Vol. 485: 57–73, 2013
doi: 10.3354/meps10307

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published June 27

Excessive spatial resolution decreases performance of quantitative models, contrary to expectations from error analyses

J. Robin Svensson*, Lisbeth Jonsson, Mats Lindegarth

Department of Biology and Environmental Sciences – Tjärnö, University of Gothenburg, Tjärnö, 452 96 Strömstad, Sweden

ABSTRACT: Increased focus on predictive aspects of ecology has recently been urged by scientists and policy makers to provide solutions to pressing societal needs. Current challenges include the large knowledge gap on the spatial distribution of marine biodiversity, and its associated goods and services, and the dependence of model performance on spatial resolution. We evaluated the importance of resolution on the predictive power and precision of empirical models of distributions of marine sessile invertebrates and macroalgae along the Swedish west coast. This was done by simulating the limits to prediction, based on 2 independent simulated proportions of biological variables, and comparing these limits to observed models at different resolutions. Simulations showed the highest achievable predictive power (r^2) and precision (RMSE) of models at fine resolutions (~1 m). In contrast to the simulations, the performance of quantitative models was better at relatively coarse resolutions (~100 m). Increased model performance at coarse resolutions could not be explained by differences in sampling or spatial variability. Instead, the improvement is likely caused by the mechanistic coupling (direct or indirect) between predictor variables, depth and hard substratum cover and patterns at coarser scales, whereas complex processes, e.g. biological interactions, shape patterns at finer scales. This match between resolution and the scale at which environmental variables operate may differ among systems, which could explain the discrepancy in outcomes between our study and previous studies. Furthermore, we provide an approach for error analysis that identifies contributions of different model components to the total uncertainty, thus facilitating model optimization.

KEY WORDS: Distribution · Sessile · Marine · Model · Quantitative · Spatial · Scale

INTRODUCTION

Formulation of predictive statements (i.e. hypotheses) is an important part of the scientific process because it ensures that existing models and theories are critically assessed and refined (Hilborn & Mangel 1997, Underwood 1997, Ford 2000). Consequently, models that withstand repeated critical tests and continue to produce accurate predictions under a wide range of conditions can be considered successful. Although these ideas are in some way embraced by most ecologists, calls for the development of a more 'predictive ecology' have often come from a perspective of challenging the scientific community to offer solutions to pressing societal needs (e.g. Peters 1991, Evans et al. 2012 and references therein).

In the marine realm, these challenges include predicting the consequences of climate change (e.g. Phillips et al. 2006), management of fish stocks (Punt & Hilborn 1997) and options for marine spatial and conservation planning (Sundblad et al. 2011). One tangible piece of evidence of the societal need for predictive capacity is the large and recognized knowledge gap on the spatial distribution of marine

*Email: robin.svensson@marecol.gu.se

biodiversity and its associated goods and services. The spatial distribution is needed as (1) indicators of status in important European policy documents (e.g. the European Habitat Directive [European Commission 1992], the Marine Strategy Framework Directive [European Commission 2008] and the Integrated Maritime Policy [European Commission 2007]) and (2) as a basis for sustainable marine spatial planning. In the benthic marine environment, data on biological structure and function are in the majority of cases collected using small-scale sampling techniques (e.g. cores, dredges, diving or photography). Because these techniques sample a very small part of the area of interest, inferences about global and local distribution of species and structures need to be based on traditional statistical procedures as well as predictive modelling (e.g. Ducrotoy 1999, Diesing et al. 2009, Sundblad et al. 2011).

Important tools in this context are empirical models of species or habitats (e.g. Guisan & Zimmermann 2000, Elith & Leathwick 2009). These models represent an increasingly diverse set of statistical techniques and approaches that use empirical relationships between the biology and associated environmental factors to predict the distribution of the particular biological variable of interest in geographic space. Knowledge about the distribution of species and their relations to environmental gradients is a fundamental component of traditional natural history. Recently, species distribution models (SDMs; Peters 1991, Guisan & Thuiller 2005, Elith & Leathwick 2009) have emerged and seen an outstanding diversification addressing issues in ecology (Anderson et al. 2002), biogeography (Leathwick 1998), invasion biology (Peterson 2003), evolution (Graham et al. 2004), conservation biology (Pearce & Lindenmayer 1998) and climate change (Thuiller 2004).

The predictive performance of empirical models is affected by 3 types of uncertainty: (1) uncertainty in the estimation of the response variable, (2) uncertainty in the estimation of the predictor variable and (3) uncertainties caused by deficiencies in model formulation and structure (Håkanson 1999). The fact that all of these sources of uncertainty, including their relative importance, is highly dependent on spatial and temporal scales means that the performance, optimum structure and sampling design will vary according to the spatial and temporal context (Levin 1992, Schneider 2001, Håkanson & Duarte 2008, Elith & Leathwick 2009, Beale & Lennon 2012). Focusing on the spatial dimension, 2 aspects of scale are relevant: extent (i.e. 'outer scale') and resolution (Turner et al. 1989). While the spatial extent defines the scope and applicability of an empirical model, and extrapolation beyond the extent of a model is not advisable, perhaps even more central to the modelling process are the considerations to do with the spatial resolution.

Firstly, because the importance of processes that structure populations and communities may vary among spatial scales, the usefulness of these factors as predictors will vary with the resolution. Therefore, it is vital to understand the processes that are used to predict distributions of species in order to avoid a mismatch between the scale used for modelling and the scale at which an environmental predictor operates (Thuiller et al. 2003). However, while some authors advocate the importance of finding one appropriate scale for a specific model (Guisan & Thuiller 2005), others argue that there is no single correct spatial scale (Wiens 1989) because species respond to different environment processes at different scales (Levin 1992). Secondly, the intensity of spatial patterns (i.e. patchiness and spatial variability) can be expected to vary strongly among resolutions (see e.g. Morrisey et al. 1992, Underwood & Chapman 1996, Norén & Lindegarth 2005, Gonzalez-Mirelis et al. 2011 for examples from marine benthic environments). Therefore, it follows that the uncertainty in estimates of biological variables to be predicted (within grid-cells) and the amount of variability to be explained by a model (among grid-cells) are also dependent on the resolution. Thus, decisions about spatial resolution can be expected to have fundamental consequences for the success of any attempts to model distribution or abundance using empirical methods. Furthermore, knowledge about the importance of ecological processes and the scale-dependence of spatial variability may be used to optimise sampling programmes in a way that will improve models significantly.

Several studies have emphasised the importance of sampling designs and quality of data for the precision and predictive power of SDMs (e.g. Guisan & Zimmermann 2000, Reese et al. 2005, Araujo & Guisan 2006, Elith & Leathwick 2009, Robinson et al. 2011, Stokland et al. 2011). In contrast, studies that address optimisation, sampling design and uncertainty of quantitative models on spatial patterns of species' abundances are scarce. Analyses of uncertainty in empirical or process-based ecological models, however, offer several insights and methodological frameworks that can also apply to spatial predictions of abundance (e.g. Håkanson 1999, Parysow et al. 2000, Jager & King 2004, Refsgaard

et al. 2007, Håkanson & Duarte 2008, Renard et al. 2010). Furthermore, there is a reasonable wealth of literature exploring the consequences of spatial and temporal variability on the design and precision of sampling programmes in aquatic environments (e.g. Morrisey et al. 1992, Underwood 1994, Benedetti-Cecchi 2001, Clarke et al. 2002). In combination, these different lines of thought suggest that the uncertainty, and therefore the predictive power, of models that predict patterns of abundance may vary strongly depending on spatial resolutions and also that the design and intensity of sampling at relevant scales affects our ability to construct accurate models and to test them critically.

In the present study, we evaluated the importance of spatial resolution on the precision and predictive power of empirical models using video-data from marine subtidal habitats off the Swedish west coast. This was done by (1) estimating spatial variability of a number of biological variables at different resolutions; (2) simulating the limits to prediction, measured as average deviation (RMSE) and proportion of the variance accounted for (r^2), set by spatial variability and measurement error; (3) comparing these limits to observed empirical (linear and non-linear) models at different resolutions; and (4) based on this information, evaluating the relative importance of sampling error compared to other sources of uncertainty.

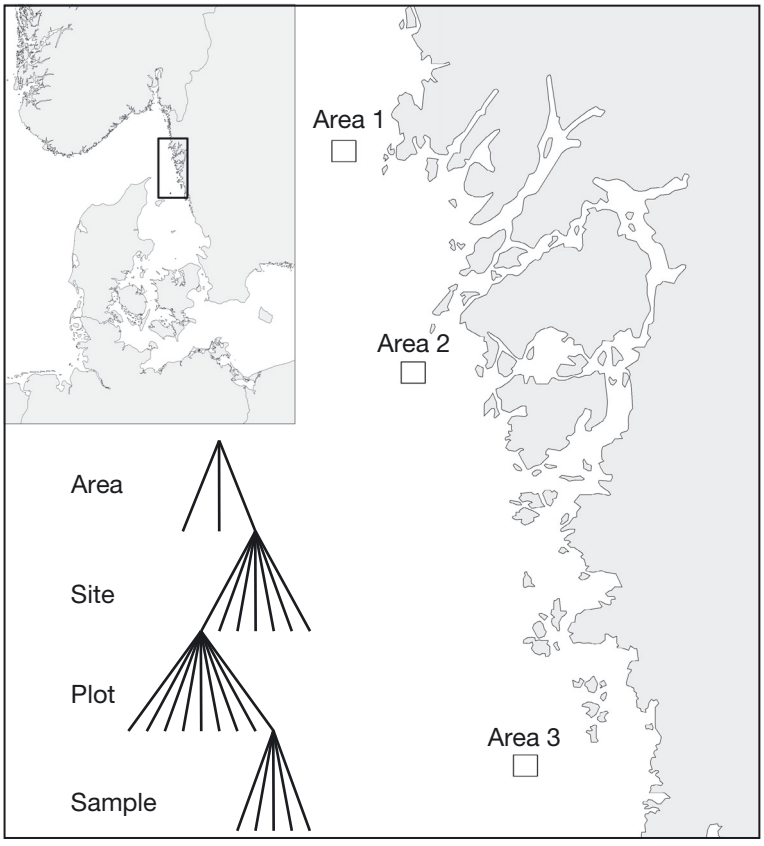


Fig. 1. Study area on the Swedish west coast and factor levels of the experimental design used in the present study. Areas 1 to 3 refer to the off-shore banks called 'Svaberget', 'Makrillbåden' and 'Vanguards grund', respectively

MATERIALS AND METHODS

Study site

The data used in the present study were collected on the west coast of Sweden at 3 offshore banks (Areas 1, 2 and 3 in Fig. 1). The substratum is dominated by flat stone in Area 1 and by rocky benthos in Areas 2 and 3, although all areas have a varying degree of mobile substrata including mearl beds and shell gravel banks. The benthic community in Area 1 is composed mainly of sessile invertebrates, whereas Areas 2 and 3 have a high degree of macroalgal cover. In a survey by the Swedish Environmental Protection Agency, a total of 62 species was observed for Areas 1, 2 and 3, and the most common species in all 3 areas were the soft coral *Alcyonium digitatum*, the bryozoan *Flustra foliacea* as well as the red macroalgae *Phycodrys rubens*, *Dellesreria sanguinea* and unidentified encrusting red algae (Lindblad & Nikolopoulos 2012).

