

Received Date : 02-Mar-2015  
Revised Date : 11-Jun-2015  
Accepted Date : 17-Jun-2015  
Article type : Review  
Editor : Marc Cadotte

## PREDICTIVE ECOLOGY IN A CHANGING WORLD

Nicolas Mouquet (1)\*, Yvan Lagadeuc (2), Vincent Devictor (1), Luc Doyen (3), Anne Duputié (4), Damien Eveillard (5), Denis Faure (6), Eric Garnier (7), Olivier Gimenez (7), Philippe Huneman (8), Franck Jabot (9), Philippe Jarne (7), Dominique Joly (10), Romain Julliard (11), Sonia Kéfi (1), Gael J. Kergoat (12), Sandra Lavorel (13), Line Le Gall (14), Laurence Meslin (1), Serge Morand (1), Xavier Morin (7), Hélène Morlon (15), Gilles Pinay (2), Roger Pradel (7), Frank M. Schurr (1,16), Wilfried Thuiller (13) and Michel Loreau (17).

- (1) Institut des Sciences de l'Evolution, Université de Montpellier, CNRS, IRD, EPHE, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France.
- (2) ECOBIO, UMR 6553, CNRS - Université de Rennes 1, F-35042 Rennes Cedex, France.
- (3) Groupement de Recherche en Économie Théorique et Appliquée (GREThA), CNRS UMR 5113, Université de Bordeaux, Avenue Léon Duguit, 33608 Pessac cedex, France.
- (4) Unité Evolution Ecologie Paléontologie, UMR CNRS 8198, Université de Lille 1 - Sciences et Technologies, Villeneuve d'Ascq, France.
- (5) Computational Biology group, LINA, UMR 6241, CNRS - EMN - Université de Nantes, Nantes, France.
- (6) Institut for Integrative Biology of the Cell (I2BC), CNRS CEA Université Paris-Sud, Saclay Plant Sciences, Avenue de la Terrasse, 91198 Gif-sur-Yvette, France.
- (7) Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, CNRS – Université de Montpellier - Université Paul-Valéry Montpellier - EPHE, 1919 Route de Mende, 34293 Montpellier Cedex 05, France.
- (8) Institut d'Histoire et de Philosophie des Sciences et des Techniques, UMR 8590 CNRS, Université Paris 1 Sorbonne, 13, rue du Four 75006 Paris, France.
- (9) Laboratoire d'Ingénierie des Systèmes Complexes, UR, IRSTEA, 9 avenue Blaise Pascal, F-63178 Aubière, France.
- (10) Laboratoire Evolution, Génomes, Comportement, Ecologie, UMR9191 CNRS, 1 avenue de la Terrasse, bâtiment 13, 91198 Gif-sur-Yvette Cedex, and Université Paris-Sud, 91405 Orsay, France.
- (11) Centre d'Ecologie et des Sciences de la Conservation, UMR 7204, MNHN-CNRS-UPMC, 55 rue Buffon, 75005, Paris, France.
- (12) Centre de Biologie pour la Gestion des Populations, UMR 1062, INRA - IRD – CIRAD - Montpellier SupAgro, 755 Avenue du campus Agropolis, 34988 Montferrier/Lez, France.
- (13) Laboratoire d'Ecologie Alpine (LECA), Univ. Grenoble Alpes, CNRS F-38000 Grenoble, France.
- (14) Institut de Systématique, Evolution, Biodiversité, Muséum National d'Histoire Naturelle, UMR 7205, CNRS-EPHE-MNHN-UPMC, 57 rue Cuvier, 75231 Paris, France.
- (15) Institut de Biologie, Ecole Normale Supérieure, UMR 8197 CNRS, 46 rue d'Ulm, 75005 Paris.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1365-2664.12482

This article is protected by copyright. All rights reserved.

- (16) Institute of Landscape and Plant Ecology, University of Hohenheim, 70593 Stuttgart, Germany.
- (17) Centre for Biodiversity Theory and Modelling, Station d'Ecologie Expérimentale, CNRS, 09200 Moulis, France.

\* Correspondence should be addressed to nmouquet@univ-montp2.fr

## SUMMARY:

1. In a rapidly changing world, ecology has the potential to move from empirical and conceptual stages to application and management issues. It is now possible to make large-scale predictions up to continental or global scales, ranging from the future distribution of biological diversity to changes in ecosystem functioning and services. With these recent developments, ecology has a historical opportunity to become a major actor in the development of a sustainable human society. With this opportunity, however, also comes an important responsibility in developing appropriate predictive models, correctly interpreting their outcomes, and communicating their limitations. There is also a danger that predictions grow faster than our understanding of ecological systems, resulting in a gap between the scientists generating the predictions and stakeholders using them (conservation biologists, environmental managers, journalists, policymakers).
2. Here we use the context provided by the current surge of ecological predictions on the future of biodiversity to clarify what prediction means, and to pinpoint the challenges that should be addressed in order to improve predictive ecological models and the way they are understood and used.
3. *Synthesis and applications.* Ecologists face several challenges to ensure the healthy development of an operational predictive ecological science: (i) clarity on the distinction between explanatory and anticipatory predictions; (ii) developing new theories at the interface between explanatory and anticipatory predictions; (iii) open data to test and validate predictions; (iv) making predictions operational and (v) developing a genuine ethics of prediction.

**Keywords:** anticipation, big data, biodiversity, ecological prediction, ecosystems, ethics, forecast, global change, modelling, theory.

## INTRODUCTION

Prediction is not new in ecology (Fig. 1). Growing from a purely observational discipline (the so-called “natural history”) to a modern scientific field, ecology has often relied upon predictions to test its fundamental theories (Hilborn & Mangel 1997). However, accumulating predictions to test theories is no longer sufficient. As in other fields such as climate or health sciences, the increasing social and political awareness of the importance of global environmental changes has prompted a shift from explanatory to anticipatory predictions (see section 2 for definitions) of the trajectories of complex ecological systems. Projections of biodiversity loss or changes and ecosystem functioning at the global scale (e.g., Thuiller *et al.* 2011; Barnosky *et al.* 2012) are increasingly common (Fig. 2). While a consensus has yet to be reached among ecologists regarding the drivers of species richness and ecosystem functioning (Loreau 2010), this sudden rise of ecological predictions (Fig. 1) has become a strong incentive for an entire research community in ecology (Coreau *et al.* 2009; Bellard *et al.* 2012; Gerrish & Sniegowski 2012 ; Evans *et al.* 2013a; Thuiller *et al.* 2013; Harfoot *et al.* 2014b).

The shift in the type of ecological predictions and their scales of applicability is also motivated by the unprecedented quantity of ecological data and the complexity of the statistical and modelling tools now available (Purves *et al.* 2013). Because ecology shares principles and methods with many other disciplines, from mathematics and computer science to environmental and social sciences, the origin and scale of data are more mixed than before, with elements from (among others) biogeography, ecophysiology, ecosystem functioning, environmental sciences, genetics, metagenomics, networks ecology, and socio-economics. This heterogeneity in the nature and origin of data requires the building of new tools to integrate those various scales. While the accessibility and quality of the data is still an issue (Costello *et al.* 2013), this “dataavalanche” has inevitably modified our way to conduct research in ecology. It is very likely that this trend will be amplified by the rise of biodiversity surveys, citizen science programmes and metagenomic sampling programmes. Interaction between various disciplines and the unprecedented amount of data has opened the way to what has recently been referred to as “predictive systems ecology” (Evans *et al.* 2013a). However, this trend also raises the critical

Accepted Article

issue that the predictions provided by the ecological scientific community might go far beyond our actual understanding of ecological systems. Clarifying the concept of prediction in ecology is also important if ecologists want to build predictive models upon appropriate theory and data, and for society to interpret these predictions with their underlying limits and uncertainties.

Ecological predictions concern a variety of objects and biological scales, from genetic variability to species richness, community composition, ecosystem functioning, and biogeochemical cycles. However, the processes at stake at each of these levels are not necessarily commensurable. For instance, predicting the invasion speed of a particular invasive species (Richter *et al.* 2013) is different from predicting how many species may be lost on Earth in the upcoming century (Pereira *et al.* 2010), the potential impact of global changes on ecosystem functioning (Harfoot *et al.* 2014b) and services (Worm *et al.* 2006), or whether and when planetary regime shifts might occur (Barnosky *et al.* 2012). While there has been some progress in developing ‘unified’ theories and models across temporal and spatial scales (McGill 2010; Chave 2013; Harfoot *et al.* 2014b), the data required to calibrate the models and test each of these predictions are not always available. For example, predicting the short-term fate of a species requires models in both population genetics and dynamics along with life-history and ecophysiological trait data (Coulson *et al.* 2001), while predicting the pace of adaptive evolution and the speciation and extinction dynamics of whole clades requires micro- and macro-evolutionary models along with phylogenetic and fossil data (e.g., Condamine, Rolland & Morlon 2013). Additionally, the accuracy of a prediction is also highly dependent on its scale and scope. For example, while it is possible to predict successful invader fish species in the Great Lakes with more than 80% accuracy (Kolar & Lodge 2002), the uncertainty around estimates of metrics as ‘simple’ as forest cover by year 2050 can be disconcerting (Pereira *et al.* 2010).

Our review aims to clarify what prediction implies in ecology, and to pinpoint some of the challenges in empirical and theoretical ecology that need to be addressed to improve predictive ecological models and the way they are understood and used by human society. Reviewing the nature and limits of predictions in ecology per se, would be too ambitious for a single review. We have rather chosen to

Accepted Article

illustrate some of the limitations and future directions concerning the question of predicting the consequences of global change on biodiversity and ecosystem functioning. We have organized our main points along five axes that we believe are essential for improving predictive ecology: (i) the distinction between explanatory and anticipatory predictions, (ii) the need for the development of new theories at the interface between explanatory and anticipatory predictions, (iii) the need for data to test and validate predictions, (iv) the challenge of operationality, and (v) the importance of developing a genuine ethics of prediction. While some of these axes have already been discussed independently in the past (e.g., Peter 1991), we feel it is important to develop them jointly as ecology is now trying to move from empirical and conceptual stages towards operationality.

## EXPLANATORY VERSUS ANTICIPATORY PREDICTIONS

“To predict” means to make a statement on what should be observed in a particular system before making the actual observation. It is important to acknowledge that there are at least two different kinds of predictions, which we call *explanatory* and *anticipatory* predictions.

Theories in science are based on hypotheses, which are general propositions about the systems under study; they formulate what should be expected if the assumptions stemming from general theoretical constructs are correct. Such expectations about individual systems, outcomes or properties are called here *explanatory predictions*. When they are consistently and repeatedly corroborated by data, we consider the hypothesis or the theory to be provisionally valid (Popper 1959). In contrast, when they differ or diverge from the data, the hypotheses and theory need to be modified. In this hypothetico-deductive reasoning, the validation process determines the explanatory role of predictions to test or compare theories. These predictions are therefore logical consequences of the models built on the hypotheses; they are not in principle limited by what is currently observed or observable. In physics, for instance, the Higgs boson was a prediction long before it became an observation.

However, in some cases, predictions are not elaborated in the hypothetico-deductive reasoning loop but only in the future tense: they are about what the world will be, assuming that the theory or the putative link between causes and effects that we are considering is “true”. Climate predictions by the Intergovernmental Panel on Climate Change (IPCC) are a good example of such predictions: they do not only predict what already is in our data sets regarding temperature averages, etc., but they also provide projections on alternative future states of the Earth under different greenhouse gas emission scenarios. Accordingly, we distinguish those predictions that are essentially concerned with anticipation from those that aim to establish explanations. *Anticipatory predictions* come in various nuances, labelled forecasts, projections and scenarios (IPCC guidance terminology, Carter & La Rovere 2001). Taking the example of species distributions, explanatory predictions aim to corroborate hypotheses on the mechanisms underpinning species distributions derived from a specific theoretical corpus. By contrast, anticipatory predictions concern likely future changes in species distributions and are decoupled from the explanatory process: they are not necessarily based on a mechanistic understanding of the forces driving the observed changes (i.e. the so-called phenomenological models; Appendix S1 in Supporting Information).

While explanatory predictions are necessarily testable, anticipatory predictions need not be, essentially because a temporal trend or pattern has to be predicted that goes beyond the predictive power of phenomenological approaches. A striking example of the difficulty in generating anticipatory predictions is the current surge of predictions to estimate and anticipate the effects of environmental change on species ranges and diversity (Thuiller *et al.* 2013). Many different approaches can be taken to build such models, including (among others) species-energy theory (e.g., Storch, Marquet & Brown 2007), correlative species distribution models or “niche models” (e.g., Guisan & Thuiller 2005) or mechanistic models of demographic processes with explicit temperature dependence (e.g., Sitch *et al.* 2008). These models rely on distinct underlying hypotheses and may lead to different predictions at different temporal and spatial scales or hierarchical levels (e.g., Morin & Thuiller 2009). Interestingly, even models relying on a single theory can lead to contrasted predictions. For instance, different correlative species distribution models, all derived from a simplified version of ecological niche theory (Guisan & Thuiller 2005), are

known to yield different predictions when applied to future conditions (Thuiller 2004; Pearson *et al.* 2006).

