How citizen science could improve Species Distribution Models and their independent assessment for conservation

Authors: Matutini Florence, Jacques Baudry, Guillaume Pain, Morgane Sineau, Joséphine Pithon

MATUTINI Florence – florence.matutini@gmail.com - BAGAP, INRAE, Institut Agro, ESA, 49000, Angers, France

BAUDRY Jacques – jacques.baudry@inrae.fr - BAGAP, INRAE, Institut Agro, ESA, 35042, Rennes, France

PAIN Guillaume – g.pain@groupe-esa.com - BAGAP, INRAE, Institut Agro, ESA, 49000, Angers, France

SINEAU Morgane - msineau@cpie72.fr – URCPIE Pays de la Loire, 44000 Nantes

PITHON Joséphine - j.pithon@groupe-esa.com - BAGAP, INRAE, Institut Agro, ESA, 49000, Angers, France

Corresponding author:

Matutini Florence, florence.matutini@gmail.com - +33 603962008 - BAGAP, INRAE Institut Agro, ESA, 55 Rue Rabelais, 49000 Angers, France

Abstract

1. Species distribution models (SDM) have been increasingly developed in recent years but their validity is questioned. Their assessment can be improved by the use of independent data but this can be difficult to obtain and prohibitive to collect. Protocoled data from citizen science may be used to establish external evaluation datasets and to improve SDM validation and applicability.
2. We used opportunistic presence-only data along with presence-absence data from a standardized citizen science program to establish and assess habitat suitability maps for 9 species of amphibian in western France. We assessed Generalized Additive and Random Forest Models' performance by (1) cross-validation using 30% of the opportunistic dataset used to calibrate the model or (2) external validation using different independent data sets derived from citizen science monitoring. We tested the effects of applying different combinations of filters to the citizen data and of complementing it with additional protocoled fieldwork.
3. Cross-validation resulted in higher AUC (Area Under the receiver operating Curve) than external evaluation causing overassessment and did not select the same models; models integrating sampling effort performed better with external validation. AUC, specificity and sensitivity of models calculated with different filtered external datasets differed for some species. However, for most species, complementary fieldwork was not necessary to obtain coherent results, as long as the citizen science data was strongly filtered.
4. Since external validation methods using independent data are considered more robust, filtering data from citizen sciences may make a valuable contribution to the assessment of SDM. Limited complementary fieldwork with volunteer's participation to complete ecological gradients may also possibly enhance citizen involvement and lead to better use of SDM in decision processes for nature conservation.

Keywords

Amphibians, citizen science, data filtering, data culling, external evaluation, habitat suitability modelling, sampling effort, species distribution modelling.

Introduction

In the current context of biodiversity loss, a stronger relationship between conservation science and citizen participation could help to make conservation actions more effective (Forrester et al., 2017; Lewandowski & Oberhauser, 2017). Availability of data from citizen sciences has considerably increased over the past few decades (Dickinson, Zuckerberg & Bonter, 2010; McKinley et al., 2017). This data has great potential because (i) large quantities of data can be collected over large areas, which would be difficult and expensive for researchers to collect; (ii) data may be collected over long time periods, which is especially useful for studying the effects of climate and landscape changes on population dynamics at large scales; (iii) citizens are involved in the research process, thereby gaining knowledge, and their involvement might lead to improved implementation of biodiversity conservation action (Dickinson, Zuckerberg & Bonter, 2010; McKinley et al., 2017). However, quality of data from participatory sciences is heterogeneous and different methods have been developed to boost data accuracy and account for bias, including interactive project development, volunteer training, expert data validation and statistical modelling improvement (Kosmala et al., 2016). Although researchers have been skeptical about the value of datasets from citizen science, recent publications show that some could be as valid as data collected by professional scientists (Kosmala et al., 2016). This is conditional on such data being judged in context (i.e. according to the sampling methods used, program objectives and applications) on the use of rigorous data sorting and analyses (Isaac et al., 2014; Steen, Elphick & Tingley, 2019; Robinson et al., 2020).

Opportunistic presence-only data collected by citizens at large scales have contributed to the expansion of species distribution models (SDM) over the past twenty years, particularly for biological conservation applications (Guisan & Thuiller, 2005). The validity of presence-only SDM is however increasingly questioned as well as their applicability (Barve et al., 2011). Presence-only data are abundant but have poor quality, few metadata and come from different sources (Robinson et al., 2020). This introduces numerous sources of bias that need to be assessed and accounted for in modelling processes (Phillips et al., 2009, Guillera-Arroita et al., 2015). Common problems are heterogeneous sampling effort, conditions and methods, imprecise spatial and temporal resolutions and different levels of expertise among observers (Schulman, Toivonen & Ruokolainen, 2007; Phillips et al., 2009, Dickinson, Zuckerberg & Bonter, 2010; McKinley et al., 2017). Different methods have been developed to correct these biases, including sorting or weighting presence-data to reduce identification errors and pseudo-replication linked to sampling effort (Guisan & Theurillat, 2000; Phillips et al., 2009) and/or using sampling effort assumptions to establish pseudo-absence sampling strategies. In order to overcome the problem of missing absence data needed for most SDM, pseudo-absence selection strategies have been developed to select absence data where real absence is most likely (Barbet-Massin et al., 2012, Phillips et al., 2009). Understanding the structure and intensity of sampling effort in space is essential to determine whether an undetected species is truly absent. For example, it may be conditioned by site accessibility (Kadmon, Farber & Danin, 2004; Phillips et al., 2009), site attractiveness or observer distribution (Phillips et al., 2009; Robinson, Ruiz-Gutierrez & Fink, 2018). Not accounting for heterogeneous sampling effort or using erroneous assumptions to define it can lead to over-assessment of model accuracy and/or false interpretation (Schulman, Toivonen & Ruokolainen, 2007; Phillips et al., 2009; Guillera-Arroita et al., 2015).

SDM validation is challenging (Vaughan & Ormerod, 2005) but is a crucial step for applying results to conservation objectives. There is still debate about SDM validity, especially when presence-only data is used to calibrate models (Barve et al., 2011). Using data with the same spatial bias to calibrate and assess a model tends to over-estimate prediction accuracy, by modelling observation

processes more than ecological processes therefore producing erroneous results. Currently, testing model accuracy with a fully independent dataset is considered the most robust method for assessing SDM (Araujo et al., 2005; Guisan, Thuiller & Zimmermann, 2017). However, obtaining an external data set for large-scale studies is often cost prohibitive and exploiting data from protocoled citizen science programmes may in some cases provide the solution. For example, Robinson et al. 2020 have shown that using filtered large-scale citizen science data for SDM calibration can improve model accuracy. Alternatively, detection-nondetection data from more standardized citizen sciences programs which are rarer than opportunistic data but have higher quality could provide presence-absence sets for external validation of presence-only SDM. In addition, using presence-only and presence-absence data at different stages of the modelling process could be a method for combining different data sets with heterogeneous quality which is a current challenge to improve SDM validity (Zipkin & Saunders, 2018; Robinson et al., 2020).

Amphibians are among the most threatened taxa in the world with rapid and widespread population declines (Stuart et al., 2004). They are particularly sensitive to fragmentation and habitat loss (Cushman, 2006) because they need different resources during their life cycle involving movements (seasonal migration and dispersion) between aquatic sites (usually ponds) and terrestrial areas (Sinsch, 1990; Cushman, 2006). Many citizen science programmes have been initiated for monitoring amphibian species (e.g. De Solla et al., 2005; Schmeller *et al.*, 2008) and data collected have been used in some conservation studies to describe population trends (Petrovan & Schmidt, 2016), road effects (Cosentino et al., 2014), climate change (Préau et al., 2019) and large-scale species distributions (Brown et al., 2016). Despite abundant literature on amphibian ecology and the availability of several citizen science databases, few studies have attempted predictive amphibian distribution models at large scales (Graham & Hijmans, 2006; Brown et al., 2016; Préau et al., 2019). Therefore, amphibian data could be suitable for testing the capacity of different types of citizen data (presence-absence or opportunist) for calibrating and assessing SDM.

