



RESEARCH
PAPER

Choices of abundance currency, community definition and diversity metric control the predictive power of macroecological models of biodiversity

Grégoire Certain^{1*}, Carsten F. Dormann² and Benjamin Planque¹

¹Institute of Marine Research, Postboks 6404, 9294 Tromsø, Norway, ²Biometry and Environmental System Analysis, University of Freiburg, Tennenbacher Strasse 4, 79104 Freiburg, Germany

ABSTRACT

Aim This study focuses on the influence of methodological choices on the predictive performance of macroecological models (MEMs), i.e. statistical models designed to predict patterns of biodiversity using environmental predictors. We emphasize the influence of three methodological choices: (1) the choice of the currency in which the abundance of each species is measured, i.e. numbers of individuals or biomass; (2) the rules used to define the species assemblage under focus, i.e. broad communities or refined ecological guilds; and (3) the influence of rare over common species in the biodiversity measure.

Location The effects of these choices are investigated using an exhaustive dataset on the fish fauna of the continental shelf of the Barents Sea.

Methods We conducted an analysis of 220 models resulting from all possible combinations of the three methodological choices. For each, we evaluated the predictive performance through an iterative cross-validation process.

Results All methodological choices we investigated strongly affected the predictive performance of MEMs. High predictive performances were obtained when using biomass instead of numbers of individuals, when focusing on narrow ecological guilds composed of species sharing the same ecological traits and when using diversity measures that give high weight to rare species.

Main conclusions We recommend that future projections of biodiversity pay more attention to abundance currency, ecological homogeneity of focal species assemblages and the diversity metric used, and systematically investigate these methodological choices prior to producing biodiversity forecasts. Splitting a whole set of species into ecological guilds appears to be a promising practice, leading to a selected set of MEMs with high predictive performances and more detailed forecasts on the fate of diversity.

Keywords

Barents Sea, biodiversity, ecological guild, fish, forecast, macroecological models, predictive power.

*Correspondence: Grégoire Certain, Institute of Marine Research, Postboks 6404, 9294 Tromsø, Norway.
E-mail: gregoire.certain@imr.no

INTRODUCTION

Forecasting patterns of biodiversity under climate change has become a major field of research in applied ecology (Thuiller, 2003; Thomas *et al.*, 2004; Ferrier & Guisan, 2006; Botkin *et al.*, 2007; Cheung *et al.*, 2009; Sommer *et al.*, 2010; Austin & Van Niel, 2011; Mokany & Ferrier, 2011). Models employed for that purpose can be purely correlative (Sommer *et al.*, 2010) or

may include a mechanistic part (Dormann *et al.*, 2012) to account for limitations due to ecological processes (Boulangeat *et al.*, 2012a; Mokany *et al.*, 2012). These models may be called macroecological models (MEMs; Guisan & Rahbek, 2011) and their projections of biodiversity under climate change scenarios strongly rely on the statistical relationships between biodiversity patterns and a set of environmental predictors.

Biodiversity forecasts are often used to draw the attention of stakeholders to the magnitude of the biodiversity crisis (see, for example, Worm *et al.*, 2006; Sommer *et al.*, 2010). However, the predictive capabilities of MEMs are often poorly known and the accuracy of their projections will only be determined decades from now. An objective assessment of the predictive power of MEMs and a better understanding of what controls their predictive accuracy are therefore critically needed. Several modelling approaches have been developed and combined to reproduce reliably observed data (Thuiller, 2003; Wood, 2006; Elith *et al.*, 2008; see Beale *et al.*, 2010, for a comparison of spatially explicit methods), but the selection and correct implementation of these methods should be the last stage of the analytical process. When biodiversity is the focal response, several important methodological choices have to be made with respect to the size, type and content of the biological dataset, the set of environmental covariates, and the choice of the biodiversity metric to be modelled before applying a method. Depending on these choices, the predictive power and the resulting forecast of MEMs are likely to vary. In the field of species distribution models, for example, forecasts may vary greatly merely as a consequence of the somewhat arbitrary choice of different sets of predictors (Synes & Osborne, 2011).

In this study, we conducted a thorough investigation of three important methodological choices, searching for those that maximize the predictive power of MEMs and evaluating how they affect the conclusions drawn from biodiversity forecasts. These three choices relate to: (1) the currency used to measure species abundance; (2) the choice of the species assemblage for which a biodiversity metric is computed; and (3) the choice of the biodiversity metric itself. Changes in currency alter the frequency distributions of species, i.e. the species abundance distribution (SAD), which in turn affects the value of any biodiversity metric other than species richness (Morlon *et al.*, 2009; Magurran & Henderson, 2012). The species composition of the modelled community might affect the biodiversity–environment relationship, leading to different numerical models and predictions. Finally, diversity metrics can be sorted according to the weight they give to frequent over rare species (Hill, 1973) and the predictive power of MEMs might change along this gradient.

Any possible realization of these three methodological choices will henceforth be referred to as a ‘modelling option’. The current study investigates the predictive power of 220 such modelling options, encompassing conventional choices for abundance currency, community composition and biodiversity metrics. The comparison is based on a comprehensive dataset of 81 fish species collected over 5 years by bottom trawling during an ecosystem survey in the Barents Sea (Olsen *et al.*, 2011). This data set has several features that make it well suited for such a study. First, the data have a fine spatial resolution and can be viewed as point-samples of demersal communities. Second, over the years a large number of samples have been collected, allowing us to split the data into train and test subsets. Third, data have been simultaneously recorded in two

abundance currencies, namely number and biomass, allowing a direct comparison without relying on allometric relationships. Fourth, an extensive ecological knowledge on the ecology of the 81 fish species considered in our study has accumulated and been compiled, allowing for an easy classification of the sampled species into ecological guilds based on different species attributes.

METHOD

General approach

Analyses were based on three main data matrices: a $N \times S$ ‘species’ table summarizing information on the relative abundance of the S species across N sampling sites, a $N \times P$ ‘predictor’ table summarizing information on the P environmental predictors across the same N sampling sites and a $S \times T$ ‘trait’ matrix summarizing information about T ecological traits across the S species. The trait matrix was used to construct communities of species according to specific ecological criteria, i.e. guilds. A total of 220 MEMs were constructed resulting from individual modelling options with specific choices of (1) abundance currency, (2) guild composition and (3) biodiversity metric. For each modelling option, the spatio-temporal pattern of the biodiversity metric was modelled as a function of environmental predictors, and the predictive power of the resulting MEM was assessed. The different steps of our methodological approach are summarized in Fig. 1.

