A road map for integrating eco-evolutionary processes into biodiversity models

Wilfried Thuiller,†* Tamara Münkemüller,† Sébastien Lavergne,1 David Mouillot,2,3 Nicolas Mouquet,† Katja Schiffers† and Dominique Gravel‡

Abstract
The demand for projections of the future distribution of biodiversity has triggered an upsurge in modelling at the crossroads between ecology and evolution. Despite the enthusiasm around these so-called biodiversity models, most approaches are still criticised for not integrating key processes known to shape species ranges and community structure. Developing an integrative modelling framework for biodiversity distribution promises to improve the reliability of predictions and to give a better understanding of the eco-evolutionary dynamics of species and communities under changing environments. In this article, we briefly review some eco-evolutionary processes and interplays among them, which are essential to provide reliable projections of species distributions and community structure. We identify gaps in theory, quantitative knowledge and data availability hampering the development of an integrated modelling framework. We argue that model development relying on a strong theoretical foundation is essential to inspire new models, manage complexity and maintain tractability. We support our argument with an example of a novel integrated model for species distribution modelling, derived from metapopulation theory, which accounts for abiotic constraints, dispersal, biotic interactions and evolution under changing environmental conditions. We hope such a perspective will motivate exciting and novel research, and challenge others to improve on our proposed approach.

Keywords
Biotic interactions, hybrid modelling, metacommunity, rapid adaptation, species distribution.

INTRODUCTION
Biodiversity models are here defined as models that provide simulated projections of the distribution and abundance of multiple species based on a set of environmental conditions (Pereira et al. 2010; Thuiller et al. 2011). The demand for such simulations in the context of global environmental change has pushed ecology, traditionally focused on describing patterns and understanding processes, towards a more predictive science. Accurate modelling tools are needed to supply managers and stakeholders with potential species distributions and community structure in response to changing environmental conditions (e.g. change in land use) and facilitate decision-making processes in conservation planning. The quality and reliability of existing tools is questionable, partly because the most widely used approaches overlook fundamental processes shaping species ranges and community structure (i.e. composition and relative abundance, Dormann 2007). We argue here that a key part of the provision of reliable projections is to upgrade modelling tools using recent theoretical developments about eco-evolutionary dynamics (e.g. Chesson 2000; Leibold et al. 2004; Solé & Bascompte 2006; Fussmann et al. 2007).

The first simulations of species’ responses to global changes were provided by ‘species distribution models’ (SDMs), developed to project large-scale species distributions from correlations between occurrence and environmental conditions (Guisan & Thuiller 2005). Even though they are often nick-named ‘niche models’, SDMs are only weakly connected to niche theory, since key processes and elements involved in the niche’s contemporary definition are not explicitly modelled. For instance, SDMs do not model net population growth rate, source-sink dynamics or competition (Chase & Leibold 2003; Holt 2009; Soberon & Nakamura 2009). Consequently, SDMs have been criticised for being over-simplistic and largely phenomenological, and therefore their reliability in making predictions about distributions under different abiotic and biotic conditions has been questioned (Davis et al. 1998). On the other side, dynamic vegetation models (DGVM, Smith et al. 2001), based on biogeochemical processes have been very successful in predicting change in vegetation with respect to global change (Smith et al. 2001). However, the focus of DGVMs on a limited number of plant functional types leads to a level of abstractness that strongly reduces their value in the context of biodiversity modelling (Thuiller et al. 2008). As an alternative, ‘hybrid’ models that include basic mechanisms, such as dispersal and demography into SDMs (Thuiller et al. 2008; Dullinger et al. 2012), have provided more realism and better predictive performance than traditional SDMs (Brotons et al. 2012; Pagel & Schurr 2012). However, like SDMs, they inherently
suffer from their weak underpinning by ecological theory. Despite being closely allied to niche theory, hybrid models do not account for biotic interactions like competition, positive interactions or trophic relationships (Gilman et al. 2010; Van der Putten et al. 2010), and only rarely do they include the capacity of species to adapt to novel conditions (but see Kearney et al. 2009).

The lack of integration of some fundamental eco-evolutionary processes into biodiversity models is surprising given that we have seen a stimulating progression of theory in the last decade. Biodiversity models were essentially developed in isolation from important theoretical contributions to species coexistence theory (Chesson 2000), metacommunity ecology (Leibold et al. 2004), metabolic theory (Brown et al. 2004), functional ecology (McGill et al. 2006), network ecology (Solé & Bascompte 2006) and niche evolution studies (Lavergne et al. 2010; Gonzalez et al. 2013). In this article, we discuss this discrepancy by first reiterating the main processes shaping species ranges and community structure, and more importantly their interplay. By contrasting this eco-evolutionary conceptual background with the current state of biodiversity modelling, we then identify the most important limitations, including available data, approaches to quantitative analyses and associated eco-evolutionary theory. These gaps are preventing the integration of multiple processes into biodiversity models. We show that the limited number of existing biodiversity models that successfully integrate several processes and the interplay among such processes are all inspired by theory. On the basis of this appraisal and synthesising recent developments in various fields, we show how a well-established theoretical background (i.e. metapopulation theory) accounting for physiological constraints, dispersal, biotic interactions and evolution can be integrated into a simple biodiversity model. Such a framework provides a potential pathway towards the next generation of biodiversity models. Recognising page space limitations, we provide illustrative examples rather than attempting a comprehensive review.