Here we compare the predictive performance of presence-only SDM for nine amphibian species using different types of data (*internal* presence-only or *external* presence-absence) from citizen science programs for their assessment. We also test the opportunity to use filtered standardized citizen science data to constitute the independent data set for external evaluation. We hypothesized that (1) the type of data used for validation (internal or external) would influence the assessment of model accuracy; (2) standardized citizen science datasets might be used as independent data for external evaluation of SDM using data filters and/or complementary fieldwork.

Materials and Methods

1.1. Study area

Our study was performed in Pays de la Loire (western France), a region covering 32 082 km² with low relief and bordering on the Atlantic Ocean to the west. The region has an extensive hydrographic network organized around the River Loire and its tributaries, influencing local climate and landscape configuration. Agricultural landscapes dominate the region and traditional hedgerow network landscapes associated with extensive livestock farming are recognized for their conservation value. Such mosaics of small pastures delimited by hedgerows and small woods and generally associated with dense pond systems (Baudry, Bunce & Burel, 2000) are favourable for many organisms including endangered species such as amphibians (Boissinot, Besnard & Lourdais, 2019). With 21 known species (for 43 species recorded in France), the region has a high responsibility for the preservation of amphibians and their habitats, including traditional hedgerow landscape and wetlands.

1.2. Biological data

141 Two types of amphibian data were used: (1) opportunistic data from a citizen database with presence-
142 only records to fit and test models (2) standardized detection-nondetection data from a citizen science
143 programme and complementary field work for external evaluations. A more detailed description of the
144 data sets and complementation strategies is available in Appendix 1.

145 *Opportunistic presence data (calibration and cross-validation dataset)*

146 We accessed presence-only occurrences of 9 species of amphibian from a regional database for the
147 period 2013-2019. We retained only species with enough data according to number of predictors used
148 (i.e. at least 477 presence cells; see Appendix 1). We selected only precise GPS records (precision of
149 the observation under 50 meters) and we checked all data for anomalies in geographical location or
150 species identification.

151 For each species, we sorted data to reduce spatial autocorrelation by projecting presence data on a
152 500m resolution grid and retaining only cells containing at least one occurrence as presence cells for
153 the analyses (Guisan & Theurillat, 2000). We chose a 500m resolution as it is the mean size of species'
154 home ranges (Semlitsch & Bodie, 2003). Finally, we excluded all opportunist data intersecting cells
155 used for external validation described below to increase independency between calibration and
156 validation sets.

157 *Standardised detection-nondetection data (external validation datasets)*

158 For external validation, we firstly extracted *detection-nondetection* amphibian data from a regional
159 citizen science database for the period 2013-2019. We retained 576 sites which were monitored at
160 least 3 times one year between February and June, during good weather conditions and following a
161 standard method commonly used for amphibian community surveys (Boissinot, Besnard & Lourdais,
162 2019) (See Appendix 1). We called this dataset *CS.0* (see Figure 1).

163 To complete and filter *CS.0*, different strategies were used. Firstly, we organised complementary
164 fieldwork in 2018 and 2019 to complete two landscape gradients (woody element density and pond

density) which are two variables known to strongly affect amphibian distribution and which are relevant in our regional context. All sites were selected randomly but so as to maximise and decorrelate the two landscape gradients in different areas. Three data collection approaches were developed: (1) 108 sites were proposed for monitoring by volunteers involved in the existing citizen science program; (2) 259 sites were proposed to the regional herpetologist group; (3) we ourselves monitored 132 ponds. In total, 263 sites were monitored: 132 sites by experts in 2018-2019 (called *PRO*, see Figure 1) and 131 by 75 volunteers in 2019 (called *VOL*, see Figure 1). All data (*CS.0*, *PRO* and *VOL*) were projected on the same 500m resolution grid. One further problem, common in citizen science programmes (Geldmann et al., 2016), is that only aquatic sites are surveyed while areas known to be very unsuitable for amphibians such as urban areas and intensive agriculture are generally excluded. To reduce this source of bias, we randomly selected 5% more 500m grid cells in totally urbanized areas without aquatic sites and 5% more grid cells in homogeneous croplands without trees or ponds and we attributed "absence" values to each after field checks (called *ABS*, see Figure 1). These landscapes represent 9% of the total area of the region.

Secondly, we applied different filter combinations to establish sub-sets from *CS.0*, *PRO* and *VOL* (see Figure 1):

(1) a minimum distance of 1km between two grids containing data;

(2) threshold values to validate non-detection as absence data, defined as a minimum sampling effort required to detect a species based on the species' detectability group and observer level of expertise. Four species detectability groups were defined from occupancy studies in France (Boissinot, 2009; Petitot et al., 2014) and the UK (Sewell, Beebee & Griffiths, 2010). Observers were classed as either *expert*, *intermediate* or *novice* using 3 criteria: number of years of participation, number of species observed and permit-holder for amphibian capture. A "*novice*" observer was considered more likely to miss or misidentify a species which was actually present than an "*expert*" observer for the same considered survey effort. In addition, novice observers did not use sampling

190 nets, influencing detectability, in particular of Urodeles. Based on our observer classes and
191 sampling methods used (e.g. acoustic, visual and/or fishing), we set threshold values for the
192 minimal sampling effort needed to validate absence data, depending on species detectability (see
193 Appendix 1 for details).

194 (3) target species to valid nondetection as absence, as recommended by Phillips and co-workers in
195 2009. So, if *species A* with the same detectability as *species B* is detected at a site, then *species B* is
196 likely to be truly absent (see Appendix 1 for target species list).

197 (4) stratified sampling on final prediction maps. We established independent datasets, stratified by
198 model predictions for each species (Newbold et al., 2010; Guisan, Thuiller & Zimmermann, 2017)
199 with equal random sampling of grid cells with data from 4 prediction classes (i.e. [0;0.3[, [0.3;0.5[,
200 [0.5;0.7[and [0.7;1]).

201

202 **1.3. Environmental dataset and variables**

203 We assembled environmental data relevant to amphibian ecology and of importance in the study
204 region (Guisan, Thuiller & Zimmermann, 2017). A more detailed description of variables with associate
205 references is available in Appendix 2.

206 Bioclimatic variables were accessed from a compilation of climate data for the period 1950-2000 at a
207 spatial resolution of 2.5 arc-min (Hijmans et al., 2005). An altitude variable was derived from the U.S.
208 Geological Survey's Hydro-K data set, at the same spatial resolution. We performed a principal-
209 components analysis (PCA) on 11 bioclimatic variables relevant for amphibians and the altitudinal layer
210 to produce 2 uncorrelated axes (see Appendix 2). Land cover data were downloaded from the highly
211 detailed vector database OCS GE 1.1 (IGN 2019), the Theia OSO Land Cover Map 2017 (available at
212 www.theia-land.fr) and from BDTopo (IGN 2019). This was coupled with a more detailed regional

213 inventory of hedgerows (from 2005 to 2008) and ponds (2012) and a national farming database from
214 the EU LPIS (Land Parcel Identification System 2016) used to classify agricultural areas (see Table 1).