The modelling options

We contrasted two commonly used currencies for species abundance: number of individuals and biomass (Morlon *et al.*, 2009). As this choice affects the resulting shape of the SAD, we refer to species individual distribution (SID) when the currency is number and species biomass distribution (SBD) when the currency is biomass.

We considered five ways of forming ecological guilds. The first one corresponded to the most common approach, i.e. a broad taxonomic guild including all species of a given phylum, here fish ($n = 81$). The four other approaches are attempts to split the broad taxonomic guild into ecological guilds whose species share common ecological properties. Using a life-history trait matrix (Table S1 in Supporting Information) gathered from ‘FishBase’ (<http://www.fishbase.org>) and its Norwegian equivalent ‘Fiskipedia’ (<http://www.fiskipedia.no/>), the fish species were grouped according to their habitat (strictly demersal species, $n = 54$, versus species possibly occurring in pelagic environment, $n = 27$), feeding behaviour (piscivorous fishes, $n = 27$, strictly benthivorous species, $n = 29$, and species including planktonic prey in their diet, $n = 25$), reproductive performance (low-fecundity fishes, $n = 42$, versus high-fecundity fishes, $n = 39$) and a cluster analysis of the complete life-history trait table, resulting in three groups of species (termed group 1, $n = 43$, group 2, $n = 19$, and group 3, $n = 19$; see Text S2 for details on the cluster analysis).

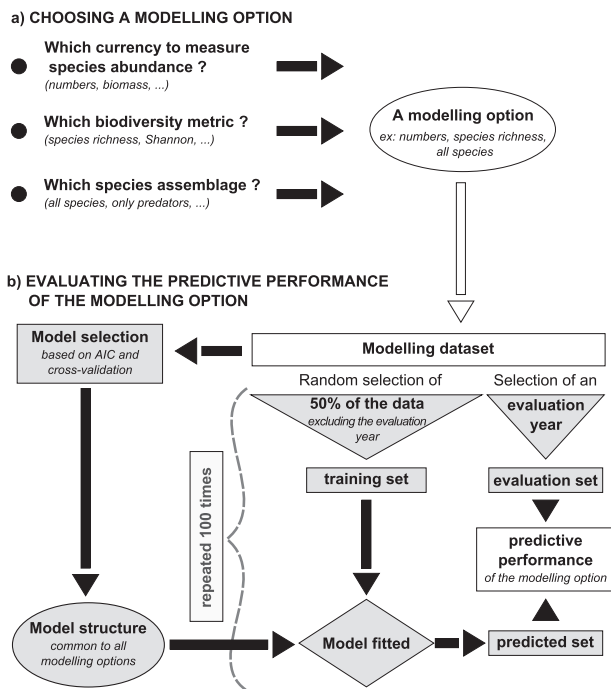


Figure 1 Schematic representation of our methodological approach. (a) Methodological choices leading to the building of modelling options. (b) Evaluation of the predictive performance of a given modelling option (AUC, area under the curve). Note that the ‘model selection’ process referred to in the upper-left box is described in Text S3.

As biodiversity metrics we considered 10 measures taken along the Hill diversity profile (Hill, 1973). A Hill diversity of order q , denoted qD , integrates most of the well-known diversity measures in a single framework: qD equals species richness when $q = 0$; e^H (with H being the Shannon diversity index) when q approaches 1 (the Hill diversity is not defined when $q = 1$; see Hill, 1973); the inverse Simpson index when $q = 2$; and the dominance index when q is approaching infinity. Decreasing q increases the weight of rare species in the biodiversity measures. The Hill diversity profile is simply the graph showing the evolution of qD as q increases from 0 to large values, and is more informative about biodiversity than any individual index (Hill, 1973). The Hill diversity additionally has the advantage of measuring biodiversity as ‘effective species numbers’ (Jost, 2006; Tuomisto, 2010). It is based on species frequencies, which can be measured in different currencies. The value qD is variable for low values of q , while it quickly reaches an asymptote when $q > 2$ (Hill, 1973). In this study, qD was measured with q taking the values of 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75, 1, 2 and 10.

We used the following notation to describe a modelling option: currency/guild/ q . For example ‘SBD/low-fecundity/0.5’ corresponds to the modelling option where the abundance currency for species is biomass, the guild under focus comprises only species with low fecundity and the diversity measure used is ${}^{0.5}D$. The combination of two currencies, 11 guilds (1 broad and

10 splits) and 10 biodiversity metrics resulted in 220 modelling options (Fig. 1a).

MEM formulation and evaluation

Each modelling option corresponds to a statistical MEM that regresses observed biodiversity patterns against a set of environmental predictors. The model formulation was carefully chosen after a thorough model selection process (Text S3), which identified a narrow set of covariates suitable across all modelling options considered. For the sake of simplicity, we decided on a common model formulation for all the modelling options, incorporating the main environmental drivers of our study area, the Barents Sea. The MEM was formulated as a generalized additive model (GAM), which allows the modelling of nonlinear relationships between diversity patterns and environmental drivers. Spatial autocorrelation was weak in model residuals (see Text S3), meaning that we did not have to use spatial GAMs that had an unrealistically long computation time for the needs of our analysis. The model is an additive series of six smooth functions, each specific to one type of covariate in the Barents Sea: bottom depth, bottom temperature, bottom salinity, surface temperature, surface salinity, surface chlorophyll a concentration and bottom current strength. A detailed description of the model formulation and associated diagnostic plots is available in Text S3.

The predictive performance of a modelling option is evaluated through the following process (Fig. 1b): (1) a year is chosen and separated from the data; (2) a ‘training’ dataset is selected using half of the remaining data (without the evaluation year); (3) the model is fitted using that ‘training’ dataset; (4) the fitted model is used to predict the ‘evaluation’ year; (5) the match between observed and predicted values is measured through Pearson’s correlation and through linear regression where observations are the response. For each year in the data, steps 2–5 are repeated 100 times, so as to get 100 evaluations. A perfect model would have a correlation of 1, a slope of 1 and an intercept of 0. As we wished to test whether selecting ecological guilds offers a consistently higher predictive power than randomly assembling species, 100 random guilds were assembled for each of the 10 guild splits, and the evaluation process was applied again, once for each random guild. Random guilds were assembled by drawing randomly the same number of species as in the guild split under focus; for example, a random guild for the demersal guild was built by randomly selecting 54 species from the 81 fish species available. No random splits were available, however, for the whole fish community.