In summary, anticipatory predictions differ from explanatory predictions in that they do not aim at testing models and theory. They rely on the assumption that underlying hypotheses are valid while explanatory predictions are based on hypotheses to be tested. Anticipatory predictions are also not necessarily supposed to be true. Instead they intend to deduce, from the models, future states of reality (forecasts), to extrapolate these models to domains where there is some uncertainty about the main parameter values (projections), and to describe possible trajectories or behaviours of the real system, depending upon a choice of parameter values that are likely to be impacted by human action (scenarios). These anticipatory predictions are therefore not meant to describe the actual future; the fact that they do not match reality does not count against the validity of the underlying hypotheses; rather they should be regarded as a guide for present action (Harfoot *et al.* 2014a).

The two types of predictions sketched here are the two ends of a continuum along which predictive practices in ecology are positioned. They define distinct uses and requirements for the same logical entities, namely, conditional statements deduced from models based on specific hypotheses. Because of their differences, the two types of predictions should not be assessed in the same way; their scope, strengths and weaknesses are different. There is today an ambiguity in the label “predictive ecology”; it is often used to mean anticipatory predictions, but the specificities of this type of predictions are often overlooked.

## **FROM EXPLANATORY TO ANTICIPATORY PREDICTION AND BACK**

Ecology is often blamed for its weakness at generating predictions, despite the accumulation of specific tools to account for complex dynamics (Levin 1992; Solé & Goodwin 2000; Anand *et al.* 2010). Surprises, i.e. the occurrence of unexpected responses in an experiment or observation are frequent and useful because they may yield new ideas. As discussed in the previous section, although some predictions can result from mere empirical correlations, all predictions are based on an implicit or explicit theoretical framework that determines their scope and

limitations. In the absence of theory, no prediction is possible. Rudimentary theories may be qualitative, but more advanced theories involve mathematical models and quantitative explanatory predictions.

### **Linking explanatory to anticipatory predictions**

Although theories are often first developed to connect and integrate concepts, hypotheses, models and data, they virtually always lead to new explanatory predictions, which in turn can be used to develop anticipatory predictions. Thus, there is a continuum from explanatory to anticipatory predictions in the process of theory development. A good example of this process is provided by the theory of trophic cascades. This theory originated from empirical observations and a conceptual hypothesis about the prevalence of top–down control in food chains (Hairston, Smith & Slobodkin 1960). This hypothesis was then turned into heuristic theoretical models (e.g., Oksanen *et al.* 1981) that made new predictions about the expected responses of ecosystems to perturbations at the bottom or at the top of the food chain. After several successful experimental tests of these explanatory predictions, trophic cascade theory was applied in lake restoration programmes in the form of biomanipulation. Although only mixed results were obtained from biomanipulation experiments, this example illustrates that reaching the operational stage of anticipatory predictions does not preclude the need for further theoretical developments. Quite the contrary, the failure of some biomanipulation programmes led to the development of new theoretical models that took into account the functional complexity of food webs (e.g., Hulot & Loreau 2006; Wollrab, Diehl & De Roos 2012). Similar processes took place in many other areas of ecology, such as in the theory of host–parasite interactions, which started with simple heuristic models and was later successfully applied to predict the propagation of human diseases and define vaccination thresholds in public health programmes (May & Anderson 1991). In these examples, we see the same initial progression from concepts to heuristic models and from explanatory to anticipatory predictions, followed by a stage in which the theory is re-examined and developed in new directions.



### **The need for more complexity...**

Ecological systems are complex; they typically include various components interacting with each other in different ways at various spatial and temporal scales. Moreover, the biotic components of ecological systems can evolve in a Darwinian way (Ings *et al.* 2009), and interactions between the biotic and abiotic components may create feedback loops at the origin of nonlinear, hence unexpected, ecosystem responses (Scheffer *et al.* 2001). Intuitive reasoning suggests that taking into account an increasing number of details would improve the accuracy of ecological predictions, but it is not necessarily the case: complex computer codes are prone to various artefacts that need to be taken care of (Galán *et al.* 2009; Augusiak, Van den Brink & Grimm 2014). This highlights an important gap in our current understanding of the link between model complexity (see Appendix S1) and predictive accuracy, although statistical criteria exist to compare the predictive accuracy of sets of models (e.g., Burnham & Anderson 2002).

Filling this gap is critical to understand when additional model complexity should be sought, and of what kind. Such research is timely as some researchers advocate more complex modelling approaches (e.g., Evans *et al.* 2013b). According to them, progress in ecology has been hampered by an excessive focus on simple models, which fail to adequately capture important processes driving ecosystem dynamics. Better integration of adjacent organization levels has been considered as a pathway to better theory and models (Allen & Hoekstra 1992). This implies on the one hand the incorporation of processes acting at several organization levels (Grimm *et al.* 2005; Chevin, Lande & Mace 2010; Thuiller *et al.* 2013) into integrative frameworks and on the other hand, a more explicit integration of responses and feedbacks to external drivers, such as the dynamics of ecosystems at larger spatial scales (Loreau, Mouquet & Gonzalez 2003; Massol *et al.* 2011), or the socio-economic drivers of ecosystem change (Liu *et al.* 2007; Kleijn *et al.* 2009).

The use of complex models, however, brings other challenges. Understanding the behaviour of these models becomes difficult, since they commonly lead to emergent effects that could not have been predicted from the knowledge of their building blocks alone (Grimm *et al.* 2005). Calibrating the models, assessing their sensitivity to some assumptions (Saltelli *et al.* 2008; Augusiak, Van den Brink &

Grimm 2014), and measuring data uncertainty are also challenging (Hartig *et al.* 2011). The analysis of complex models is the focus of intense research, not only in biology (Wilkinson 2009), but also in climatology (Edwards & Marsh 2005), industry (Lorenzo *et al.* 2011) and statistics (Kennedy & O'Hagan 2001), with an increasing number of software facilities to disseminate state of the art techniques (e.g., Jabot, Faure & Dumoulin 2013). Ecological research could benefit from current advances in other complex system fields to deal with its current inclination towards complexity (Borgatti *et al.* 2009).

### **... or for less complexity?**

While there is a need for including some ecological complexity in predictive ecology, this should not come at the cost of understanding and tractability. The paradox is that to address complexity, we need to simultaneously simplify our ecological understanding. Genes and species have been the most studied units of organization in ecology and evolution, and as a result of the wealth of knowledge about the mechanisms underpinning their dynamics, they remain at the heart of predictive ecology. However, the focus on these facets of biodiversity poses several problems for prediction, as the knowledge of, say, species responses to environmental change is not sufficient to predict the assembly of novel communities (Suding *et al.* 2008). Recently, complementary approaches have been proposed to use other units of organization in predictive models. Here we provide two examples where new units of biodiversity have been used to provide large-scale predictions of biodiversity and ecosystem functioning.

Functional traits and their quantification through functional diversity (Lavorel & Garnier 2002; McGill *et al.* 2006) are potentially powerful tools for the prediction of future patterns of biodiversity and ecosystem functioning as they link the successive steps that go from the projection of species distributions to the assembly of novel communities and ecosystem functioning. For instance, functional traits have been used to model responses to large-scale environmental changes in plant species distributions (e.g., Reu *et al.* 2011). Once potential distributions are known according to environmental factors, trait-based models of community assembly can be applied to predict community composition within a given trophic level (e.g., de Bello *et al.*

2012), as well as with multi-trophic interactions (Lavorel *et al.* 2013). Dynamic models that incorporate trait-based species responses to the abiotic environment, biotic interactions and dispersal limitation (Boulangeat *et al.* 2012), are able to predict current vegetation regional distribution and could thus be applied to project climate and land-use change scenarios (Boulangeat, Georges & Thuiller 2014). Dynamic models based on the understanding of functional trade-offs in plants have also allowed new insights into ecosystem functioning (Falster *et al.* 2011). Such insights are now incorporated into new models of vegetation and biogeochemistry, which use explicit representations of plant functional traits and their trade-offs rather than a few fixed functional types (e.g., Pavlick *et al.* 2013).

Like functional diversity, phylogenetic diversity has been recently proposed as an indirect way to approach community assembly rules and ecosystem functioning (Mouquet *et al.* 2012; Srivastava *et al.* 2012). The reasoning behind this approach is that: (i) phylogenetic relationships within a community of interacting species result from the joint effects of environmental and interaction filters and thus inform about the processes of community assembly and (ii) phylogenetic diversity is correlated with functional diversity and thus is a good proxy for the potential effects of species diversity on ecosystem functioning. While this shortcut suffers from several limitations (e.g., Gravel *et al.* 2012) it offers a unique opportunity for ecologists to scale up ecological functioning to biogeographical scales at which collecting functional trait data is almost impossible while extensive phylogenies are available (e.g., Thuiller *et al.* 2011).

### **Feeding back to explanatory predictions**

An open question is whether the current theoretical corpus of ecology and evolution is mature and sophisticated enough to warrant any kind of anticipatory prediction about biodiversity and ecosystems over the next few decades to centuries. Scaling up and down models across organizational levels is currently one of the main challenges of theoretical ecology (Chave 2013). The driving factors and processes of ecological systems can interact in such a way as to produce non-predictable outcomes, even in simple models (Scheffer *et al.* 2001).

Accepted Article

There are entire areas where theory is currently lacking or highly fragmentary, and where new theory should greatly improve our understanding of the effects of global environmental changes upon ecosystems and human societies. For instance, there is still limited understanding of the structure, dynamics and functioning of ecological networks (Bascompte 2009). These have been mostly studied so far as isolated pieces, focusing on specific types of interactions among species (e.g. food, host-parasite, or mutualistic webs). In reality, all types of interactions occur simultaneously in ecosystems, generating multiple coupled ecological networks (Olff *et al.* 2009; Kéfi *et al.* 2012). Building a more integrative theory of ecological networks will be key to predict the response of ecosystems to environmental changes. Furthermore, some processes, (e.g., positive interactions among species such as facilitation, Martorell & Freckleton 2014) remain understudied. Another example is ecosystem stability. Ecosystem stability has been studied for a very long time in ecology, yielding unremitting debates (Ives & Carpenter 2007), but research in this area has addressed a wide range of different issues with little integration and predictive power (Pimm 1984; Loreau *et al.* 2002). Recent research has begun to build a mechanistic, predictive theory of ecosystem stability (Loreau & de Mazancourt 2013; Morin *et al.* 2014) but we are still far from a comprehensive theory that integrates the multiple components of stability (e.g. variability, resilience, persistence, resistance, reactivity) as well as multiple trophic levels.

Another example of an area where more theoretical development is critically needed concerns the interaction between humans and the biosphere (e.g., Liu *et al.* 2007; Collins *et al.* 2011). This interaction is arguably one of the most important ecological interactions on Earth since it drives most of the current environmental changes and will be key to determining the future of the Earth system (Chapin *et al.* 2011; Rounsevell *et al.* 2012), and yet it is still largely understudied from an ecological perspective (Harfoot *et al.* 2014a). Admittedly, humans are a difficult species to study and model because of their complex plastic behaviour, but this is no reason not to devote significant efforts to develop theory on their interactions with ecological systems (e.g., Taylor 2009; Reuveny 2012). Such a theory should allow us to make new explanatory, and perhaps even anticipatory, predictions from a different angle than that provided by economics and other social sciences.

Accepted Article

These examples call for a significant theoretical effort, either through a theoretical paradigm shift, or through integration, strengthening and extension of current theoretical insights. While rough anticipations can be obtained from mere empirical correlations, reliable anticipatory predictions, especially predictions outside the range of conditions experienced so far, require both a solid interdisciplinary theoretical and statistical background and a robust mechanistic understanding of the phenomena to be predicted. This mechanistic, theory-based approach has the advantage of not only improving predictions themselves, but also identifying their sources of uncertainty and improving both theory and predictions as knowledge accumulates (Thuiller *et al.* 2013).

### **The need for surprises in ecology**

The need for fundamental research in ecology should thus be considered a central objective in the development of predictive ecology and promoted. Research agencies are too often requesting immediate operationality while understanding should be the first target, with enough flexibility for surprises to happen along the way. Surprises, unexpected and often counterintuitive results, play an important role in the advancement of science in general, where they may contribute to initiate what has been called paradigm change/ scientific revolution (Kuhn 1962). Yet, they seem to be particularly common in ecology (Doak *et al.* 2008). It is unclear whether this is because ecology is a young science, or because of the very nature of its object, complex, ever-evolving under Darwinian dynamics and environmental change, and subject to many non-linear phenomena (e.g., Suding, Gross & Houseman 2004). Even well accepted concepts may be shaken. A spectacular example is the role reversal in the lobster–whelk predator–prey couple (Fig. 3). Such surprises are very difficult to publish, especially as long as a clear explanation has not been found (Doak *et al.* 2008); which may bias the development of the field. Indeed, the scientific community, like any community, tends to aggregate around theories and paradigms and tends to shunt currently inexplicable results. In many cases, however, surprises occur because of a missing link in our chain of understanding. Addressing the question of why they occur may help us identify missing links of knowledge and sometimes even fill in these gaps. This process seems therefore essential to the progress of science and should be encouraged. While the development of a coherent theoretical corpus should be our ultimate goal, leeway must be left for ‘out-of-the-

track' studies, which are particularly likely to produce novelties, later to become part of normal science. This means access to high-profile scientific journals but also some support from funding agencies that should integrate the need for risk and for the unexpected at the basis of their evaluation criteria.