215 We calculated land-cover variables in windows composed of a 500m grid cell with a buffer of 300m
216 (see Table 1). This took into account landscape context based on species' dispersal capacities as well
217 as the resolution of the species data set (Guisan & Thuiller, 2005). Distance and home range differ
218 among amphibian species but a 1km circle may be accepted as an average maximum range (Collins &
219 Fahrig, 2017). Collins and Fahrig (2017) and Boissinot et al. (2019) show that landscape variables affect
220 Anuran occupancy and diversity at this scale in agriculture-dominated regions. We use the same
221 environmental variables for all species (see Table 2) except *Bufo spinosus* for which "pond density"
222 (water point <5000m²) was substituted by "water point density" because of this species' ability to
223 reproduce in larger water bodies with fish (Boissinot, Besnard & Lourdais, 2019).

224 All predictive variables were centred and scaled. The spatial correlation between environmental
225 predictors was investigated using the variance inflation factor (VIF) as a measure of multicollinearity
226 and Pearson correlation tests with VIF<6 and r<0.6 as advised by O'Brien, 2007 (see Appendix 2).

227

228 **1.4. Habitat suitability modelling**

229 1.4.1. Statistical models

230 Different modelling algorithms can lead to varying results according to heterogeneous sensitivities and
231 calculation processes (Thuiller et al., 2009). Therefore, consensus models based on multi-modelling
232 approaches (ensemble-modelling) can improve final results by reducing 'noise' associated with
233 individual model errors (Araujo et al., 2005; Thuiller et al., 2009; Meller et al., 2014). For each species,
234 we used one regression-based approach (Generalized Additive Models, GAM) and one machine
235 learning algorithm (Random Forest, RF) to predict and assess habitat suitability within the studied
236 region with 50 bootstrap replicates. Presence points were randomly split 50 times into a training set

237 (70% of the whole dataset) and the remaining 30% were used as testing set for internal evaluation (see
238 1.4.). To construct these models, we used biomod2 package (Thuiller et al., 2009) in R environment v.
239 3.5.3 (R Development Core Team, 2019).

240 1.4.2. Background data and pseudo-absence selection

241 Modelling habitat suitability for a species with GAM or Random Forest requires both presence and
242 absence data. We tested three strategies for generating artificial absence points: (s1) simple random
243 selection of background points within the studied region (Guisan, Thuiller & Zimmermann, 2017); (s2)
244 random pseudo-absence selection excluding known presence points (Engler, Guisan & Rechsteiner,
245 2004); (s3) s3: random pseudo-absence selection constrained to take sampling effort into account. The
246 latter strategy aimed to select pseudo-absences where true absences were more likely. For this
247 strategy we considered three main sources of bias in pseudo-absence selection: accessibility, linked
248 to distance from roads or urban areas (Kadmon, Farber & Danin, 2004; Barbet-Massin et al., 2012),
249 attractiveness, relating to oversampling in protected sites or tourist areas (Phillips et al., 2009;
250 Robinson, Ruiz-Gutierrez & Fink, 2018) and observer effort, because certain administrative areas are
251 covered by particularly active nature protection organisations. For each strategy, the number of
252 artificial absences was fixed equal to the number of presence data (Barbet-Massin et al., 2012; Liu,
253 Newell & White, 2019) and we performed 10 replicates of the artificial absence generation processes.

254 1.4.3. Ensemble modelling

255 Finally, we conducted ensemble modelling using the median value of (1) all generated individual
256 maps (i.e. 500 maps/algorithm) (Thuiller et al., 2009) to compare internal *versus* external evaluation
257 and (2) maps calibrated with 100% of presence-only data to compare different external evaluations
258 sets (i.e. 10 maps/algorithm).

259

260 **1.5. Internal and external model validation**

261 We firstly use a cross-validation method using a 30% random split of the whole set to asses each model
262 (for pseudo-absence selection strategies s1, s2 and s3) with 50 bootstraps repeated 10 times. We
263 calculated the area under the curve (AUC) of a receiver operating characteristic (ROC) plot of the
264 predicted model habitat suitability scores with (1) the 30% test dataset and (2) with the larger filtered
265 external independent dataset (e.g. *CS.2+ABS+SUP*) using Biomod2 package (Thuiller et al., 2009). AUC
266 is the most common metric used in SDM studies, as it has the advantage of being threshold and
267 prevalence independent and has been accepted as the standard measure for assessing SDM accuracy
268 (Guisan, Thuiller & Zimmermann, 2017). The SDM with the highest AUC value is the most accurate
269 (Raes & Aguirre-Gutiérrez, 2018) and AUC scores range from "fail" 0.32 to "excellent" 0.99 (Swets,
270 1988). AUC > 0.50 signifies that the model has better prediction than a random model.

271 Secondly, we calculated AUC values, specificity (true negative rate) and sensitivity (true positive rate)
272 of ensemble models calibrated with 100% of the presence-only data using different evaluation sets
273 derived from the global external dataset used in the previous stage (see Figure 1). These calculations
274 (with 100 bootstraps) were performed using PresenceAbsence package (Freeman 2012) with a
275 standard threshold value for presence-absence discrimination fixed at 0.5.

276

277 **Results**

278 *Model performance and selection*

279 For each species, the median AUC was higher with internal validation than external validation for all
280 three pseudo-absence selection strategies (s1, s2 and s3), both for GAM and Random Forest (see Figure
281 2 and appendix 3) with a delta-AUC ranging from 0.05 (*Triturus marmoratus*) to 0.21 (*B. spinosus*). With
282 internal evaluation (cross-validation), all models had excellent (AUC>0.90) very good (0.80-0.90) or
283 good accuracy (0.70-0.80) except for the model of *B. spinosus* and *Hyla arborea* including sampling
284 effort parameters (s3). However, with external evaluation, only four species had a high level of

accuracy (AUC>0.70): *Salamandra salamandra*, *T. marmoratus*, *Pelodytes punctatus* and *Rana temporaria*. For *Rana dalmatina*, *T. cristatus* and *Lissotriton helveticus*, model accuracies were poor (0.60<AUC<0.70) and for *B. spinosus* and *H. arborea* even poorer (AUC<0.60). The strategy s1 (background data) was not selected neither with internal nor external evaluation.

The method used for pseudo-absence selection influenced the predictive performance of models but differences between AUC values were minimal (Figure 2). However, s2 (uncorrected sampling bias) was the best strategy for six species when internal validation was used, while s3 was best for seven species when external validation was used. Results for RF can be found in the supplementary material but do not differ greatly.

Impact of model selection on final habitat suitability map

Final maps of habitat suitability resulting from models selected by internal or external evaluation and comparative maps s2 and s3 are available Figure 3. The models selected by internal or external evaluation produced different habitat suitability maps which would lead to different interpretation for conservation decisions. Maps for *H. arborea* and *B. spinosus* are not shown because of poor accuracy (see supplementary results appendix 3). All response curves and associated variable contributions can be found in the supplementary material (Appendix 3).

Comparison of external evaluation sets

Values of AUC, sensitivity and specificity to four species are shown in Table 2 (two Anurans and two Urodeles; one forest specie and one generalist specie each). Results for other species and CS.1 (similar to CS.2) are presented in appendix 3. Considering AUC values, evaluation with the external dataset from participative science without filter data (*CS.0*) show more similar model selection results than internal cross-validation except for *R. temporaria*. Sorting presence data led to decreased sensitivity and increased specificity for all species except for *S. Salamandra*. The model evaluation we obtained

308 were similar whether using stratified data from volunteers' only (STRAT_CS) or stratified data with
309 added professional observations (STRAT_ALL), or professional data only (PRO).