Given a modelling option, an evaluation year and an evaluation criterion, we can compare the set of 100 measures of predictive performance obtained with a defined guild with the corresponding set obtained with a random guild. By subtracting each element of the former from each element of the latter, we obtain 10,000 estimates of the differences in predictive performance between the modelling option under focus and its counterpart based on random guild. In that way we estimated the probability of increased MEM predictive performance

when guilds were assembled according to specific ecological criteria.

Investigating the relevance of guild splitting

To illustrate how biodiversity forecasts can vary depending on the modelling options, we compared the biodiversity patterns predicted from a MEM built on the broad 'fish' community with summed predictions from MEMs based on guild splitting presenting good predictive performance. For that exercise, we are restricted to setting $q = 0$ as it is the only case when the sum of the diversity measure for each guild equals the diversity measure for the broad community. Biodiversity patterns will be predicted for a set of hypothetical situations corresponding to a particular environmental change in the Barents Sea.

Study area and model: demersal fauna of the Barents Sea

The Barents Sea continental shelf (1,600,000 km²) is delimited by the Norwegian and Russian coast, Novaya Zemlya Island, Franz Joseph Land, the Svalbard Archipelago and a shelf break toward the deep ocean in the west (Jakobsen & Ozhigin, 2012). Warm and salty Atlantic waters are brought from the south by the Norwegian Atlantic Current and extend over the western and central parts of the Barents Sea, while cold and fresh Arctic water dominates in the northern part. The area where these two main water masses meet is called the Polar Front. The northern areas are seasonally covered in ice, and in winter the edge of the ice usually follows the oceanic fronts.

The demersal community of the Barents Sea has been sampled each summer (August–September) since 2004 following a regular sampling scheme of stations separated by 30–40 km (Anonymous, 2010). Data from 2004 to 2008 were used in this analysis. At each station, biological samples were taken with a Campelen 1800 bottom trawl towed on double warps (Johannesen *et al.*, 2012). The standard towing time was 15 min at 3 knots, equivalent to a towing distance of 0.75 nautical miles. On board, the catch was sorted by species, counted and weighed, so that numbers and biomasses are documented for each species collected. For each trawl, numbers and biomasses were standardized by towing distance. The dataset we used for this study comprised 2316 sampling stations in which 81 fish species (Table S1) were identified. The number of sampling stations per year ranged between 355 and 538.

RESULTS

Choosing an evaluation year and an evaluation criterion

The results from this study cannot all be presented in the main text. We chose to present them for one evaluation year (2008) and one evaluation criterion (Pearson's correlation), which are well representative of the analyses. The results for all evaluation years and evaluation criteria are available in Appendix S4.

Predictive power of MEMs

Figure 2 displays a few selected examples to show how the predictive power of MEMs from different modelling options can be compared. Figure 2(a) shows that the predictive power of the modelling option SBD/all fishes/ $q = 1$ was higher than the predictive power of the modelling option SID/all fishes/ $q = 1$. Figure 2(b) illustrates that restricting the analysis to the 54 demersal species gave a slightly better result in terms of predictive performance than selecting 54 species randomly, but it was still in the same range. Figure 2(c) shows that the predictive power of the modelling options SID/all fishes decreased as q increased. Figure 2(d) repeats the comparison shown in Fig. 2(a) but with a guild composed of species living partly or fully in the pelagic area. The use of this guild clearly led to an increase in predictive power when compared with a random guild (Fig. 2e). Finally, Fig. 2(f) shows that the predictive power of the modelling options SBD/pelagic reached a maximum for q values between 0.2 and 0.4. These are only a few comparisons among the possible 220 modelling options, but they illustrate the most common patterns: a higher predictive performance is usually reached with SBD-based MEMs and with biodiversity metrics of low- q order. In addition, large heterogeneities of predictive performance are revealed across guilds.

These general patterns can be seen from the comparison between the 220 modelling options (Fig. 3, Appendix S4). The predictive power of SBD-based MEMs was generally higher than that of the corresponding SID-based MEMs. Six out of 10 ecological guilds (pelagic, ichtyophagous, low fecundity, high fecundity, group 1 and group 3) led to MEMs with a higher predictive power compared with a random assemblage of fish species. A less pronounced improvement was observed for the group 2 and the planktophagous guild, and for the demersal and strictly benthivorous guilds predictive performances were no better than with random guilds. Predictive power also changed along the diversity profile, according to one of the two patterns illustrated in Fig. 2(c) and (f): either a strictly decreasing predictive power with increasing q or a dome-shaped variation with an optimal predictive power for q values between 0.1 and 0.5, the latter being more common in SBD-based MEMs than in SID-based MEMs.

In summary, when taking 2008 as the evaluation year and the Pearson correlation as the evaluation criterion, the highest predictive power was achieved when the abundance currency is biomass, when fish guilds were based on fecundity and when the diversity metric was the Hill diversity of order q between 0.2 and 0.5. The same general pattern with some moderate variations across years and guilds is observed with the other evaluation years and when using the slope and intercept of the linear regression as evaluation criterion (Appendix S4).

Choosing a relevant model for fish diversity in the Barents Sea

Our investigation across the 220 modelling options revealed that splitting the fish community into low- and high-fecundity

