

# Niche Dynamics of Invasive Species

Olivier Broennimann

2023-11-10

## b. Niche dynamics of invasive species

In this section, we will use the `ecospat` package to investigate the niche dynamics of an invasive species across its native and invasive ranges. We will develop the example of the spotted knapweed (*Centaurea stoebe*), an Asteraceae native to Central and Eastern Europe. In Europe, both diploid and tetraploid cytotypes can be found, but only the tetraploids have spread to North America, where it is considered an invasive species, most presumably because of its ability to produce perennial polycarpic plants (Treier et al. 2009 ; Mráz et al. 2011). This species served as study case upon which most of the methodological development on niche dynamics presented in this section have been based (Broennimann et al. 2007 ; Broennimann et al. 2012 ; Petitpierre et al. 2012 ; Guisan et al. 2014). Note however the analyses presented here use occurrence and environmental data at a lower resolution than used in the original analyses, which can lead to slightly different results. The main idea of the niche dynamics analyses is to quantify environmental niches by investigating the occurrence density of species in a gridded climatic space allowing for direct pixel quantification and permutation tests.

### i. Data preparation

First we need to load the R packages need for the analyses. In addition to the package `ecospat`, the package `terra` is need for handling geo-referenced datasets, `geodata` to download the climatic dataset and `dplyr` to query, subset and summarize data, and `ade4` to perform the ordination analyses.

```
# Load the needed packages
library(ecospat)
library(terra)
library(tidyterra)
library(geodata)
library(dplyr)
library(ade4)
library(viridis)
```

Then we download the climatic data, here the bioclim variables of the worldclim 2.1 dataset at 10' resolution using `worldclim_global` from the `geodata` package. Note that the data is downloaded to a randomly attributed temporary local folder using `tempdir`. For computational efficiency, we aggregate the climatic data at 30' resolution (~50 km) using `aggregate` from `terra`. The names of the worldclim layers can be renamed for easier labeling in following plots. To delimit the study area for the analyses we also download (and unzip) the geo-referenced polygons of the Terrestrial Ecoregions of the World distributed by the WWF. Function `vect` from `terra` allows to convert the polygons in a convenient vector object for further analyses, and `set` coordinate reference system (argument `crs`).

```

set.seed(123)
tempdir<-tempdir()

## download worlclim 2.1 climate data at 10' resolution and biomes data

clim.world<-geodata::worldclim_global("bio",res=10,version="2.1",path=tempdir)
clim.world<-terra::aggregate(clim.world,3,fun=mean,na.rm=TRUE)
names(clim.world)<-sub("wc2.1_10m_", "", names(clim.world))
names(clim.world)<-sub("_", "", names(clim.world))

## download Terrestrial Ecoregions of the World dataset

teow.URL<-"https://storage.googleapis.com/teow2016/Ecoregions2017.zip"
# official website for manual download:
# https://files.worldwildlife.org/wwfcmprod/files/Publication/file
# /6kcchn7e3u_official_teow.zip
download.file(teow.URL,destfile = paste0(tempdir,"/TEOW.zip"))
unzip(paste0(tempdir,"/TEOW.zip"),exdir = tempdir)
biomes<-terra::vect(paste0(tempdir,"/Ecoregions2017.dbf"),crs="+proj=longlat")