## **ESSENTIAL TO THE CONCEPT OF PREDICTION IS THE NEED FOR DATA**

As predicting the fate of biodiversity and ecosystem functioning in a changing environment will require integrating data from very different origins (Fig. 4), the need for building a common framework and common methods for data acquisition, storage and sharing in ecology is acknowledged by a growing number of scientists and managers.

### **Entering the era of data intensive science**

Ecology is currently undergoing a major transformation to become a “big data” science (Kelling *et al.* 2009; Michener & Jones 2012; Hampton *et al.* 2013). Recent large public data bases are covering different temporal and spatial scales for thousands of organisms and from genes to ecosystems (Appendix S2).

Although these data sets are undoubtedly of great value for predictive ecology, their interoperability (Jones *et al.* 2006) and accessibility remains relatively limited. Besides the large-scale studies for which clear data management plans are implemented (leading to so-called “big data”), the vast majority of data in ecology comes from independent studies. Although highly valuable *per se*, this leads to small, uncoordinated data sets (so called “dark data”, Heidorn 2008) whose form and content can be highly specific to a particular research question or researcher (Heidorn 2008; Hampton *et al.* 2013). The issue is that, to integrate this information into big data sets, we have to account for the distribution of observation effort. To some extent, for coarse-grain species distributions, the simple accumulation of data might counter-balance the lack of information on sampling effort. However, for other purposes (e.g. quantitative assessment of biodiversity variation in space and time) it is unlikely that observation biases will ever become negligible.

Accepted Article

An important limitation of large-scale data collection is that the link between the motivation for data collection and the very process of collecting data is often not explicit. Different sampling designs can lead to radically different answers to the same question (e.g., Courbois *et al.* 2008). Explicitly integrating a priori knowledge in sampling design can reduce these discrepancies. An interesting step forward in ecology comes from large-scale theory-driven data collection, usually the kind of data used in anticipatory predictions (Albert *et al.* 2010; Dengler & Oldeland 2010). Here, target populations, sampling space and sampling units are simulated based on the theory underlying the expected analysis (e.g. population modelling, community dynamics or range dynamics) along with expert knowledge. This approach should help to: (i) construct acceptable hypotheses on the expected patterns; (ii) simulate different sampling designs including, for instance, sampling costs or the difficulty to reach sampling sites and (iii) test for the effect of sampling effort and design (distribution of samples in space) on estimates (e.g. number of populations or individuals). Obviously theory-driven data collection should not create circularity (whereby data only reflect what we already know) but rather delineate the relevant scale and sampling effort necessary to address a specific question.

### **The need for a common ontology?**

The recent development of eco-informatics (Jones *et al.* 2006) results from the acknowledgement that integration of very different data sets must become a priority in ecology (Costello *et al.* 2013). The challenge is to develop a common ontology to move ecology forwards into the information era (Madin *et al.* 2008). Ontology is defined as a formal representation or classification of concepts and their relationships within a domain of interest (e.g. a population is part of a metapopulation). By definition, it is derived from a previous understanding of the system described and thus is contingent on the relevant research fields (i.e. standards, terminologies and thesauri are domain-specific).

As integrating data sets based on different ontologies might be very challenging, two major types of ontologies are used to formalize knowledge within a domain: “generic” and “domain-specific” ontologies. The first describes very general concepts and both facilitates and guides the integration of information coming from

more specific vocabularies and from domain-specific ontologies (Madin *et al.* 2008). For example OBOE (Extensible Observation Ontology) appears to be well adapted to represent biodiversity and ecosystem data (Madin *et al.* 2007). The second type of ontologies (“domain-specific”) is based on formal definitions of concepts and how these are related within a narrower field of research. In the case of biodiversity science, domain ontologies might be developed for each organization level recognized in the Essential Biodiversity Variables proposed by Pereira *et al.* (2013): genetic composition, species and populations, species traits, community composition, ecosystem structure, and ecosystem function. At the lowest organization level, molecular ontologies are nowadays widely adopted by geneticists (Ashburner *et al.* 2000), and efforts are underway to develop ontologies at higher levels, e.g. for organismic traits (Laporte, Mougnot & Garnier 2012) and for the structure and function of ecosystems (Porter *et al.* 2011). These initiatives, however, remain isolated and the resulting tools are very far from being used by the whole scientific community of ecologists.

A side effect of the development of ontologies might be to “freeze” or “canalize” the conceptual understanding of ecological systems. A classification system and the search for interoperability between data sets inevitably leads to strong simplifications of the spatial and temporal contexts in which the data were collected, which could be done at the expense of regional specificities (Turnhout & Boonman-Berson 2011) and even modify the understanding of the phenomena considered (Lindenmayer & Likens 2013). By definition, ontology is based on an understanding of the system considered and thus the question might also be: do we have a sufficient understanding of ecological diversity and ecosystem functioning to be able to propose a common ontology in ecology?

### **The need for data availability**

Predictive ecology most often depends on data collected by other scientists, and sharing material is therefore an important issue (Costello *et al.* 2013). This will be particularly true for aggregated data that are valid at larger temporal and spatial scales than the observation point. Data sharing goes beyond releasing data in a publicly accessible data base. It also includes the constraint that data should be



reliable and, therefore, peer-reviewed, in a format that is meaningful for putative users, and finally, easy to find and access (Costello *et al.* 2013). Metadata in particular are central to this process (Michener *et al.* 1997). Since its inception, the Convention on Biological Diversity (<http://www.cbd.int>) has stressed the need to “*maintain and organize by any mechanism, data derived from identification and monitoring activities*” (article 7d). Many projects have been conducted with this aim at the local, regional, national or international level by either private or public institutions, with in some cases, incentive measures from non-governmental organizations and/or private foundations to share their data.

Despite these initiatives, sharing is the exception rather than the rule, and a number of authors have recently condemned the lack of a data-sharing culture among ecologists and advocated that all ecological data should be released in open-access data bases and eventually shared and reused (Michener & Jones 2012; Costello *et al.* 2013; Hampton *et al.* 2013). Some measures from editors have been undertaken and several journals now have data-sharing policies, publication being contingent upon data release (for example this journal). Other measures also come from funding agencies: for instance, the US NSF now requires that data management plans are embedded in research proposals and that data are made publicly available after a certain period of time. Finally, some initiatives encourage data sharing via meta-analyses (e.g. <http://www.nceas.ucsb.edu/meta/index.html>).

Once ecologists have agreed to share their data, they still need to decide where to submit them in the arcana of web-based data bases. They have the choice between international and publicly managed data bases, which are fairly constraining in terms of data format, and less searchable and less reusable web portals mainly managed by publishers (with looser format policy). Globally, open data enhances manuscript citation (Piwowar & Vision 2013) but the next step is to provide peer review to guarantee data quality and interoperability (Costello *et al.* 2013). Data acquisition could thus become an objective in itself and should be promoted by the creation of data-focused journals. This would reward the individuals who collected data and agreed to release them freely through the citation system of standard science. The next step may also be to share the knowledge linked to each data set, such as existing analyses with related computer programmes, ongoing projects, and

even aborted analyses so that one can build on previous experience (Poisot, Mounce & Gravel 2013). The key to progress in sharing data is a win–win situation in which everyone will benefit from sharing data, and where the quality of the data information will increase over time (Fig. 5).

## **IMPROVING OPERATIONALITY**

Ecology has always been connected to applied science through ambitious programmes of management, conservation biology or restoration ecology (e.g., Isaac *et al.* 2007), but current anticipatory predictions are made at very large spatial and temporal scales, with potentially strong consequences for human society (Fig. 6). This scaling up of ecological prediction has been made possible for scientific reasons (data availability, conceptual and modelling progress), but it also results from societal pressure (funding agency policies, societal paradigm of operability). Although this need for operability can catalyze scientific development, scientists should still objectively evaluate their ability to make predictions and communicate their limitations.

### **Defining operational scales**

It is now common to see continental-scale predictions on the future of biodiversity (Fig. 2) or even ecosystem services (Fig. 6). This tendency, while responding to important needs, might also overestimate the spatial and temporal scales at which reliable predictions can be made (Fig. 4). The question of scale in ecological prediction is not new (Chave 2013). But despite advances to go beyond organismal scales (e.g., Lavorel & Grigulis 2012; Mouquet *et al.* 2012; Srivastava *et al.* 2012) issues associated with spatial and temporal scales remain a frustrating problem in ecological prediction.

By definition, models based on coarsely resolved environmental data cannot fully account for fine-grained complexity. For example, species distribution models do not incorporate enough data at distribution limits (and either over- or underestimate extinction risks); they are also unable to account for local heterogeneities (thus they would tend to overestimate extinction risks, Scherrer & Korner 2010) and they do not

Accepted Article

take intraspecific variability into account. Temporal scales also pose serious challenges to ecological prediction. Indeed, long-term predictions amplify small deviations among models and among climatic or land-use scenarios. Moreover, short-term predictions with immediate socio-economic implications, such as those required by regional managers and/or decision makers, carry a great uncertainty about specific changes due to idiosyncratic effects of local circumstances and surprises, including extreme climatic or socio-economic events (Walker & Salt 2012). Such issues support the need to couple different types of models, from mechanistic fine-scale models to large-scale species distribution models (McMahon *et al.* 2011; Bellard *et al.* 2012). Practice in ecological prediction also reveals that the selection of appropriate spatial and temporal scales is a particularly sensitive issue for communication with stakeholders and transfer to decision makers. In particular, large-scale, spatially explicit predictions and their underlying limitations can be particularly difficult to understand by regional managers and local policymakers.

### **Improving the statistical toolbox of ecological prediction**

The development of predictions in ecology has long been hindered by the methodological division between phenomenological and mechanistic models (Appendix S1). Mechanistic models are used for understanding, but no general formal framework is available for parameterizing complex mechanistic models from data, quantifying the uncertainty of predictions, and comparing alternative models (Clark & Gelfand 2006). While statistics provides such a framework (e.g., Burnham & Anderson 2002), its application in ecology has been largely restricted to simple phenomenological models. The latter have been in turn criticized for not representing ecological processes, being poorly linked to theory and potentially yielding biased forecasts.

In recent years, however, the long-standing division between mechanistic and statistical models has begun to wane (Clark & Gelfand 2006). This is due to developments in computational statistics that increasingly enable ecologists to apply statistical principles of parameter estimation, model selection, and uncertainty analysis to mechanistic models. Prominent examples are methods developed in Bayesian statistics such as Markov chain Monte Carlo (e.g., Clark 2005) and

Approximate Bayesian Computation (e.g., Csillery *et al.* 2010). These methods have been widely employed in phylogenetic inference for 15 years (Yang & Rannala 1997). In ecology, they have been used to analyse time series of population abundance with theoretical models of population dynamics (Clark & Bjornstad 2004) or to estimate models of invasion dynamics from abundance variation in space and time (Hooten *et al.* 2007). Emerging applications use multiple data types to estimate complex mechanistic models of range dynamics (Pagel & Schurr 2012) or vegetation dynamics (Hartig *et al.* 2012). Bayesian methods such as Kalman filters also lend themselves to data assimilation where posterior distributions of parameters and the resulting predictions are regularly updated as new data become available. Data assimilation has a long tradition in weather forecasting and is now increasingly applied in epidemiology, fish stock assessment, and ecosystem science (Niu *et al.* 2014).

Statistical inference with mechanistic models has the potential to profoundly transform ecology (Clark & Gelfand 2006) for several reasons. First, the process-explicit and hierarchical nature of mechanistic models means that their parameters and state variables can be measured independently (although this may not always be easy). Hence, mechanistic models can be linked to a greater diversity of data types than simple correlative models. The statistical estimation of the parameters of these hierarchical models fosters understanding by synthesizing knowledge from disparate sources (Schaub *et al.* 2007; Cressie *et al.* 2009). Second, hierarchical statistical methods enable integration of mechanistic models and ecological data with complex structures (e.g., Cressie *et al.* 2009). This should increase the value of 'noisy' ecological data that are now available in big data bases for mechanistic understanding and prediction. Third, the statistical estimation of parameters in mechanistic models should increase integration of ecological theory and applications (Marie & Simioni 2014). Models that were developed in a purely theoretical context now start to be parameterized from large-scale data and might soon be used to forecast dynamics of a wide range of phenomena (Schurr *et al.* 2012; Thuiller *et al.* 2013). Similarly, parameters of small-scale mechanistic models can now be estimated from large-scale data. For example, parameters from a physiological model describing how the uptake and allocation of carbon and nitrogen determines plant growth (Thornley 1998) were estimated from data on the geographical

distribution of tree species in Europe (Higgins *et al.* 2012). Clearly, the statistical estimation of mechanistic model parameters is no panacea for ecological prediction but it will entice theoretical ecology to build general models capable of predicting a broad range of real-world phenomena (Evans *et al.* 2013a).

### **Validation at the heart of predictive ecology**

Validation is at the basis of model reliability but is also essential to the credibility of the scientific community towards stakeholders and thus should be considered as a top priority. Validation is the process of testing the behaviour of a model using a data set different from the one used for calibration (Rykiel 1996). There are inherent limitations in the validation process: (i) the perception of the predictive performance of the model depends on the techniques used (Araujo *et al.* 2005; Bahn & McGill 2013) and (ii) by definition validation also strongly depends on the availability and quality of data sets (Duputié, Zimmermann & Chuine 2014).

