Combining correlative and mechanistic habitat suitability models to improve ecological compensation

Eric Meineri\textsuperscript{1,2,*,*}, Anne-Sophie Deville\textsuperscript{1,3,*}, David Grémillet\textsuperscript{3,4}, Michel Gauthier-Clerc\textsuperscript{1,5} and Arnaud Béchet\textsuperscript{1}

\textsuperscript{1}Centre de recherche de la Tour du Valat, Le Sambuc, 13200, Arles, France
\textsuperscript{2}Department of Ecology, Environment and Plant Sciences, Stockholm University, Lilla Frescati, SE-10691, Stockholm, Sweden
\textsuperscript{3}Centre d’Ecologie Fonctionnelle et Evolutive, CEFE-CNRS UMR 5175, 1919 rue de Mende, 34293, Montpellier Cedex 5, France
\textsuperscript{4}Percy FitzPatrick Institute of African Ornithology, DST/NRF Centre of Excellence, Department of Biological Sciences, University of Cape Town, Rondebosch, 7701, South Africa
\textsuperscript{5}Departement Chrono-Environnement, UMR UFC/CNRS 6249 USC INRA, Université de Franche-Comté, Besançon, France

ABSTRACT

Only a few studies have shown positive impacts of ecological compensation on species dynamics affected by human activities. We argue that this is due to inappropriate methods used to forecast required compensation in environmental impact assessments. These assessments are mostly descriptive and only valid at limited spatial and temporal scales. However, habitat suitability models developed to predict the impacts of environmental changes on potential species’ distributions should provide rigorous science-based tools for compensation planning. Here we describe the two main classes of predictive models: correlative models and individual-based mechanistic models. We show how these models can be used alone or synoptically to improve compensation planning. While correlative models are easier to implement, they tend to ignore underlying ecological processes and lack accuracy. On the contrary, individual-based mechanistic models can integrate biological interactions, dispersal ability and adaptation. Moreover, among mechanistic models, those considering animal energy balance are particularly efficient at predicting the impact of foraging habitat loss. However, mechanistic models require more field data compared to correlative models. Hence we present two approaches which combine both methods for compensation planning, especially in relation to the spatial scale considered. We show how the availability of biological databases and software enabling fast and accurate population projections could be advantageously used to assess ecological compensation requirement efficiently in environmental impact assessments.

Key words: animal spatial dynamics, biodiversity, compensation, conservation, energy expenditure, habitat loss, mitigation, offsets, stakeholders.

CONTENTS

I. Introduction ................................................................. 2
II. Ecological compensation: where is science? ................................. 2
III. Habitat suitability models: relative payoffs of correlative and individual-based mechanistic approaches ........................................ 4
(1) Correlative approach ................................................. 4
(2) Individual-based mechanistic approach ................................... 6

* Address for correspondence (Tel: +46 739162288; E-mail: meineri.eric@gmail.com).
† These authors contributed equally to this work.
I. INTRODUCTION

It is estimated that 83% of all land surface has been affected by human activities (Sanderson et al., 2002). Change in land use is an important driver of biodiversity loss, and probably the most severe within terrestrial ecosystems (Sala et al., 2000). Land-use change detrimental to biodiversity is mainly caused by land clearing for agriculture and land transformation to infrastructure building such as urban and industrial sprawl, highway, harbour or airport building. In most regions of the world, policies have been implemented to balance anthropogenic alteration of biodiversity with land conservation or restoration in order to reach an objective of ‘no net loss’ of biodiversity, and if possible of ‘net gain’ (Slootweg et al., 2010).

Impacts of development projects affecting land use are generally subject to Environmental Impact Assessments (EIAs). Mitigation is at the core of the EIA process (Wood, 2003) and is generally implemented sequentially, first with attempts to avoid negative impacts, followed by minimization of degradation, rectification of negative impacts and, finally, compensation for unavoidable loss. Compensation consists in creating offset areas by restoring, enhancing, and sometimes preserving habitats harbouring equivalent biodiversity elsewhere. Approximately 16000 mitigation programs (or EIAs) are conducted across the European Union each year (GHK for DG Environment, 2010) and the US market for environmental mitigation is worth hundreds of millions of US dollars (Burgin, 2010). Environmental policy decisions widely influence natural systems and compensation could be a major avenue for limiting the negative impacts of human action on biodiversity (Kiesecker et al., 2010), even though the very efficiency of biodiversity trading may raise skepticism (e.g. Walker et al., 2009).

The achievement of a ‘no net loss’ objective depends on a sound estimation of project impacts upon species, of the offset required to compensate for the predicted loss, and of the restoration potential of the compensation area. This implies quantifying the size of the impact and of the need for compensation. The exchangeability between the two then is generally mediated by one or several metrics (or currencies), such as the surface of habitat affected by the project, the number of species and the size of the impacted populations, or any population traits that may be affected, such as body condition or survival probability at different life stages. Estimating how changes in habitat suitability will alter or improve the status of targeted populations is thus at the heart of the offset mechanism.

In principle, habitat suitability models (HSMs) developed to predict species distributions in the context of environmental change allow an assessment of such metrics in wild populations. HSMs have been developed along two main approaches. First, phenomenological or correlative models use species occurrences to infer their environmental requirements and to identify suitable habitats. These models are usually applied from worldwide- to landscape-scale predictions. Second, mechanistic models establish a causal relationship between species’ distributions and their environments. This modelling approach is most often applied from landscape to local scale (see Pearson & Dawson, 2003, for definition of landscape domain). In theory, correlative, mechanistic or hybrid HSMs could all offer a relevant contribution to conservation and decision-making within the process of ecological compensation. In practice however, despite major scientific achievements and prestigious publications, one is often left wondering how predictive modelling is actually being transferred into environmental policy and practice (Braunisch et al., 2012). As Meffe & Viederman (1995) pointed out, advances in conservation biology research are worthless if they are not translated into effective policy.

