Supporting Information
Earth and field observations underpin metapopulation dynamics in complex landscapes: case study on Carabids

Jonathan Giezendanner[a], Damiano Pasetto[a], Javier Perez-Saez[a], Cristiana Cerrato[b], Ramona Viterbi[b], Silvia Terzago[c], Elisa Palazzi[c], and Andrea Rinaldo[a,d,1]

[a]Laboratory of Ecohydrology, École Polytechnique Fédérale de Lausanne, 1015 Lausanne CH; [b]Gran Paradiso National Park, 10135 Torino, Italy; [c]CNR, Institute of Atmospheric Sciences and Climate (ISAC-CNR), 10133 Torino, Italy; [d]Dipartimento ICEA, Università di Padova, 35131 Padova, Italy

The Supporting Information (SI) is organized as follows. Section SI 1 provides a small review of carabids in the literature, explaining the chosen modeling constrains and approach. Section SI 2 details the processing of the earth observation (EO) data used in the modeling framework. Section SI 3 shows the details of a data analysis performed both on the in-situ and EO data, and the relations between them. Section SI 4 reports additional modeling details which complement and expand the methods as reported in the main document. Finally, section SI 5 expands the results section of the main document with additional details, larger figures and advanced interpretation clues.

SI 1. Carabids in the literature

A. Study reasons and interests. Carabids have been, and continue to be, widely studied, and continue to be, for two principal reasons. 1) In agriculture they are often seen as useful in pest controls (1, 2). Natural enemies of pests, they are cost effective, sustainable and ecological. 2) They serve as good bioindicators (2–6). Due to their sensitivity to change in environmental conditions, carabids are used as bioindicators of environmental, ecological or biodiversity status, which rely on species presence, number of individuals, and diversity (number of species)(2). They are commonly seen as sentinels of environmental pollution, soil-nutrient status, habitat classification for nature protection, habitat quality, change, and alteration, and used to assess the state of communities of other species. Certain reservations on their use as bioindicators remain because of their seasonal variation and because standard sampling methods often favor capture of active and/or generalist species (activity-dependent densities) (2, 4).

B. Models. Mathematical models considered in the literature to describe carabid populations are various, but they can basically be divided into two categories: statistical models which aim at predicting the presence or the abundance of a certain species, and dynamical models which aim at reproducing the observed processes. Statistical models include index based models (e.g. mean individual biomass) (7), scaling laws (8), linear regressions (9) and habitat suitability models (or species distribution models) (10). Dynamical models range from individual based models (5) to patch occupancy models (11), over probabilistic and spatially-implicit dynamical models (12), networks (13) and demographic (14–16), or their suitable combinations thereof, often called hierarchical models. These models can be either spatially-implicit or explicit and static or dynamic in time.

The most frequently used models are the demographic models, coupled with landscape models (5, 15–17). These models explicitly consider the age structure of populations of carabids (according to the life cycle, see below) and their internal dynamics (14), and use a landscape model to link them together through, for instance dispersal (5). The top-level landscape model can in some cases be a classical metapopulation model (11) with typical patch dynamics.

C. Life cycle. Carabids have four developement stages (18): egg, larvae, pupa and finally adult beetle. Depending on the speed of larval development, carabids can hibernate as larvae or adults (6), with reproduction occurring in spring or early summer (3, 6, 18). They usually develop from egg to adult in less than a year, but for certain species this can last up to four years (3). After reproduction, they then normally perish, but certain species can last longer (3, 18). The activity period in the year for mountain carabids is usually of 3-4 months, distributed from spring to fall depending on the species (6). The reproduction rate is strongly affected by temperature and external factors such as food availability. Depending on carabid species, females lay a dozen eggs if they guard them, and up to a couple of hundreds if they do not guard them.

D. Pitfall trapping, data gathering and sampling. For gathering data on carabids, pitfall trapping has been one of the most commonly used field method, both for their ease to use and cost effectiveness (2, 19). In literature, pitfall trapping has been criticized at times, because it poses certain issues regarding: the scattered distribution of traps versus the heterogenous spatial distribution of a species, as well as the passivity of the trap versus the activity of the species. Traps usually fail to capture small and scattered species (19), which influence the sampling and the interpretation of the results. Additionally, spatially and temporally scattered sampling fails to capture many species (20). Traps depend on individuals to make their way toward them, and so the sampling is heavily biased by the relative activity of a species, coupled with its actual abundance (2, 19–22). Even species with high abundance might not be sampled accurately if mostly inactive (20). Generally, pitfall trapping seems to

over-represent generalist species compared to specialized species. It has therefore been argued that instead of talking about abundance resulting from pitfall traps, one should call them "activity-abundance" (20, 21), or "activity-density" (22).

Nevertheless, pitfall trapping can be very instructive. Year-long data can be used to accurately estimate yearly density, and between year fluctuations are well captured (3, 20, 23). Data from spatially-scattered subpopulations can be assembled to gain an overview of the metapopulation structure in the landscape (4).

E. Population dynamics and ecology. **Metapopulation dynamics** Most carabid species function as metapopulations (3, 6), operating at two scales (24), the local scale composed of subpopulations, driven by habitat quality, and the landscape or large scale, at which exchanges between subpopulations occur (19, 25), maintaining declining, isolated subpopulations through dispersal and colonization processes, and even colonize new favorable habitats. Without these landscape-wide metapopulation processes, certain species of carabids would go extinct (13, 26). At the landscape scale, fragmentation and connectivity play a major role in the survival of a species as metapopulation.

Local scale - habitat quality and ecology At the local scale, biotic and abiotic factors influence habitat quality and carabids abundance (18). Carabids can be divided into three groups of tolerance to habitat quality (2, 27): geographically wide-ranged species present in the whole landscape (ubiquitous), species adapting to a wide range of habitats (eurytopic), and species that prove sensitive to a particular habitat quality (stenotopic). Different factors influencing habitat quality are mentioned in the literature: temperature (2, 3, 6, 21, 23, 28, 29); humidity (2, 3, 6, 23); light (23); presence of vegetation (2, 29); food conditions (2, 3); presence of competitors (2, 3). Among these factors, temperature emerges as a the key factor, along with humidity and presence of vegetation (and height of the herbaceous layer), and to a lesser extent, light. Activity, in the sense of mobility, seems to be most influenced by rainfall, temperature and seasonality (6, 22, 28).

Landscape scale - fragmentation and connectivity At the landscape scale, or large scale, there seems to be a general agreement that fragmentation, the "*partitioning of a continuous habitat into many small remnants*" (4), and landscape connectivity strongly influence the presence of carabids (2, 4, 5, 13, 19, 21, 22, 24, 25, 30), as well as the maintenance of populations and subpopulations in heterogenous landscapes, although some argue that fragmentation does not seem to have an effect on carabids diversity (see e.g. (31)), and limited to no effect on certain species (e.g. (4)). Fragmentation, connectivity and habitat availability operate at the landscape level (24). Subpopulations further away from sources of colonizers may struggle to maintain their population (30), such that the distribution of individuals in the landscape directly reflects the history (although lagged, i.e. extinction debt, adaptation lag (13, 30)), and the current state of the network (connectivity to suitable habitat) (2, 21, 30). Species organized as metapopulations, i.e. composed of subpopulations, and connected by dispersal, seem to have a better chance of maintaining their species in the landscape (4). These species, of course, take advantage of the landscape connectivity.

Landscape scale - dispersal In the literature, carabid dispersal, defined as the "*undirected movement away from the habitat of origin*", emerges as the key-component for maintenance of a population on the landscape level (3–5, 23, 25, 26, 32). Without dispersal, the survival of these species would not be possible. The exchange of individuals, even doomed to go extinct in unsuitable habitat, is the main mechanism of the persistence of a metapopulation type of species in a landscape. Dispersal has two primary functions: to re-populate habitat with low presence, and to colonize empty habitats. The two mechanisms jointly ensure that a turnover of extinction and recolonization exists which guarantees the survival of the species in a heterogeneous landscape (23, 26). For carabids, dispersal is not only a mean to escape unsuitable habitat, as individuals are also observed to explore areas away from suitable habitat, even entering sites endowed with unfavorable conditions. In certain species, females even carry eggs on their back while dispersing (3). The dispersal mechanism, driven by connectivity (4, 25, 26, 32), is what maintains the population.

Dispersal occurs in two different types of movements: directed and undirected (or random) walk (3, 23). Directed walk occurs when an individual finds itself in the middle of unfavorable habitat, moving in a straight line as a strategy to move away as fast as possible from the unsuitable conditions until favorable habitat is found, where undirected walk is then again undergone. Undirected walk is mainly performed to find food in favorable habitat. Much greater distances are covered in directed walk. Carabids are believed to be able to cross long distances. Field studies report average movements between 2.5 and 10 meters per day (32), while others go as far as up to 50 meters per day depending on species and individuals (23), mostly driven by hunger which drives higher mobility (directed walk).

SI 2. Earth Observation Data

EO data derived from Landsat have an original resolution of 30 m, while those derived from the digital elevation model have a resolution of 90 m. These raster data have been upscaled to 180 meters resolution in order to match the dimension of the plots used for the in-situ observations. Data of each raster have been linearly scaled between -1 and 1 by selecting the minimum and maximum value over the considered period of time (2006-2013). This operation facilitates the interpretation of the parameters and improves the calibration process. This section provides more details about the data acquisition, source and, where needed, processing.

A. Temperature data. The temperature dataset originates from the E-OBS gridded dataset (33), which has been downscaled to a 90 meter resolution using stochastic downscaling (34).

The temperature data originates from the E-OBS gridded dataset, a European land-only daily gridded dataset for precipitation and minimum, maximum, and mean surface air temperature at 25 km resolution starting in 1950 and regularly updated. The dataset builds on daily observations at more than 2,000-point locations available through the European Climate Assessment and Data set portal (ECA&D*, (35)).

In mountain areas, meteo-climatic features show strong spatial variability over short distances, especially along elevational gradients. Reconstructing reliable elevation gradients of the important variables based on observations is challenging because weather stations are often sparse and do not offer adequate/homogeneous coverage. In particular, high elevations are often under-sampled or not monitored at all (36).

The rate at which air cools with elevation may vary in the range between -9.8°C / km (i.e., the dry-air adiabatic lapse rate) for unsaturated air and -4°C /km (i.e., the saturated adiabatic lapse rate, (36)) for saturated air. When air is saturated, the cooling rate for a rising air parcel is lower due to the release of latent heat by the condensation process. Average temperature lapse rates show strong variability in relation to the climatic zone, as well as to the season. The highest values are reached in summer, over tropical deserts; the strongest negative rates are reached in winter over Siberia, Canada, and polar regions (37).

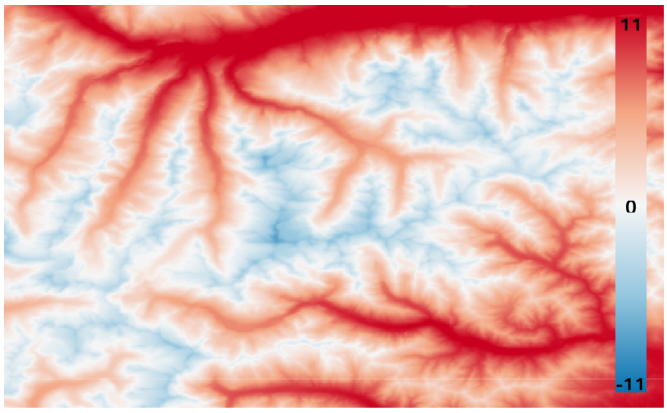

Fig. S1. Median temperature for the year 2006 in the domain of interest.

In order to reproduce the small-scale patterns of temperature, we correct the E-OBS large scale (25 km) fields for the elevation mismatch between their original smooth orography and the elevation in a fine-resolution (90 m) digital elevation model (DEM) by applying a simple correction based on the air temperature lapse rate. The correction has been applied on a monthly basis, employing monthly air temperature lapse rates observed in the Alpine region ((38), see Table 1). The temperature downscaling software has been published here: https://github.com/jhardenberg/RainFARM.jl.

Table 1. Monthly temperature lapse rates for the Alpine region [°C (1000m)$^{-1}$] (based on (38))

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Lapse Rate	4.5	5	5.8	6.2	6.5	6.5	6.5	6.5	6	5.5	5	4.5

For the purpose of this study, the median temperature over the study period (April to October) is then computed for each pixel thus obtaining one temperature map per year for from 2006 to 2012 (see Fig. S1 for an example of median temperature for the year 2006).

B. Tasseled Cap Indices. The aim of the EO data is to accurately represent the mean conditions on the ground during the sampling period, from April to October. For each tasseled cap product and for each study year, a composite image is crated covering the region of interest and the time period.

The tasseled cap map is derived from Landsat 7 images, the only high-resolution satellite covering the whole study period from 2006 to 2013 (39). Only images covering the entire study region where considered, which results in 5 - 9 images per year. Landsat 7 has one major flaw, in that the Scan Line Corrector (SLC) failed on May 31 2003. All the satellite images taken after this period therefore contain gaps which need to be filled or compensated.

