of our choice to consider "likely rare" and "potentially rare"
356 species as threatened, the identification of a decline or not does not influence the performance
357 of a method.



358
359 **Fig. 1** Methodological summary of the study

360

361 **3. Results**

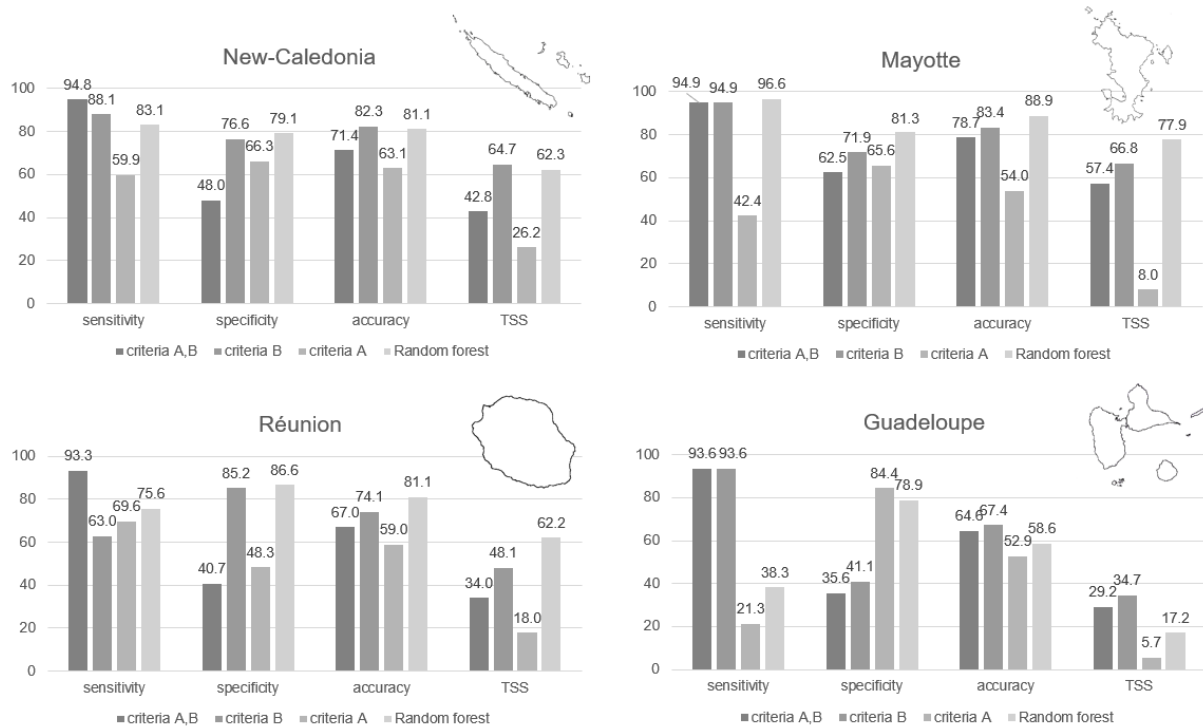
362 **3.1. Pre-assessment methods using criterion B and Random Forest performed best**

363

364 The pre-assessment method with the best TSS and accuracy was the Random Forest method
 365 for Réunion, Mayotte, and it also had a good performance in New Caledonia (Fig. 2). It gave
 366 a sensitivity of 75-96% for these territories and a specificity between 79% and 87% (Fig. 2).
 367 In contrast, the use of Random Forest for the endemic flora of Guadeloupe had a low
 368 sensitivity (38.3%) but a high specificity (78.9%). Due to the absence of Red Lists in the
 369 French sub-Antarctic islands, the low number of endemic species with a conservation status in
 370 Martinique, and the absence of pressure maps in the Scattered Islands, the Random Forest
 371 method was not applied for these 3 territories. The other machine learning method we tested
 372 (Zizka et al. 2020) also performed well but with lower accuracy than Random Forests
 373 (Supplementary Information 2). Pre-assessment of status using criteria A and B together
 374 provided the best sensitivity for New Caledonia and Réunion, but the specificity of this
 375 method was low. As for criterion B used alone, it had the best accuracy and TSS for

376 Guadeloupe and New-Caledonia and a relatively high specificity and sensitivity in the
377 Scattered Islands (local and regional scales). The high TSS in Guadeloupe was due to high
378 sensitivity whereas specificity was very low for this territory, i.e. criterion B assigned greater
379 extinction risks to Guadeloupe's endemic tracheophytes than the assessments in the 2019 Red
380 List (UICN Comité français et al. 2019). On the other hand, criterion B assigned lower risks
381 to Réunion endemic taxa compared to the 2013 Red List (UICN France et al. 2013). In order
382 to improve the performance of methods and take into account the differences in surface area
383 between territories, different grid cell sizes and thresholds were used for the calculation of the
384 AOO and the number of locations in Mayotte, the French sub-Antarctic Islands, Guadeloupe,
385 Martinique and the Scattered Islands (Table 1, Supplementary Information 2). The use of
386 criterion A alone, obtained with different pressure layers (Table 1), gave the lowest TSS and
387 accuracy for all the territories.

388 Finally, the number of specimens significantly distinguished LC species, which had a large
389 number of observations, from threatened categories (CR, EN, VU) which had a low number
390 of observations, in all territories.



391 **Fig. 2** Accuracy, TSS, sensitivity and specificity of the four pre-assessment methods for four
 392 FOTs: Guadeloupe, Mayotte, Réunion, and New Caledonia. Methods are represented from
 393 dark to light grey: criteria A,B; criteria B; criteria A; Random Forest. For graphical reason the
 394 TSS score was multiplied by 100. Performance analysis could not be applied to territories
 395 where no or few endemic taxa were assessed with a IUCN conservation status.
 396

397

398 3.2 Based on the best performing methods, up to sixty percent of endemic species are
 399 potentially threatened

400 We used each method in each territory but chose to present results only for the best
 401 performing ones (all results are in Supplementary Information 5). Using the most appropriate
 402 pre-assessment method for each territory, 2 135 taxa (60.0%) were found to be threatened
 403 (Table 2; Supplementary Information 5 and 6). In Réunion and Mayotte, the method with the
 404 best precision and TSS, i.e. the Random Forest algorithm, indicated that in total 262 taxa
 405 endemic to these territories would be potentially threatened and 273 would be non-threatened
 406 (Table 2). Adding the criterion B results obtained in territories where it performed the best
 407 (New-Caledonia, Guadeloupe), or where the Random Forest algorithm could not be applied
 408 (Martinique, Scattered islands, French sub-Antarctic islands), we pre-assessed an additional 1
 409 873 as potentially threatened and 1 147 as non-threatened. However, the number of threatened

410 taxa estimated with criterion B is probably a high estimate due to the low specificity of this
411 method in Guadeloupe (i.e. many NT or LC species have been pre-assessed as threatened) but
412 also to the lower specificity of Criterion B in New Caledonia compared to the Random Forest
413 method. When the Random Forest method, which obtains good precision and TSS in all
414 territories, is used in New Caledonia and Guadeloupe, the total proportion of threatened
415 endemic taxa is 56.4%.