**CONCEPTUAL BASIS OF SPECIES RANGE DYNAMICS AND COMMUNITY STRUCTURE**

The past decades of ecological research have thoroughly established and described the main eco-evolutionary processes and drivers of range dynamics and community structure. Because they have been comprehensively discussed in the literature (Sexton et al. 2009; HilleRisLambers et al. 2012), we will only mention them briefly here. Instead, we highlight the importance of the interplay among major processes, which may cause non-trivial effects on species range dynamics and community patterns (e.g. Norberg et al. 2012).

These main processes are species’ physiology (e.g. Hutchinson 1959), dispersal (e.g. Snyder & Chesson 2003), biotic interactions (e.g. MacArthur 1972) and evolutionary adaptation (e.g. Polechova et al. 2009; Sexton et al. 2009). Our understanding of how the processes affect species range dynamics and community structure can be illustrated by the concept of ecological filters (Fig. 1, Weihir & Keddy 1999; Soberon & Nakamura 2009).

Following this conceptual diagram, the selection of species composing local communities first depends on their dispersal capacity (A) and their biogeographic history (B). These factors define the geographic range a species can reach in a given time and thus – under environmental change – its ability to track favourable conditions. The ‘Niche filter’ selects species that can establish and maintain positive population growth under the given environmental conditions. This selection process is driven by the match between species physiology (C) on the one hand and the environmental conditions shaped by the abiotic environment such as climate or habitat quality (E), and by inter- and intraspecific biotic interactions (D), on the other (Chase & Leibold 2003). Biotic interactions may act by reducing available resources or ameliorating harsh conditions. They may also have an effect on demographic rates due to trophic or mutualistic interactions. Finally, evolutionary adaptation affects the geographic limits of species ranges (Polechova et al. 2009; Sexton et al. 2009) by shaping species’ physiology (G), dispersal characteristics (F) and biotic interactions (H).

**Interplay among processes**

Although the different processes of range dynamics and community structure have been intensively studied, their interplay has been emphasised only recently (Sexton et al. 2009; HilleRisLambers et al. 2012). One reason for this may be that the processes act at different spatial scales and that the patterns observable at one scale may...

---

**Figure 1** (a) Conceptual representation of ecological filters selecting species from the global pool and shaping the realised local communities. Filters operate at different dimensions (geographic space and ecological space) and are not hierarchical. (b) Main processes involved in shaping species range dynamics and community structure and their direct (A–E) and indirect (F–K) effects on the filtering process. Interactions between abiotic environment, physiology and dispersal can also be important but are omitted here to avoid the figure becoming too complex. See main text for more details.
be driven by processes at other scales (Levin 1992; Chave 2013). Here, we detail how previously overlooked interplays between these key processes may affect species range dynamics.

**Interplay between dispersal and biotic interactions** (Leibold et al. 2004 and Fig. 1.J). The importance and the effect of dispersal on species distribution may vary with community composition and the resulting biotic interactions. Density-dependent dispersal is an obvious example, but there are more complex phenomena that can also impact species range dynamics and biodiversity patterns. For instance, following climate change, enemy-victim interactions may affect the migration rate of a victim by several orders of magnitude, because the enemy tends to lag behind the victim, causing an enemy release at the edge of the range and thus much faster population growth (Moorcroft et al. 2006). A striking example highlighted by Van der Putten et al. (2010) concerns the North American tree species *Prunus serotina* (black cherry). Seeds are dispersed by birds and can land far away from the parent trees and thus escape the soil pathogens accumulated by adult trees. These soil-pathogen released seeds have theoretically a better chance to establish, but they will miss symbiotic mutualists, such as arbuscular and ectomycorrhizal fungi that are accumulated around the parent plants. This interplay between two processes forcing in opposing directions will prevent accurate predictions of the distribution of the species when the feedbacks are ignored.

**Interplay between the abiotic environment, biotic interactions and physiology** (Pulliam 2000 and Fig. 1.K). The structure of interaction networks is expected to vary over space and with the environment (e.g. Poisot et al. 2012). The ‘stress-gradient hypothesis’ conceptualises the observation that the frequency of positive vs. negative interactions tends to vary inversely across gradients of physical stress (Bertness & Callaway 1994). For instance, the negative effects of soil biota on European beech survival have been shown to decrease with altitude, being mostly restricted to lower elevations (Defossez et al. 2011). Pairwise interactions are also affected by metabolic rates, which are known to scale with temperature, and thus ultimately influencing the ecosystem functioning and resource availability (Yvon-Durocher & Allen 2012). Physiological limitations (and therefore variability in environmental conditions) will influence biotic interactions since they determine the pool of species that can potentially co-occur at a given location and consequently the structure of interaction networks.