We focus on generating one composite image per year. The composite image requires a distribution of values of each pixel in time, which requires one to fill the gaps for the years when gaps are found by computing the focal median in a given neighborhood for each missing pixels. This method is applied to all images covering the region of interest in the period (see Fig. S2 for an example of gap filling using this technique). Once the time-series of images is gap-filled, a composite image which accounts for cloud cover is created through Google Earth Engine's `simple-composite` algorithm†. This algorithm computes

* http://eca.knmi.nl/

† https://developers.google.com/earth-engine/landsat#simple-composite

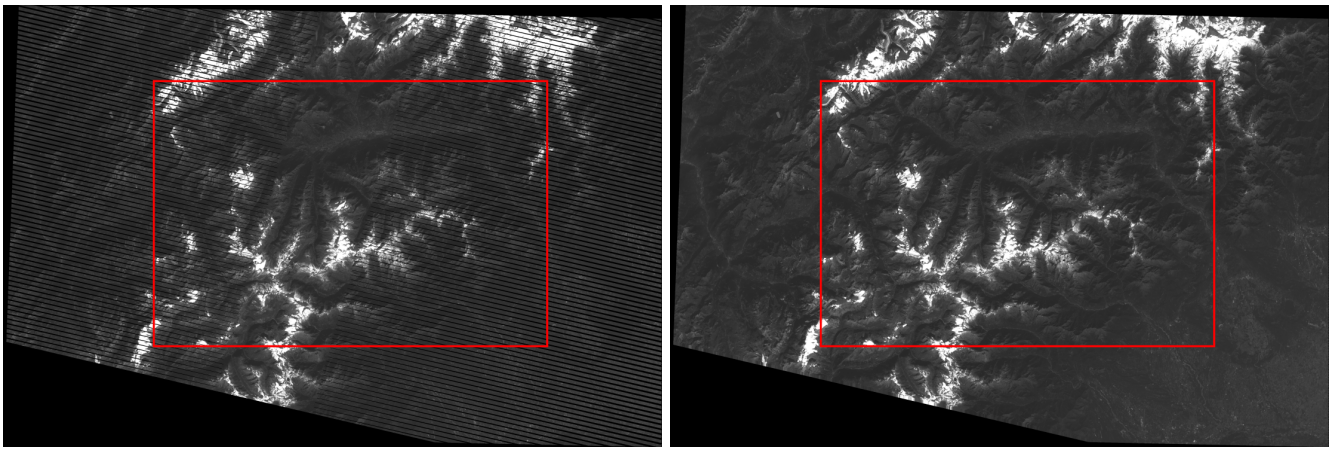

Fig. S2. Example of gap filling. Left: original Landsat 7 image acquired on October 10 2006, with gaps left from SLC failure. Right: same image after gap filling.

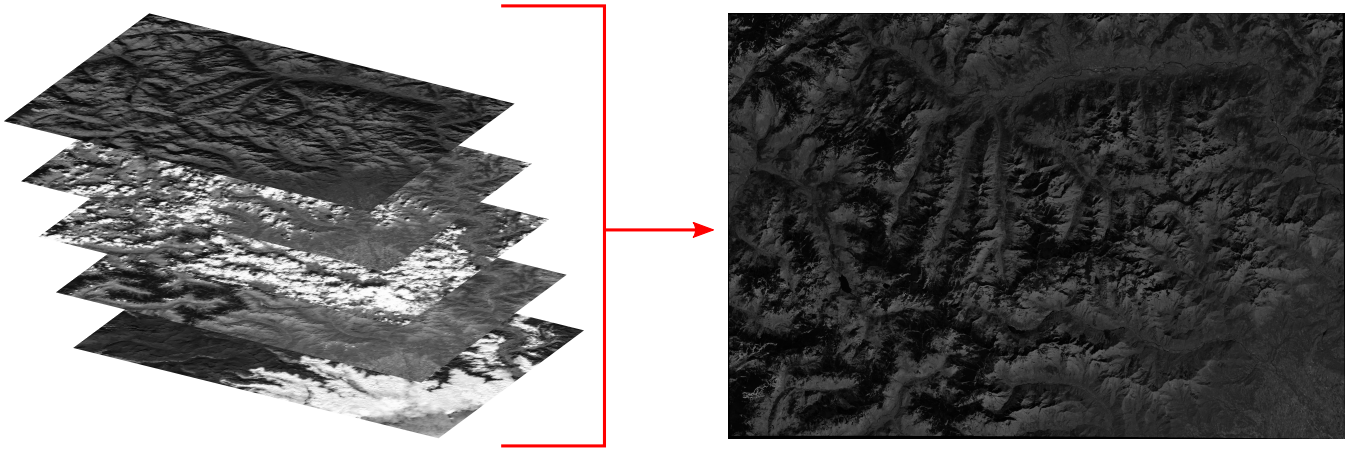

Fig. S3. Example of simple composite. From a series of images, the algorithm composes a median image taking into account the cloud score. The example shows the 'red' bands of a series of images of 2006 on the left, and the resulting composite 'red' band on the right.

a cloud score for each pixel, which is then used as weight when combining the images (the higher the cloud score, the less important the contribution of that pixel). The resulting image contains the weighted median value of the time series for each pixel and each band (see Fig. S3).

Finally, the three tasseled cap transformations are computed from the composite image.

Table 2. Coefficients for the linear combination of bands for computing the different tasseled cap indexes (based on (40)).

Index	Red	Green	Blue	NIR	SWIR1	SWIR2
Brightness	0.3561	0.3972	0.3904	0.6966	0.2286	0.1596
Greenness	-0.3344	-0.3544	-0.4556	0.6966	-0.0242	-0.2630
Wetness	0.2626	0.2141	0.0926	0.0656	-0.7629	-0.5388

The three indexes, brightness, greenness and wetness, are derived from a linear combination of the bands (table 2) contained in the Landsat 7 composite images (see Fig. S4).

This same process is applied for each year in order to obtain one image describing the median conditions in each pixel for each year.

C. Topographic data. All topographic data are generated from a digital elevation model (DEM) of 90 m resolution of the GPNP region. The DEM is extracted from the online Earth explorer tool, courtesy of the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (https://Earthexplorer.usgs.gov/). The DEM is used to generate slope and aspect, which is in turn divided in northness and eastness.

The slope is given as the largest difference in elevation from one cell to the eight neighbors. The slope is obtained for each

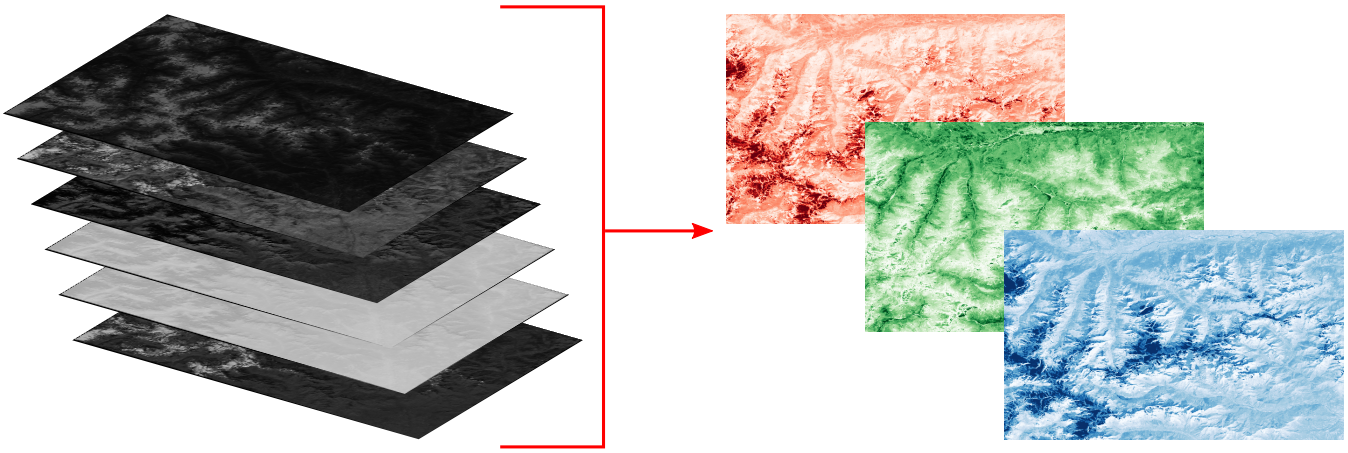

Fig. S4. Example of creation of the three tasseled cap products, brightness, greenness and wetness (right) from the bands of the composite image (left). The example shows the composite image of 2006.

pixel as[‡]:

$$\text{slope} = \arctan\sqrt{\left(\frac{dz}{dx}\right)^2 + \left(\frac{dz}{dy}\right)^2} \quad [1]$$

with $\frac{dz}{dx}$ and $\frac{dz}{dy}$ the gradients to the neighboring pixels in the direction of x (the columns of the raster) and y (the rows of the raster) computed with the moving windows:

$$\frac{dz}{dx} : \frac{1}{8}\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad \frac{dz}{dy} : \frac{1}{8}\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad [2]$$

with the factors of the matrix multiplying the elevation values in the neighboring cells to obtain the difference gradient.

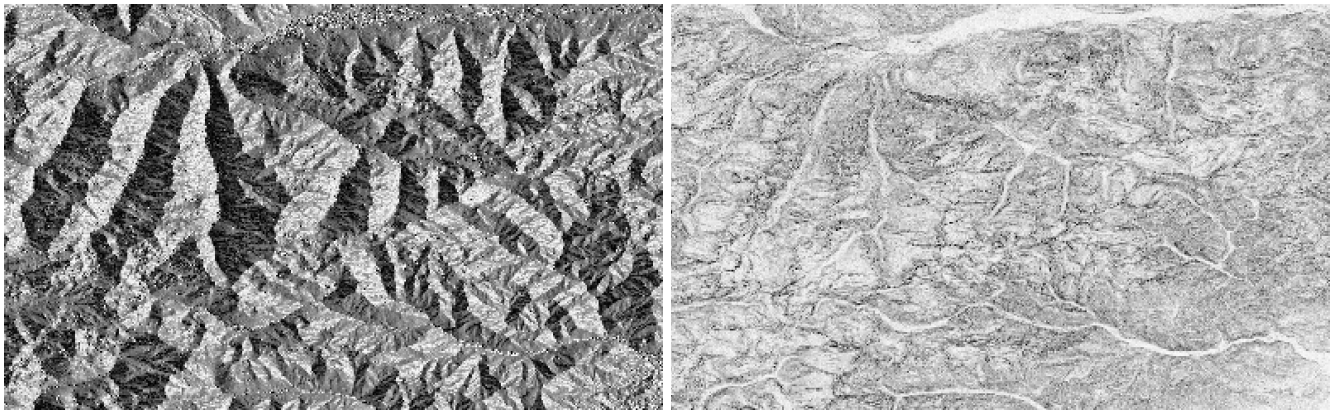

Fig. S5. Map of aspect and slope derived from the digital elevation model.

Aspect is defined as the cardinal direction of the maximum slope from one cell to its neighbors. Similarly to the slope, the aspect is obtained by using a moving window computing a gradient on its neighbors. The angle of the steepest slope is then computed from these informations[§]:

$$\text{aspect} = \arctan2\left(\frac{dz}{dy}, -\frac{dz}{dx}\right) \quad [3]$$

Aspect is then translated in easteness and northness, which are computed by applying the cosine and the sine respectively. This transformation ensures continues values bounded between -1 and 1.

D. Forest Cover. The forest cover in each pixel is here assigned to the final product estimated by the European Commission for the year 2006: https://forest.jrc.ec.europa.eu/en/past-activities/forest-mapping/. The classification proposes to fuse data from different sources (Land Cover Data from CORINE, EO data from MODIS) in order to produce maps of forest presence (41).

[‡] http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=How%20Slope%20works

[§] http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=How%20Aspect%20works

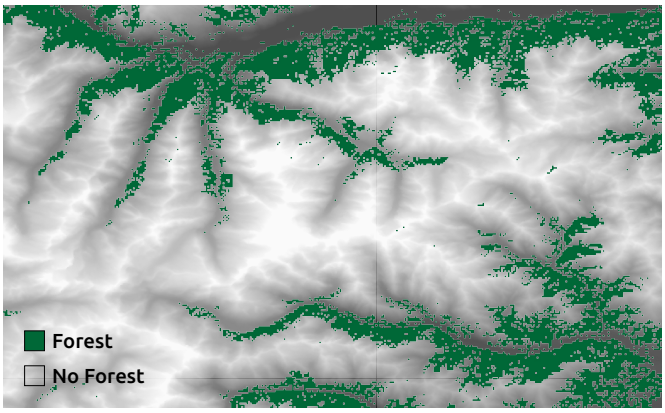

Fig. S6. Map of forest cover for the year 2006 in the domain of interest.

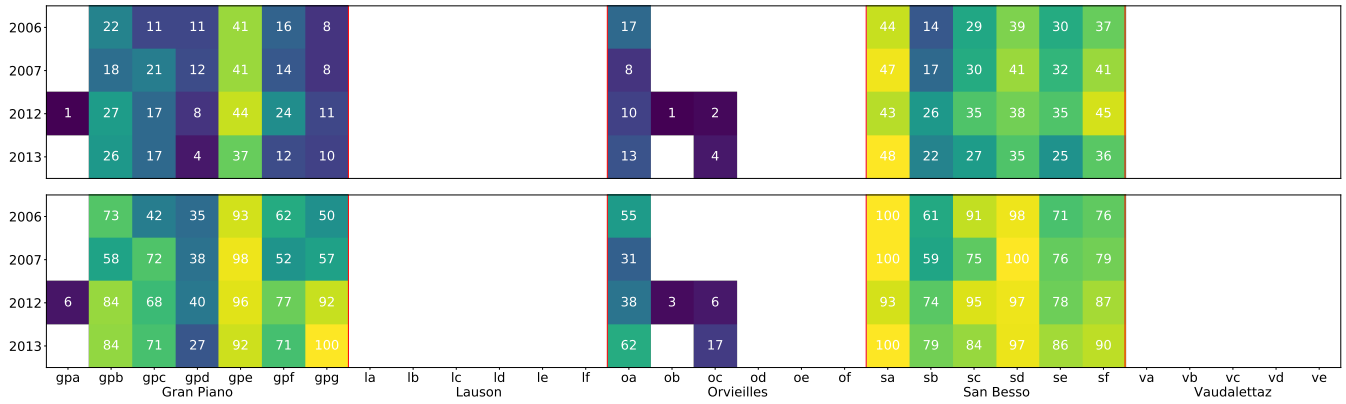

Fig. S7. Number of times *Pterostichus flavofemoratus* was observed (top) and ratio of number of times observed vs number of times sampled (in %)(bottom) in each plot and year. In each valley the plots are ordered from lowest to highest elevation. The blank pixels indicate that the species was not found in this plot/year.

The product is only available for the year 2006, we here assume that the forest cover does not drastically change in the considered time span (8 years) and chosen resolution (180 meters).

SI 3. Data Analysis

Common data analysis operations have been performed in order to gain a basic understanding on the relations between the in-situ data and the environmental drivers, and about the dynamics of the species themselves.