```

We now load and prepare the data for the two regions of analysis, namely Europe, where the species is native, and North America, where the species is invasive. The occurrence data are retrieved from the `extdata` folder included in the *ecospat* package and converted to a convenient vector format with *vect*. We thus use only occurrences of tetraploid cytotypes to ensure a fair comparison between ranges, as diploids are absent from the North American continent. We restrict the regions of analysis (background) to the biomes where the species is present in each realm using a pipe of *dyplr* functions. The biomes selected are Temperate Broadleaf Forests (4), Temperate coniferous forests (5), Boreal Forests (6), Temperate Grasslands (8) Mediterranean Forests and woodlands (12) and Deserts (13). We further restrict the background to areas distant of less than 500 km from existing occurrences (this avoids including large regions of the Palearctic realm where the species does not occur in the native range). The background vector can then be used to *crop* and *mask* the climatic data. Note that the choice of the background area has important consequences on niche quantification, especially in the similarity tests presented below. The general advice in selecting the background is to include areas that are, or have been accessible to the species through dispersal during its evolutionary life (Barve et al. 2011). We recommend to perform a sensitivity analysis on the parameters chosen for the background selection to ensure the stability and reliability of results. The plots show the background area and overlaid occurrences to visualize the final data to be analyzed.

```

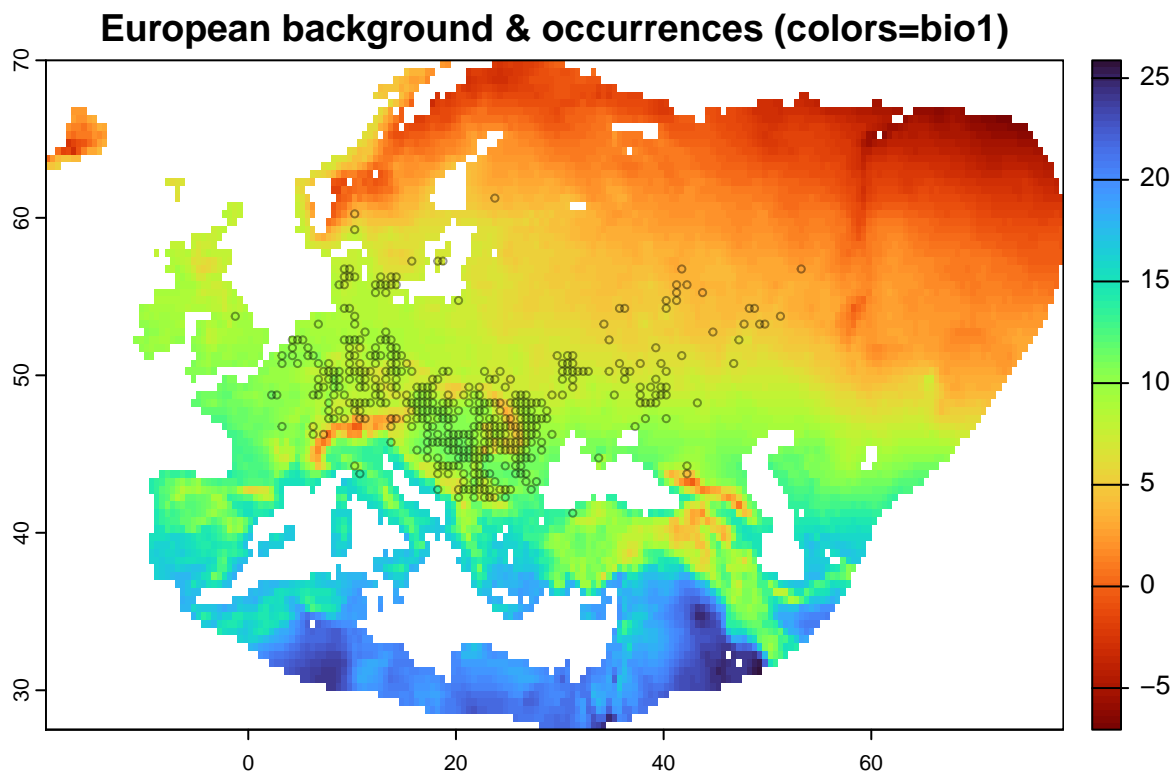
## data for Europe

occ.EU.df<-read.delim(system.file("extdata/Csto4xEU.txt", package="ecospat"))
occ.EU<-vect(occ.EU.df,geom=c("x", "y"),crs="+proj=longlat")

# biomes

buf<-buffer(occ.EU,1500000) %>% aggregate() # 1500km buffer around occurrences
bkg<-biomes %>% filter(BIOME_NUM%in%c(4,5,6,8,12,13)) %>% group_by(REALM) %>%
  summarize() %>% crop(buf) # biomes in the realm where Centaurea is present
clim.EU<-crop(clim.world,bkg)
clim.EU<-mask(clim.EU,bkg)
plot(clim.EU[[1]], col=viridis(100,option="turbo",direction=-1),
     main="European background & occurrences (colors=bio1)")
#plot annual temperature (bio1)
plot(occ.EU,cex=0.5,pch=21,alpha=0.4,add=TRUE)

```



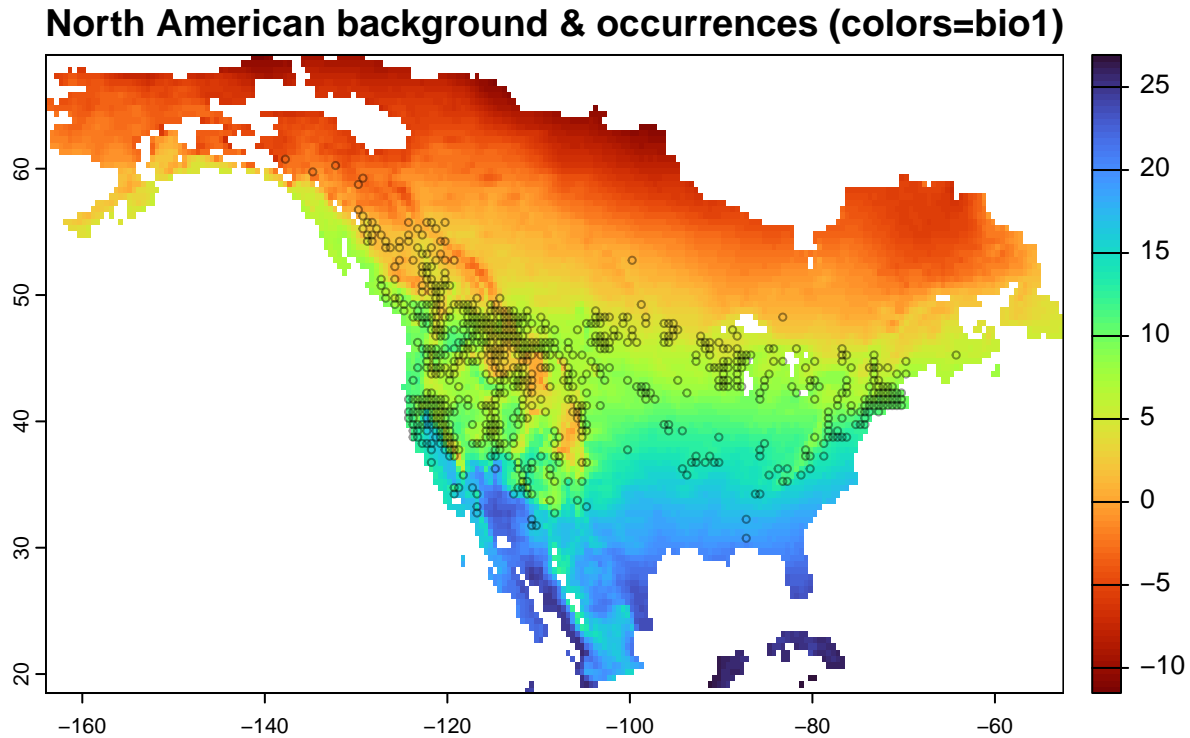
```
## data for North America

# occurrences

occ.NAM.df<-read.delim(system.file("extdata/Csto4xNAM.txt", package="ecospat"))
occ.NAM<-vect(occ.NAM.df,geom=c("x", "y"),crs="+proj=longlat")

# biomes

buf<-buffer(occ.NAM,1500000) %>% aggregate()
bkg<-biomes %>% filter(BIOME_NUM%in%c(2,4,5,6,8,12,13)) %>% group_by(REALM) %>%
  summarize() %>% crop(buf)
clim.NAM<-crop(clim.world,bkg)
clim.NAM<-mask(clim.NAM,bkg)
plot(clim.NAM[[1]], col=viridis(100,option="turbo",direction=-1),
  main="North American background & occurrences (colors=bio1)"
  #plot annual temperature (bio1)
plot(occ.NAM,cex=0.5,pch=21,alpha=0.4,add=TRUE)
```



