

The iterative ensemble modelling approach increases the accuracy of fish distribution models

Christine Lauzeral, Gaël Grenouillet and Sébastien Brosse

C. Lauzeral (christine.lauzeral@univ-tlse3.fr), G. Grenouillet and S. Brosse, Univ. de Toulouse, UPS, ENFA, UMR5174 EDB (Laboratoire Évolution et Diversité Biologique), 118 route de Narbonne, FR-31062 Toulouse, France, and CNRS, UMR5174 EDB, FR-31062 Toulouse, France. CL also at: Laboratoire Evolution et Diversité Biologique, U.M.R 5174, C.N.R.S – Univ. Paul Sabatier, 118 route de Narbonne, FR-31062 Toulouse cedex 4, France.

Methodological absences, i.e. when a species is not detected although it is actually present, are known to reduce the prediction accuracy of species distribution models (SDMs). To deal with this problem, we assessed whether a new iterative ensemble modelling (IEM) approach better predicts the spatial distribution of a set of 31 freshwater fish species, exhibiting a wide range of prevalence and methodological absences. Model efficiency was compared using one threshold-independent (AUC) and three threshold-dependent indicators of model predictive performance: the percentage of misclassified sites; the Kappa index; and the True Skill Statistic. We then reconstructed species assemblages from individual species predictions and compared observed assemblages to those predicted using EM and IEM using the Jaccard index. Compared to an EM approach, IEM improved model predictive performance for most difficult-to-detect species. The iterative approach outperformed EM at modelling the distribution of difficult-to-detect species, provided that presence data are representative of the niche of the species. At the assemblage level, the discrepancy between observed and IEM predicted assemblages was significantly lower than that between observed and EM predicted assemblages, showing that IEM can be used to predict the distribution of entire species assemblages. The IEM approach provides a way to consider difficult-to-detect species in species distribution models.

Freshwaters are among the most anthropogenically threatened ecosystems. They are negatively impacted by habitat destruction, biological invasions, pollution and overexploitation (Wilcove et al. 1998, Butchart et al. 2010). They are predicted to face serious threats through climate changes, because aquatic organism dispersal is constrained by the structure of the river networks (Hugueny 1989, Fagan 2002). Hence, a high proportion of strictly freshwater species could soon be at risk of extinction due to their restricted abilities to track shifting climatic conditions (Fausch et al. 2002). In Europe, this is particularly true for coldwater fishes that are predicted to lose a large part of their distribution through climate change (Buisson et al. 2008a). Increased efforts are thus needed to identify freshwater species responses to environmental change and to fully use existing data to guide conservation efforts for freshwater ecosystems.

Species distribution models (SDMs) are increasingly applied as predictive tools for conservation planning and management purposes. They are usually based on the use of presence–absence data, but although the presence of a species is factual (except for misidentifications or georeferencing errors), absence can have multiple meanings. Lobo et al. (2010) listed three distinct types of absences: 1) environmental absences, when the environmental conditions do not allow the presence of the species; 2) contingent absences,

when the environmental conditions are favorable but other factors such as biotic interactions, barriers to dispersion or local extinction are responsible for the absence of the species; 3) methodological absences, when the species is present but not detected. Environmental absences are informative and contribute to SDM's accuracy as much as presences. On the contrary, methodological absences that do not reflect species environmental preferences, reduce the reliability of SDM predictions (Lobo 2008). Unfortunately, fish in general are difficult to detect and their detectability depends on the species, the fishing method and the river size (Murphy and Willis 1996). Methodological absences are thus frequent in fish occurrence databases.

Binomial likelihood models are specially designed to account for potential sampling errors and have recently been used in the context of global change (Moritz et al. 2008, Kéry et al. 2010, Rowe et al. 2010). Unfortunately, these models are designed to be computed using long-term survey data that are rarely available. Presence-only SDMs consider only the presence of a species to determine its niche (Hirzel et al. 2002, Farber and Kadmon 2003), providing an alternative to the problem of absence uncertainty. However, several assumptions are required (i.e. non-biased sampling, constant detectability, Yackulic et al. 2013) and do not allow the computation of species prevalence (Ward et al. 2009). The

use of presence–absence SDMs is therefore recommended when absence data are available (Yackulic et al. 2013).