Sampling

Samples of benthic flora and fauna were collected at water depths of 10 to 30 m in 3 off-shore areas (3 to 5 km diameter) using a video-camera mounted to a remotely operated vehicle (ROV). In each area, 7 sites (50 to 200 m diameter) were sampled, and 9 plots (mean ± SD: 11 ± 3 m diameter) were sampled at each site (see Fig. 1). During sampling at a site, the ROV was operated at a constant speed (0.05 to 0.1 m s^{-1}) for 2 to 3 min, and the bottom was filmed from a nearly constant height over the bottom (1 to 2 m). In the laboratory, the video was stopped at 5 randomly

selected times, and each of these fields of view (~0.12 m^2) was analysed using point-samples in 100 regularly spaced points (n_{point} = 100). The points were overlaid on the computer screen, and the species (or genus) of macroalgae or sessile animal as well as the type of substrate (hard substrate or sediment) was recorded under each point.

Estimates of spatial variability

To model precision and predictive power at different spatial resolutions, variance components associated with different spatial resolutions were estimated using the restricted maximum likelihood method with the program R (specifically the library 'lme4' and the function 'lmer()'; R Development Core Team 2008; Table 1). As examples, we used 6 conspicuous components of these benthic assemblages: total cover of algae, total cover of invertebrates, cover of 2 red algae (*Phycodrys rubens* and *Delesseria sanguinea*) and cover of 2 types of sessile invertebrates (*Alcyonium digitatum* and *Flustra* sp.). Furthermore, similar analyses were done using data on depth and substrate (the proportion of hard substrate, i.e. rock or boulders), in which proportion of substrate was calculated from the 100 n_{points} per sample. At a later stage, these were used as environmental predictors in the empirical models (see 'Comparing maximum precision to observed regressions at different resolutions').

Simulating maximum precision and predictive power at different resolutions

The total uncertainty of any predictive model can be conceptually described as consisting of 3 main components: model structure, parameter estimation and data variability (Chatfield 1995, Refsgaard et al. 2007). A fundamental presumption of the simulations presented here is that the limits to precision of a model are ultimately set by the precision of the data used for model building and testing (e.g. Håkanson & Duarte 2008). One way to empirically assess how data uncertainty limits model precision is to plot replicate samples against each other and calculate the 'predictive power' of one sample in relation to the other (e.g. Håkanson & Peters 1995). In essence, this means that uncertainties due to model structure and parameter estimation are eliminated and all deviations are due to data variability. This empirical procedure can be described by a linear model:

$$Y_i = \beta_0 + \beta_1 X_i + e_i \quad (1)$$

where Y_i and X_i are the values of the first and second sample, respectively, β_0 and β_1 are the regression coefficients, and e_i is the residual deviation. Because the samples are repeated samples of the same units, $\beta_0 = 0$, $\beta_1 = 1$, and the linear model can be simplified as follows:

$$Y_i = X_i + e_i \quad (2)$$

Noting that X is also sampled with error and allowing for j samples to be taken for each Y and k samples of X ($X_{ik} = X_i\ \eta_{ik}$), the model can be written as follows:

$$Y_{ij} = X_i + \eta_{ik} + e_{ij} \quad (3)$$

where η_{ik} is the deviation of the kth measurement in the ith level of X. This means that if data variability is the only source of uncertainty in a model, the variance can be estimated from the following relation:

$$\mathrm{Var}(\hat{Y}_i) = \frac{s^2_\eta}{n_X} + \frac{s^2_\varepsilon}{n_Y} \quad (4)$$

where $\mathrm{Var}(\hat{Y}_i)$ is the expected variance of fitted values of Y, s^2_η and s^2_ε are the variances of the random terms associated with sampling X_i and Y_i, and n_x and n_y are the number of samples for X_i and Y_i, respectively. The average deviation from the model, the root mean square error (RMSE), can be estimated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{s^2_\eta}{n_X} + \frac{s^2_\varepsilon}{n_Y}} \quad (5)$$

Table 1. Analysis of variance of design for benthic sampling on the Swedish west coast. Levels are the number of units within each hierarchical scale, MS estimates show the components of variability associated with each mean square estimate, and variances around means within a scale show variances used to simulate precision at various resolutions (n_{point} is the number of points used to estimated cover in an individual photo)

Source	Levels	MS estimates	Variance within scale
Area = A	a = 3	MSA = $s^2_e + ns^2_{P(S,A)} + cns^2_{S(A)} + bcns^2_A$	$V_{Area} = (s^2_e + ns^2_{P(S,A)} + cns^2_{S(A)})$
Site = $S(A)$	b = 7	MSS = $s^2_e + ns^2_{P(S,A)} + cns^2_{S(A)}$	$V_{Site} = (s^2_e + ns^2_{P(S,A)})$
Plot = $P(S, A)$	c = 9	MSP = $s^2_e + ns^2_{P(S,A)}$	$V_{plot} = s^2_e$
Residual (Res)	n = 5	MSRes = s^2_e	$V_{sample} = n_{point}p(1-p)$

Following the arguments above, we assessed the highest achievable model precision (i.e. lowest RMSE) and predictive power (i.e. highest r^2) by Monte Carlo simulations. Instead of using duplicate measurements at identical sites, we used estimated variance components and ranges of percent coverage for a set of taxa at specific spatial resolutions (Table 1). Thus, in a first set of simulations, we sampled N mean values of percent coverage from a uniform distribution within the observed range of a particular variable and resolution. To assess how accurately these proportions can be estimated in 2 independent times of sampling (p_1 and p_2), we added errors by sampling from a set of distributions. For resolutions ranging from 10s to 1000s of meters (plots, sites and areas), errors were added by sampling of n_{sim} values from a normal distribution with the appropriate variance, $N(0, V_{scale})$. Note that for all resolutions, the n_{sim} samples were allocated randomly at that particular spatial scale. This means that:

$$V_{Site} = n_{sim}\left(s_e^2 + s_{P(S,A)}^2\right) \quad (6)$$

and

$$V_{Area} = n_{sim}\left(s_e^2 + s_{P(S,A)}^2 + s_{S(A)}^2\right) \quad (7)$$

where V_{Site} and V_{Area} are the variance within the resolutions of site and area, respectively, and s_e^2, $s_{P(S,A)}^2$ and $s_{S(A)}^2$ are the variance components for the residual, plot and site, respectively (for calculations of variance components, see Table 1). To simulate uncertainty as a function of the number of point samples within individual photos (samples), we sampled n_{point} samples from a binomial distribution, $B(n_{point}, p_i)$ (Table 1). A total of 1000 simulations were run for each combination of resolution, N and n_{sim} or n_{point}. For each run, r^2 and RMSE were calculated using the 2 independent simulated proportions, $\hat{p}_1$ and $\hat{p}_2$. All simulations were done using purpose-built scripts in R (R Development Core Team 2008).

Comparing maximum precision to observed regressions at different resolutions

While the aim of the first set of simulations was to explore the theoretical limits to precision and predictive power using a generic sampling design (i.e. completely randomised), a second set of simulations was done to assess the limits using the particular hierarchical design used in this experiment. This second set of simulations was done at the resolution of samples (a total of 945 samples), plots (189 plots) and sites (21 sites) but not on the resolution of areas (only 3 areas available). The aim was to compare the theoretical limits of the performance of 2 very different types of regression models (multiple regression and random forest) using depth and cover of hard substratum as environmental predictors at different resolutions.

Two types of comparisons were made: (1) performance of model fit and (2) performance during testing (validation) of models. This was done because these factors constitute 2 distinct parts of the modelling procedure, which are affected in different ways by variability in the response variable. In the latter type of comparison, variability in the response variable affects the precision in the fitting phase (training data) and in the test phase (test data). Therefore, simulations of limits for tests of models (SIM_{test}) were identical to those described in the previous section (Eq. 4). In the model-fitting phase, however, measurements of the response variable are only used as training data.

Regression models were fitted and tested using default configurations of the 'stats' package v. 2.13.1 (simple linear regression [e.g. 'function **lm()**'], R Development Core Team 2008) and the package 'randomForest' v. 4.6-2 ('function randomForest()'; Liaw & Wiener 2002). These regression methods were selected because they represent one simple (i.e. linear, parametric) and one highly flexible (i.e. non-linear, non-parametric) example of possible techniques for modelling. This allows us to assess whether any pattern in relation to spatial resolutions appear to be sensitive to modelling technique.

