

# A trait-based approach for predicting species responses to environmental change from sparse data: how well might terrestrial mammals track climate change?

LUCA SANTINI<sup>1</sup>, THOMAS CORNULIER<sup>2</sup>, JAMES M. BULLOCK<sup>3</sup>, STEPHEN C. F. PALMER<sup>2</sup>, STEVEN M. WHITE<sup>3,4</sup>, JENNY A. HODGSON<sup>5</sup>, GRETA BOCEDI<sup>2</sup> and JUSTIN M. J. TRAVIS<sup>2</sup>

<sup>1</sup>Department of Biology and Biotechnologies, Sapienza Università di Roma, Viale dell'Università 32, 00185 Rome, Italy, <sup>2</sup>Institute of Biological and Environmental Sciences, University of Aberdeen, Zoology Building, Tillydrone Avenue, Aberdeen AB24 2TZ, UK, <sup>3</sup>NERC Centre for Ecology & Hydrology, Benson Lane, Wallingford OX10 8BB, UK, <sup>4</sup>Wolfson Centre for Mathematical Biology, Mathematical Institute, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK, <sup>5</sup>Department of Evolution, Ecology and Behaviour, University of Liverpool, Biosciences Building, Crown Street, Liverpool L69 7ZB, UK

## Abstract

Estimating population spread rates across multiple species is vital for projecting biodiversity responses to climate change. A major challenge is to parameterise spread models for many species. We introduce an approach that addresses this challenge, coupling a trait-based analysis with spatial population modelling to project spread rates for 15 000 virtual mammals with life histories that reflect those seen in the real world. Covariances among life-history traits are estimated from an extensive terrestrial mammal data set using Bayesian inference. We elucidate the relative roles of different life-history traits in driving modelled spread rates, demonstrating that any one alone will be a poor predictor. We also estimate that around 30% of mammal species have potential spread rates slower than the global mean velocity of climate change. This novel trait-space-demographic modelling approach has broad applicability for tackling many key ecological questions for which we have the models but are hindered by data availability.

**Keywords:** climate change velocity, demographic models, dispersal, integrodifference equations, life-history traits, population spread rate, range shift, rangeShifter, trait space, virtual species

Received 20 November 2015 and accepted 29 February 2016

## Introduction

The rate at which a population can spread across space is likely a key determinant of how well species are able to respond to climate change (Pacifi *et al.*, 2015). Until recently, almost all projections of species' future distributions have largely ignored the process of population spread (Travis *et al.*, 2013). For more than a decade, the field of climate envelope modelling relied almost exclusively on projections that made one of two extreme assumptions in relation to population spread: no dispersal vs. unlimited dispersal (Bateman *et al.*, 2013), implying either that a species would be unable to colonise any newly suitable regions or that it would be able rapidly to reach all of the newly available suitable climate space. Recognising the potential limitations of this approach, some authors have considered 'partial dispersal scenarios' (Bateman *et al.*, 2013) that rely on average dispersal distance and the number of dispersal events in a given time frame (Hannah *et al.*, 2005; Schloss *et al.*, 2012; Visconti *et al.*, 2015) in order to make predictions for how well large

numbers of species are likely to be able to track a shifting climate.

However, over the last few years, increasing recognition of the importance of ecological and evolutionary dynamics of range shifts has resulted in calls for the development of a new generation of models for forecasting biodiversity futures (Dormann *et al.*, 2012; Schurr *et al.*, 2012; Travis *et al.*, 2013), and dispersal has been highlighted as a critical process for inclusion (Huntley *et al.*, 2010). This call is being met, and there has already been a proliferation of models for biodiversity forecasting that incorporate increased biological realism [see Lurgi *et al.* (2015) for a recent review of such models]. These models represent ecological and evolutionary processes in differing degrees of detail. Thus, we already possess a good theoretical understanding of key determinants of spread rate. The main reason for the continued incorporation of reduced ecological realism in models forecasting the dynamics of large numbers of species is likely the lack of sufficient high-quality ecological data for parameterisation, rather than the lack of appropriate, and sufficiently efficient, modelling approaches. A key challenge will be to use

Correspondence: Luca Santini, e-mail: luca.santini.eco@gmail.com

these models for anything more than a relatively small set of species for which we have the required data for parameterisation (e.g. Nathan *et al.*, 2011; Bullock *et al.*, 2012). We need approaches for making the best possible use of the considerable available ecological data that exist across many species, given that they are sparse and patchy in nature. Here, we introduce the concept of using a trait-space approach for understanding how spread rates will vary across a wide and realistic range of life histories.

Biological traits are not assembled at random in species, but show various degrees of covariation which reflect evolutionary optimal strategies and physical constraints (Bielby *et al.*, 2007). An understanding of how traits are combined may enable one to make inferences about the biological traits of poorly known species, while accounting for the biological variation observed in nature. Trait-based approaches have become used increasingly in several ecological fields including biodiversity-provisioned ecosystem services (Suding *et al.*, 2008; Diaz *et al.*, 2013), assessing species' intrinsic vulnerability to extinction (Purvis *et al.*, 2000; Cardillo *et al.*, 2006; González-Suárez & Revilla, 2013) and phylogenetic comparative analyses (FitzJohn *et al.*, 2009; Santini *et al.*, 2015). All these fields have made a different use of traits, but share a focus on biological traits rather than species.

In this study, we develop a novel approach to determine which life-history traits are the best predictors of spread rate and also illustrate how we can use the method to determine the proportion and types of species within a defined (e.g. taxonomic) group that are likely to have insufficiently high rates of spread to keep pace with climate change. We use the life-history data available across terrestrial mammal species to fit a multivariate trait-space model. Terrestrial mammals exhibit very diverse ecologies, and are one taxon for which a good amount of ecological information is available (Jones *et al.*, 2009). Yet, we have the complete data needed to model spread for few terrestrial mammal species. While certain traits (e.g. body mass) are better documented, ecological variables related to dispersal or demographic parameters are lacking or poorly known for most species, and when available are often uncertain. The model that we develop is able to predict missing trait value combinations based on our knowledge of traits' covariation in mammals. Having a large number of spread rates for virtual species, representing life histories that are realistically constrained, offers opportunities for addressing important fundamental and applied questions. Crucially, adopting this approach removes the need to have complete sets of life-history data for many species; instead, a statistical description of trait

space, including the covariation between different traits, can be derived from the patchy data that are available across many species.