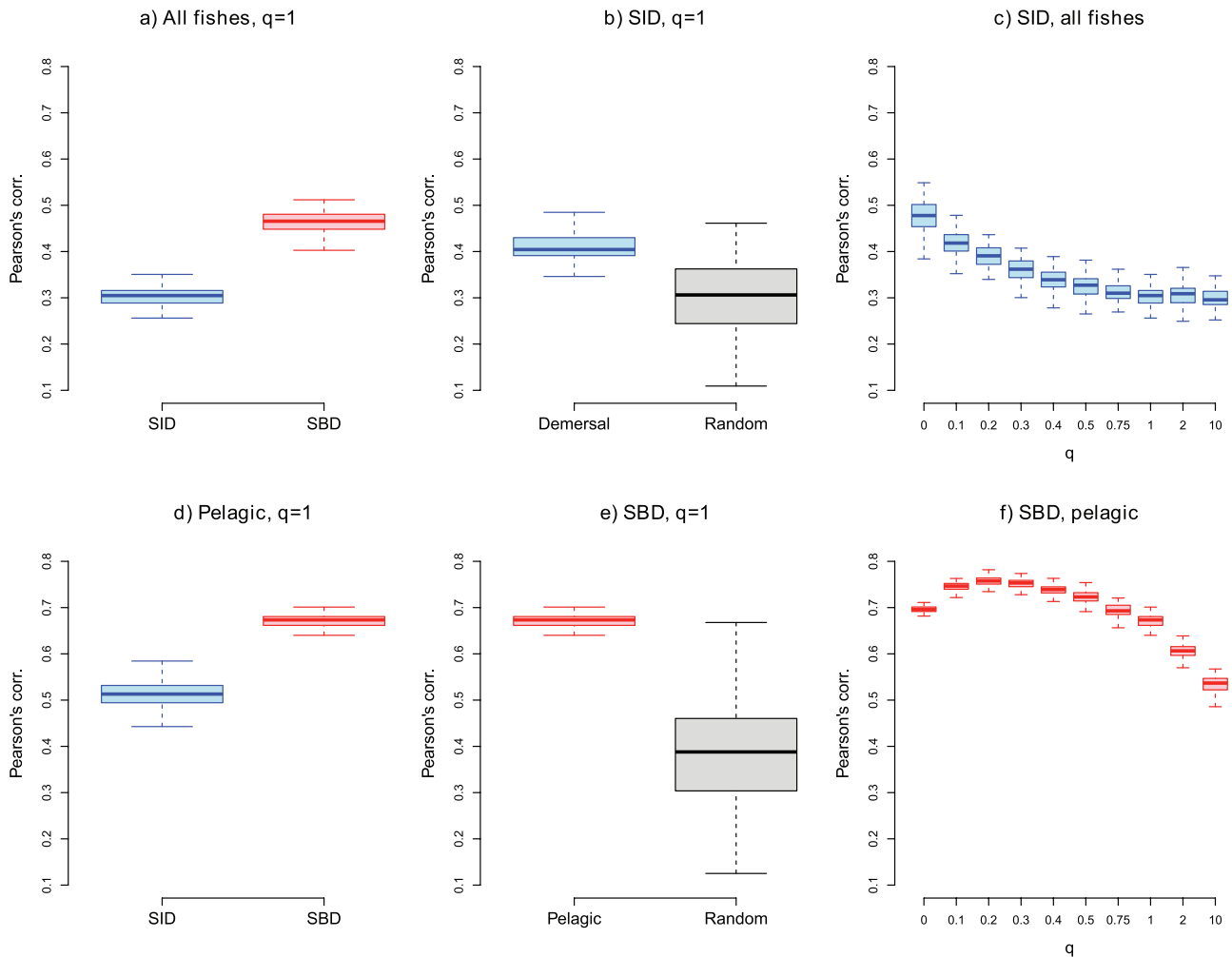


Figure 2 Predictive power (y -axis) of some selected modelling option (x -axis). Boxplots show the extent of 100 predictive power estimates (boxes show the median and quartiles, whiskers extend to the minimum and maximum values). The upper row (a, b, c) focuses on the all fishes and demersal guilds, the lower row (d, e, f) focuses on the pelagic fish guild. The first column (a, d) shows example of results when contrasting two abundance currency (left, species individuals distribution SID; right, species biomass distribution SBD) and the second column compares modelling options based on defined (left) versus random (right) communities; the third column (c, f) shows the evolution of predictive power along the Hill diversity profile. The order of Hill's diversity is represented by the parameter q .

species provided two MEMs (SBD/low fecundity/0.1–0.5 and SBD/high fecundity/0.1–0.5) with high predictive performances. They were used to predict biodiversity patterns for the current time period (Fig. 4, Text S3). The maps show that both guilds exhibited a very different and almost reversed spatial pattern of biodiversity, with high biodiversity of the low-fecundity guild clearly associated with the north-eastern Arctic region of the Barents Sea, while biodiversity for the high-fecundity guild peaked in the south-western Atlantic area. These differences can be explained by the shapes of some of the diversity–predictor relationships (Fig. 5a) that are clearly reversed between the low- and high-fecundity guild, while the shape of the diversity–predictor relationship obtained for the all-fishes guild is simply a mixture of the response of the two fecundity-based guilds.

Biodiversity forecast: total versus guild split approach

To further illustrate the advantage of using guild splitting when forecasting diversity patterns under an environmental change scenario, we contrasted species richness ($q = 0$) predictions from a MEM based on the all-fishes guild with the summed predictions of two MEMs focusing on the high- and low-fecundity guilds. We focused on species richness because it is insensitive to the choice of the abundance currency and it can be summed across split guilds to ease the comparison between split and non-split options. The environmental scenarios we used for the comparison were based on the qualitative assessment of predicted changes for the Arctic area (ACIA, 2004) that suggests an increase in air temperature of approximately 4–8 °C and an

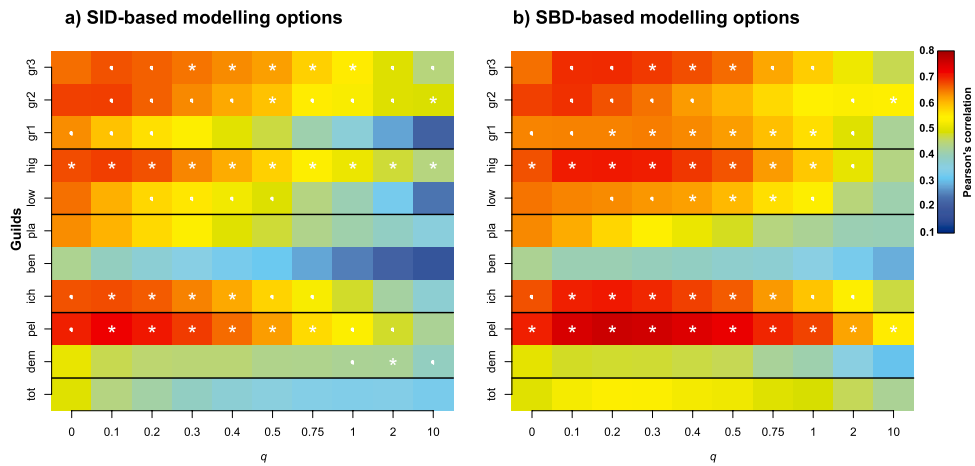


Figure 3 Expected predictive power (median of the 100 estimates) for the 220 modelling options. Left panel, species individuals distribution (SID)-based macroecological models (MEMs). Right panel, species biomass distribution (SBD)-based MEMs. *y*-axis, guild definition, *x*-axis, Hill diversity profile. White stars show modelling options where the probability that the predictive power is higher than if guilds were randomly assembled is < 0.05. White dots correspond to the same probability < 0.1. Abbreviations for guilds: tot, all fishes; dem, demersal fishes; pel, pelagic fishes; icht, ichtyophagous fishes; ben, benthivorous fishes; pla, planktivorous fishes; low, low-fecundity fishes; hig, high-fecundity fishes; gr1, group 1; gr2, group 2; gr3, group 3.