**Interplay between dispersal and evolution** (Kirkpatrick & Barton 1997 and Fig. 1-F). Several studies demonstrated how the interplay between movement and evolutionary dynamics affect range dynamics. For example, dispersal can prevent adaptation of marginal populations by gene swamping (Hendry et al. 2001), but can also lead to demographic and genetic rescue effects (e.g. Holt 2003). Modelling studies have shown that heterogeneous habitat may result in local adaptation of populations, which can either hamper (Schiffers et al. 2013) or facilitate (e.g. Polechova et al. 2009) adaptation to temporally changing conditions. Complex dynamics also arise when dispersal itself is under selection (e.g. Travis et al. 2010), sometimes leading to counter-intuitive results, such as the formation of stable range edges due to previous range shifts (e.g. Phillips 2012). An empirical example for a positive feedback between dispersal and evolution is the invasion process of the invasive cane toad in Australia. Toads with high dispersal ability tend to gather at the front of the invasion range. The interbreeding of these individuals speeds up the evolution of fast dispersal even without a directed selection (sometimes termed the ‘Olympic Village Effect’). The evolutionary dynamics in turn affect the species’ ability to expand its range, accelerating invasion dynamics and thus creating an evolutionary ‘runaway effect’ (Shine et al. 2011).

**Interplay between biotic interactions, dispersal and evolution** (Urban et al. 2008 and Fig. 1-F,H,1). Metacommunity theory has emerged as a logical framework to approach the interaction between these processes (Urban et al. 2008). An interesting example of the effects of this interplay was provided by Norberg et al. (2012) in their study on the effects of dispersal, adaptation and biotic interactions on population rescue in the face of climate change. Using a theoretical eco-evolutionary model, they showed that population responses to climate warming are far from being trivial when these processes are considered in concert. In their study, high dispersal allowed species to track their niche, but did not always reduce extinction risk and could even prevent evolutionary rescue. Moreover, extinction was only observed with competition, suggesting the prominent role of biotic interactions in the response to climate change. The relative effect of dispersal and local adaptation also depends strongly on the type of ecological interaction and the strength of the environmental gradient (Holt & Barfield 2009). Empirical studies confirm the relevance of the interplay between dispersal, evolution and biotic interactions by showing that the spatial evolutionary processes that generate trait variation among populations can in turn influence local community structure (Crutsinger et al. 2008) and even ecosystem dynamics (Palkovacs et al. 2009).

### IMPLEMENTATION OF THEORY INTO BIODIVERSITY MODELS

Given the conceptual basis detailed above, the question arises why integrated biodiversity models do not yet account for all of the fundamental processes and their interplay. Several requirements for integrating processes into predictive models may be limiting: (1) even though the appropriate background theory might be able to explain species distributions, it might not be mature enough to make quantitative predictions; (2) there is no established framework for process description and parameter estimation balancing the trade-off between complexity and tractability (Box 1); and (3) data availability does not allow for a reliable parameterisation of the simulation tool. In the following, we detail the current limitations in knowledge and data and then explore the current state and future capacity of existing modelling frameworks to facilitate further process integration.

**Current limitations in knowledge and data**

Although SDMs or other more process-based physiological models do exist (Guisan & Thuiller 2005; Higgins et al. 2012), the abiotic factors that restrict the distributions are not widely known for a large number of species due to the difficulty in measuring their physiological limits. While controlled-experiments are commonly used for short-lived species (Hooper et al. 2008; Kearney et al. 2009), this approach is much more challenging for long-lived organisms. Physiological requirements for plant survival are considered under the heading of resource availability, which involves biological properties (e.g. carbon balance) and can be differentiated from purely physical limits to viability such as heat, cold and drought tolerance (Crawford 2008). A remarkable difficulty in measuring these
Box 1 Balancing complexity and tractability

The gap between a comprehensive and integrated framework and established eco-evolutionary knowledge might be partly explained by mathematical tractability and parameterisation. Integrating all relevant processes and associated models necessarily implies a dramatic increase in parameters and assumptions. While ecology has historically looked for universal laws of species distribution (e.g. Lawton 1999), to create operational biodiversity models, we may sometimes need to get back to case-specific approaches, and use a different set of rules for different groups of organisms. We see two main directions of research that will help to reduce complexity in biodiversity models.

The first approach is a forward approach classifying organisms along some ‘response’ axes such as temperature dependency, dispersal potential, generation time or body size, and to adapt the model complexity accordingly. For example, it is likely that endotherms will be less affected by climate change than ectotherms and that organisms with short generation times are more likely to show evolutionary responses than those with long generation times. For each given process modelled (e.g. demography, dispersal, evolution or physiology), expert knowledge from empiricists will provide sets of acceptable simplifications, while theoreticians will set up the level of tractable complexity they can afford to add into their models (e.g. Boulangeat et al. 2012b). The scaling issue will be crucial in this context as the dimension of the response variables will be relative to the spatial and temporal extent of environmental changes. Positioning organisms in the multi-dimensional response variables space will allow the separation of groups of organisms for which a common modelling framework can be used. Most of the integrated models described in this review follow this idea. However, this type of simplification relies on a very good understanding of species biology, which is unfortunately not always available.