310

311 **Discussion**

312 External evaluation with independent data generated lower AUC values than cross-validation, which
313 calls into question the validity of models validated by commonly used selection threshold values such
314 as AUC> 0.70. According to Araujo et al. 2005, internal evaluation with non-independent data always
315 leads to over-optimistic assessment of model performance. Even if cross-validation is better than
316 substitution procedures (Araujo et al., 2005; Vaughan & Ormerod, 2005; Edwards et al., 2006), split
317 data for internal validation are non-independent and do not avoid the main limits of correlative models
318 in SDMs because of spatial or temporal autocorrelation, especially when sampling effort is highly
319 heterogeneous (Edwards et al., 2006, Roberts et al. 2017). Our result supports criticisms of certain
320 types of SDM and further highlights the need to be careful in their general interpretation and
321 assessment (Lobo, Jiménez-valverde & Real, 2008).

322 The difference between internal-AUC and external-AUC is particularly pronounced for the most
323 common and generalist species in spite of the large number of available data, especially for *B. spinosus*
324 and *H. arborea* whose models failed to attain an acceptable level of accuracy with external evaluation.
325 Brotons et al. 2004 highlighted the difficulty of predicting distributions of the most generalist species.
326 However, for such species, the use of filters increases specificity values considerably and the results
327 are coherent with these species' wide ranging and ubiquitous distributions. Using external presence-
328 absence data also makes it possible to exploit the whole presence-only data set for calibration and to
329 use stricter filters to reduce sampling bias or data culling to retain higher quality data (Steen et al.
330 2019, Isaak et al. 2014). Our study shows that it is possible to apply strong filters (e.g. STRAT_CS) but
331 finally retain reasonable sample sizes for most species.

332 It should be noted that, for four species, *R. temporaria*, *T. marmoratus*, *S. salamandra* and *P.*
333 *punctatus*, our results were ambiguous. For the first three, all forest-dwelling species or very closely
334 related to woodlands (Boissinot, Besnard & Lourdais, 2019), both internal and external validation
335 methods selected models with sampling effort integrated (s3). Two main reasons could explain these
336 results: firstly, presence data may have been insufficient for *R. temporaria*, and secondly these species'
337 affinity for forest habitats. *R. temporaria* is a rare species and is more dependent on wet forest, flood
338 meadows and small streams as breeding sites than other species (Boissinot et al., 2015). This species
339 has a patchy distribution (i.e. locally abundant but regionally rare) and is difficult to detect. Hence,
340 presence data are few in the studied area both in the opportunist dataset (N=477 presence-cells) and
341 in the independent dataset (N=13 presence-cells). As highlighted by Vaughan & Ormerod (2005) such
342 factors can lead to model over fit, even with a relatively small number of variables, resulting in high
343 AUC values. In addition, according to Brotons et al. (2004), low-density habitat (i.e. forest habitat in
344 our region) may be over-weighted and it can be difficult to assess between good or bad suitability
345 without adapted presence-absence data. Monitored forest sites are rare in our validation dataset and
346 the assumptions we used to define sampling effort may not be well adapted for forest specialists.
347 Finally, *P. punctatus* is a rare species but abundant on the Atlantic coast. Unlike the other species, it is
348 a pioneer, adapted to open areas, especially primary unvegetated habitats such as sand dunes and
349 mudflats with frequent physical disturbance (Joly et al., 2005). These habitats are mainly located near
350 the coast and along the main regional floodplains (Loire Valley), with a high local density of presence
351 data. So this species is also patchily distributed and models may be affected by the same bias as *R.*
352 *temporaria*. Alternatively, the similarity between AUC values may also relate to sampling effort bias
353 along the Atlantic coast (e.g. Fithian et al., 2015). There results highlight the need to adapt methods
354 and filters used for each species.

355 **Using heterogeneous data from citizen science in SDM**

356 The reduction of all available data to presence-only leads to loss useful information from the original
357 data and the combined use of different data type and quality could improve SDM (Isaac et al., 2019;
358 Robinson et al., 2020). Our results show that it is possible to obtain useful external and independent
359 datasets for model validation from filtered protocoled citizen science data. Indeed, the use of filters
360 have successfully reduce bias and noise in citizen science data sets for SDM in others studies (Robinson,
361 Ruiz-Gutierrez & Fink, 2018; Isaac et al., 2019; Steen, Elphick & Tingley, 2019).

362 Since external independent data is necessary for more robust assessment of SDM (Araujo et al., 2005),
363 but prohibitive to collect, filtering low quality but large datasets from monitoring to obtain more
364 standard and independent data may be worthwhile. In addition, AUC appears to be more informative
365 when presence-absence data is used to assess and compare models than when presence-background
366 data alone is used (Jiménez-Valverde, 2012). However, using detection-nondetection citizen data
367 without filters may also lead to erroneous results because of overlapping sources of bias in both
368 datasets (e.g. *CS.0* selects the same model as cross-validation). The large amount of available data
369 allows us to strongly select data according to our research objective. In addition, the use of citizen
370 science data was useful to cover larger spatial and temporal space, to get more data about rare species
371 and also to use the whole presence-only dataset for model calibration. Our results using *PRO* datasets
372 are inconclusive for rare species perhaps because their detection was insufficiently frequent (e.g. only
373 two observations of *R. temporaria* for 132 sampled sites). Finally, we found that general rules to guide
374 data sorting were difficult to define. Our results were sensitive to the type of data used, and the species
375 studied, reinforcing the need to define filters according to available data and species' ecology (Steen,
376 Elphick & Tingley, 2019).

377 **Involved stakeholders and citizens in the research process**

378 In our studies, 75 volunteers were involved and good communication between us (researchers),
379 program coordinators and volunteers was important to facilitate exchange of information and
380 knowledge. Where volunteer citizens are involved, communication is essential for both project success

and ethical reasons (Ellwood, Crimmins & Miller-Rushing, 2017) and regular feedback keeps volunteers motivated (Newman et al., 2017). Information letters and talks at meetings of local nature groups were used to inform the people involved in our project of its progress i.e. to present general aims, to recruit volunteers and to present preliminary and final results. According to Zapponi et al. (2017) mapping the distribution of species or habitat suitability can be an especially a good aim for a citizen science programme. Indeed, many nature organisations are already involved in distribution atlas projects and naturalists are aware of data collection methods and local species distributions. They seek out ways to prioritise field observations; making a useful contribution to developing SDMs to guide conservation action can be a source of motivation, making scientist-volunteer interactions easier. Involving citizens at different stage of the mapping process may make action easier to implement, through both better shared knowledge and stronger personal involvement. Indeed, Forrester et al. (2017) and Lewandowski and Oberhauser (2017) highlight an increase in conservation advocacy among participants of citizen science projects that might improve access to evidence for conservationist and decision makers (Sutherland & Wordley, 2017).

Acknowledgements

This work would not have been possible without the support of the Pays de la Loire Herpetological Group, the CPIE Regional Union and the French BirdLife International partner (LPO). We are especially grateful to Benoit Marchadour (LPO) who coordinates the regional distribution atlas. We also acknowledge the many naturalists involved, for access to data and for additional fieldwork, especially Dorian Angot, Anne-Lise Charpentier, Baptiste Gaboriau, Ludovic Aubry, Jeanne Abbou, Flore Gamet, François Varenne, Adrien Maitrepierre, Leslie Michal, and Martin Bonhomme. We thank Xavier Lozupone for his help for mapping of environmental variables at regional level and Véronique Beaujouan, Adeline Bulot, Hervé Daniel, Gilles Martel and Aurélien Besnard for providing helpful

discussion. Our work was supported by funding from Angers Loire Metropole, The French Society for Ecology and Evolution and “Humanity and Biodiversity”.