The method 'random forest' was selected from a range of newly developed techniques of equivalent flexibility and potential performance (e.g. GAM, MARS, GLM and CART; Guisan & Zimmermann 2000, Prasad et al. 2006, Elith & Leathwick 2009). It is an ensemble method in which a large number of decision trees are built and predictions in regression problems are based on the average prediction for all of these (Breiman 2001, Liaw & Wiener 2002). Each tree is based on a bootstrap sample (the remaining samples are used for 'out-of-bag' [OOB] estimation of error), and at each node, only a fraction of predictors (randomly selected) are used. The method maintains the automatic treatment of interactive effects of traditional tree-based methods but avoids the risk of suboptimal solutions due to local maxima and does not require pruning. Because of its OOB estimation, the method does not require that the data be split into training and test datasets (e.g. Liaw & Wiener 2002, Prasad et al. 2006, but see Segal 2004 for instances when error may be slightly inflated). This was partic-

Table 2. Range and variability (s^2) within resolutions of environmental and biological variables at the different spatial resolutions: Sample (1 × 1 m), Plot (10 × 10 m), Site (100 × 100 m) and Area (1000 × 1000 m)

Scales	Sample		Plot			Site			Area		
	Min. value	Max. value	Min. average	Max. average	s^2_e	Min. average	Max. average	$s^2_{P(S,A)}$	Min. average	Max. average	$s^2_{S(A)}$
Biological variables											
Total invertebrates	0.0	100.0	0.0	82.8	203.0	3.2	43.6	120.6	12.7	34.0	39.4
Flustra sp.	0.0	85.0	0.0	54.4	67.3	0.0	11.6	35.4	0.3	6.1	9.0
Alcyonium digitatum	0.0	63.0	0.0	36.6	48.4	0.0	15.1	23.6	3.5	9.6	19.1
Total macroalgae	0.0	105.0	0.0	90.8	127.6	2.4	56.1	194.5	15.3	31.9	253.2
Phycodrys rubens	0.0	78.0	0.0	50.8	63.5	0.1	27.6	52.0	3.9	9.4	44.0
Delesseria sanguinea	0.0	92.0	0.0	28.4	47.8	0.0	12.8	7.2	0.3	3.4	8.6
Environmental variables											
Depth	9.0	33.0	10.2	33.0	0.2	14.5	27.4	7.1	20.8	23.7	16.5
Hard substratum	0.0	100.0	0.0	100.0	720.5	4.2	100.0	355.6	23.9	90.9	196.4

ularly useful here because the empirical data had only 21 samples available at the scale of sites. The low number of samples at this resolution is also why multiple regression was not evaluated in the test phase.

RESULTS

Estimates of spatial variability

Variance components of the biological response variables and environmental predictor variables showed 2 types of scale dependence (Table 2). For all variables except the total cover of algae and depth, the largest variance component was associated with the highest resolution (s^2_e, among samples meters apart within plots). In terms of relative importance, the variability associated with the highest resolution amounted to 40–75% of the sum of all sources of variability. For these variables, the variability among plots within sites, $s^2_{P(S,A)}$, was the second most important source (11–34%), while the variability among sites within areas, $s^2_{S(A)}$, was the least important (8–28%). Thus, the variability of most biological features and the type of substrate was dominated by variability at fine spatial resolutions. The total cover of algae and depth showed the opposite pattern. The variability among samples contributed only 22 and 1% for algae and depth, respectively, while at the most coarse resolution, $s^2_{S(A)}$, the corresponding valves were 44 and 69% (Table 2). Furthermore, calculation and comparison of deviations at different spatial scales showed that normal distribution is a close approximation at all spatial scales, which is an important assumption for methods used in the present study (see Appendix 1).

Best achievable precision and predictive power at different resolutions

Simulations of the highest possible predictive power (r^2) and precision (RMSE) show that large differences are to be expected depending on which resolution is chosen for a particular model and species (Figs. 1 & 2). Importantly, the performance of models in terms of predictive power and precision is strongly affected by how well the abundance is estimated within individual units, i.e. by the number of samples (n_{sim}). Comparisons among resolutions at equivalent sampling intensity also show that the predictive power decreases strongly at coarse resolutions (Fig. 2). As an example, the best achievable predictive power (r^2) varies from roughly 0.65 to 0.95 if the total cover of algae is modelled using 10 samples for resolutions that represent areas (~1 × 1 km) and plots (10 × 10 m), respectively (Fig. 2d). The corresponding levels of precision (RMSE) would be 10 and 5% cover (Fig. 3d). Thus, if the performance of predictive models was determined by sampling errors alone, we would expect performance to improve at successively more detailed resolution. This is consistent with the observed combination of larger spatial variability and decreased range of mean values at coarse resolutions.

Among the possible sources of uncertainty for the simulations of the highest achievable predictive power and precision are the number of points that are used in the analyses of sample photos (at the highest resolution) as well as the range of the response variables. The predictive power increased with increasing number of points in a photo, and similarly, the RMSE decreased with increasing number of points in a sample photo (Fig. 4). The range of

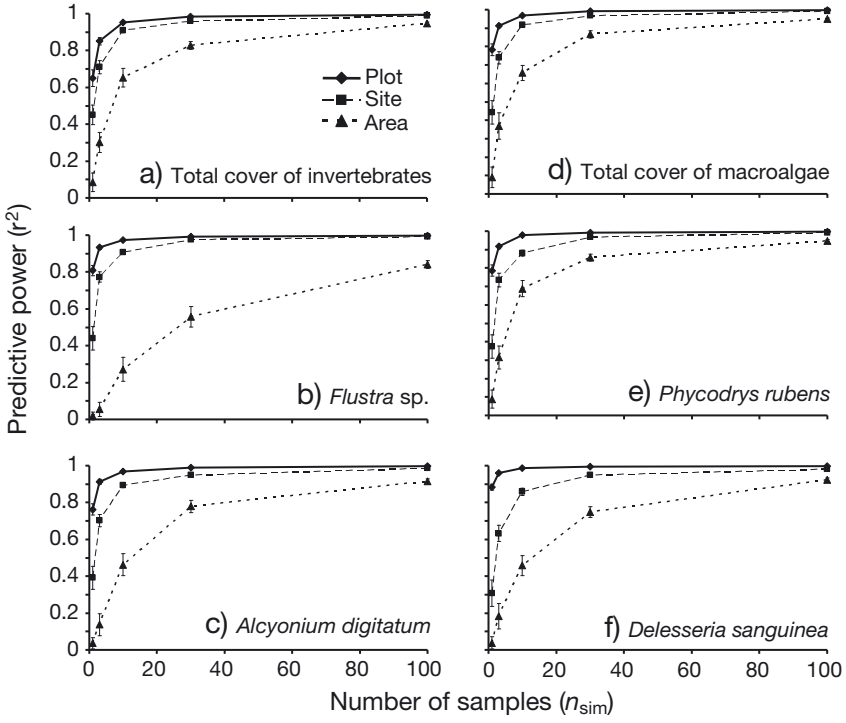


Fig. 2. Simulations of the highest achievable predictive power (r^2) at the different spatial resolutions Plot, Site and Area for cover of (a) invertebrates, (b) *Flustra* sp., (c) *Alcyonium digitatum*, (d) macroalgae, (e) *Phycodrys rubens* and (f) *Delesseria sanguinea*. Data are presented as mean ± SD

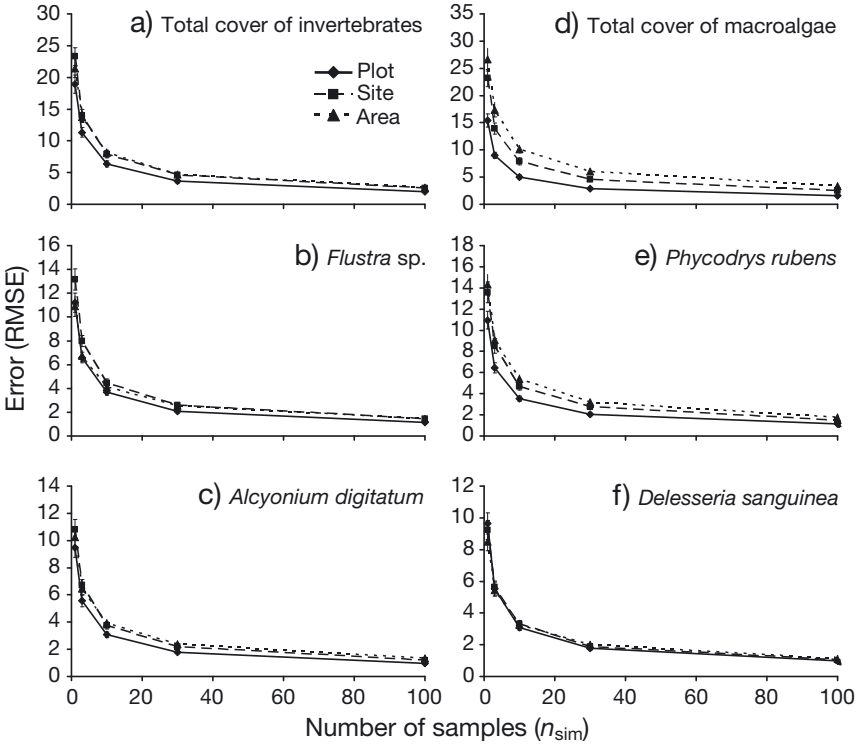


Fig. 3. Simulations of the highest achievable precision (i.e. lowest possible RMSE) at the different spatial resolutions Plot, Site and Area for cover of (a) invertebrates, (b) *Flustra* sp., (c) *Alcyonium digitatum*, (d) macroalgae, (e) *Phycodrys rubens* and (f) *Delesseria sanguinea*. Data are presented as mean ± SD

response variables also affected precision and predictive power. Large ranges reached high r^2 values already at 100 points, while the lower range (0 to 10) would require several hundred points to reach the same predictive power (Fig. 4a). The precision of the 2 larger ranges was very similar and showed distinctly higher RMSE values than the smallest range (Fig. 4b), likely because the larger ranges include more variation, yielding less precision (higher RMSE) in absolute measures.

Comparing best achievable and observed performance at different resolutions

Fitted models

Simulations of the best achievable predictive power and precision (SIM_{fit}) were compared to observed results from simple linear regressions (LM_{fit}) and regressions using the more flexible random forest approach (RF_{fit}). Models were fit for the biological variables using averages of depth and cover of hard substratum as predictors at resolutions of samples (~1 × 1 m), plots (10 × 10 m) and sites (100 × 100 m). Coarser resolutions were deemed irrelevant in the context of habitat models. A number of general patterns emerged from these analyses.