Recently, it has been shown that consensus methods are able to cope with prediction variability by combining an ensemble of predictions from different modelling methods (Araújo and New 2007). By calculating the general trend among various statistical methods, ensemble modelling (EM) generally provides more accurate predictions, and is therefore recommended as an approach for dealing with inter-model variability in predictive performance (Marmion et al. 2009, Grenouillet et al. 2011). Here using 31 freshwater fish species, an iterative ensemble approach (hereafter called IEM; Lauzeral et al. 2012) was designed to reduce the effect of non-environmental absences on EM. The method is based on the idea that outputs from ensemble models can be used to correct the observed data before they are used as a new EM calibration set. Converting non-environmental absences to presences hence increases the fit of model predictions to corrected occurrences. That fit should theoretically tend towards a perfect fit that represents the end of the iterative process (i.e. predictions do not differ between iterations). This method has been proven to be efficient for dealing with non-environmental absences using virtual species (Lauzeral et al. 2012), built using a threshold approach which eliminates contingent absences (Meynard and Kaplan 2013). In that study, contingent and methodological absences were generated randomly or mostly at the edge of the environmental niche. Although these virtual data provided initial proof of the advantages of IEM, they did not encompass the complexity of ‘real world’ data. The IEM method has thus not been validated using real-world species distribution data. In this context, the main objective of this study was to assess whether IEM approaches well predict the spatial distribution of individual fish species and the composition of assemblages, and to compare these results with predictions from classical EM approaches. To do this, we used an extended dataset containing climate and physical characteristics and stream fish occurrences in France. We modelled the spatial distribution of 31 fish species in 1110 stream sections, and we compared the values predicted by EM and IEM techniques with observed values obtained from thoroughly surveyed sites.

Material and methods

Data

An extensive dataset of freshwater fish occurrences covering the entire French territory was used. It was provided by the French National Agency for Water and Aquatic Environments (Onema; see Poulet et al. 2011 for more details). We selected sites sampled from 1980 to 2005 from the dataset. We obtained a total of more than 10 000 occurrence records for 31 fish species in 1110 stream sections (hereafter referred to as ‘sites’). The temporal changes in fish assemblage composition during the sampling period were limited. No site experienced major changes in human activity (impoundment, pollution) during the sampling period. Moreover, assemblage modifications due to climate variations occurring during the sampling period mostly affected the abundance

and not the occurrence of individual species (Daufresne and Boet 2007). For each site, the environmental characteristics known to be the most relevant descriptors of habitat requirements of fish were measured. Five environmental descriptors were available for each site: distance from the headwater source (DIS, km); surface area of the drainage basin above the sampling site (SDB, km²); slope (SLO, ‰); mean stream width (WID, m) and water depth (DEP, m). These variables were used to construct two summary variables: one (G) describing the position of sites along the upstream–downstream gradient, the other (V) the velocity of the stream. G was obtained following Buisson et al. (2008b), using a principal component analysis (PCA) on DIS and SDB to eliminate the colinearity between these variables. The first axis of the PCA was then kept as a synthetic variable describing the longitudinal gradient. The velocity of the stream (V) was derived from the Chezy formula:

$$V = \log WID + \log DEP + \log SLO - \log(WID + 2DEP).$$

The CRU CL 2.0 (Climatic Research Unit Climatology ver. 2.0) dataset (New et al. 2002) with a resolution of 10' × 10' was used to describe the current climate. Three climatic variables related to ecological requirements of fish were retained: the mean annual air temperature, the mean annual air temperature range and the annual cumulated precipitation. Air temperatures were used as a substitute for water temperatures, which are currently not available for all French streams. Indeed, streams and rivers are reasonably well-mixed water bodies that easily exchange heat with the atmosphere. Thus, air and river water temperatures show a strong positive correlation (Caissie 2006).

Selecting calibration and test data

The entire dataset was first split into two independent parts, acting as calibration and testing subsets. Following Pineda and Lobo (2009), we chose well-monitored sites as test sites to reduce the number of methodological absences. Well-monitored sites were those satisfying the two following criteria: 1) at least 15 successive fish collections performed (once or twice a year) and 2) an average species saturation curve that stabilized at the end of the sampling effort. The average saturation curve resulted from a mean value of 100 saturation curves generated by sampling the data randomly. We retained as test sites the locations where less than one new species appeared on the average saturation curve during the last 5 surveys. This provided a testing subset of 191 well-monitored sites, where fish presences accounted for all the species observed during all the surveys. These test sites were distributed throughout the country (Fig. 1). The remaining 919 sites were kept for calibration data. In the calibration data set, fish occurrence came from a single recent survey (during the last 5 yr), and hence contained a substantial proportion of methodological absences (i.e. undetected species).

Species detectability

The detectability of each species was evaluated on the 191 well-monitored sites. It was calculated as the ratio between