Author contributions

F.M., M.S., G.P., J.B. and J.P. contributed to the original idea. F.M. collected biodiversity data, M.S. and F.M. coordinated the voluntary work. F.M. analysed the data and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Supplementary materials

Appendix 1 –

https://www.biorxiv.org/content/biorxiv/early/2020/06/03/2020.06.02.129536/DC1/embed/media-1.pdf?download=true

Appendix 2 –

https://www.biorxiv.org/content/biorxiv/early/2020/06/03/2020.06.02.129536/DC2/embed/media-2.pdf?download=true

Appendix 3 –

https://www.biorxiv.org/content/biorxiv/early/2020/06/03/2020.06.02.129536/DC3/embed/media-3.pdf?download=true

Bibliographie

Araujo, M. *et al.* (2005) 'Validation of species-climate impact models under climate change', Global Change Biology, 11(9), pp. 1504–1513. doi: 10.1111/j.1365-2486.2005.001000.x.

Barbet-Massin, M. *et al.* (2012) 'Selecting pseudo-absences for species distribution models: How, where and how many?', Methods in Ecology and Evolution, 3(2), pp. 327–338. doi: 10.1111/j.2041-210X.2011.00172.x.

Barve, N. *et al.* (2011) 'The crucial role of the accessible area in ecological niche modeling and species distribution modeling', Ecological Modelling, 222(11), pp. 1810–1819. doi: 10.1016/j.ecolmodel.2011.02.011.

Baudry, J., Bunce, R. G. H. and Burel, F. (2000) 'Hedgerows: An international perspective on their origin, function and management', Journal of Environmental Management, 60(1), pp. 7–22. doi: 10.1006/jema.2000.0358.

Boissinot, A. (2009) *Influence de la structure du biotope de reproduction et de l'agencement du paysage, sur le peuplement d'amphibiens d'une région bocagère de l'ouest de la France*, Diplôme de l'Ecole Pratique des Hautes Etudes. Laboratoire de Biogéographie et d'Ecologie des Vertébrés (EPHE/CEFE/CNRS).

Boissinot, A. *et al.* (2015) 'Small woods positively influence the occurrence and abundance of the common frog (Rana temporaria) in a traditional farming landscape', Amphibia Reptilia, 36(4), pp. 417–424. doi: 10.1163/15685381-00003013.

Boissinot, A., Besnard, A. and Lourdais, O. (2019) 'Agriculture , Ecosystems and Environment Amphibian diversity in farmlands : Combined in fl uences of breeding-site and landscape attributes in western France', Agriculture, Ecosystems and Environment. Elsevier, 269(September 2018), pp. 51–61. doi: 10.1016/j.agee.2018.09.016.

Brotons, L. *et al.* (2004) 'Presence-absence versus presence-only modelling methods for predicting bird habitat suitability', Ecography, 27(4), pp. 437–448. doi: 10.1111/j.0906-7590.2004.03764.x.

449 Brown, J. L. *et al.* (2016) 'Spatial biodiversity patterns of Madagascar's amphibians and reptiles', PLoS
450 ONE, 11(1), pp. 1–26. doi: 10.1371/journal.pone.0144076.

451 Collins, S. J. and Fahrig, L. (2017) 'Responses of anurans to composition and configuration of
452 agricultural landscapes', Agriculture, Ecosystems and Environment. Elsevier B.V., 239, pp. 399–409.
453 doi: 10.1016/j.agee.2016.12.038.

454 Connor, T. *et al.* (2018) 'Effects of grain size and niche breadth on species distribution modeling',
455 Ecography, 41(8), pp. 1270–1282. doi: 10.1111/ecog.03416.

456 Cosentino, B. J. *et al.* (2014) 'Citizen science reveals widespread negative effects of roads on amphibian
457 distributions', Biological Conservation, 180, pp. 31–38. doi: 10.1016/j.biocon.2014.09.027.

458 Cushman, S. A. (2006) 'Effects of habitat loss and fragmentation on amphibians: A review and
459 prospectus', Biological Conservation, 128(2), pp. 231–240. doi: 10.1016/j.biocon.2005.09.031.

460 Dickinson, J. L., Zuckerberg, B. and Bonter, D. N. (2010) 'Citizen Science as an Ecological Research Tool:
461 Challenges and Benefits', Annual Review of Ecology, Evolution, and Systematics, 41(1), pp. 149–
462 172. doi: 10.1146/annurev-ecolsys-102209-144636.

463 Edwards, T. C. *et al.* (2006) 'Effects of sample survey design on the accuracy of classification tree
464 models in species distribution models', 9, pp. 132–141. doi: 10.1016/j.ecolmodel.2006.05.016.

465 Ellwood, E. R., Crimmins, T. M. and Miller-Rushing, A. J. (2017) 'Citizen science and conservation:
466 Recommendations for a rapidly moving field', Biological Conservation. Elsevier Ltd, 208, pp. 1–4.
467 doi: 10.1016/j.biocon.2016.10.014.

468 Engler, R., Guisan, A. and Rechsteiner, L. (2004) 'An improved approach for predicting the distribution
469 of rare and endangered species from occurrence and pseudo-absence data', The Journal of Applied
470 Ecology, 41, pp. 263 –274. doi: 10.1111/j.0021-8901.2004.00881.x.

471 Fithian, W. *et al.* (2015) 'Bias correction in species distribution models: Pooling survey and collection
472 data for multiple species', Methods in Ecology and Evolution, 6(4), pp. 424–438. doi: 10.1111/2041-
473 210X.12242.

Forrester, T. D. *et al.* (2017) 'Creating advocates for mammal conservation through citizen science', Biological Conservation. Elsevier Ltd, 208, pp. 98–105. doi: 10.1016/j.biocon.2016.06.025.

Geldmann, J. *et al.* (2016) 'What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency requirements', Diversity and Distributions, 22(11), pp. 1139–1149. doi: 10.1111/ddi.12477.

Graham, C. H. and Hijmans, R. J. (2006) 'A comparison of methods for mapping species ranges and species richness', Global Ecology and Biogeography, 15(6), pp. 578–587. doi: 10.1111/j.1466-8238.2006.00257.x.

Guillera-Arroita, G. *et al.* (2015) 'Is my species distribution model fit for purpose? Matching data and models to applications', Global Ecology and Biogeography, 24(3), pp. 276–292. doi: 10.1111/geb.12268.

Guisan, A. and Theurillat, J.-P. (2000) 'Assessing alpine plant vulnerability to climate change: a modeling perspective', Integrated Assessment, 1(4), pp. 307–320. doi: 10.1023/A:1018912114948.

Guisan, A. and Thuiller, W. (2005) 'Predicting species distribution: Offering more than simple habitat models', Ecology Letters, 8(9), pp. 993–1009. doi: 10.1111/j.1461-0248.2005.00792.x.

Guisan, A., Thuiller, W. and Zimmermann, N. E. (2017) *Habitat Suitability and Distribution Models: With Applications in R*, Ecology, biodiversity and conservation. The University of Chicago Press. doi: doi: 10.1017/9781139028271.

Hijmans, R. J. *et al.* (2005) 'Very high resolution interpolated climate surfaces for global land areas', International Journal of Climatology, 25(15), pp. 1965–1978. doi: 10.1002/joc.1276.

