

# Environmental niche modelling

## **Introduction**

Environmental niche modelling, also called climate envelope modelling, is a common approach to predict the spatial distribution of species. It uses the ecological niche that a species occupies, i.e. the range of environmental conditions in which species can live and be observed, to predict the future distribution of a species given the projected changes in the environment, such as climatic changes.

The environmental niche modelling approach differs in a fundamental way from the approaches introduced in the previous assignments. In those models, a set of rules (physical laws or empirical relations) was implemented into a set of dynamical equations. Given such equations, and provided a set of initial conditions, the models could be used to study the temporal and spatial dynamics of a system. Environmental niche models, on the other hand, are not dynamical models but statistical models, and thus they are not based on mechanisms and laws but on statistical, observational relationships. More specifically, the approach has two main steps:

- i) observations of species distribution are linked to environmental variables through statistical fitting.
- ii) Such relationships are used to predict future species distribution when the environmental variables are changed (as projected by climate models).

It is important to keep in mind that the modelling approach is based on different principles. Statistical models are very powerful in predictions within the measured range and for the current conditions. A disadvantage of statistical methods is however that predictions beyond the measured range (e.g. extrapolation) can be erroneous. Nonetheless, due to a lack of mechanistic understanding, environmental niche modelling is often applied to study species range shifts triggered by climate change.

With environmental niche modelling the possible implications of climate change for biodiversity can be assessed: biodiversity can decrease as a result of disappearance of a set of environmental conditions (e.g. due to physical barriers such as ocean coasts). One limitation of environmental niche modelling is the fact that displacement rates are not accounted for. That is, it can only indicate which areas will be suitable for species in future climates, without considering whether displacement rates of species are high enough to move to these areas.

In this assignment we will apply environmental niche modelling to a set of fictive species using real global climate maps of current and future environmental conditions. The approach we use is based on a paper by Leroy et al. (2015) in which a package for the statistical software R is introduced. Here we will not establish ourselves the relationship between species observation and environmental variables (step i) of the modelling approach), but we will assume the relationships are known (e.g. from literature). We will focus only on the step ii), i.e. we will use the established relationship to reproduce current species ranges and predict them for the future (Part I of the assignment). In the second part of the assignment we will study the displacement rates required for species to track the projected changes in mean annual temperature, as done previously by Burrows et al. (2011).

## PART I

Download “EnvNiche.zip” from Blackboard and extract the file in your desired working directory. The file contains four files: three .m files and one .mat file. The .mat file can be opened by typing:

```
load('EnvNiche.mat')
```

This loads six variables into the workspace.

BIO is a layered matrix (i.e. it is a 3-D matrix instead of the 1-D vectors and the 2-D matrices you've worked with so far). Each layer is a global map of a bioclimatic variable averaged over the period 1950-2000 for all land area excluding Antarctica (Hijmans et al., 2005). The bioclimatic variables are listed in Table 1. The units of `BIO(:, :, 1:11)` are degrees Celsius. The units of `BIO(:, :, 12:19)` are mm. More details on the bioclimatic variables can be found at <http://www.worldclim.com/bioclim>. The maps were composed as described on this page <http://www.worldclim.com/methods>. The maps have a resolution of 10×10 arc minutes. Matrices `pixX` and `pixY` give the width and height respectively of the grid cells of the maps in km. Matrix `mask` is a map filled with zeros for ocean cells and ones for land cells. The maps cover the earth from -180 (column 1) to 180 (column 2160) degrees longitude and 90 (row 1) to -60 (row 900) degrees latitude, as given by vector `x` and vector `y`.

*Table 1. Bioclimatic variables*

<code>BIO(:, :, 1)</code>	= Annual Mean Temperature
<code>BIO(:, :, 2)</code>	= Mean Diurnal Range (Mean of monthly (max temp - min temp))
<code>BIO(:, :, 3)</code>	= Isothermality ( <code>BIO2/BIO7</code> ) (* 100)
<code>BIO(:, :, 4)</code>	= Temperature Seasonality (standard deviation *100)
<code>BIO(:, :, 5)</code>	= Max Temperature of Warmest Month
<code>BIO(:, :, 6)</code>	= Min Temperature of Coldest Month
<code>BIO(:, :, 7)</code>	= Temperature Annual Range ( <code>BIO5-BIO6</code> )
<code>BIO(:, :, 8)</code>	= Mean Temperature of Wettest Quarter
<code>BIO(:, :, 9)</code>	= Mean Temperature of Driest Quarter
<code>BIO(:, :, 10)</code>	= Mean Temperature of Warmest Quarter