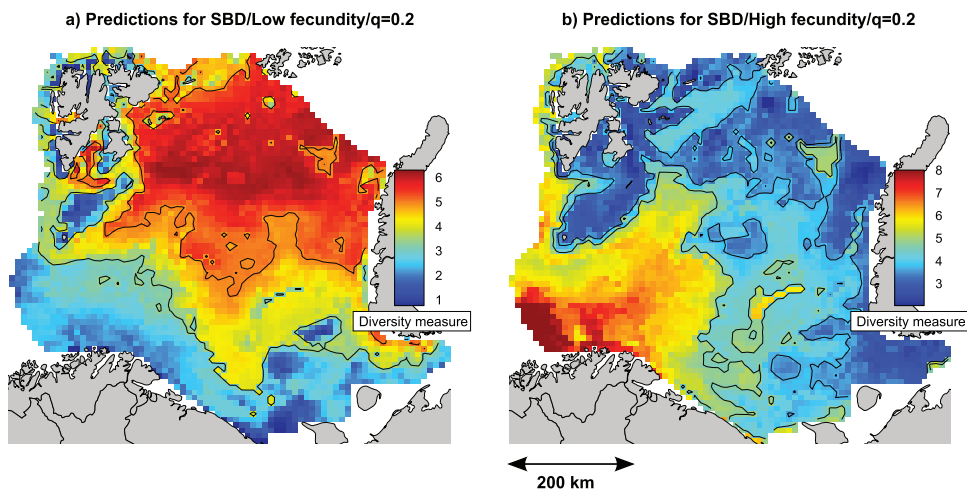


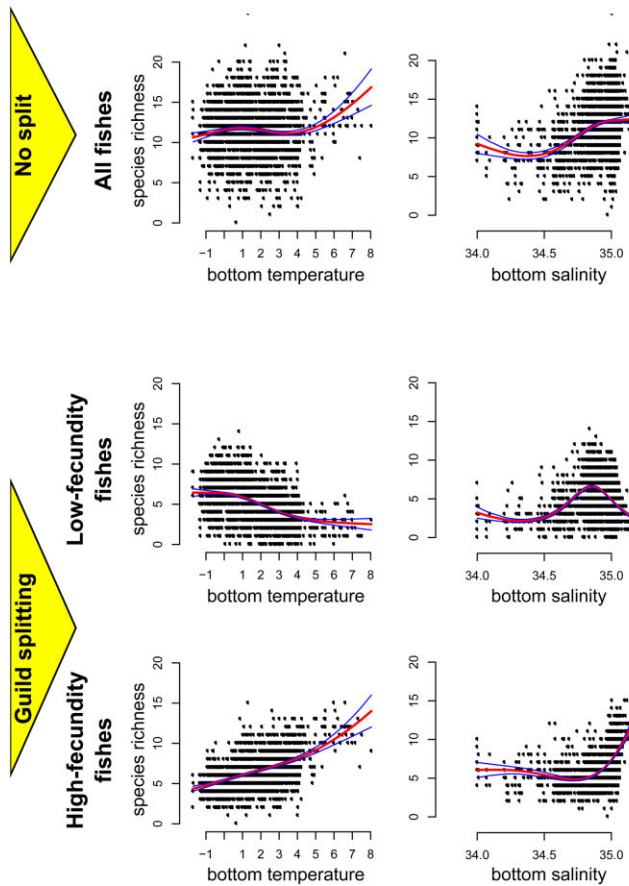
Figure 4 Predicted biodiversity patterns (Hill diversity of order 0.2) for the current time period (2004–08) for the low-fecundity fish community (left panel) and the high-fecundity fish community (right panel).

increase in ice melting and precipitation that will reduce salinity. Since reliable quantitative predictions for sea temperature and salinity are not available at our spatial resolution, we simply built a set of four scenarios by adding 1, 2, 3 or 4 °C to the sea temperature fields and reducing salinity fields by 0.05, 0.1, 0.15 and 0.2‰, respectively. We stress that these scenarios are hypothetical, and that their purpose is to compare the behaviour of the modelling approaches. Figure 5(b) shows the results obtained for scenario 4, and the results for all scenarios are available in Fig. S5.

Focusing only on the total species richness, predictions from the total and split approaches produced similar spatial patterns (Fig. 5b), with an important increase predicted in the

south-western Barents Sea. With the guild split, however, the predicted increase saturates more quickly (Fig. 5b). More importantly, guild splitting allows us to follow the fate of species richness for low- and high-fecundity fishes separately, revealing an important reconfiguration of the fish community in the central and northern Barents Sea that could not be predicted if the analysis was restricted to a single MEM. As low- and high-fecundity diversity exhibit contrasting responses to temperature (Fig. 5a), the apparent stability of species richness in the central and northern Barents Sea predicted by the total model results from a large decrease in the number of low-fecundity species compensated by an increase in the number of high-fecundity species.

a) Relationship between species richness (y-axis) and environmental parameters (x-axis)



b) Predicted change in species richness under environmental change (scenario 4)

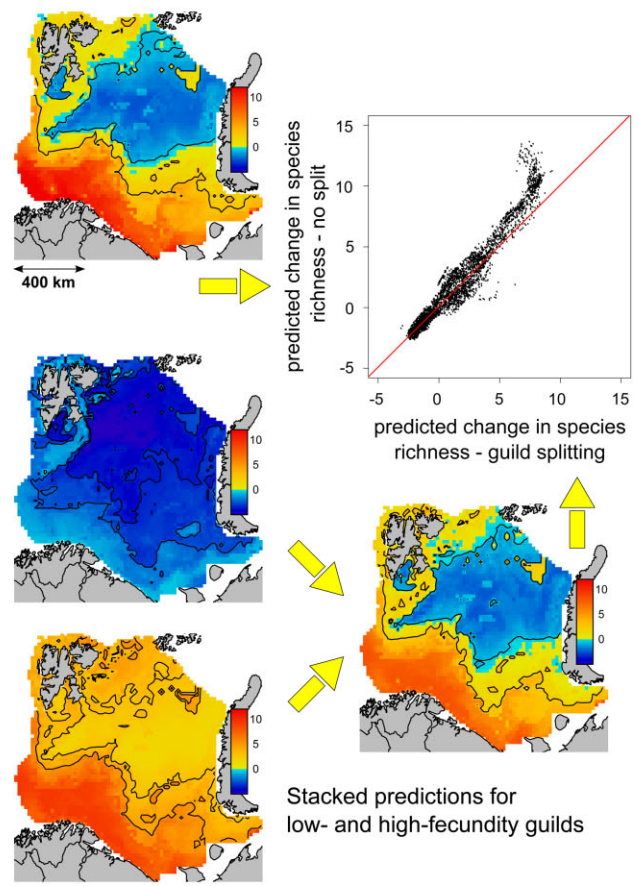


Figure 5 Expected change in species richness under environmental change (scenario 4) as predicted by two macroecological models (MEMs) for low- and high-fecundity fishes and one MEM focusing on all fishes: (a) example of statistical relationships between biodiversity and environmental predictors modelled by the MEM; (b) expected change in species richness according to the all-fishes and guild-split approaches.