In this paper, we argue that HSMs could offer a way to ground ecological compensation into a robust scientific approach allowing a better-quantified assessment of what may be impacted and of the need for compensation. We thereby describe the relative pay-offs of correlative and mechanistic approaches and propose to combine these two classes of models to help implement efficient compensation measures.

II. ECOLOGICAL COMPENSATION: WHERE IS SCIENCE?

Ecological compensation generally requires metrics to define reference points (e.g. surface areas of habitats, number of species or ecological functions impacted), and to identify suitable replacement sites. Assessment methods to estimate impacts are multiple. For instance, a dozen assessment methods exist for wetland mitigation
Category: Distribution models and ecological compensation

Select the preferred level of detail for summarization.

Level 1: General overview

Level 2: Detailed discussion

Level 3: Technical analysis

---

**Distribution models and ecological compensation**

alone (Bartoldus, 1999). Most of these methods aim at determining the level of impacts on biodiversity, depending on the conservation status of the species impacted, the nature of the impact, the surface impacted, the species’ colonization ability, the ecological and geographical equivalence, and more rarely the uncertainty in the effectiveness of restoration actions. Offset ratios are then defined to achieve an appropriate exchange regarding the metrics used and to ensure no net loss in the long term. Ratios are set to reflect the number of offset units that must be provided to compensate one unit of loss at the project site (McKenney & Kiesecker, 2010). For example, a ratio of 3:1 means that 3 ha of wetland must be restored for each ha of natural wetland impacted or lost. An abundant literature discusses how ratios should be defined and used (see, for instance, Robb, 2002; Moilanen et al., 2009), yet no agreement has been reached upon a standardized approach. In practice, ratios are often the result of negotiation between regional policy makers, project managers and conservationists.

Synthetic ecological indicators were proposed as a standardized approach to assess biodiversity richness in habitats where a project is planned, and then define required ratios to compensate for biodiversity loss. These indicators describe ecosystems in simple terms, while being sensitive to ecosystem functioning (Griffith, 1997). They can be based on (i) the relative abundance of species in a community, reflecting its evenness, richness or diversity (e.g., Ludwig & Reynolds, 1988; Legendre & Legendre, 1998), (ii) the complexity of an ecosystem via the number of trophic levels (e.g., Kantoussan et al., 2010), (iii) the naturalness level of a patch as measured by comparing it with a long-undisturbed stand of the same plant community (Parkes, Newell & Cheal, 2003), or (iv) indicators based on remote-sensing, as for chlorophyll concentration reflecting biomass (e.g., Platt & Sathyendranath, 2008).

For example, levels of impacts and ratios chosen to compensate for the destruction of a 1800 ha wetland intended to become a new airport at Notre-Dame-des-Landes (near Nantes, Western France) were provided by experts of the Environmental Agency instructed to prepare the EIA. Levels of impact on a 1 ha surface area were converted to scores of compensation needs (e.g., major impact: score = 2 compensation units; moderate impact: score = 1; minor impact: score = 0.5) and were matched to units of response to the need for compensation (e.g., restoration of 1 ha of habitat = 2 units of response to the compensation need; improvement of an already existing habitat = 1 unit; and so on) (Courtejoie, 2014). Such ratios were said to depend on the ecological importance of the patch of impacted wetland, with respect to its hydrological function and its biodiversity. However, no scientific method was provided to assess the real impact on biodiversity, in particular in terms of reduction in population sizes of the species considered. Moreover, target sites where restoration was planned were chosen by satellite imagery crossing data from hydrographic systems and topography, independently from any prospective modeling of the potential carrying capacity of these sites for impacted species.

Overall, the few studies that monitored the outcome of compensation efforts demonstrated considerable variability in their success, and substantial room for improvement. A recent review showed the weak performance of wetland restoration both in terms of biodiversity and functionality (Moreno-Mateos et al., 2012). Several other evaluations revealed failures of compensation measures (Teels, Mazanti & Rewa, 2004; Wolters, Garbutt & Bakker, 2005; Quigley & Harper, 2006; Reiss, Hernandez & Brown, 2009; Maron et al., 2010; Tischew et al., 2010; Curran, Hellweg & Beck, In Press).

Given this accumulating evidence that compensation fails efficiently to offset biodiversity loss, concerns were expressed regarding the importance of respecting the mitigation hierarchy (Kiesecker et al., 2010; Quetier & Lavoire, 2011) and calls were made for long-term monitoring of offset measures (Race & Fonseca, 1996; Walker et al., 2009). Criticisms of current mitigation practices also focused on the lack of scientific background in decision-making (Harding et al., 2001; Rahn, Doremus & Eijsackers, 2006; Burgin, 2008). First, conservationists generally rely on expert opinion to assess human impacts and then determine ratios (Sutherland & Watkinson, 2001). Yet forecasted impacts are rarely based on scientifically sound evaluation (Sutherland, 2006). Second, when an environmental expertise study is carried out, it is usually mostly descriptive and valid at restricted spatial and temporal scales, with no visibility over longer periods (e.g. Fero et al., 2008). Further, a limited set of ecological criteria are considered and little attention is paid to the dynamic character of natural areas and their role for far-ranging species (i.e., foraging area, breeding site or resting place) (Robb, 2002). Moreover, ecological indicators rarely match the reality of biological systems (Turnhout, Hisschemoller & Eijsackers, 2007), can seldom be used and compared at different scales (Noss, 1990), and lack fixed quantitative reference values and methodological standards (Tolkamp & Van Rooy, 1990). Finally, species may show contrasting reactions to land-use changes due to different physiological thresholds and niche characteristics. Although this clearly makes the matter of compensation far more complex than the simple use of indicators, this issue is rarely considered (Morin & Thuiller, 2009).