By generating large sets of virtual species (trait value combinations), each with its complete life-history data, we then use two well-established demographic modelling approaches – analytical integrodifference equations (IDEs; Neubert & Caswell, 2000) and the individual-based model (IBM) RangeShifter (Bocedi *et al.*, 2014a) – to project spread rates for a large number of species. We use the two, quite different modelling approaches (Travis *et al.*, 2011) to ascertain the robustness of our trait-space method. To demonstrate the utility of this novel method, we then:

1. Test relationships between traits that are more widely available (e.g. body mass) and our modelled rate of spread to establish the degree to which these 'proxy' traits may be used as first-order estimators of a species ability to shift its range under a changing environment.
2. Provide an estimate for the proportion of terrestrial mammal species that are likely to have spread rates slower than the global mean velocity of climate change (Loarie *et al.*, 2009), also highlighting which types of species are likely to be those that fail to keep pace.
3. Establish, across mammalian trait space, the degree of consistency in estimates of spread rate obtained between a rapid analytical approach and a much more computationally demanding individual-based simulation.

## Materials and methods

### *Modelling trait covariation and virtual species simulation*

We compiled data on 10 life-history traits for terrestrial mammals. These were chosen to represent either traits directly affecting population dynamics: age at sexual maturity (SxMat), litters per year (NLit), litter size (LitS), median Euclidean dispersal distance (DDist), adult annual survival (Surv) and average longevity (Long); or traits that can be used as proxies for the former: home range size (HR), population density (Dens), body mass (Mass) and trophic level (Diet). Mass in mammals is known to be related to many other traits in a log-linear fashion, and Diet often influences these relationships (Hendriks *et al.*, 2009). SxMat, NLit, LitS, DDist, Long and Surv were used to parameterise the two models of spread rate (IDE and IBM). Dens was used to parameterise the IBM, and Mass, Dens and HR were assessed as proxy traits of spread rate. We define traits broadly here, as features that can be considered as characteristic of a species and which can be measured at the individual or population level. A full description of the compiled data, their sources, units and sample

sizes, and rationale for their inclusion are available in Table S1. While traits such as body mass and trophic level are widely available for a large number of species, traits such as dispersal distance and annual survival are sparse and are population- and context-dependent.

To simulate virtual species presenting complete and realistic combinations of life-history traits, we started by constructing a model of allometric relationships between all ten traits from the compiled data. We used a multivariate Gaussian (i.e. multiresponse) mixed model to estimate correlations between traits and covariates while accounting for broad phylogenetic structures. We adopted a Bayesian approach employing latent variables (predictors representing the unknown true value of a process which may or may not be directly observed) to deal with missing trait values while retaining information provided by species for which only partial data were available. For each single-response trait, we used body mass and diet as fixed-effect predictors. We treated body mass as a covariate rather than another trait/response because it is strongly related to many other traits (Bielby *et al.*, 2007) and we were interested in simulating virtual species according to broad species categories. All response variables were transformed to ensure approximate normality of residuals and finally centred and scaled prior to fitting the model. We chose the following model structure:

$$\begin{aligned} \text{Lit}_i &= \alpha_{11} \cdot \text{Diet}_i^c + \alpha_{12} \cdot \log(\text{Mass}_i) + \alpha_{13} \cdot \log(\text{Mass}_i) \cdot \text{Diet}_i^c \\ &\quad + \alpha_{14} \cdot \log(\text{Mass}_i)^2 + \alpha_{15} \cdot \log(\text{Mass}_i)^2 \cdot \text{Diet}_i^c + \omega_{k,1} + \varepsilon_{i,1} \\ \text{SxMat}_i &= \alpha_{21} \cdot \text{Diet}_i^c + \alpha_{22} \cdot \log(\text{Mass}_i) + \alpha_{23} \cdot \log(\text{Mass}_i) \cdot \text{Diet}_i^c \\ &\quad + \omega_{k,2} + \varepsilon_{i,2} \\ \text{NLit}_i &= \alpha_{31} \cdot \text{Diet}_i^c + \alpha_{32} \cdot \log(\text{Mass}_i) + \alpha_{33} \cdot \log(\text{Mass}_i) \cdot \text{Diet}_i^c + \omega_{k,3} \\ &\quad + \varepsilon_{i,3} \\ \text{Dens}_i &= \alpha_{41} \cdot \text{Diet}_i^c + \alpha_{42} \cdot \log(\text{Mass}_i) + \alpha_{43} \cdot \log(\text{Mass}_i) \cdot \text{Diet}_i^c \\ &\quad + \omega_{k,4} + \varepsilon_{i,4} \\ \text{HR}_i &= \alpha_{51} \cdot \text{Diet}_i^c + \alpha_{52} \cdot \log(\text{Mass}_i) + \alpha_{53} \cdot \log(\text{Mass}_i) \cdot \text{Diet}_i^c + \omega_{k,5} \\ &\quad + \varepsilon_{i,5} \\ \text{DDist}_i &= \alpha_{61} \cdot \text{Diet}_i^c + \alpha_{62} \cdot \log(\text{Mass}_i) + \alpha_{63} \cdot \log(\text{Mass}_i) \cdot \text{Diet}_i^c \\ &\quad + \omega_{k,6} + \varepsilon_{i,6} \\ \text{Surv}_i &= \alpha_{71} \cdot \text{Diet}_i^c + \alpha_{72} \cdot \log(\text{Mass}_i) + \alpha_{73} \cdot \log(\text{Mass}_i) \cdot \text{Diet}_i^c + \omega_{k,7} \\ &\quad + \varepsilon_{i,7} \\ \text{Gest}_i &= \alpha_{81} \cdot \text{Diet}_i^c + \alpha_{82} \cdot \log(\text{Mass}_i) + \alpha_{83} \cdot \log(\text{Mass}_i) \cdot \text{Diet}_i^c + \omega_{k,8} \\ &\quad + \varepsilon_{i,8} \\ \text{Long}_i &= \alpha_{91} \cdot \text{Diet}_i^c + \alpha_{92} \cdot \log(\text{Mass}_i) + \alpha_{93} \cdot \log(\text{Mass}_i) \cdot \text{Diet}_i^c \\ &\quad + \omega_{k,9} + \varepsilon_{i,9} \end{aligned}$$

where  $\alpha$  is an array of coefficients to be estimated,  $i$  and  $k$  are respectively the species and Order indices, and  $\text{Diet}^c$  is an indicator variable for carnivore diet.