Isaac, N. J. B. *et al.* (2014) 'Statistics for citizen science: Extracting signals of change from noisy ecological data', Methods in Ecology and Evolution, 5(10), pp. 1052–1060. doi: 10.1111/2041-210X.12254.

Isaac, N. J. B. *et al.* (2019) 'Data Integration for Large-Scale Models of Species Distributions', Trends in Ecology and Evolution, 35(1), pp. 56–67. doi: 10.1016/j.tree.2019.08.006.

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

Joly, P. *et al.* (2005) 'Canalization of size at metamorphosis despite temperature and density variations in Pelodytes punctatus', Herpetological Journal, 15(1), pp. 45–50.

Kadmon, R., Farber, O. and Danin, A. (2004) 'Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models', Ecological Applications, 14(2), pp. 401–413. doi: 10.1890/02-5364.

Kosmala, M. *et al.* (2016) 'Assessing data quality in citizen science', Frontiers in Ecology and the Environment, 14(10), pp. 551–560. doi: 10.1002/fee.1436.

Lewandowski, E. J. and Oberhauser, K. S. (2017) 'Butterfly citizen scientists in the United States increase their engagement in conservation', Biological Conservation. Elsevier B.V., 208, pp. 106–112. doi: 10.1016/j.biocon.2015.07.029.

Liu, C., Newell, G. and White, M. (2019) 'The effect of sample size on the accuracy of species distribution models: considering both presences and pseudo-absences or background sites', Ecography, 42(3), pp. 535–548. doi: 10.1111/ecog.03188.

Lobo, J. M., Jiménez-valverde, A. and Real, R. (2008) 'AUC: A misleading measure of the performance of predictive distribution models', Global Ecology and Biogeography, pp. 145–151. doi: 10.1111/j.1466-8238.2007.00358.x.

McKinley, D. C. *et al.* (2017) 'Citizen science can improve conservation science, natural resource management, and environmental protection', Biological Conservation. Elsevier Ltd, 208, pp. 15–28. doi: 10.1016/j.biocon.2016.05.015.

Meller, L. *et al.* (2014) 'Ensemble distribution models in conservation prioritization: From consensus predictions to consensus reserve networks', Diversity and Distributions, 20(3), pp. 309–321. doi: 10.1111/ddi.12162.

Newman, G. *et al.* (2017) 'Leveraging the power of place in citizen science for effective conservation decision making', Biological Conservation. The Authors, 208, pp. 55–64. doi: 10.1016/j.biocon.2016.07.019.

O'Brien, R. M. (2007) 'A caution regarding rules of thumb for variance inflation factors', Quality and Quantity, 41(5), pp. 673–690. doi: 10.1007/s11135-006-9018-6.

Petitot, M. *et al.* (2014) 'Optimizing occupancy surveys by maximizing detection probability: Application to amphibian monitoring in the Mediterranean region', Ecology and Evolution, 4(18), pp. 3538–3549. doi: 10.1002/ece3.1207.

Petrovan, S. O. and Schmidt, B. R. (2016) 'Volunteer conservation action data reveals large-scale and long-term negative population trends of a widespread amphibian, the common toad (Bufo bufo)', PLoS ONE, 11(10), pp. 1–12. doi: 10.1371/journal.pone.0161943.

Phillips, S. J. *et al.* (2009) 'Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data', Ecological Applications, 19(1), pp. 181–197. doi: 10.1890/07-2153.1.

Préau, C. *et al.* (2019) 'Predicting suitable habitats of four range margin amphibians under climate and land-use changes in southwestern France', Regional Environmental Change. Regional Environmental Change, 19(1), pp. 27–38. doi: 10.1007/s10113-018-1381-z.

Raes, N. and Aguirre-Gutiérrez, J. (2018) 'A Modeling Framework to Estimate and Project Species Distributions in Space and Time', Mountains, Climate and Biodiversity, pp. 309–320.

Robinson, O. J. *et al.* (2020) 'Integrating citizen science data with expert surveys increases accuracy and spatial extent of species distribution models', Diversity and Distributions, (October 2019), p. 806547. doi: 10.1101/806547.

Robinson, O. J., Ruiz-Gutierrez, V. and Fink, D. (2018) 'Correcting for bias in distribution modelling for rare species using citizen science data', Diversity and Distributions, 24(4), pp. 460–472. doi: 10.1111/ddi.12698.

549 Schmeller, D. S. *et al.* (2009) 'Ventajas del monitoreo de biodiversidad basado en voluntarios en
550 Europa', Conservation Biology, 23(2), pp. 307–316. doi: 10.1111/j.1523-1739.2008.01125.x.

551 Schulman, L., Toivonen, T. and Ruokolainen, K. (2007) 'Analysing botanical collecting effort in Amazonia
552 and correcting for it in species range estimation', Journal of Biogeography, 34(8), pp. 1388–1399.
553 doi: 10.1111/j.1365-2699.2007.01716.x.

554 Semlitsch, R. D. and Bodie, J. R. (2003) 'Biological Criteria for Buffer Zones around Wetlands and
555 Riparian Habitats for Amphibians and Reptiles', Conservation Biology, 17(5), pp. 1219–1228. doi:
556 10.1046/j.1523-1739.2003.02177.x.

557 Sewell, D., Beebee, T. J. C. and Griffiths, R. A. (2010) 'Optimising biodiversity assessments by
558 volunteers: The application of occupancy modelling to large-scale amphibian surveys', Biological
559 Conservation, 143(9), pp. 2102–2110. doi: 10.1016/j.biocon.2010.05.019.

560 Sinsch, U. (1990) 'Migration and orientation in anuran amphibians', Ethology Ecology and Evolution,
561 2(1), pp. 65–79. doi: 10.1080/08927014.1990.9525494.

562 De Solla, S. R. *et al.* (2005) 'Effect of sampling effort and species detectability on volunteer based
563 anuran monitoring programs', Biological Conservation, 121, pp. 585–594. doi:
564 10.1016/j.biocon.2004.06.018.

565 Steen, V. A., Elphick, C. S. and Tingley, M. W. (2019) 'An evaluation of stringent filtering to improve
566 species distribution models from citizen science data', Diversity and Distributions, 25(12), pp. 1857–
567 1869. doi: 10.1111/ddi.12985.

568 Stuart, S. N. *et al.* (2004) 'Status and trends of amphibian declines and extinctions worldwide', Science,
569 306(5702), pp. 1783–1786. doi: 10.1126/science.1103538.

570 Sutherland, W. J. and Wordley, C. F. R. (2017) 'Evidence complacency hampers conservation', Nature
571 Ecology and Evolution. Springer US, 1(9), pp. 1215–1216. doi: 10.1038/s41559-017-0244-1.

572 Swets, J. (1988) 'Measuring the accuracy of diagnostic systems', Science, 240(4857), pp. 1285–1293.
573 doi: 10.1126/science.3287615.

Thuiller, W. *et al.* (2009) 'BIOMOD - A platform for ensemble forecasting of species distributions', Ecography, 32(3), pp. 369–373. doi: 10.1111/j.1600-0587.2008.05742.x.

Vaughan, I. P. and Ormerod, S. J. (2005) 'The continuing challenges of testing species distribution models', Journal of Applied Ecology, 42(4), pp. 720–730. doi: 10.1111/j.1365-2664.2005.01052.x.

Zapponi, L. *et al.* (2017) 'Citizen science data as an efficient tool for mapping protected saproxylic beetles', Biological Conservation, 208, pp. 139–145. doi: 10.1016/j.biocon.2016.04.035.