## ii. Niche quantification

The niche quantification will be performed in two-dimensional space represented by the two first components (PC1 and PC2) of a principal component analysis (PCA). We chose to use two niche axes that maximize the variation contained in the original 19 bioclim variables from the worldclim dataset, but the user could as well chose two specific variables relevant for the ecology of the species based on expert assessment. Niche quantification in one dimension is presented in a following section. To calibrate the PCA on the full extent of climatic values present in both background areas, we first join the European and North American datasets using *merge* from *terra*. We then extract the climatic values for each pixels of the global background (*clim.glob*) and for each continental background (*clim.glob.EU*, *clim.glob.NAM*), and for each set of continental occurrences (*clim.occ.EU*, *clim.occ.NAM*) using *extract* from *terra*. The PCA is calibrated on the values of the global background using *dudi.pca* from *ade4*. The function *ecospat.plot.contrib* allows to plot the correlation circle indicating the correlation of original variables with the principal component axes with arrows. Here PC1 is mostly (but not exclusively) correlated with precipitation variables (bio12 to bio19) while PC2 is more correlated with temperature variables (bio1 to bio11). The inertia of each principal component (similar to the variance explained in a regression model context) is also indicated. Here the combined contribution of PC1 and PC2 explains almost 80% of the climatic variation in the original dataset. The calculation of PCA scores (*li* vectors in *dudi.pca* objects) for the pixels of the backgrounds (*scores.glob*, *scores.glob.EU*, *scores.glob.NAM*) and for the occurrences (*scores.occ.EU*, *scores.occ.NAM*) using *suprow* from *ade4* will be useful for the next step.

```
# extraction of climate values

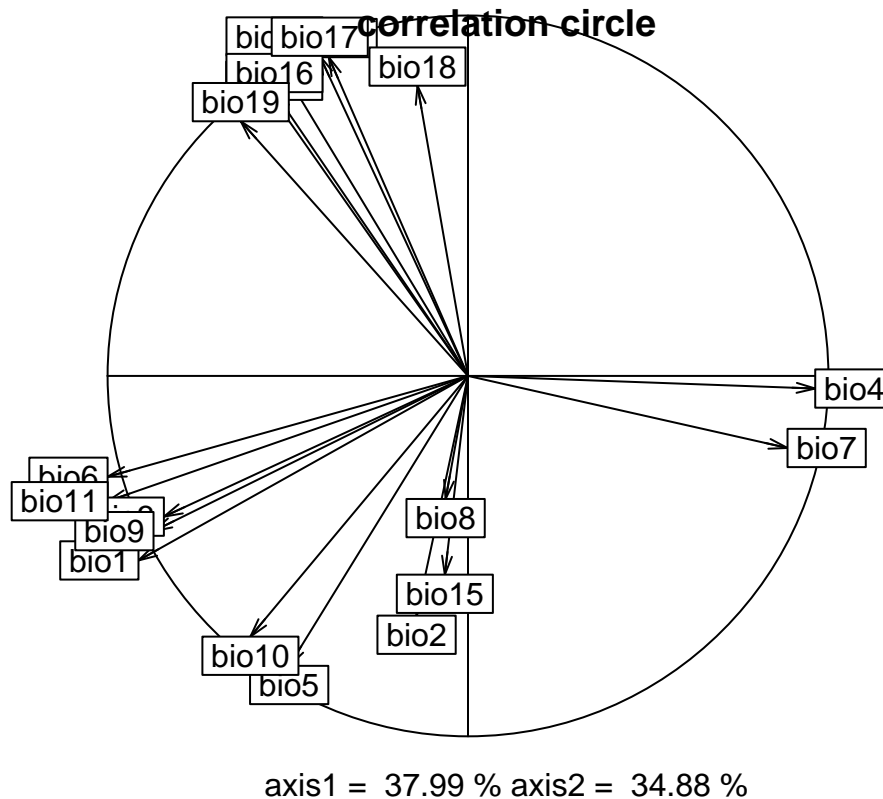
clim<-terra::merge(clim.EU,clim.NAM)
clim.glob<-na.omit(terra::values(clim))
```

```

clim.glob.EU<-na.omit(terra::values(clim.EU))
clim.glob.NAM<-na.omit(terra::values(clim.NAM))
clim.occ.EU<-terra::extract(clim.EU,occ.EU,ID=FALSE)
clim.occ.NAM<-terra::extract(clim.NAM,occ.NAM,ID=FALSE)

## PCA calibration
pca.env<-dudi.pca(clim.glob,scannf=FALSE,nf=2)
ecospat.plot.contrib(contrib=pca.env$co,eigen=pca.env$eig)