A. In-Situ data. One of the first tasks is to understand the dynamics of the observed species in space and time. In order to do so, we plotted the number of times the species of interest where observed, respectively the ratio between the number of times observed and the number of times sampled in each plot per year (see figures S7 and S8). *Pterostichus flavofemoratus* was selected in order to confront the modeling framework to a dynamically stable species. Figure S7 shows how, albeit not present everywhere, the plots where the species is observed are relatively constant and consistent in presence over time. In the San Besso and Gran Piano valleys, the species is consistently present in all plots and years, with a gradient in presence to be observed in both valleys. In Orvieilles the species is present in the lowest plot and almost always absent in the other plots. If *Pterostichus flavofemoratus* was selected for its stability, *Carabus depressus* was selected because of its variations in time. The species is present in all valleys at times, appearing and disappearing between years. In Lauson for instance, the species has been observed in 2006, disappeared from most plots in 2007, slowly reappeared in 2012 to finally return in 2013. Additionally, the number of times the species was observed per year is constantly lower than *Pterostichus flavofemoratus*. In the case of *Pterostichus flavofemoratus*, the number of times the species was observed is fairly high, and this permits a stronger reliance on the species' actual presence or absence. This is not the case for *Carabus depressus*, where the species was on average observed less than half of the times in each plot. This, of course, does not necessary have anything to do with the accuracy of the sampling, since, as mentioned before, pitfall trapping measures activity-density.

B. EO data. In this modeling work EO data are fundamental to characterize the species habitat and its possible changes during the years of simulation.

Considering that the model is calibrated on the sampled data, it is first interesting to understand whether these plots are representative of the environmental conditions in the modeling regions, i.e. the GPNP. Figure S9 shows the distribution of each

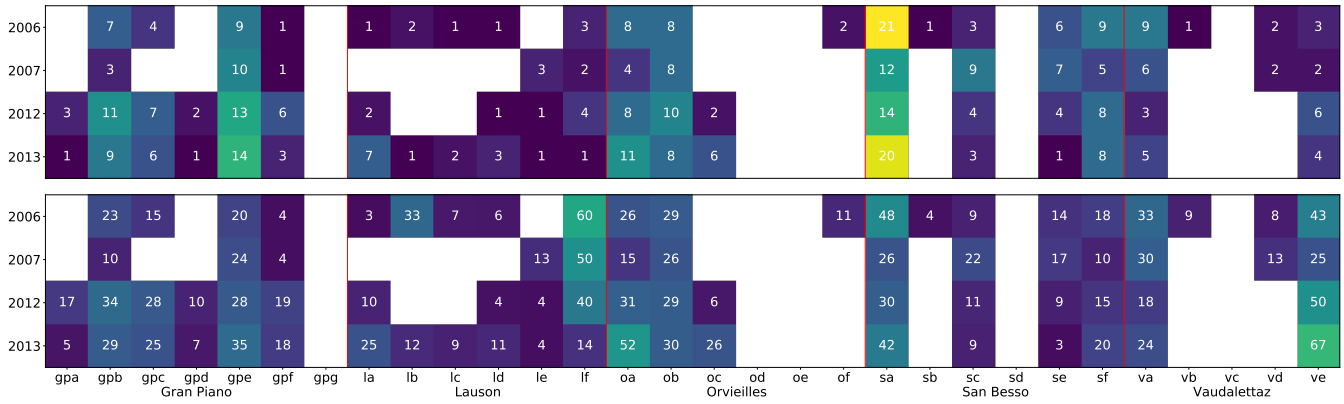

Fig. S8. Number of times *Carabus depressus* was observed (top) and ratio of number of times observed vs number of times sampled (in %)(bottom) in each plot and year. In each valley the plots are ordered from lowest to highest elevation. The blank pixels indicate that the species was not found in this plot/year.

variable and the relation between them, for each pixel value in the GPNP (in blue), and at the sampling points (in orange). The histograms in the diagonal show the orange and blue density plots seem to overlap in most cases, which implies that the data at the sampling points is a good representation of the whole park. In isolated cases, for instance for temperature when paired with most of the variables, the values recorded at the plots are not covering the whole range of values at the GPNP. This is especially evident when temperature is considered in relation with the tasseled cap indexes, where for lower temperatures the sampled values do not seem to be as well represented. A reasonable explanation for this fact is that lower temperatures are mostly found at higher elevations, where the sampling was not performed as few species survive above a certain altitude, especially in areas with year-long snow covers or on glaciers. In terms of correlations, the chosen variables seem to be independent of each other. Even in cases with relatively strong correlation (for example greenness vs temperature), the spread around the trend is very large suggesting that the combination of variables can have more explanatory power than the variable on its own. The relation between brightness and wetness depict a particular behavior, having a positive correlation when considering the whole GPNP, but negative when only selecting the data in the sampled plots. To a lesser extent the inverse can be observed between greenness and brightness. By looking at the density plots of these variables, one might speculate that these behaviors are artifacts induced by outliers, as in the density plots most data seem to follow the same trend. The pairs plot indicates that all chosen variables are needed and that none could be discarded.

The relevance of each variable is further tested by performing a principal component analysis (PCA). PCA is technique widely adopted to understand the degree of correlation in a given data set (42). PCA performs an transformation of the data matrix (i.e. here the EO data) into an orthogonal uncorrelated data set (42, 43). The technique projects the data onto a base of orthogonal axes called principal components. The eigenvalues associated with each axis, or the fraction of the eigenvalue compared to the sum of all eigenvalues, determine the proportion of variance of the data explained by each axis. A large eigenvalue can be interpreted as a large correlation between the data in the original data set. Apart from analyzing eigenvalues, the result of a PCA are usually interpreted by projecting the original data set onto the two first axes of the PCA (see figure S10). The original vectors corresponding to the original data set are then displayed in direction and magnitude relative to the degree to which they contribute to each axis, which permits to identify unnecessary covariates. Figure S10 shows the outcome of one PCA per sampling year for the data in the sampling locations. In each year, the proportion of variance does not exceed 35%, indicating that the various EO axes can not be projected on a single strong component. In 2006 and 2012, the two first components seem to explain most of the data variance compared to the other years, but even in this case most axes are needed. This means that each variable is needed in order to accurately depict the environmental conditions in each plot, and no variable can be removed. The composition of the principal component (loadings plot) shows that most variables contribute roughly equally, with variations in time: for example in 2012 where tasseled cap data and northness explain most of the first component's variance.

From the projection of the variables on the two first axes (third line), a clear difference can be seen regarding the composition of the different valleys. Especially Gran Piano and Vaudalettaz seem to differentiate themselves, with the others closer together. In all years, the variable axes are well distributed and rarely overlapping, which again confirms the importance of keeping all variables to characterize the habitat. Another important information is the trend and distribution of the variables in time. Over the GPNP domain the trends for the time-varying variables (temperature, wetness, brightness and greenness, figure S11) appear relatively stable. With a slope of 1.01, greenness is the variable with the strongest increasing trend, followed by brightness with 0.64. The trajectories of the plots (in grey in figure S11) show that temperature homogeneously follows the median, with most of the lines above the median. These homogeneous lines can be explained by the nature of the temperature data, stemming from a downscaled model relying on elevation. The plot lines for the other variables are more heterogeneous. Wetness and brightness in the plots lie below the median, and greenness above. All of them follow the trend set by the slope of the distribution, meaning the trends observed in the plots are representative of the trends observed in the landscape.

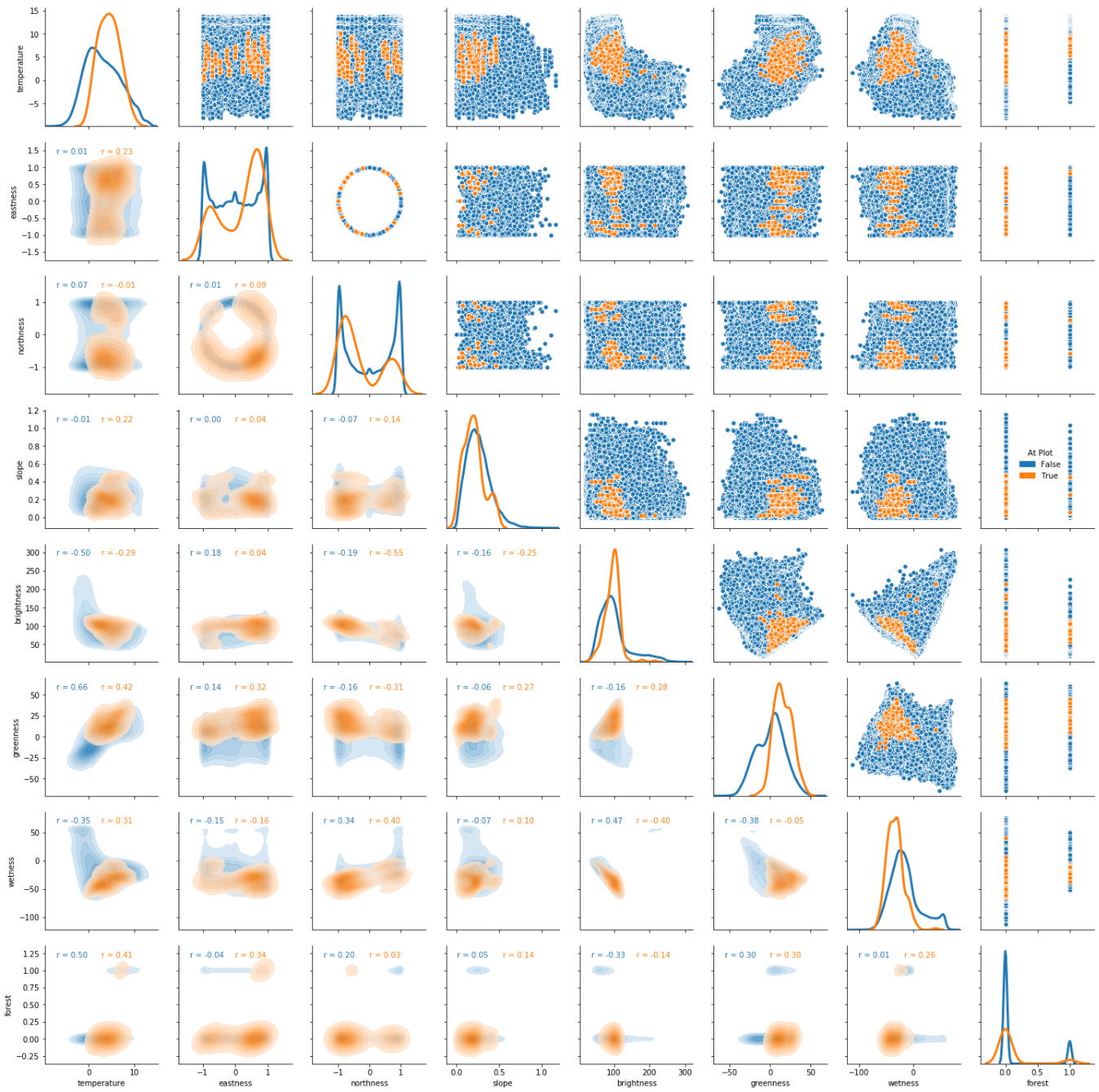

Fig. S9. Pairs plot of the EO data in the sampling plots (in orange), and in the whole GPNP (in blue). The diagonal shows the distribution of the variable. The upper grid shows the scatter plot of each variable plotted against all others. This permits to display the combined distribution of the variables, one pair at a time. The lower grid shows the density plots derived from the scatter plots, along the computed correlation coefficient. Note that there are two correlation coefficients, one for the data at the plot positions, and one in the whole GPNP.

C. Combined in-situ and EO data analysis. To model the two species of interest, the relation between the presence of the species in the landscape and the descriptors, here the EO data, has to be understood. Figures S12 and S13 show the same PCA as in figure S10, but grouping the plots recording species presence (blue) or absence (orange) *Pterostichus flavofemoratus* and *Carabus depressus*, respectively. The PCA in figure S12 shows that the covariates can not entirely differentiate the plots with presence form the plots with absence, which was to be expected as two first axes only explain around 60% of the variance each year. Still, certain plots are not overlapping, especially in 2012. For these plots, the prediction of presence or absence might be easier solely based on fitting the EO data to the in-situ data. The PCA in figure S13 shows a relatively strong overlapping of the presence and absence group. Given the relatively similar EO conditions for present and absent plots, spatial dynamics have to be considered in order to explain the spatial distribution of the species, solely relying on EO is not enough.