With  $j$  being the Trait index, the random taxonomic Order effects  $\omega_{k,j}$  follow a normal distribution with mean zero and variance  $\theta_j$  and errors  $\varepsilon_{i,j}$  follow a multivariate-normal distribution with mean 0 and full symmetric  $9 \times 9$  variance-

covariance matrix  $\Sigma_\varepsilon$ :

$$\varepsilon_{i,j} \sim \text{MVN}(0, \Sigma_\varepsilon)$$

$$\Sigma_\varepsilon = \begin{bmatrix} \sigma_1^2 & \sigma_{2,1} & \cdots & \sigma_{9,1} \\ \sigma_{2,1} & \sigma_2^2 & \cdots & \sigma_{9,2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{9,1} & \sigma_{9,2} & \cdots & \sigma_9^2 \end{bmatrix}$$

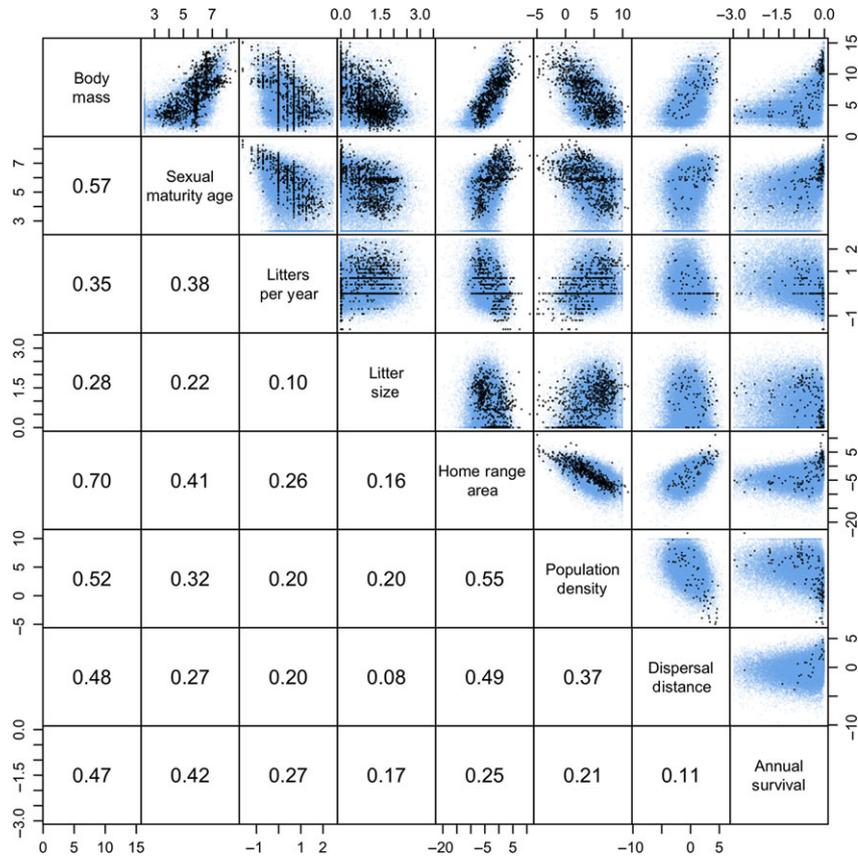
The model was fitted using the MCMCglmm package for R 3.0.2 (Hadfield, 2010), using 6 000 000 iterations and a burn-in of 200 000 iterations.

### Drawing sets of virtual species from trait space

Having fitted the model of life-history space, accounting for correlation between traits, the next task was to draw sets of virtual species (realistic combinations of trait values) for which demographic modelling can be used to determine spread rates. For a given body mass and diet, a virtual species was drawn by (i) predicting all mean trait values from the model, (ii) adding normal variation (between-Order random effect) to these predictions with mean 0 and variance  $\theta_j$  and (iii) adding multivariate-normal variation (corresponding to model residuals) with mean 0 and variance  $\Sigma_\varepsilon$ . For our spread modelling, we drew two sets of virtual species. The first set comprised 15 000 species simulated with body masses sampled from a log-uniform distribution  $U [1.5, 15 \log(\text{g})]$ , reflecting the range of body masses (in g) observed in terrestrial mammals and a diet, either carnivore or omnivore/herbivore, sampled from a binomial ( $P = 0.5$ ) distribution (see Fig. 1). Figure 1 shows that the simulated traits of the virtual species well capture the main features of trait distribution and covariation observed in the empirical data set. This allowed us to fill gaps of information for less known traits while considering their variability and relationships with other traits. We used this set of virtual species to parameterise both IDE and IBM approaches (see Appendix S2) to derive spread rates for species representing the full range of body mass (in herbivores and carnivores), and to determine the relationships between different life-history traits and spread rates, and to make a thorough comparison of outcomes from the two modelling approaches. A second set of >50 000 species was sampled from the observed distribution of mammalian body mass and diet (Wilman *et al.*, 2014); this enabled us to estimate the proportion of real mammal species that have modelled spread rates lower than current estimates for climate velocity. Because of the stochastic nature of the virtual species sampling, we sampled 10 replicates for each of >5000 species' body masses and diet for a total of >50 000 species, thus maintaining the observed proportions across species body mass and diets.

### Analytical model

We modelled the range expansion velocity of the virtual species using a stage-structured IDE as derived by Neubert & Caswell (2000). A full description of how population projection matrices were built for each virtual species using the trait



**Fig. 1** Correlations between log-transformed biological traits in terrestrial mammals, both in empirical data (black dots) and in simulated data (light blue dots). The lower panels show the Pearson's correlation coefficients of the relationships.

values for litter size, litters per year, age at sexual maturity, longevity and annual survival is given in Appendix S2.

The model is given by Eqn (1) and describes how the population density  $n$  (vector representing all of the life stages) at each location  $x$  in continuous, infinite space changes from time  $t$  to  $t + 1$  (which represents a year in this study)

$$n(x, t + 1) = \int_{-\infty}^{\infty} [K(x - y) \circ B_n] n(y, t) dy \quad (1)$$

Here,  $\circ$  indicates elementwise multiplication,  $B_n$  is a stage-structured population projection matrix that describes density-dependent population growth at location  $y$ , and  $K(x - y)$  is a matrix of dispersal kernels that describes the set of probabilities of the relocation from  $y$  to  $x$  of individuals undergoing each demographic transition. In summary, over a time step the population grows at each location  $y$  and individuals are dispersed. The population at location  $x$  is given by integrating this process over all locations  $y$ . Calculation of the population spread rate requires a population projection matrix describing demography at low density (i.e. at the forefront of the spreading population;  $B_0$ ) and a matrix  $M(s)$ , which describes the dispersal kernel for each demographic transition in terms of a moment-generating function (mgf). In the absence of good information on mammal dispersal kernels, and for simplicity, we assumed an exponential kernel for each dispersive stage,

which has a mgf =  $1/(1 - \alpha s)$ , where  $\alpha$  is the mean dispersal distance (where mean DDist =  $\frac{\text{median DDist}}{\log(2)}$ ), as derived for each virtual species, and  $s$  is the wave shape parameter. Under this model, a population forms a wave of constant shape that advances at constant speed, and this asymptotic wavespeed  $c^*$  can be derived analytically (Neubert & Caswell, 2000) as

$$c^* = \min_{0 < s < \infty} \left[ \frac{1}{s} \ln \rho_1(s) \right] \quad (2)$$

where  $\rho_1$  is the dominant eigenvalue of the matrix that is the product of the demographic and dispersal matrices [ $B_n \circ M(s)$ ]. This approach includes simplifying assumptions such as no temporal variation or Allee effects, isotropic dispersal, and the environment is treated as spatially homogeneous.