Such high levels of uncertainty about the outcome of compensation measures is largely explained by methodological difficulties in assessing the complex functioning of ecosystems at varying spatial and temporal scales and the wide range of possible impacts. Hence, reliance on a general method to determine offset ratios seems inadequate. Offset calculation methods thus lack a
structured and transparent framework grounded in up-to-date ecological knowledge. We argue here that HSMs are efficient tools to predict how species’ suitable habitat can be altered by anthropogenic changes, and should therefore be considered an opportunity to meet the requirements of compensation planning.

III. HABITAT SUITABILITY MODELS: RELATIVE PAYOFFS OF CORRELATIVE AND INDIVIDUAL-BASED MECHANISTIC APPROACHES

Gathering the relevant scientific data, generating predictive scenarios, and ultimately maintaining the functioning of natural systems or restoring areas to offset losses are the three main challenges to managers involved in mitigation implementation. Compensation implementation should therefore include more quantitative, model-based decision-making (Rahn et al., 2006), ensuring well-articulated and standardized scientific and legal standards. Predictive modelling could be an essential tool within this process, integrating field measurements and experiments, and supporting expert judgment (Shugart, Smith & Post, 1992).

(1) Correlative approach

Correlative habitat suitability models (hereafter CHSMs), also known as species distribution models, bioclimatic envelope models, ecological niche models or climatic envelope models (Elith & Graham, 2009), are widely used in ecology to predict potential species’ distributions (mammals and birds: e.g. Peterson et al., 2002; butterflies: e.g. Beaumont & Hughes, 2002; amphibians and reptiles: e.g. Araujo & New, 2007; plants: e.g. Thuiller et al., 2005; Meineri, Skarpaas & Vandvik, 2012).

CHSMs are derived from Hutchinson’s (1957) niche theory and the gradient analysis of Whittaker (1956). They link current species observations with environmental variables, to yield predictions on current, past or future worldwide to landscape-scale species distributions (Pearson & Dawson, 2003; Elith & Graham, 2009; Fig. 1). More precisely, these models rely on the establishment of statistical relationships between environmental predictors (most often climatic and land-use variables) and observed species distributions (Pearson & Dawson, 2003; Austin, 2007; Beaumont et al., 2009; Elith & Graham, 2009) to infer species-specific environmental requirements. CHSMs rely on two major assumptions. First, species occurrences and explanatory variables (e.g. climate) are in a steady state with the environment, and second, niche is conserved through time and space (Elith & Leathwick, 2009). CHSMs are generally species-specific, although several single-species models can be combined to predict net biodiversity loss (e.g. Thomas et al., 2004; Thuiller et al., 2005) or can be coupled with ordination techniques to yield predictions at the community level [see Ferrier & Guisan (2006) and Baselga & Araujo (2009) for review on community modelling].

CHSMs are often based on presence/absence data [generalized linear model (GLM): Thuiller, 2003; Araujo et al., 2005; generalized additive model (GAM): Leathwick, Whitehead & McLeod, 1996; Midgley et al.,
Distribution models and ecological compensation

Kitching & Beck, 2011; Barbet-Massin et al.

Most CHSMs (e.g. Araujo et al., currently one of the most widely used tools to implement logistic regression (i.e. GAM, MARS), tree analyses [Generalized Boosted Regression Models (GBM), random forest analysis (RTA)], or on the principle of maximum entropy (e.g. MAXENT). As phenomenological models, the fit of a CHSM to empirical data is generally assessed by P values indicating the probability of an effect compared to a null model, but greater emphasis is currently made on Akaike information criterion (AIC) and multimodel inference (Burnham & Anderson, 2002). Model evaluation often consists in comparing model predictions with a subset of species observations excluded from the modelling process or in cross-validation techniques (Jeschke & Strayer, 2008). Model agreement is then assessed using area under the receiving operating curve (AUC), Kappa and/or true skill statistic (TSS) statistics. AUC has the advantage of directly comparing the probability of species occurrence predicted from the model with species observations. By contrast, predicted probabilities of species occurrence need to be transformed to presence/absence values before Kappa or TSS can be applied (Elith & Leathwick, 2009). Liu et al. (2005) detailed and compared several strategies that can be applied to transforming the results of habitat suitability models to presence/absence data.

However, predicted distributions can differ according to the modelling technique used (Thuiller, 2003; Elith et al., 2006). Hence, the biodiversity modelling package BIOMOD (Thuiller, 2003) has been developed to capitalize on the most widely used modelling techniques and produce more reliable predictions. BIOMOD is currently one of the most widely used tools to implement CHSMs (e.g. Araujo et al., 2004; Ballesteros-Mejía, Kitching & Beck, 2011; Barbet-Massin et al., 2012). More recently, ENMTools (Warren, Glor & Turelli, 2010) and ModEco (Guo & Liu, 2010) have been proposed as user-friendly software packages to facilitate quantitative comparisons of different CHSMs.