DISCUSSION

Methodological considerations

To test the effects of modelling choices on predictive performance, 220 different modelling options were evaluated. The evaluation process was thorough, testing several evaluation years separately and using three evaluation criteria. Our results are consistent whatever the evaluation year and statistic chosen (Appendix S4). However, our evaluation process is bounded to our study area and reveals nothing about the predictive ability of our MEM in other areas or for communities other than fishes in the Barents Sea. The fact that we used non-spatially explicit models is a limitation. The investigation of the spatial structure in the data and in the model residuals (Text S3) revealed that response variables were strongly correlated in space over large scales (up to 200–300 km) and in time (across years). Model residuals were only moderately spatially correlated, but at a smaller spatial scale (*c.* 50–100 km), which corresponds to clusters of two or three sampling stations. Furthermore the model

residuals were no longer correlated through time, which suggests that our model was capturing well the large-scale patterns and year-to-year variability of diversity. We think that our use of non-spatial models does not invalidate the result of our comparison across modelling options. However, once a set of suitable MEMs has been identified, the implementation of their spatially explicit counterpart, if possible, should constitute a natural follow-up.

Similarly, for the sake of simplicity, we restricted our analysis to purely correlative MEMs. But our results are also relevant for their semi-mechanistic counterpart, even though the latter may achieve higher predictive performance due to their better account of ecological processes.

Appropriate methodological choices improve the relevance of MEMs

Many studies concerned with the prediction of biodiversity patterns focus on the development of modelling techniques (Thuiller, 2003; Gotelli *et al.*, 2009; Guisan & Rahbek, 2011;

Mokany & Ferrier, 2011). Our work clearly demonstrates that simple methodological choices such as abundance currency, biodiversity metric and community assemblage also lead to different models with different predictions and varying predictive powers.

First, the predictive power of MEMs was affected by the choice of the abundance currency. The superiority of SBD-based MEMs was clear for most of the guilds considered (Fig. 3, Appendix S4). Recent theoretical work (Morlon *et al.*, 2009; Henderson & Magurran, 2010; Magurran & Henderson, 2012) revealed that observed differences between SID and SBD in various communities are mediated through the currency upon which the selection pressure operates most. In the case of the Barents Sea, our results suggest that environmental pressure is primarily exerted on biomasses, that is, environmental effects on energy allocation between species might contribute more to the shaping of the species frequencies than stochastic demographic processes controlling the numbers of individuals.

Second, biodiversity metrics giving more weight to rare species (low q values) led to a higher predictive power than biodiversity metrics giving more weight to abundant species. This result emphasizes the role played by rare species in total biodiversity patterns and contrasts somehow with studies focusing on species range (e.g. Jetz & Rahbek, 2002; Lennon *et al.*, 2004), which found that geographical patterns of biodiversity were largely predictable from occurrences of common and widespread species, while rare species were rather erratic, unpredictable elements. Species richness (diversity of order $q = 0$) is usually well predicted, but can be outperformed by diversity metrics with a slightly higher Hill number (i.e. $0.1 \leq q \leq 0.5$), as in the case of the SBD-based MEMs of the pelagic guild, ichtyophagous guild and low- and high-fecundity guilds. Diversity of order $q > 0$ is sensitive to both the number of species present and the shape of the SAD, making it more informative than species richness alone. It is therefore interesting to note that these metrics, despite their higher complexity, can still be predicted with similar or better performance than species richness.

Third, guild splitting led to more detailed prediction of the fate of a community than a single MEM encompassing all species. The relationships between biodiversity and environmental predictors can be very different between guilds. Modelling these differences explicitly, rather than pooling all species, resulted in distinct models predicting distinct biodiversity patterns, a more detailed description of the relationships between environmental predictors and biodiversity, and ultimately more detailed predictions. Furthermore, as the prediction relies on a single model, it may be more sensitive to boundary conditions than a set of carefully chosen guild-based MEMs, as suggested by the differences in predicted species richness between the guild splitting approach and the total approach in the species-rich area. The comparison between the predictions of an all-fishes MEM and the summed predictions of a low-fecundity and a high-fecundity MEM (Fig. 5b) illustrates well the additional information obtained with guild splitting. Without it, the emphasis would be on the increase in species richness in the south-western area. With it, one important piece of information

is added: reconfiguration between the low- and high-fecundity communities in the central and northern Barents Sea. Such information is of crucial interest, as it may trigger large changes within the food web that would stay unnoticed with a single all-fishes MEM.

To be efficient, guild splitting requires extensive investigation: MEMs based on the benthivorous guild, for example, had a rather poor predictive power (Fig. 3), even though they focus on a well-defined guild of organisms that share a precise trophic trait. We do not have a clear explanation for why some groupings such as the fecundity-based split outperformed other ones, but we think this could be a good starting point for further study aiming at understanding the ecology of the demersal community in the Barents Sea. Relevant guild splits may be very difficult to identify a priori, and therefore we recommend the exploration of modelling options as a standard practice for MEM-based projections of biodiversity. Our study shows that it is feasible to carry out such a careful examination, and how to test statistically the efficiency of guild splitting using random splits.