### Stochastic individual-based model

We used RangeShifter, a single-species, spatially explicit, individual-based simulation platform (Bocedi *et al.*, 2014a). RangeShifter integrates complex population dynamics with dispersal behaviour which can be modelled in either a phenomenological (dispersal kernels) or mechanistic (movement models) way. Particularly, for stage-structure population dynamics, RangeShifter translates classic population

projection matrices (Caswell, 2001) into equivalent parameters for the IBM (see Appendix S2). In our simulations, species were allowed to expand their range across strips of homogeneous gridded landscapes where all cells were considered suitable for the species. At the beginning of each simulation, the first row of the landscape was initialised with a number of individuals equal to the total row carrying capacity as derived from the species population density. Demography was determined by the population matrix equivalent to that parameterised in the analytical model. All individuals dispersed at the end of their juvenile stage for a distance drawn randomly from a negative exponential dispersal kernel as derived for the IDE. As RangeShifter is stochastic, each species simulation was replicated 10 times. We calculated the rate of spread by dividing the distance covered by years of simulations. The distance covered was measured as the mean of the distance of the farthest five rows of cells weighted by the number of individuals present. The distance covered was then averaged across the ten replicates and divided by the years of simulation. See Appendix S1 for a more detailed explanation of the IBM simulations.

### Analyses

To assess which traits best predict modelled IDE spread rate, we fitted a generalised additive model (GAM), which included all other biological traits as a predictor and a smooth term for SxMat as showing a nonlinearity ( $GM_{allTraits}$ ). To assess the variance explained by individual life histories, we performed a GLM for each biological predictor separately, and a GAM for SxMat ( $GM_{indTraits}$ ). Similarly, we used a GAM to assess the relative importance of demography and dispersal in predicting spread rates and included dispersal distance and population growth rate (the dominant eigenvalue of the population matrix) ( $GM_{demo}$ ). We fitted a smooth term for population growth rate as it presented an asymptotic relationship with the rate of spread. We fitted two GLMs and a GAM to establish relationships between the rate of spread and body mass ( $GM_{proxyMass}$ ), population density ( $GM_{proxyDens}$ ) and home range area ( $GM_{proxyHR}$ ), respectively. Because these variables are considered to be important predictors of dispersal distance in mammals (Whitmee & Orme, 2012; Santini *et al.*, 2013) and are correlated with all other traits in the virtual data set, we explored their possible value as a proxy for predicting spread rate in mammals. All GLMs and GAMs assumed a Gaussian error and an identity link function. The rate of spread (IDE) and all predictors were log-transformed. The predictors of multivariate models ( $GM_{allTraits}$  and  $GM_{demo}$ ) were also standardised prior to fitting the models in order to compare their effect sizes.

To estimate the proportion of mammalian species that are likely to have spread rates slower than the global mean velocity of climate change, we first differentiated mammal species according to their distribution in each biome as defined by Olson *et al.* (2001). We identified those mammal species occurring in each biome by overlaying species geographic ranges (IUCN, 2015) with biomes: species whose majority (>50%) of range overlapped with a specific biome were considered as

present. According to the observed distribution of body mass and trophic level by biomes, we divided the set of virtual species into subsets representing the observed distribution of mammalian body masses and trophic levels in each biome. We then compared the spread rate predictions for all virtual species, and for each subset, with the geometric mean of the distribution of predicted climate change velocity as estimated in Loarie *et al.* (2009) both globally and for individual biomes. We limited this analysis to biomes with >50 species.

Finally, to establish, across mammalian trait space, the degree of consistency in estimates of spread rate obtained between the analytical approach (IDE) and the individual-based simulation (IBM), we compared the two modelling approaches and the effect sizes of different biological traits on spread rate using a MCMCglmm multiresponse model (Hadfield, 2010; see Appendix S1 for more details).

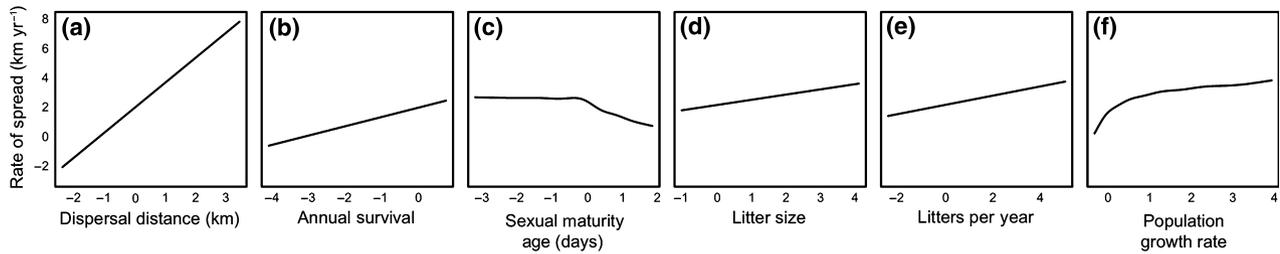
Integrodifference equation modelling was performed in MatLab (MATLAB and Statistics Toolbox Release, 2012), all data analyses were conducted in R (R Core Team, 2014), and GIS analyses were performed in GRASS GIS (GRASS Development Team, 2012).