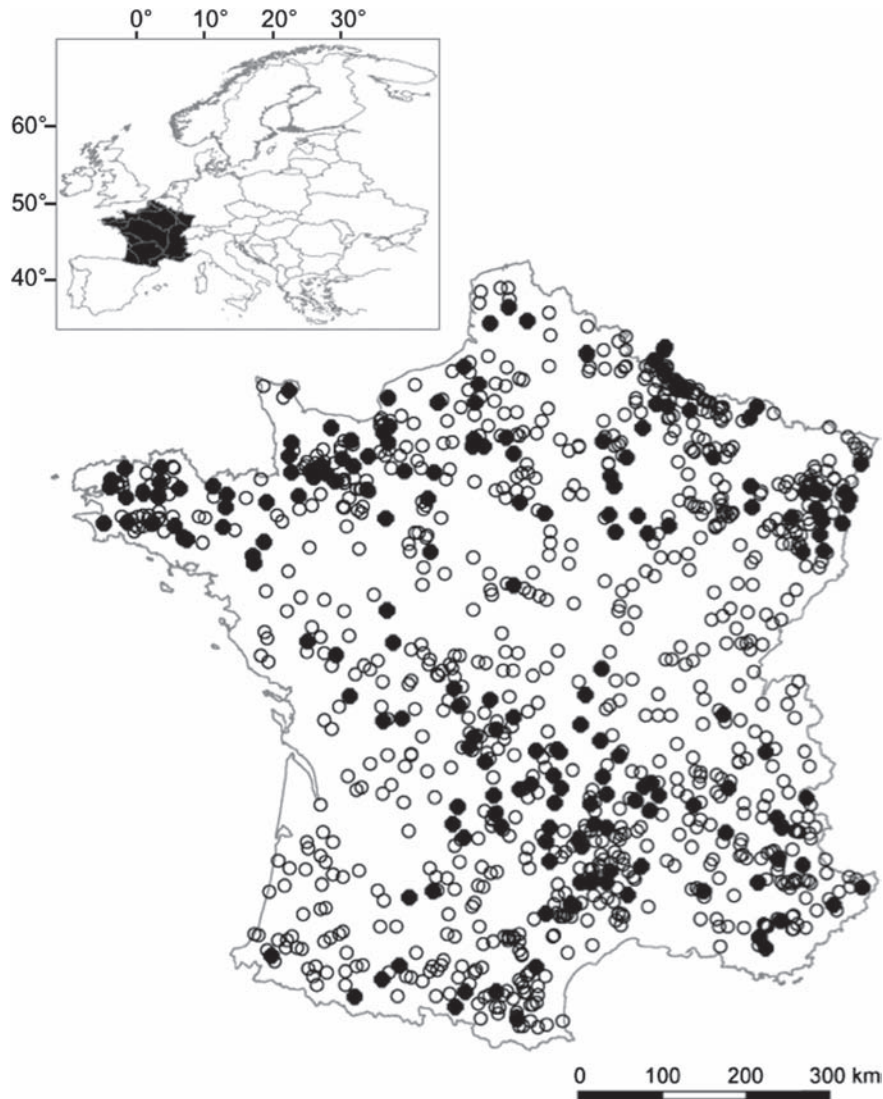


Figure 1. Geographical distribution over France of the 919 fish sampling sites (white dots) used as calibration dataset, and the 191 sites (black dots) identified as well-monitored sites and used as testing dataset (see Material and methods).

the number of surveys where the species was detected (cumulated over sites where the species was detected at least once) and the total number of surveys at the considered sites (Table 1).

EM and IEM modelling

As in an EM framework, we used six predictive modelling methods: generalized linear models (GLM); generalized additive models (GAM); random forest (RF); classification and regression trees (CART); generalized boosted regression models (GBM) and linear discriminant analysis (LDA). All the models were built using the 5 environmental and climatic variables described above, except for GLM and LDA models in which quadratic terms of the variables were also included to deal with non-linearity. The six statistical methods were used to build models using the calibration data set. For each site, the six resulting suitability levels (one per modelling method) were then averaged, giving rise to a per-site suitability level (Marmion et al. 2009) for each

species. We refrained from selecting models or weighting the model outputs using an accuracy measurement like the AUC (area under the ROC curve) that could favor models that overfit data and hence reduce the correction rate of methodological absences.

For each species, the per-site suitability level vector was converted into a presence–absence response using a cut-off threshold determined by maximizing the Kappa index which is a chance-corrected measure (Santika 2011) of SDM accuracy. This approach was preferred to the maximisation of the sum of sensitivity and specificity (or MST), which gives less accurate prevalence predictions (Freeman and Moisen 2008, Mouton et al. 2009). Indeed, the MST reduces specificity, especially for rare species (Liu et al. 2013), and hence increases the risk of niche overestimation through the iterative process.

In the IEM procedure, the presence–absence vector that was predicted using the EM framework (as described above, corresponding to one IEM iteration) was in turn used to update the observed presence–absence vector.