In this context, explanatory and anticipatory predictions face different challenges. Explanatory predictions are generally derived from simple conceptual models with few parameters that are often difficult to derive from observation. For instance, following Hubbell's neutral theory (Hubbell 2001), a particular distribution of relative species abundances (RSA) is expected in local communities and its analytical formula is governed by simple parameters (the so-called "biodiversity parameter" and dispersal, Volkov *et al.* 2003). The RSA of an empirical system can then be compared to the expected RSA derived from neutral theory (e.g., with rainforests and coral reefs data, Hubbell 2006). However, it is difficult to imagine, for instance, that the biodiversity parameter (mostly a combination between the size of the metacommunity and speciation rate) will be common to all species. Moreover, the simplicity of the associated conceptual models often leads to improper generalization in the nature of the prediction and to potential contradictions.

Anticipatory predictions are even more concerned by these limitations as they are marked by a strong difference between the scales of predictions and the scales of measurement (Fig. 4). To overcome this limitation cross-scale validation has been proposed (using data generated at a lower scale to validate models built for a larger

Accepted Article

scale), but even here the question of the interchangeability of processes between scales has not been truly addressed (Morozov & Poggiale 2012; Chave 2013). Another approach is to calibrate models on a data set and validate them on a spatially or temporally independent data set. In practice, however, two independent data sets might not be available, and the calibration and validation data sets are often defined as random subsets of the original data set (Araujo *et al.* 2005). As a result, the calibration and validation data sets are often not independent because abiotic variables show strong spatial or temporal autocorrelation. Non-independence of these data sets thus yields overly optimistic estimates of the accuracy of model projections (Heikkinen, Marmion & Luoto 2012; Bahn & McGill 2013). Moreover, even though spatially distinct samples are frequent, examples of temporally distinct samples are scarce because of the lack of long-term data sets at the temporal resolution needed for anticipatory predictions (but see Dobrowski *et al.* 2011). Although fruitful initiatives were launched in the past decades (e.g. see ILTER network: <http://www.ilternet.edu>), long-term data sets are still rare. Resolving these issues will require special efforts in long-term data collection, coordination, and sharing.

### **Acknowledging uncertainty and limitations**

With the blossoming of probabilistic phenomenological and mechanistic predictive models, apprehending uncertainty has become a central challenge. There are two main sources of uncertainty in predictive modelling: model uncertainty and parameter uncertainty, the latter being the most active field of research, the former being often ignored.

Model uncertainty refers to the inherent quality and justification of the model itself. Complex mechanistic models require a certain number of approximations, hypotheses and critical choices that are often hard to justify. Simulation experiments should be more widely used to test the importance and relevance of a given process or mechanism in a mechanistic model (Pagel & Schurr 2012). Alternatively, scenario analyses could be further developed where different mechanisms or hypotheses are used to give the range of predictions for a given model. Extensive benchmarking and comparative analyses between models will also provide insights into the influence of incorporating certain mechanisms or not (e.g., Cheaib *et al.* 2012).

Predictive models typically require inputs, the values of which being not known with certainty. Uncertainty analysis aims to quantify the overall uncertainty of a model, in order to estimate the range of possible output values, including error propagation in the case of complex models. Uncertainty and dependence modelling, model inferences, efficient sampling, screening and sensitivity analysis, and probabilistic inversion are among the most active research areas (Kurowicka & Cooke 2006). To date, despite few examples (e.g., Hartley, Harris & Lester 2006) and the awareness that different algorithms are likely to give different scenarios (Thuiller 2004; Buisson *et al.* 2010), uncertainty in parameter estimation or input data is still rarely reported (but see, Lobo 2008; Duputié, Zimmermann & Chuine 2014).

Better integration of statistical analyses into mechanistic fitting framework should foster appropriate reporting of uncertainty (e.g., Jabot & Chave 2011; Marion *et al.* 2012). So far, however, a full treatment of uncertainty has been too time-consuming and complex to be achieved. To meet this challenge, there is a need for mathematical, statistical and computational skills that extend beyond the range of standard ecological expertise towards unusual techniques likely to mix concepts of determinism and randomness that are usually considered independent (Anand *et al.* 2010). Despite these caveats, pragmatism should be encouraged, for instance by sub-sampling alternative climate projections for the same scenario to still give a basic representation of uncertainty and by considering that parameters in mechanistic models should not be fixed to one value but rather obey a probability density function based on prior knowledge. Progress will also come from integrating methods from other fields that already consider uncertainty, for instance through the use of robust simulations of probabilistic models (Gillespie 2007) or automatic verification of dynamical properties using dedicated probabilistic model-checkers (Fisher, Harel & Henzinger 2011).

Following this blossom of probabilistic extensions of both phenomenological and mechanistic models, we also advocate a better definition of the various sources of uncertainty. It remains important to disentangle epistemic uncertainty and human decision uncertainty as a consequence of model predictions (Kujala, Burgman & Moilanen 2013), and to promote a better communication of the resulting limitations to stakeholders.

## Interfacing ecological prediction with social sciences

Global changes have profound consequences for human societies through the damage, threats and risks they induce on biodiversity, ecosystem functioning and services, and economic goods (MEA 2005). This interface between biodiversity and human society requires connecting ecological predictions with human and social sciences (e.g, Chapin *et al.* 2011). Economics, sociology, anthropology, law and management should typically be mobilized. The challenge is to examine the coupled dynamics and management of the social-ecological systems at play. For instance, in marine ecosystems, accounting for fleet dynamics, fishing strategies, behaviours and goals can be decisive for improving the scenarios or management of both biodiversity and ecosystem services (e.g., Doyen *et al.* 2012). The current shortcomings of public policies and regulations for the management of biodiversity and renewable resources can be explained partly by an insufficient consideration of the complexity at play. In fisheries, for instance, there is a need for ecosystem-based predictive approaches for fishery management (FAO 2003). However, designing an operational ecosystem approach to fisheries remains controversial (Plagányi 2007; Doyen *et al.* 2012).

The development of scenario planning is a promising tool to fill the gap between ecological and social approaches. It lies at the crossroad between ecological prediction, social science and policymaking and might thus provide an appropriate ground for improving communication between the scientific community, stakeholders and decision makers (Bennett *et al.* 2003; Walz *et al.* 2007; Coreau *et al.* 2009). Scenarios can also be combined with mechanistic models. This mixed approach called story and simulation (Alcamo 2001) combines narratives about possible changes as input parameters and mechanistic models that can quantify the consequences of these changes (Biggs *et al.* 2007). Along this line, the development of Integrated Assessment Models (IAM) that combine both natural (i.e. terrestrial vegetation model or physical models of the ocean–atmosphere) and human subsystems (i.e. energy supply and demand) will be key to providing futures of human development (Harfoot *et al.* 2014a).

The need to integrate social sciences into ecological predictions also brings questions about governance. The heterogeneity of agents involved in ecological processes is high and contributes to complexity in the design of public decision and



management. Agents such as fishermen, farmers, hunters, conservation agencies, regulation agencies, politicians and tourists often differ largely in their preferences, strategies, level of information and inputs in the dynamics of socio-ecosystems. Social scientists will be key actors here to build consensus and coordination using, for instance, participatory methods or scenario models.

## **TOWARD AN ETHICS OF ECOLOGICAL PREDICTION**

Last but not least, the rise of predictive ecology stresses the need for an appropriate ethical framework. Ethics will concern several aspects of ecological predictions, such as the acquisition of data and knowledge (from local people to communities and state organizations) and their free access, the transparency of the models and scenarios (through the use and development of open-source software in statistics, modelling and geographic information systems), and the dissemination of the scientific results, outcomes and limitations.

Large-scale anticipatory predictions of ecosystem services are becoming available (e.g. Fig. 6), and they typically have important consequences for human societies, with impacts far beyond landscape management and the design of natural reserves. These predictions are often used by stakeholders and communicated to the public well before any consensus has been reached within the scientific community. Moreover, even though the “culture” of uncertainty is deeply rooted in the scientific community, it is still lacking in the general public and policy circles for whom anticipatory predictions are generally made. The general public often gets a mixed, sometimes confusing, message from ecologists and journalists on a number of highly debated topics such as biodiversity loss and climate change, the impacts of invasive species (e.g., Lodge & Shrader-Frechette 2003), and emerging diseases (Lafferty & Wood 2013), among many others. Misunderstandings often exist regarding science, expertise, values, and public policy, between ecologists and a more general audience as well as among ecologists themselves.

Two distinct aspects are involved in an ethics of ecological prediction: (i) the ethics of the scientific process generating predictions, up to its communication to the broader public, and (ii) the ethics of the use of these predictions in the public debate.

Accepted Article

Regarding the first point, it is critical to include the dissemination of scientific results in scientific ethics. Scientists can no longer ignore to whom they disseminate their predictions and why a particular result becomes publicized. Critically, they have the responsibility to communicate about the limitations and uncertainties associated with their predictions to avoid unwarranted generalizations and uses as far as possible. Although we do recognize the difficulty of this task (again, if variance and uncertainty are relatively clear statistical concepts for scientists, what about their perception by a given land planner or citizen?), ignorance of the consequences of one's findings should no longer be viewed as ethically acceptable scientific practice

Assuming that scientists integrate their public roles and the consequences of their findings in their standard scientific practice, a second, more difficult issue then arises: can they predict the social representations accompanying the scientific debate they initiate (Pielke 2007)? These representations have a causal impact on reality through the behaviour of those who adopt them, and this in turn will impinge on the state of the environment and therefore on the validity of the ecological predictions themselves. Scientists produce theories that are tested through falsification of hypotheses and confrontation with empirical results. Although the production of scenarios for the long term will hardly be refuted during the short period of a scientific grant, many hypotheses are not confronted to the social perception, construction and representation of the environment. Social scientists have an opportunity here to move at the forefront of predictive ecology in both the design of the ecological study generating predictions, by investigating the social perception and representation of the questions and hypotheses, and in the dissemination of its results. In principle, this should allow us to integrate, within ecological predictions, the very effects of these predictions on social behaviour, and therefore, to increase predictive accuracy. Ethical issues arise in this process. For instance, what ethical status should be given to the "beliefs" of local people regarding the "ecosystem services" that scientists seek to predict and in the operational policy tools that may result from these predictions, such as schemes of payment for ecosystems services (Beatley 1994; Callicott 2003)? These difficult questions should be addressed if predictive ecology is to be socially operational.

## CONCLUSION

The growing societal need for predictions of current and future anthropogenic environmental changes, the growing interaction between different disciplines and the unprecedented accumulation of ecological data are pushing ecology to become increasingly “predictive”. It is important to distinguish between at least two different kinds of predictions with different functions, which we call explanatory and anticipatory predictions. Far from reducing the need for new theory, the current trend toward a more predictive ecology makes this need all the more compelling. Reliable anticipatory predictions require a solid theoretical background based on a robust mechanistic understanding of the phenomena to be predicted and an iterative theoretical process in which explanatory predictions are generated and tested. Essential to the concept of prediction is also the need for data to test or validate predictions. Ecology is undergoing a major transformation to become a “big data” science, which implies an urgent need for building a common ontology and standards for data acquisition, storage and sharing.

The need for anticipatory predictions has triggered a blossoming of models to predict the future of biodiversity. This diversity of models can be seen as an opportunity because making anticipatory predictions at different scales and organizational levels necessarily calls for an integrated approach and the development of new theories. Anticipatory predictions are also pushing ecology to become more operational. Operationality comes with important challenges for ecology, in particular: (i) identifying the appropriate spatial and temporal scales and developing specific models, (ii) developing statistical inference methods based on mechanistic models, (iii) developing appropriate model validation procedures, and (iv) integrating inputs from social sciences and decision-making. The current trend toward anticipatory predictions is an opportunity for ecologists to become key societal actors, but with this opportunity also comes an important responsibility in the way the results and their limitations are communicated to society. An important challenge for ecology is to develop a genuine ethics of prediction to gain full credibility (Fig. 7).

## ACKNOWLEDGMENTS

This paper is dedicated to the memory of the late Robert Barbault, who strongly contributed to renew scientific ecology in France and whose heritage will root deep, beyond our scientific field, into our society. This paper is a contribution of the CNRS working group ECOPRED who met during the two-day conference "Prospective de l'Institut Ecologie et Environnement" held in Avignon on 24–25 October 2012. The authors would particularly like to thank Françoise Gaill and Stéphanie Thiébault (former and current directors of the Institute Ecology and Environment of the CNRS) for their support and encouragement. The authors would like to thank the CNRS for letting researchers in ecology work in complete independence beyond the sometimes short-term constraints imposed by research funding agencies. ML was supported by the TULIP Laboratory of Excellence (ANR-10-LABX-41).

## DATA ACCESSIBILITY

Data have not been archived because this article does not contain data.