### Results

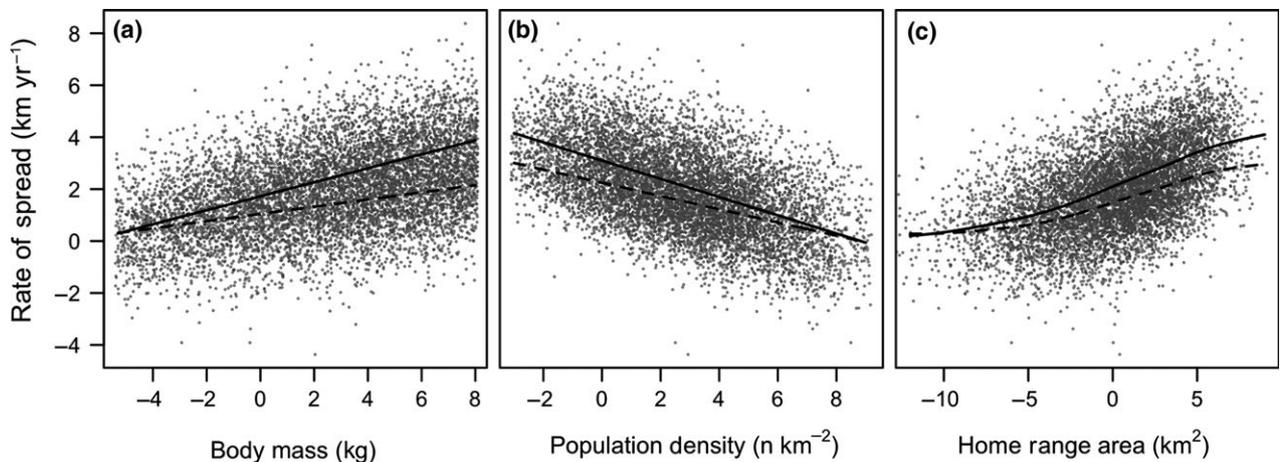
#### *Life-history determinants of spread rates in terrestrial mammals*

The relative importance of different traits in driving spread rate can be inferred from the steepness of their relationships (Fig. 2a–e). Spread rate is primarily related to changes in median dispersal distance, followed in order by annual survival, sexual maturity age (inversely), litter size and litters per year ( $GM_{allTraits}$ :  $R^2 = 0.97$ ; Table S2; Fig. 2a–e), although sexual maturity age only affects spread rate for values higher than 1 year of age. Except for dispersal distance, individual life-history traits explained a low proportion of the variance ( $R^2$  for  $GM_{indTraits}$ : Dispersal Distance = 0.77; Litters per Year = 0.17; Annual Survival = 0.08; Sexual Maturity Age = 0.04; Litter Size = 0.02). A curvilinear effect is evident in the partial dependence on the population growth rate, which shows a strong effect for slight increase in growth rate at low growth rates, and smaller effects for higher growth rates ( $GM_{demo}$ :  $R^2 = 0.97$ ; Fig. 2f).

Considering the proxy predictors, spread rate tends to increase with increasing species body mass ( $GM_{proxyMass}$ :  $R^2 = 0.30$ ) and home range area ( $GM_{proxyHR}$ :  $R^2 = 0.32$ ) and to decrease with increasing population density ( $GM_{proxyDens}$ :  $R^2 = 0.33$ ), and at higher rates in carnivores than in omnivores/herbivores (Fig. 3). While body mass and population density yield a log-linear relationship with spread rate, the home range relationship is sigmoidal with a slower increase for small and large home range sizes (Fig. 3). A key



**Fig. 2** Partial dependence of rate of spread (as predicted by the analytical integrodifference equation) on biological traits (a–e) and population growth rate (f) based on two generalised additive models. In the first (GMallTraits), modelled rate of spread is predicted using all biological variables used for building population matrices (a–e), while in the second (GMdemo), it is predicted using median dispersal distance and population growth rate (f; dominant eigenvalue of the population matrix). All variables were log-transformed and standardised prior to fitting the model.



**Fig. 3** Relationship of log-transformed body mass, population density and home range area with log-transformed integrodifference equation-predicted rate of spread fitted with GLM (a and b) and generalised additive model (c). Solid line = carnivores; dashed line = omnivores and herbivores.

conclusion to be drawn from this analysis is that any one proxy trait on its own provides a very weak predictor of a species spread rate (see the substantial scatter in all three panels of Fig. 3).

#### *Estimates of the proportion of 'at-risk' species*

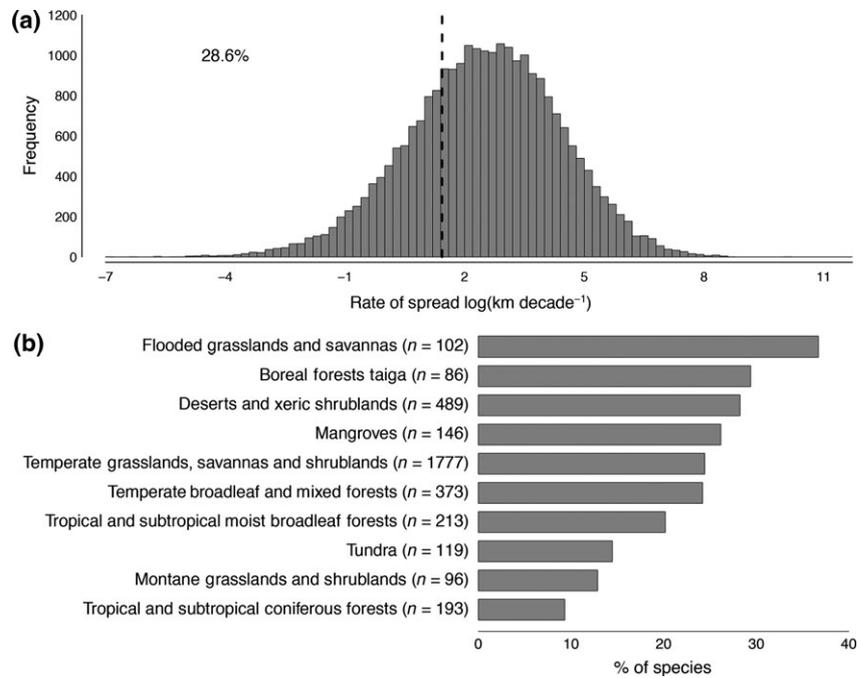
Globally, almost 30% of species' spread rates fall below the geometric mean of the predicted climate change velocity (Fig. 4a). This proportion greatly varies across biomes, as a function of both local assemblage body mass and trophic level distribution and local climate velocity (Fig. 4b). At one extreme, we find tropical and subtropical coniferous forests, and montane grassland and shrublands, where about 10% of the species are predicted to spread slower than climate velocity. At the other extreme, we find flooded grassland and savannas, boreal forests (taiga), mangroves, and deserts and xeric shrublands with percentages reaching 36% of species not able to keep pace with climate change.

