Uncertainties in projecting spatial distributions of marine populations

Benjamin Planque^{1*}, Edwige Bellier², and Christophe Loots³

Planque, B., Bellier, E., and Loots, C. 2011. Uncertainties in projecting spatial distributions of marine populations. – ICES Journal of Marine Science, 68: 1045 – 1050.

Received 31 May 2010; accepted 15 December 2010; advance access publication 8 March 2011.

Projection of future spatial distributions of marine populations is a central issue for ecologists and managers. The measure of projection uncertainty is particularly important, because projections can only be useful if they are given with a known and sufficiently high level of confidence. Uncertainties can arise for the observation process, conceptual and numerical model formulations, parameter estimates, model evaluation, appropriate consideration of spatial and temporal scales, and finally the potential of adaptation of living systems. Comprehensive analyses of these multiple sources of uncertainty have not been carried out so far, and how these uncertainties are considered in current studies has not yet been described. To analyse how these different sources of uncertainty are currently considered in marine research, we did a survey of published literature during the period 2005 – 2009. From the 75 publications selected, we calculated how frequently each type of uncertainty was considered. We found that little attention is given to most sources of uncertainty, except for uncertainty in parameter estimates. As a result, most current projections are expected to be far less reliable than usually assumed. The conclusion is that, unless uncertainty can be better accounted for, such projections may be of limited use, or even risky to use for management purposes.

Keywords: complex adaptive systems, conceptual models, model evaluation, multiple hypotheses, scales, spatial distribution, statistical modelling.

Introduction

Anticipating the effects of global change—including climate change—on ecosystems has become of primary importance. More specifically, describing and understanding the factors that currently affect the spatial distribution of marine populations is a central issue for marine ecologists. Therefore, projecting how these spatial distributions may change in the future is of chief concern for managers, conservationists, and human communities that depend on marine resources. The term spatial distribution can be used to define the geographical extent of a marine population, as well as the abundance of the individuals (or the density) within these geographical boundaries. Here, we indiscriminately refer to these two meanings.

Predictive models of spatial distribution must provide some quantification of the uncertainty associated with their projections. The measure of uncertainty is particularly important, because projections can only be useful if they are given with a known and sufficiently high level of confidence (equivalent to a sufficiently low level of uncertainty). Very uncertain projections are expected to be far less informative and useful than less uncertain ones.

We briefly present the major sources of uncertainty when projecting spatial distributions of marine populations. These are related to the observation process, conceptual and numerical model formulations, parameter estimates, model evaluation,

appropriate consideration of spatial and temporal scales, and finally the potential for adaptation of living systems. In predictive models, uncertainty can arise from all these sources, during the successive steps of modelling (Figure 1). The combination of the uncertainties at each of these steps determines the degree of confidence we can have about which projections will bear some relationship with the future state of the world (the "question mark" in Figure 1).

Observation uncertainty

Our perception of the spatial distribution of marine populations, and of the factors that can control it, is uncertain and depends on the observation process. There are no direct means to access the real world, and the way we perceive the marine world is filtered through the lenses of observation methodology, such as observation instruments (e.g. trawl, plankton net, hydroacoustics, video, ...), sampling strategy (e.g. research survey, commercial catch data, market sampling, historical collections), and sampling design (number of samples, replicates, spatial and seasonal distribution of samples, etc.). Because of the limitations of the observation methodology and because there is no single method to observe the marine world adequately, our representation of this world is necessarily incomplete and uncertain.

¹Institute of Marine Research, Postboks 6404, 9294 Tromsø, Norway

²Norwegian Institute for Nature Research (NINA), 7485 Trondheim, Norway

³Formerly Fisheries and Oceans Canada, PO Box 5030, Moncton, NB, Canada E1C 96B; now Ifremer, Laboratoire Ressources Halieutique, F-62321 Boulogne sur Mer, France

^{*}Corresponding Author: tel: +47 77 609721; e-mail: benjamin.planque@imr.no

1046 B. Planque et al.

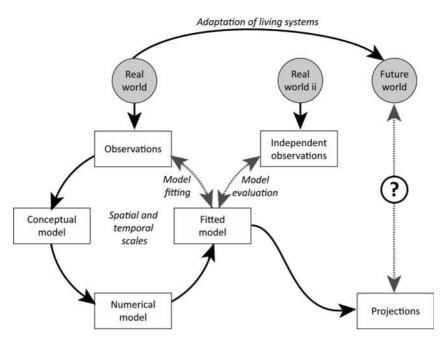


Figure 1. A schematic of the modelling methodology, including major steps and sources of uncertainty. Black arrows indicate the succession of the different steps. Dotted arrows represent the model confrontation with data. The question mark on the right indicates the resulting uncertainty in the projections of future spatial distributions. Adapted from Anderson (2010).