Numerous criticisms have been raised regarding the validity of CHSMs (Pearson & Dawson, 2003; Guisan & Thuiller, 2005; Heikkinen et al., 2006; Dormann, 2007). First, records of species absences which are often necessary to fit the models cannot be inferred with certainty (Elith et al., 2006). In response, several methods have been proposed to infer ‘pseudo-absences’ (e.g. Lutolf, Kienast & Guisan, 2006), as well as the use of modelling algorithms that do not require absence data. Still, absence records remain problematic as modelling algorithms often extract pseudo-absences internally using one of the many strategies described in the literature, therefore keeping uncertainty in the predictions. Similarly, public data sources must be used with caution for developing CHSMs, especially for subspecies that are morphologically similar but with distinct ecological requirements and geographic distributions (Lozier, Aniello & Hickerson, 2009). Second, CHSMs generally ignore spatial autocorrelation (SAC). SAC can be caused by distance-related biological processes, non-linear relationships modelled as linear, or simply by the absence of spatially structured environmental variables important for species distributions (Dutilleul, 2011). SAC can be problematic when present in model residuals as it may inflate type I errors or invert the slope of a response curve (Dormann et al., 2007). Although Dormann et al. (2007) proposed several solutions to account for SAC in correlative species distribution models, SAC issues are ignored in most recent publications. Third, because correlations between species’ distributions and environmental factors are obtained through presence/absence data, CHSMs are assumed to provide a spatial representation of the realized niche in space and time, and assume the realized niche to represent a species’ distribution (i.e. the environmental space where the species does occur; Pulliam, 2000; Sillero, 2011). However, the realized niche does not necessarily reflect the physiological limits of a species’ range, as individuals can occur across wider environmental ranges than have been recorded during a particular area/time. By contrast, species can be seen in areas which do not suit the set of conditions and resources necessary for a viable population (Pulliam, 2000). This can be explained either by the presence of transient individuals, or by artificial management allowing a species to be present (for instance birds receiving supplementary food; Robb et al., 2008). Source–sink dynamics can also contribute to define erroneous ecological niches and generate flawed CHSMs. Source–sink effects occur when some habitats of good quality (habitat source) allow recruitment that exceeds mortality, and export individuals to ‘sink’ areas where individual fitness is lower because of suboptimal environmental conditions. This phenomenon can greatly influence patterns of population dynamics (see for instance Gill et al., 2001). Identifying source and sink habitats is thus essential to define adequate species niches for CHSM parameterization.

Until recently, most CHSMs did not explicitly consider biotic interactions, evolutionary processes and dispersal ability (Pearson & Dawson, 2003; Hampe, 2004; Araujo & Guisan, 2006; Thuiller et al., 2008; Engler & Guisan, 2009; Huéy et al., 2012). However, some recent CHSMs included dispersal ability through modelling packages which combine correlations with mechanistic
cellular automata based on dispersal kernels (Engler & Guisan, 2009; Engler et al., 2009; Carvalho et al., 2010; Midgley et al., 2010; Pagel & Schurr, 2012). It can also be argued that because CHSMs predict the realized niche, biotic interactions may be implicitly accounted for within the abiotic predictor of the model. Yet, biotic interactions may differ in another environmental context and individual species responses to environmental change may be expected (Lenoir et al., 2008). Hence, such correlations may not apply in the future, especially as interspecies relationships may change (Prentice & Solomon, 1991; Woodward & Rocafort, 1991; Prentice et al., 1992; Pearson & Dawson, 2003). Recent improvements have been made in this regard and some studies have now explicitly included biotic interactions into regional- and landscape-scale models (see for instance Leathwick et al., 1996; Araujo & Luoto, 2007; Preston et al., 2008; Pellissier et al., 2010; Harris et al., 2012; Meineri et al., 2012). Yet, none of these studies managed to account for potential non-analogue biotic interactions when performing model projections in another spatio-temporal context.

Remote-sensing scientists designed accurate techniques to describe ecosystems (Horning et al., 2010). Combining remote-sensing parameters with species occurrence has been proved to be useful to define species’ niches more accurately, and to predict habitat suitability in the context of ecological compensation. This approach could become common following the model of other conservation-planning activities, such as the identification of suitable habitats for rare species (Gaubert, Papes & Peterson, 2006) and the design of reserves (Brito et al., 2009). For instance, incorporating primary production [Normalized Difference Vegetation Index (NDVI)] within a CHSM improved the assessment of habitat quality required for an endangered population of brown bears (Ursus arctos) in Spain (Wiegand et al., 2008). Also, a CHSM considering the wide dynamic range vegetation index (WDRVI) has been shown to constitute a suitable tool for characterizing giant panda (Ailuropoda melanoleuca) habitat and monitoring its temporal dynamics (Tuanmu et al., 2011). Finally, Brito et al. (2009) used remote sensing and presence data to identify biogeographic patterns for three species of North African canids in the Sahara. This work allowed the mapping of species distribution, highlighting suitable refuges for these endangered species.

Despite numerous recent improvements, CHSMs are still criticized for their rationale which relies on the link between species occurrence and environmental data without taking underlying mechanisms into account. They are nonetheless useful, easy to use and provide fairly good predictions at regional spatial scales (Pearson & Dawson, 2003; Austin, 2007). The required information is often freely available from museum and climate databases, so that they do not necessarily require time-consuming fieldwork.