#### *How the analytical and individual-based estimates of spread rate compare*

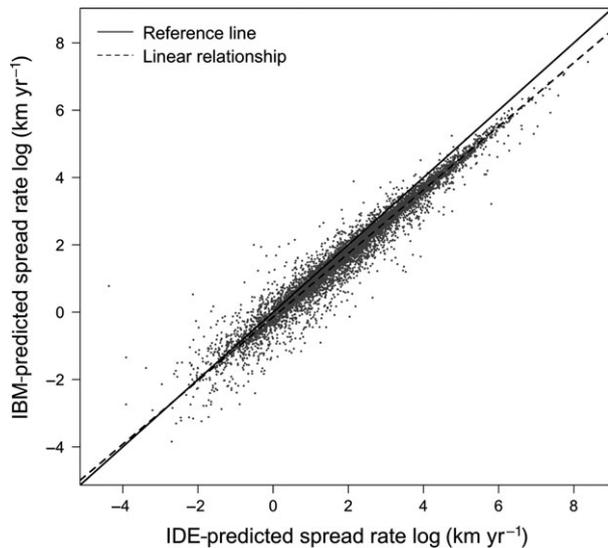
The predicted rates of population spread of the two models were strongly correlated (Pearson's  $r = 0.73$ ) (Fig. 5; Table S3), but the slope of the relationship deviated from 1 : 1. The IDE was characterised by higher predicted spread rates and lower variance ( $\sigma^2 = 0.20$ ) than the IBM ( $\sigma^2 = 0.24$ ). Biological predictors of spread rate have comparable effects in the two models, with the only exception being somewhat lower contributions of litter size and dispersal distance variables in the IBM (Table S2), a difference likely due to the intrinsic stochasticity in these parameters in the latter model.

#### **Discussion**

The trait-space modelling approach presented here can be applied for exploring a wide range of processes when species' data are limited, while using the



**Fig. 4** (a) Distribution of predicted log-transformed mammalian spreading rates. The dashed line represents the global geometric mean of climate change velocity as predicted by Loarie *et al.* (2009), and the percentage represents the species that are estimated to have a projected spread rate slower than the climate change velocity. (b) Percentages of mammal species that are projected to have a potential spread rate slower than predicted climate change velocity in Loarie *et al.* (2009) divided among the world's biomes. The number of species considered (species range overlapping  $\geq 50\%$  with the biomes) is reported in brackets.



**Fig. 5** Relationship between the projected rate of spread from the analytical IDE (integro-difference equation) and the IBM (RangeShifter). Dashed line = 1 : 1 relationship; solid line = linear relationship between the two models' output (major axis regression:  $IBM \sim -0.13 * IDE - 0.15$ ).

available empirical data to constrain simulations within ecologically realistic scenarios. It therefore addresses the very real challenge currently faced by ecologists in

predicting how species' populations will spread in response to environmental change.

#### Biological predictors of spread rate

The method adopted to simulate virtual species allowed us to test a large number of trait value combinations which encompass the variation observed across terrestrial mammals. The most important predictor of the projected rate of spread is dispersal distance, followed by the annual survival, the age at which species disperse and are able to reproduce, the size of the litter and the interval between successive reproductive events. The major role of dispersal distance and age at reproduction in spread rate has also been found for invasive plants (Coutts *et al.*, 2011). However, except for dispersal distance which is rarely known, single traits are not good predictors and the velocity at which populations spread is better described by a multidimensional predictor describing both the distance dispersed by individuals and the overall growth rate of the population.

We found a positive diet-dependent relationship between spread rate and body mass, and an even stronger positive relationship with population density and home range size. These probably reflect the

well-documented relationships between these traits and dispersal distance in mammals (Whitmee & Orme, 2012; Santini *et al.*, 2013). Yet, the scatter around these relationships limits their usefulness in making predictions.

#### *Mammal abilities to track shifting climate*

Worryingly, the models projected that many mammal species may spread at a slower velocity than that predicted for the shifting climate (i.e. Loarie *et al.*, 2009). This endorses previous studies, which suggest that a majority of mammal species are likely to lose parts of their ranges in the near future (Thomas *et al.*, 2004; Thuiller *et al.*, 2006; Levinsky *et al.*, 2007; Schloss *et al.*, 2012) and that measures to mitigate this effect will be necessary. Species that will be mostly affected by range loss are small species that have short dispersal distances and that occur in biomes where the climate is shifting more rapidly. The mammalian assemblages of flooded grasslands and savannas, taiga, deserts and xeric shrublands, and mangroves are projected to be particularly threatened. Overall, close to 30% of species are projected to be unable to spread faster than future climate change. Realised spread rates could be dramatically slower where the natural habitat of the biome is largely converted and fragmented, as is certainly the case for some of the biomes mentioned above.

It is important to note that we focussed on only one of the factors – albeit a very important one, population spread rate – that determines vulnerability to climate change (Pacifi *et al.*, 2015). We compared spread rate to a simplified climate change velocity that only considers average temperature changes, while ignoring other changes in climatic variables, land use changes and species interactions (Lenoir & Svenning, 2014). Also, although a low velocity of spread in relation to local climate velocity is an indication of future range loss, only spatially explicit models allow one to provide quantitative estimates of range loss (Thomas *et al.*, 2004; Thuiller *et al.*, 2006; Levinsky *et al.*, 2007; Schloss *et al.*, 2012; Travis *et al.*, 2013). However, comparing spread rates to climate change velocity is a straightforward and simple metric, which allows one to start focussing on those taxa most at risk (Nathan *et al.*, 2011; Bullock, 2012).

#### *Stochastic and deterministic models of spread rate*

In agreement with Travis *et al.* (2011), the two approaches to modelling population spread yielded concordant results, although the analytical model consistently predict higher rates of spread. Furthermore, the relative contributions of individual variables to projected spread rates were similarly described by

the two approaches. Analytical models are of great use for predicting spread rate for a large number of species, which can become computationally demanding if using simulations. Simulations, employing either IBMs or numeric realisations of IDEs, are useful when modelling population spread in real and complex landscapes and sufficient information is available to parameterise them. Here, we used simplified individual-based simulations to allow the comparison with the analytical IDEs and to use the limited set of variables for which sufficient ecological data were available. Such simple models can also be good approximations of spread across moderately varying landscapes (Dewhurst & Lutscher, 2009; Gilbert *et al.*, 2014a) and can be modified straightforwardly to represent more complex variation (Gilbert *et al.*, 2014b). It is clear that dispersal and demography vary within species due to a range of factors such as sex, genetics or landscape context (e.g. Bocedi *et al.*, 2014b), but the full potential of realistic simulations for risk estimation is even more hampered by a lack of data quantifying such variation (e.g. movement rules; Palmer *et al.*, 2011). As general information becomes available on how certain individual or population processes vary according to these intrinsic and extrinsic factors (e.g. butterfly dispersal behaviour; Stevens *et al.*, 2010), it will become possible to make general predictions using more realistic simulations.

A further simplification is that we modelled dispersal distance as a negative exponential function, which may not be representative of real dispersal distributions (Nathan *et al.*, 2012), especially underestimating long distance events. Although there are other, potentially more accurate, dispersal functions (Nathan *et al.*, 2012) and alternative mechanistic approaches (Palmer *et al.*, 2011), the exponential function is commonly used (e.g. Schloss *et al.*, 2012; Santini *et al.*, 2016) as it can be derived from the average dispersal distance, which is often the only metric available.