		<i>Spermatophytes</i>		<i>Pteridophytes</i>		
	<i>Pre-assessment Method</i>	<i>Strict endemics</i>	<i>Regional endemics</i>	<i>Strict endemics</i>	<i>Regional endemics</i>	<i>All taxa</i>
New-Caledonia	Criterion B	1495 (60.1%)	0	53 (51.9%)	0	1548 (59.7%)
Réunion	Random Forest	119 (49.5%)	45 (35.1%)	5 (26.3%)	12 (46.1%)	181 (43.8%)
Mayotte	Random Forest	41 (75.5%)	38 (62.2%)	1 (100%)	1 (100%)	81 (69.2%)
Guadeloupe	Criterion B	18 (90%)	99 (71.2%)	1 (100%)	24 (88.8%)	142 (75.9%)
Martinique	Criterion B	25 (86.2%)	121 (81.3%)	0	17 (100%)	163 (89.1%)
Scattered Islands	Criterion B, adapted as in [Hivert et al. 2018], local scale)	2 (25%)	15 (39.4%)	0	0	17 (36.9%)
French sub-Antarctic islands	<u>Kerguelen:</u> Criterion B, <u>Crozet:</u> <u>Amsterdam-St Paul:</u> Criterion B adapted as in [Hivert et al. 2018], local scale)	1 (11.1%)	1*(20%)	0	1 (50%)	3 (16.6%)
TOTAL		1701 (59.6%)	319 (62.3%)	60 (48.3%)	55 (73.9%)	2135 (60.0%**)

416 **Table 2:** Number of species pre-assessed as potentially threatened using the "Random Forest"
417 (Réunion, Mayotte), criterion B (New Caledonia, Martinique, Guadeloupe, Kerguelen),
418 criteria A+B (Martinique) or criterion B based on Hivert et al. (2018) (Supplementary
419 Information 3: Scattered Islands, Crozet and Amsterdam-Saint-Paul). Percentages of the total
420 taxa in these categories are given in parentheses. **Colobanthus kerguelensis* is potentially VU
421 in Crozet but LC in Kerguelen.** A few endemic taxa of the Scattered islands and French

422 sub-Antarctic islands were counted several times as they have a conservation status in several
423 islands of these territories.

424

425 **4. Discussion**

426 4.1. How many endemic species are potentially threatened and why?

427 Using pre-assessment methods with the highest performance for each territory we estimated
428 that up to 2 135 endemic taxa are potentially threatened and 1 420 are potentially non-
429 threatened for seven FOTs. The proportion of potentially threatened endemic tracheophytes is
430 thus 60.0%, which is higher than recent estimates based on published Red Lists (51% for
431 spermatophytes and 37% for pteridophytes; Véron et al. 2021). Of these, 91 of the 138
432 previously categorized DD taxa and 1 160 of the 1 745 previously categorized NE taxa are
433 potentially threatened. This shows that a high proportion of DD species may be potentially
434 threatened as found in mammals (Bland et al. 2015), reptiles (Bland and Böhm 2016) or
435 sharks and rays (Walls and Dulvy 2020). However, many taxa have few records, e.g. 443 taxa
436 have three or fewer records, and further studies, local expertise or inventories would be
437 necessary to understand whether this represents a genuine extinction risk or a lack of
438 knowledge. Moreover, it must be considered that not all factors leading to the DD
439 classification of a taxon can be taken into account by the pre-assessment methods, e.g.
440 uncertain occurrences, taxonomy, misidentification, nomenclature (e.g. lost or uncertain type
441 specimen), unknown provenance (Hochkirch et al. 2021). For these reasons we did not define
442 a DD pre-status category in this study and suggest that further knowledge is necessary for
443 some taxa assessed as DD in a Red List to evaluate their conservation status. Therefore,
444 identifying the causes of data-deficiency will facilitate the use of pre-assessment methods to
445 predict the extinction risk of these species, include them in biodiversity indices and species
446 conservation status analyses, and avoid unnoticed extinctions (Bland et al. 2017). Besides, we
447 acknowledge that in New-Caledonia and Guadeloupe the number of potentially threatened

448 taxa may have been over-estimated (the method with the highest accuracy had a relatively low
449 specificity in these territories, see section 4.3.2), and using a method with a higher specificity
450 in these territories reduced the proportion of threatened taxa in the FOTs to 56.4%.

451 Based on the methods with the highest accuracy and TSS, Martinique has the highest
452 proportion of potentially threatened taxa but method performance could not be assessed in this
453 case and there is a lack of occurrence data for many taxa (pers. comm. G.Viscardi).
454 Guadeloupe also has a high proportion of potentially threatened endemic tracheophytes,
455 however risks may have been overestimated in this territory as the specificity of the method is
456 low. This is likely due to the lack of comprehensive occurrence data for taxa previously
457 classified as non-threatened and to their natural range-restriction. Excluding Guadeloupe and
458 Martinique, Mayotte is the territory with the highest proportion of potentially threatened
459 endemic species due to the narrow range of endemic species there and high levels of threats,
460 especially agriculture and deforestation. In Réunion, plants occurring in areas with high alien
461 plant invasion pressure are expected to be at risk (Fenouillas et al. 2021). Pressure from
462 invasive species is likely a main threat to endemic plants in all FOTs due to competition for
463 resources (alien plants) and grazing by introduced animals (e.g. cattle, rabbits) (Soubeyran et
464 al. 2015, but see Caujapé-Castells et al. 2010). Thereby, endemic plants from the Scattered
465 islands are threatened by invasions (plants, goats), cyclones (for woody species) and the small
466 size of their populations make them highly vulnerable to extreme climatic events and/or sea
467 level rise. Yet, reduced human activity, flora monitoring and conservation actions explain
468 why the proportion of threatened endemic plant species is lower in these islands compared to
469 other FOTs. In New-Caledonia other important pressures on plant diversity are fires and
470 mining activities (Meyer et al. 2021). Moreover, it is important to consider the interactions
471 between these threats, which are expected to have a greater impact than taken individually
472 (Krupnick 2013; Leclerc et al. 2018; Rojas- Sandoval et al. 2020). For example, in the French

473 sub-Antarctic islands, recent climatic changes, the drop in precipitation, combined with
474 warmer temperatures, have led to the decline of certain native species, leaving bare soil,
475 which has benefited the most dispersive invasive species and changed the structure of
476 communities (Chapuis et al. 2004; Robin et al. 2011).

477 4.2. What are the implications of pre-assessments for biodiversity conservation?