Table 1. Fish species prevalence in the calibration (n = 919) and test (n = 191) data set and species detectability (%).

| Code | Species name | Calibration prevalence | Test prevalence | Detectability |
|------|------------------------------------|------------------------|-----------------|---------------|
| Abb | <i>Abramis brama</i> | 75 | 41 | 39.9 |
| Ala | <i>Alburnus alburnus</i> | 160 | 56 | 66.7 |
| Alb | <i>Alburnoides bipunctatus</i> | 107 | 32 | 43.2 |
| Amm | <i>Ameiurus melas</i> | 48 | 17 | 39.4 |
| Ana | <i>Anguilla anguilla</i> | 310 | 98 | 81.2 |
| Bab | <i>Barbatula barbatula</i> | 508 | 140 | 77.8 |
| Bam | <i>Barbus meridionalis</i> | 58 | 9 | 89.4 |
| Bar | <i>Barbus barbus</i> | 177 | 49 | 60.2 |
| Blb | <i>Blicca bjoerkna</i> | 59 | 36 | 44.2 |
| Chn | <i>Chondrostoma nasus</i> | 73 | 32 | 43.2 |
| Cog | <i>Cottus gobio</i> | 446 | 129 | 78.3 |
| Cyc | <i>Cyprinus carpio</i> | 58 | 38 | 20.1 |
| Esl | <i>Esox lucius</i> | 149 | 66 | 57.2 |
| Gaa | <i>Gasterosteus aculeatus</i> | 88 | 49 | 28.3 |
| Gog | <i>Gobio gobio</i> | 446 | 126 | 74.1 |
| Gyc | <i>Gymnocephalus cernua</i> | 44 | 32 | 38.6 |
| Lap | <i>Lampetra planeri</i> | 218 | 94 | 59.1 |
| Leg | <i>Lepomis gibbosus</i> | 121 | 61 | 41.7 |
| Lel | <i>Leuciscus leuciscus</i> | 214 | 71 | 57.8 |
| Lol | <i>Lota lota</i> | 33 | 14 | 30.9 |
| Pat | <i>Parachondrostoma toxostoma</i> | 40 | 9 | 55.1 |
| Pef | <i>Perca fluviatilis</i> | 227 | 85 | 64.8 |
| Php | <i>Phoxinus phoxinus</i> | 502 | 151 | 75.8 |
| Pup | <i>Pungitius pungitius</i> | 64 | 37 | 39.9 |
| Rha | <i>Rhodeus amarus</i> | 68 | 27 | 39.5 |
| Rur | <i>Rutilus rutilus</i> | 332 | 102 | 69.9 |
| Sas | <i>Salmo salar</i> | 37 | 41 | 47.9 |
| Sce | <i>Scardinius erythrophthalmus</i> | 87 | 65 | 32.7 |
| Sqc | <i>Squalius cephalus</i> | 418 | 105 | 77.3 |
| Tes | <i>Telestes souffia</i> | 61 | 16 | 74.5 |
| Tit | <i>Tinca tinca</i> | 111 | 66 | 37.6 |

We considered non-environmental absences to be the false presences predicted by the EM model (i.e. the cases where the model predicted a species presence while it was actually absent from the observed calibration set). These absences were then considered as presences and the resulting new vector of presence-absence was modeled as a new calibration vector. The entire procedure was repeated 150 times (see Lauzeral et al. 2012 for more detail). The modelling procedure was implemented in R 2.13 (R Development Core Team), using the packages GAM, GBM, MASS, random Forest, ROCR and rpart. The R code of the IEM is provided in Supplementary material Appendix 1, and a working example is provided in Supplementary material Appendix 2.

To evaluate the prediction variability of the two non-deterministic statistical methods (i.e. GBM and RF), we ran the EM 150 times for each species. We observed that the maximum number of sites with variable predictions was reached in less than 30 runs and that less than 3% of the sites had variable predictions. We thus considered that our IEM model had stabilized when less than 3% of the sites provided variable predictions in 30 successive iterations. At the end of

the 150 iterations, predictions were stable, therefore providing a potential distribution of the species.

Variability among the six SDM predictions was evaluated at each iteration of the IEM calculation process. Following Thuiller (2004), we performed a standardized principal component analysis (PCA) on the data matrix consisting of the 6 probability-of-presence vectors at the 191 test sites, and we evaluated the consensus among the predictions by calculating the percentage of variance accounted for by the first axis of the PCA.

Comparing IEM and EM

The IEM was run on the 31 fish species and the results obtained after the 150 iterations were compared to those obtained using the classical EM (i.e. those obtained at the end of the first IEM iteration).

For each of the 31 fish species, we first evaluated both EM and IEM on the 191 independent test sites by measuring four indicators of predictive performance. We used one threshold-independent (AUC) and three threshold-dependent indicators of performance: the percentage of mispredicted sites (i.e. both false absences and false presences); the Kappa index, a chance corrected measure; and the True Skill Statistic (TSS), which takes into account both the sensitivity and specificity of the models.

We determined the predicted species richness per site by summing occurrence predictions of the 31 species. The efficiencies of EM and IEM in predicting species richness were compared by assessing the relationship between observed and predicted species richness on the test sites. Finally, we compared the predicted species assemblages to the observed assemblages by calculating a Jaccard similarity index between observed and predicted assemblages. Wilcoxon tests were used to make pairwise comparisons between the two modelling methods.

Results

Prevalence in the calibration data set was highly variable among fish species, ranging from 0.036 for burbot *Lota lota* to 0.55 for stone loach *Barbatula barbatula*. Detectability also showed high variability (i.e. 4-fold variation among species) and ranged from 20.1% (*Cyprinus carpio*) to 89.4% (*Barbus meridionalis*) (Table 1). This large variation in species detectability suggested large variations in the occurrence of methodological absences in the calibration data set.