Sometimes the extent to which these methodological investigations have been carried out is not fully documented. Often future projections of biodiversity rely solely on species richness (i.e. $q = 0$) and consider a single species pool mixing hundreds to thousands of species (Diniz-Filho & Bini, 2005; Kreft & Jetz, 2007; Woodward & Kelly, 2008; Sommer *et al.*, 2010; Belmaker & Jetz, 2011). More generally, studies attempting to link biodiversity to environmental conditions consider only one species assemblage (e.g. Lobo & Martín-Piera, 2002; Cayuela *et al.*, 2006; Kaboli *et al.*, 2006; Kreft & Jetz, 2007; Woodward & Kelly, 2008; Algar *et al.*, 2009; La Sorte *et al.*, 2009; Sommer *et al.*, 2010; Dunstan & Foster, 2011) or different species assemblages broadly taxonomically defined (e.g. 'birds', 'reptiles', 'mammals', 'amphibians': Graham *et al.*, 2006; Araújo *et al.*, 2008; Axmacher *et al.*, 2011, Belmaker & Jetz, 2011). Exceptions exist, though. In more localized studies, communities are usually more narrowly defined, either taxonomically (Beck *et al.*, 2011; Ekschmitt *et al.*, 2003) or functionally (Henry *et al.*, 2010). Steinmann *et al.* (2009), for example, made a thorough comparison by contrasting the predictive efficiency of models of species richness for 40 functional groups of perennial herbs. They did not find clear improvements for species richness modelling when species were split into functional groups. Still, they remarked on the potential of such an approach for increasing understanding. Boulangeat *et al.* (2012b) proposed a method based on functional similarity to split a whole community into functional subunits that are appropriate for the purposes of diversity forecasting. It is notable that the most detailed studies in this field focus on plants – well-known, sampled and documented study models. Our work suggests that these approaches are also highly appropriate and can be efficiently implemented in the case of much less well-documented communities. In our study, splitting communities according to a well-defined ecological trait proved to be at least as effective as splitting according to hierarchical clustering of life-history traits. But hierarchical clustering may still be useful when analysing communities of hundreds of species with known traits.

CONCLUSION

Biodiversity metrics that account for the relative frequency of species ($q > 0$) might be more predictable than species richness ($q = 0$), provided that they still give an important weight to rare species ($q < 0.5$). Choices related to abundance currency and guild splitting should be systematically investigated as part of biodiversity modelling. One can expect that guild splitting, i.e. the use of demographically, physiologically or ecologically homogeneous communities, will provide sound and detailed predictions of change in diversity patterns, based on a relatively restricted set of MEMs with high predictive power. Hopefully, such practice will increase the ability of scientists to anticipate the effect of global change on biodiversity.

ACKNOWLEDGEMENTS

First, we wish to thank all the participants in the joint Russian–Norwegian ecosystem survey. Thanks to Sigrid Lind and Frode Vikebø for providing oceanographic data. Thanks to Magnus Aune Wiedemann for providing information on fish life-history traits. Thanks to Edwige Bellier for inspiring discussions about community ecology. Thanks to the Freiburg University, and especially to Gita Benadi, Lara Budic and Frédéric Clément, for providing a stimulating working environment. Thanks to Virginie Ramasco for her skills in R language, and for having carefully proofread the manuscript. Finally, thanks to the editors and referees that evaluated the manuscript and provided useful comments to improve it. G.C. was funded by the BarEcoRe (NFR project 200796/41) project.

REFERENCES

- ACIA (2004) Impacts of a warming arctic: Arctic Climate Impact Assessment. Cambridge University Press. Available at: <http://ACIA.uaf.edu> (accessed 15 February 2012).
- Algar, A.C., Kharouba, H.M., Young, E.R. & Kerr, J.T. (2009) Predicting the future of species diversity: macroecological theory, climate change, and direct tests of alternative forecasting methods. *Ecography*, **32**, 22–33.
- Anonymous (2010) Survey report from the joint Norwegian/Russian ecosystem survey in the Barents Sea in August–September 2009. ISSN 1502-8828, IMR/PINRO. Available at: http://www.imr.no/filarkiv/2010/09/imr-pinro_2-2010_til_web.pdf/nb-no (accessed 1 January 2011).
- Araújo, M.B., Nogués-Bravo, D., Diniz-Filho, J.A.F., Haywood, A.M., Valdes, P.J. & Rahbek, C. (2008) Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography*, **31**, 8–15.
- Austin, M.P. & Van Niel, K.P. (2011) Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography*, **38**, 1–8.
- Axmacher, J.C., Liu, Y., Wang, C., Li, L. & Yu, Z. (2011) Spatial α -diversity patterns of diverse insect taxa in northern China: lessons for biodiversity conservation. *Biological Conservation*, **144**, 2362–2368.
- Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J. & Elston, D. (2010) Regression analysis of spatial data. *Ecology Letters*, **13**, 246–264.
- Beck, J., Brehm, G. & Fiedler, K. (2011) Links between the environment, abundance and diversity of Andean moths. *Biotropica*, **43**, 208–217.
- Belmaker, J. & Jetz, W. (2011) Cross-scale variation in species richness–environment associations. *Global Ecology and Biogeography*, **20**, 464–474.
- Botkin, D.B., Saxe, H., Araújo, M.B., Betts, R., Bradshaw, R.H.W., Cedhagen, T., Chesson, P., Dawson, T., Etterson, J.R., Faith, D.P., Ferrier, S., Guisan, A., Skoldbjorg Hansen, A., Hilbert, D.W., Loehle, C., Margules, C., New, M., Sobel, M.J. & Stockwell, D.R.B. (2007) Forecasting the effects of global warming on biodiversity. *BioScience*, **57**, 227–236.
- Boulangéat, I., Gravel, D. & Thuiller, W. (2012a) Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. *Ecology Letters*, **15**, 584–593.
- Boulangéat, I., Philippe, P., Abdulhak, S., Douzet, R., Garraud, L., Lavergne, S., Lavorel, S., Van Es, J., Vittoz, P. & Thuillier, W. (2012b) Improving plant functional groups for dynamic models of biodiversity: at the crossroads between functional and community ecology. *Global Change Biology*, **18**, 3464–3475.
- Cayuela, L., María, J., Benayas, R., Justel, A. & Salas-rey, J. (2006) Modelling tree diversity in a highly fragmented tropical montane landscape. *Global Ecology and Biogeography*, **15**, 602–613.
- Cheung, W.W.L., Lam, V.W.Y., Sarmiento, J.L., Kearney, K., Watson, R. & Pauly, D. (2009) Projecting global marine biodiversity impacts under climate change scenarios. *Fish and Fisheries*, **10**, 235–251.
- Diniz-Filho, J.A.F. & Bini, L.M. (2005) Modelling geographical patterns in species richness using eigenvector-based spatial filters. *Global Ecology and Biogeography*, **14**, 177–185.
- Dormann, C.F., Schymanski, S.J., Cabral, J., Chuine, I., Graham, C., Hartig, F., Kearney, M., Morin, X., Römermann, C., Schröder, B. & Singer, A. (2012) Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, **39**, 2119–2131.
- Dunstan, P.K. & Foster, S.D. (2011) RAD biodiversity: prediction of rank abundance distributions from deep water benthic assemblages. *Ecography*, **34**, 798–806.
- Ekschmitt, K., Stierhof, T., Dauber, J., Kreimes, K. & Wolters, V. (2003) On the quality of soil biodiversity indicators: abiotic and biotic parameters as predictors of soil faunal richness at different spatial scales. *Agriculture, Ecosystems and Environment*, **98**, 273–283.
- Elith, J., Leathwick, J.R. & Hastie, T. (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**, 802–813.
- Ferrier, S. & Guisan, A. (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, **43**, 393–404.