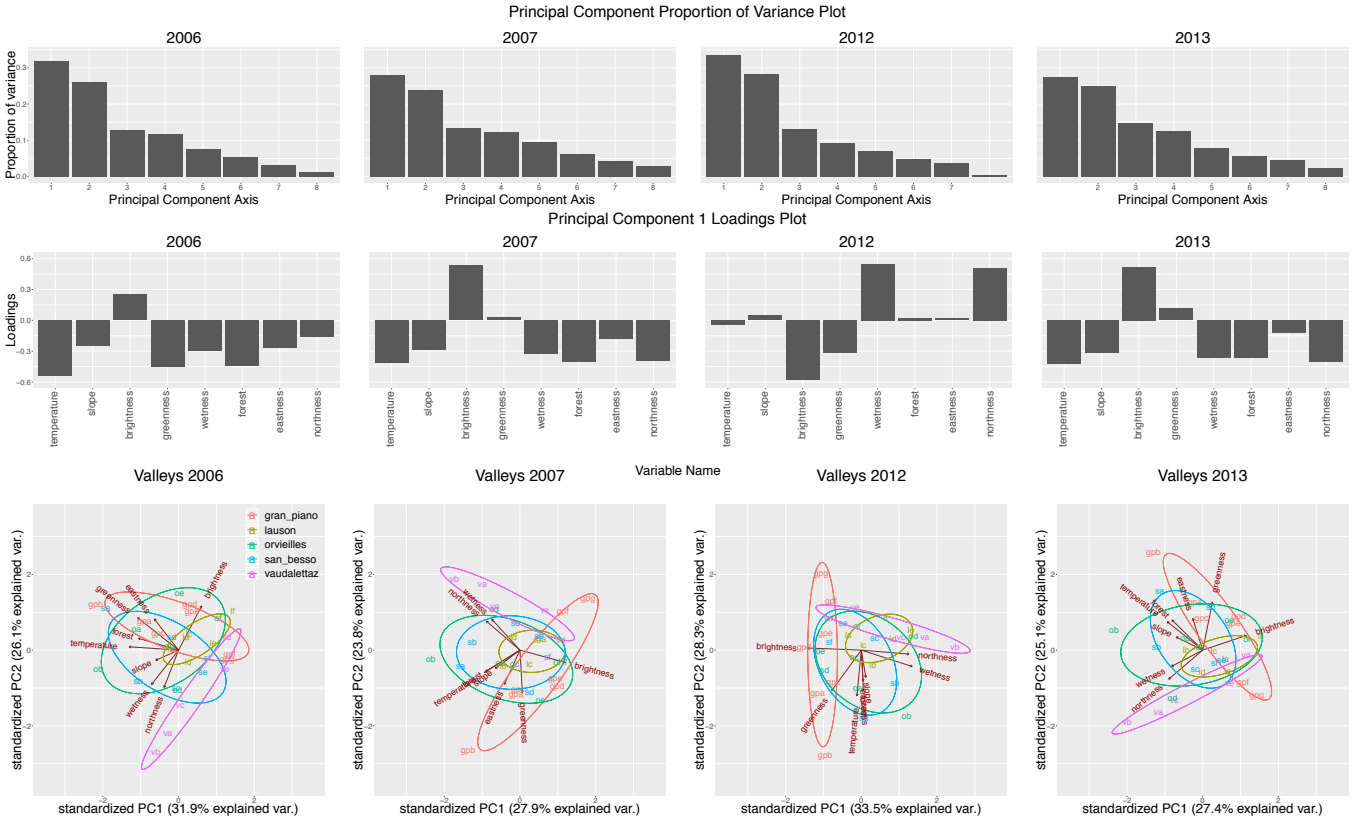

Fig. S10. Proportion of variance explained by each principal component for each sampled year (top). Loadings plot showing the contribution of each variable to the first principal component for each sampled year (middle). Projection of the EO data on the two first principal component, and grouping of the data by valley (bottom).

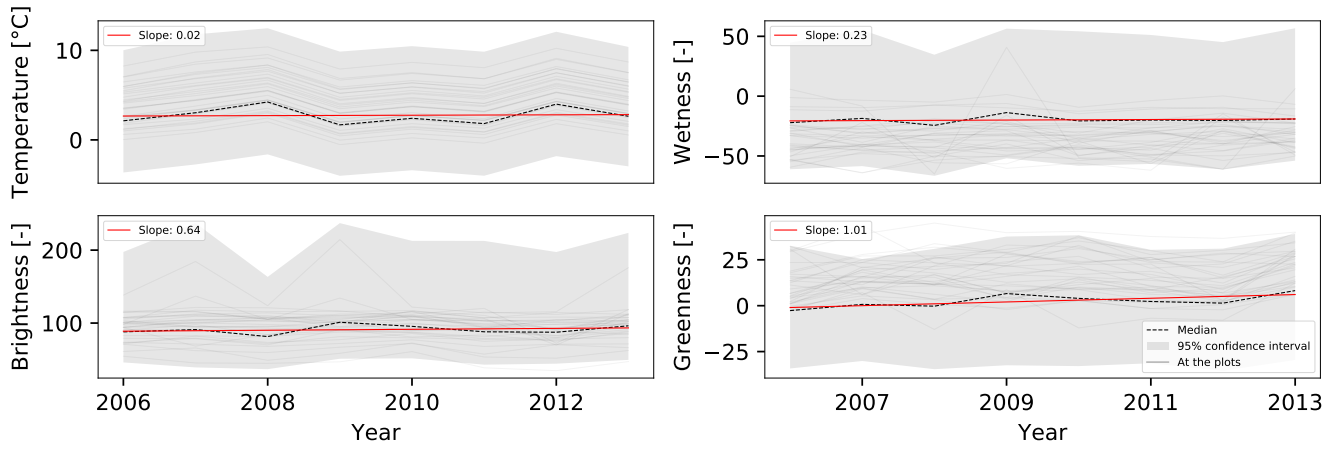

Fig. S11. Median and 95% confidence interval of the time-varying variables, as well as the slope of the linear regression over the 8 years. The light grey lines show the the path for the variables at the plot locations.

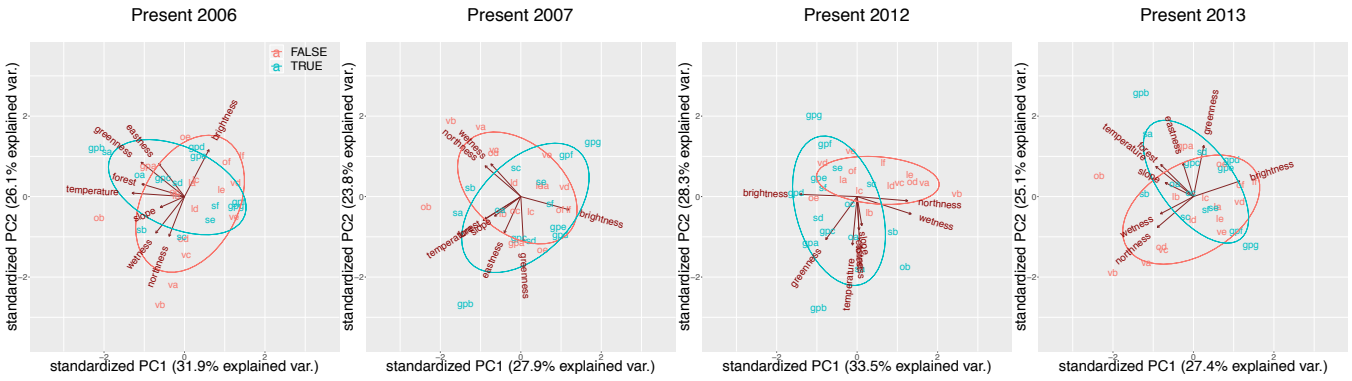

Fig. S12. Projection of the EO data on the two first principal component, and grouping of the data by presence/absence of the species *Pterostichus flavofemoratus*.

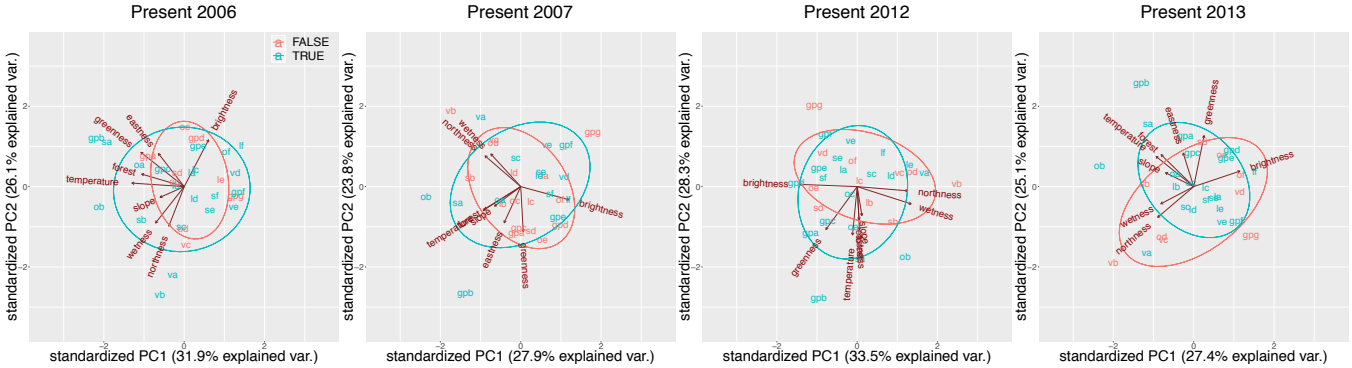

Fig. S13. Projection of the EO data on the two first principal component, and grouping of the data by presence/absence of the species *Carabus depressus*.

SI 4. Model details

A. Metapopulation model. At every simulation time t, SPOM simulates a possible distribution of occupied cells in a domain having N cells by computing the number of cells where the species went extinct during the time-step and the newly occupied cells due to colonization. Extinction and colonization rates depend on the species traits and on the landscape features. A binary state variable $w_i(t)$ is set to 1 when the cell i is occupied and 0 when empty ($i = 1, \ldots, N$). Starting from a given initial distribution of occupied cells, at each time step, the model allows unoccupied cells to be colonized by surrounding occupied cells and species in occupied cells can go extinct.

B. Initial Presence. We here set the hypothesis that the absence of a species from a valley is explained by the species not having yet colonized that valley rather than by the differences in niche definition, given the similar EO data in the different sampling sites (see SI 3, PCA). The pixels in the valleys where no individuals were observed during the whole year are thus set to have probability zero. Fig. S15 shows the initial presence after calibration.

C. Spin-up and equilibrium assumption. Starting from a certain spatial configuration, a dynamic model such as SPOM is almost always far from equilibrium, and needs to be relaxed to reach a quasi-equilibrium. In SPOM, this is the case while more extinction than colonization events occur, or vice-versa (44). From this initial configuration, it is possible to tend to a pseudo-steady state, a situation where the occupancy oscillates around a stable state (45). This phase, which requires to run a certain number of Markov steps (see main document (MD), equation 1), is called the spin-up period. The model is simply run without changing the parameters nor the cofactors until the simulations stabilize. This can take a certain time depending on the parameter values, especially the values of c and e. When keeping the ratio between c and e constant, the larger their values, the shorter the spin-up period. While small values of c and e will require long spin-up periods, large values while tend faster to this equilibrium.

Note that incorporating this spin-up period in the model is not equivalent to making the equilibrium assumption (44), but merely a necessary step in order to be certain that the observed trends don't stem from a modeling bias, but rather from the simulation stochasticity and changes in explanatory variables.

Here, starting from an initial distribution, we run the spin-up at each iteration of the calibration in order to make sure the model oscillates around the quasi-equilibrium state.

D. IF2 applied to SPOM (details). This section gives further details to the main method section, with most steps refering to the algorithm at the end of the section and figure S14.

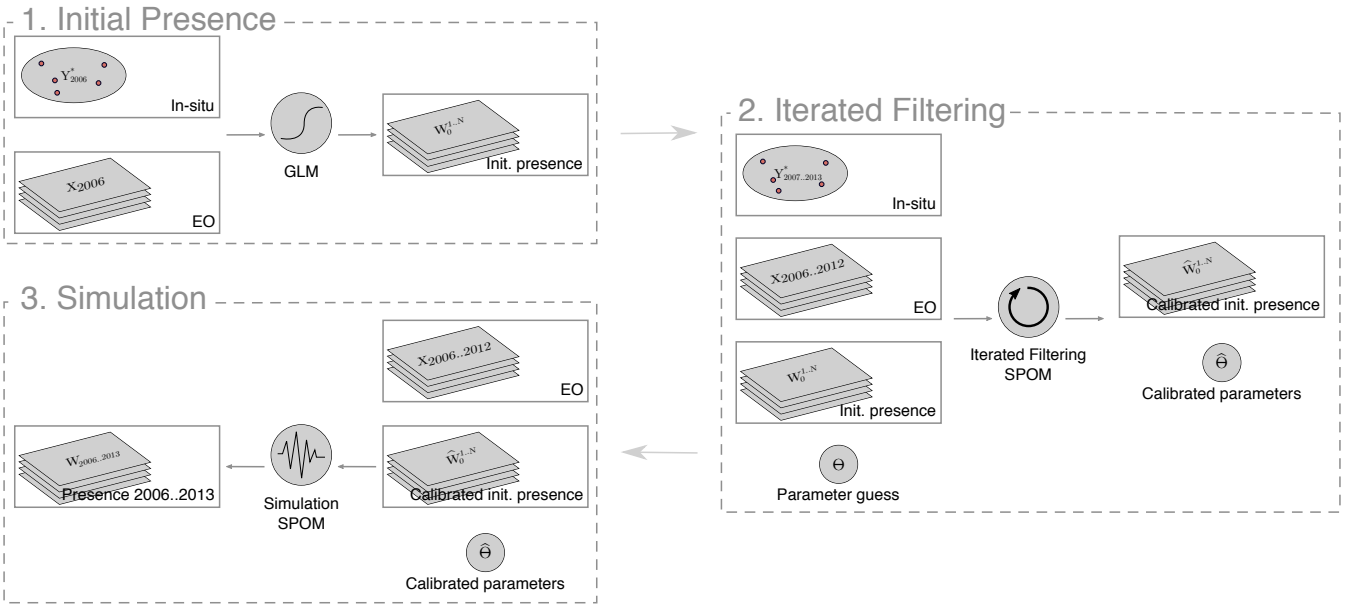

Fig. S14. Overview of the proposed framework as detailed in the methods: 1) The species' initial spatial occupancy is estimated by a generalized linear model (GLM) based on earth observation data (EO) and calibrated with in-situ data of the first sampled year (2006 in this case); 2) The metapopulation model (SPOM) parameters and the initial presence are calibrated by iterated filtering (IF), with the EO data of each year used for the species' fitness, and the in-situ data for the likelihood. 3) The calibrated initial presence and parameters are used to run the metapopulation model as a simulation in order to predict the focus species' presence in the landscape.

Each particle corresponds to a set of parameter values and model states in each pixel. The initial parameter values are drawn by means of latin hypercube sampling (LHS, (46)). LHS is a technique designed to randomly sample from a multi-dimensional distribution, but which, unlike the uniform random sampling, takes into consideration the previously sampled values. This techniques improves the sampling efficiency of the solution space. For each initial value drawn via LHS, the calibration is performed.