(2) In contrast, the backward approach will first build a complex mechanistic model for a species for which good knowledge of present and past distribution is available. Complexity will then be reduced by removing processes step by step and comparing predictions to observed data until an acceptable complexity-tractability balance is reached. This approach has the advantage of not requiring a prior understanding of species distributions. However, the flipside is that whenever different combinations of processes lead to the same response pattern it is impossible to identify the true underlying mechanisms (e.g. Münkemüller et al. 2012). Also, this strategy is restricted to very few groups of species for which a sufficient amount of distribution data is available.

It is likely that ecologists will use both forward and backward approaches depending on the amount of knowledge and data available for each group of organisms. Consensus might be reached, if species can be classified into simple response groups (for example, based on body size) prior to forward and backward procedures (backward approach) in which at least a few species will provide enough data to build up complex models (forward approach).

limits by means of observational studies is that individuals at range margins may be exposed to special environmental conditions as a consequence of particular relationships with their neighbours (e.g. facilitation). Indirect measures can be achieved by comparing the realised environmental space inferred from observed distributions and in situ niches from botanical garden and herbarium data (Vetaas 2002), or by measuring key physiological parameters (e.g. carbon daily balance) at different locations of a species’ range (Körner & Paulsen 2004). Global monitoring and experimental set-ups will surely provide a valuable source of data for parameterising physiological models in the coming years (e.g. The Long Term Ecological Research Network (LTER), The National Ecological Observatory Network (NEON), The European Biodiversity Observation Network (EU BON)]. More experimental studies analysing the true environmental (e.g. soil for plants), and not only climatic, boundaries of species are also needed to critically assess and model their physiological response to environmental changes.

Sound methodological approaches exist to implement dispersal into biodiversity models, but parameterisation is often limited by data availability, particularly for rare long-distance dispersal events (Nathan et al. 2008). Little remains known about the dispersal capability of a majority of species, except for a few generalisations concerning very mobile vs. sessile organisms and the scaling of space use with body size (Jetz et al. 2004). A fundamental measure in many models is the relative importance of short vs. rare long-distance dispersal events (Kot et al. 1996), which may have fundamentally different effects on rates of range shifts under changing environments (Le Galliard et al. 2012). For instance, contemporary colonisation of the plant Lactuca serriola in the Netherlands over the past 50 years shows that long-distance dispersal occurs from populations far away from the expansion zone, which then colonise the surrounding area by short-distance dispersal (Hooffman et al. 2006).

Biotic interactions seem to be quite well understood and supported by considerable empirical data at least for pairwise interactions (Gilman et al. 2010), but theory on the biogeography of network structure remains to be developed (Gravel et al. 2011b; Poissonet al. 2012). Even in well-established domains such as food web theory, there is debate, for instance, between proponents of ratio-dependent functional responses, and prey-dependent functional responses (Andrzej & Ginzburg 2012). The co-variation between the structure of ecological interactions and the abiotic requirements is still largely unknown (Hellmann et al. 2012). There is also a rather poor quantitative understanding of community interactions outside of resource competition and predation (and perhaps mutualism), for instance with pathogens and parasites, and for various kinds of indirect and ‘engineering’ interactions. These issues need to be addressed as a priority, first to increase our understanding of basic processes, and second to incorporate them into quantitative modelling tools.

Finally, in the field of evolutionary ecology several important theoretical predictions await empirical validation before proper integration into a predictive modelling framework is possible (Box 2). How labile evolutionary adaptations in physiological traits are on short time-scales is, for example, unknown (e.g. Sgrò et al. 2010), but determines whether or not ignoring adaptation in modelling the response of a given species would impact our ability to forecast its persistence in a changing environment (Atkins & Travis 2010). If evolutionary adaptation is relevant for persistence, we still need to
determine under which conditions it can rescue species that are pushed out of their niche by environmental change (Schiffers et al. 2013). More work is also needed on the role of phenotypic plasticity and its interplay with local adaptation (e.g. Chevin et al. 2013). More work is also needed on the role of phenotypic plasticity and its interplay with local adaptation (e.g. Chevin et al. 2013). More work is also needed on the role of phenotypic plasticity and its interplay with local adaptation (e.g. Chevin et al. 2013). More work is also needed on the role of phenotypic plasticity and its interplay with local adaptation (e.g. Chevin et al. 2013). More work is also needed on the role of phenotypic plasticity and its interplay with local adaptation (e.g. Chevin et al. 2013).

Current state of integrated modelling frameworks

When looking at a range of recent publications, it becomes clear that only a small part of the existing eco-evolutionary theory has been implemented in biodiversity models (Table S1). Nonetheless, despite the above-mentioned limitations, the ecological modelling arena has made tremendous progress in the last few years. Most of the approaches now at least account implicitly for dispersal and abiotic constraints, while a few account for three or more processes simultaneously. However, their interplay is still often modelled less explicitly or simply ignored. Most of the developments so far have dealt with integrating abiotic constraints, dispersal and population demography (e.g. Dullinger et al. 2012). Interestingly, as we will demonstrate in the following paragraphs, the approaches with the highest level of integration (i.e. explicit consideration of several interacting processes) are those that have been the most inspired by formalised theory (e.g. metabolic theory, mass-energy theory).