## FIGURE LEGENDS

**Figure 1: Prediction is not new in ecology.** In January 1862, Charles Darwin receives orchids from Madagascar. A particular species, the Madagascan Comet Orchid *Angraecum sesquipedale* with a surprisingly long nectar spur (20–35 cm) catches his attention. No insect with a proboscis of this length has been described but Darwin affirms his existence, as the plant cannot reproduce without a suitable pollinator. The pollinator, a hawk moth, was indeed discovered in 1903, 41 years after Darwin's prediction. The discoverers added the name *praedicta* ("predicted") to the species name, *Xanthopan morgani praedicta*, in honour of this prediction. In the foreground we illustrate the growing importance of predictive ecology in recent years (especially through prediction of species distribution), we show the number of annual citations for articles that have "prediction" and "ecology" in their keywords (source: Web of Science, the search criteria used: Topic = prediction and ecology; Timespan = All Years). Illustration Laurence Meslin.

**Figure 2: The upscaling of ecological predictions.** Predictions on future biodiversity have become available at the continental scale. (A) Global patterns of projected mammal loss, obtained from projected future changes in suitable habitat, in

relation to global Biodiversity Hotspots (hatched). This projection is obtained for a scenario where at least 30% of suitable habitat will globally be lost by 2050 (worst-case Millennium Ecosystem Assessment scenario). Reprinted from Visconti *et al.* (2011) by permission of the Royal Society. (B) Map of the projected future of phylogenetic diversities (scenario A1FI for 2080 from the GIEC) and their relative differences with 1961–1990 for Plants, Birds and Mammals. Maps represent average phylogenetic diversity (PD; colour scale) across the sample of 100 phylogenetic trees used for each study group. Reprinted by permission from Macmillan Publishers Ltd: Nature (Thuiller *et al.* 2011), copyright (2011).

**Figure 3: The need for surprise in Ecology.** The tentative reintroduction of rock lobsters in the South African Marcus Island failed because the released lobsters were immediately attacked and consumed by the overabundant whelks, which used to be their prey (Barkai & McQuaid 1988). Illustration Laurence Meslin.

**Figure 4: The scaling of data collection in ecology.** Data in ecology are organized along two constraints of “control” and “scale” of observation. These two axes trade off and allow addressing either ecological processes or patterns. This compromise limits our ability to address ecological complexity at particular spatial and temporal scales: the zone (a) is not informative and the zone (b) is technically unreachable. Explanatory predictions are by definition concerned by the process axis while the anticipatory prediction concerns both axes. The scales of projections needed to forecast the future of biodiversity and ecosystem functioning (in green) concern scales that are not often reachable. LTER (The Long Term Ecological Research) programme was launched by the USA National Science Foundation in 1980 to conduct research on ecological issues that can last decades and span large areas (<http://www.lternet.edu>).

**Figure 5: Data life span in ecology.** Data information content as a function of time. After being published by researchers, information content in “dark data” is “naturally” declining with time (lower curve). Inversely, information content in “open data” is continuously increasing with time (upper curve). Figure modified from Michener *et al.* (1997). The various steps described here relate to the “data life cycle” formalized in

the DataONE project (Michener & Jones 2012), in which steps from data acquisition to analysis are described and subjected to specific treatments for which tools have been especially developed.

**Figure 6: Consequences of predictive ecology for human society.** As illustrated in these two examples, predictive ecology has a strong interface with human economy and society development (respectively on fish diversity and vegetation productivity). (A) Projected rate of range shifts in marine pelagic species caused by climate change from 2005 to 2050. The colour scale represents the poleward shift (in km per year). The projections are based on bioclimatic envelope models for 1066 species of fish and invertebrates, under IPCC scenario SRES A1B. Reprinted from Pereira *et al.* (2010) with permission from AAAS. (B) Simulated net vegetation primary production changes by 2071–2100 compared with a control period (1961–1990). The LPJ-GUESS ecosystem model has been parameterized under three regional climate model-generated climate scenarios (from the European Union project PRUDENCE). Contrasted scenarios have been chosen for illustration: from left to right RCAO/ECHAM-OPYC/A2, HIRHAM/ECHAM-OPYC/B2, HIRAM/Had/AM3H/A2). Reprinted from Morales *et al.* (2007) with permission from John Wiley and Sons.

**Figure 7: The need for deontology to gain in credibility.** In 1986, James Lighthill (illustration), president of the International Union of Theoretical and Applied Mechanics (<http://www.iutam.net>) made this statement: “Here I have to pause, and to speak once again on behalf of the broad global fraternity of practitioners of mechanics. We are all deeply conscious today that the enthusiasm of our forebears for the marvellous achievements of Newtonian mechanics led them to make generalizations in this area of predictability which, indeed, we may have generally tended to believe before 1960, but which we now recognize were false. We collectively wish to apologize for having misled the general educated public by spreading ideas about the determinism of systems satisfying Newton's laws of motion that, after 1960, were to be proved incorrect” (Lighthill 1986). This example of scientific integrity should motivate ecologists to build predictive ecology upon a strong deontological background to avoid having to make a similar statement in 20 years. Illustration Laurence Meslin.

## REFERENCES

- Albert, C.H., Edwards, J.T., Graham, C.H., Yoccoz, N.G., Zimmermann, N.E. & Thuiller, W. (2010) Applied sampling in ecology and evolution - integrating questions and designs. *Ecography*, **33**, 1028-1037.
- Alcamo, J. (2001) Scenarios as tools for international environmental assessment. *Experts Corner Report, Prospects and Scenarios*, pp. 31. European Environment Agency Copenhagen.
- Allen, T.F.H. & Hoekstra, T.W. (1992) *Toward a Unified Ecology*. Columbia University Press.
- Anand, M., Gonzalez, A., Guichard, F., Kolasa, J. & Parrott, L. (2010) Ecological Systems as Complex Systems: Challenges for an Emerging Science. *Diversity*, **2** 395–410.
- Araujo, M.B., Pearson, R.G., Thuiller, W. & Erhard, M. (2005) Validation of species-climate impact models under climate change. *Global Change Biology*, **11**, 1504-1513.
- Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G. & Gene Ontology, C. (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics*, **25**, 25-29.
- Augusiak, J., Van den Brink, P.J. & Grimm, V. (2014) Merging validation and evaluation of ecological models to 'evaluation': a review of terminology and a practical approach. *Ecological Modelling*, **280**, 117-128.
- Bahn, V. & McGill, B.J. (2013) Testing the predictive performance of distribution models. *Oikos*, **122**, 321–331.
- Barkai, A. & McQuaid, C. (1988) Predator–prey role reversal in a marine benthic ecosystem. *Science*, **242**, 62–64.
- Barnosky, A.D., Hadly, E.A., Bascompte, J., Berlow, E.L., Brown, J.H., Fortelius, M., Getz, W.M., Harte, J., Hastings, A., Marquet, P.A., Martinez, N.D., Mooers, A., Roopnarine, P., Vermeij, G., Williams, J.W., Gillespie, R., Kitzes, J., Marshall, C., Matzke, N., Mindell, D.P., Revilla, E. & Smith, A.B. (2012) Approaching a state shift in Earth's biosphere. *Nature*, **486**, 52-58.
- Bascompte, J. (2009) Disentangling the Web of Life. *Science*, **325**, 416-419.
- Beatley, T. (1994) *Ethical land use: principles of policy and planning*. Johns Hopkins University Press.
- Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W. & Courchamp, F. (2012) Impacts of climate change on the future of biodiversity. *Ecology Letters*, **15**, 365-377.
- Bennett, E.M., Carpenter, S.R., Peterson, G.D., Cumming, G.S., Zurek, M. & Pingali, P. (2003) Why global scenarios need ecology. *Frontiers in Ecology and the Environment*, **1**, 322-329.
- Biggs, R., Raudsepp-Hearne, C., Atkinson-Palombo, C., Bohensky, E., Boyd, E., Cundill, G., Fox, H., Ingram, S., Kok, K., Spehar, S., Tengo, M., Timmer, D. & Zurek, M. (2007) Linking futures across scales: a dialog on multiscale scenarios. *Ecology and Society*, **12**, 17.
- Borgatti, S.P., Mehra, A., Brass, D.J. & Labianca, G. (2009) Network Analysis in the Social Sciences. *Science*, **323**, 892-895.
- Boulangéat, I., Georges, D. & Thuiller, W. (2014) FATE-HD: A spatially and temporally explicit integrated model for predicting the vegetation structure and diversity at regional scale. *Global Change Biology*, **20**, 2368-2378.

- Boulangeat, I., Philippe, P., Abdulhak, S., Douzet, R., Garraud, L., Lavergne, S., Lavorel, S., van Es, J., Vittoz, P. & Thuiller, W. (2012) Improving plant functional groups for dynamic models of biodiversity: at the crossroads between functional and community ecology. *Global Change Biology*, **18**, 3464-3475.
- Buisson, L., Thuiller, W., Casajus, N., Lek, S. & Grenouillet, G. (2010) Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, **16**, 1145-1157.
- Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multimodel inference : a practical information-theoretic approach*. Springer, New York.
- Callicott, J.B. (2003) *Beyond the Land Ethic: More Essays in Environmental Philosophy*. State University of New York Press, Albany, NY.
- Carter, T.R. & La Rovere, E.L. (2001) Developing and Applying Scenarios. *Climate Change 2001: Impacts, Adaptation, and Vulnerability* (eds M. J.J., O.F. Canziani, N.A. Leary, D.J. Dokken & K.S. White). Cambridge University Press.
- Chapin, F.S., Power, M.E., Pickett, S.T.A., Freitag, A., Reynolds, J.A., Jackson, R.B., Lodge, D.M., Duke, C., Collins, S.L., Power, A.G. & Bartuska, A. (2011) Earth Stewardship: science for action to sustain the human-earth system. *Ecosphere*, **2**, 20.
- Chave, J. (2013) The problem of pattern and scale in ecology: what have we learned in 20years? *Ecology Letters*, **16**, 4-16.
- Cheaiab, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrêne, E., François, C., Gritti, E.S., Legay, M., Pagé, C., Thuiller, W., Viovy, N. & P., L. (2012) Climate change impacts on tree ranges: model intercomparison facilitates understanding and quantification of uncertainty. *Ecology Letters*, **15**, 533-544.
- Chevin, L.M., Lande, R. & Mace, G.M. (2010) Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. *PLoS Biology*, **8**, 8.
- Clark, J.S. (2005) Why environmental scientists are becoming Bayesians. *Ecology Letters*, **8**, 2-14.
- Clark, J.S. & Bjornstad, O.N. (2004) Population time series: Process variability, observation errors, missing values, lags, and hidden states. *Ecology*, **85**, 3140-3150.
- Clark, J.S. & Gelfand, A.E. (2006) A future for models and data in environmental science. *Trends in Ecology and Evolution*, **21**, 375-380.
- Collins, S.L., Carpenter, S.R., Swinton, S.M., Orenstein, D.E., Childers, D.L., Gragson, T.L., Grimm, N.B., Grove, M., Harlan, S.L., Kaye, J.P., Knapp, A.K., Kofinas, G.P., Magnuson, J.J., McDowell, W.H., Melack, J.M., Ogden, L.A., Robertson, G.P., Smith, M.D. & Whitmer, A.C. (2011) An integrated conceptual framework for long-term social-ecological research. *Frontiers in Ecology and the Environment*, **9**, 351-357.
- Condamine, F.L., Rolland, J. & Morlon, H. (2013) Macroevolutionary perspectives to environmental change. *Ecology Letters*, **16**, 72-85.
- Coreau, A., Pinay, G., Thompson, J.D., Cheptou, P.O. & Mermet, L. (2009) The rise of research on futures in ecology: rebalancing scenarios and predictions. *Ecology Letters*, **12**, 1277-1286.
- Costello, M.J., Michener, W.K., Gahegan, M., Zhang, Z.Q. & Bourne, P.E. (2013) Biodiversity data should be published, cited, and peer reviewed. *Trends in Ecology and Evolution*, **28**, 454-461.
- Coulson, T., Catchpole, E.A., Albon, S.D., Morgan, B.J.T., Pemberton, J.M., Clutton-Brock, T.H., Crawley, M.J. & Grenfell, B.T. (2001) Age, Sex, Density, Winter Weather, and Population Crashes in Soay Sheep. *Science*, **292**, 1528-1531.
- Courbois, J.Y., Katz, S.L., Isaak, D.J., Steel, E.A., Thurow, R.F., Rub, A.M.W., Olsen, T. &