We here add the initial occupancy of the landscape to the particle definition. At the beginning of each iteration, for each particle the spin-up is run with the initial presence contained in each particle, and the final state is conserved for the next iteration. The spin-up period assures that the observed trends in species presence stem from stochastic fluctuations or environmental forcing, i.e. temporal changes in the cofactors, and not from the model reaching an equilibrium imposed by the parameters (see SI 4.1). Since the parameters change at each step and iteration, it is important to re-run the spin-up at the beginning of the iteration to ensure the pseudo-equilibrium, but with the parameter change being slow over the calibration process, the final state of the previous iteration is already close to the new equilibrium imposed by the sampled parameters, and thus this state is used as initial occupancy for the new parameters and spin-up, which final state then becomes the new initial presence for the particle. To summarize, each particle contains: parameter values $\Theta(D,\, c,\, e,\, \alpha_0,\, \beta_j)$ and an initial presence W_0.

For each species, we first perform a broad calibration process which takes into account the spin-up period's length. Since the spin-up period is dependent on the scale of c and e, and as the two parameters' scale influences the speed of the process after the spin-up, the different species require different spin-up length. The optimal spin-up period is chosen based on the likelihood and the effective sample size which measures the relative importance of each particle. The effective sample size is computed as (47) $\hat{N}_{\text{eff}}^{\text{ens}} = \frac{1}{\sum_{i=1}^{N^{\text{ens}}} (w_i^2)}$, with N^{ens} the number of particles, and w_i the weight of particle i computed from the likelihoods. Small values of effective sample size indicate that one particle is absorbing all others, and the parameter space is not efficiently sampled. Only spin-up periods yielding an acceptable effective sample size are considered.

Once the spin-up length determined, a first LHS is performed on a broad range for each parameter value. The fitness parameters are sampled in the interval $[-4, 4]$, D in $[50, 500]$, c in $[100, 500]$ and e in $[10^{-3}, 10^{-2}]$ (or for c and e a factor of these values depending on the spin-up period). A second LHS is then performed around the parameter values with the best likelihood values from the previous sampling. The final calibrated values are the values generated from IF with the best likelihood. The process is described in further details in SI 4.2.

With the calibrated parameters and initial occupancy, the model is then run without filtering and parameter propagation in order to generate the species distribution. This simulation is run with 480 independent repetitions to obtain occupancy distributions for each pixel.

The calibration is summarized here:

1 The initial spatial distribution of the species in the landscape is determined:
- **a.** The initial probability of presence is computed via GLM based on in-situ and EO data from 2006.
- **b.** For each of the N^{ens} particles of the IF a presence-absence map is drawn from the probability map ($W_0^{1..N}$).

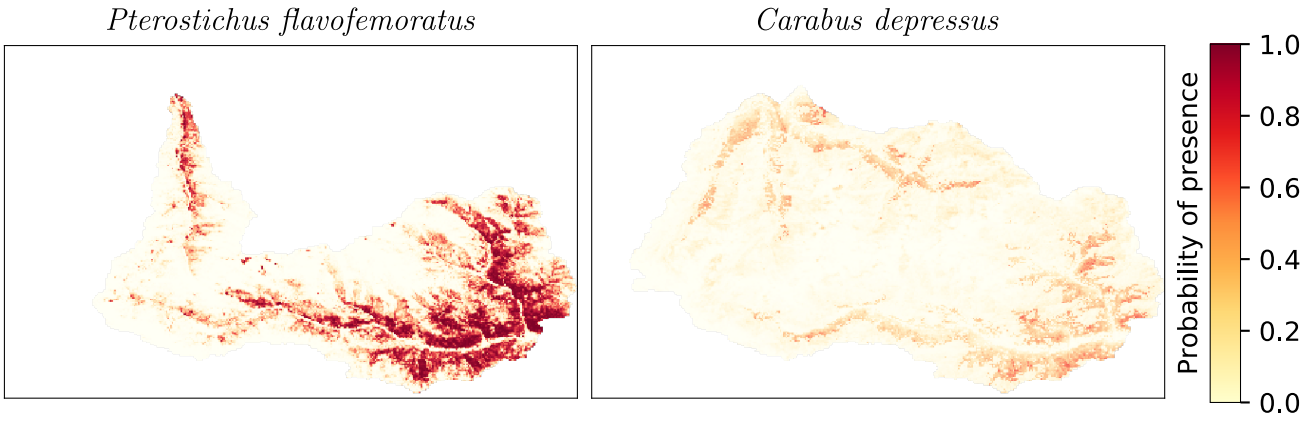

Fig. S15. Initial probability of presence estimated by a species distribution model for two species of interest *Pterostichus flavofemoratus* and *Carabus depressus*.

2 The first phase is designed to determine the spin up period:
- **a.** A number of initial parameter guesses are drawn via LHS (θ_0), including different spin up periods and scales of c and e.
- **b.** For each LHS guess, IF2 is performed.
- **c.** The process with the best likelihood and effective sample size is selected, and determines the spin up period for the species.

3 The second phase is designed to narrow the range of the parameters:
- **a.** Keeping the scale of c and e fixed and fixing the length of the spin up period, a number of new parameter guesses are drawn from LHS with the initial broad range of parameters.
- **b.** For each guess, IF2 is performed.
- **c.** The guess with the best likelihood is kept.

4 The last phase is designed to select the optimal values of the parameters around the already narrowed parameters:
- **a.** A final LHS drawing is performed around the values of the best performing guess.
- **b.** For each guess, IF2 is performed.
- **c.** The final parameter set is taken from the maximum likelihood over all guesses.

The IF2 algorithm applied to SPOM reads as follow (adapted from (48)):

Inputs:
- Number of iterations M
- Number of iterations without propagation M'
- Number of particles N^{ens}
- Initial spatial species presence for each particle $W_0^1, \dots, W_0^{N^{\text{ens}}}$
- LHS parameter guess θ_0
- In-situ data $Y^*_{2006,2007,2012,2013}$, with Y^* containing the observations k and n for the year.
- EO data $X^j_{2006\dots2012}$

For $m = 0$ to M:
Propagate parameters with the density function h and perturbation sequence σ_m: $\Theta_{1\dots N^{\text{ens}}} \sim h_{1\dots N^{\text{ens}}}(\theta|\Theta_{1\dots N}; \sigma_m)$
Compute fitness function (equation 4 (MD))
Perform the spin up with EO data X^j_{2006} to produce presence $W^1_{2006}, \dots, W^{N^{\text{ens}}}_{2006}$
Set new initial presence: $W^j_0, = W^j_{2006}$, $j = 1, \dots N^{\text{ens}}$
Compute likelihood based on in-situ data of 2006: $\mathcal{L}^j(\theta^j|Y^*_{2006}, W^j_t)$, $j = 1, \dots N^{\text{ens}}$ (see equation 8 (MD))
Filter particles based on weight of each particle: $w_j = \frac{\mathcal{L}^j}{\sum_n^{N^{\text{ens}}} \mathcal{L}^n}$ (see algorithm 2 in (47))

For $y = 2006$ to 2012:
Propagate parameters: $\Theta_j \sim h_j(\theta|\Theta_j; \sigma_m)$, $j = 1, \dots N^{\text{ens}}$
Compute fitness function
Perform Markov step with EO data of year y X^j_y to produce presence W^j_{y+1} (see section 1 (MD))

If $(y + 1)$ is in in-situ:
Compute likelihood based on in-situ data of $(y + 1)$: $\mathcal{L}^|(\theta^j|Y^*_{2006}, W^j_t)$
Filter particles based on weights

end if

end for

end for
Select best parameter set based on maximum likelihood, set parameters in all particles to the best particle.
For $m = 0$ to M':
Perform same operations but without parameter propagation, compute average likelihood over the iterations.
end for

Outputs: - Calibrated parameters
- Calibrated initial presence
- Associated average likelihood

Fig. S16 shows an example of 30 iterations of the calibration process described above.

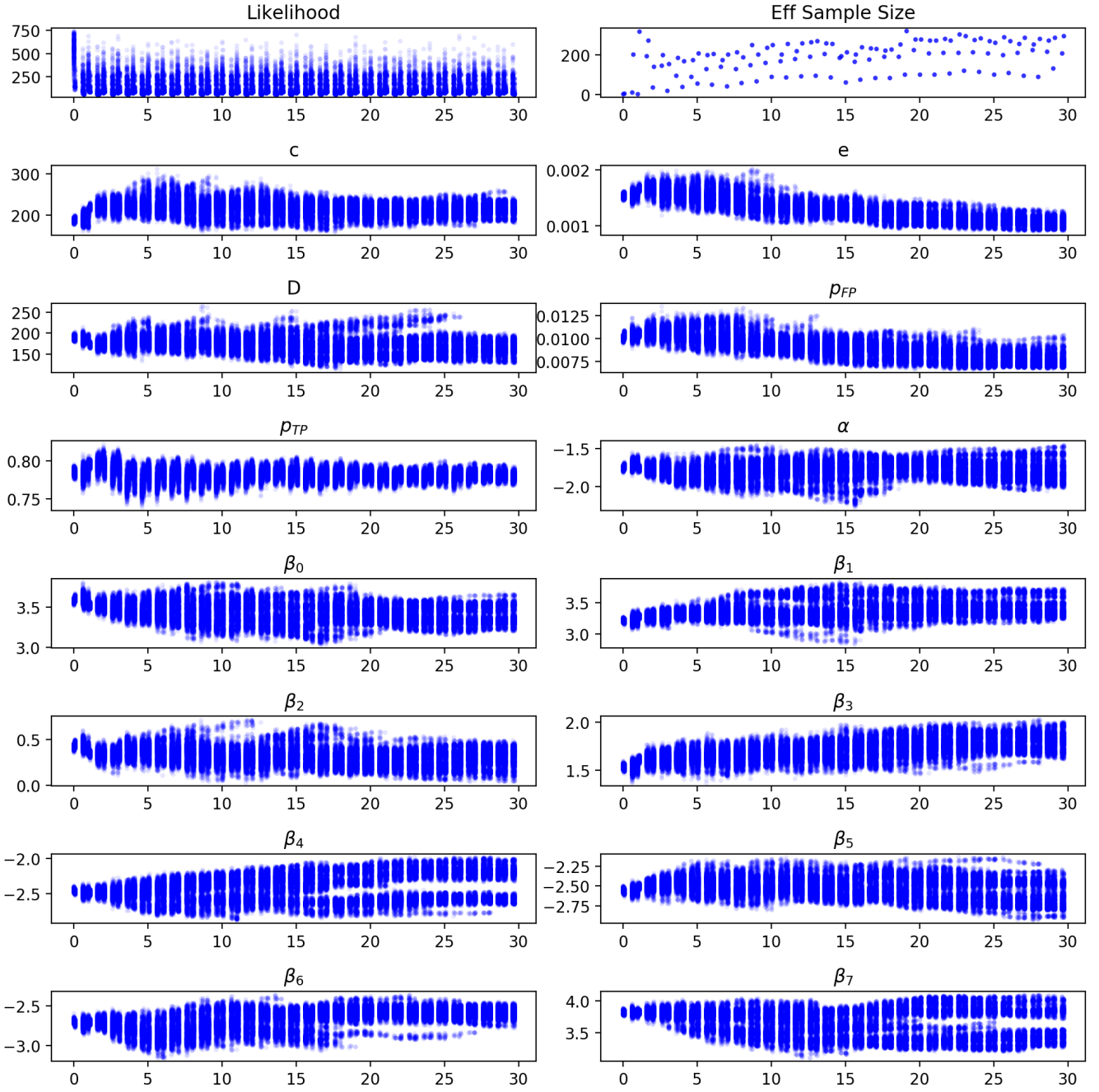

Fig. S16. Example of 30 iterations of IF for the species *Pterostichus flavofemoratus*. This example shows the algorithm after three IF layers. The graph on the top left shows the negative log-likelihood, and on the top right the effective sample size which measures the relative importance of each particle. The effective sample size is computed as (47) $\hat{N}_{\text{eff}}^{\text{ens}} = \frac{1}{\sum_{i=1}^{N^{\text{ens}}} (w_i^2)}$, with N^{ens} the number of particles, and w_i the weight of particle i computed from the likelihoods. To ensure a correct calibration process, it is of advantage to have large effective sample size values, which ensures no particle depletion, a phenomenon where all but a few particles have very small weights and are discarded at each filtering. The remaining plots display the values of the parameters at each step and iteration.

SI 5. Additional results and details

The spin-up periods for *Pterostichus flavofemoratus* and *Carabus depressus* have been set respectively to 400 and 120 steps. The uneven durations are due to the differences between the calibrated colonization and extinction rates of the two species resulting in different ecological timescales.

A. *Pterostichus flavofemoratus*. Figs S17 shows the simulated average occupancy of *Pterostichus flavofemoratus* in the landscape for each year, computed from 480 repetitions. To better appreciate the temporal changes, the relative difference in average occupancy in each cell computed as the realtive difference between two consecutive years is highlighted in Fig. S18.

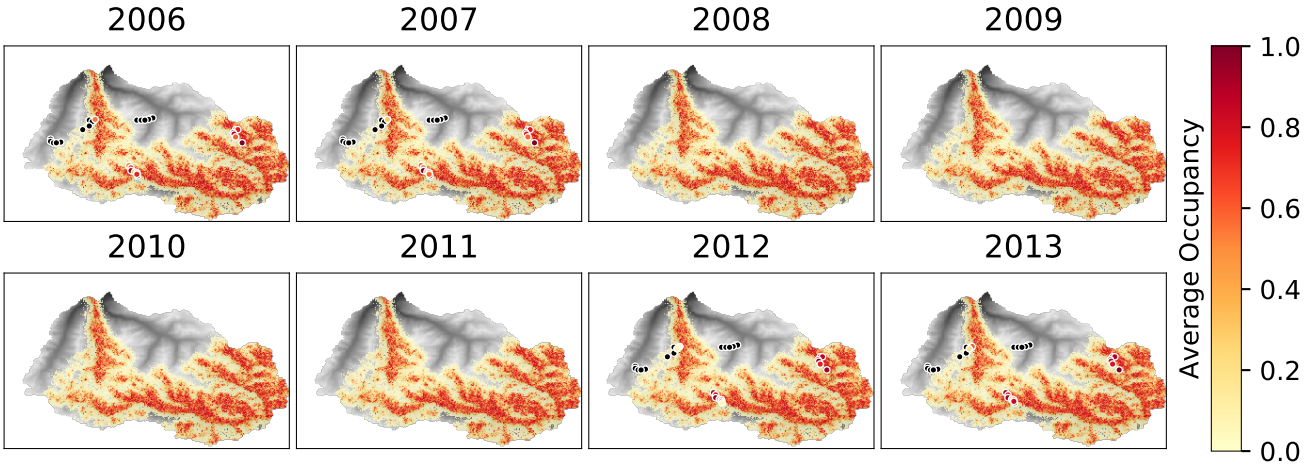

Fig. S17. Output of the filtering algorithm for the calibrated parameters: average occupancy of *Pterostichus flavofemoratus* in time. For the sampling years (2006, 2007, 2012 and 2013) the ratio of number of times a species was observed vs the number of times it was sampled is displayed at the sampling location.