IEM modelling

For all 31 species, the predictions stabilized after 2 to 35 iterations, depending on the species. However, for two species (i.e. bleak *Alburnus alburnus* and gudgeon *Gobio gobio*), model predictions showed a temporary stabilization (less than 6 sites had variable predictions during more than 10 iterations) before stabilizing permanently. For these species, the final predictions differed from intermediate stable predictions for less than 5% of the sites.

After a few iterations, the 6 different methods provided consensual predictions for the 191 test sites. Indeed, at the first iteration (i.e. the EM), the mean percentage of variance accounted for by the first axis of the PCA was 85.6%. Using IEM, consensus increased after 15 iterations up to 93.2%, and then reached a plateau up to the end of the iterative procedure.

Predictive performance

At the species level, IEM significantly reduced false absences (Wilcoxon test, $p < 0.001$, Fig. 2). Two out of the three threshold dependent indices showed that IEM performed better than EM (Kappa: Wilcoxon test, $p < 0.01$, TSS: Wilcoxon test, $p < 0.001$) (Fig. 3). In particular, the Kappa index calculated for IEM resulted in a good score (> 0.6) for 12 species and a moderate score (between 0.4 and 0.6) for 17 species. Our predictions were thus good-to-moderate for 29 out of the 31 species (i.e. 94% of the species). EM performance was clearly lower with only 8 species reaching a Kappa score above 0.6 and 23 species for which the predictions were good-to-moderate (i.e. 74% of the species). Species that benefited most from iterations were some of the rare ones (nase *Chondrostoma nasus*; salmon *Salmo salar* and stickleback *Gasterosteus aculeatus*). Species that did not benefit from iterations were some of the most common (stone loach *B. barbatula* and gudgeon *G. gobio*). More generally, poorly detectable species benefited from iterations: the variation of the Kappa index and of the TSS significantly decreased as species detectability increased ($p < 0.05$ and $p < 0.01$ respectively) (Fig. 4). Replacing non-environmental absences by presences nevertheless increased false presences significantly (Wilcoxon test, $p < 0.001$, Fig. 2). Lowering false absences and increasing false presences led to a percentage of errors that did not differ significantly between EM and IEM (Wilcoxon test, $p = 0.18$; $15.9 \pm 7.3\%$ and $15.3 \pm 6.6\%$ respectively). The AUC, a threshold independent index, showed a significant ($p < 0.01$) but limited decrease using IEM (Fig. 3), which did not depend on species detectability (Fig. 4).

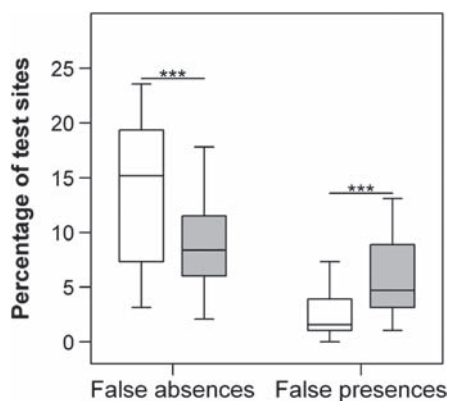


Figure 2. Proportion of the two types of mispredicted sites across all species after the first iteration (EM, white) and at the end of the process (IEM, grey). False absence means that the species was detected but was predicted as absent. False presence means that the species was not detected but was predicted as present.

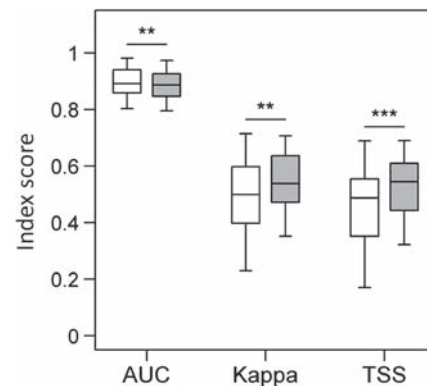


Figure 3. Model evaluation of the two SDM approaches, EM (white) and IEM (grey), using a threshold independent measure (AUC) and two threshold dependent measures (Kappa and TSS) across all species.

At the assemblage level, the relationship between observed and predicted richness was highly significant for EM ($r^2 = 0.90$; $p < 0.001$, Fig. 5a) which reliably predicted species richness in sites containing few species. However, EM tended to underestimate species richness at sites with the highest diversity, where it predicted on average 68% of the observed richness. This bias was reduced using IEM, where predicted species richness at the same sites was on average 89% of the observed one, with a highly significant relationship between observed and predicted values ($r^2 = 0.91$; $p < 0.001$; Fig. 5b). Finally, IEM also increased the similarity (i.e. Jaccard index) between observed and predicted fish assemblages from 0.45 ± 0.25 to 0.50 ± 0.26 (Wilcoxon test, $p < 0.001$).