- Jordan, C.E. (2008) Evaluating probability sampling strategies for estimating redd counts: an example with Chinook salmon (*Oncorhynchus tshawytscha*). *Canadian Journal of Fisheries and Aquatic Sciences*, **65**, 1814-1830.
- Cressie, N., Calder, C.A., Clark, J.S., Hoef, J.M.V. & Wikle, C.K. (2009) Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, **19**, 553-570.
- Csillery, K., Blum, M.G., Gaggiotti, O.E. & Francois, O. (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology and Evolution*, **25**, 410-418.
- de Bello, F., Price, J.N., Munkemuller, T., Liira, J., Zobel, M., Thuiller, W., Gerhold, P., Gotzenberger, L., Lavergne, S., Leps, J., Zobel, K. & Partel, M. (2012) Functional species pool framework to test for biotic effects on community assembly. *Ecology*, **93**, 2263-2273.
- Dengler, J. & Oldeland, J. (2010) Effects of sampling protocol on the shapes of species richness curves. *Journal of Biogeography*, **37**, 1698-1705.
- Doak, D.F., Estes, J.A., Halpern, B.S., Jacob, U., Lindberg, D.R., Lovvorn, J., Monson, D.H., Tinker, M.T., Williams, T.M., Wootton, J.T., Carroll, I., Emmerson, M., Micheli, F. & Novak, M. (2008) Understanding and predicting ecological dynamics: Are major surprises inevitable? *Ecology*, **89**, 952-961.
- Dobrowski, S.Z., Thorne, J.H., Greenberg, J.A., Safford, H.D., Mynsberge, A.R., Crimmins, S.M. & Swanson, A.K. (2011) Modeling plant ranges over 75 years of climate change in California, USA: temporal transferability and species traits. *Ecological Monographs*, **81**, 241-257.
- Doyen, L., Thébaud, O., Martinet, V., Gourguet, S., Béné, C., Bertignac, M., Fifas, S. & Blanchard, F. (2012) Stochastic viability to ecosystem-based management of multi-species fisheries. *Ecological Economics*, **75**, 32-42.
- Duputié, A., Zimmermann, N.E. & Chuine, I. (2014) Where are the wild things? Why we need better data on species distribution. *Global Ecology and Biogeography*, **23**, 457-467.
- Edwards, N. & Marsh, R. (2005) Uncertainties due to transport-parameter sensitivity in an efficient 3-D ocean-climate model. *Climate Dynamics*, **24**, 415-433.
- Evans, M.R., Bithell, M., Cornell, S.J., Dall, R.X., Díaz, S., Emmott, S., Ernande, B., Grimm, V., Hodgson, D.J., Lewis, S.L., Mace, G.M., Morecroft, M., Moustakas, A., Murphy, E., Newbold, T., Norris, K.J., Petchey, O., Smith, M., Travis, J.M.J. & Benton, T.G. (2013a) Predictive systems ecology. *Proceedings of the Royal Society B*, **280**, 20131452.
- Evans, M.R., Grimm, V., Johst, K., Knuuttila, T., Langhe, R.d., Lessells, C.M., Merz, M., O'Malley, M.A., Orzack, S.H., Weisberg, M., Wilkinson, D.J., Wolkenhauer, O. & Benton, T.G. (2013b) Do simple models lead to generality in ecology? *Trends in Ecology and Evolution*, **28**, 578-583.
- Falster, D.S., Brannstrom, A., Dieckmann, U. & Westoby, M. (2011) Influence of four major plant traits on average height, leaf-area cover, net primary productivity, and biomass density in single-species forests: a theoretical investigation. *Journal of Ecology*, **99**, 148-164.
- FAO (2003) The ecosystem approach to fisheries. *FAO Technical Guidelines for Responsible Fisheries*, pp. 112.
- Fisher, J., Harel, D. & Henzinger, T.A. (2011) Biology as Reactivity. *Communications of the Acm*, **54**, 72-82.
- Galán, J.M., Izquierdo, L.R., Izquierdo, S.S., Santos, J.I., Del Olmo, R., López-Paredes, A. & Edmonds, B. (2009) Errors and artefacts in agent-based modelling. *Journal of*

- Artificial Societies and Social Simulation*, **12**, 1.
- Gerrish, P.J. & Sniegowski, P.D. (2012) Real time forecasting of near-future evolution. *Journal of the Royal Society Interface*, **9**, 2268-2278.
- Gillespie, D.T. (2007) Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, pp. 35-55.
- Gravel, D., Bell, T., Barbera, C., Combe, M., Pommier, T. & Mouquet, N. (2012) Phylogenetic constraints on ecosystem functioning. *Nature Communication*, **3**.
- Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.H., Weiner, J., Wiegand, T. & DeAngelis, D.L. (2005) Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science*, **310**, 987-991.
- Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993-1009.
- Hairston, N.G., Smith, F.E. & Slobodkin, L.B. (1960) Community Structure, Population Control, and Competition. *The American naturalist*, **94**, 421-425.
- Hampton, S.E., Strasser, C.A., Tewksbury, J.J., Gram, W.K., Budden, A.E., Batcheller, A.L., Duke, C.S. & Porter, J.H. (2013) Big data and the future of ecology *Frontiers in Ecology and the Environment*, **11**, 156-162.
- Harfoot, M., Tittensor, D.P., Newbold, T., McInerny, G., Smith, M.J. & Scharlemann, J.P.W. (2014a) Integrated assessment models for ecologists: the present and the future. *Global Ecology and Biogeography*, **23**, 124-143.
- Harfoot, M.B.J., Newbold, T., Tittensor, D.P., Emmott, S., Hutton, J., Lyutsarev, V., Smith, M.J., Scharlemann, J.P.W. & Purves, D.W. (2014b) Emergent Global Patterns of Ecosystem Structure and Function from a Mechanistic General Ecosystem Model. *PLoS Biology*, **12**, 24.
- Hartig, F., Calabrese, J.M., Reineking, B., Wiegand, T. & Huth, A. (2011) Statistical inference for stochastic simulation models - theory and application. *Ecology Letters*, **14**, 816-827.
- Hartig, F., Dyke, J., Hickler, T., Higgins, S.I., O'Hara, R.B., Scheiter, S. & Huth, A. (2012) Connecting dynamic vegetation models to data - an inverse perspective. *Journal of Biogeography*, **39**, 2240-2252.
- Hartley, S., Harris, R. & Lester, P.J. (2006) Quantifying uncertainty in the potential distribution of an invasive species: climate and the Argentine ant. *Ecology Letters*, **9**, 1068-1079.
- Heidorn, P.B. (2008) Shedding Light on the Dark Data in the Long Tail of Science. *Library Trends*, **57**, 280-299.
- Heikkinen, R.K., Marmion, M. & Luoto, M. (2012) Does the interpolation accuracy of species distribution models come at the expense of transferability? . *Ecography*, **35**, 276-288.
- Higgins, S.I., O'Hara, R.B., Bykova, O., Cramer, M.D., Chuine, I., Gerstner, E.-M., Hickler, T., Morin, X., Kearney, M.R., Midgley, G.F. & Scheiter, S. (2012) A physiological analogy of the niche for projecting the potential distribution of plants. *Journal of Biogeography*, **39**, 2132-2145.
- Hilborn, R. & Mangel, M. (1997) *The Ecological Detective: Confronting Models with Data* Princeton University Press, Princeton.
- Hooten, M.B., Wikle, C.K., Dorazio, R.M. & Royle, J.A. (2007) Hierarchical spatiotemporal matrix models for characterizing invasions. *BIOMETRICS*, **63**, 558-567.
- Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton.
- Hubbell, S.P. (2006) Neutral theory and the evolution of ecological equivalence. *Ecology*,

- 87**, 1387-1398.
- Hulot, F.D. & Loreau, M. (2006) Nutrient-limited food webs with up to three trophic levels: Feasibility, stability, assembly rules, and effects of nutrient enrichment. *Theoretical Population Biology*, **69**, 48-66.
- Ings, T.C., Montoya, J.M., Bascompte, J., Bluthgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E., Lewis, H.M., Olesen, J.M., van Veen, F.J.F., Warren, P.H. & Woodward, G. (2009) Ecological networks - beyond food webs. *Journal of Animal Ecology*, **78**, 253-269.
- Isaac, N.J., Turvey, S.T., Collen, B., Waterman, C. & Baillie, J.E. (2007) Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny. *PloS one*, **2**, e296.
- Ives, A.R. & Carpenter, S.R. (2007) Stability and diversity of ecosystems. *Science*, **317**, 58-62.
- Jabot, F. & Chave, J. (2011) Analyzing Tropical Forest Tree Species Abundance Distributions Using a Nonneutral Model and through Approximate Bayesian Inference. *The American naturalist*, **178**, E37-E47.
- Jabot, F., Faure, T. & Dumoulin, N. (2013) EasyABC: performing efficient approximate Bayesian sampling schemes using R. *Methods in Ecology and Evolution*, **4**, 684-687.
- Jones, M.B., Schildhauer, M.P., Reichman, O.J. & Bowers, S. (2006) The new bioinformatics: Integrating ecological data from the gene to the biosphere. *Annual Review of Ecology, Evolution and Systematics*, **37**, 519-544.
- Kéfi, S., Berlow, E., Wieters, E., Navarrete, S., Petchey, O., Wood, S., Boit, A., Joppa, L., Lafferty, K., Williams, R., Martinez, N., Menge, B., Blanchette, C., Iles, A. & Brose, U. (2012) More than a meal... Integrating non-feeding interactions into food webs. *Ecology Letters*, **15**, 291-300.
- Kelling, S., Hochachka, W.M., Fink, D., Riedewald, M., Caruana, R., Ballard, G. & Hooker, G. (2009) Data-intensive Science: A New Paradigm for Biodiversity Studies. *Bioscience*, **59**, 613-620.
- Kennedy, M.C. & O'Hagan, A. (2001) Bayesian calibration of computer models. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, **63**, 425-450.
- Kleijn, D., Kohler, F., Baldi, A., Batary, P., Concepcion, E.D., Clough, Y., Diaz, M., Gabriel, D., Holzschuh, A., Knop, E., Kovacs, A., Marshall, E.J.P., Tscharntke, T. & Verhulst, J. (2009) On the relationship between farmland biodiversity and land-use intensity in Europe. *Proceedings of the Royal Society B*, **276**, 903-909.
- Kolar, C.S. & Lodge, D.M. (2002) Ecological predictions and risk assessment for alien fishes in North America. *Science*, **298**, 1233-1236.
- Kuhn, T. (1962) *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago.
- Kujala, H., Burgman, M.A. & Moilanen, A. (2013) Treatment of uncertainty in conservation under climate change. *Conservation Letters*, **6**, 73-85.
- Kurowicka, D. & Cooke, R. (2006) *Uncertainty Analysis with High Dimensional Dependence Modelling*. John Wiley & Sons.
- Lafferty, K.D. & Wood, C.L. (2013) It's a myth that protection against disease is a strong and general service of biodiversity conservation: Response to Ostfeld and Keesing. *Trends in Ecology and Evolution*, **28**, 503-504.
- Laporte, M.A., Mougnot, I. & Garnier, E. (2012) ThesauForm - Traits: a web based collaborative tool to develop a thesaurus for plant functional diversity research. *Ecological Informatics*, **11**, 34-44.
- Lavorel, S. & Garnier, E. (2002) Predicting changes in community composition and

- ecosystem functioning from plant traits: revisiting the Holy Grail. *Functional Ecology*, **16**, 545-556.
- Lavorel, S. & Grigulis, K. (2012) How fundamental plant functional trait relationships scale-up to trade-offs and synergies in ecosystem services. *Journal of Ecology*, **100**, 128-140.
- Lavorel, S., Storkey, J., Bardgett, R.D., de Bello, F., Berg, M.P., Le Roux, X., Moretti, M., Mulder, C., Pakeman, R.J., Diaz, S. & Harrington, R. (2013) A novel framework for linking functional diversity of plants with other trophic levels for the quantification of ecosystem services. *Journal of Vegetation Science*, **24**, 942-948.
- Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology*, **73**, 1943-1967.
- Lighthill, J. (1986) The recently recognized failure of predictability in newtonian dynamics. *Proceedings of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences*, **407**, 35-50.
- Lindenmayer, D.B. & Likens, G.E. (2013) Analysis: Don't do big-data science backwards. *Nature*, **499**, 284.
- Liu, J., T, D., Carpenter, S.R., Alberti, M., Folke, C. & Moran, E. (2007) Complexity of Coupled Human and Natural Systems. *Science*, **317**, 1513-1516.
- Lobo, J. (2008) More complex distribution models or more representative data? *Biodiversity and Informatics*, **82**, 14-19.
- Lodge, D.M. & Shrader-Frechette, K. (2003) Nonindigenous species: Ecological explanation, environmental ethics, and public policy. *Conservation Biology*, **17**, 31-37.
- Loreau, M. (2010) Linking biodiversity and ecosystems: towards a unifying ecological theory. *Philosophical Transactions of the Royal Society B*, **365**, 49-60.
- Loreau, M. & de Mazancourt, C. (2013) Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecology letters*, **16**, 106-115.
- Loreau, M., Downing, A., Emmerson, M., Gonzalez, A., Hughes, J., Inchausti, P., Joshi, J., Norberg, J. & Sala, O. (2002) A new look at the relationship between diversity and stability. *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives* (eds M. Loreau, S. Naeem & P. Inchausti), pp. 79-91. Oxford University Press, Oxford.
- Loreau, M., Mouquet, N. & Gonzalez, A. (2003) Biodiversity as spatial insurance in heterogeneous landscapes. *Proceedings of the National Academy of Sciences of the U.S.A.*, **100**, 12765-12770.
- Lorenzo, G., Zanicco, P., Gimenez, M., Marques, M., Iooss, B., Lavin, R.B., Pierro, F., Galassi, G., D'Auria, F. & Burgazzi, L. (2011) Assessment of an Isolation Condenser of an Integral Reactor in View of Uncertainties in Engineering Parameters. *Science and Technology of Nuclear Installations*.
- Madin, J., Bowers, S., Schildhauer, M., Krivov, S., Pennington, D. & Villa, F. (2007) An ontology for describing and synthesizing ecological observation data. *Ecological Informatics*, **2**, 279-296.
- Madin, J.S., Bowers, S., Schildhauer, M.P. & Jones, M.B. (2008) Advancing ecological research with ontologies. *Trends in Ecology and Evolution*, **23**, 159-168.
- Marie, G. & Simioni, G. (2014) Extending the use of ecological models without sacrificing details: a generic and parsimonious meta-modelling approach. *Methods in Ecology and Evolution*, **5**, 934-943.
- Marion, G., McInerney, G.J., Pagel, J., Catterall, S., Cook, A.R., Hartig, F. & O'Hara, R.B. (2012) Parameter and uncertainty estimation for process-oriented population and distribution models: data, statistics and the niche. *Journal of Biogeography*, **39**, 2225-2239.