The modeled average occupancy of *Pterostichus flavofemoratus* (Fig. S17) well match the in-situ observed ratio at most of the plots. This fact suggests that the calibration process identified a parameter combination that results in an appropriate fitness function and metapopulation dynamics.

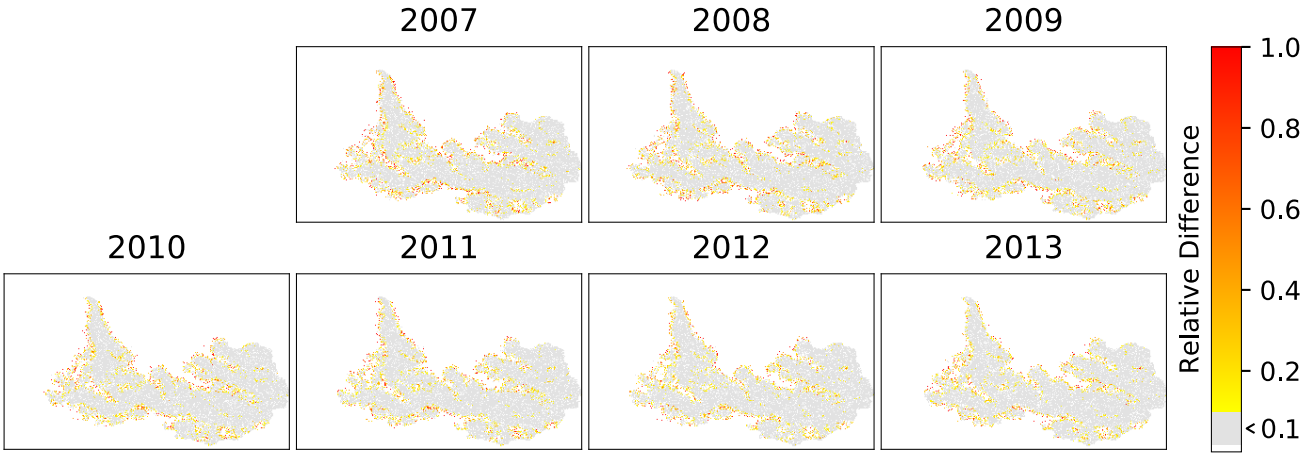

Fig. S18. Relative difference in average occupancy computed from one year to the other for the *Pterostichus flavofemoratus*. The relative difference is computed as $|\hat{x}_y - \hat{x}_{y-1}|/\hat{x}_{y-1}$ which highlights large changes compared to the initial value. The color scale is shown in white where no change occur, grey where less than 10% of difference is observed, and a linear gradient caped at 100% for values above.

The occupancy in the landscape of *Pterostichus flavofemoratus* appears to be relatively stable, although with a slight increase over time (Fig. 3c (MD)). This is evident also from Fig. S18, where the species' core presence does not change in time, and most of the variations in occupancy occur at the edge of this stable core.

To better understand the cause of the temporal changes in presence, Figs S19 and S20 show the fitness for *Pterostichus flavofemoratus* at each year and its temporal variations, which are only due to the changes in the co-factors. The variations in the *Pterostichus flavofemoratus* presence are, to a certain extent, related to the variations in fitness which are higher at the

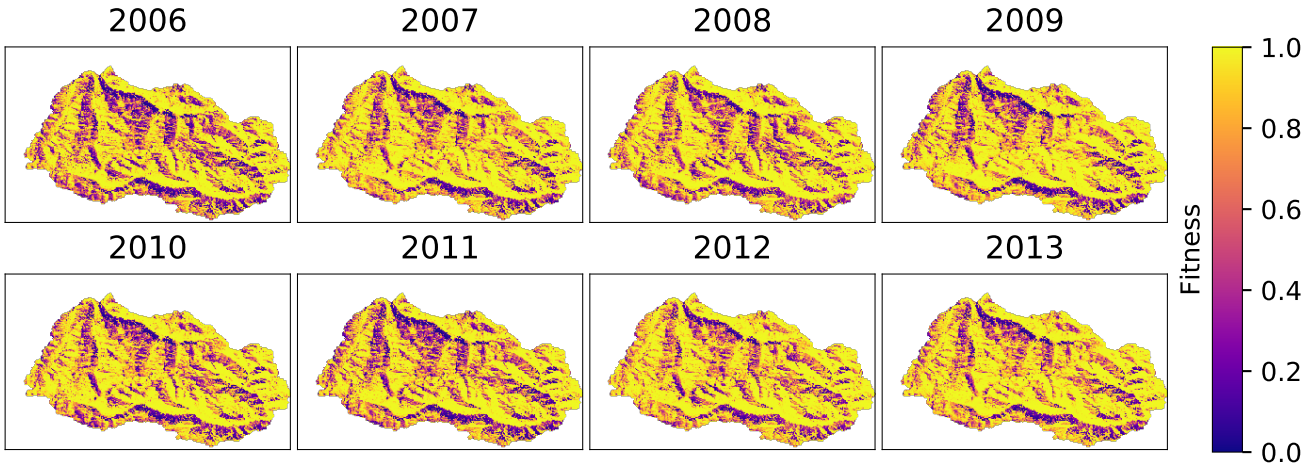

Fig. S19. Fitness of *Pterostichus flavofemoratus* in the GPNP computed from the calibrated parameters (equation 4 (MD), parameters shown in table 1 (MD)) and the EO data for the simulated years.

border of the occupied area (compare Figs S18 and S20). Another important process that generates variations in presence is related to the changes in colonization pressure from surrounding occupied cells, which might lead to extinction and possibly later to recolonisation, i.e. metapopulation dynamics.

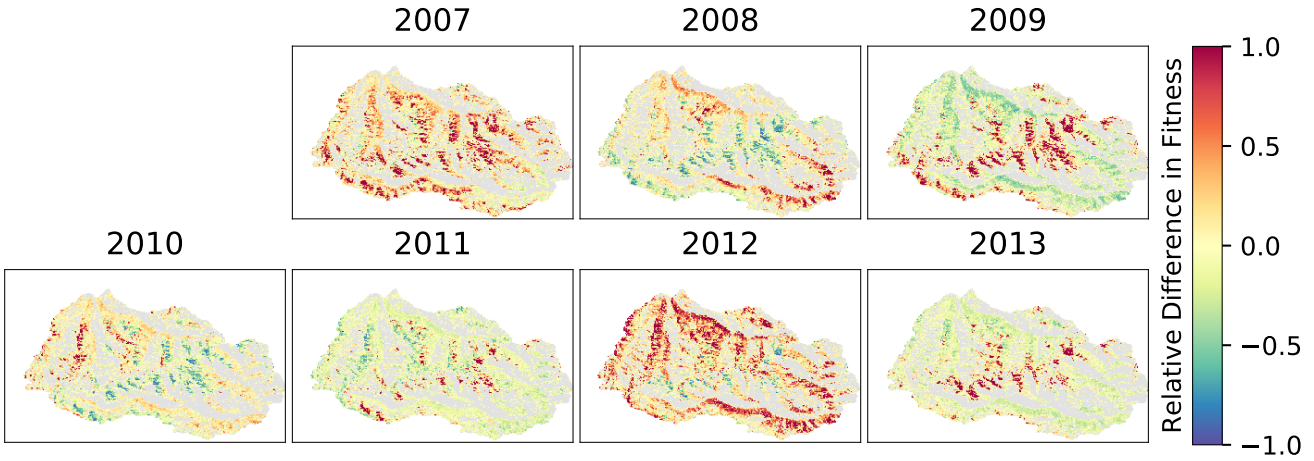

Fig. S20. Relative change in fitness value of *Pterostichus flavofemoratus* from one year to the other.

The species' fitness appears to be generally high in large parts of the park (Fig. S19). The logit-like fitness used here (equation 4 (MD)), permits to understand the contribution of each parameter to the overall fitness (49), where the change in the parameter's value reflects the change in fitness. As all cofactors have been feature scaled (SI 2), a unit change in the parameter directly influences the change in the contribution of the covariate to the fitness relative to the other explanatory variables. For *Pterostichus flavofemoratus*, most of the variables contribute to the fitness, although brightness and greenness to a lesser extent. Temperature, wetness and forest presence contribute the most, but greenness eastness, northness and slope are not to be neglected. As eastness, northness, slope and forest presence do not vary in time, these factors, along with the intercept (α_0) represent the base fitness of the species, which is then modified in time by the time-dependent covariates. In time, a general increasing trend in the fitness can be observed, which could in part explain the trend in increasing number of occupied cells observed in Fig. 3c (MD). The variation of the fitness in time is related to the high positive values of the parameters multiplying temperature, brightness and greenness.

Fig. S21 shows a comparison between the modeled and in-situ number of times the species was observed in each plot (k) and year. The modeled k is derived by drawing 1'000 samples from the binomial distribution $P(k = K|m) = (m = 1)B(p_{TP}, n) + (m = 0)B(p_{FP}, n)$ for each particle, plot and year (with the values of p_{TP} and p_{FP} displayed in table 1 (MD)). The displayed median k and the confidence interval are computed from the joint distribution. The modeled k of *Pterostichus flavofemoratus* encapsulates the in-situ data k within its confidence interval in almost all of the plots. Plots b to g in the valley of Orvieilles where no to little traps have been sampled positively appear to have a large confidence interval overestimating the

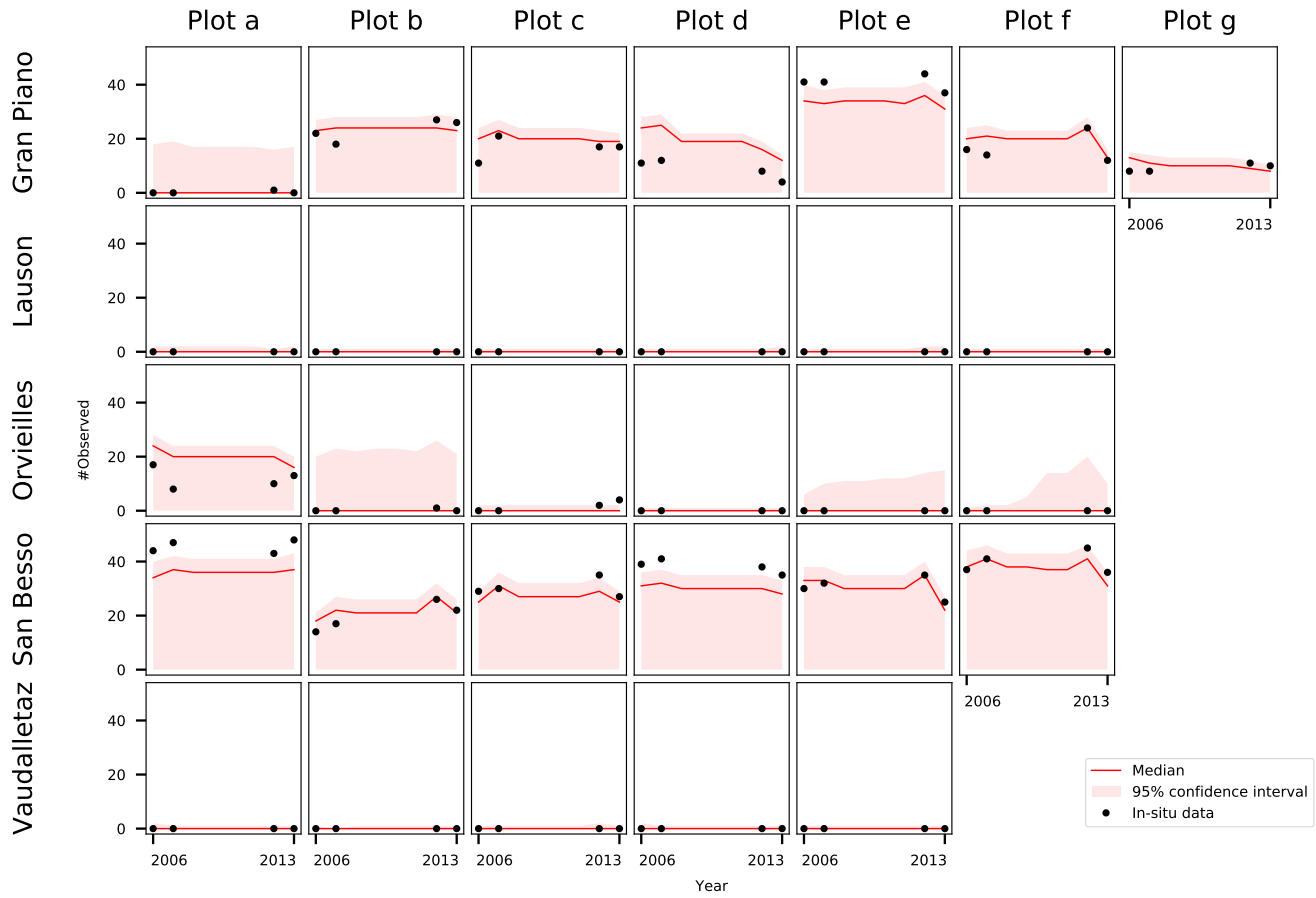

Fig. S21. Modeled (red with confidence interval) and in-situ (black dots) number of times *Pterostichus flavofemoratus* was observed (k) in each plot and year. The modeled k is generated by sampling from a binomial distribution $P(k = K|m) = (m = 1)B(p_{TP}, n) + (m = 0)B(p_{FP}, n)$. For the years 2008 to 2011 the average n is taken.

presence, but the median is close to the in-situ data. The median k in the plots in the valleys of Gran Piano and San Besso follows well the in-situ data, showing a good match in increasing and decreasing trends, although slightly underestimated in some plots.