Zipkin, E. F. and Saunders, S. P. (2018) 'Synthesizing multiple data types for biological conservation using integrated population models', Biological Conservation. Elsevier, 217(March 2017), pp. 240–250. doi: 10.1016/j.biocon.2017.10.017.

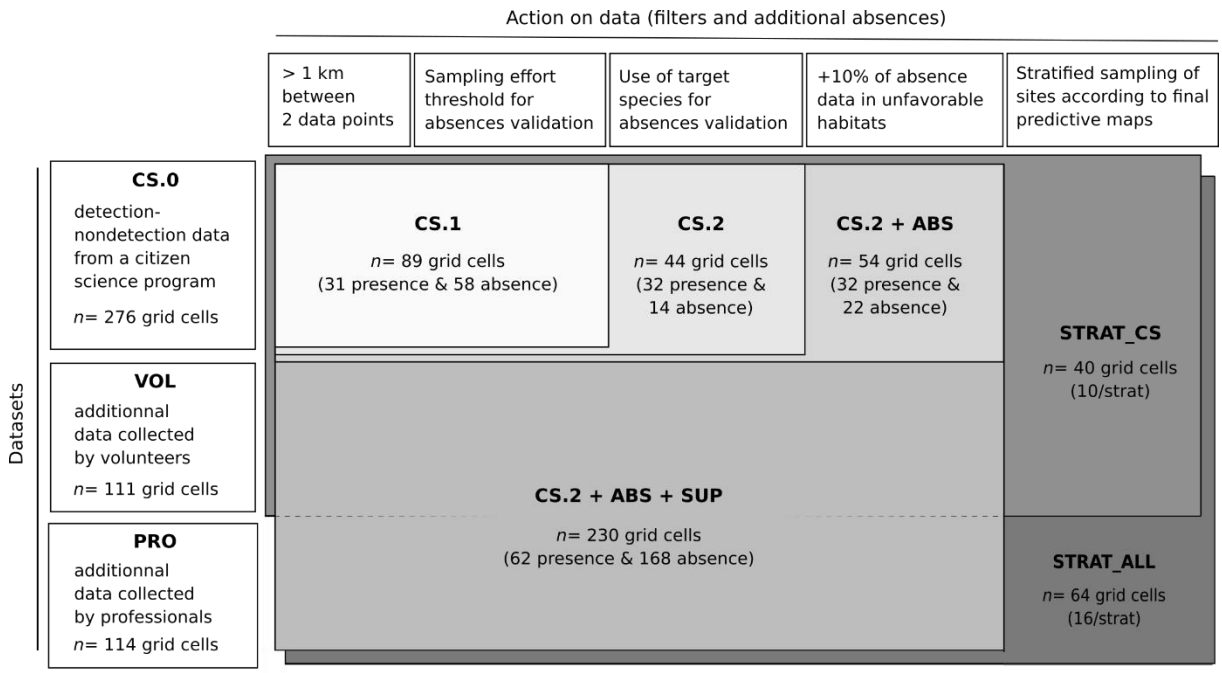

Figure 1. Description of datasets, filters and complementation used for external evaluation derived. Number of data (*n* grid cells) given for *T. marmoratus*, for 1 iteration only and s3. CS: data derived from the citizen science data set; SUP: additional data collected by volunteers (VOL) and professionals (PRO).

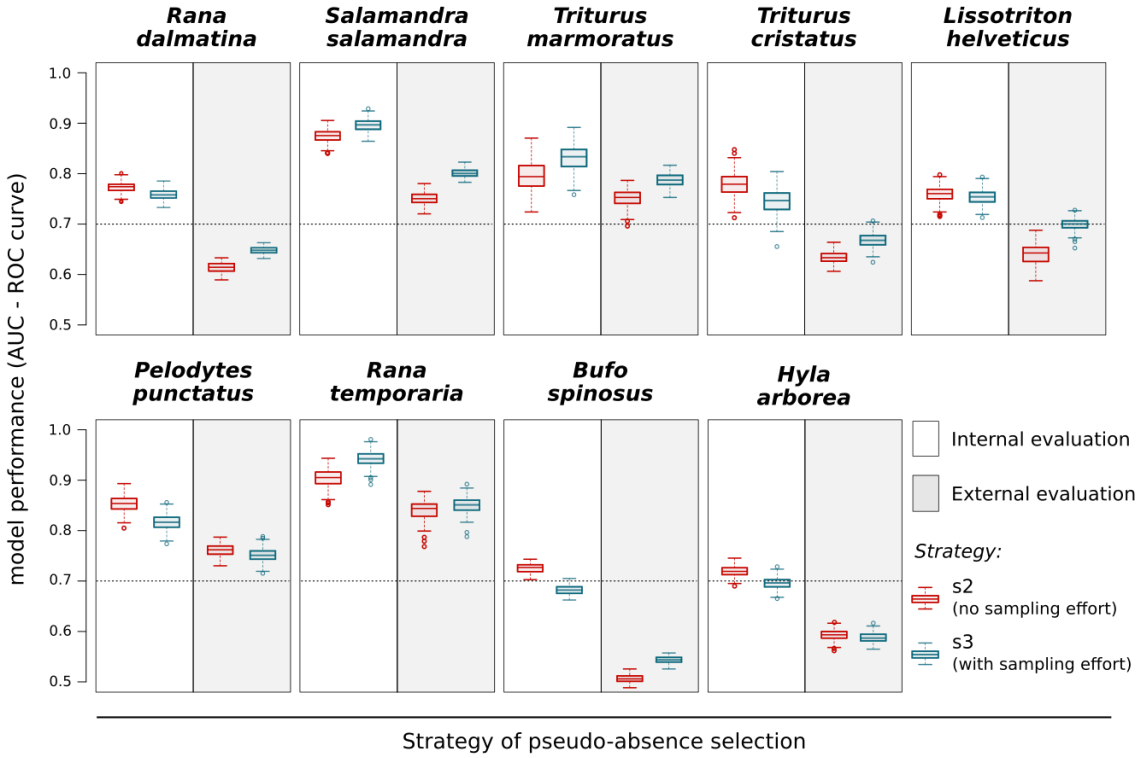

Figure 2. Model performance for the 9 studied species assessed by external or internal data using different pseudo-absence selection strategies. Assessment by AUC under the ROC for GAM only are shown (see supplementary material). Artificial absence sampling strategies shown are s2 (random pseudo-absence selection excluding known presence points) and s3 (random pseudo-absence selection excluding known presence points and adjusted to consider site accessibility and sampling effort). Per strategy, 10 replicates of the artificial absence points generation processes with 50 bootstraps for the random selection of the straining set (70%) and the internal testing set (30%). Black dotted line indicates the 0.70 threshold above which models have an acceptable level of accuracy (Swets 1988).