478 Pre-assessments are not intended to replace Red Lists, as it is essential to follow the
479 methodology developed by the IUCN to assess the genuine conservation status of species in a
480 standardized and reproducible way. In particular, we acknowledge that the variables used do
481 not rigorously match the criteria of the IUCN Red List, although they try to conform to the
482 criteria as much as possible and are all related to extinction risks (see 4.3.2). Yet our aim was
483 to automatically pre-assess the potential extinction risks of species, and many criteria and sub-
484 criteria are difficult (if not impossible) to evaluate without the knowledge of experts. We see
485 several benefits of the pre-assessment methods used in this study to assist Red Lists. First,
486 these methods can help to prioritize first-time assessments (e.g., many taxa in Martinique,
487 Kerguelen) and re-assessments (e.g., Réunion, Mayotte). Especially, automated methods can
488 be used during the pre-evaluation step of the IUCN Red List process. They give an indication
489 of which taxa are most likely to be threatened and which should attract closer attention during
490 IUCN assessment workshops when resources and time are limited. They also allow the
491 identification of species which are likely not threatened (Bachman et al. 2020), especially by
492 using pre-assessment methods with high specificity, and which need less data compared to
493 threatened species to be published on the Red List (IUCN 2013). This allows the reduction of
494 resources spent on predicted-to-be non-threatened taxa. Second, the variables used in
495 automatic pre-assessments can contribute to evaluate the Red List criteria, for example
496 calculating the proportion of the range size of a taxon that intersects newly degraded areas can
497 help to estimate a decline in species habitat quality. Third, pre-assessments can help prioritize

498 data collection, for example by identifying data-deficient species that are likely to be
499 threatened and may require further knowledge to validate their conservation status (Cazalis et
500 al. 2022). From the uses presented above, pre-assessment methods can support the
501 development of Red Lists by facilitating the evaluation of the conservation status of more
502 taxa, allowing one to cover more taxonomic groups, geographical regions and more up to date
503 assessments. Furthermore, this will support the use of Red Lists as a powerful tool to evaluate
504 the state of biodiversity, to prioritize species in conservation planning, to support conservation
505 policies such as the Convention on Biological Diversity and to raise awareness of biodiversity
506 loss. We therefore agree with Cazalis et al. (2022) who advocated that pre-assessment
507 methods should be better used in IUCN Red Lists. Yet, as stated above, it is essential to
508 follow the methodology developed by the IUCN to assess the genuine conservation status of
509 species and it is not recommended to use pre-assessment methods alone to implement
510 conservation actions. We therefore believe that the main aim of the pre-assessment methods is
511 to facilitate the development of Red Lists which will further contribute to conservation actions
512 and policies. Finally, researchers in biodiversity conservation often need information on
513 conservation status that are not available yet for many taxonomic groups and geographic
514 places. Therefore, using automatic methods can help to approach species conservation
515 statuses with a high confidence and include the results in research studies.

516

517 4.3. Quality and shortcomings of pre-assessment methods and the necessity to adapt them to 518 the studied territory

519

520 4.3.1. Data quality

521 This pre-assessment of the conservation status of the flora of the FOTs benefited from a first
522 study on the taxonomy and endemic status of tracheophytes and bryophytes, a necessary step
523 for this new study (Véron et al. 2021). In addition, recent field collections have allowed us to
524 precisely geolocate endemic species in the FOTs. Yet, there are still taxa with imprecise
525 coordinates, especially in Mayotte and Guadeloupe, which may require further inventories by
526 botanical experts. We also compiled herbarium data that can be used to determine the change
527 of a species' range over time, to guide new collections, to revise taxonomy, to discover new
528 species or to document extinctions. Information on the rarity or commonness of the taxa is
529 also sometimes directly written on the herbarium sheet. More importantly, herbarium
530 specimens allow reproducibility (the base of the scientific method) as only these permanent
531 collections allow for the reassessment of the identification of taxa at any time from the initial
532 observation in the field. Herbarium data thus provide a historical context of the floristic
533 diversity of a territory and its evolution. For example, in Réunion, the use of herbarium data
534 alone informs us that many collections have been made in places that are now degraded.
535 Many studies have thus recommended using herbarium data for conservation purposes (Rivers
536 et al. 2010, Rivers et al. 2011, Nic Lughadha et al. 2019, Albani Rocchetti et al. 2021).

537

538 4.3.2. Performance by criteria, limitations and improvements

539 Criterion B based on AOO, EOO, number of locations, is relatively easy to automatically
540 measure for pre-assessing IUCN conservation status (e.g. Dauby et al. 2017; Moat 2017). The
541 size of a species' range is strongly correlated with its extinction status (Bland et al. 2015), and
542 we have shown the importance of range and population structure variables on the
543 classification of species into one of the Red List threat categories (Supplementary Information
544 4). We acknowledge that rarity and conservation status are two different concepts, as a
545 narrow-range species may have stable populations and be secure while a widespread species

546 could be threatened on a high proportion of its range and at risk of extinction. Yet, narrow
547 range-size is an important indication of extinction risk. In the “largest and statistically better
548 supported review of the relation be-
549 tween traits and extinction risk to date”, Chichorro et al. (2019) showed that species with
550 small range and narrow habitat breadth were the most vulnerable to extinctions. Previous pre-
551 assessments also showed that species pre-assessed with Criterion B-based methods had a high
552 probability of matching their IUCN Red List status (Nic Lughadha et al. 2019; Zizka et al.
553 2020). In line with these results, we found that for the majority of territories the simplified use
554 of criterion B had a high accuracy and TSS, especially in New Caledonia. The high
555 performance of criterion B and its tight relationship with extinction risk means that species
556 pre-assessed as potentially threatened require specific attention during IUCN workshops and
557 that only few of them are expected to be secure in these territories. On the contrary, criterion
558 B specificity was low in Guadeloupe while sensitivity was high, meaning that pre-assessed
559 conservation statuses have likely been overestimated. This indicates that it is important to
560 look at all performance measures to fully assess the reliability of pre-assessment methods.
561 Interestingly, criterion B was not the most frequently used criterion in existing Red Lists (with
562 the exception of New Caledonia) which have often used criterion D and sometimes criteria C
563 and A (Supplementary Information 1). This shows that a simplified version of criterion B
564 assessment can perform well to assess whether a taxon is potentially threatened or not, even if
565 it was classified in a Red List based on another criterion. This could be explained by the
566 expected positive relationship between AOO and the number of individuals (criterion C) and
567 of the use of AOO in criterion D2. However, Walker et al. (2022)) showed that a simplified
568 criterion B method may not perform as well in pre-assessing the conservation status of species
569 previously classified in a Red List based on criterion A.

570 A limitation of this range criterion is, however, that the threshold values of EOO and AOO
571 defined by the IUCN (2012) are not always adapted to the small size of some territories,
572 leading to an overestimation of the extinction risk (Martín 2009). The IUCN therefore does
573 not recommend using the Red List criteria in very small territories. Consequently, for the
574 Scattered Islands and French sub-Antarctic islands (except Kerguelen), where assessing
575 conservation status is an important conservation challenge to respond to management
576 objectives (Hivert et al. 2018, see also Molloy et al. 2002; Martín 2009), it was necessary to
577 adapt the size of the grids used to calculate the AOO, but also the conservation status
578 thresholds defined by the IUCN. Such adaptations are not recommended in the IUCN Red
579 List but in the context of pre-assessments they improve the performance of methods using
580 criterion B for small territories.