This is confirmed in Fig. 4c, d (MD), which shows the distribution of the data within the modeled confidence interval as displayed in Fig. S21. Most of the data are around the 50% quantile, with 88.3% of the observations within 95% of the model envelope. The highest source of deviation from the median is at quantile 100%, where the model underestimates the certainty of occupancy of the plot, which can for instance be observed in plot e of Gran Piano and plots a and d of San Besso. Overestimation of the presence (left hand side of the plot in Fig. 4c, d (MD)), is less frequent, with plots d and f of Gran Piano being good examples.

B. *Carabus depressus*. As before, Fig. S22 show the simulated average occupancy of *Carabus depressus* computed from 480 repetitions, along the relative difference in average occupancy in each cell shown in Fig. S23.

The average occupancy of *Carabus depressus* well agrees with the in-situ data (Fig. S22). Overall, the average occupancy is more patchy, with clusters of high occupancy and surrounding lower occupancy maintained by the colonization process.

Similarly to *Pterostichus flavofemoratus*, the relative differences in occupancy of *Carabus depressus* (Fig. S23) are mostly within the borders of the high occupancy clusters, although a certain activity also happens within the clusters, to a higher extent than for *Pterostichus flavofemoratus*.

The fitness of *Carabus depressus* (Fig. S24) is less general than the one of *Pterostichus flavofemoratus* in the sense that it highlights some distinct features on the valley sides . *Carabus depressus* appears to be more sensitive to the presence of vegetation (greenness, table 1 (MD)), with a base fitness consisting in the intercept, eastness, northness, and, to a lesser extent, slope. The fitness parameters related to forest cover, temperature, wetness and brightness are not relevant for this species.

In time, the fitness increases in certain part of the domain, and decreases in others, mostly driven by greenness in the southern valleys. The patterns in fitness differences shown in Fig. S25 are mostly in agreement with the trends in occupancy observed in Fig. 3c (MD).

Fig. S26 shows a comparison between the in-situ and modeled k at the plot level, as done for *Pterostichus flavofemoratus*. Since the uncertainty on the false negative data (1 - p_{TP}, table 1 (MD)) is higher for the *Carabus depressus*, this translates in

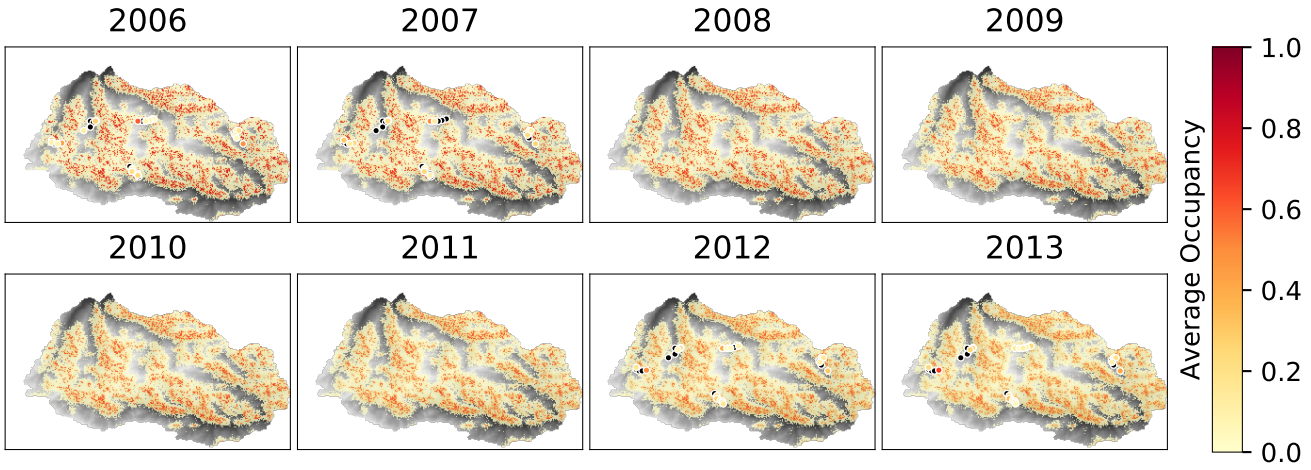

Fig. S22. Same as Fig. S17 but for *Carabus depressus*. Average occupancy in time with sampling years additionally displaying in-situ data.

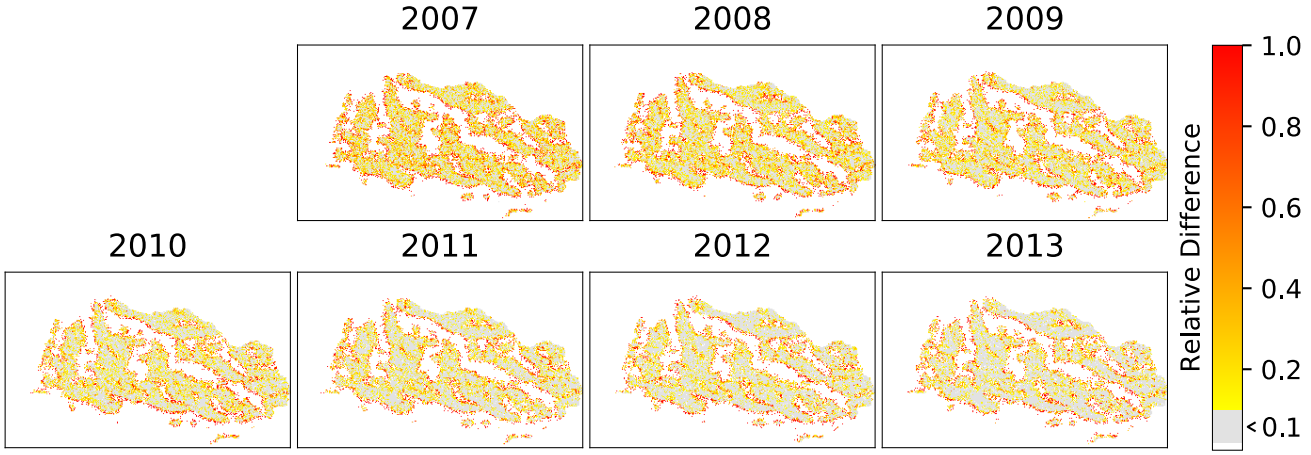

Fig. S23. Same as Fig. S18 but for *Carabus depressus*. Relative difference in average occupancy.

higher uncertainty above the median in the modeled k (Fig. S26). The modeled uncertainty captures most of the trends in the Gran Piano valley, and performs quite well in the Lauson and Vaudalletaz valley, but fails most of the times in Orvieilles. Plots a and f of Gran Piano, f of Lauson and a and e of Vaudalletaz are particularly interesting. They show that the decreasing and successively increasing trends in the observations are well capture by the modeled through changes in the confidence interval and, in some cases, by strong variations in the median. The valley of San Besso is interesting as well, where the model is capable to predict the complete absence of observations in plot b, but encapsulates within the uncertainty all the observations in the other plots. This is confirmed as well in Fig. 4c, d (MD), where the distribution of observations in the model envelope is displayed. The modeled median (50% quantile) explains a good part of the data, with 85% of the observations within the 95% confidence interval, but there appears to be a strong underestimation of the occupancy (right hand side). This is for instance to be seen, as mentioned before, in the valley of Orvieilles.

From table 1 (MD) we can to some extend understand the importance of the explanatory variables. Both species appear to have a preferential orientation (south-east facing slopes). Temporally varying covariates appear to be important for both species, but *Carabus depressus* relies only on greenness.

C. Trends in presence and sensitivity to parameters. The two species appear to be very stable. The fluctuations happen around one or multiple stable cores, where the pressure of colonization lessens. The trends observed in the occupancy (Fig. 3c (MD)) can mostly be related to the changes in fitness, but are only observed at the landscape level, with little to no trends observed at the pixel level. The average occupancy appears to be lower for *Carabus depressus* than for *Pterostichus flavofemoratus*, which is related to a more clustered fitness function along with a lower base fitness.

To further explore the link between occupancy and the model parameters, a sensitivity analysis of the mean final occupancy with respect to these parameters is performed. Each parameter is varied by 20% of itself (for the model parameters c, e and D) or its range (for the fitness parameters α_0 and β_j) in a positive and negative direction. The relative changes in the occupancy

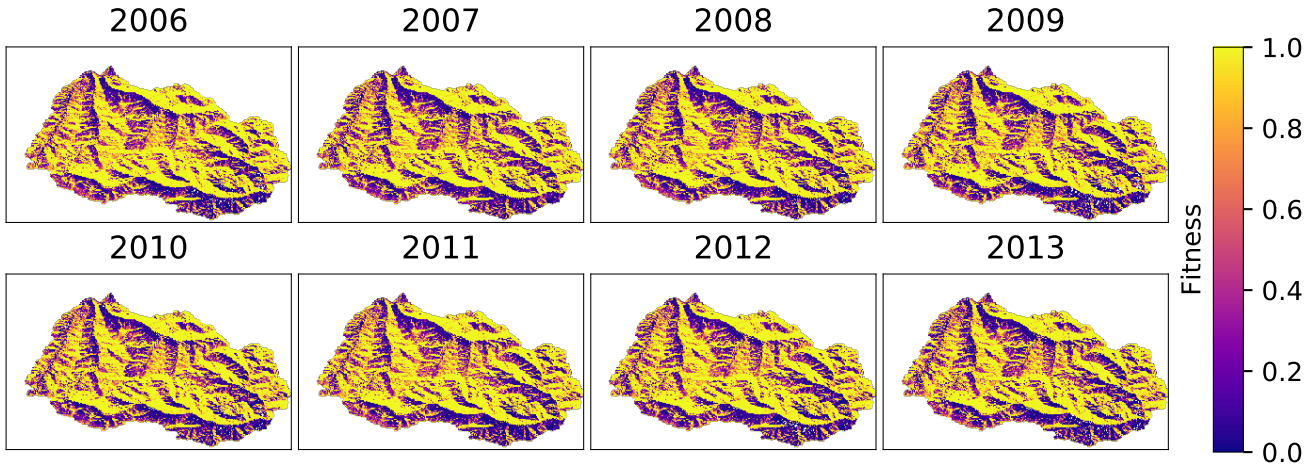

Fig. S24. Same as Fig. S19 but for *Carabus depressus*. Fitness of the species for the simulated years.

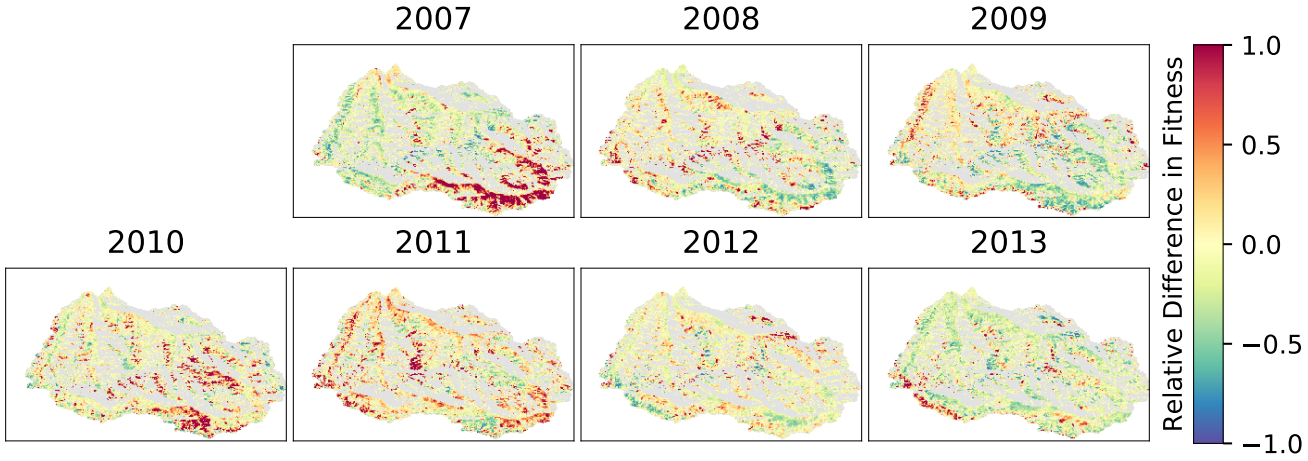

Fig. S25. Relative change in fitness value of *Carabus depressus* from one year to the other.

computed through this analysis are shown in Fig. S27.

Both species are fairly sensitive to changes in the intrinsic model parameters, c, e and D, which define how the buffer around the core of occupied cells behave. Of course, a larger value of e will lead to a lower average occupancy, and a higher value of c will increase the average occupancy. Interestingly, a change in the base fitness value (α_0) has a strong impact on the occupancy, and decreasing its value seems to have a much higher impact than increasing it. For both species, brightness is the most sensitive time-dependent covariate. Changes in occupancy are inversely proportional to the changes in the brightness value, displaying similar behaviors as the extinction coefficient. This is true as well for slope and the presence of forest. Eastness and northness both change proportionally to the change in parameter value, but show a much stronger reaction for *Pterostichus flavofemoratus*, even though the initial parameter values are in the same order of magnitude for both species.