Discussion

Modeling methods are often tested on virtual species and then validated using real-world species (Bean et al. 2012, Hanberry et al. 2012), because virtual species never capture the real-world complexity that influences species distributions (Meynard and Kaplan 2013).

Here we report the ability of IEM to model real-world species distributions. This study complements a simulation study (Lauzeral et al. 2012) where it was shown that IEM approaches improved EM model predictions. We provide evidence that IEM better predicts species distributions than EM for 27 out of the 31 real species considered in this study. This is particularly true for the less detectable species that are characterized by abundant false absences. These species are often difficult to predict (Pearson et al. 2007, Wisz et al. 2008), and the IEM thus provides a powerful alternative to existing methods. Most of species having a higher detectability were also well predicted using IEM, with similar predictions than those obtained using EM. This testifies that the IEM provides a way to consider almost all species irrespective of their detectability. Although EM reliably predicted the species richness in sites containing few species, it under-predicted species richness at sites with the highest diversity, in which the detectability of some species is low (Kéry and Schmid 2006). These hard-to-detect species increase the

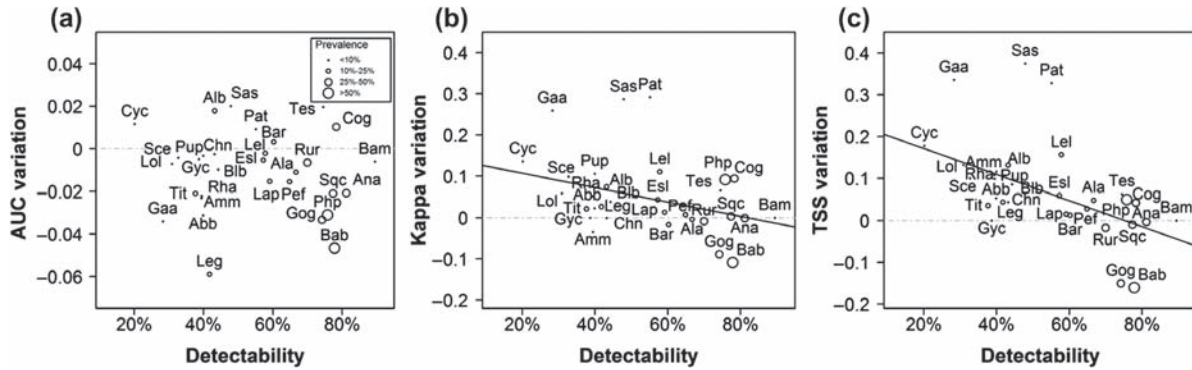


Figure 4. Effect of species detectability on the differences in performance indices between IEM and EM (i.e. IEM – EM) for the 31 species (see Material and methods): (a) AUC; (b) Kappa index; (c) TSS. Lines indicate significant linear relationships ($p < 0.05$). The size of the dots is proportional to the prevalence of the species in the calibration data set. Species codes as in Table 1.

percentage of methodological absences and thus decrease the prediction accuracy of the models. IEM therefore provides a more realistic prediction of both observed richness and assemblage composition, irrespective of the diversity of the sites. This corroborates the idea that the IEM is efficient in filling in methodological absences.

Some differences in prediction performance can occur depending on the species. A general pattern is that species with small geographical extent and strict ecological requirements (i.e. habitat specialists) yield models with higher accuracy than habitat generalists with larger areas of occupancy (Kadmon et al. 2003, Segurado and Araújo 2004, Hernandez et al. 2006, Franklin et al. 2009). Such a tendency also holds for IEM and two out of the 12 species having a prevalence higher than 0.65 did not benefit from iterations (i.e. gudgeon *G. gobio* and stone loach *B. barbatula*, occurring in 73.3 and 66.0% of the test data sites, respectively).

For difficult-to-detect and rare species, low detectability and low prevalence generally decrease the consensus between modeling methods in EM (Pearson et al. 2007, Wisz et al.

2008). Such a tendency was lowered by the iterative process of IEM that increases consensus between modeling methods (Lauzeral et al. 2011). Among 10 poorly-detectable species (i.e. detectability < 0.4), only two (black bullhead *Ameiurus melas* and ruffe *Gymnocephalus cernua*) did not markedly benefit from iterations. For these two species, too much information about their environmental preferences was probably missing due to their rarity in the calibration dataset (Table 1), and the percentage of non-environmental absences exceeded a critical percentage making the IEM unable to fill the gaps in the dataset as also observed in virtual species (Lauzeral et al. 2012).