Conceptual uncertainty

For any marine population studied, there can be a high degree of uncertainty about which conceptual model(s) to select. Conceptual models are mental representation of the processes that control the spatial distribution of marine populations. In practice, conceptual models should encapsulate the essential entities and processes of the system of interest. Conceptual models are constructed from intuition and genius, as well as knowledge and experience (Anderson, 2010); therefore, there is no unique way to build them. On the contrary, a range of conceptual models can be constructed to explain observed data. For population spatial distribution, the most common conceptual models have been reviewed in Planque et al. (2011) and include: (i) geographical attachment, (ii) environmental control, (iii) density-dependent habitat selection, (iv) spatial dependence, (v) demographic structure, (vi) species interactions, (vii) persistence, or any combination of the above.

An overview of critical assumptions commonly used when projecting future distribution can also be found in Brander (2009). In the context of climate change, conceptual models are becoming increasingly difficult to construct with confidence. Because environmental conditions are beginning to move outside the window of observed conditions during the past half million years (e.g. oxygen, pH), it is difficult to state to what future populations will become more sensitive. Although such concerns about the incompleteness of models for projecting distribution shifts are not easy to deal with, they cannot be ignored.

Numerical model uncertainty

The choice of the numerical implementation of a conceptual model brings an additional source of uncertainty in the modelling methodology. The development of numerical methods has been rapid in recent years (see Elith and Leathwick, 2009, for a recent and very comprehensive review). Numerical implementations

can vary in the way they represent functional relationships, deal with interactions, non-linearity, and complexity in general, and can accommodate various statistical distributions. Comparative studies of the performance of these different numerical implementations reveal that (i) based on identical datasets and identical conceptual models, distinct numerical implementations can result in different results and different prediction performances and (ii) no numerical method seems to outperform others under every circumstance (Marmion *et al.*, 2009).

Parameter uncertainty

Uncertainty in the estimated parameter values adds to the already mentioned uncertainties in observation, conceptual, and numerical models. For any numerical model, one or several model parameters must be estimated. This is generally achieved during the "model fitting" process. Rather than point estimates, analytical methods or iterative optimization techniques provide confidence intervals for these parameters. This is commonly given in conventional regression models and other less conventional methods, which would be too many to list here (many of these are presented in Hastie *et al.*, 2001, but the list is continually increasing). As the confidence intervals widen, the parameter uncertainty increases.

Model evaluation

Misuse of evaluation methods and current uncertainty about the performance of these methods can result in the uncertain identification of the "best" predictive model(s). Ultimately, model evaluation provides an objective way of measuring model performance, a central step in building predictive models. There are, however, many approaches to implementing model evaluation. These include visual comparison between observations and predictive maps, measures of data-fitting performances, cross-validation techniques, and validation on new sets of independent data. Validation on independent datasets apparently is the most

robust approach (Loots *et al.*, 2010). Within each particular validation approach, numerical implementations have specific properties and requirements (Hastie *et al.*, 2001). There is an ongoing debate on the assumptions underlying the methods and metrics used to evaluate model performance, and methods that were commonly used up to the recent past were inadequate in many instances (Lobo *et al.*, 2007; Jiménez-Valverde *et al.*, 2008). Model evaluation therefore remains an uncertain step.

Spatial and temporal scales

The recognition of scale dependence of ecological processes is necessary to understand the distribution and abundance of organisms (Levin, 1992; McGill, 2010). Conversely, ignoring scale dependence and multiple-scale structures in ecological processes can result in erroneous projections of future species distribution under environmental change (Dormann, 2007; Zurell *et al.*, 2009). Scaling mismatch between the grain size of environmental variables and that of distributional data (i.e. species data) can amplify the uncertainties inherent in each of the datasets (Seo *et al.*, 2009; Wiens *et al.*, 2009). Inference will generally be weaker when based on vague notions of scale than if precise notion of scale is used (Bellier *et al.*, 2007; McIntire and Fajardo, 2009).