Cheung et al. (2012), for instance, developed an integrated model based on eco-physiology, dispersal, distribution and population dynamics to predict the climate change impact on more than 600 species of marine fishes due to changes in distribution, abundance and body size. The authors assumed from theory that the maximum body weight of marine fishes and invertebrates was fundamentally limited by the balance between catabolism and anabolism, which both depend on temperature through the Arrhenius equation. Using their integrated model, they show that the averaged maximum body weight of marine assemblages is expected to shrink by 14–24% from 2000 to 2050 under a high-emission scenario, with half of this shrinkage due to physiology and the other half to range shifts. This result predicts a major economic impact, since it may act in synergy with resource over-exploitation and change in primary productivity.

Box 2 Critical questions regarding species’ evolutionary responses to climate change

Recent reviews and syntheses have pointed at a number of basic, critical questions that still remain to be addressed empirically before integrating community eco-evolutionary dynamics into biodiversity models (Lavergne et al. 2010; Hoffmann & Sgrò 2011).

Can evolution proceed fast enough?

Often, adaptive evolution will be the only possible mechanism for population rescue under environmental change. There is accumulating evidence that contemporary changes have triggered genetic responses in many organisms, sometimes over very few generations (e.g. reviewed in Lavergne et al. 2010). It is not clear, however, whether this adaptation is fast enough to prevent species from extinction. Even if some models predict populations could persist in a maladaptive state (Urban et al. 2012), this has not been empirically tested yet.

Will adaptive evolution impact ecological dynamics?

Theory predicts that adaptive evolution can alter ecological dynamics, such as population demography or biotic interactions (Gomulkiewicz & Houle 2009), but empirical evidence for this effect has remained scarce. Recently, Ellner et al. (2011) showed from empirical data on fish, birds and zooplankton that the proportional contribution of rapid evolution to ecological dynamics can vary enormously, with sometimes large effects. However, no framework is currently able to predict the effect of evolutionary dynamics given a measured heritability of phenotypic traits, the strength of natural selection and the amplitude of ecological changes. Most of our current understanding of eco-evolutionary dynamics of ecological systems comes from theoretical models that await empirical validation (Urban et al. 2012).

Which niche-related traits show significant genetic variation?

It is essential that phenotypic traits be considered in future models of biodiversity since they constitute the link between an individual’s fitness in a given environment, biotic interactions and evolutionary change. To advance towards more realism, future model parameterisation should build on the theory and applications of quantitative genetics (Shaw & Etterson 2012), and on studies quantifying heritability of functional traits or other fitness components (e.g. Geber & Griffen 2003).

Can trait correlations impede populations’ response to selection?

Genetic correlations between traits can impede adaptive evolution to changing environments. Recent models help to better understand the expected evolutionary trajectories of populations when several intercorrelated traits are under selection: for instance, under certain conditions, some traits may evolve in the opposite direction than expected from a single-trait study (Duputié et al. 2012). Even when traits are not genetically correlated adaptive evolution to climatic change may be impossible when local stabilising selection on a second trait also impacts on individual fitness (Schiffers et al. 2013).

Can phenotypic plasticity enhance or counteract responses to selective pressures?

Phenotypic plasticity is certainly a widespread phenomenon in nature, and it seems straightforward that it could trigger population persistence in the absence of adaptive evolution (Hoffmann & Sgrò 2011). Phenotypic plasticity may, however, pull trait values in a direction incompatible with the adaptation to environmental (e.g. climate driven) selective pressures (e.g. Eckhart et al. 2004). Eventually, models should separate genetic from plastic trait changes, and recent progress has been made towards this aim (Hoffmann & Sgrò 2011).
As another example, Kearney et al. (2009) integrated evolution, dispersal and abiotic constraints and their interplay with biophysical models of energy and mass transfer. The authors suggested that solving the energy balance equation for an ectotherm provides an estimate of the core body temperature under a given set of environmental conditions, further defining physiological function and survival. This approach requires information on essential physiological parameters such as thermal dependence of egg, larval and pupal development for ectotherms, or basal metabolic rate or physiological response curves for endotherms. The approach, which was developed from first principles and experimental data, gave congruent results with a traditional SDM fitted with observed distributional data (Kearney et al. 2010). This result could seemingly justify the use of the simpler SDM approach. However, the strong advantage of mechanistic niche modelling, as proposed by Kearney et al. (2009), is the integration of dispersal and the evolution of some of the modelled traits linked to the distribution (Kearney et al. 2009). Using a standard quantitative genetic model, the authors simulated the evolutionary change in egg desiccation resistance and consequently the occurrence and spreading rate of *Aedes aegypti* in northern Australia (Fig. 2). The model was run with and without climate change. Such an integrated model accounting for range dynamics and evolution is not unique (see Kramer et al. 2008, 2010), but has rarely been applied to biodiversity modelling and is limited to well-studied taxa allowing model parameterisation.