581 Estimations of population decline (here comprising decline in EOO or number of individuals)
582 are of moderate importance in classifying taxa as threatened. Nevertheless, the observation of
583 a decrease in the number of occurrences or the range of a species or a low degree of
584 protection may indicate that special attention should be paid to the conservation status of the
585 species. Indeed, biodiversity losses cannot only be assessed in terms of extinction, and
586 population decline is another indicator to analyze the impact of anthropogenic activities (e.g.
587 Bissessur et al. 2017). This is particularly the case for strict island endemics where there are
588 no populations outside the territory to replenish the island population. Furthermore, in the
589 case of regional endemics, this reinforcement may be made difficult by the need to cross the
590 ocean, which acts as a barrier to dispersal for many species. There are also species for which
591 there is no estimated decline following the method of Carrington et al. (2017) but for which
592 the area of occurrence or number of occurrences is very low. These species have been
593 classified as 'likely rare' or "potentially rare" which can also correspond to a VU status under
594 criterion D2. These species are not currently endangered but future pressures could strongly

595 impact them and quickly move them to EX, CR or EN. Due to these limitations in estimating
596 the number of locations and a continuous decline, we acknowledge that some of the species
597 pre-assessed as potentially threatened under criterion B may not be at risk. However, as this
598 method has a high accuracy we expect that such species are few and that the criterion B
599 method remains useful to identify species potentially at risk. Improving on the automatic
600 measure of the number of locations and population decline is a future challenge in the
601 development of pre-assessment methods.

602 The criterion A pre-assessment method makes it possible to estimate the proportion of a
603 species' distribution area affected by anthropogenic pressures. Yet, it assumes that threats are
604 uniform across the species range and are fixed in time and space. The strength of the criterion
605 A assessment according to Stévant et al. (2019) is that it greatly increases the sensitivity of
606 pre-assessments when combined with criterion B. The joint use of criteria A and B therefore
607 helps to identify species that are potentially at risk because they are range-restricted and/or
608 found in a degraded area. In addition, habitat degradation variables are relatively important in
609 the classification of a taxon in one of the Red List threat categories. The available pressure
610 maps allow a first assessment of the severity of the threat (Fenouillas et al. 2021), whereas
611 criterion B focuses on geographical criteria. However, the severity of anthropogenic pressures
612 could vary greatly depending on the taxa. For example, some species can be found in urban
613 environments – such as *Clusia major* in Guadeloupe or *Aloe mayottensis* in Mayotte – or in
614 areas where fires occur (depending on fire intensity). On the contrary, some taxa could also be
615 threatened despite their expected resilience to the anthropization of habitats (Bissessur et al.
616 2017) or when they are on the periphery of a degraded area. Here we considered that a taxon
617 must have been observed within a currently degraded area for it to be potentially impacted,
618 therefore the conservation status of taxa found close to degraded areas may have been

619 underestimated. In Mayotte, for example, many occurrences are found close to deforested
620 areas but have not been considered as impacted, which underestimates the risk they face.

621 The Random Forest algorithm provides the best accuracy, specificity and TSS in Réunion and
622 Mayotte and also performs well in New-Caledonia. With the exception of Guadeloupe, this
623 method also has a good sensitivity, and is therefore particularly well adapted to the pre-
624 assessment of NE species. We also tested different parameters of the algorithm in order to
625 make the method as reliable as possible. One of the difficulties in using classification trees is
626 the complexity of interpreting the results and the choice of model parameters, although recent
627 publications have improved our understanding and use of these models in the context of pre-
628 assessments (Walker et al. 2022). The application of these models also requires that a
629 conservation assessment has taken place in the past and that enough data are available to train
630 the model. In addition, we used range, degraded habitat, and protected area variables, but
631 considering other factors such as climate, habitat, could improve these methods (Walker et al.
632 2022; but see Bland et al. 2015). In particular, the impact of climate change has not been
633 considered: while many studies focus on changes in phenology or range, the impacts of
634 climate change on population survival are still poorly known (Albani Rocchetti et al. 2021).
635 In small territories, such as the Scattered Islands, sea level rise due to climate change could
636 lead endemic species to extinction. Besides, future climate modeling can greatly help us
637 understand how species ranges, and their future conservation statuses, are expected to change
638 (Dubos et al. 2021).

639

640 **5. Conclusion**

641 In this study we showed the value of pre-assessment methods to estimate potential
642 conservation statuses and support Red Lists. Especially, Random Forest and criterion B pre-

643 assessments methods were generally successful in classifying species as threatened or non-
644 threatened, with up to 88% accuracy. Therefore, we support the idea that automatic pre-
645 assessment methods can support IUCN Red Lists to fasten and prioritize assessments by
646 identifying predicted-to-be threatened taxa. We suggest that the choice of the pre-assessment
647 method and its parameters should be adapted to the characteristics of each territory and
648 especially its size in order to fully consider conservation goals and achieve the best possible
649 performance. For example, changes in pre conservation status thresholds in very small
650 territories as in the Scattered Islands or some of the French sub-Antarctic islands could help to
651 better tackle conservation challenges. Based on the best performing pre-assessment method
652 for each FOT we estimated that up to 60% of tracheophytes endemic to the FOTs could be
653 threatened. The future assessments using the standardized IUCN Red List methodology could
654 focus first on the species pre-assessed as threatened in this study. Once the species confirmed
655 as threatened are identified we recommend that conservation measures be put in place as soon
656 as possible to protect these threatened species.

657

658

659 **References**

- 660 Albani Rocchetti G, Armstrong CG, Abeli T, et al (2021) Reversing extinction trends: new uses of (old)
661 herbarium specimens to accelerate conservation action on threatened species. *New Phytol*
662 230:433–450. <https://doi.org/10.1111/nph.17133>
- 663 Bachman S, Walker B, Barrios S, et al (2020) Rapid Least Concern: towards automating Red List
664 assessments. *BDJ* 8:e47018. <https://doi.org/10.3897/BDJ.8.e47018>
- 665 Bachman SP, Field R, Reader T, et al (2019) Progress, challenges and opportunities for Red Listing.
666 *Biological Conservation* 234:45–55. <https://doi.org/10.1016/j.biocon.2019.03.002>
- 667 Bar-On YM, Phillips R, Milo R (2018) The biomass distribution on Earth. *Proc Natl Acad Sci USA*
668 115:6506–6511. <https://doi.org/10.1073/pnas.1711842115>
- 669 Bissessur P, Baider C, Florens FBV (2017) Rapid population decline of an endemic oceanic island plant
670 despite resilience to extensive habitat destruction and occurrence within protected areas.
671 *Plant Ecology & Diversity* 10:293–302. <https://doi.org/10.1080/17550874.2017.1402382>