Adaptability of living systems

How useful the knowledge gained from past observations will be to predict future changes in marine population distributions is highly uncertain. Anticipated effects of perturbations on ecosystems are commonly derived from past observation of the effects of similar perturbations. However, ecosystems are both complex and adaptive. They present a high degree of non-linearity and a strong dependence on historical contingencies (Levin, 1998, 2002, 2005). The assumption that future responses will resemble past ones is therefore unlikely to hold usually, at least beyond a certain time horizon.

In terrestrial and freshwater ecology, recent works, such as those by Thuiller (2004), Heikkinen *et al.* (2006), Dormann *et al.* (2008), and Buisson *et al.* (2010), have examined the uncertainty associated with spatial projection, but these analyses were restricted to observation uncertainties, numerical formulation, or uncertainties in future climate based on current climate scenarios (an issue not discussed here). Comprehensive analyses of the multiple sources of uncertainty presented here have not been carried out so far, and how these uncertainties are considered in current studies has not yet been described. It is clear whether the marine research community shares a common approach to this issue or rather specific individual approaches and whether all sources of uncertainty are well understood and adequately taken into consideration.

In this paper, we analyse how these different sources of uncertainty are currently considered in marine research. More specifically, we measure how different sources of uncertainty are accounted for in the literature published during the past 5 years in the field of marine population spatial distribution.

Material and methods

We did a literature survey using the Thomson–Reuters ISI Web of Knowledge database with the following set of criteria: (spatial or geograph* or distribution* or habitat) and (fish* or benth*) and (sea, ocean, or coast* or marin*) and model*. The selection was restricted to articles published from 2005 to early March 2010

and within the fields of "marine and freshwater biology" or "oceanography" or "fisheries". In all, 1137 articles were found to match these criteria. From these 1137 articles, we selected those presenting models that were (or could be) used for the projection of spatial distribution of marine populations. This amounted to 75 publications.

For each article selected, we assessed whether the following criteria were addressed.

Observation uncertainty

We checked whether the uncertainty linked to observations was explicitly measured or modelled. We further checked whether an observation model had been considered, e.g. whether the observation process was included in the model design.

Conceptual uncertainty

We considered which of the seven common conceptual models (geographical attachment, environmental control, density-dependent habitat selection, spatial dependence, demographic structure, species interactions, and persistence) had been considered. We also verified whether other types of conceptual model were presented. In addition, we checked whether the uncertainty linked to the choice of particular conceptual model was considered in the model construction and evaluation.

Numerical model uncertainty

We checked, for each of the conceptual models that was considered, whether several numerical modelling methodologies had been compared (e.g. linear models, non-linear models, smoothing functions, regression trees).

Parameter uncertainty

We checked whether the uncertainty linked to the estimation of the parameters was presented, e.g. by providing values or plots of confidence intervals.

Model evaluation

We checked, which method(s) was (were) used to evaluate model performance: visual comparison between observations and predictive maps, measures of data-fitting performances, cross-validation techniques, and validation on new sets of independent data.

Spatial and temporal scales

We reported whether the scale of investigation was considered implicitly or explicitly. Implicit account of scale was reported when the authors defined the scale of investigation before carrying out the modelling. Explicit account of scale was reported when appropriate scale(s) was (were) quantitatively determined during the modelling process, rather than defined in advance.

Adaptability of living systems

We verified whether the authors discussed possible adaptations of the populations to future external conditions. We also reported whether such adaptation was explicitly accounted for.

The results were summarized by calculating the proportion of research articles dealing with each of the above points.

Results and discussion

Results from the literature survey (Table 1) indicate that, on average, little attention is given to the various sources of uncertainties in models and consequently to uncertainties in the resulting

B. Planque et al.

Table 1. Results of the literature survey: the right column reports the percentage of publications (n = 75) fitting each criterion related to different types of uncertainty.