First, if the predictive power was only limited by the precision of sampling, we would expect more powerful models at fine resolution. This is despite the fact that more samples were taken at successively larger scales. Interestingly, this was in sharp contrast to what was observed for the fitted models (Fig. 5). Both the models based on linear regression and those based on random forest showed strong tendencies toward better explana-

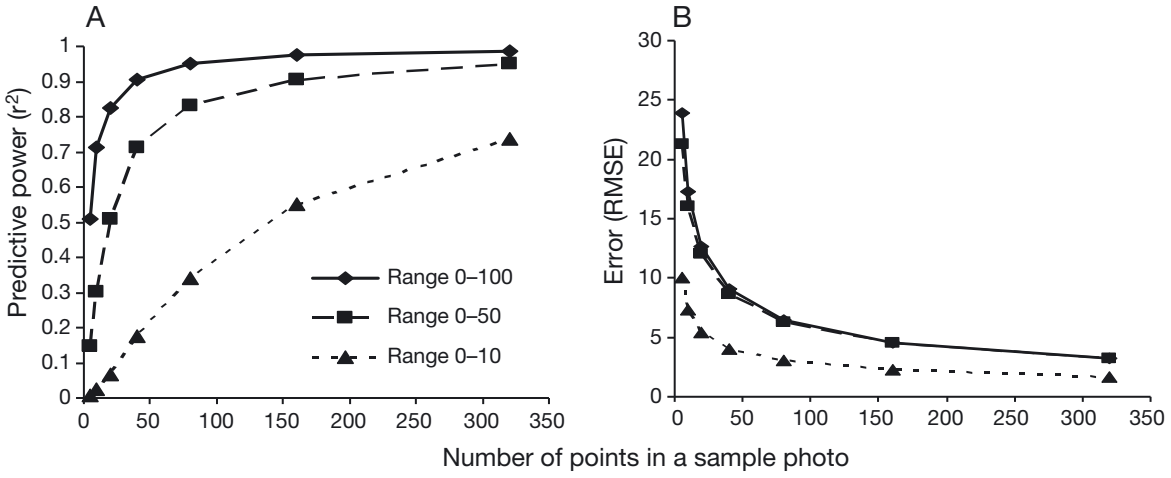


Fig. 4. Simulations of the effects of the number of points in one picture (n_{points}) on the predictive power (r^2) and precision (RMSE) for the 3 different ranges of percentage cover of marine sessile species

tory power at coarser resolutions (Fig. 5). In particular, high predictive power was observed for resolutions at the scale of sites, where r^2 ranged from 0.4 to 0.8 for linear regressions and were constantly >0.8 for the random forest model. Apart from improvements in predictive power, plots of observed and predicted data also show that random forest provided a better representation of the expected 1:1 relationship at all resolutions (Fig. 6). Although none of the models generally reached the best achievable predictive power for any variable, models fitted by random forest occasionally showed higher r^2 values than expected at the coarsest spatial resolution. This is most likely a result of overfitting, which calls for a high degree of caution in the interpretation of the power of fitted models. Nevertheless, the qualitative patterns are consistent among methods and variables.

Second, if precision measured as mean deviation from models (RMSE) were determined only using errors in the measurement of the biological features, the expected deviation would be on the order of 5% cover for most species and resolutions (Fig. 7). Given

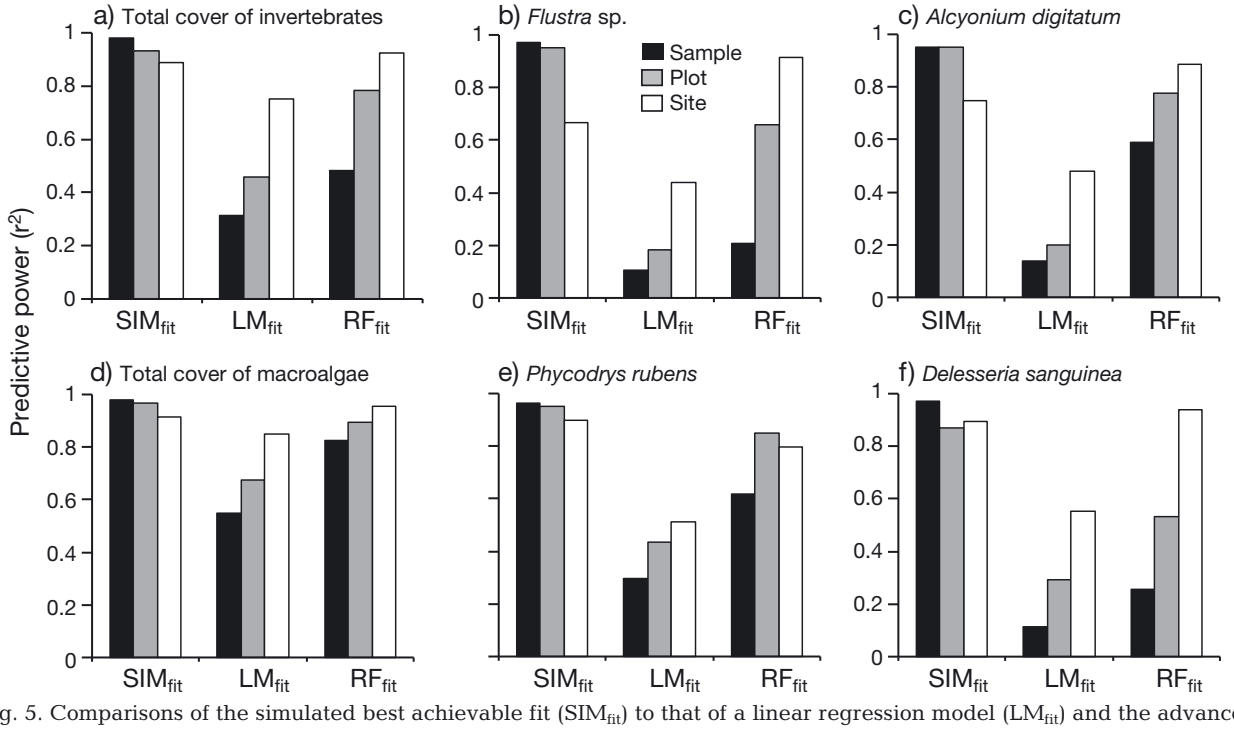


Fig. 5. Comparisons of the simulated best achievable fit (SIM_{fit}) to that of a linear regression model (LM_{fit}) and the advanced model random forest (RF_{fit}) at the different spatial resolutions of Sample, Plot and Site for cover of (a) invertebrates, (b) *Flustra* sp., (c) *Alcyonium digitatum*, (d) macroalgae, (e) *Phycodrys rubens* and (f) *Delesseria sanguinea*

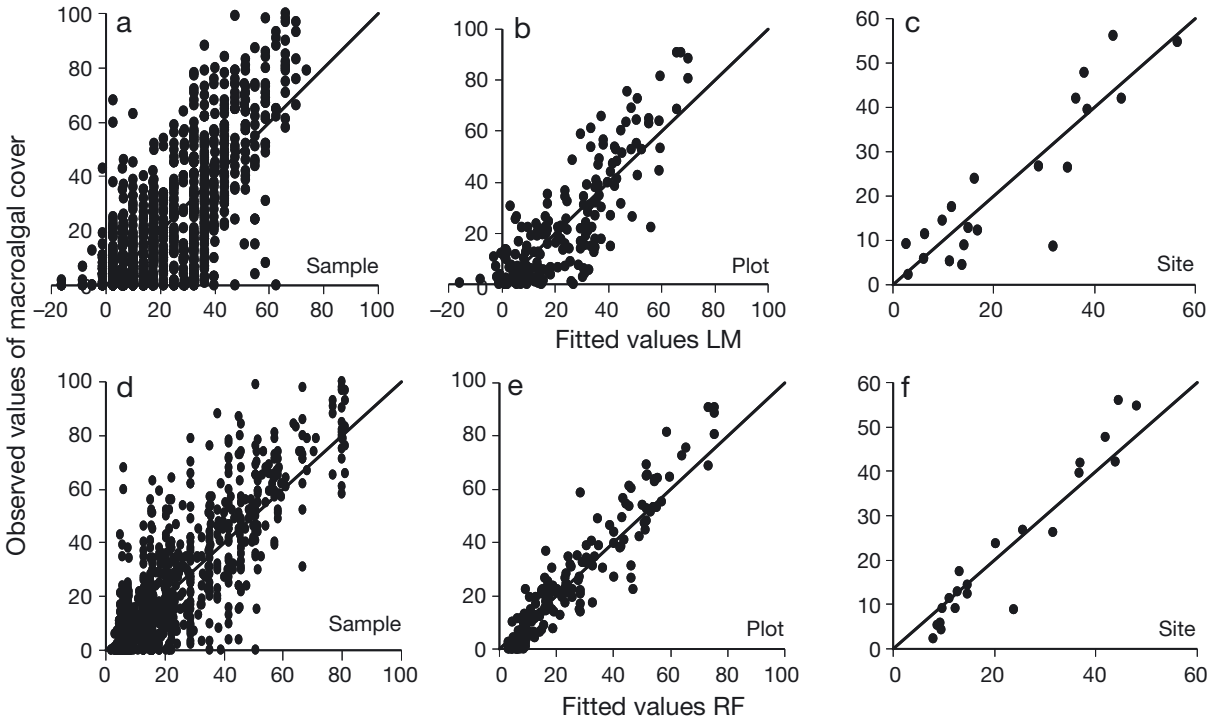


Fig. 6. Examples of the patterns between observed and predicted values of the fitted (a–c) linear model (LM) and (d–f) random forest model (RF) at the 3 different spatial resolutions of Sample (n = 945), Plot (n = 189) and Site (n = 21), with imposed 1:1 lines, for the response variable total cover of macroalgae

the actual sampling design, the simulations show that there is a tendency for smaller deviations at coarser resolution. However, there were some minor deviations from this general pattern, i.e. the simulations of precision showed the highest RMSE at the intermediate resolution for the total cover of invertebrates (Fig. 7a) and the lowest RMSE for the finest resolution for the total cover of macroalgae (Fig. 7d). In accordance with the conclusions from analyses of predictive power, the fits of the regression models were all better at coarser resolution than at fine resolution. This is true for both types of regression, although the RMSE of the random forest approach was generally lower. One interesting pattern is that the difference between the simulated precision, which is expected from information on sampling variability and design, and the precision observed from fitted models is much larger at finer resolution. This indicates that at fine resolution, the precision of a model is generally not limited by measurement error but by imperfections of the model. At coarser resolution, however, the precision of the model fit approaches the precision expected from sampling errors. Hence, at this scale, an increase in the quality of the samples, i.e. reducing the sampling error, may have larger impacts on model fit than increasing the number of samples.

Tests of models

The test, or validation, of random forest models (RF_{test}) was performed using the OOB estimate of error and the 'pseudo R-squared' implemented in 'randomForest()'. These measures were compared to best achievable performance from simulated data (SIM_{test}). The outcome of these analyses was similar to those of the fitted models (i.e. RF_{fit} and LM_{fit} compared to SIM_{fit}). Simulations suggested that the best achievable predictive power (r^2) is generally expected to decrease from an average of 0.94 at a resolution of samples to 0.73 for sites (Fig. 8). In contrast to this, models of the total cover of invertebrates *Flustra* sp. and *Delesseria sanguinea* showed maximum predictive power at the resolution of sites. Models of the total cover of algae and *Phycodrys rubens* were the most powerful at the intermediate resolution of plots, although models of the total cover of algae were consistently powerful ($r^2 \approx 0.7$). The predictive power for models of *Alcyonium digitatum* was consistently poor ($r^2 \approx 0.2$).