```



```

## PCA scores

scores.glob<-pca.env$li
scores.glob.EU<-suprow(pca.env,clim.glob.EU)$li
scores.glob.NAM<-suprow(pca.env,clim.glob.NAM)$li
scores.occ.EU<-suprow(pca.env,clim.occ.EU)$li
scores.occ.NAM<-suprow(pca.env,clim.occ.NAM)$li

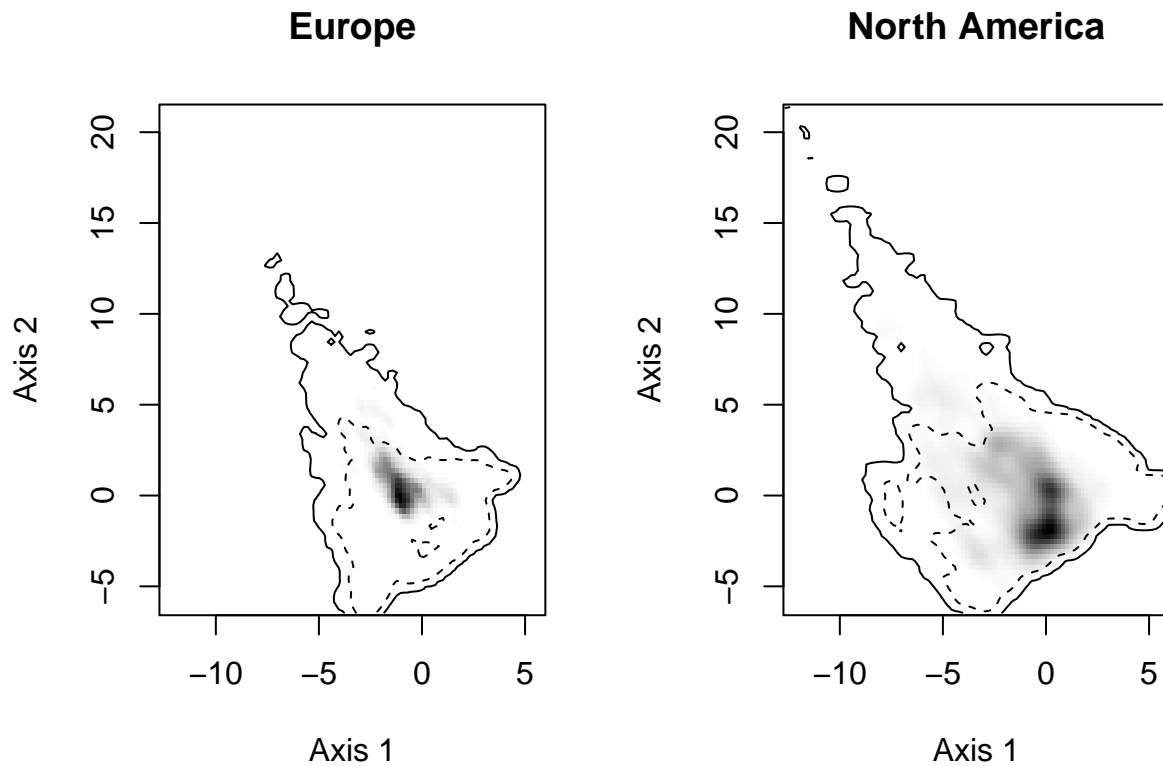
```

The following step is the core of the niche quantification analysis. The function *ecospat.grid.clim.dyn* first creates a gridded two-dimensional climatic space of  $R \times R$  pixels bounded by the minimum and maximum values of the PCA scores of the background points (*score.glob*). It then uses kernel density estimation functions to calculate the density of occurrence in each pixels. The result is a grid of occurrence densities in a two-dimensional climatic space, with high values indicating the core of species niche, low values indicating the margin of the niche, and zero values indicating conditions outside of the niche. For comparison between ranges, occurrences densities are standardized between zero and one (object *z.uncor*). An alternative version of occurrences densities is also calculated by dividing in each pixel the occurrence density by the density of