Uncertainty type	Evaluation criteria	Percentage
Observation	Observation uncertainty	7
	Observation model	1
Conceptual	Geographic attachment	19
models	Environmental conditions	95
	Density-dependent habitat selection	11
	Spatial dependence	27
	Demographic structure	9
	Species interactions	7
	Persistence	7
	Conceptual model uncertainty	3
Numerical model	Uncertainty	24
Parameter	Uncertainty	69
Model evaluation	Visual	8
	Data-fitting performance	45
	Cross-validation	36
	Prediction performance	24
Scale	Implicit	45
	Explicit	12
Adaptability	Discussed	4
	Explicit	0

projections, despite the recent and rapid increase in the number of publications presenting models of spatial distribution.

Observation uncertainty

This was considered in only 5 of the 75 studies we reviewed and only a single study explicitly accounted for the observation process in the model design. This is rather surprising, given that observation process modelling is designed to represent the sampling process, therefore allowing for accurate estimates of animal densities (Lewy and Kristensen, 2009). Recent advances in state–space modelling that have found application for observation modelling in animal movements (Patterson *et al.*, 2008) have not yet been applied to population distribution models.

Conceptual model uncertainty

This was only accounted for in one of the studies surveyed. Conceptual uncertainty is associated with mental representation of the processes that control the spatial distribution of marine populations and may consequently be perceived as a qualitative uncertainty (compared with more quantitative parameter uncertainties; see below). However, uncertainties in conceptual models form the basis of the entire modelling process and model evaluation procedures (Jiménez-Valverde et al., 2008). We argue that it is possible to evaluate such uncertainties quantitatively (see, for example, suggestions in Planque et al., 2011), and this has been demonstrated in a few case studies in the North Sea, in the literature more recent than that analysed in our survey (Loots et al., 2010). Detailed analysis of the conceptual models commonly used indicates that models based on environmental controls vastly dominate marine literature (95%). Except spatial dependence—often analysed as spatial autocorrelation all other hypotheses are considered in less than one-quarter of the studies. Therefore, most studies assume (at least implicitly) that environmental drivers must override all other drivers of population distribution, although such assumption is by no means obvious and usually not demonstrated.

Numerical model uncertainty

Uncertainty in the appropriateness of the numerical formulation is addressed in one-quarter of the articles. Although this proportion remains relatively low, it is a greater proportion than the above-mentioned sources of uncertainty. This may be partly explained by the recent publication of comparative studies of numerical models (Elith *et al.*, 2006; Meynard and Quinn, 2007; Tsoar *et al.*, 2007, among others) that have raised attention to this specific issue.

Parameter uncertainty

This was accounted for in 69% of the literature analysed. This is by far the most commonly considered source of uncertainty. It could be attributed to the statistical methods used to model species distribution. These are usually parametric methods that provide confidence intervals of the parameters (Rushton *et al.*, 2004; Guisan and Thuiller, 2005). Such confidence intervals are easily interpretable and provide a direct quantitative estimate of uncertainty that can be carried over to the model projections.

Model evaluation

Most of the studies we analysed were based on statistical models; therefore, model evaluation based on the visual comparison of predicted and observed distributions was infrequent. Rather, authors often used quantitative evaluations based on fitting, crossvalidation, or prediction performance. Fitting performance remains widely used (45%), despite known problems associated with spatially autocorrelated data and overfitting (Hastie et al., 2001), which are rarely accounted for. Cross-validation has been advocated as an efficient alternative to fitting performance metrics (Hastie et al., 2001) and is now available in many modelling packages (e.g. the general cross-validation score in Wood, 2006). However, when data are strongly autocorrelated, crossvalidation suffers from problems similar to data fitting and will usually result in the selection of complex models with low bias and high variance, i.e. with low predictive power (Telford and Birks, 2005). Model evaluation on independent data was only recorded in one quarter of the studies, a moderate score, considering that this approach is likely to be the most robust (Planque et al., 2011).

Spatial and temporal scales

These were defined before modelling in 45% of the literature reviewed. Although this may appear as a high percentage, it indicates that more than half the papers reviewed did not report the scale on which processes were modelled or why such scale(s) had been selected. Explicit account of scale during the modelling process was obviously less common and was only found in 12% of the articles reviewed. This is again a surprisingly low percentage, given the importance of scale definition and investigation in understanding and modelling ecological processes (Levin, 1992; McGill, 2010).

Adaptability of living systems

We found that only a small fraction of published studies (4%) discussed or simply mentioned possible implications of ecological adaptability for projected changes in population distribution.

Not surprisingly, none of the articles reviewed implemented such processes in the modelling methodology.