Furthermore, simulations of precision for the 4 individual species show that, assuming the uncertainty of models was only due to sampling errors (in combination with sampling design), we would expect a decrease in RMSE from finer to coarser

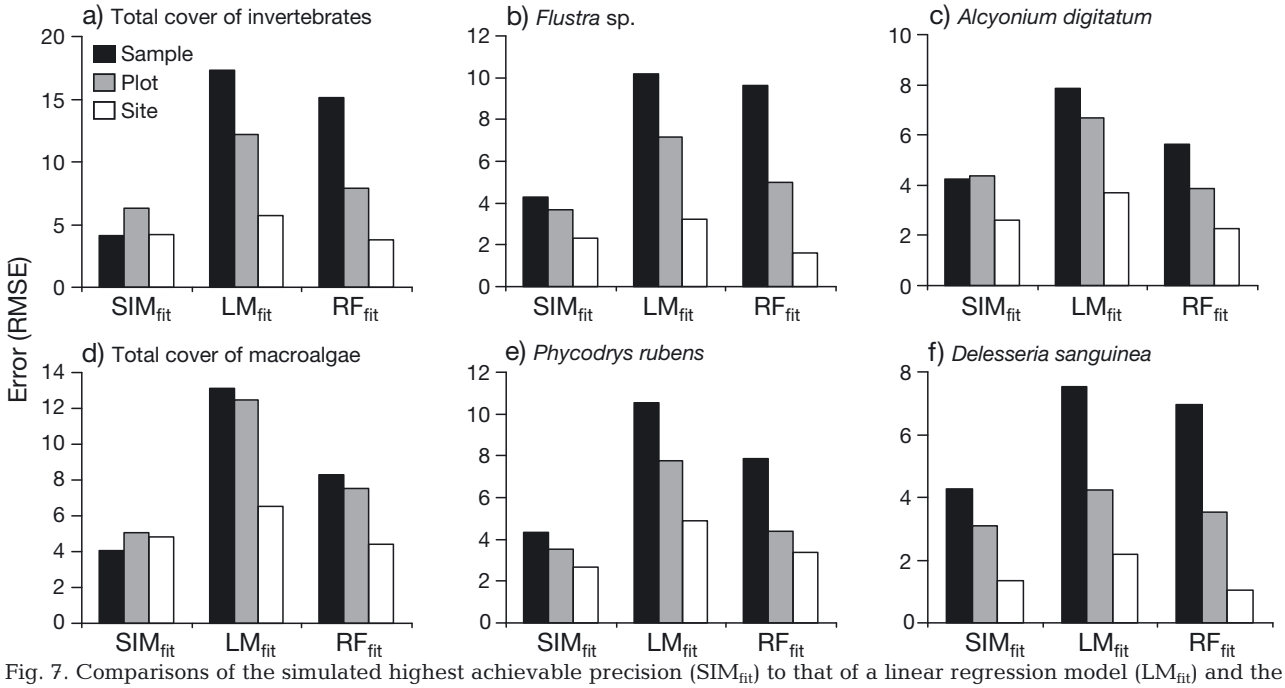


Fig. 7. Comparisons of the simulated highest achievable precision (SIM_{fit}) to that of a linear regression model (LM_{fit}) and the advanced model random forest (RF_{fit}) at the different spatial resolutions of Sample, Plot and Site for cover of (a) invertebrates, (b) *Flustra* sp., (c) *Alcyonium digitatum*, (d) macroalgae, (e) *Phycodrys rubens* and (f) *Delesseria sanguinea*

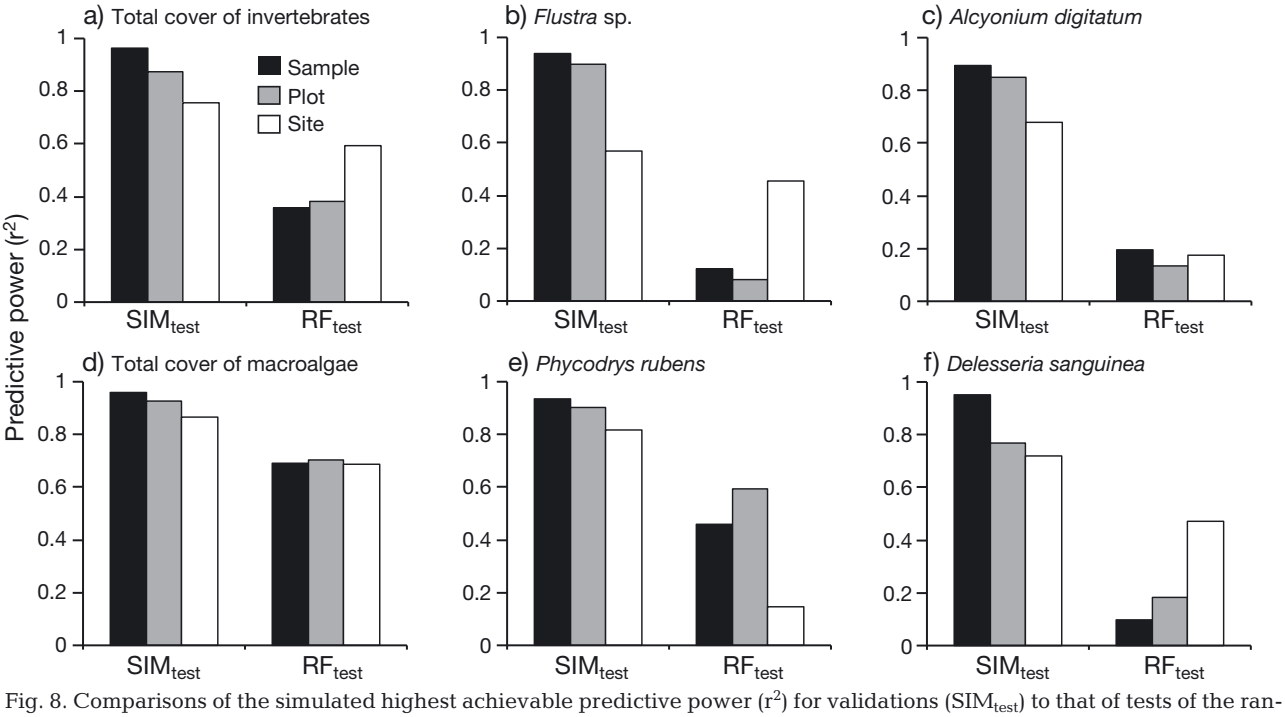


Fig. 8. Comparisons of the simulated highest achievable predictive power (r^2) for validations (SIM_{test}) to that of tests of the random forest model (RF_{test}) at the different spatial resolutions of Sample, Plot and Site for cover of (a) invertebrates, (b) *Flustra* sp., (c) *Alcyonium digitatum*, (d) macroalgae, (e) *Phycodrys rubens* and (f) *Delesseria sanguinea*

resolutions (Fig. 9). This is also consistent with what is observed for the OOB errors. For the total cover of invertebrates and algae, the patterns were slightly more complex. Simulations indicated that maximum precision could be expected at the finest resolution, while the deviations observed for the models were smaller at the coarsest resolution (Fig. 9). One important observation is that the sim-

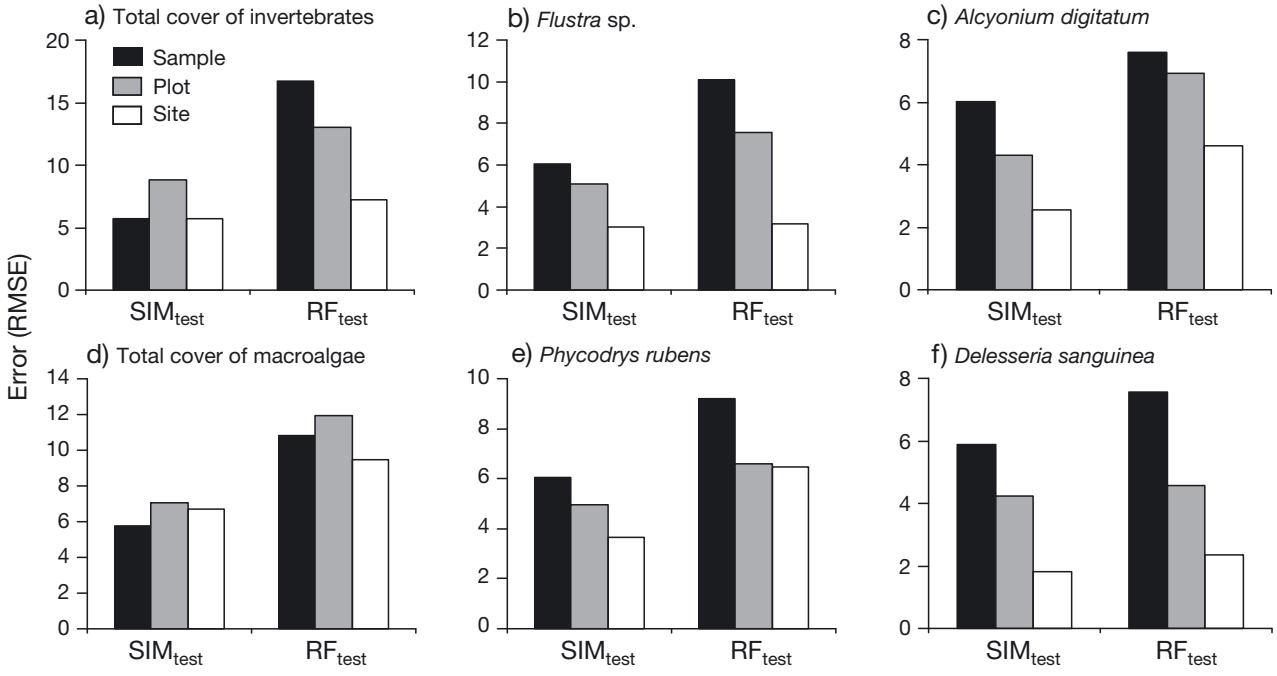


Fig. 9. Comparisons of the simulated highest achievable precision (i.e. lowest possible RMSE) for validations (SIM_{test}) to that of tests of the random forest model (RF_{test}) at the different spatial resolutions of Sample, Plot and Site for cover of (a) invertebrates, (b) *Flustra* sp., (c) *Alcyonium digitatum*, (d) macroalgae, (e) *Phycodrys rubens* and (f) *Delesseria sanguinea*

ulated best achievable precision was not exceeded by the observed precision, measured as OOB estimates, for any of the biological variables or resolutions (Fig. 9).