While these factors preclude highly accurate predictive models, this lack of knowledge applies to all approaches undertaken so far (Thuiller *et al.*, 2006; Schloss *et al.*, 2012; Visconti *et al.*, 2015). The primary purpose of our approach is heuristic, allowing us to disentangle the relative contribution of different life-history traits to species spreading abilities, and thus to make broad-brush predictions about the relative abilities of mammals to keep pace with climate change given their traits.

#### *Benefits of the approach and future directions*

Empirical data describing spread are sparse and rarely comparable among different contexts, and disentangling the contributions of species biology from

those of landscape and other environmental factors is generally difficult, if not impossible. Simulation of virtual species provides a valuable tool for deriving general predictions about ecological processes, assessing the determinants of variation among species in these processes, and projecting risks from environmental change. Previous approaches to simulating virtual species with multiple traits ignore either uncertainty or trait covariation. Many species' traits scale with body mass, and even when the effect of body mass is controlled, other life-history traits are significantly correlated, due to phylogeny, evolutionary strategies and physical constraints (Bielby *et al.*, 2007). If such covariation in traits is ignored, random sampling can lead to virtual assemblages of trait values that are unrealistic, leading to (i) uncertainty about the role of each individual trait in the process investigated, (ii) the creation of artificial trait combinations that are outside those found in nature and (iii) limited applicability of any results with respect to real species. To overcome this problem, allometric relationships might be used to generate idealised species (e.g. Kitzes & Merenlender, 2013), or to select real species representative of target groups (e.g. Schippers *et al.*, 2011), or representative life-history categories (e.g. Coutts *et al.*, 2011). However, these approaches ignore real biological variability and uncertainty around life-history trait relationships and so constrain our ability to investigate the full range of biological possibilities, and hitherto have limited our ability to provide reliable analyses and modelling projections.

In this study, we have developed and demonstrated the use of virtual species that represent the trait values and covariations observed in nature, which can provide a deeper understanding of important ecological processes, such as the ability of species to track shifting climate. Our approach is applicable to address many other ecological questions, for which mechanistic models are available, but where data availability hampers our capacity to apply them to real species. This approach allows one to use available information while accounting for the uncertainty due to our limited knowledge of other parameters. Given the diversity of life, the lack of knowledge for most species and the increasing threats to biodiversity, it is important to develop a strong theoretical underpinning that can be generalised at the species level in order to provide guidelines for management in applied ecology and conservation biology.

## Acknowledgements

LS was supported by two STSMs by the COST Action ES1101 "Harmonising Global Biodiversity Modelling" (Harmbio),

supported by COST (European Cooperation in Science and Technology). JMB and SMW were funded by CEH projects NEC05264 and NEC05100. JMJT and SCFP are grateful for the support of the Natural Environment Research Council UK (NE/J008001/1).

LS, JAH and JMJT conceived the original idea. LS, JAH, JMB, TC & JMJT designed the study; LS collected the data; LS and TC performed the statistical analyses; LS conducted the integrodifference modelling assisted by JMB and SMW. LS conducted the individual-based modelling assisted by SCFP. LS led the writing supported by JMJT, JMB, SCFP, SMW, TC, JAH and GB.

## References

- Bateman BL, Murphy HT, Reside AE, Mokany K, VanDerWal J (2013) Appropriateness of full-, partial- and no-dispersal scenarios in climate change impact modelling. *Diversity and Distributions*, **19**, 1224–1234.
- Bielby J, Mace GM, Bininda-Emonds ORP *et al.* (2007) The fast-slow continuum in mammalian life history: an empirical reevaluation. *The American Naturalist*, **169**, 748–757.
- Bocedi G, Palmer SCF, Pe'er G, Heikkinen RK, Matsinos YG, Watts K, Travis MJM (2014a) RangeShifter: a platform for modelling spatial eco-evolutionary dynamics and species' responses to environmental changes. *Methods in Ecology and Evolution*, **5**, 388–396.
- Bocedi G, Zurell D, Reineking B, Travis MJM (2014b) Mechanistic modelling of animal dispersal offers new insights into range expansion dynamics across fragmented landscapes. *Ecography*, **37**, 1240–1253.
- Bullock JM (2012) Plant dispersal and the velocity of climate change. In: *Dispersal Ecology and Evolution* (eds Clobert J, Baguette M, Benton TG, Bullock JM), pp. 366–377. Oxford University Press, Oxford.
- Bullock JM, White SM, Prudhomme C, Tansey C, Perea R, Hooftman DAP (2012) Modelling spread of British wind-dispersed plants under future wind speeds in a changing climate. *Journal of Ecology*, **100**, 104–115.
- Cardillo M, Mace GM, Gittleman JL, Purvis A (2006) Latent extinction risk and the future battlegrounds of mammal conservation. *Proceedings of the National Academy of Sciences USA*, **103**, 4157–4161.
- Caswell H (2001) *Matrix Population Models: Construction, Analysis and Interpretation*, 2nd edn. Sinauer Associates, Sunderland, Massachusetts.
- Coutts RS, van Klinken RD, Yokomizo H, Buckley YM (2011) What are the key drivers of spread in invasive plants: dispersal, demography or landscape: and how can we use this knowledge to aid management? *Biological Invasions*, **13**, 1649–1661.
- Dewhurst S, Lutscher F (2009) Dispersal in heterogeneous habitats: thresholds, spatial scales, and approximate rates of spread. *Ecology*, **90**, 1338–1345.
- Diaz S, Purvis A, Cornelissen JHC *et al.* (2013) Functional traits, the phylogeny of function, and ecosystem service vulnerability. *Ecology and Evolution*, **3**, 2958–2975.
- Dormann CF, Schymanski SJ, Cabral J *et al.* (2012) Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, **39**, 2119–2131.
- FitzJohn RG, Maddison WP, Otto SP (2009) Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic Biology*, **58**, 595–611.
- Gilbert MA, Gaffney EA, Bullock JM, White SM (2014a) Spreading speeds for plant populations in landscapes with low environmental variation. *Journal of Theoretical Biology*, **363**, 436–452.
- Gilbert MA, White SM, Bullock JM, Gaffney EA (2014b) Spreading speeds for stage structured plant populations in fragmented landscapes. *Journal of Theoretical Biology*, **349**, 135–149.
- González-Suárez M, Revilla E (2013) Variability in life-history and ecological traits is a buffer against extinction in mammals. *Ecology Letters*, **16**, 242–251.
- GRASS Development Team (2012) Geographic Resources Analysis Support System (GRASS) Software, Version 6.4.2.
- Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, **33**, 1–22.
- Hannah L, Midgley G, Hughes G, Bomhard B (2005) The view from the Cape: extinction risk, protected areas, and climate change. *BioScience*, **55**, 231–242.
- Hendriks AJ, Willers BJC, Lenders HJR, Leuven RSEW (2009) Towards a coherent allometric framework for individual home ranges, key population patches and geographic ranges. *Ecography*, **32**, 929–942.