- Martorell, C. & Freckleton, R.P. (2014) Testing the roles of competition, facilitation and stochasticity on community structure in a species-rich assemblage. *Journal of Ecology*, **102**, 74-85.
- Massol, F., Gravel, D., Mouquet, N., Cadotte, M., Fukami, T. & Leibold, M. (2011) Linking community and ecosystem dynamics through spatial ecology. *Ecology Letters*, **14**, 313-323.
- May, R.M. & Anderson, B. (1991) *Infectious Diseases of Humans: Dynamics and Control* Oxford University Press, Oxford.
- McGill, B.J. (2010) Towards a unification of unified theories of biodiversity. *Ecology Letters*, **13**, 627-642.
- McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology and Evolution*, **21**, 178-185.
- McMahon, S.M., Harrison, S.P., Armbruster, W.S., Bartlein, P.J., Beale, C.M., Edwards, M.E., Kattge, J., Midgley, G., Morin, X. & Prentice, I.C. (2011) Improving assessments of climate-change impacts on global biodiversity. *Trends in Ecology and Evolution*, **26**, 249-259.
- MEA (2005) The Millennium Ecosystem Assessment (<http://www.MAweb.org>).
- Michener, W.K., Brunt, J.W., Helly, J.J., Kirchner, T.B. & Stafford, S.G. (1997) Nongeospatial metadata for the ecological sciences. *Ecological Applications*, **7**, 330-342.
- Michener, W.K. & Jones, M.B. (2012) Ecoinformatics: supporting ecology as a data-intensive science. *Trends in Ecology and Evolution*, **27**, 85-93.
- Morales, P., Hickler, T., Rowell, D.P., Smith, B. & T Sykes, M. (2007) Changes in European ecosystem productivity and carbon balance driven by regional climate model output. *Global Change Biology*, **13**, 108-122.
- Morin, X., Fahse, L., De Mazancourt, C., Scherer-Lorenzen, M. & Bugmann, H. (2014) Temporal stability in forest productivity increases with tree diversity due to asynchrony in species dynamics. *Ecology Letters*, **17**, 1526-1535.
- Morin, X. & Thuiller, W. (2009) Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology*, **90**, 1301-1313.
- Morozov, A. & Poggiale, J.C. (2012) From spatially explicit ecological models to mean-field dynamics: The state of the art and perspectives. *Ecological Complexity*, **10**, 1-11.
- Mouquet, M., Devictor, V., Meynard, C.N., Munoz, F., Bersier, L.F., Chave, J., Couteron, P., Dalecky, A., Fontaine, C., Gravel, D., Hardy, O.J., Jabot, F., Lavergne, S., Leibold, M., Mouillot, D., Münkemüller, T., Pavoine, S., Prinzing, A., Rodrigues, A.S.L., Rohr, R.P., Thébaud, E. & Thuiller, W. (2012) Ecophylogenetics: advances and perspectives. *Biological Reviews*, **87**, 769-785.
- Niu, S., Yiqi, L., Michael, C.D., Trevo, F., Keenan, T.F., Zheng, S., Jianwei, L. & Chapin, F.S. (2014) The role of data assimilation in predictive ecology. *Ecosphere*, **5**, art65.
- Oksanen, L., Fretwell, S.D., Arruda, J. & Niemela, P. (1981) Exploitation ecosystems in gradients of primary productivity. *The American naturalist*, **118**, 240-261.
- Olf, H., Alonso, D., Berg, M.P., Eriksson, B.K., Loreau, M., Piersma, T. & Rooney, N. (2009) Parallel ecological networks in ecosystems. *Philosophical Transactions of the Royal Society B*, **364**, 1755-1779.
- Pagel, J. & Schurr, F.M. (2012) Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology & Biogeography*, **21**, 293-304.
- Pavlick, R., Drewry, D.T., Bohn, K., Reu, B. & Kleidon, A. (2013) The Jena Diversity-

- Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs. *Biogeosciences*, **10**, 4137-4177.
- Pearson, R.G., Thuiller, W., Araújo, M.B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T.E. & Lees, D.C. (2006) Model-based uncertainty in species' range prediction. *Journal of Biogeography*, **33**, 1704-1711.
- Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., Coops, N.C., Dulloo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Hoft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M. & Wegmann, M. (2013) Essential Biodiversity Variables. *Science*, **339**, 277-278.
- Pereira, H.M., Leadley, P.W., Proenca, V., Alkemade, R., Scharlemann, J.P.W., Fernandez-Manjarres, J.F., Araujo, M.B., Balvanera, P., Biggs, R., Cheung, W.W.L., Chini, L., Cooper, H.D., Gilman, E.L., Guenette, S., Hurtt, G.C., Huntington, H.P., Mace, G.M., Oberdorff, T., Revenga, C., Rodrigues, P., Scholes, R.J., Sumaila, U.R. & Walpole, M. (2010) Scenarios for Global Biodiversity in the 21st Century. *Science*, **330**, 1496-1501.
- Peter, R.H. (1991) *A Critique for Ecology*. Cambridge University Press, Cambridge.
- Pielke, R. (2007) *The honest broker: making sense of science in policy and politics*. Cambridge University Press, Cambridge.
- Pimm, S.L. (1984) The complexity and stability of ecosystems. *Nature*, **307**, 321-326.
- Piwowar, H.A. & Vision, T.J. (2013) Data reuse and the open data citation advantage. *PeerJ*, **1**, e175.
- Plagányi, É.E. (2007) Models for an ecosystem approach to fisheries. *FAO Fisheries Technical Paper*. FAO, Rome.
- Poisot, T., Mounce, R. & Gravel, D. (2013) Moving toward a sustainable ecological science: don't let data go to waste ! *Ideas in Ecology & Evolution*, **6**, 11-19.
- Popper, K. (1959) *Logic of scientific discovery*. Routledge, London.
- Porter, J., O'Brien, M., Costa, D., Henshaw, D., Gries, C., Melendez, E., Vanderbilt, K., Downing, J. & Laundre, J. (2011) A controlled vocabulary for LTER data keywords. . *Environmental Information Management Conference 2011* (eds M.B. Jones & C. Gries), pp. 168-169. . University of California, Santa Barbara, CA.
- Purves, D., Scharlemann, J., Harfoot, M., Newbold, T., Tittensor, D.P., Hutton, J. & Emmott, S. (2013) Time to model all life on Earth. *Nature*, **493**, 295-297.
- Reu, B., Zaehle, S., Proulx, R., Bohn, K., Kleidon, A., Pavlick, R. & Schmidlein, S. (2011) The role of plant functional trade-offs for biodiversity changes and biome shifts under scenarios of global climatic change. *Biogeosciences*, **8**, 1255-1266.
- Reuveny, R. (2012) Taking stock of Malthus : modeling the collapse of historical civilizations. *Annual Review of Resource Economics*, **4**, 303-329.
- Richter, R., Berger, U.E., Dullinger, S., Essl, F., Leitner, M., Smith, M. & Vogl, G. (2013) Spread of invasive ragweed: climate change, management and how to reduce allergy costs. *Journal of Applied Ecology*, **50**, 1422-1430.
- Rounsevell, M.D.A., Pedrolib, B., Erbc, K.H., Grambergerd, M., Gravsholt-Buscke, A., Haberlc, H., Kristensene, S., Kuemmerlef, T., Lavorel, S., Lindnerh, M., Lotze-Campenj, H., Metzgera, M.J., Murray-Rusta, D., Poppj, A., Pérez-Sobab, M., Reenberge, A., Vadineanuk, A., Verburgl, P.H. & Wolfslehneri, B. (2012) Challenges for Land System Science *Land Use and Policy*, **29**, 899-910.
- Rykiel, E.J. (1996) Testing ecological models: The meaning of validation. *Ecological*

- Modelling*, **90**, 229-244.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. & Tarantola, S. (2008) *Global sensitivity analysis: the primer*. John Wiley & Sons.
- Schaub, M., Gimenez, O., Sierro, A. & Arlettaz, R. (2007) Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. *Conservation Biology*, **21**, 945-955.
- Scheffer, M., Carpenter, S.R., Foley, J.A., Folke, C. & Walker, B. (2001) Catastrophic shifts in ecosystems. *Nature*, **413**, 591-596.
- Scherrer, D. & Korner, C. (2010) Infra-red thermometry of alpine landscapes challenges climatic warming projections. *Global Change Biology*, **16**, 2602-2613.
- Schurr, F.M., Pagel, J., Cabral, J.S., Groeneveld, J., Bykova, O., O'Hara, R.B., Hartig, F., Kissling, W.D., Linder, H.P., Midgley, G.F., Schroeder, B., Singer, A. & Zimmermann, N.E. (2012) How to understand species' niches and range dynamics: a demographic research agenda for biogeography. *Journal of Biogeography*, **39**, 2146-2162.
- Sitch, S., Huntingford, C., Gedney, N., Levy, P.E., Lomas, M., Piao, S.L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C.D., Prentice, I.C. & Woodward, F.I. (2008) Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Global Change Biology*, **14**, 2015-2039.
- Solé, R. & Goodwin, B. (2000) *Signs Of Life How Complexity Pervades Biology*. Basic Books, New York.
- Srivastava, D.S., Cadotte, M.W., MacDonald, A.A.M., Marushia, R.G. & Mirotnick, N. (2012) Phylogenetic diversity and the functioning of ecosystems. *Ecology Letters*, **15**, 637-648.
- Storch, D., Marquet, P. & Brown, J. (2007) *Scaling Biodiversity*. Cambridge University Press, Cambridge.
- Suding, K.N., Gross, K.L. & Houseman, G.R. (2004) Alternative states and positive feedbacks in restoration ecology. *Trends in Ecology and Evolution*, **19**, 46-53.
- Suding, K.N., Lavorel, S., Chapin, F.S., Cornelissen, J.H.C., Diaz, S., Garnier, E., Goldberg, D., Hooper, D.U., Jackson, S.T. & Navas, M.L. (2008) Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. *Global Change Biology*, **14**, 1125-1140.
- Taylor, M.S. (2009) Environmental crises : past, present, and future. *Canadian Journal of Economics*, **42**, 1240-1275.
- Thornley, J.H.M. (1998) Modelling shoot : root relations: the only way forward? *Annals of Botany*, **81**, 165-171.
- Thuiller, W. (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020-2027.
- Thuiller, W., Lavergne, S., Roquet, C., Boulangeat, I. & Araujo, M.B. (2011) Consequences of climate change on the Tree of Life in Europe. *Nature*, 531-534.
- Thuiller, W., Münkemüller, T., Lavergne, S., Mouillot, D., Mouquet, N., Schifffers, K. & Gravel, D. (2013) A road map for integrating eco-evolutionary processes into biodiversity models. *Ecology Letters*, **16**, 94-105.
- Turnhout, E. & Boonman-Berson, S. (2011) Databases, Scaling Practices, and the Globalization of Biodiversity. *Ecology and Society*, **16**, 35.
- Visconti, P., Pressey, R.L., Giorgini, D., Maiorano, L., Bakkenes, M., Boitani, L., Alkemade, R., Falcucci, A., Chiozza, F. & Rondinini, C. (2011) Future hotspots of terrestrial mammal loss. *Philosophical Transactions of the Royal Society B*, **366**, 2693-2702.

- Volkov, I., Banavar, J.R., Hubbell, S.P. & Maritan, A. (2003) Neutral theory and relative species abundance in ecology. *Nature*, **424**, 1035-1037.
- Walker, B. & Salt, D. (2012) *Resilience Practice: building capacity to absorb disturbance and maintain function*. Island Press.
- Walz, A., Lardellia, C., Behrendt, H., Grêt-Regamey, A., Lundströma, C., Kytziac, S. & Bebia, P. (2007) Participatory scenario analysis for integrated regional modelling. *Landscape and Urban Planning*, **81**, 114-131.
- Wilkinson, D.J. (2009) Stochastic modelling for quantitative description of heterogeneous biological systems. *Nature Reviews Genetics*, **10**, 122-133.
- Wollrab, S., Diehl, S. & De Roos, A.M. (2012) Simple rules describe bottom-up and top-down control in food webs with alternative energy pathways. *Ecology Letters*, **15**, 935-946.
- Worm, B., Barbier, E.B., Beaumont, N., Duffy, J.E., Folke, C., Halpern, B.S., Jackson, J.B.C., Lotze, H.K., Micheli, F., Palumbi, S.R., Sala, E., Selkoe, K.A., Stachowicz, J.J. & Watson, R. (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science*, **314**, 787-790.
- Yang, Z.H. & Rannala, B. (1997) Bayesian phylogenetic inference using DNA sequences: A Markov Chain Monte Carlo method. *MOLECULAR BIOLOGY AND EVOLUTION*, **14**, 717-724.

### Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Appendix S1:** Characteristics of mechanistic and phenomenological models in ecology.