In general, the advantage of integrating several processes simultaneously is not only to provide more informative models of biodiversity but also to raise new ecological questions or hypotheses and to give invaluable insights into the drivers of species distributions. For instance, by integrating abiotic constraints, dispersal and biotic interactions in a single framework, Boulangeat et al. (2012a) managed to quantify the effects of dispersal and plant interactions on the abiotic niche. The unbiased estimation of the niche allowed them to identify potential source-sink areas and the environmental conditions where positive and negative interactions were most important (Fig. 3).

The downside of integrating multiple processes simultaneously is the intricate balance between complexity and tractability (Levins 1966). Reducing complexity by identifying unimportant processes and interactions to minimise the number of free parameters will thus remain a key challenge in biodiversity modelling (Box 1). However, the above examples and the approaches listed in Table S1 show that the integration of multiple processes into a modelling framework becomes possible when building the approach on a strong theoretical background. We believe that a theory driven development of simulation tools is necessary for building next-generation biodiversity models. Such an approach should help, among other things, managing complexity and providing more tractable statistical models. In parallel, theoretical simulations also provide intuition of the most important mechanisms by means of sensitivity analyses and provide some mechanistic understanding of parameters and predictions.

A THEORY-BASED BIODIVERSITY MODEL – AN EXAMPLE

Although the approaches described above (Kearney et al. 2009; Cheung et al. 2012) have been successful in integrating several processes, they are not flexible enough to allow for future elaborations and additional integration of processes. They are also quite specific to one type of organism. As a more flexible alternative to a theory-based biodiversity model, metapopulation theory may be used in a broader context, as the cornerstone for a new generation of biodiversity models. This approach has the advantage of being explicit about the processes involved, provides interpretable parameters and has solid foundations in population dynamics. It also builds on the recent addition of environmental heterogeneity, dispersal limitation and biotic interactions into the incidence function. The addition of an evolutionary perspective is the next challenge.

The backbone of metapopulation ecology: the incidence function

Metapopulation theory, built around patch occupancy (Levins 1969), describes patch colonisation and extinction dynamics:

$$\frac{dp_i}{dt} = C_i(H_i – p_i) – e(p_i)\; , \quad \text{(1)}$$

where $p_i$ is the probability of a species $i$ to be present in a patch (and incidentally the regional abundance of the species), $H_i$ is the proportion of suitable habitat patches in the landscape for that specie...
Figure 3 Effects of dispersal limitation and plant interactions on the distribution of an alpine plant species, Bromus erectus, depicted in abiotic space (Boulangeat et al. 2012a). (a) Observed distribution in abiotic space. Left: density of predicted presences normalised by the number of sample plots within each grid cell. Right: third quartile of predicted abundance class within each grid cell. (b) Left / right: Proportion of sources / sinks among predicted presences. Middle: abundances in sources and sinks. (c) Effect of biotic interactions. Left: predicted presence density without biotic interactions. Right: negative and positive effects of biotic interactions on the abiotic niche of the species. Redrawn from Boulangeat et al. (2012a).
species, \( C_i \) is the colonisation rate of empty patches (simply \( C_i = f \phi_i \) in Leivin's model, where \( f \) is the fecundity per patch, but other formulations exist, see below) and \( e_i \) is the extinction rate. Solving this model at equilibrium yields the incidence function:

\[
p_i^* = \frac{C_i}{C_i + e_i}.
\]

This equation provides some basic information about a species' distribution: even if the species could establish a viable population in a given environment (i.e. it has a positive local rate of increase, namely a source Pulliam 1988), it might not persist, if the local extinction rate is too high relative to the colonisation rate (e.g. when the local carrying capacity is small and there is high demographic stochasticity). There are therefore three alternative explanations for range limits: habitat availability, colonisation dynamics and local extinctions (Holt & Keitt 2000). That way, the approach also proposes an ecological explanation for the absence of a species in a favourable location.

**Environmental heterogeneity**

The incidence function could be easily transformed into a species distribution model. For this, the incidence and all rates are made specific to a location \( x \). The environment could have an effect on both the colonisation rate, through an effect on propagule production and the establishment success of offspring, and on extinction rate, through any disturbance (natural or human driven) that causes temporary and localised shifts in demographic rates (Molilan & Hanksi 1998). A fraction \( 1 - Hi \) of the landscape may also be unsuitable for the species (e.g. human land use). For illustration, we will consider here that \( Hi = 1 \), that is, that the carrying capacity of a species is a function of the local environmental conditions, \( K(E_x) \), and that the extinction rate is inversely related to the population size \( (e_x = K(E_x)^{-1}) \). After some manipulations and linearisation of the incidence function (Eqn 2, Hanski 1999), we can express it as a species distribution model of the form:

\[
\ln\left(\frac{p_i^x}{1 - p_i^x}\right) = \ln(C_i) + \ln(K(E_x)),
\]

which could be easily fitted to data by maximum likelihood approaches and compared to other species distribution models (Guisan & Thullier 2005). Although its particular form might reduce the fit to empirical data compared to traditional SDMs, it is worth consideration because of its explicit theoretical foundations and parameter interpretability.