(2) Individual-based mechanistic approach
(a) Description of models
Correlative models are at one extreme of a correlative/process model continuum, whilst mechanistic models are at the other extreme. In purely correlative models, processes are implicit. They remain unknown but are thought to cause the observed correlation. By contrast, mechanistic models rely on established causal relationships between species distributions and environmental variables, independently of data on species distributions (Kearney & Porter, 2004, 2009; Kearney et al., 2008; see also Fig. 1). These models are therefore described as ‘process-based’. Thus, mechanistic models do not use environmental variables as predictors to infer forthcoming distributions or processes. Rather, environmental variables are used as proximate information and input values. Such mechanistic models are bottom-up models since they aim at inferring the emergent properties of a system (population, community and ecosystem) from interactions among the individual components of this system (Grimm, 1999). This contrasts with correlative approaches, within which the modeller applies a correlative framework to any particular species or community, without explicitly taking into account processes underlying the causal relationship between species distributions and environmental variables. Mechanistic models thus provide a spatial and temporal representation of fundamental species niche (Sillero, 2011). More precisely, they attempt to simulate the mechanisms considered to underlie observed correlations with environmental attributes (Beerling, Huntley & Bailey, 1995), by using detailed knowledge of the ecophysiological responses of target species to environmental variables (Stephenson, 1998). Mechanistic models have also been referred to as ecophysiological models (Stephenson, 1998), process-oriented models (Carpenter, Gillison & Winter, 1993), process-based models (Morin & Thuiller, 2009), agent-based models (McLane et al., 2011), or thermodynamic models (Fort, Porter & Gremillet, 2009; Kearney et al., 2013). Indeed, studying behavioural and physiological adaptations of organisms to environmental conditions (e.g. thermoregulation and homeostasis) is fundamental in a changing world, to determine limitations to adaptive processes and conditions favourable to species maintenance (Le Maho, 2002). Mechanistic models were shown to deliver more accurate distribution predictions than CHSMs (Yates, Kittel & Cannon, 2000), and to offer great potential for improving their flexibility and performance (Buckley et al., 2010). However, these models are usually applied to smaller areas (extent and grain) than CHSMs, making them more local- to landscape-scale models (but see Yates et al., 2000).

Early mechanistic models had two major weaknesses, which have been addressed in more elaborated versions. First, they used to consider all individuals as identical in their behaviour and physiology. Second, individuals...
Distribution models and ecological compensation

were not spatially distributed in the environment. This means that all individuals had the same probability to interact, independently of the distance separating them. Such shortcomings led Huston, Deangelis & Post (1988) to propose an individual-based approach. Compared to CHSMs, such individual-based mechanistic models (hereafter IBMMs) present the fundamental advantage of including individual learning, adaptation and evolution (Grimm, 1999). Indeed, individual learning is a key component influencing individual fitness, which evolves as a function of habitat characteristics and inter- or intraspecific relationships. Each individual is informed on its local environment (e.g. resources, competitors). Information is processed by a set of decision rules, to produce an action that may in turn impact other individuals and/or the local environment. Hence, IBMMs may include game theory (e.g. Broom & Ruxton, 1998), as well as genetic aspects, to allow selective and/or stochastic population processes. IBMMs therefore allow individuals to vary in competitive ability, and to maximize their fitness. In this context, individuals use optimal decision rules, which means that they are likely to respond to environmental changes in the same way as ‘real’ organisms. Finally, IBMMs account for source–sink processes by estimating species range dynamics in both areas (Schurr et al., 2012). They are therefore expected to provide a reliable means of predicting how animal populations will be influenced by environmental change (see also Fig. 2). These models could be very useful to advise managers and policy makers in the context of ecological compensation, as they provide quantitative predictions on the effects of a project on biodiversity.

Because of these many advantages and of the improvement of computing facilities, IBMMs have generated increasing interest in recent years (Fig. 3) and have been subject to important developments. Nevertheless, IBMMs require abundant input data, sometimes difficult to obtain at regional scales. Fortunately, long-term studies have now accumulated vast amounts of information on species ecology. Table 1 describes the most widely used IBMMs, their main domain of use and availability (we do not present a similar table for CHSMs as this has been provided previously in numerous reviews, see e.g. Guisan & Zimmermann, 2000; Jeschke & Strayer, 2008).

(b) How individual-based mechanistic models can combine animal energetics, metapopulation processes and demography with environmental information to yield compensation planning

‘Energy is the common currency of life, as it fuels biological processes at every level of organization’ (Careau et al., 2008, p. 641). Hence, understanding foraging habitat selection as a function of animal energy requirements is a major goal in a world strongly and rapidly modified by human activities. We therefore argue that animal energetics should be taken into account when
establishing compensation measures. Mechanistic models including energy requirements and food availability as key input values are thus particularly relevant in the context of foraging habitat loss (Conover & Caudell, 2009; Sibly et al., 2013). On the basis of these two key parameters, IBMMs can predict habitat suitability and population parameters (e.g. survival rates, breeding rates), allowing the assessment of the impacts of a project before an offset area can truly meet compensation objectives.