- 672 Bland LM, Bielby J, Kearney S, et al (2017) Toward reassessing data-deficient species. *Conservation*
673 *Biology* 31:531–539. <https://doi.org/10.1111/cobi.12850>
- 674 Bland LM, Böhm M (2016) Overcoming data deficiency in reptiles. *Biological Conservation* 204:16–22.
675 <https://doi.org/10.1016/j.biocon.2016.05.018>
- 676 Bland LM, Collen B, Orme CDL, Bielby J (2015) Predicting the conservation status of data-deficient
677 species: Predicting Extinction Risk. *Conservation Biology* 29:250–259.
678 <https://doi.org/10.1111/cobi.12372>
- 679 Bossard M, Feranec J, Otahel J (2000) CORINE land cover technical guide: Addendum 2000. European
680 Environment Agency, Copenhagen
- 681 Breiman L (2001) Breiman and Cutler’s Random Forests for Classification and Regression. Package
682 “randomForest”. *Machine Learning* 45:5–32. <https://doi.org/10.1023/A:1010933404324>
- 683 Brummitt NA, Bachman SP, Griffiths-Lee J, et al (2015) Green Plants in the Red: A Baseline Global
684 Assessment for the IUCN Sampled Red List Index for Plants. *PLoS ONE* 10:e0135152.
685 <https://doi.org/10.1371/journal.pone.0135152>
- 686 Carrington CM, Edwards RD, Krupnick GA (2018) Assessment of the Distribution of Seed Plants
687 Endemic to the Lesser Antilles in Terms of Habitat, Elevation, and Conservation Status.
688 *Caribbean Naturalist* 2:30–47
- 689 Carrington CMS, Krupnick GA, Acevedo-Rodríguez P (2017) Herbarium-Based Preliminary
690 Conservation Assessments of Lesser Antillean Endemic Seed Plants Reveal a Flora at Risk. *Bot*
691 *Rev* 83:107–151. <https://doi.org/10.1007/s12229-017-9182-5>
- 692 Caujapé-Castells J, Tye A, Crawford DJ, et al (2010) Conservation of oceanic island floras: Present and
693 future global challenges. *Perspectives in Plant Ecology, Evolution and Systematics* 12:107–
694 129. <https://doi.org/10.1016/j.ppees.2009.10.001>
- 695 Cazalis V, Di Marco M, Butchart SHM, et al (2022) Bridging the research-implementation gap in IUCN
696 Red List assessments. *Trends in Ecology & Evolution* S0169534721003372.
697 <https://doi.org/10.1016/j.tree.2021.12.002>
- 698 Chapuis J-L, Frenot Y, Lebouvier M (2004) Recovery of native plant communities after eradication of
699 rabbits from the subantarctic Kerguelen Islands, and influence of climate change. *Biological*
700 *Conservation* 117:167–179. [https://doi.org/10.1016/S0006-3207\(03\)00290-8](https://doi.org/10.1016/S0006-3207(03)00290-8)
- 701 Chichorro F, Juslén A, Cardoso P (2019) A review of the relation between species traits and extinction
702 risk. *Biological Conservation* 237:220–229. <https://doi.org/10.1016/j.biocon.2019.07.001>
- 703 Dauby G, Stévant T, Droissart V, et al (2017) *ConR*: An R package to assist large-scale multispecies
704 preliminary conservation assessments using distribution data. *Ecol Evol* 7:11292–11303.
705 <https://doi.org/10.1002/ece3.3704>
- 706 Dubos N, Montfort F, Grinand C, et al (2021) Are narrow-ranging species doomed to extinction?
707 Projected dramatic decline in future climate suitability of two highly threatened species.
708 *Perspectives in Ecology and Conservation* S2530064421000894.
709 <https://doi.org/10.1016/j.pecon.2021.10.002>

- 710 Endemia, RLA Flore NC (2019) La Liste rouge de la flore menacée de Nouvelle-Calédonie (synthèse
711 mars 2019)
- 712 Fenouillas P, Ah-Peng C, Amy E, et al (2021) Quantifying invasion degree by alien plants species in
713 Reunion Island. *Austral Ecology* aec.13048. <https://doi.org/10.1111/aec.13048>
- 714 Freiberg M, Winter M, Gentile A, et al (2020) LCVP, The Leipzig catalogue of vascular plants, a new
715 taxonomic reference list for all known vascular plants. *Sci Data* 7:1–7.
716 <https://doi.org/10.1038/s41597-020-00702-z>
- 717 Gargominy O, Tercerie S, Regnier C, et al (2021) TAXREF v15, référentiel taxonomique pour la
718 France : méthodologie, mise en oeuvre et diffusion. Muséum national d’Histoire naturelle,
719 Paris
- 720 Gouvernement de la Nouvelle-Calédonie (2014) Cartographie - occupation des sols.
721 <https://georep.nc/>
- 722 Hansen MC, Potapov PV, Moore R, et al (2013) High-Resolution Global Maps of 21st-Century Forest
723 Cover Change. *Science* 342:850–853. <https://doi.org/10.1126/science.1244693>
- 724 Hivert J, Bouillet V, Féraud J, et al (2018) Démarche d’évaluation collégiale du statut de menace
725 régionale de la flore vasculaire terrestre des îles Eparses. Rapport non publié., Conservatoire
726 Botanique national et centre permanent d’Initiatives pour l’Environnement de Mascarin. île
727 de la Réunion
- 728 Hochkirch A, Samways MJ, Gerlach J, et al (2021) A strategy for the next decade to address data
729 deficiency in neglected biodiversity. *Conservation Biology* 35:502–509.
730 <https://doi.org/10.1111/cobi.13589>
- 731 IPNI (2021) International Plant Names Index. The Royal Botanic Gardens, Kew, Harvard University
732 Herbaria & Libraries and Australian National Botanic Gardens.
- 733 IUCN (2012) IUCN Red List Categories and Criteria: Version 3.1. Second edition. Gland, Switzerland
734 and Cambridge, UK
- 735 IUCN (2013) Documentation standards and consistency checks for IUCN Red List assessments and
736 species accounts. Version 2. Adopted by the IUCN Red List Committee and IUCN SSC Steering
737 Committee
- 738 IUCN Red List version 2022-2 (2022) Summary statistics: Table 1a Last updated: 08 December 2022
- 739 Kougioumoutzis K, Kokkoris IP, Panitsa M, et al (2021) Extinction Risk Assessment of the Greek
740 Endemic Flora. *Biology* 10:195. <https://doi.org/10.3390/biology10030195>
- 741 Krupnick GA (2013) Conservation of Tropical Plant Biodiversity: What Have We Done, Where Are We
742 Going? *Biotropica* 45:693–708. <https://doi.org/10.1111/btp.12064>
- 743 Krupnick GA, Kress WJ, Wagner WL (2009) Achieving Target 2 of the Global Strategy for Plant
744 Conservation: building a preliminary assessment of vascular plant species using data from
745 herbarium specimens. *Biodivers Conserv* 18:1459–1474. <https://doi.org/10.1007/s10531-008-9494-1>
746