**Dispersal limitation**

A major contribution to metapopulation ecology is the spatially explicit version of the incidence function (Hanski 1999). Using the same approach as above, one could consider that the propagule pressure at a location \( x \) is proportional to seed rain from neighbouring sites \( y \) with the following definition: \( C_{xy} = f \sum p_j \exp(-a d_{xy}) \), where \( a \) is a parameter for the dispersal kernel and \( d_{xy} \) the distance between locations \( x \) and \( y \). The model could then be evaluated with a predefined \( a \), based on known traits (e.g. Boulangeat et al. 2012a), or estimated from data. The latter would correspond to a hypothesis-driven approach with a spatial autocorrelation term added to a species distribution model. In both cases, accounting for dispersal limitations would contribute reducing bias in parameter estimation and improve model fit (Dormann et al. 2007).

**Biotic interactions**

One of the major issues for biodiversity models is the complexity of biotic interactions. While several attempts have been published to integrate biotic interactions into biodiversity models, most of them still use a phenomenological description of interactions, which, in essence, prevents the proper application to transient dynamics (Boulangeat et al. 2012a). Biotic interactions have nonetheless been included into the metapopulation framework using several approaches (reviewed in Hanski 1999). Gravel et al. (2011b) extended the island biogeography theory to account for trophic interactions. The model was derived from two simple assumptions: a species can colonise a patch only if it has at least one prey present, and it goes extinct if its last prey goes extinct. Consequently, Gravel et al. modified the colonisation rate so that \( Ci = Ci q_j \), where \( Ci \) is the colonisation probability when there is at least one prey already present in the patch and \( q_j \) is the probability that at least one prey species from the diet in the regional species pool is already present (defined with the incidence of all species \( i \) in the diet of species \( j \), i.e. \( q_j = 1 - \prod (1 - p_i^j) \)). The extinction rate was similarly modified, with \( e_j = e_0 + e_a \), where \( e_0 \) is the community-independent extinction rate and \( e_a \) is the additional extinction related to the probability of losing the last prey (defined as \( e_a = \sum e_b p_j \prod (1 - p_i^j) \)). The model was solved at equilibrium and was found to fit the distribution of consumers of pelagic food webs in freshwater lacks from the Adirondacks in Upstate New York in the United States. Although it did not account for physiological constraints, this study was the first to propose and parameterise a species distribution model accounting for complex interaction network structures in species rich systems (see Fig. 4, for a simulation example of an interaction network, and see Gravel et al. 2011a for more details).

Although the above definitions of the effect of biotic interactions on colonisation and extinction were derived for predator–prey interactions, other formulations could also be implemented when other interactions prevail, such as the competition-colonisation trade-off (Tilman 1994), source-sink dynamics (Mouquet & Loreau 2003) and mutualism (Klaussmeier 2001). The model could also be extended to meta-ecosystems (Gravel et al. 2010). The determinants of range limits in this framework are likely to be more complicated than in a framework considering only a single population. For instance, the range of a predator might be limited indirectly by the response of its prey (either through colonisation or extinction rates). The formalism highlights that the empirical covariance of predator and prey responses to environmental gradients needs to be evaluated.

**Local adaptation**

The different determinants of range limits could obviously have different evolutionary implications (Holt & Keitt 2000). Even though evolution in metapopulation models has received some attention (Orsini et al. 2009), local adaptation along environmental gradients has been overlooked. A way to integrate the effects of local adaptation could be through the adaptation of eqn 3. In this case, the fecundity or the extinction rate of a species at location \( x \) could not only be a function of the environmental conditions but also of the measured...
Local maladaptation could be measured as the deviation from an optimal phenotype (e.g. Duputi et al. 2012), and is expected to vary depending on local environmental conditions. In addition, using the same theoretical background, \( f_c \), the fecundity per patch, could also become location-specific and be a function of the local adaptation of the population.

**Link to landscape simulation model to include habitat dynamics**

The standard metapopulation approach assumes a static landscape structure. However, patches themselves could be dynamic, for instance the spatial arrangement of mature forest stands necessary for several bird species, might change over time owing to different land uses (e.g. forestry, pasture, intensive agriculture). Substantial work has been carried out in landscape ecology to build mechanistic plant succession models based on light competition and life-history traits (Noble & Slatyer 1980). Such models could be coupled with a metapopulation approach for dependent organisms. For example, the original model of Noble & Slatyer (1980) has been extended to include disturbances (e.g. grazing), and dispersal (e.g. spatially explicit seed production and seed dispersal on a lattice; Albert et al. 2008). The vegetation dynamics could be used to determine the environmental conditions influencing species richness per community, ranging from red (low) to blue (high); black lines denote dispersal between patches. (b) Relationship between species richness and the environment. (c) Interaction matrix for the whole network (top left) and three selected communities. The interactions between predators (columns) and preys (rows) are denoted by black squares. Some interactions do not occur locally owing to the absence of the predator or the prey (light grey). (d) The distribution of a randomly selected species along the environmental gradient (dots) and the fundamental niche (line), as determined by the relationship between the colonisation probability and the environment. The discrepancy between occurrence and the fundamental niche arises from the impact of biotic interactions.