The main result from this review is that uncertainty in spatial projections has been poorly considered in marine ecological research. Although we did not carry out a similar review for other ecological work, we found that the literature dealing with the types of uncertainty we discussed here almost exclusively pertains to terrestrial ecology (as attested by the bibliographic references). This is worrying, because it indicates that the current projections of changes in marine biota distributions are likely poorly reliable. It is also promising because many concepts and methods are readily available from terrestrial ecology, for implementation in marine systems. Based on the presentations at the PICES Conference on Climate Change Effects on Fish and Fisheries in Sendai, Japan, we believe that there is an increasing trend in explicitly handling various sources of uncertainty in model projections. However, a more extensive study or a repetition of the current study would be required to confirm this

In this limited literature survey, we voluntarily left out uncertainties in climate models, but this component must also be considered in any serious attempt to project realistic future spatial distributions, particularly when downscaling climate models. Recent analyses of ocean–climate-projection uncertainties have been presented by Hollowed *et al.* (2009) and Wang *et al.* (2010). An example of inclusion of climate model uncertainties in freshwater fish distribution forecasts can be found in Buisson *et al.* (2010).

Highly uncertain or inaccurate projections could prove detrimental rather than useful to supporting management decisions. Therefore, the many sources of uncertainty should be considered carefully and explicitly for projections of future spatial distributions of marine populations to be useful for managers, but we found that this has not yet been the case. Model developments should include a systematic qualitative description of uncertainties for all the steps highlighted in Figure 1, as well as an evaluation of the relative contributions of the different sources of uncertainty. In addition, models developed to produce projections should at least be evaluated on their ability to predict (not just fit) observed past changes in population distributions. This would provide an essential objective assessment of model performance and potential predictive capabilities.

Conclusions

The majority of current modelling efforts to project future spatial distribution of marine populations largely ignore most sources of uncertainties. As a result, most currently available projections are presented as much more precise that they actually are. We conclude that, unless uncertainty can be better accounted for, such projections may be of limited use, or even risky to use for management purpose.

Acknowledgments

The authors thank ICES, PICES, and FAO for allowing them to present the results of their work at the Symposium on Climate Change Effects on Fish and Fisheries and in the current volume. EB was supported by the Norwegian Research Council (WFR 185109). Frida Lasram, who contributed to early discussions and analysis presented in this work, and two anonymous reviewers are thanked for their expertise and advice that helped to improve earlier versions of this article.

References

- Anderson, T. R. 2010. Progress in marine ecosystem modelling and the "unreasonable effectiveness of mathematics". Journal of Marine Systems, 81:4-11.
- Bellier, E., Monestiez, P., Durbec, J-P., and Candau, J-N. 2007. Identifying spatial relationships at multiple scales: principal coordinates of neighbour matrices (PCNM) and geostatistical approaches. Ecography, 30: 385–399.
- Brander, K. 2009. Impacts of climate change on marine ecosystems and fisheries. Journal of the Marine Biological Association of India, 51: 1–13.
- Buisson, L., Thuiller, W., Casajus, N., Lek, S., and Grenouillet, G. 2010. Uncertainty in ensemble forecasting of species distribution. Global Change Biology, 16: 1145–1157.
- Dormann, C. F. 2007. Promising the future? Global change projections of species distributions. Basic and Applied Ecology, 8: 387–397.
- Dormann, C. F., Schweiger, O., Arens, P., Augenstein, I., Aviron, S., Bailey, D., Baudry, J., *et al.* 2008. Prediction uncertainty of environmental change effects on temperate European biodiversity. Ecology Letters, 11: 235–244.
- Elith, J., Graham, C. H., Anderson, R. P., Dudìk, M., Ferrier, S., Guisan, A., Hijmans, R. J., et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography, 29: 129–151.
- Elith, J., and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology Evolution and Systematics, 40: 677–697.
- Guisan, A., and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8: 993–1009.
- Hastie, T., Tibshirani, R. J., and Friedman, J. H. 2001. The Elements of Statistical Learning. Springer-Verlag, New York.
- Heikkinen, R. K., Luoto, M., Araujo, M. B., Virkkala, R., Thuiller, W., and Sykes, M. T. 2006. Methods and uncertainties in bioclimatic envelope modelling under climate change. Progress in Physical Geography, 30: 751–777.
- Hollowed, A. B., Bond, N. A., Wilderbuer, T. K., Stockhausen, W. T., A'mar, Z. T., Beamish, R. J., Overland, J. E., et al. 2009. A framework for modelling fish and shellfish responses to future climate change. ICES Journal of Marine Science, 66: 1584–1594.
- Jiménez-Valverde, A., Lobo, J. M., and Hortal, J. 2008. Not as good as they seem: the importance of concepts in species distribution modelling. Diversity and Distributions, 14: 885–890.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology, 73: 1943–1967.
- Levin, S. A. 1998. Ecosystems and the biosphere as complex adaptive systems. Ecosystems, 1: 431–436.
- Levin, S. A. 2002. Complex adaptive systems: exploring the known, the unknown and the unknowable. Bulletin of the American Mathematical Society, 40: 3–19.
- Levin, S. A. 2005. Self-organization and the emergence of complexity in ecological systems. BioScience, 55: 1075–1079.
- Lewy, P., and Kristensen, K. 2009. Modelling the distribution of fish accounting for spatial correlation and overdispersion. Canadian Journal of Fisheries and Aquatic Sciences, 66: 1809–1820.
- Lobo, J. M., Jimenez-Valverde, A., and Real, R. 2007. AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography, 17: 145–151.
- Loots, C., Vaz, S., Planque, B., and Koubbi, P. 2010. What controls the spatial distribution of the North Sea plaice spawning population? Confronting ecological hypotheses through a model selection framework. ICES Journal of Marine Science, 67: 244–257.
- Marmion, M., Luoto, M., Heikkinen, R. K., and Thuiller, W. 2009. The performance of state-of-the-art modelling techniques depends on geographical distribution of species. Ecological Modelling, 220: 3512–3520.