- 747 Leclerc C, Courchamp F, Bellard C (2018) Insular threat associations within taxa worldwide. *Sci Rep*
748 8:6393. <https://doi.org/10.1038/s41598-018-24733-0>
- 749 Mace GM, Collar NJ, Gaston KJ, et al (2008) Quantification of Extinction Risk: IUCN's System for
750 Classifying Threatened Species. *Conservation Biology* 22:1424–1442.
751 <https://doi.org/10.1111/j.1523-1739.2008.01044.x>
- 752 Martín JL (2009) Are the IUCN standard home-range thresholds for species a good indicator to
753 prioritise conservation urgency in small islands? A case study in the Canary Islands (Spain).
754 *Journal for Nature Conservation* 17:87–98. <https://doi.org/10.1016/j.jnc.2008.10.001>
- 755 Meyer S, Birnbaum P, Bruy D, et al (2021) The New Caledonia Plants RLA: bringing botanists together
756 for the conservation of the archipelago's crown jewel. *Imperiled : The Encyclopedia of*
757 *Conservation*
- 758 Miller JS, Porter-Morgan HA, Stevens H, et al (2012) Addressing target two of the Global Strategy for
759 Plant Conservation by rapidly identifying plants at risk. *Biodivers Conserv* 21:1877–1887.
760 <https://doi.org/10.1007/s10531-012-0285-3>
- 761 MNHN, OFB (2021) 2003-2021. National inventory of natural heritage (INPN).
762 <https://inpn.mnhn.fr/accueil/index?lg=en>
- 763 Moat J (2017) rCAT: Conservation Assessment Tools
- 764 Moat J, Bachman SP, Field R, Boyd DS (2018) Refining area of occupancy to address the modifiable
765 areal unit problem in ecology and conservation: Area of Occupancy. *Conservation Biology*
766 32:1278–1289. <https://doi.org/10.1111/cobi.13139>
- 767 Molloy J, Bell B, Clout M, et al (2002) Classifying species according to threat of extinction. A system
768 for NewZealand. *Threatened species occasional publication* 22:26
- 769 Myers N, Mittermeier RA, Mittermeier CG, et al (2000) Biodiversity hotspots for conservation
770 priorities. *Nature* 403:853–858
- 771 Nic Lughadha E, Bachman SP, Leão TCC, et al (2020) Extinction risk and threats to plants and fungi.
772 *Plants People Planet* 2:389–408. <https://doi.org/10.1002/ppp3.10146>
- 773 Nic Lughadha E, Walker BE, Canteiro C, et al (2019) The use and misuse of herbarium specimens in
774 evaluating plant extinction risks. *Phil Trans R Soc B* 374:20170402.
775 <https://doi.org/10.1098/rstb.2017.0402>
- 776 OEIL (2021) Observatory of environment in New-Caledonia - pressures and threats - MODIS data.
777 <https://oeil.nc/fr/page/les-outils-de-surveillance-des-incendies>
- 778 Pelletier TA, Carstens BC, Tank DC, et al (2018) Predicting plant conservation priorities on a global
779 scale. *Proc Natl Acad Sci USA* 115:13027–13032. <https://doi.org/10.1073/pnas.1804098115>
- 780 R Core Team (2020) R: A language and environment for statistical computing
- 781 Rivers MC, Bachman SP, Meagher TR, et al (2010) Subpopulations, locations and fragmentation:
782 applying IUCN red list criteria to herbarium specimen data. *Biodivers Conserv* 19:2071–2085.
783 <https://doi.org/10.1007/s10531-010-9826-9>

- 784 Rivers MC, Taylor L, Brummitt NA, et al (2011) How many herbarium specimens are needed to detect
785 threatened species? *Biological Conservation* 144:2541–2547.
786 <https://doi.org/10.1016/j.biocon.2011.07.014>
- 787 Robin M, Chapuis J-L, Lebouvier M (2011) Remote sensing of vegetation cover change in islands of
788 the Kerguelen archipelago. *Polar Biol* 34:1689–1700. <https://doi.org/10.1007/s00300-011-1069-z>
789
- 790 Rodrigues A, Pilgrim J, Lamoreux J, et al (2006) The value of the IUCN Red List for conservation.
791 *Trends in Ecology & Evolution* 21:71–76. <https://doi.org/10.1016/j.tree.2005.10.010>
- 792 Rojas-Sandoval J, Ackerman JD, Tremblay RL (2020) Island biogeography of native and alien plant
793 species: Contrasting drivers of diversity across the Lesser Antilles. *Divers Distrib* 26:1539–
794 1550. <https://doi.org/10.1111/ddi.13139>
- 795 Roskov Y, Ower G, Orrell T, et al (2019) Species 2000 & ITIS Catalogue of Life, 2019 Annual Checklist,
796 Species 2000: Naturalis. Leiden, The Netherlands
- 797 Russell JC, Kueffer C (2019) Island Biodiversity in the Anthropocene. *Annu Rev Environ Resour* 44:31–
798 60. <https://doi.org/10.1146/annurev-environ-101718-033245>
- 799 Soubeyran Y, Meyer J-Y, Lebouvier M, et al (2015) Dealing with invasive alien species in the French
800 overseas territories: results and benefits of a 7-year Initiative. *Biol Invasions* 17:545–554.
801 <https://doi.org/10.1007/s10530-014-0766-2>
- 802 Stévant T, Dauby G, Lowry PP, et al (2019) A third of the tropical African flora is potentially
803 threatened with extinction. *Sci Adv* 5:eaax9444. <https://doi.org/10.1126/sciadv.aax9444>
- 804 UICN Comité français, MNHN, CBIG (2019) La Liste rouge des espèces menacées en France – Chapitre
805 Flore vasculaire de Guadeloupe. Paris, France
- 806 UICN France (2018) Guide pratique pour la réalisation de Listes rouges régionales des espèces
807 menacées - Méthodologie de l’UICN & démarche d’élaboration. Seconde édition. Paris,
808 France
- 809 UICN France, CBNM, FCBN, MNHN (2014) La Liste rouge des espèces menacées en France - Chapitre
810 Flore vasculaire de Mayotte. Paris, France
- 811 UICN France, CBNM, FCBN, MNHN (2013) La Liste rouge des espèces menacées en France – Chapitre
812 Flore vasculaire de La Réunion. Paris, France
- 813 Véron S, Rodrigues-Vaz C, Lebreton E, et al (2021) An assessment of the endemic spermatophytes,
814 pteridophytes and bryophytes of the French Overseas Territories: towards a better
815 conservation outlook. *Biodiversity and Conservation*. <https://doi.org/10.1007/s10531-021-02186-8>
816
- 817 Walker BE, Leão TCC, Bachman SP, et al (2022) Evidence-based guidelines for automated
818 conservation assessments of plant species. *Conservation Biology*.
819 <https://doi.org/10.1111/cobi.13992>
- 820 Walls RHL, Dulvy NK (2020) Eliminating the dark matter of data deficiency by predicting the
821 conservation status of Northeast Atlantic and Mediterranean Sea sharks and rays. *Biological*
822 *Conservation* 246:108459. <https://doi.org/10.1016/j.biocon.2020.108459>

823 Zizka A, Silvestro D, Vitt P, Knight TM (2020) Automated conservation assessment of the orchid family
824 with deep learning. Conservation Biology *cobi.13616*. <https://doi.org/10.1111/cobi.13616>

825

826

827

828 *Statements and Declarations*

829 **Funding** The research was supported by the Office Français de la Biodiversité (OFB), the
830 Ministère des Outre-mer and the Muséum National d'Histoire naturelle (MNHN)

831 **Competing interests** The authors have no competing interests to declare that are relevant to
832 the content of this article.