BIO (:, :, 11) = Mean Temperature of Coldest Quarter
BIO (:, :, 12) = Annual Precipitation
BIO (:, :, 13) = Precipitation of Wettest Month
BIO (:, :, 14) = Precipitation of Driest Month
BIO (:, :, 15) = Precipitation Seasonality (Coefficient of Variation)
BIO (:, :, 16) = Precipitation of Wettest Quarter
BIO (:, :, 17) = Precipitation of Driest Quarter
BIO (:, :, 18) = Precipitation of Warmest Quarter
BIO (:, :, 19) = Precipitation of Coldest Quarter

**Exercise 1**

a) Explore the BIO maps using the `imagesc()` function. Use the following function to visualize the maps:

```
imagesc(x,y,BIO(:, : , <VAR>));
```

where here `<VAR>` is the bioclimatic variable number (i.e. a value from 1 to 19). After visualizing the map, enter the commands `colorbar`, `axis xy` and `axis image` one by one and explain what each command does.

b) Use the commands of (a) to plot the matrices `mask`, `pixX` and `pixY`. Explain why `pixX` decreases towards the poles.

c) Calculate the percentage of land cells in the `mask` map. TIP: one way to do this is to use the function `mean()`. Explain why this percentage need not be equal to the actual percentage of land area on the map. Use matrices `mask`, `pixX` and `pixY` to show that the percentage of land area on the map equals 29.67 % and include the commands you used in the report.

**Exercise 2**

The maps of the bioclimatic variables can be used to create maps of environmental suitability for a given species. Environmental suitability has a value between 0 and 1. The more suitable environments have a value close to 1 and the less suitable environments have a value close to 0. Two of the included `.m` files are response functions `logistic(x,alpha,beta)` (equation 1) and `gauss(x,mu,sig)` (equation 2). These functions define the relation between a certain bioclimatic variable and the environmental suitability for a species.

$$f(x) = \frac{1}{1 + e^{\frac{x-\beta}{\alpha}}} \quad (1)$$

$$f(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2)$$

In equations 1 and 2,  $x$  is a bioclimatic variable and  $\alpha$ ,  $\beta$ ,  $\mu$ , and  $\sigma$  are parameters that control the shape of the response curves as described in the comments of the `.m` files. Explore the behavior of the functions `logistic()` and `gauss()` by applying the functions to a theoretical bioclimatic variable for different values for  $\beta$ ,  $\mu$ , and  $\sigma$ , as done in figure 1. Use `plot()` to visualize the functions and `hold on` to show multiple plots in one graph. For each plot include a title, axis labels and a legend. Let the bioclimatic variable range from 0 to 50 (e.g. `x=0:0.1:50`). Describe in your own words how the different parameters affect the shapes of the response curves.

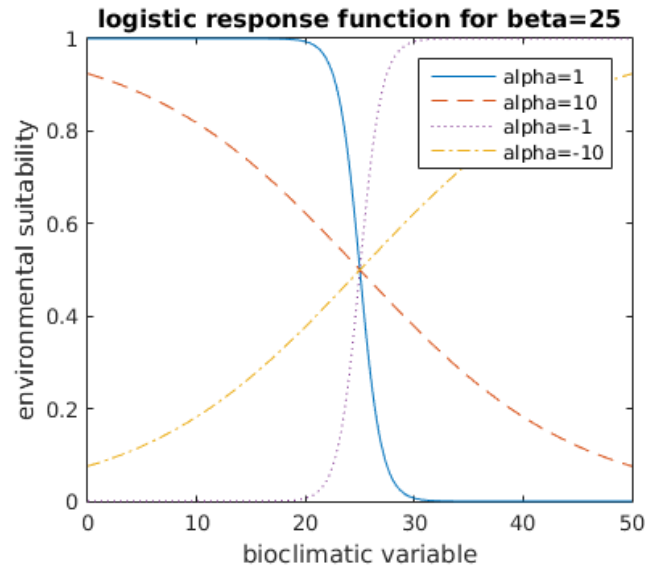


Figure 5.1: Logistic function for different values of  $\alpha$ .