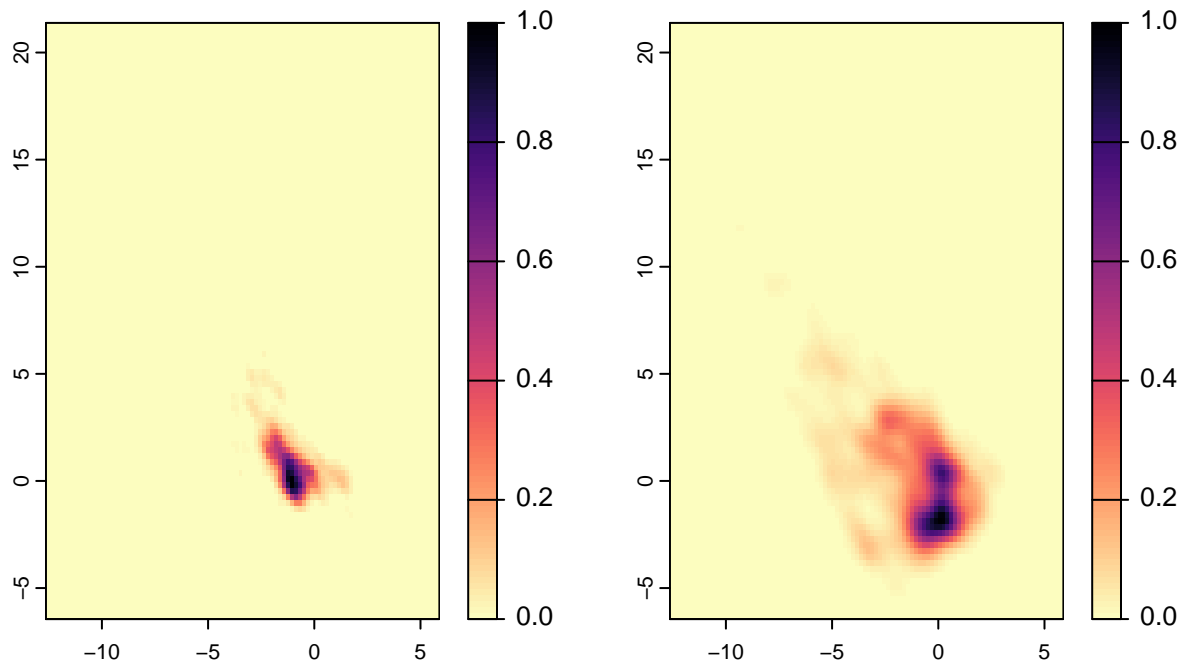
the corresponding climatic condition in the background (object `z.cor` return by *ecospat.grid.clim.dyn*; defined as occurrence occupancy in Broennimann et al. 2012). The function *ecospat.plot.niche* allows to plot the calculated niche as a gradient of occurrence densities (shades of gray pixels) along the two first PCA axes. The contours of all climates (black solid line) and 50% most common climates (black dashed line) are also shown. Alternatively, the niches can be directly plotted with the *plot* function as *rast* objects.

```
z.EU <- ecospat.grid.clim.dyn(glob=scores.glob,
                             glob1 = scores.glob.EU,
                             sp= scores.occ.EU,
                             R=100, th.sp=0,th.env=0, kernel.method = "ks")
z.NAM <- ecospat.grid.clim.dyn(glob=scores.glob,
                               glob1 = scores.glob.NAM,
                               sp= scores.occ.NAM,
                               R=100, th.sp=0,th.env=0, kernel.method = "ks")

par(mfrow=c(1,2))
ecospat.plot.niche(z.EU,title="Europe")
ecospat.plot.niche(z.NAM,title="North America")
```



```
plot(z.EU$z.uncor,col=viridis(100,option="magma",direction = -1))
plot(z.NAM$z.uncor,col=viridis(100,option="magma",direction = -1))
```



```
par(mfrow=c(1,1))
```

Now we can analyze the results of the niches quantification more in depth. First we want to know how much the two niches overlap. This can be achieved with the *ecospat.niche.overlap*, which calculates the Schoener's D and Hellinger I overlap indices. The parameter *cor=TRUE* allows to assess the overlap when correcting for the density of environmental conditions in the background. In our case, using *cor=FALSE*, we detect an overlap D of 0.36.

```
ecospat.niche.overlap(z.EU,z.NAM,cor=FALSE)
```

```
## $D
## [1] 0.3287533
##
## $I
## [1] 0.5597273
```

Now we can ask the question whether the overlap measured is higher than random. This can be achieved with *ecospat.niche.equivalency.test* and *ecospat.niche.similarity.test* functions, based on (Warren, Glor, and Turelli 2008). The test of niche equivalency asks whether the overlap observed between the two niches is significantly higher than the overlap that one would get if the occurrences between the two niches would be randomly reallocated. On the other hand, the test of similarity asks whether the overlap observed between the two niches is significantly higher than the overlap one would get if the two niches would be randomly moved in the background space. Both tests have been implemented in *ecospat* for correspondence with (Warren, Glor, and Turelli 2008), but the niche equivalency test almost always provide significant results even when the overlap is very low and is thus in general not informative. We thus only run the niche

similarity test, with parameters *rand.type* = 1 to randomize both niches in the background, *rep*=100 to set the number of replications, *intersection*=NA to perform the randomization in global background without removing marginal environments, and using *overlap.alternative* = "higher" as the alternative hypothesis for the test (i.e. a p-value < 0.05 indicate that the observed overlap is significantly higher than the simulated overlap from randomized niches). Here, with a p-value of 0.25, the niche similarity test indicate that the niche of the spotted knapweed in Europe and in North America do not significantly depart from similarity.

```
sim.test<-ecospat.niche.similarity.test(
  z.EU,z.NAM, rand.type=1,rep=100,intersection=NA, overlap.alternative="higher")

sim.test$obs$D
```

```
## [1] 0.2537814
```