DISCUSSION

In the present study, we evaluated the importance of spatial resolution on the precision (RMSE) and predictive power (r^2) of quantitative models of marine benthos by comparing simulations of highest achievable performance to the observed performance of models using 2 types of statistical approaches. The observed performance as well as the highest achievable predictive power and precision of models depended strongly on spatial resolution. The simulations of the highest achievable model performance suggested that both predictive power and precision are expected to increase at fine resolutions (1 × 1 m). In contrast, the observed performance of both simple and flexible modelling approaches was better at coarse (1 × 1 km) to intermediate (10 × 10 m) spatial resolutions.

To elucidate this discrepancy in performance between simulations and observed models, we must first understand why the best achievable performance is expected to be better at fine resolution. The main reason is that the spatial variability within a sampling unit increases, or at least does not decrease, as the resolution gets coarser. This is because nested components of variability successively add uncertainty to mean estimates at successively larger scales (e.g. Morrisey et al. 1992, Norén & Lindegarth 2005). Despite this, it is possible from traditional sampling theory to achieve precise mean estimates at larger scales and thus precise models at coarse resolution by extensive sampling (e.g. Underwood 1997, Quinn & Keough 2002). Our simulations consequently showed a strong effect of sample size on precision for all variables as well as differences among resolutions for a given sample size. As a general recommendation, we conclude that achieving an average standard error of ±10% cover of the marine benthic species in our study area requires >3 samples, while ≥10 samples are required for an error of ±5% cover. Furthermore, because the precision depends on sample size, it is possible to achieve a precision comparable to that of fine resolutions by more extensive sampling per modelling unit at coarser resolution. One example is the simulations of cover of macroalgae (Fig. 3), where an error of ±10% requires <3 samples at the plot (10 × 10 m) resolution but 5 to 10 samples at the site (100 × 100 m) resolution and >10 samples at the area (1000 × 1000 m) resolution.

The difference in maximum achievable predictive power (r^2) among resolutions was even more pronounced than that observed for maximum achievable precision. As a general rule of thumb, <3 samples were required to achieve a predictive power of $r^2 > 0.8$ at the finest resolution, while approximately 10 and 30 samples were required at intermediate and coarse resolutions, respectively (Fig. 2). Similarly to the differences in precision, these large differences among resolutions are explained by increased spatial variability within sampling units but importantly also exist in combination with differences in variability and range among sampling units. The range of sampling units defines the scope of a model by representing the variability 'to be explained' by the model. This scope decreases at coarser resolutions because spatial patterns are averaged out at larger scales. As an example, the range of the total cover of algae decreased from 100% at the level of individual photos to 91, 54 and 16% at resolutions of plots, sites and areas, respectively. Thus, the loss of contrast among sampling units may explain, partly, why the best achievable predictive power decreases with coarsening resolution. However, the impact of the measurement error and spatial variability should not be neglected as there were also differences among resolutions for the best achievable precision, which is not affected by range. In conclusion, there were larger differences among resolutions for the best achievable predictive power compared to the highest possible precision, largely due to range, and decreased precision at coarser resolutions can be compensated for by a larger number of samples. This is in accordance with previous studies showing that small sample size can be a significant source of instability and errors in models (Stockwell & Peterson 2002, Guisan & Thuiller 2005). Consequently, it may be worthwhile to consider the benefits of an increased number of samples, and increased quality of samples, in optimizations of model performance (Stockwell & Peterson 2002, Guisan & Thuiller 2005, Lobo 2008).

In direct contrast to the results from the simulations, the observed models showed increasing performance with coarsening spatial resolution. Both the observed linear models and the more flexible random forest models showed higher predictive power at coarser resolutions for all variables, whereas the simulations showed the opposite pattern (Fig. 5). Random forest models showed higher predictive power and accuracy (although they also showed overfitting for the coarsest resolution), than the linear models, but all models showed the same pattern of increased performance at coarser resolutions. However, higher predictive power of observed models at coarse resolutions is likely unrelated to the sampling design because both types of simulations of best model performance, i.e. based on the hierarchical or the random sampling design, showed decreasing predictive power with increasing resolution. Because the simulations showed decreasing predictive power at a smaller sample size and more narrow range, the increased performance of the observed models at coarse resolutions occurred in spite of the decreasing number of samples and range.

The increase in model performance at coarser resolution is also in contrast to previously reported results in the literature. Studies that have tested differences among resolutions in model performance are scarce, particularly in marine environments, and show either very low effects or the direct opposite pattern. Gottschalk et al. (2011) showed that the performance of species-habitat models of 13 bird species, measured as the explained deviance, decreased with increasing resolution and that the best model results were observed at a resolution (grain size) of 1 to 3 m. They also showed in a meta-analysis (based on 6 studies) that previous studies show no or very weak effects of resolution (although see Tobalske 2002) but that the pattern of the weak effects still indicated increased performance at higher resolutions. The fact that only weak effects were reported in the meta-analysis was explained by biases among the tested resolutions (e.g. Chust et al. 2004, Seoane et al. 2004) and the fact that the high resolutions were still relatively coarse (i.e. ≥100 m; Guisan et al. 2007a,b). However, other studies not included in this meta-analysis showed the best model performance at the fine resolutions. For instance, Kendall et al. (2011) observed higher r-values at high spatial resolutions for correlations between resident fish abundance and sediment area as well as for correlations between hard bottom edge and resident species richness. Heikkinen et al. (2007) found the most accurate models (highest AUC and kappa values) of the distribution of the clouded apollo butterfly *Parnassius mnemosyne* at high to intermediate resolutions. However, Heikkinen et al. (2007) suggested that not only the quality of the data at different resolutions is important (i.e. the accuracy of species records and precise locations of habitats) but also the ecological relevance, because the butterfly species in their study operates at the same spatial scales as the highest level of resolution in the study. Similarly, Graf et al. (2005) found the best species habitat model of capercaillie *Tetrao urogallus* at resolutions equivalent to the size of an annual home range (250 ha),

which was of intermediate resolution in their study. Thompson & McGarigal (2002) investigated both elements of spatial scale, i.e. the extent as well as the resolution, and found different scale-dependent patterns on the habitat selection of bald eagle *Haliaeetus leucocephalus*. At fine resolution, the eagles avoided disturbed areas and chose nesting sites based on canopy structure, whereas at coarse resolution, they showed a strong selection for areas with human activity and avoided areas with no activity. This shows that important patterns can emerge at multiple scales, which cannot be identified in studies using only one resolution. Hence, it is important to acknowledge that there might be more than one 'best' resolution at which to model (Wiens 1989, Levin 1992, Elith & Leathwick 2009) and that the ecological relevance is vital for assigning appropriate resolutions (Heikkinen et al. 2007).

The importance of ecological relevance could possibly explain both the discrepancies in outcomes between our study and previous studies and the differences between the simulations and the observed models. It is likely that the scale at which the natural processes occur, which may be used to determine the distribution of species, can differ among ecological assemblages and systems. In the benthic marine system of our study, it may not be realistic to expect high predictability in species' distributions at very fine resolutions (i.e. ≤1 m^2). This is because overall patterns based on depth and cover of hard-substratum are easily obscured by processes that are not easily measured at very small scales, such as small-scale hydrodynamics, the orientation of the surface and sporadic occurrence of predators, antagonists or local disturbances. Instead, coarser resolutions may for certain environmental variables better reflect the natural processes that shape sessile species distributions. Hence, the simulations show the best achievable model performance if measurement error in the response variables and spatial variability are the sole limiting factors (Håkanson 1999, Håkanson & Duarte 2008), whereas the observed models also take into account the error in estimates of the environmental variables as well as the model structure and how it relates to pattern of natural processes (Guisan & Zimmermann 2000, Elith & Leathwick 2009). For instance, the linear model used the depth and cover of hard substratum as predictor variables and showed higher predictive power and accuracy at coarser resolutions, suggesting that more general patterns of these variables yield better predictions than the highly localized conditions identified at very fine resolutions. Hence, it is possible that overly detailed data of environmental variables can obscure more general ecologically relevant patterns, which may result in lower performance of qualitative models of species distributions. Consequently, the ecological relevance at different resolutions in different systems may explain why the best model performance was found at fine resolutions at which butterflies operate (Heikkinen et al. 2007), at intermediate resolutions equivalent to home ranges of birds (Graf et al. 2005) and at coarse resolutions that better reflect the patterns between sessile species and environmental variables in the present study.

One interesting consequence of these findings is that they create possibilities to assess the importance of different sources of variability and, more importantly, to evaluate different options for designing sampling programs and modelling efforts in order to optimise their predictive power, precision and applied benefits. One possibility is to perform error (uncertainty) analyses of simulations and models (e.g. Håkanson & Peters 1995, Jager & King 2004), noting again that the total uncertainty of a model ($\sigma^2_{e,\mathrm{Total}}$) can be partitioned among the uncertainty caused by errors in estimates of the response variable ($\sigma^2_{e,y}$), errors in estimates of predictor variables ($\sigma^2_{e,x}$) and by insufficiencies in the structure of the model ($\sigma^2_{e,\mathrm{model}}$):

$$\sigma^2_{e,\mathrm{Total}} = \sigma^2_{e,y} + \sigma^2_{e,x} + \sigma^2_{e,\mathrm{model}} \quad (8)$$

Using estimates of precision from simulations and fitted models of the total cover of algae, the scope for model improvements can be analysed at different resolutions (Table 3). These analyses illustrate previous conclusions of decreasing total uncertainty at successively larger scales for both methods and that the random forest approach more than halves the component that includes model structure at all resolutions. Another interesting observation, however, is that the relative importance of uncertainties due to

Table 3. Analysis of error in models of total cover of algae using simulated uncertainty associated with the *y*-variable, $\sigma^2_{e,y}$ (=$RMSE^2$ from Fig. 6d), at different spatial resolutions and the observed total uncertainty of fitted models, $\sigma^2_{\mathrm{total}}$, allows calculation of the combined uncertainty associated with model structure and estimation of predictor variables. SIM: simulation; LM: linear model; RF: random forest