833 **Data availability** The datasets generated during and/or analysed during the current study are
834 available from the corresponding author on reasonable request.

835 **Authors' contributions** Simon Véron, Arthur Bernard, Carlos Rodrigues-Vaz, Elise
836 Lebreton, Serge Muller, Vanessa Invernón, Germinal Rouhan, Olivier Gargominy, Sébastien
837 Leblond, Marc Pignal, Sandrine Terceirie and Guillaume Gigot contributed to the study
838 conception and design. Material preparation and data collection were performed by all
839 authors. Analyses were performed by Simon Véron, Arthur Bernard and Elise Lebreton. The
840 first draft of the manuscript was written by Simon Véron, Arthur Bernard and Elise Lebreton
841 and all authors commented on previous versions of the manuscript. All authors read and
842 approved the final manuscript.

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859 **Figure captions:**

860

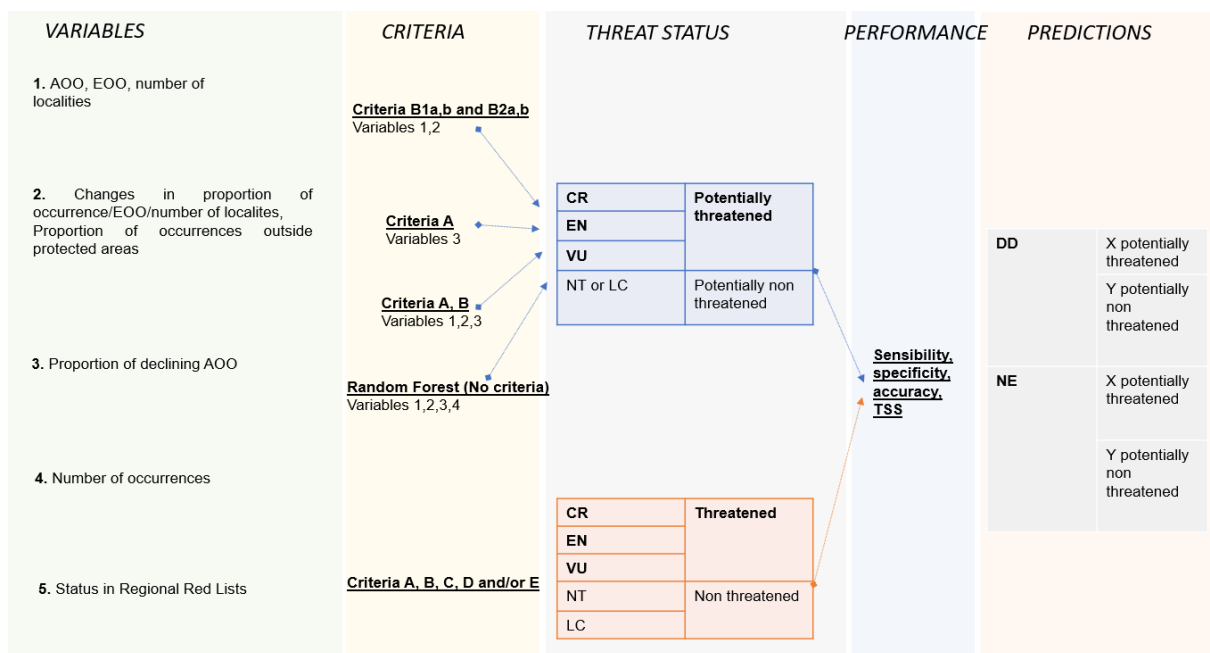
861 **Fig. 1** Methodological summary of the study

862

863 **Fig. 2** Accuracy, TSS, sensitivity and specificity of the four pre-assessment methods for four
864 FOTs: Guadeloupe, Mayotte, Réunion, and New Caledonia. Methods are represented from
865 dark to light grey: criteria A,B; criteria B; criteria A; Random Forest. For graphical reason the
866 TSS score was multiplied by 100. Performance analysis could not be applied to territories
867 where no or few endemic taxa were assessed with a IUCN conservation status.

868

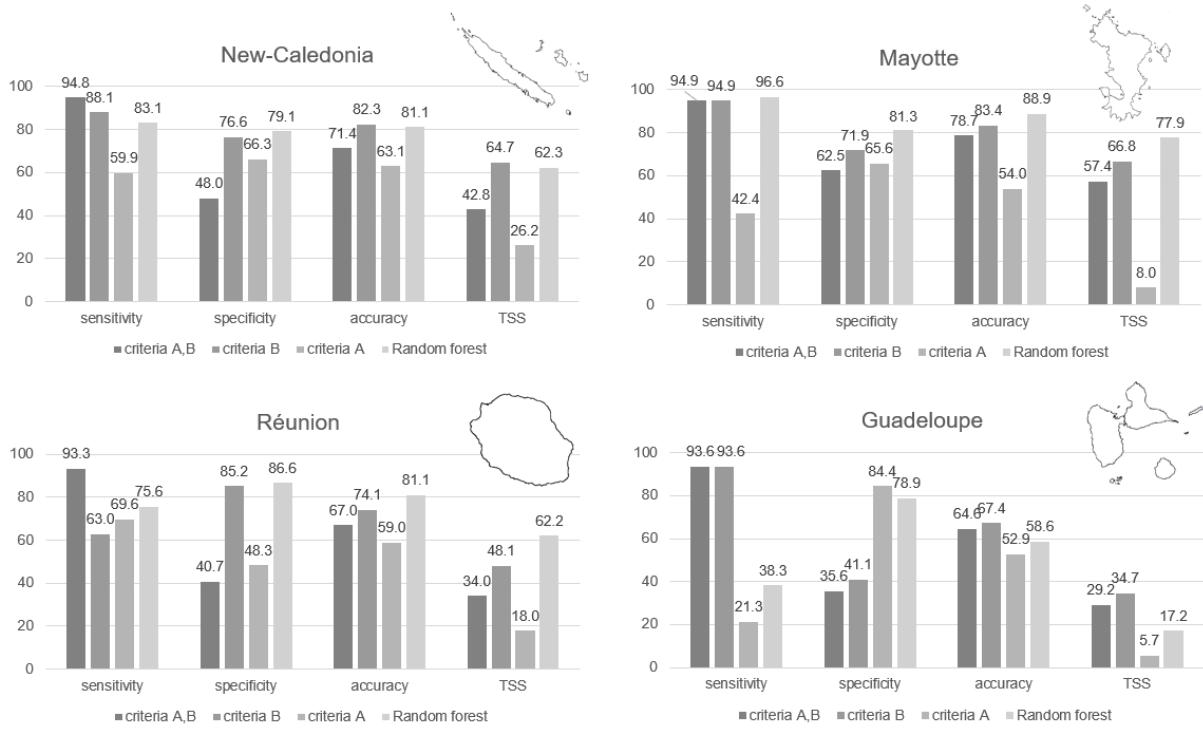
869 **Figures:**



870

871 **Fig. 1 (not colored)**

872



873

874 **Fig. 2 (not colored)**

875

876

877 **Tables:**

Table 1: Data used for each territory. Criterion B parameters vary between territories in order to reach the best performance possible for this pre-assessment method. CLC=Corinne Land Cover (Bossard et al. 2000)