We can now investigate whether some parts of the native niche are not occupied in the invaded range, and conversely, if the species occupies new conditions in the invasive range. This can be done with the *expansion*, *stability* and *unfilling* indices return by the *ecospat.niche.dyn.index* function. Stability and *expansion* correspond to the percentages of pixel of the invasive niche which conditions present, and not present in the native range, respectively. *Unfilling* corresponds to the percentage of the native niche which is not present in the invaded range. The results indicate that the invasive niche of the spotted knapweed expands largely beyond conditions present in the native niche, while only a small fraction of the native niche is not occupied in the invaded ranges. It is possible to plot the *expansion*, *stability* and *unfilling* indices in the environmental space using the function *ecospat.plot.niche.dyn*. The plot shows the niche stability in blue, niche expansion in orange, and niche unfilling in yellow. The solid green and red contour lines indicate the extent of environmental conditions that exists in the native and invaded ranges, respectively. The dotted contour lines indicate the quantile of most abundant environment in both ranges (here 50% most abundant, with *quant*=0.5). The densities of occurrences in the native range (argument *interest*=1) are displayed using gray shading. T

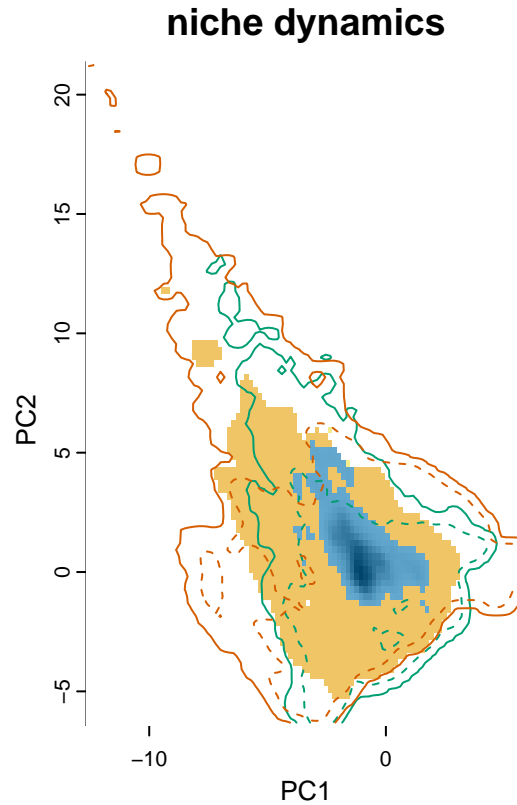
```
dyn<-ecospat.niche.dyn.index(z.EU,z.NAM,intersection=NA)
dyn$dynamic.index.w
```

```
##      expansion      stability      unfilling
## 0.5719737854 0.4280262146 0.0007058004
```

```
# colorblind friendly palette, Wong, B. (2011). Nature Methods
pal<-c("#D55E00", "#E69F00", "#F0E442", "#CC79A7", "#009E73", "#0072B2", "#56B4E9")

ecospat.plot.niche.dyn(z.EU,z.NAM,quant=0.5,title="niche dynamics",name.axis1 ="PC1",
  name.axis2="PC2",interest=1,col.unf="#F0E442", col.exp="#E69F00",
  col.stab="#0072B2", colZ1="#009E73", colZ2="#D55E00",transparency = 40)
```