	SIM	LM		RF	
	$\sigma^2_{e,y}$	$\sigma^2_{\mathrm{total}}$	$\sigma^2_{\mathrm{x}} + \sigma^2_{\mathrm{model}}$	$\sigma^2_{\mathrm{total}}$	$\sigma^2_{\mathrm{x}} + \sigma^2_{\mathrm{model}}$
Sample	16.7	171.3	154.7	68.4	51.7
Plot	25.5	156.0	130.6	56.3	30.8
Site	23.3	42.7	19.4	19.5	–3.8

errors in the estimation of total cover is low at fine resolution (10 and 25% for linear regression and random forest, respectively) but larger at coarse resolution (50 and ~100%, respectively). These analyses clearly illustrate that efficient strategies for improving the precision of predictions depend on the spatial resolution, despite the fact that the extreme fit observed for random forest at the coarsest resolution (see Fig. 5) is an obvious case of overfitting. Random forest models are in some respects protected against overfitting (Breiman 2001, Liaw & Wiener 2002, Prasad et al. 2006), but this is only true for the number of trees in a forest, not the length of trees (Liaw & Wiener 2002, Segal 2004). Consequently, introducing tree pruning or limiting the number of splits may further improve the performance of this technique (Segal 2004). Therefore, the absolute uncertainties in the random forest models (see Fig. 5) need to be interpreted with caution, even though the pattern is consistent among variables. At fine resolution (i.e. ~1 m), improving precision in measurements by taking more samples will not likely reduce the overall uncertainty of models substantially. Instead, at these scales, a more fruitful approach would be to improve model structure by including more dynamic factors and small-scale processes. At coarse resolution (10 or 100 m), however, where the model structure (i.e. the factors and the estimated model parameters) appears to have reduced uncertainty to a larger extent, the most efficient strategy for further refinement of predictions may be different. As an example, the uncertainty associated with estimating the coverage of algae amounts to 50–100% at these scales for random forest models. Therefore, the most efficient strategy here is probably to collect more samples and thus to obtain more precise estimates. Note, however, that the uncertainty of models at the 100 m resolution was typically 70–75% smaller than that at 1 m resolution. Despite the general desire to improve models and predictions, it is possible that this precision may suffice for the purposes of the study. Consequently, we believe that much can be gained by investigating whether the largest source of uncertainty is associated with the predictor variable or the model structure before deciding how to improve models on species distributions.

CONCLUSIONS

In the present study, we found that the spatial resolution is vital for the observed performance as well as the highest achievable predictive power and precision of quantitative models. Simulations of the highest achievable model performance showed that the predictive power of models should increase with increasing resolution, but the observed models instead showed higher predictive power with decreasing resolution. Similarly, the precision of the observed models increased with increasing spatial resolution, whereas the simulations showed decreasing precision at coarse resolutions. The simulations show the limits of model performance that are based on the uncertainty associated with the response variable. Therefore, it is instead the structure of the models that can explain the increased performance of the observed models at coarse resolutions. One important aspect of the model structure is how well the spatial resolution matches the scale at which the response and environmental variables operate, i.e. ecological relevance. This could explain the discrepancy in outcomes between our study and previous studies, because the match between resolution and variable is likely to differ among systems. In our study, we found that excessive resolution of environmental variables can obscure more general ecologically relevant patterns that can cause lower performance of models. Furthermore, we showed a technique for determining which component of a model contributes the most to the total amount of uncertainty and therefore has the highest potential for improvement in optimizations of models.

Acknowledgements. This research was funded by the European Community's Seventh Framework Programme (FP/2007-2013) under grant agreement no. 217246 made with the joint Baltic Sea research and development programme BONUS, by the Swedish Environmental Protection Agency from contracts 08/391 PREHAB and 10/179, by FORMAS from contract 217-2006-357 and by Water Management through a contract for the research programme WATERS. We thank 4 anonymous reviewers and the editor for improving this manuscript.

LITERATURE CITED

Anderson RP, Peterson AT, Gomez-Laverde M (2002) Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. Oikos 98:3–16

Araujo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. J Biogeogr 33: 1677–1688

Beale CM, Lennon JJ (2012) Incorporating uncertainty in predictive species distribution modelling. Philos Trans R Soc Lond B 367:247–258

Benedetti-Cecchi L (2001) Beyond BACI: optimization of environmental sampling designs through monitoring and simulation. Ecol Appl 11:783–799

Breiman L (2001) Random forests. Mach Learn 45:5–32

Chatfield C (1995) Model uncertainty, data mining and statistical inference. J R Stat Soc Ser A Stat Soc 158:419–466

Chust G, Pretus JL, Ducrot D, Ventura D (2004) Scale dependency of insect assemblages in response to landscape pattern. Landsc Ecol 19:41–57

Clarke RT, Furse MT, Gunn RJM, Winder JM, Wright JF (2002) Sampling variation in macroinvertebrate data and implications for river quality indices. Freshw Biol 47: 1735–1751

Diesing M, Coggan R, Vanstaen K (2009) Widespread rocky reef occurrence in the central English Channel and the implications for predictive habitat mapping. Estuar Coast Shelf Sci 83:647–658

Ducrotoy JP (1999) Protection, conservation and biological diversity in the North East Atlantic. Aquat Conserv 9: 313–325

Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst, Book 40:677–697

European Commission (1992) Council Directive 92/43/EEC. Conservation of natural habitats and of wild flora and fauna. Off J Euro Comm L 206:7–49

European Commission (2007) An integrated maritime policy for the European Union. COM (2007) 2575, European Commission, Brussels

European Commission (2008) Council Directive 2008/56/EC. A framework for community action in the field of marine environmental policy. Off J Euro Comm L 164:19–40

Evans MR, Norris KJ, Benton TG (2012) Predictive ecology: systems approaches. Philos Trans R Soc Lond B 367: 163–169

Ford ED (2000) Scientific method for ecological research. Cambridge University Press, Cambridge

Gonzalez-Mirelis G, Bergstrom P, Lindegarth M (2011) Interaction between classification detail and prediction of community types: implications for predictive modelling of benthic biotopes. Mar Ecol Prog Ser 432:31–44

Gottschalk TK, Aue B, Hotes S, Ekschmitt K (2011) Influence of grain size on species-habitat models. Ecol Model 222: 3403–3412

Graf RF, Bollmann K, Suter W, Bugmann H (2005) The importance of spatial scale in habitat models: capercaillie in the Swiss Alps. Landsc Ecol 20:703–717

Graham CH, Ron SR, Santos JC, Schneider CJ, Moritz C (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. Evolution 58:1781–1793

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8: 993–1009

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecol Model 135:147–186

Guisan A, Graham CH, Elith J, Huettmann F, Distri NS (2007a) Sensitivity of predictive species distribution models to change in grain size. Divers Distrib 13:332–340

Guisan A, Zimmermann NE, Elith J, Graham CH, Phillips S, Peterson AT (2007b) What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? Ecol Monogr 77:615–630

Håkanson L (1999) On the principles and factors determining the predictive success of ecosystem models, with a focus on lake eutrophication models. Ecol Model 121: 139–160

Håkanson L, Duarte CM (2008) Data variability and uncertainty limits the capacity to identify and predict critical changes in coastal systems—a review of key concepts. Ocean Coast Manag 51:671–688

Håkanson L, Peters RH (1995) Predictive limnology. Methods for predictive modelling, Vol 1. SPB Academic Publishing, Amsterdam

Heikkinen RK, Luoto M, Kuussaari M, Toivonen T (2007) Modelling the spatial distribution of a threatened butterfly: impacts of scale and statistical technique. Landsc Urban Plan 79:347–357

Hilborn R, Mangel M (1997) The ecological detective, Vol 1. Princeton University Press, Princeton, NJ

Jager HI, King AW (2004) Spatial uncertainty and ecological models. Ecosystems 7:841–847

Kendall MS, Miller TJ, Pittman SJ (2011) Patterns of scale-dependency and the influence of map resolution on the seascape ecology of reef fish. Mar Ecol Prog Ser 427: 259–274

Leathwick JR (1998) Are New Zealand's *Nothofagus* species in equilibrium with their environment? J Veg Sci 9:719–732

Levin SA (1992) The problem of pattern and scale in ecology. Ecology 73:1943–1967

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2:18–22

Lindblad C, Nikolopoulos A (2012) Survey of off-shore banks, Report 6385, Vol 1. Swedish Environmental Protection Agency, Stockholm (in Swedish)

Lobo JM (2008) More complex distribution models or more representative data? Biodivers Informatics 5:14–19

Morrisey DJ, Howitt L, Underwood AJ, Stark JS (1992) Spatial variation in soft-sediment benthos. Mar Ecol Prog Ser 81:197–204

Norén K, Lindegarth M (2005) Spatial, temporal and interactive variability of infauna in Swedish coastal sediments. J Exp Mar Biol Ecol 317:53–68

Parysow P, Gertner G, Westervelt J (2000) Efficient approximation for building error budgets for process models. Ecol Model 135:111–125

Pearce J, Lindenmayer D (1998) Bioclimatic analysis to enhance reintroduction biology of the endangered helmeted honeyeater (*Lichenostomus melanops cassidix*) in southeastern Australia. Restor Ecol 6:238–243

Peters RH (1991) A critique for ecology. Cambridge University Press, Cambridge

Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. Q Rev Biol 78: 419–433

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190:231–259

Prasad AM, Iverson LR, Liaw A (2006) Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9:181–199

Punt AE, Hilborn R (1997) Fisheries stock assessment and decision analysis: the Bayesian approach. Rev Fish Biol Fish 7:35–63

Quinn GP, Keough MJ (2002) Experimental design and data analysis for biologists, Vol 1. Cambridge University Press, Cambridge

R Development Core Team (2011) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Reese GC, Wilson KR, Hoeting JA, Flather CH (2005) Factors affecting species distribution predictions: a simulation modeling experiment. Ecol Appl 15:554–564

Refsgaard JC, van der Sluijs JP, Hojberg AL, Vanrolleghem

PA (2007) Uncertainty in the environmental modelling process — a framework and guidance. Environ Model Softw 22:1543–1556

Renard B, Kavetski D, Kuczera G, Thyer M, Franks SW (2010) Understanding predictive uncertainty in hydrologic modeling: the challenge of identifying input and structural errors. Water Resour Res 46:1–22

Robinson LM, Elith J, Hobday AJ, Pearson RG, Kendall BE, Possingham HP, Richardson AJ (2011) Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. Glob Ecol Biogeogr 20:789–802

Schneider DC (2001) The rise of the concept of scale in ecology. BioScience 51:545–553

Segal MR (2004) Machine learning benchmarks and random forest regression. eScholarship Repository University of California, San Francisco, CA, available at http://repositories.cdlib.org/cbmb/bench_rf_regn