Three main methods, reviewed in Fort, Porter & Gremillet (2010), are currently used to assess energy requirements in animals. First, time–energy budgets, combining time budgets with the energetic costs of each activity, are the most accurate method. However, they require extensive field data and time-consuming observations. Second, allometric equations are based on the relationship between field metabolic rates (FMRs) and body mass. This method requires very few field data. Yet allometric equations are not species-specific and do not integrate environmental factors which can strongly affect energy budgets. Third, thermodynamic models assess the theoretical energy requirements of animals through calculations based on the fundamental principles of thermodynamics (Kearney, 2012; Kearney et al., 2013). Such thermodynamic equations present three main advantages: (i) they are species-specific; (ii) they include environmental factors, such as wind or temperature; and (iii) they allow the assessment of energy expenditure with no or limited need for time-consuming observations or manipulations. Hence, thermodynamic equations represent the best trade-off when field data are scarce (Fort et al., 2009, 2010).

Another crucial advantage of IBMMs in the context of compensation implementation is that they allow forecasting impacts on the different individual components of an animal population. This proves useful when compensation measures can only focus on identified key stages of a population, such as reproductive adults or juveniles. For instance, Piou & Prevost (2012) developed an individual-based demo-genetic model to simulate population dynamics of the Atlantic salmon (Salmo salar) within southern European populations (individual-based Atlantic salmon model, IBASAM; see also Table 1). This model highlighted the importance of parameters related to juvenile growth in structuring the population, which is crucial to the assessment of extinction risk of the species, as well as the efficiency of mitigation strategies.

Finally, IBMMs may also incorporate habitat spatial configuration, and the impact of metapopulation processes on species dynamics, which are crucial features in conservation and for the management of patchy environments (Van Teeffelen, Vos & Opdam, 2012). This is an important point as EIAs generally only study direct on-site effects and neglect possible negative effects off the impacted site (Lenzen et al., 2003). For instance, disturbance is often an indirect negative impact influencing species at nearby sites (e.g. Summers, Cunnington & Fahrig, 2011). Considering nearby sites is also fundamental when defining an offset area. An offset area can contain appropriate food resources or appropriate resting places, but be inappropriate because of its configuration, for instance because it is surrounded by dense physical boundaries preventing dispersal or vigilant behaviour. Crucially, IBMMs take into account both the quality and the location of the disturbed site or the offset area, which are the two key elements conditioning species survival.