The ability of IEM to deal with both low prevalence and non-environmental absences allows reliable predictions to be made of the distribution of a wide range of species, and hence extends the usefulness and range of application of EMs. For instance, it might provide more robust estimates of the potential distribution of species that are not in equilibrium with their environment, such as non-native species. Indeed, since many non-native species have been recently

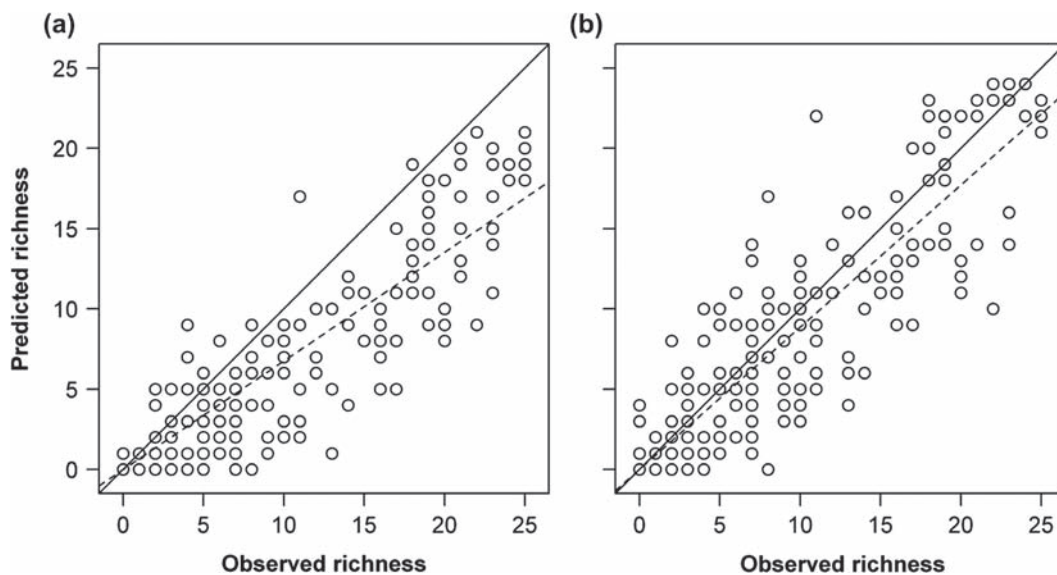


Figure 5. Relationship between observed and predicted fish species richness in the 191 test sites: (a) EM; (b) IEM). The dashed line represents the linear relationship between observed and predicted richness ($y = 0.68x$, $r^2 = 0.90$, $p < 0.001$ for EM; $y = 0.89x$, $r^2 = 0.91$, $p < 0.001$ for IEM). The solid line represents the perfect fit line ($y = x$).

introduced and do not occupy their whole potential niche in their exotic range (Villéger et al. 2011, Lauzeral et al. 2012), a substantial part of the absences in the exotic range are hence non-environmental. In addition, the spatial range of endangered species extirpated from a substantial part of their historical niche is challenging to reconstruct (Di Domenico et al. 2012), and here the IEM offers an efficient way to address this issue. This is because it can be used to deal with absences linked to local extirpations through human or natural disturbances.

Finally, the efficiency of IEM at dealing with both common and rare species allows all the species within a community to be taken into consideration. Aggregating all individual species predictions will hence permit predictions of community changes, under various anthropogenic pressures or future scenarios. Such a community approach will enable changes in community metrics (e.g. species richness, functional diversity) to be predicted and incorporated into species conservation management.

Acknowledgements – This study was supported by the BIOFRESH European project (FP7-ENV-2008). We are indebted to the French National Agency for Water and Aquatic Environment (Onema) for providing fish data. EDB is part of the ‘Laboratoires d’Excellence’ (LABEX) entitled TULIP (ANR-10-LABX-41) and CEBA (ANR-10-LABX-25-01).