Seoane J, Bustamante J, Diaz-Delgado R (2004) Are existing vegetation maps adequate to predict bird distributions? Ecol Modell 175:137–149

Stockwell DRB, Peterson AT (2002) Effects of sample size on accuracy of species distribution models. Ecol Model 148: 1–13

Stokland JN, Halvorsen R, Stoa B (2011) Species distribution modeling — effect of design and sample size of pseudo-absence observations. Ecol Model 222:1800–1809

Sundblad G, Bergström U, Sandström A (2011) Ecological coherence of marine protected area networks: a spatial assessment using species distribution models. J Appl Ecol 48:112–120

Thompson CM, McGarigal K (2002) The influence of research scale on bald eagle habitat selection along the lower Hudson River, New York (USA). Landsc Ecol 17:569–586

Thuiller W (2004) Patterns and uncertainties of species' range shifts under climate change. Glob Change Biol 10: 2020–2027

Thuiller W, Araujo MB, Lavorel S (2003) Generalized models vs. classification tree analysis: predicting spatial distributions of plant species at different scales. J Veg Sci 14: 669–680

Tobalske C (2002) Effects of spatial scale on the predictive ability of habitat models for the green woodpecker in Switzerland. In: Scott JM, Heglund PJ, Samson F, Haufler J, Morrison M, Raphael M, Wall B (eds) Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, CA, p 197–204

Turner MG, Dale VH, Gardner RH (1989) Predicting across scales: theory development and testing. Landsc Ecol 3:245–252

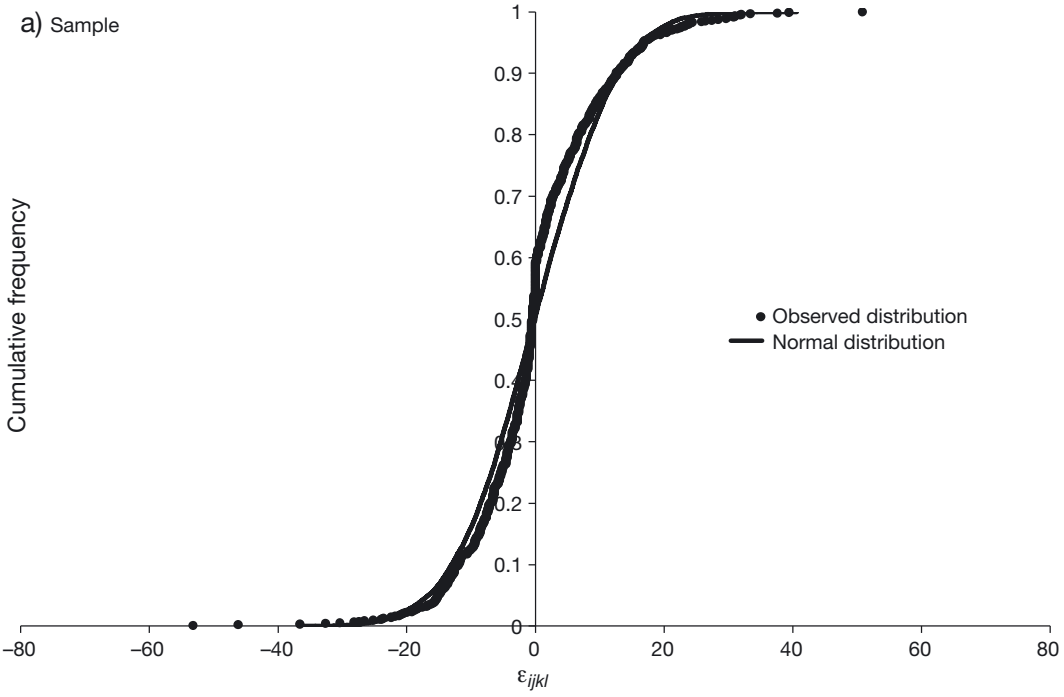
Underwood AJ (1994) On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecol Appl 4:3–15

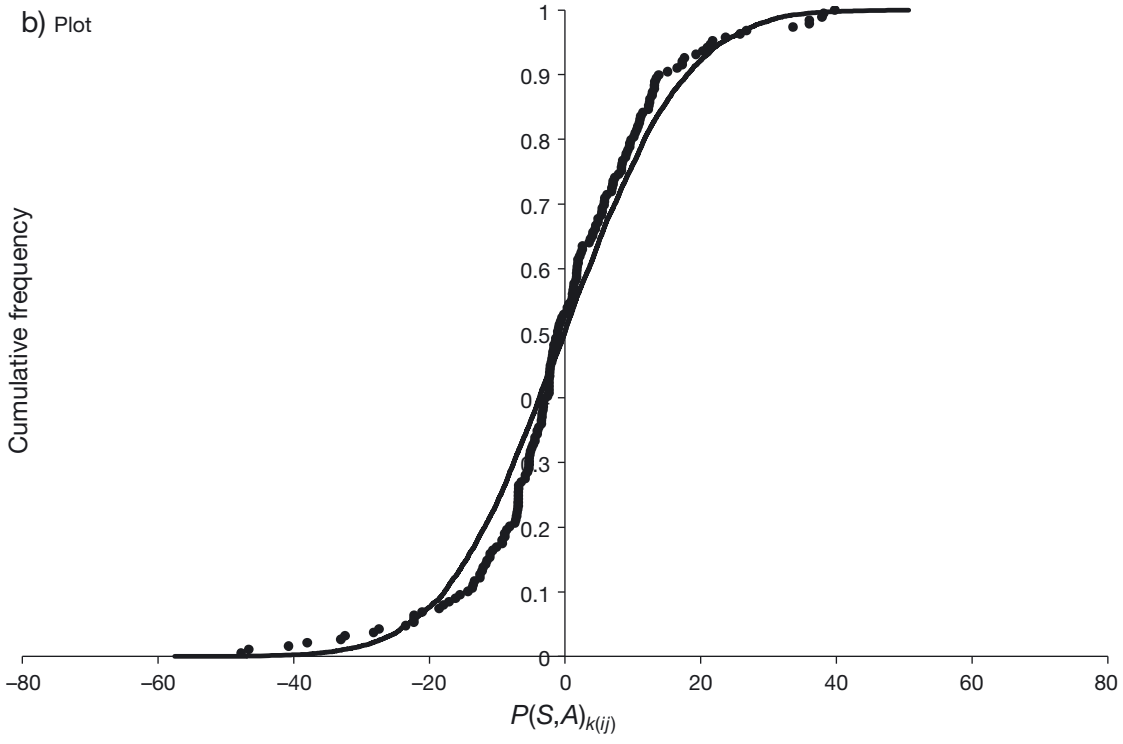
Underwood AJ (1997) Experiments in ecology — their logical design and interpretation using analysis of variance, Vol 1. Cambridge University Press, Cambridge

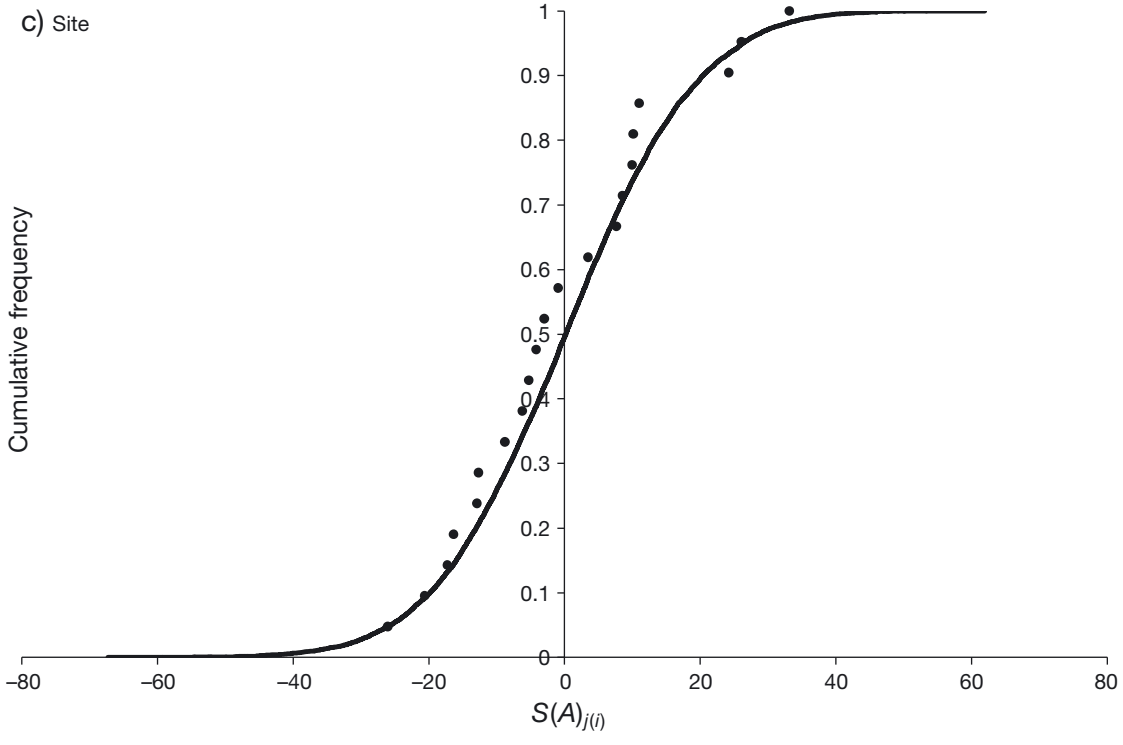
Underwood AJ, Chapman MG (1996) Scales of spatial patterns of distribution of intertidal invertebrates. Oecologia 107:212–224

Wiens JA (1989) Spatial scaling in ecology. Funct Ecol 3: 385–397

Appendix 1. Observed cumulative distributions of estimated deviations at different spatial resolutions using the model $Y_{ijkl} = \mu + A_i + S(A)_{j(i)} + P(S,A)_{k(ij)} + \varepsilon_{ijkl}$. The distributions of observed deviations, i.e. (a) ε_{ijkl}, (b) $P(S,A)_{k(ij)}$ and (c) $S(A)_{j(i)}$, are compared to normal distributions $N(s^2_x, 0)$, where s^2_x is the estimated variance component at the relevant scale for total algal cover.







Editorial responsibility: Lisandro Benedetti-Cecchi, Pisa, Italy

Submitted: February 6, 2012; Accepted: February 8, 2013
Proofs received from author(s): June 12, 2013