### Exercise 3

We can apply the functions `gauss()` and `logistic()` on bioclimatic maps to create environmental suitability maps. We will do this for three terrestrial species:

- Species 1 prefers annual temperatures of 20 degrees Celsius.
- Species 2 does not tolerate temperatures below 0 degrees Celsius.
- Species 3 cannot live in climates where the average monthly precipitation drops below 50 mm.

Use the functions `gauss()` (for species 1) and `logistic()` (for species 2 and 3) on the maps `BIO(:, :, 1)`, `BIO(:, :, 6)` and `BIO(:, :, 14)` to create the environmental suitability maps for the three species. For species 1 use `sigma=5` degrees C, for species 2 `alpha=-10` degrees C, and for species 3 `alpha=-10` mm. Store the maps as matrices and include plots of the maps in your report. For each map, state the function, the bioclimatic maps and parameter values of that you used.

TIP: organize your commands in a script from now on, instead of using the command window.

### Exercise 4

The maps generated in Exercise 3 are not very realistic, for example, species 2 has suitable environments both in Indonesia and on parts of Greenland. The environmental suitability for species generally depends on more than one bioclimatic variable. We can create suitability maps that depend on more than one bioclimatic variable by multiplying suitability maps. For example, by multiplying the suitability maps of species 1 and 2 we obtain a suitability map for a species that requires temperatures above 0 degrees Celsius and  $>50$  mm of monthly precipitation. In order to obtain values between 0 and 1 such a map would have to be rescaled. This can be done with the `rescale()` function.

- a) Open `rescale.m` and explain how this function works.
- b) Next, create a suitability map for a new species (species 4) that needs annual

temperatures of around 10 degrees ( $\sigma=5$ ) and annual rainfall of around 600 mm ( $\sigma=100$ ). Store the map as a matrix and include the used commands and a plot of the map in your report.

TIP I: use `gauss()` and `rescale()` to create this map. TIP II: Be aware that you need to multiply the matrices element by element (i.e. by using `.*`).

### **Exercise 5**

Even when a set of environmental conditions is very suitable for a given species, the species may not be observed in locations with those environmental conditions, because of factors acting at smaller spatial scales, such as biotic pressure (competition with other species, predation), disturbances, stochastic events. The same can occur vice versa: species may be found in areas of low suitability because of very particular conditions allowing their occurrence (microclimatic/microhabitat conditions, facilitation by other species, etc). To account for this, the suitability maps should be converted into probability maps. Probability maps reflect the likelihood of encountering a species at certain location given the bioclimatic variables. The probability can be calculated using the `rescale` function: `rescale(y, mmin, mmax)`, where `y` is the suitability map and `mmin` and `mmax` are the likelihoods of encountering a species in its least and most suitable environments respectively.

a) Create two probability maps for species 4 of Exercise 4, one map with `mmin=0` and `mmax=0.5` and one map with `mmin=0.2` and `mmax=0.5`. Include the commands and the maps in your report, and store the probability maps as matrices.

The probability maps can be used to create actual presence/absence maps by making a matrix of random numbers that has the size of the probability map. This can be done using the function `rand(m, n)`, which gives a random matrix with `m` rows and `n` columns with values between 0 and 1. A species is present in a cell if the value in the random matrix in that cell is lower than the probability value in the same cell, and absent if the opposite is true. In Matlab the values of matrices can be compared element by element using *conditional statements* (e.g. `>`, `<` or `==`). For example `matrix1>matrix2`, gives a matrix with ones (true) and zeros (false) where `matrix1` is greater than `matrix2` respectively.

b) Use conditional statements to create presence/absence maps for both probability maps and include the maps and the commands in your report. Describe and explain the differences between the two presence/absence maps.

c) Each of the two presence/absence maps reflect one possible way a species is distributed, i.e. they are *realizations* of a statistical model. Is it likely that these presence/absence maps correspond with the actual distribution of the species? Why (not)?

### **Exercise 6**

We will now study how the suitability map of species 4 will change over the coming decades. A number of global circulation models (GCMs) have been developed to assess the future climate under elevated greenhouse gas concentrations. For the last IPCC report (AR5), four greenhouse gas concentration trajectories, called representative concentration pathways (rcp), have been used as inputs for these GCMs. The outputs of the models can be found on [http://www.worldclim.com/cmip5\\_10m](http://www.worldclim.com/cmip5_10m).