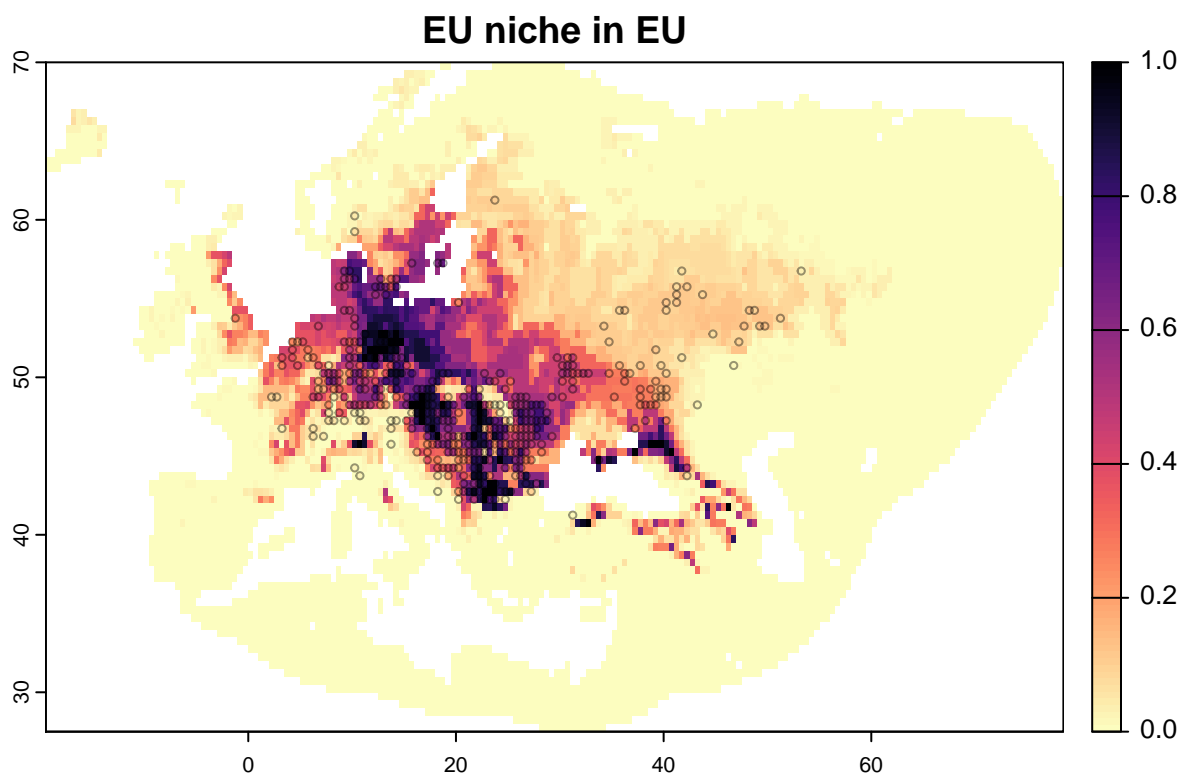




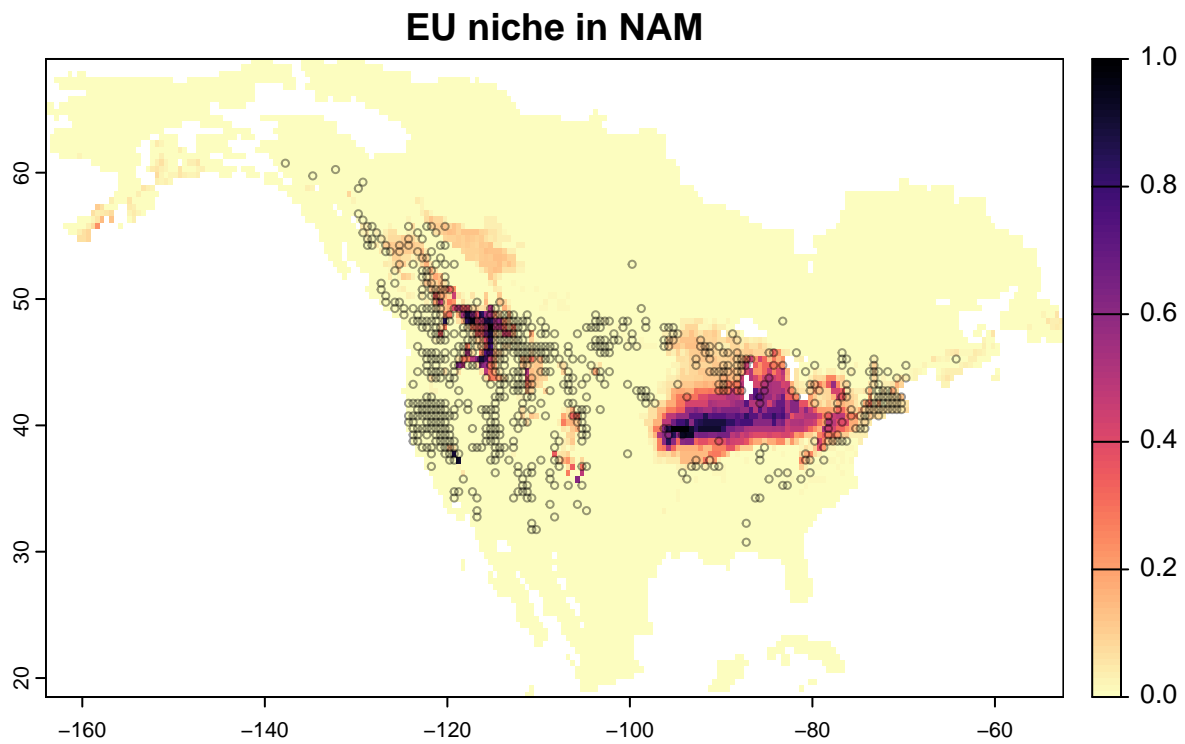
Finally, we might wonder where are the areas predicted with a high density of occurrence by the niche quantification models. Indeed, every pixel in the geography corresponds to a predicted density of occurrence in the environmental space. The function `ecospat.niche.zProjGeo` allows to do that. It is possible to project the density of occurrence of one niche (`z` object) across its own range of calibration (e.g. for Europe using `zproj=NULL` and `env=clim.EU`) but also to project the density of occurrence to another range, here for example to project the European niche of the spotted knapweed in North America.

Similarly, we can also project niche dynamic indices in geography using the function `ecospat.niche.dynIndexProjGeo`. The argument `proj` controls the range of projection (i.e. `proj=2` for the North American range), and `env` sets the environmental `rast` object used for the projection. The results here show that a large part of the North East and Midwest, as well as smaller fragmented areas of the Rocky Mountains in the West have climatic conditions corresponding to the native niche of spotted knapweed, but that large areas, especially in the South correspond to an expansion of the niche.

```
geo.z.EU<-ecospat.niche.zProjGeo(z.EU,zproj=NULL,env=clim.EU)
plot(geo.z.EU,main="EU niche in EU",col=viridis(100,option="magma",direction=-1))
plot(occ.EU,add=TRUE,cex=0.5,pch=21,alpha=0.4)
```