	Main sources of species records	Number of occurrence records compiled	Median age of records	Pre-assessed endemic species	Anthropogenic pressures (criterion A)	Criterion B parameters
New Caledonia	Endemia and RLA Flore NC 2019; ERMINES project (for ultramafic substrates); Herbaria P and NOU; Other international herbaria; INPN; GBIF	126 329 records of which 121 657 have been georeferenced	1978	2 591 (out of 2 758 endemic species)	Urban and agricultural environments (Gouvernement de la Nouvelle-Calédonie 2014); Loss of forest cover (Hansen et al. 2013); Fires (OEIL 2021); Active mines (Gouvernement de la Nouvelle-Calédonie 2014)	AOO, EOO according to recommendations of (IUCN 2012) Number of locations = 10km*10km square grid
Réunion	Mascarine; Herbarium REU; Herbarium P; INPN; GBIF	102 406 records of which 102 208 were georeferenced	2010	413 (out of 489 endemic species)	Urban and agricultural environments (Bossard et al. 2000); Land use change 2012-2018 (Bossard et al. 2000); loss of forest cover (Hansen et al. 2013); invasive species (Fenouillas et al. 2021)	AOO, EOO (IUCN 2012) Number of locations = 10km*10km square grid
Mayotte	Mascarine; Herbarium MAO; Herbarium P; INPN; GBIF	7 178 records of which 7 088 georeferenced	2012	117 (out of 124 endemic species)	Urban and agricultural environments	AOO [EOO not considered] (IUCN 2012)

					(Bossard et al. 2000); Land use change 2012-2018 (Bossard et al. 2000); loss of forest cover (Hansen et al. 2013)	Number of locations = 5km*5km square grid
Martinique	CBN Martinique; Herbarium MTK; Herbarium P; Dataset in Carrington et al. (2017); INPN; GBIF	4 514 records including 3640 georeferenced	1945	183 (out of 221 endemic species)	Urban and agricultural environments (Bossard et al. 2000); Land use change 2012-2018 (Bossard et al. 2000); loss of forest cover (Hansen et al. 2013)	AOO, EOO (IUCN 2012) Number of locations = 10km*10km square grid
Guadeloupe	Gwada Botanica (https://www.gwadabotanica.fr/); Herbarium GUAD; Herbarium P; Dataset in Carrington et al. (2017); INPN; GBIF	8 551 records including 7 985 georeferenced	1981	187 (out of 206 endemic species)	Urban and agricultural environments (Bossard et al. 2000); Land use change 2012-2018 (Bossard et al. 2000); loss of forest cover (Hansen et al. 2013)	AOO, EOO (IUCN 2012) Number of locations = 5km*5km square grid
French sub- Antarctic islands (Kerguelen, Crozet, Saint- Paul,	National Nature Reserve of the French sub-Antarctic Islands; Project IPEV 136 Subanteco; GBIF; Herbarium P.	12 652 records of which 12 372 geolocated	2017	16 (out of 15 endemic species + 2 patrimonial species)	Not assessed	<u>Kerguelen</u> : AOO, EOO (IUCN 2012) Number of locations = 10km*10km square

Amsterdam)						<p>grid</p> <p><u>Saint Paul, Amsterdam, Crozet</u> AOO, 100m*100m square grid, Number of locations = 100m*100m square grid.</p> <p>CR = $\leq 0.5\%$ of grid cells; EN = $\leq 3.5\%$; VU = $\leq 7.5\%$</p>
Scattered Islands (Europa, Glorioso, Juan de Nova, Tromelin)	Conservatoire Botanique National Mascarin (Hivert et al. 2018)	Hivert et al. (2018)	Hivert et al. (2018)	38 (out of 38 endemic species)	Not assessed	<p>AOO, 100m*100m square grid. CR = $\leq 0,35$ to $0,96\%$ of grid cells; EN = $\leq 2,88$ to $3,5\%$; VU = $\leq 6,73$ to $7,48\%$.</p> <p>Besides , different thresholds are used to estimate extinction risks at local and regional scales (see Hivert et al. 2018)</p> <p>Number of locations = number of islands</p>

* Occurrence records include preserved specimens, living specimens, observations, photographs, and material samples

		<i>Spermatophytes</i>		<i>Pteridophytes</i>		
	<i>Pre-assessment Method</i>	<i>Strict endemics</i>	<i>Regional endemics</i>	<i>Strict endemics</i>	<i>Regional endemics</i>	<i>All taxa</i>
New-Caledonia	Criterion B	1495 (60.1%)	0	53 (51.9%)	0	1548 (59.7%)
Réunion	Random Forest	119 (49.5%)	45 (35.1%)	5 (26.3%)	12 (46.1%)	181 (43.8%)
Mayotte	Random Forest	41 (75.5%)	38 (62.2%)	1 (100%)	1 (100%)	81 (69.2%)
Guadeloupe	Criterion B	18 (90%)	99 (71.2%)	1 (100%)	24 (88.8%)	142 (75.9%)
Martinique	Criterion B	25 (86.2%)	121 (81.3%)	0	17 (100%)	163 (89.1%)
Scattered Islands	Criterion B, adapted as in [Hivert et al. 2018], local scale)	2 (25%)	15 (39.4%)	0	0	17 (36.9%)
French sub-Antarctic islands	<u>Kerguelen:</u> Criterion B, <u>Crozet:</u> <u>Amsterdam-St Paul:</u> Criterion B adapted as in [Hivert et al. 2018], local scale)	1 (11.1%)	1*(20%)	0	1 (50%)	3 (16.6%)
TOTAL		1701 (59.6%)	319 (62.3%)	60 (48.3%)	55 (73.9%)	2135 (60.0%**)

Table 2: Number of species pre-assessed as potentially threatened using the "Random Forest" (Réunion, Mayotte), criterion B (New Caledonia, Martinique, Guadeloupe, Kerguelen), criteria A+B (Martinique) or criterion B based on Hivert et al. (2018) (Supplementary Information 3: Scattered Islands, Crozet and Amsterdam-Saint-Paul). Percentages of the total taxa in these categories are given in parentheses. **Colobanthus kerguelensis* is potentially VU in Crozet but LC in Kerguelen.** A few endemic taxa of the Scattered islands and French sub-Antarctic islands were counted several times as they have a conservation status in several islands of these territories.