References

- Araújo, M. B. and New, M. 2007. Ensemble forecasting of species distributions. – *Trends Ecol. Evol.* 22: 42–47.
- Bean, W. T. et al. 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. – *Ecography* 35: 250–258.
- Buisson, L. et al. 2008a. Modelling stream fish species distribution in a river network: the relative effects of temperature versus physical factors. – *Ecol. Freshwater Fish* 17: 244–257.
- Buisson, L. et al. 2008b. Climate change hastens the turnover of stream fish assemblages. – *Global Change Biol.* 14: 2232–2248.
- Butchart, S. H. M. et al. 2010. Global biodiversity: indicators of recent declines. – *Science* 328: 1164–1168.
- Caissie, D. 2006. The thermal regime of rivers: a review. – *Freshwater Biol.* 51: 1389–1406.
- Daufresne, M. and Boet, P. 2007. Climate change impacts on structure and diversity of fish communities in rivers. – *Global Change Biol.* 13: 2467–2478.
- Di Domenico, F. et al. 2012. *Buxus* in Europe: Late Quaternary dynamics and modern vulnerability. – *Perspect. Plant Ecol.* 14: 354–362.
- Fagan, W. F. 2002. Connectivity, fragmentation, and extinction risk in dendritic metapopulations. – *Ecology* 83: 3243–3249.
- Farber, O. and Kadmon, R. 2003. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. – *Ecol. Model.* 160: 115–130.
- Fausch, K. D. et al. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. – *BioScience* 52: 483–498.
- Franklin, J. et al. 2009. Effect of species rarity on the accuracy of species distribution models for reptiles and amphibians in southern California. – *Divers. Distrib.* 15: 167–177.
- Freeman, E. A. and Moisen, G. G. 2008. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. – *Ecol. Model.* 217: 48–58.
- Grenouillet, G. et al. 2011. Ensemble modelling of species distributions: the effects of geographical and environmental ranges. – *Ecography* 34: 9–17.
- Hanberry, B. B. et al. 2012. Pseudo-absence generation strategies for species distribution models. – *PLoS One* 7: e44486.
- Hernandez, P. A. et al. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. – *Ecography* 29: 773–785.
- Hirzel, A. H. et al. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? – *Ecology* 83: 2027–2036.
- Hugueny, B. 1989. West-African rivers as biogeographic islands – species richness of fish communities. – *Oecologia* 79: 236–243.
- Kadmon, R. et al. 2003. A systematic analysis of factors affecting the performance of climatic envelope models. – *Ecol. Appl.* 13: 853–867.
- Kéry, M. and Schmid, A. 2006. Estimating species richness: calibrating a large avian monitoring program. – *J. Appl. Ecol.* 43: 101–110.
- Kéry, M. et al. 2010. Predicting species distributions from checklist data using site-occupancy models. – *J. Biogeogr.* 37: 1851–1862.
- Lauzeral, C. et al. 2011. Identifying climatic niche shifts using coarse-grained occurrence data: a test with non-native freshwater fish. – *Global Ecol. Biogeogr.* 20: 407–414.
- Lauzeral, C. et al. 2012. Dealing with noisy absences to optimize species distribution models: an iterative ensemble modelling approach. – *PLoS One* 7: e44486.
- Liu, C. et al. 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. – *J. Biogeogr.* 40: 778–789.
- Lobo, J. M. 2008. More complex distribution models or more representative data? – *Biodivers. Inform.* 5: 14–19.
- Lobo, J. M. et al. 2010. The uncertain nature of absences and their importance in species distribution modelling. – *Ecography* 33: 103–114.
- Marmion, M. et al. 2009. Evaluation of consensus methods in predictive species distribution modelling. – *Divers. Distrib.* 15: 59–69.
- Meynard, C. N. and Kaplan, D. M. 2013. Using virtual species to study species distributions and model performance. – *J. Biogeogr.* 40: 1–8.
- Moritz, C. et al. 2008. Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. – *Science* 322: 261–264.
- Mouton, A. M. et al. 2009. Prevalence-adjusted optimisation of fuzzy models for species distribution. – *Ecol. Model.* 220: 1776–1786.
- Murphy, B. R. and Willis, D. W. 1996. *Fisheries techniques*. – American Fisheries Society.
- New, M. et al. 2002. A high-resolution dataset of surface climate over global land areas. – *Clim. Res.* 21: 1–25.
- Pearson, R. G. et al. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. – *J. Biogeogr.* 34: 102–117.
- Pineda, E. and Lobo, J. M. 2009. Assessing the accuracy of species distribution models to predict amphibian species richness patterns. – *J. Anim. Ecol.* 78: 182–190.
- Poulet, N. et al. 2011. Time trends in fish populations in metropolitan France: insights from national monitoring data. – *J. Fish Biol.* 79: 1436–1452.
- Rowe, R. J. et al. 2010. Range dynamics of small mammals along an elevational gradient over an 80-year interval. – *Global Change Biol.* 16: 2930–2943.
- Santika, T. 2011. Assessing the effect of prevalence on the predictive performance of species distribution models using simulated data. – *Global Ecol. Biogeogr.* 20: 181–192.

- Segurado, P. and Araújo, M. B. 2004. An evaluation of methods for modelling species distributions. – *J. Biogeogr.* 31: 1555–1568.
- Thuiller, W. 2004. Patterns and uncertainties of species' range shifts under climate change. – *Global Change Biol.* 10: 2020–2027.
- Villéger, S. et al. 2011. Homogenization patterns of the world's freshwater fish faunas. – *Proc. Natl Acad. Sci. USA* 108: 18003–18008.
- Ward, G. et al. 2009. Presence-only data and the EM algorithm. – *Biometrics* 65: 554–563.
- Wilcove, D. S. et al. 1998. Quantifying threats to imperiled species in the United States. – *Bioscience* 48: 607–615.
- Wisz, M. S. et al. 2008. Effects of sample size on the performance of species distribution models. – *Divers. Distrib.* 14: 763–773.
- Yackulic, C. B. et al. 2013. Presence-only modelling using MAXENT: when can we trust the inferences? – *Methods Ecol. Evol.* 4: 236–243.

Supplementary material (Appendix ECOG-00554 at <www.ecography.org/readers/appendix>). Appendix 1–3.