**Figure 4** Species distribution in trophic metacommunities. The environment varies linearly along the \( X \)-axis. The colonisation probability is maximal at the niche optimum (\( f_c = 0.4 \)), the baseline extinction probability is 0.3 and increases to 0.4 with the presence of one predator. (a) Local species richness per community (from 0 – red – to 15 – blue –; black lines denote dispersal between patches). (b) Relationship between species richness and the environment. (c) Interaction matrix for the whole network (top left) and three selected communities. The interactions between predators (columns) and preys (rows) are denoted by black squares. Some interactions do not occur locally owing to the absence of the predator or the prey (light grey). (d) The distribution of a randomly selected species along the environmental gradient (dots) and the fundamental niche (line), as determined by the relationship between the colonisation probability and the environment. The discrepancy between occurrence and the fundamental niche arises from the impact of biotic interactions.
fraction $H_i$ of suitable habitats for higher levels (e.g., herbivores, pollinators or seed dispersers). In addition, species distribution modules for interacting species could be integrated to inform the recruitment probability and potential growth functions and thereby influence their spatial distribution. The next challenge is to expand this framework with other common types of biotic interactions such as coupled herbivore-predator systems and traits-mediated indirect interactions. The critical step in coupling different modelling frameworks is to identify the connecting drivers and processes. For instance, linking the physiological and the succession processes. For instance, linking the physiological and the succession models is far from trivial and raises several questions (Gallien et al. 2010). If habitat quality is known to influence recruitment, does it also impact survival and growth? What is the shape of the relationship between habitat quality and recruitment (is the best approximation function logistic, asymptotic or Gaussian)? These and other necessary decisions require both a better understanding of the effects of habitat quality on demography and novel types of data such as extensive geo-localised demographic data for a large set of species.

**CONCLUSION**

Although the development of reliable biodiversity models is an enormous task for the scientific community, it is urgently needed to provide managers and stakeholders with projections of biodiversity dynamics. Such approaches would also help to promote the emergence of questions and approaches at the crossroads between ecology and evolution. This development requires a quantitative understanding of the main processes shaping species ranges and community structure, including an assessment of the availability and suitability of required data, and the advancement of integrated simulation frameworks to explicitly model the interplay between these processes. Although former advances in biodiversity modelling at the biogeographical scale focused on the task of simplifying known relationships and interplays of processes, we advocate here that this should not come at the cost of oversimplification. We believe that such developments can be integrated into biodiversity models without losing tractability or predictive power. The right balance between realism and simplicity in integrated biodiversity models will only be achieved when underpinned by sound theoretical foundations. In addition to the challenges associated with large-scale biodiversity modelling will be an opportunity for new research directions that go well beyond adapting classical ecological paradigms to the spatial and organisational scales. To illustrate these challenges and opportunities we have proposed a framework based on multi-species metapopulation theory that incorporates some of the latest developments in landscape ecology. We hope such a perspective will motivate exciting and novel research, and challenge others to improve on our proposed approach.

**ACKNOWLEDGEMENTS**

We are grateful to Michael Hochberg, Marcel Holyoak and the CNRS-INSEE for organising the Symposium Ecological Effects of Environmental Change in June 2012 in Paris, M.R. Kearney and I. Boulangeat for allowing us to reprint illustrations, David Bourke for English proof-checking of the manuscript and Björn Reineking and five anonymous referees for their constructive criticisms. The research leading to this article had received funding from the European Research Council under the European Community’s Seven Framework Programme FP7/2007–2013 Grant Agreement no. 281422 (TEEMBIO). TM was funded by the ERA-Net BiodivERsA (ANR-11-EBID-002 CONNECT). We also acknowledge support from the ANR SCION (ANR-08-PEXT-03) and EVORANGE (ANR-08-PEXT-11) projects. DM was supported by a Marie Curie International Outgoing Fellowship (FISHECO) with agreement number IOF-GA-2009-236316 A+. NM was supported by the CNRS. DG was supported by the NSERC and the Canada Research Chair programme. We also thank the participants of the workshop entitled ‘Advancing concepts and models of species range dynamics’ funded by Danish Council for Independent Research – Natural Sciences (No. 10-085056), for insightful discussions around the topic.

**AUTHORSHIP**

WT and DG formulated the ideas, WT, TM, KS and DG designed the structure of the article and all the authors contributed substantially to the writing.

**REFERENCES**


Chave, J. (2013). The problem of pattern and scale in ecology: what have we learned in 20 years? Ecol. Lett. 16 (S1), 4–16.


Chave, J. (2013). The problem of pattern and scale in ecology: what have we learned in 20 years? Ecol. Lett. 16 (S1), 4–16.


Chave, J. (2013). The problem of pattern and scale in ecology: what have we learned in 20 years? Ecol. Lett. 16 (S1), 4–16.