The data are grouped in two tables, one for 2050 (average for 2041-2060) and one for 2070 (average for 2061-2080). Each row contains the output of one GCM. Each column contains the outputs of the models for a representative concentration pathway. The codes rcp26, rcp45, rcp60 and rcp85 stand for a radiative forcing of +2.6, +4.5, +6.0 and +8.5 W/m<sup>2</sup> respectively in the year 2100, relative to pre-industrial values.

Download the bioclimatic variables (bi) of 2050 as projected by the HadGEM2-ES model (a model by the Met Office Hadley Centre for Climate Science and Services; Martin et al., 2011) for the rcp45 pathway. Extract “he45bi50.zip” to your working directory. To import a bioclimatic variable, you can use the command

```
map=imread('he45bio50<VAR>.tif');
```

where <var> corresponds to the numbers of the bioclimatic variables in table 1. The imported matrices should be converted to the standard numerical format within MATLAB (double). This can be done using the following command:

```
map=double(map);
```

Next, the oceans need to be masked out (we are considering only terrestrial species). We can do this by setting the values of the ocean cells to NaN ('Not A Number').

```
map(mask==0)=NaN;
```

The maps of temperature (e.g. “he45bi501.tif”, “he45bi502.tif”, ..., “he45bi5011.tif”) have units of [degrees Celsius \* 10]. To convert these maps to degrees Celsius divide by 10:

```
map=map/10;
```

Use the above commands to create a suitability map for species 4 for the climate of 2050. Compare the map with the suitability map of Exercise 4. Does its suitable area expand, shrink, move? Why?

## PART II

Although the environmental niche modelling approach can indicate which areas will be suitable for species in future climates, it may fail to predict the actual future distribution of species because it does not consider that species need to move to reach new suitable areas. Thus, their displacement rates may limit their redistribution. In this second part we will study the displacement rates required for species to track the projected changes in mean annual temperature (following Burrows et al. (2011)).

### Exercise 8

#### a) Spatial gradient calculation

\* Calculate the spatial gradient in the current temperature map as follows:

```
[gx,gy]=gradient(BIO(:, :, 1));
```

Go to the MATLAB documentation to see what the function `gradient()` does (click on Algorithms). How does this approach differ from the self-organization exercise?

\* What is the unit of the gradient that you calculated? How can we convert to degrees Celsius per km?

\* Use Pythagoras' theorem to calculate the resulting vector of the spatial gradient with unit degrees Celsius per km as depicted in Figure 5.2. Store the map as a matrix, and include a figure in your report with a colorbar between 0 and 0.1. Explain the spatial differences in the temperature gradients.

#### b) Temporal gradient calculation

Generate a map of the temperature difference between the current and future climate by subtracting the current temperature map from the temperature map by the HadGEM2-ES model. Divide by the number of decades between the two maps to obtain the temporal gradient in degrees Celsius per decade. Notice that the “current” temperature map can be assumed to be valid for 1975 as the map is based on the averages from 1950-2000. Store the map as a matrix and include a figure in your report.

#### c) Velocity calculation

Combine the maps of the spatial and temporal gradients to get a velocity map in km per decade (TIP: use units of the two gradients to figure out how to obtain the velocity).

Include a figure in your report with a colorbar between 0 and 100. What does the velocity map represent? Do you think all species will be able to move with this rate? Use Figure 5.3 below to create a map that shows the areas where climate change is too rapid for split-hoofed mammals to track. TIP: Use conditional statements (see in exercise 5) to compare the velocity map with the median value of split-hoofed mammals.

BONUS: repeat exercise 8 for a different bioclimatic variable and explain the differences.

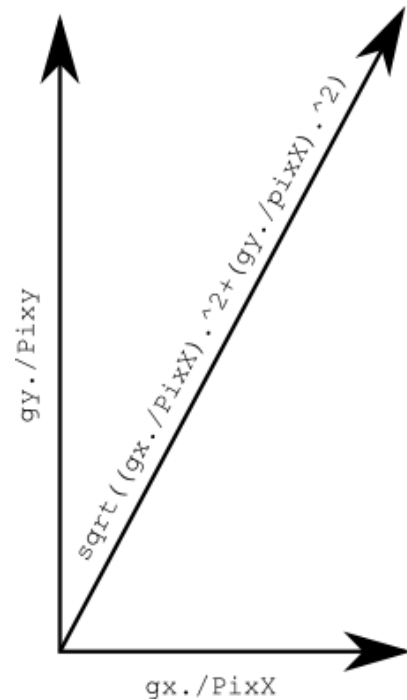


Figure 5.2: Calculation of the resulting vector.