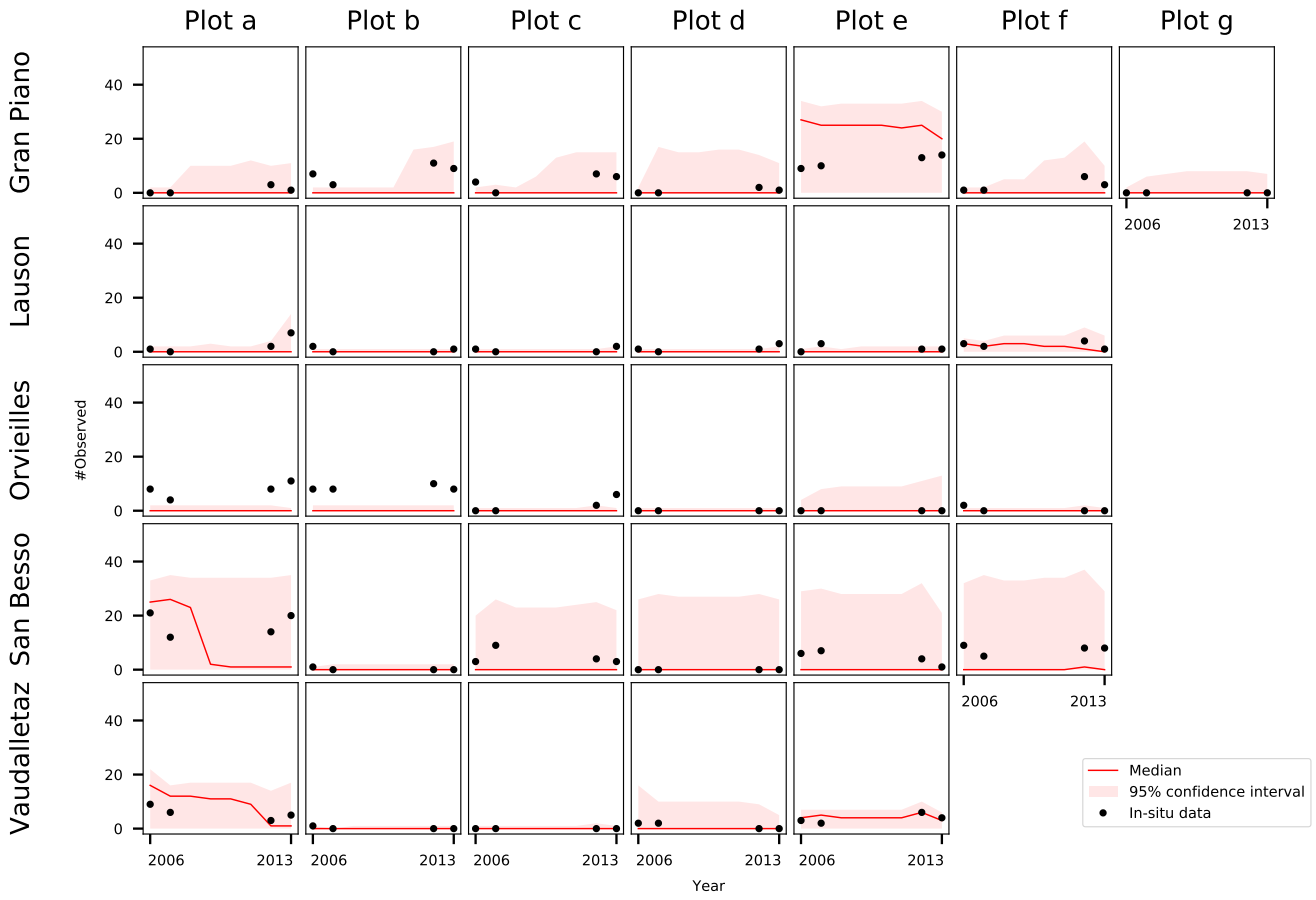

Fig. S26. Same as Fig. S21 but for *Carabus depressus*. Modeled and in-situ number of times the species was observed (k) in each plot and year.

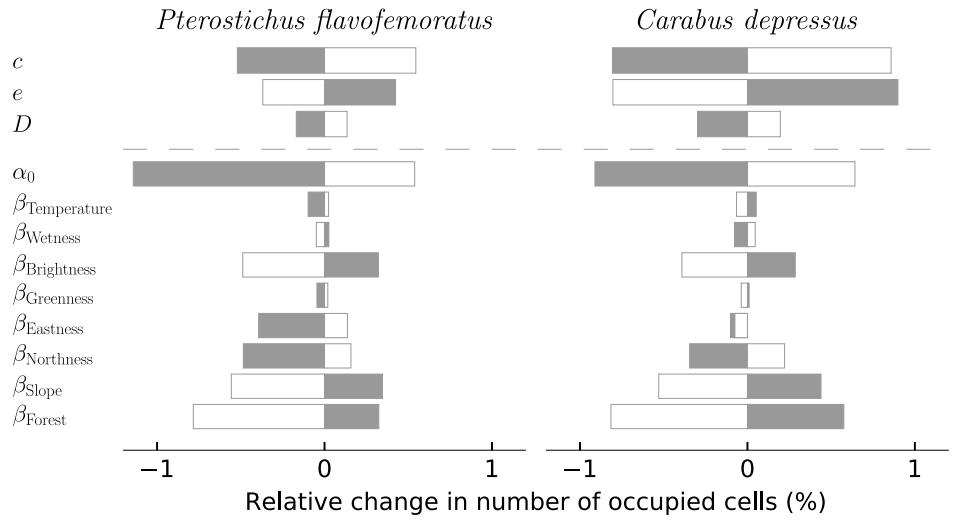

Fig. S27. Sensitivity of the calibrated model in terms of relative change in occupied cells to a 20% perturbation of the parameter values for the two species of interest. The white bars represent a positive change in the parameter value, and the grey bars a negative change. The dashed line separates the parameters intrinsic to the model (c, e and D) and the parameters employed by the fitness function (equation 4 (MD)).

1. Kromp B (1999) Carabid beetles in sustainable agriculture: a review on pest control efficacy, cultivation impacts and enhancement A2 - Paoletti, M.G. BT - Invertebrate Biodiversity as Bioindicators of Sustainable Landscapes. *Agriculture, Ecosystems and Environment* 74:187–228.
2. Rainio J, Niemelä JK (2003) Ground beetles (Coleoptera : Carabidae) as bioindicators. *Biodiversity and Conservation* 12(McGeoch 1998):487–506.
3. Lövei GL, Sunderland KD (1996) Ecology and Behavior of Ground Beetles. *Annu. Rev. Entomol.* 41(1 12):231 – 256.
4. Niemelä J (2001) Carabid beetles (Coleóptera : Carabidae) and habitat fragmentation: a review. *Eur. J. Entomol.* 98:127–132.

5. Jopp F, Reuter H (2005) Dispersal of carabid beetles - Emergence of distribution patterns. *Ecol. Modell.* 186(4):389–405.
6. Kotze DJ, et al. (2011) Forty years of carabid beetle research in Europe - from taxonomy, biology, ecology and population studies to bioindication, habitat assessment and conservation. *Zookeys* 100(SPEC. ISSUE):55–148.
7. Skłodowski J (2009) Interpreting the condition of the forest environment with use of the SCP/MIB model of carabid communities (Coleoptera: Carabidae). *Baltic Journal of Coleopterology* 9(2):89–100.
8. Gudowska A, et al. (2017) Mass scaling of metabolic rates in carabid beetles (carabidae) – the importance of phylogeny, regression models and gas exchange patterns. *The Journal of Experimental Biology* 220(18):3363–3371.
9. Kleinwächter M, Rickfelder T (2007) Habitat models for a riparian carabid beetle: Their validity and applicability in the evaluation of river bank management. *Biodiversity and Conservation* 16(11):3067–3081.
10. Schröder B (2008) Challenges of species distribution modeling belowground. *J. Plant Nutr. Soil Sci.* 171(3):325–337.
11. Akçakaya HR (2000) Viability analyses with habitat-based metapopulation models. *Researches on Population Ecology* 42(1):0045.
12. Griffith DM, Brown JS (1992) A Null Model of Patch Assessment with an Application to a Carabid Cave Beetle. *Oikos* 64(3):523–526.
13. Jordán F, Magura T, Tóthmérész B, Vasas V, Ködöböcz V (2007) Carabids (Coleoptera: Carabidae) in a forest patchwork: A connectivity analysis of the Bereg Plain landscape graph. *Landscape Ecol.* 22(10):1527–1539.
14. Rushton S, Sanderson R, Luff M, Fuller R (1996) Modelling the spatial dynamics of ground beetles (Carabidae) within landscapes. *Ann. Zool. Fenn.* 33(1):233–241.
15. Benjamin R, Cédric G, Pablo I (2008) Modeling spatially explicit population dynamics of Pterostichus melanarius I11. (Coleoptera: Carabidae) in response to changes in the composition and configuration of agricultural landscapes. *Landscape Urban Plann.* 84(3-4):191–199.
16. Pichancourt JB, Burel F, Auger P (2006) Assessing the effect of habitat fragmentation on population dynamics: An implicit modelling approach. *Ecol. Modell.* 192(3-4):543–556.
17. Reuter H, et al. (2005) The concepts of emergent and collective properties in individual-based models - Summary and outlook of the Bornhöved case studies. *Ecol. Modell.* 186(4):489–501.
18. Wachmann E, Platen R, Barndt D (1995) *Laufkäfer, Beobachtung, Lebensweise.* (Augsburg : Naturbuch-Verlag).
19. Tenan S, et al. (2016) Hierarchical models for describing space-for-time variations in insect population size and sexratio along a primary succession. *Ecol. Modell.* 329:18–28.
20. Woodcock BA (2007) Pitfall Trapping in Ecological Studies in *Insect Sampling in Forest Ecosystems.* Vol. 3, pp. 37–57.
21. Thiele HU (1977) *Carabid Beetles in Their Environments - A Study on Habitat Selection by Adaptations in Physiology and Behaviour.* (Springer, Berlin, Heidelberg).
22. Parkinson L, Marshall EJP (1998) Isolating the components of activity-density for the carabid beetle Pterostichus melanarius in farmland. 116(1):103–112.
23. Newell SC (1986) Ph.D. thesis.
24. Burel F, Butet A, Delettre YR, Millàn De La Peña N (2004) Differential response of selected taxa to landscape context and agricultural intensification. *Landsc. Urban Plan.* 67:195–204.
25. Matern A, Drees C, Meyer H, Assmann T (2008) Population ecology of the rare carabid beetle Carabus variolosus (Coleoptera: Carabidae) in north-west Germany. *Journal of Insect Conservation* 12(6):591–601.
26. den Boer PJ (1990) The Survival Value of Dispersal in Terrestrial Arthropods *. *Biol. Conserv.* 54(356):175–192.
27. Eversham BC, Roy DB, Telfer MG (1996) Urban, industrial and other manmade sites as analogues of natural habitats for Carabidae. *Ann. Zool. Fenn.* 33(c):149–156.
28. Gobbi M, et al. (2007) Environmental features influencing Carabid beetle (Coleoptera) assemblages along a recently deglaciated area in the Alpine region. *Ecological Entomology* 32(6):682–689.
29. Viterbi R, et al. (2013) Patterns of biodiversity in the northwestern Italian Alps: a multi-taxa approach. *Community Ecology* 14(1):18–30.
30. Petit S, Burel F (1998) Effects of landscape dynamics on the metapopulation of a ground beetle (Coleoptera, Carabidae) in a hedgerow network. *Agriculture, Ecosystems & Environment* 69(3):243–252.
31. Davies KF, Margules CR (1998) Effects of Habitat Fragmentation on Carabid Beetles : Experimental Evidence. *J. Anim. Ecol.* 67(3):460–471.
32. Thomas CF, Green F, Marshall EJ (1997) Distribution, Dispersal and Population Size of the Ground Beetles, Pterostichus melanarius (Illiger) and Harpalus rufipes (Degeer) (Coleoptera, Carabidae), in Field Margin Habitats. *Biological Agriculture and Horticulture* 15(1-4):337–352.
33. Cornes RC, van der Schrier G, van den Besselaar EJ, Jones PD (2018) An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research: Atmospheres* 123(17):9391–9409.
34. D'Onofrio D, Palazzi E, von Hardenberg J, Provenzale A, Calmanti S (2014) Stochastic Rainfall Downscaling of Climate Models. *J. Hydrometeorol.* 15(2):830–843.
35. Haylock MR, et al. (2008) A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research Atmospheres* 113(20).
36. Dodson R, Marks D (1997) Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Climate research* 8(1):1–20.
37. Barry RG (2008) *Mountain Weather and Climate.* (Cambridge University Press), 3 edition.
38. Rolland C (2003) Spatial and seasonal variations of air temperature lapse rates in alpine regions. *J. Climate* 16(7):1032–1046.
39. Pasetto D, et al. (2018) Integration of satellite remote sensing data in ecosystem modelling at local scales: Practices and trends. *Methods Ecol. Evol.* 9(8):1810–1821.
40. Hasan M, Baig A, Zhang L (2014) Derivation of a tasselled cap transformation based on Landsat 8 at- Derivation of a tasselled cap transformation based on Landsat 8 at- satellite reflectance. (April).
41. Kempeneers P, Sedano F, Seebach L, Strobl P, San-Miguel-Ayanz J (2011) Data fusion of different spatial resolution remote sensing images applied to forest-type mapping. *IEEE Trans. Geosci. Remote Sens.* 49(12):4977–4986.
42. Janžekovič F, Novak T (2012) PCA – A Powerful Method for Analyze Ecological Niches in *Principal Component Analysis - Multidisciplinary Applications*, ed. Sanguansat P. Intech edition, pp. 127–142.
43. Zuur AF, Ieno EN, Smith GM (2007) *Analyzing Ecological Data.* (Springer).
44. Moilanen A (2004) SPOMSIM: Software for stochastic patch occupancy models of metapopulation dynamics. *Ecol. Modell.* 179(4):533–550.
45. Moilanen A (1999) Patch Occupancy Models of Metapopulation Dynamics : Efficient Parameter Estimation Using Implicit Statistical Inference. *Ecology* 80(3):1031–1043.
46. Mckay MD, Beckman RJ, Conover WJ (1977) A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* 21(2):239–245.
47. Arulampalam MS, Maskell S, Gordon N, Clapp T (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* 50(2):174–188.
48. Ionides EL, Nguyen D, Atchadé Y, Stoev S, King AA (2015) Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proceedings of the National Academy of Sciences* 112(3):719–724.
49. Hegel TM, Cushman SA, Evans J, Huettmann F (2010) Current State of the Art for Statistical Modelling of Species Distributions in *Spatial Complexity, Informatics, and Wildlife Conservation*, eds. Cushman SA, Huettmann F. pp. 273–311.