- Gotelli, N.J., Anderson, M.J., Arita, H.T. *et al.* (2009) Patterns and causes of species richness: a general simulation model for macroecology. *Ecology Letters*, **12**, 873–886.
- Graham, C.H., Moritz, C. & Williams, S.E. (2006) Habitat history improves prediction of biodiversity in rainforest fauna. *Proceedings of the National Academy of Sciences USA*, **103**, 632–636.
- Guisan, A. & Rahbek, C. (2011) SESAM – a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *Journal of Biogeography*, **38**, 1433–1444.
- Henderson, P. & Magurran, A.E. (2010) Linking species abundance distributions in numerical abundance and biomass through simple assumptions about community structure. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 1561–1570.
- Henry, M., Cosson, J.F. & Pons, J.M. (2010) Modelling multi-scale spatial variation in species richness from abundance data in a complex Neotropical bat assemblage. *Ecological Modelling*, **221**, 2018–2027.
- Hill, M.O. (1973) Diversity and evenness?: a unifying notation and its consequences. *Ecology*, **54**, 427–432.
- Jakobsen, T. & Ozhigin, V. (2012) *The Barents Sea: ecosystem, resources, management*, p. 825. Tapir, Trondheim.
- Jetz, W. & Rahbek, C. (2002) Geographic range size and determinants of avian species richness. *Science*, **297**, 1548–1551.
- Johannesen, E., Høines, Å.S., Dolgov, A.V. & Fosshem, M. (2012) Demersal fish assemblages and spatial diversity patterns in the Arctic–Atlantic transition zone in the Barents Sea. *PLoS ONE*, **7**, e34924.
- Jost, L. (2006) Entropy and diversity. *Oikos*, **113**, 363–375.
- Kaboli, M., Guillaumet, A. & Prodon, R. (2006) Avifaunal gradients in two arid zones of central Iran in relation to vegetation, climate, and topography. *Journal of Biogeography*, **33**, 133–144.
- Kreft, H. & Jetz, W. (2007) Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences USA*, **104**, 5925–5930.
- La Sorte, F., Lee, T.M., Wilman, H. & Jetz, W. (2009) Disparities between observed and predicted impacts of climate change on winter bird assemblages. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 3167–3174.
- Lennon, J.J., Koleff, P., Greenwood, J.J.D. & Gaston, K.J. (2004) Contribution of rarity and commonness to patterns of species richness. *Ecology Letters*, **7**, 81–87.
- Lobo, J.M. & Martín-Piera, F. (2002) Searching for a predictive model for species richness of Iberian dung beetle based on spatial and environmental variables. *Conservation Biology*, **16**, 158–173.
- Magurran, A.E. & Henderson, P. (2012) How selection structures species abundance distributions. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 3722–3726.
- Mokany, K. & Ferrier, S. (2011) Predicting impacts of climate change on biodiversity: a role for semi-mechanistic community-level modelling. *Diversity and Distributions*, **17**, 374–380.
- Mokany, K., Harwood, T.D., Williams, K.J. & Ferrier, S. (2012) Dynamic macroecology and the future for biodiversity. *Global Change Biology*, **18**, 3149–3159.
- Morlon, H., White, E.P., Etienne, R.S., Green, J.L., Ostling, A., Alonso, D., Enquist, B.J. *et al.* (2009) Taking species abundance distributions beyond individuals. *Ecology Letters*, **12**, 488–501.
- Olsen, E., Michalsen, K., Ushakov, N. & Zabaznikov, V. (2011) The ecosystem survey. *The Barents Sea: ecosystem, resources, management* (ed. by T. Jakobsen and K. Ozhigin), pp. 604–609. Tapir, Trondheim.
- Sommer, J.H., Kreft, H., Kier, G., Jetz, W., Mutke, J. & Barthlott, W. (2010) Projected impacts of climate change on regional capacities for global plant species richness. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 2271–2280.
- Steinmann, K., Linder, H.P. & Zimmermann, N.E. (2009) Modelling plant species richness using functional groups. *Ecological Modelling*, **220**, 962–967.
- Synes, N.W. & Osborne, P.E. (2011) Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. *Global Ecology and Biogeography*, **20**, 904–914.
- Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N. *et al.* (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.
- Thuiller, W. (2003) BIOMOD – optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, **9**, 1353–1362.
- Tuomisto, H. (2010) A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, **33**, 2–22.
- Wood, S. (2006) *Generalised additive models: an introduction with R*, p. 401. Chapman & Hall/CRC.
- Woodward, F.I. & Kelly, C.K. (2008) Responses of global plant diversity capacity to changes in carbon dioxide concentration and climate. *Ecology Letters*, **11**, 1229–1237.
- Worm, B., Barbier, E.B., Beaumont, N., Duffy, J.E., Folke, C., Halpern, B.S., Jackson, J.B.C. *et al.* (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science*, **314**, 787–790.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Table S1 List of the species or group of species considered in our studies, and composition of the five communities considered.

Text S2 Details of the cluster analysis leading to the split of the broad fish guild into three trait-defined guilds.

Text S3 Details of the model selection, model formulation and model evaluation, including an analysis of the spatial patterns in the data and in the residuals.

Appendix S4 Exhaustive set of figures for the comparison across the modelling options, including all evaluation years and criteria. See readme for details.

Figure S5 Species richness predictions for the low-fecundity guild, the high-fecundity guild and the all-fishes guild obtained under all environmental scenarios considered in this study.

BIOSKETCHES

Grégoire Certain is a post-doc with interests in spatial ecology, statistical modelling, biodiversity monitoring and management in marine ecosystems.

Carsten F. Dormann is interested in statistical ecology and specifically species distribution analyses.

Benjamin Planque is a researcher with interests in the management of marine resources, marine ecosystem functioning and stochastic dynamic models.

Editor: José Alexandre Diniz-Filho