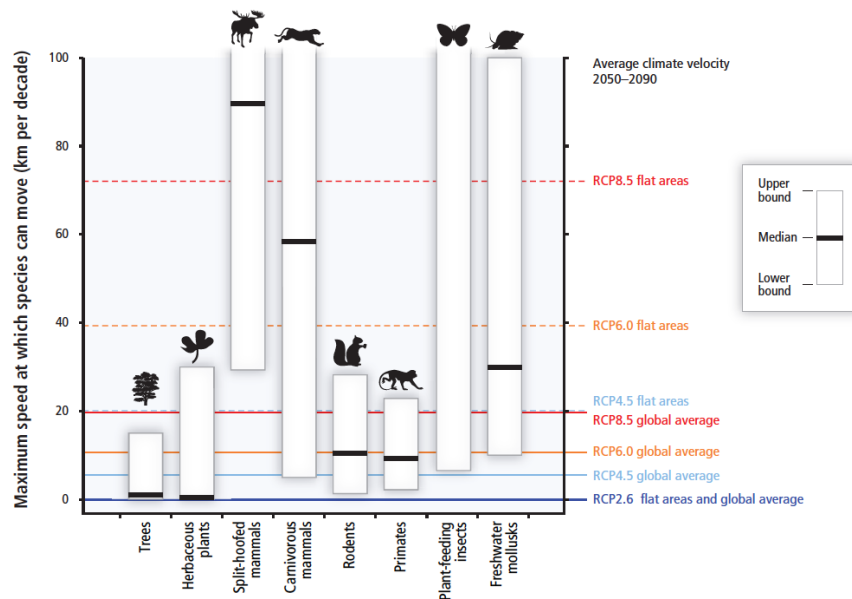


Figure 5.3: Maximum speeds at which species can move across landscapes (based on observations and models; vertical axis on left), compared with speeds at which temperatures are projected to move across landscapes (climate velocities for temperature; vertical axis on right). White boxes with black bars indicate ranges and medians of maximum movement speeds for trees, plants, mammals, plant-feeding insects (median not estimated), and freshwater mollusks. For RCP2.6, 4.5, 6.0, and 8.5 for 2050–2090, horizontal lines show climate velocity for the global-land-area average and for large flat regions. Species with maximum speeds below each line are expected to be unable to track warming in the absence of human intervention (IPCC, 2014).

## References

Burrows, M. T., Schoeman, D. S., Buckley, L. B., Moore, P., Poloczanska, E. S., Brander, K. M., Brown, C., Bruno, J. F., Duarte, C. M., Halpern, B. S., Holding, J., Kappel, C. V., Kiessling, W., O'Connor, M. I., Pandolfi, J. M., Parmesan, C., Schwing, F. B., Sydeman, W. J., and Richardson, A. J. (2011). The pace of shifting climate in marine and terrestrial ecosystems. *Science*, 334(6056): 652–5.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A. (2005), Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.*, 25: 1965–1978.

IPCC, 2014: Summary for policymakers. In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, p. 14.

Leroy, B., Meynard, C. N., Bellard, C. and Courchamp, F. (2015), virtualspecies, an R package to generate virtual species distributions. *Ecography* 38: 001-009.

Martin, G. M., Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., Hinton, T. J., Jones, C. D., McDonald, R. E., McLaren, a. J., O'Connor, F. M., Roberts, M. J., Rodriguez, J. M., Woodward, S., Best, M. J., Brooks, M. E., Brown, A. R., Butchart, N., Dearden, C., Derbyshire, S. H., Dharssi, I., Doutriaux-Boucher, M., Edwards, J. M., Falloon, P. D., Gedney, N., Gray, L. J., Hewitt, H. T., Hobson, M., Huddleston, M. R., Hughes, J., Ineson, S., Ingram, W. J., James, P. M., Johns, T. C., Johnson, C. E., Jones, A., Jones, C. P., Joshi, M. M., Keen, A. B., Liddicoat, S., Lock, a. P., Maidens, A. V., Manners, J. C., Milton, S. F., Rae, J. G L., Ridley, J. K., Sellar, A., Senior, C. a., Totterdell, I. J., Verhoef, A., Vidale, P. L. and Wiltshire, A. (2011). The HadGEM2 family of Met Office Unified Model climate configurations. *Geoscientific Model Development*, 4(3): 723–757.