Several mechanistic models have been developed to study and forecast animal species’ distributions in the context of habitat changes (Table 1). For instance, Porter & Mitchell (2006) have established a spatially explicit state-of-the-art model (NicheMapper™) that...
<table>
<thead>
<tr>
<th>Model name</th>
<th>Reference</th>
<th>Main question</th>
<th>General functioning</th>
<th>Main input parameters</th>
<th>Main outputs parameters</th>
<th>Main domain of use</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>MORPH</td>
<td>Stillman (2008)</td>
<td>Predicting the effect of environmental change on foraging animal populations.</td>
<td>The key assumptions are that individuals within populations behave in order to maximize their perceived fitness, but that perceived fitness may not always be positively related to the actual chances of survival and reproduction.</td>
<td>Forager types, food patches distribution, food resource availability in each patch, forager's component target size (ash fresh dry matter (AFDM) or energy), food component value (in AFDM or energy content), forager's diet consumption rate.</td>
<td>Forager survival probability, mean final store size of foragers, time spent in each patch.</td>
<td>Birds (particularly waterbirds)</td>
<td>Restricted</td>
</tr>
<tr>
<td>WaderMorph</td>
<td>West et al. (2011)</td>
<td>Assessing the effects of mitigation measures for shorebirds.</td>
<td>Simplified version of MORPH.</td>
<td>Forager types and diet, location.</td>
<td>Forager survival probability, mean final store size of foragers, time spent in each patch.</td>
<td>Shorebirds</td>
<td>Free access on: <a href="http://individual">http://individual</a> ecology.bournemouth.ac.uk/software.html</td>
</tr>
<tr>
<td>NicheMapper™</td>
<td>Porter &amp; Mitchell (2006)</td>
<td>Estimating local hourly warmest and coolest microenvironments and the metabolic rate and other animal-dependent requirements (such as water and food requirements) necessary for an animal to maintain energy balance, and thereby self maintenance and survival, growth and reproductive requirements in a given environment.</td>
<td>Based on first principles of thermodynamics, this model integrates climate data and species-specific parameters to solve a coupled system of mass and heat transfers between the organism and its environment.</td>
<td>Climatic data: temperature, relative humidity, wind speed. Species-specific data: morphology (e.g. size of different parts of the body), physiology (including flight parameters) and behaviour (time energy budget).</td>
<td>Field energy expenditure in J/d, water requirements (g/time), activity time (h/d), wet food needed (g/d), shade requirements (%).</td>
<td>All vertebrates</td>
<td>Patented; version for R in development</td>
</tr>
<tr>
<td>Model name</td>
<td>Reference</td>
<td>Main question</td>
<td>General functioning</td>
<td>Main input parameters</td>
<td>Main outputs parameters</td>
<td>Main domain of use</td>
<td>Availability</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>---------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Individual-based Atlantic salmon model (IBASAM)</td>
<td>Piou &amp; Prevost (2012)</td>
<td>The general purpose of IBASAM is to evaluate how environmental conditions affect the evolution of <em>Salmo salar</em> life-history strategies and population size at the southern edge of their European distribution range.</td>
<td>The model is divided into eight submodels corresponding to lifecycle events and processes. To these biological submodels, two environmental submodels are added [river climate submodel (SM9) and ocean climate submodel (SM10)]. The model considers demographic stochasticity for mortality and reproduction. A sensitivity analysis determines crucial parameters in structuring the simulated populations.</td>
<td>Fish population description (including sex, age, mass, body length, fat reserve, location, number of returns from the sea, intrinsic metabolism and genetic material coding). Habitat description (daily observations of temperature and river flow, with stationary oceanic growth conditions).</td>
<td>Survival of different life stages of the population.</td>
<td>Constructed for salmon but should be transferable to other fish species.</td>
<td>Free access on: <a href="http://www.bordeaux-aquitaine.inra.fr/stpee/umrecobiop/outils/ibasam">http://www.bordeaux-aquitaine.inra.fr/stpee/umrecobiop/outils/ibasam</a></td>
</tr>
<tr>
<td>Trophic Link</td>
<td>Caron-Lormier <em>et al.</em> (2011)</td>
<td>Assessing the trophic-functional ecosystem model of arable agriculture.</td>
<td>This model is a trophic-functional simplification of plant and invertebrate individuals, and their interactions, focused on the arable ecosystem. However, it can be applied to any ecosystem or part of an ecosystem with any number of functional types. TrophicLink captures the important trophic interactions that drive growth dynamics.</td>
<td>As each trophic-functional type is composed of many different species, with distinct values, each trophic-functional type requires a considerable list of parameter values if it is to reflect the variation and diversity found in nature. Plants and invertebrates: growth rates, mortality rates. Plants only: energy allocated to seed production.</td>
<td>Individual patterns of growth of plants and invertebrates</td>
<td>Plants and invertebrates</td>
<td>Restricted</td>
</tr>
<tr>
<td>Model name</td>
<td>Reference</td>
<td>Main question</td>
<td>General functioning</td>
<td>Main input parameters</td>
<td>Main outputs parameters</td>
<td>Main domain of use</td>
<td>Availability</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>---------------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>-------------------------</td>
<td>--------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>AQUATOX</td>
<td>US Environmental Agency (2001)</td>
<td>Predicting the fate of various pollutants, such as nutrients and organic chemicals, and their effects on the ecosystem, including fish, invertebrates, and aquatic plants.</td>
<td>AQUATOX simulates the transfer of biomass, energy and chemicals from one compartment of the ecosystem to another. It does this by simultaneously computing each of the most important chemical or biological processes for each day of the simulation period.</td>
<td>Site parameters (width, length, depth, water inflow and discharge, nutrient loading, dissolved oxygen, pH, light, temperature, organic chemical content) and biotic parameters (phytoplankton, zooplankton, zoobenthos).</td>
<td>Outputs in terms of time-varying biomass of plants and animals, chemical concentrations in water, and concentrations of organic toxicants in water, organic sediments and biota.</td>
<td>Aquatic plants, invertebrates and fish</td>
<td>Free access on <a href="http://water.epa.gov/scitech/datait/models/aquatox/download.cfm#download">http://water.epa.gov/scitech/datait/models/aquatox/download.cfm#download</a></td>
</tr>
<tr>
<td>Mapped atmosphere-plant-soil system (MAPSS)</td>
<td>Neilson (1995)</td>
<td>Simulating the potential natural vegetation that can be supported at any upland site in the world under a long-term steady-state climate.</td>
<td>This global biogeography model operates on the fundamental principle that ecosystems will tend to maximize the leaf area that can be supported at a site by available soil moisture or energy.</td>
<td>Climate and soil data (including temperature, precipitation and surface roughness).</td>
<td>Hydrological and water-balance variables and vegetation variables (distribution, life form, and leaf area).</td>
<td>Vegetation</td>
<td>Free access on: <a href="http://www.daac.ornl.gov">http://www.daac.ornl.gov</a></td>
</tr>
<tr>
<td>BIOME2</td>
<td>Haxeltine et al. (1996); Haxeltine &amp; Prentice (1996)</td>
<td>Select which plant types may potentially be present at a particular site.</td>
<td>BIOME2 uses a coupled carbon and water-flux simulation model to capture the effects of minimum temperature tolerances and chilling requirements on determining the distributions of different plant types.</td>
<td>Latitude, soil texture, and mean monthly climate data (temperature, precipitation, and sunshine hours).</td>
<td>Net primary production and leaf area for the combination of major plant types. A rule base is then used to translate the model output into vegetation structural categories.</td>
<td>Vegetation</td>
<td>Restricted</td>
</tr>
</tbody>
</table>
incorporates principles from heat and mass transfer, engineering, physiology, morphology, and behaviour within a geographical information system (GIS) interface, to allow predictions across real landscapes. This model computes heat and water mass balances for required animal expenditures (Porter, Budaraju & Stewart, 2000; Porter et al., 2002).

Concomitantly, Stillman (2008) developed an individual-based modelling approach (MORPH) with an emphasis on foraging behaviour, to assess the effect of anthropogenic and environmental change on animals. MORPH’s key assumptions are that individuals behave in order to maximize their perceived fitness, but that perceived fitness may not always be positively related to the actual chances of survival and reproduction (i.e. individuals may make sub-optimal decisions; see Battin, 2004; Kokko & Lopez-Sepulcre, 2006). MORPH was initially used in waterbirds (see for instance Durell et al., 2006), but it is extremely flexible and potentially applicable to a wider range of species. The model does not calculate the energy requirements of the species, but this could be efficiently dealt with using the aforementioned NicheMapper™ model. Furthermore, West et al. (2011) developed a user-friendly interface which is a simplified version of MORPH used to assess the effects of mitigation measures for shorebirds (WaderMORPH). This IBMM is accessible online to non-specialists with a direct interest in coastal issues. To our knowledge, this is currently the only open-access tool for advising coastal management and policy.