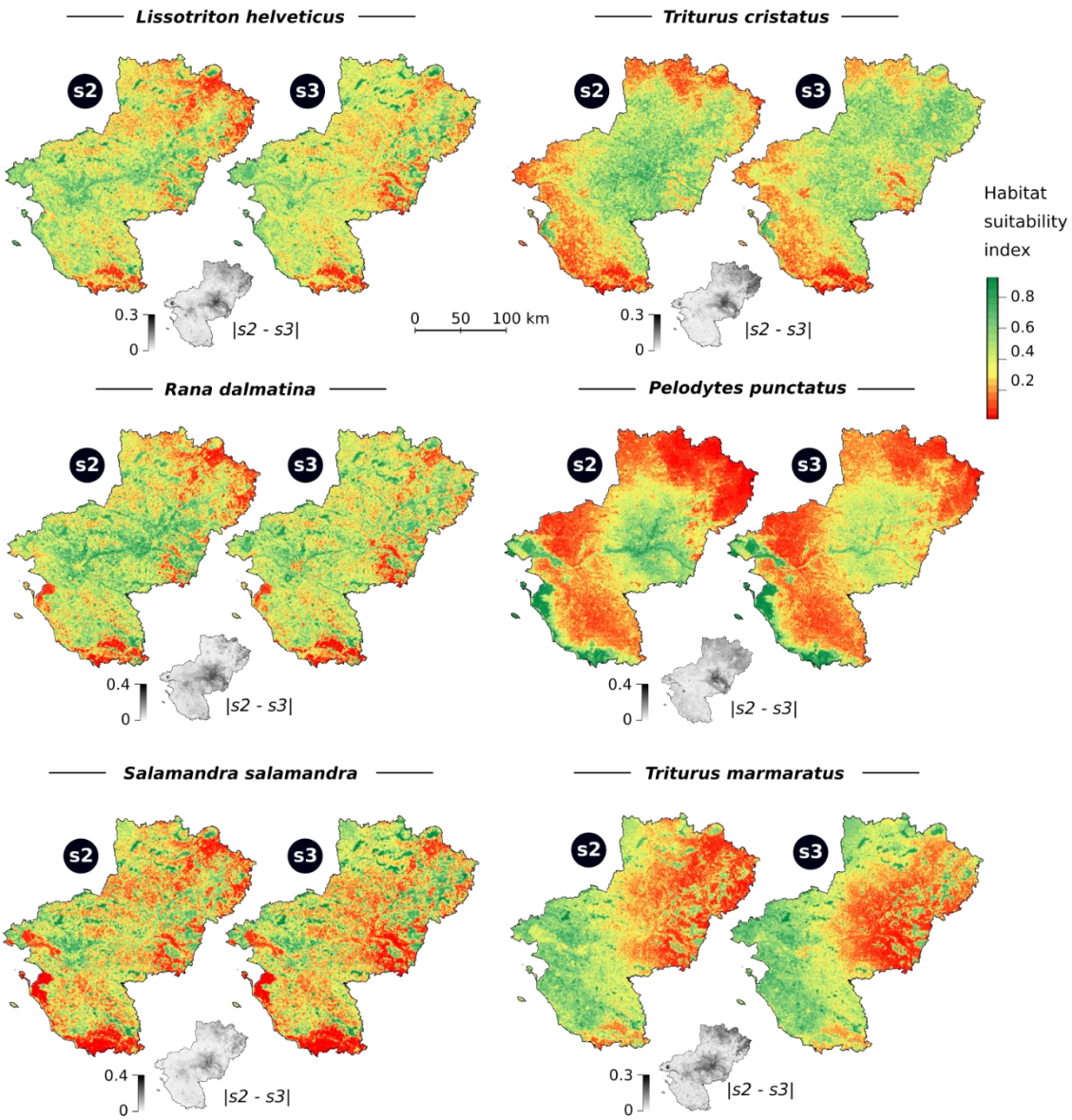

Figure 3. Habitat suitability maps for six studied species produced using two form of pseudo-absence selection: s2 (random pseudo-absence selection excluding known presence points) and s3 (random pseudo-absence selection constrained to account sampling effort). The black and white map under each pair shows net difference between s2 and s3. Map resolution is 500m.

Variable category	Code	Variable description	Original resolution
Climatic	CLIM_1	first axis from a PCA on 12 worldclim variables and altitude	2.5 arc-min
	CLIM_2	second axis from a PCA on 12 worldclim variables and altitude	2.5 arc-min
Land cover	%WOOD_DM	Proportion of deciduous and mixed forest	5m
	%WOOD_C	Proportion of coniferous forest	5m
	%CROP	Proportion of crop	20m
	%PASTURE	Proportion of permanent pasture	20m
	NB_PONDS	Pond density (or water point density)	5m
	L_HEDGE	Hedgerow density	5m
	L_ROAD_1ST	Primary road density outside urban areas	5m
	L_ROAD_2ND	Secondary road density outside urban areas	5m
	L_RIVER	Canal and river density	5m
	%URBAN	Proportion of urban area	20m

Table 1. Environmental variables used for species distribution modelling of each amphibian species in Pays de la Loire region. Associated references are available in the Appendix 2.

		Triturus marmoratus			*Lissotriton helveticus*			*Hyla arborea*			*Rana temporaria*		
		SEN	**SPE**	**AUC**	**SEN**	**SPE**	**AUC**	**SEN**	**SPE**	**AUC**	**SEN**	**SPE**	**AUC**
CS.0	*s2*	0,63	0,41	0,58	**0,69**	0,29	0,53	**0,82**	0,43	0,68	0,71	**0,87**	0,87
	s3	**0,65**	**0,45**	**0,61**	0,64	**0,34**	0,54	0,78	**0,49**	0,68	**0,86**	0,84	0,88
CS.2 + ABS	*s2*	0,59	0,78	0,81	**0,67**	0,76	0,80	**0,71**	0,61	0,73	0,64	**0,88**	0,86
	s3	0,58	**0,87**	0,81	0,62	**0,88**	**0,86**	0,68	**0,64**	0,74	**0,82**	0,84	0,86
CS.2 + ABS + SUP	*s2*	0,58	**0,82**	0,82	**0,51**	0,77	0,68	**0,63**	**0,56**	0,63	0,59	**0,91**	0,85
	s3	**0,62**	0,78	0,81	0,49	0,78	**0,74**	0,57	0,54	0,63	**0,76**	0,86	0,85
PRO	*s2*	0,58	**0,83**	**0,85**	**0,41**	0,76	0,61	**0,59**	**0,39**	0,51	0,50	0,95	0,71
	s3	**0,79**	0,77	0,83	0,37	**0,79**	**0,64**	0,51	0,33	**0,54**	0,50	0,92	0,71
STRAT_CS	s2	0,71	**0,72**	**0,81**	0,67	0,81	0,82	0,72	0,64	0,73	0,79	0,57	0,69
	s3	0,70	0,69	0,78	0,68	0,82	**0,88**	0,71	0,63	**0,75**	**0,95**	**0,60**	**0,79**
STRAT_ALL	s2	**0,80**	**0,74**	**0,87**	0,66	0,76	0,78	**0,68**	**0,62**	0,70	0,84	0,57	0,73
	s3	0,78	0,70	0,83	**0,71**	**0,83**	**0,89**	0,66	0,60	**0,72**	**0,96**	0,58	**0,81**

Table 2. Model performance according to different filters and complementary fieldwork applied to the external evaluation dataset. External datasets used were (see Figure 1): *CS.0* (all data from the standardized citizen science dataset); *CS.2 + ABS* (*CS.2* with 10% supplementary absence cells in very unfavourable habitats); *PRO* (data collected by professionals only in 2018-2019); *CS.2 + ABS + SUP* (citizen science data cited before adding all complementary fieldwork by professionals and volunteers); STRAT_CS (stratified data selection from CS.2+ABS with complementary fieldwork by volunteers); STRAT_ALL (stratified data selection from CS.2+ABS+SUP). Models assessed: s2 (random pseudo-absence selection excluding known presence points) and s3 (random pseudo-absence selection constrained to account sampling effort and correct sampling bias). SEN: sensitivity, SPE: specificity. Bold values show best values between s2 et s3 with delta > 0.02 and grey cells show species with less than 2 presence data. All analyses with a random sampling in presence selection with a distance condition or a stratified random selection were performed using 100 bootstraps (mean calculation).