- Huntley B, Barnard P, Altwegg R *et al.* (2010) Beyond bioclimatic envelopes: dynamic species' range and abundance modelling in the context of climatic change. *Ecography*, **33**, 621–626.
- IUCN (2015) The IUCN Red List of Threatened Species. Version 2015-3. Available at: <http://www.iucnredlist.org> (accessed on 9 September 2015).
- Jones KE, Bielby J, Cardillo M *et al.* (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, **90**, 2648.
- Kitzes J, Merenlender A (2013) Extinction risk and tradeoffs in reserve site selection for species of different body sizes. *Conservation Letters*, **6**, 341–349.
- Lenoir J, Svenning JC (2014) Climate-related range shifts – a global multidimensional synthesis and new research directions. *Ecography*, **37**, 15–28.
- Levinsky I, Skov F, Svenning JC, Rahbek C (2007) Potential impacts of climate change on the distributions and diversity patterns of European mammals. *Biodiversity and Conservation*, **16**, 3803–3816.
- Loarie SR, Duffy PB, Hamilton H, Asner GP, Field CB, Ackerly DD (2009) The velocity of climate change. *Nature*, **462**, 1052–1055.
- Lurgi M, Brook BW, Saltré F, Fordham DA (2015) Modelling range dynamics under global change: which framework and why? *Methods in Ecology and Evolution*, **6**, 247–256.
- MATLAB and Statistics Toolbox Release (2012). The MathWorks, Inc., Natick, Massachusetts, United States.
- Nathan R, Horvitz N, He Y, Kuparinen A, Schurr FM, Katul GG (2011) Spread of North American wind-dispersed trees in future environments. *Ecology Letters*, **14**, 211–219.
- Nathan R, Klein E, Robledo-Arnuncio JJ, Revilla E (2012) Dispersal kernels: review. In: *Dispersal Ecology and Evolution* (eds Clobert EDJ, Baguette M, Benton TB, Bullock JM), pp. 187–210. Oxford University Press, Oxford.
- Neubert MG, Caswell H (2000) Demography and dispersal: calculation and sensitivity analysis of invasion speed for structured populations. *Ecology*, **81**, 1613–1628.
- Olson DM, Dinerstein E, Wikramanayake ED *et al.* (2001) Terrestrial ecoregions of the world: a new map of life on earth. *BioScience*, **51**, 933–938.
- Pacifici AM, Foden WB, Visconti P *et al.* (2015) Assessing species vulnerability to climate change. *Nature Climate Change*, **5**, 215–225.
- Palmer SCF, Coulon A, Travis MJM (2011) Introducing a “stochastic movement simulator” for estimating habitat connectivity. *Methods in Ecology and Evolution*, **2**, 258–268.
- Purvis A, Gittleman JL, Cowlshaw G, Mace GM (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society B*, **267**, 1947–1952.
- R Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: <http://www.R-project.org/> (accessed 1 January 2014).
- Santini L, Di Marco M, Visconti P, Daniele B, Luigi B, Rondinini C (2013) Ecological correlates of dispersal distance in terrestrial mammals. *Hystrix, the Italian Journal of Mammalogy*, **24**, 181–186.
- Santini L, Rojas D, Donati G (2015) Evolving through day and night: origin and diversification of activity patterns in modern primates. *Behavioural Ecology*, **26**, 789–796.
- Santini L, Saura S, Rondinini C (2016) Connectivity of the global network of protected areas. *Diversity and Distributions*, **22**, 199–211.
- Schippers P, Verboom J, Vos CC, Jochem R (2011) Metapopulation shift and survival of woodland birds under climate change: will species be able to track? *Ecography*, **34**, 1–11.
- Schloss CA, Nuñez TA, Lawler JJ (2012) Dispersal will limit ability of mammals to track climate change in the Western Hemisphere. *Proceedings of the National Academy of Sciences USA*, **109**, 8606–8611.
- Schurr FM, Págel J, Sarmiento Cabral J *et al.* (2012) How to understand species' niches and range dynamics: a demographic research agenda for biogeography. *Journal of Biogeography*, **39**, 2146–2162.
- Stevens VM, Turlure C, Baguette M (2010) A meta-analysis of dispersal in butterflies. *Biological Reviews*, **85**, 625–642.
- Suding KN, Lavorel S, Chapin FS *et al.* (2008) Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. *Global Change Biology*, **14**, 1125–1140.
- Thomas CD, Cameron A, Green RE *et al.* (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.
- Thuiller W, Broennimann O, Hughes G, Alkemade RM, Midgley GF, Corsi F (2006) Vulnerability of African mammals to anthropogenic climate change under conservative land transformation assumptions. *Global Change Biology*, **12**, 424–440.
- Travis MJM, Harris CM, Park KJ, Bullock JM (2011) Improving prediction and management of range expansions by combining analytical and individual-based modelling approaches. *Methods in Ecology and Evolution*, **2**, 477–488.
- Travis MJM, Delgado M, Bocedi G *et al.* (2013) Dispersal and species' responses to climate change. *Oikos*, **122**, 1532–1540.
- Visconti P, Bakkenes M, Baisero D *et al.* (2015) Projecting global biodiversity indicators under future development scenarios. *Conservation Letters*, **9**, 5–13.
- Whitmee S, Orme CDL (2012) Predicting dispersal distance in mammals: a trait-based approach. *Journal of Animal Ecology*, **82**, 211–221.
- Wilman H, Belmaker J, Simpson J, de la Rosa C, Rivadeneira MM, Jetz W (2014) EltonTraits 1.0: species-level foraging attributes of the world's birds and mammals. *Ecology*, **95**, 2027–2027.

## Supporting Information

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

- Table S1.** Trait data used to simulate mammal species. Source, unit and sample size of the rational for their inclusion are reported.
- Table S2.** MCMCgmm multiresponse model estimates for the means and  $\pm 95\%$  CI of the posterior distribution of the coefficients of life-history traits on the spreading rate as estimated by the analytical model IDE and the IBM RangeShifter.
- Table S3.** Mean of the posterior distribution of the variance–covariance matrix estimate from the MCMCgmm multiresponse model.
- Appendix S1.** Extended methods.
- Appendix S2.** Building population projection matrices for IDE and comparison with matrices in RangeShifter model.