**Appendix S2:** Non-exhaustive list, of international initiatives of the scientific community aiming for sharing ecological data.



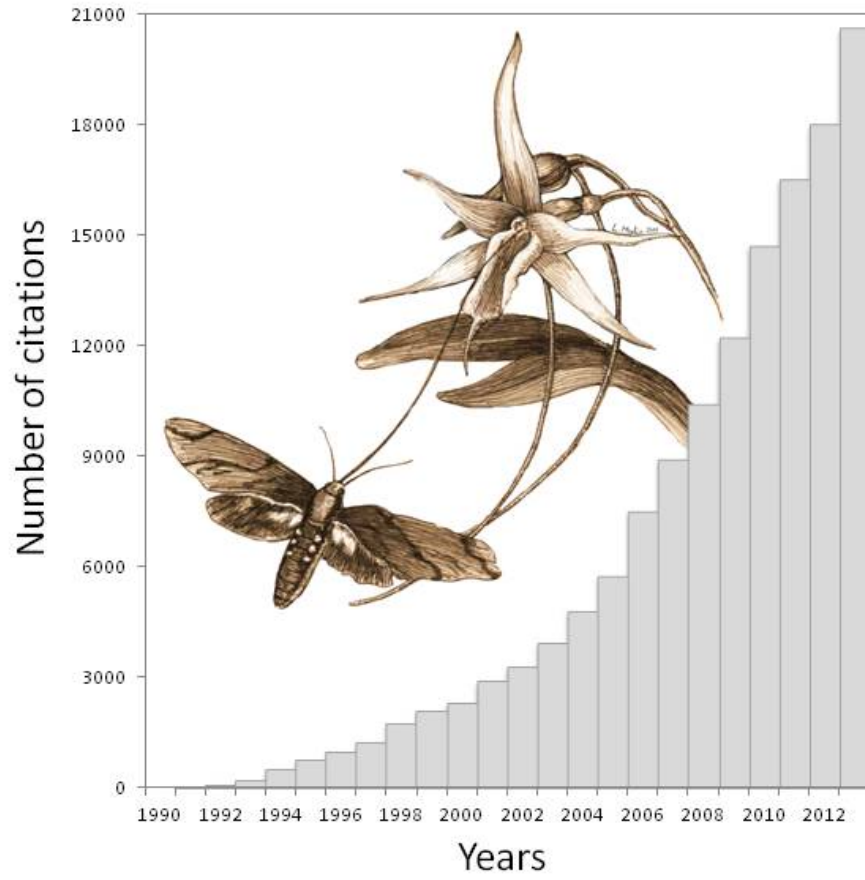


Fig.1

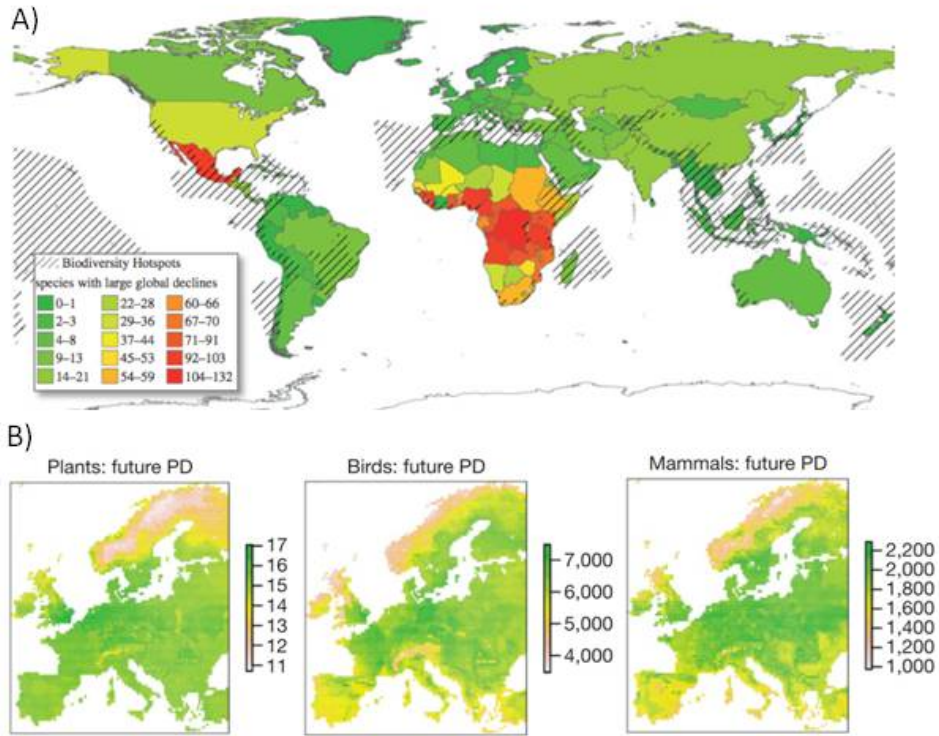


Fig.2



Fig.3

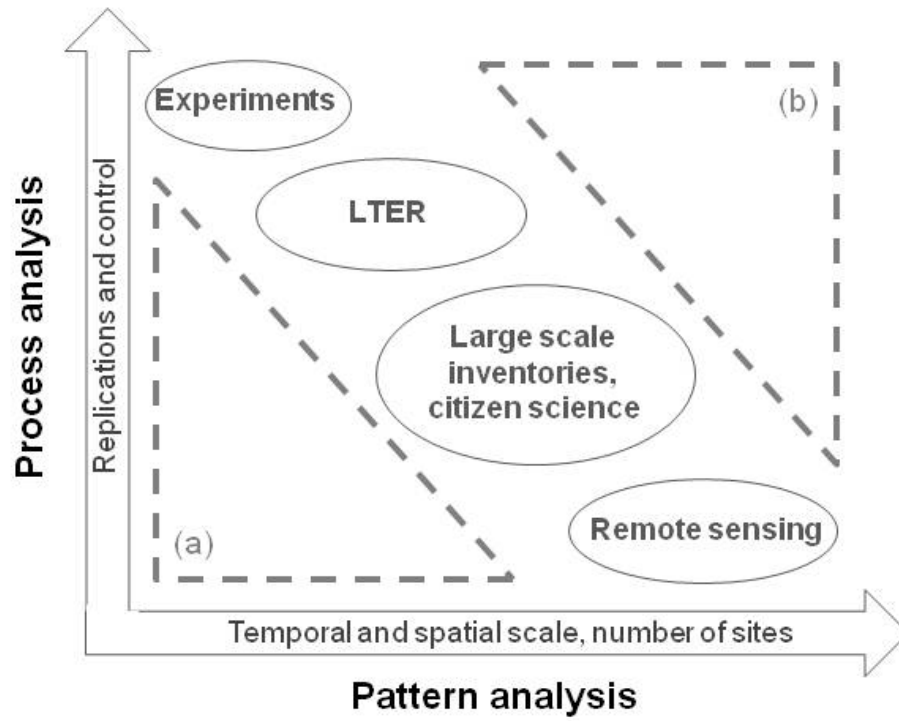


Fig.4

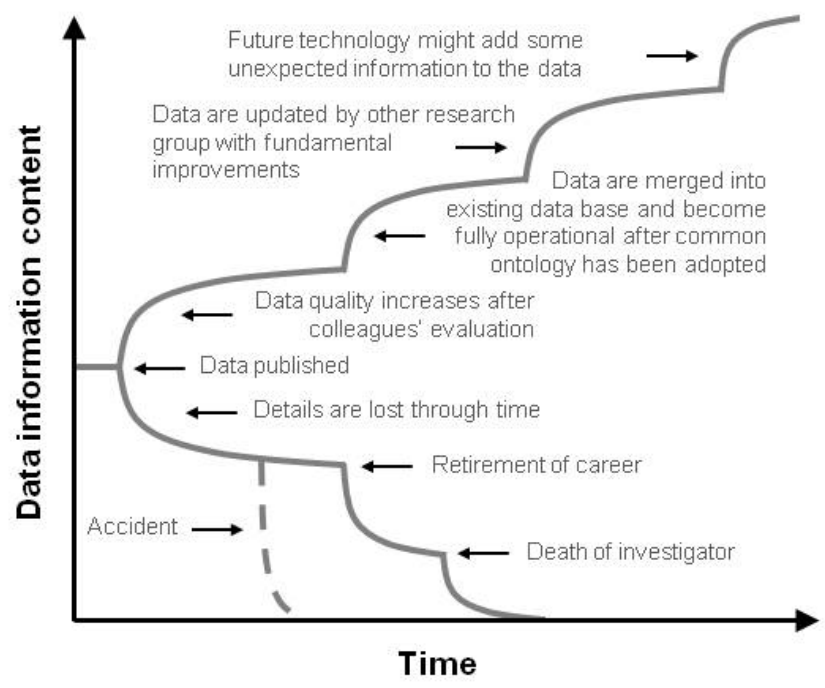


Fig.5

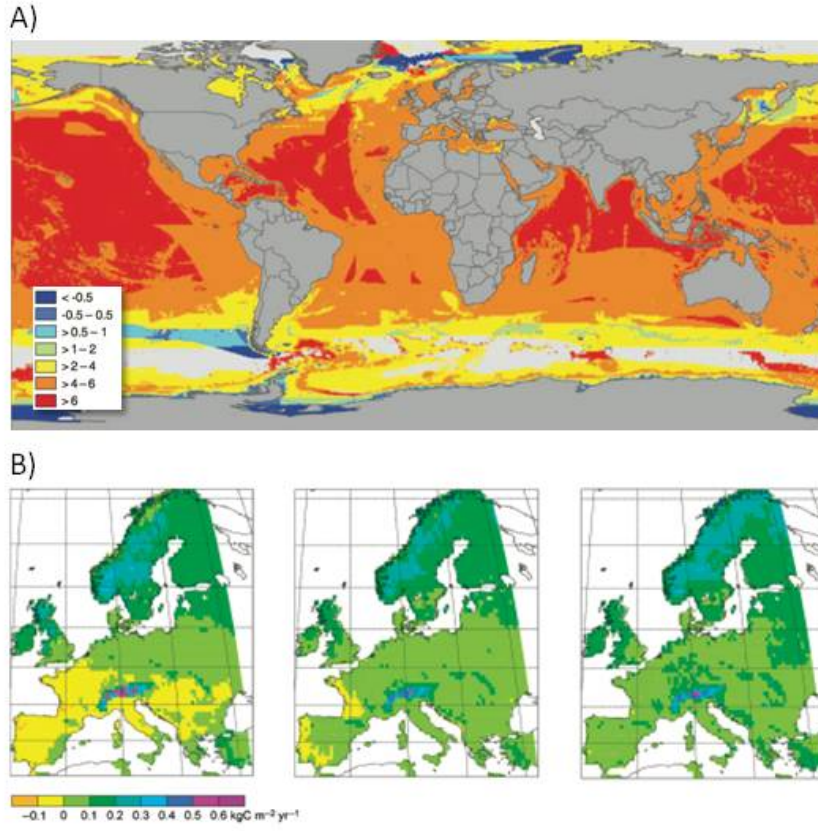


Fig.6

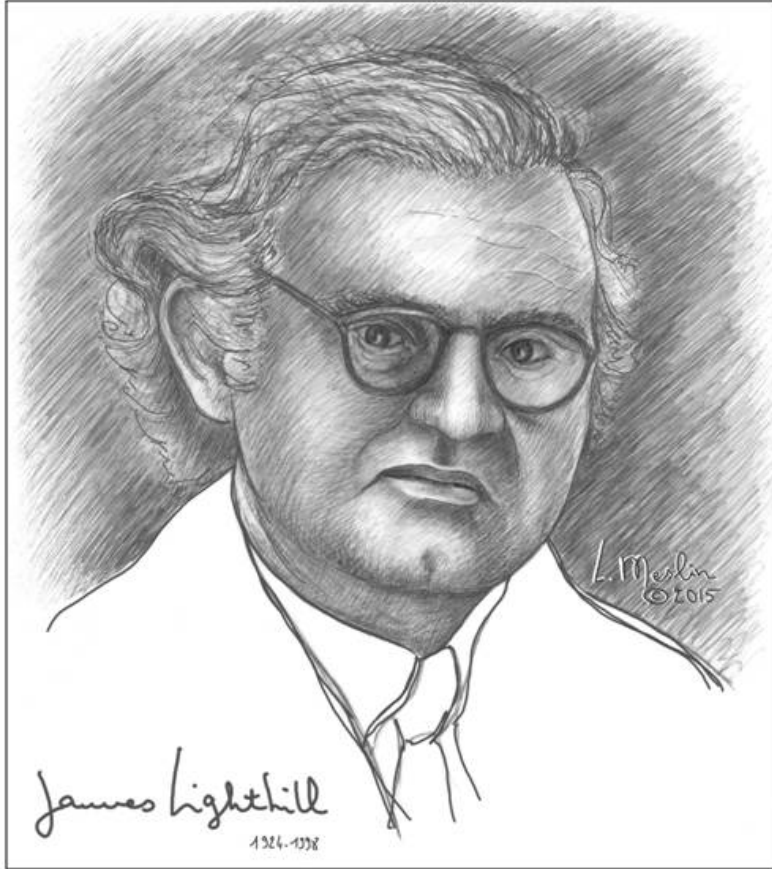


Fig.7