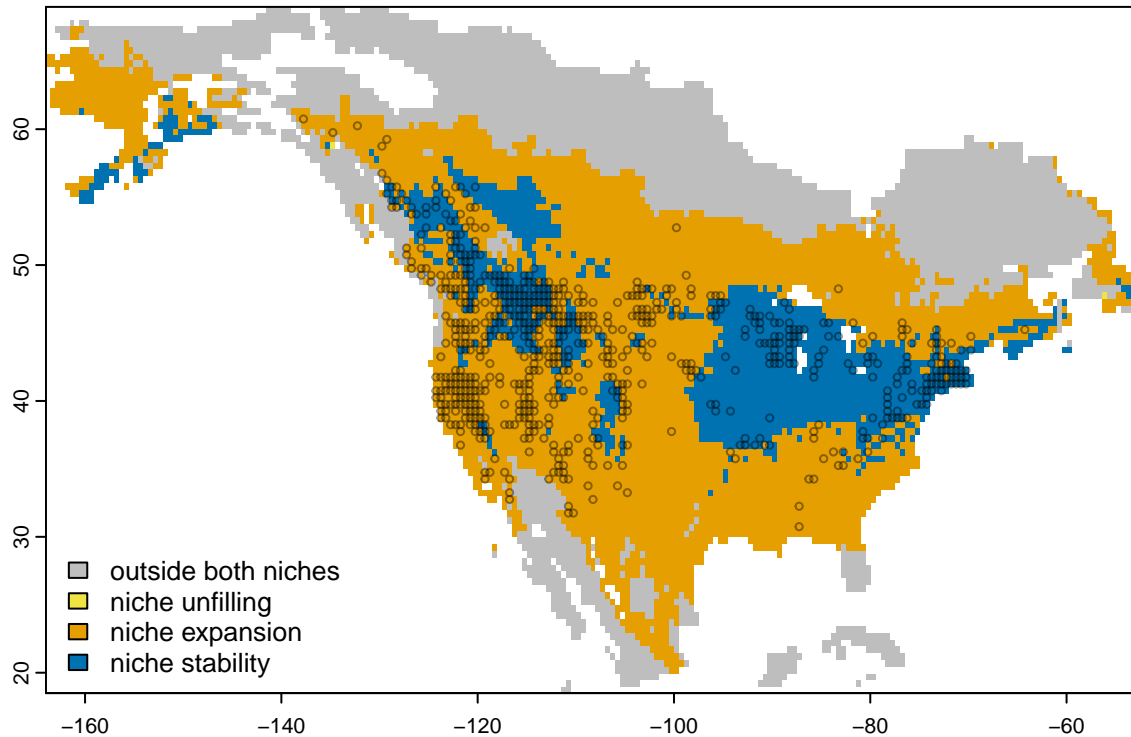


```
geo.z.EUtoNAM<-ecospat.niche.zProjGeo(z.EU,zproj=z.NAM,env=clim.NAM)
plot(geo.z.EUtoNAM,main="EU niche in NAM",col=viridis(100,option="magma",direction =-1))
plot(occ.NAM,add=TRUE,cex=0.5,pch=21,alpha=0.4)
```



```
geo.dyn<-ecospat.niche.dynIndexProjGeo(z.EU,z.NAM,proj=2,env=clim.NAM)

plot(geo.dyn,col = c("grey", "#F0E442", "#E69F00", "#0072B2"),
     plg=list(legend=c("outside both niches", "niche unfilling",
                      "niche expansion", "niche stability"), x="bottomleft"))
plot(occ.NAM,add=TRUE,cex=0.5,pch=21,alpha=0.4)
```



### iii. References

- Broennimann, O., M. C. Fitzpatrick, P. B. Pearman, B. Petitpierre, L. Pellissier, N. G. Yoccoz, W. Thuiller, et al. 2012. "Measuring Ecological Niche Overlap from Occurrence and Spatial Environmental Data." *Global Ecology and Biogeography* 21 (4): 481–97. <https://doi.org/10.1111/j.1466-8238.2011.00698.x>.
- Broennimann, O., U. A. Treier, H. Muller-Scharer, W. Thuiller, A. T. Peterson, and A. Guisan. 2007. "Evidence of Climatic Niche Shift During Biological Invasion." *Ecology Letters* 10 (8): 701–9. <https://doi.org/10.1111/j.1461-0248.2007.01060.x>.
- Guisan, A., B. Petitpierre, O. Broennimann, C. Daehler, and C. Kueffer. 2014. "Unifying Niche Shift Studies: Insights from Biological Invasions." *Trends in Ecology & Evolution* 29 (5): 260–69. <https://doi.org/10.1016/j.tree.2014.02.009>.
- Mráz, Patrik, Robert S. Bouchier, Urs A. Treier, Urs Schaffner, and Heinz Müller-Schärer. 2011. "Polyploidy in Phenotypic Space and Invasion Context: A Morphometric Study of *Centaurea Stoebe*s.l." *International Journal of Plant Sciences* 172 (3): 386–402. <https://doi.org/10.1086/658151>.
- Petitpierre, B., C. Kueffer, O. Broennimann, C. Randin, C. Daehler, and A. Guisan. 2012. "Climatic Niche Shifts Are Rare Among Terrestrial Plant Invaders." *Science* 335 (6074): 1344–48. <https://doi.org/10.1126/science.1215933>.
- Treier, Urs A., Olivier Broennimann, Signe Normand, Antoine Guisan, Urs Schaffner, Thomas Steinger, and Heinz Müller-Schärer. 2009. "Shift in Cytotype Frequency and Niche Space in the Invasive Plant *Centaurea Maculosa*." *Ecology* 90 (5): 1366–77. <https://doi.org/10.1890/08-0420.1>.
- Warren, D. L., R. E. Glor, and M. Turelli. 2008. "Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution." *Evolution* 62 (11): 2868–83. <https://doi.org/10.1111/j.1365-3113.2008.04100.x>.