Despite their obvious importance and their significant development (see Fig. 3), very little has been published on the successful use of IBMMs for implementing mitigation measures, or to assess their efficiency. One rare case study concerns the effects of an harbour extension in Le Havre (France) on shorebirds (Durell et al., 2005; Fig. 4). In this case, authorities proposed mitigation measures based on presumed negative effects while the project was already well underway. Durell et al. (2005) then used an IBMM to assess the efficiency of proposed mitigation measures, and advise harbour authorities on possible improvements. Specifically, the authors demonstrated an important negative effect of the forthcoming harbour extensions on the mortality rates and body condition of dunlins (*Calidris alpina*) from this area, and the necessity to create an additional buffer area to avoid disturbance to the birds (Fig. 4).

### IV. HOW TO MAKE THE BEST USE OF CORRELATIVE AND MECHANISTIC MODELS TO IMPLEMENT COMPENSATION

The simplicity of CHSMs has led to their popularity but also to extensive criticism (Guisan & Thuiller, 2005; Heikkinen et al., 2006; Dormann, 2007). Despite potential limitations, the importance of bioclimatic model predictions should not be underestimated (Araújo & Peterson, 2012). Following Pearson & Dawson (2003),
CHSMs predictions should rather be viewed and used as first approximations indicating the potential magnitude and broad pattern of future impacts, rather than as accurate simulations of future species distributions. Specifically, CHSMs lack of precision seems more critical at local scales (Pearson & Dawson, 2003; Austin, 2007) which is generally the scale used when implementing compensation (generally from a few ha, such as in infrastructure building, up to 10000 ha, such as in oil fields). Hence, the use of CHSMs alone does not seem appropriate to implement compensation measures. Conversely, the use of IBMMs can be time-consuming, especially at large scales where abundant data are required.

The dichotomy between CHSMs and IBMMs should be seen as the two ends of a continuum with respect to the explicit inclusion of processes (Dormann et al., 2012). For instance, only a few IBMMs are entirely developed ‘forward’, i.e. only based on known, sound, ecological knowledge. Indeed, many IBMMs use distributional data to evaluate model structure or to calibrate unmeasurable input parameters and are also correlative to a certain degree. Along the same lines, although CHSMs may suffer from spurious correlations, they generally rely on explanatory variables that are expected to represent causal mechanisms (Dormann et al., 2012), and can incorporate a few mechanistic processes to produce improved and informed predictions (Helmuth, 2009; Kearney & Porter, 2009; La Sorte & Jetz, 2010).

For instance, a new class of spatial models describe population growth rates by combining occurrence/abundance data with processes such as adaptation and competition, but also fecundity and survival probabilities (see for instance Schurr et al., 2012; Pagel & Schurr, 2012). These models were initially developed for plants (e.g. BioMove; Midgley et al., 2010), and are currently being adapted to animals, or even to the study of plant–animal interactions (Merow et al., 2011).

IBMMs and CHSMs could also be used sequentially and not necessarily through a hybrid approach. Indeed, CHSMs perform well at making broad predictions over regional scales and can be implemented more rapidly, whereas IBMMs are more difficult to implement at such large scales because of the wide range of required information and the extent of necessary computing resources. Thus CHSMs may help to identify a suitable geographic zone for a species or a set of species at current and future time periods. These zones can be used as offsets within a compensation process. IBMMs can be used thereafter to implement accurate compensation measures at higher spatio-temporal scales.

V. CONCLUSIONS

(1) Although ecological compensation has been implemented for more than 30 years on a worldwide scale, its objective of ‘no net loss’ has rarely been met, compromising biodiversity conservation. This may be due to lax environmental policies (Quétier, Regnery & Levrel, 2014), but is also strongly linked to methodological difficulties in assessing the impact of land-use change, and in proposing sound compensation measures.

(2) Habitat suitability models (HSMs) appear as an adapted and objective tool to advise compensation policy, since they allow predictions of potential species distributions following land-use changes. Moreover, in contrast to the current use of compensation ratios which use ad hoc assessment charts, HSMs provide a scientific approach which can be implemented to assess site-specific and long-term impacts.

(3) The broad types of HSMs (CHSMs and IBMMs) constitute two research avenues with ongoing development. We believe that these models should not be looked at in isolation, but rather in tandem. Overall, the use of HSMs needs to be adapted to the different stages of the compensation process, and to the particular aspect being tackled (e.g. spatial and temporal scale, target species to conserve, key stage within a population). At a fine spatial scale, IBMMs incorporating animal energetics are particularly relevant when the area to be lost is a foraging habitat, and combinations of models such as MORPH and NicheMapper™ appear to be extremely powerful.

(4) Although the use of HSMs is currently restricted to the scientific community, we posit that the recent development of user-friendly interfaces will contribute significantly to making them available to the much larger community involved in operative ecological compensation.

VI. ACKNOWLEDGEMENTS

We thank Vincent Devictor and two reviewers for constructive comments on the manuscript. This work was supported by the Foundation MAVA and a doctoral grant from Salinalgue, the Région Provence-Alpes-Côte d’Azur (South of France), and the French Fond Unique Interministériel (FUI).

VII. REFERENCES


Biological Reviews (2014) 000–000 © 2014 The Authors. Biological Reviews © 2014 Cambridge Philosophical Society
Distribution models and ecological compensation


Biological Reviews (2014) 000–000 © 2014 The Authors. Biological Reviews © 2014 Cambridge Philosophical Society