B. Planque et al.

- McGill, B. J. 2010. Matters of scale. Science, 328: 575-576.
- McIntire, E. J. B., and Fajardo, A. 2009. Beyond description: the active and effective way to infer processes from spatial patterns. Ecology, 90: 46–56.
- Meynard, C. N., and Quinn, J. F. 2007. Predicting species distributions: a critical comparison of the most common statistical models using artificial species. Journal of Biogeography, 34: 1455–1469.
- Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O., and Matthiopoulos, J. 2008. State-space models of individual animal movement. Trends in Ecology and Evolution, 23: 87–94.
- Planque, B., Loots, C., Petitgas, P., Lindstrøm, U., and Vaz, S. 2011. Understanding what controls the spatial distribution of fish populations using a multi-model approach. Fisheries Oceanography, 20: 1–17.
- Rushton, S. P., Ormerod, S. J., and Kerby, G. 2004. New paradigms for modelling species distributions? Journal of Applied Ecology, 41: 193–200.
- Seo, C., Thorne, J. H., Hannah, L., and Thuiller, W. 2009. Scale effects in species distribution models: implications for conservation planning under climate change. Biology Letters, 5: 39–43.

- Telford, R. J., and Birks, H. J. B. 2005. The secret assumption of transfer functions: problems with spatial autocorrelation in evaluating model performance. Quaternary Science Reviews, 24: 2173–2179.
- Thuiller, W. 2004. Patterns and uncertainties of species' range shifts under climate change. Global Change Biology, 10: 2020–2027.
- Tsoar, A., Allouche, O., Steinitz, O., Rotem, D., and Kadmon, R. 2007. A comparative evaluation of presence-only methods for modelling species distribution. Diversity and Distributions, 13: 397–405.
- Wang, M., Overland, J. E., and Bond, N. A. 2010. Climate projections for selected large marine ecosystems. Journal of Marine Systems, 79: 258–266.
- Wiens, J. A., Stralberg, D., Jongsomjit, D., Howell, C. A., and Snyder, M. A. 2009. Niches, models, and climate change: assessing the assumptions and uncertainties. Proceedings of the National Academy of Sciences of the USA, 106: 19729–19736.
- Wood, S. N. 2006. Generalized Additive Models. An Introduction with R. Chapman and Hall, London.
- Zurell, D., Jeltsch, F., Dormann, C. F., and Schröder, B. 2009. Static species distribution models in dynamically changing systems: how good can predicitions really be? Ecography, 32: 733–744.