

Exercise ! Run your first SDM !

1/ Open RStudio

2/ First of all, make sure that 'ncdf4', 'raster', 'dismo' and 'gbm' R packages are installed on your RStudio

3/ Load the directory EXERCICE from the GitHub page

<https://github.com/SCAR/EGABICourse19>

4/ Download the folder, and set it as your working directory (you can create a new RStudio project)

5/ Explore the files you have in your folders

data : contains occurrences and environmental descriptors datasets

- Occurrences data for the Antarctic sea urchin *Sterechinus neumayeri*.

You can find the metadata at :

https://ipt.biodiversity.aq/resource?r=echinoids_occurrences_southern_ocean

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5673840/>

- You also have a folder containing an ensemble of environmental covariates at the extent of the Southern Ocean (that were downloaded at :

https://data.aad.gov.au/metadata/records/environmental_layers). You can also access to their metadata in this link.

- You have an ascii file (KDE.asc), which is the KDE layer (Kernel Density Estimator)

We will see exactly what it is in following courses ! you don't really need to use it right now ☺

scripts : contains R files

Run_yOur_SDM.R

This is the main script you will use. **Open it.** It contains all the things you need to generate, calibrate your model and create the outputs.

Functions_gbm.R

R file that contains functions modified from the 'gbm' R package to run BRT models and provide models outputs. You don't need to explore/understand this file, it is just called automatically by the *run_yOur_SDM.R* code. It is necessary to have it close to the *run_yOur_SDM.R* file to make this last one work.

clock4_crossValidation.R

R file that contains a function to partition the occurrence dataset into 4 groups following a random partition of the space into 4 areas. Then, each time a model is run, three of these four groups are used to train the model and the remaining one to test the model predictions. All combinations are generated, providing 4 model results for each random partitioning of the occurrence dataset. Same, this function it is automatically called by the *run_yOur_SDM.R* code

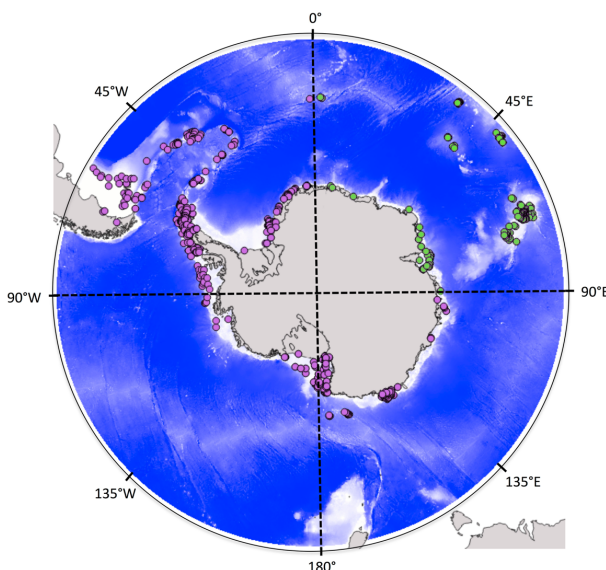


Fig. Four groups are randomly defined, and 3 groups are used to train the model (pink), while the remaining is used to test model predictions (green, data spatially independent to test the model)

calibration_BRT.R

R file that was used to define the inner parameters of BRT (learning rate *lr*, bag fraction *bf*, tree complexity *tc*). BRT being the SDM algorithm that was chosen in this exercise to compute the models (BRT= Boosted Regression Trees). Following Elith et al. (2008) and its supplementary material. This file compares model outputs calibrated with different values. The combination of these 3 parameters that reduces the predictive deviance was selected. You have nothing to do with this file, you can explore it if you want to see the calibration procedure or use it to calibrate your models in the future.

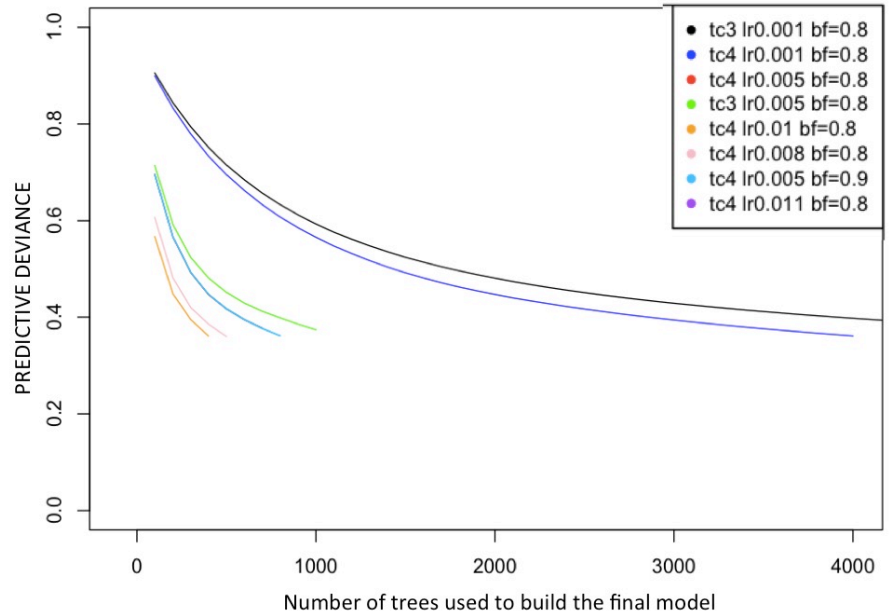


Figure : Trials of model and comparisons with different set of parameters. Figure computed with the *calibration_BRT.R* script

6/ Select all the code of the *run_yOur_SDM.R* and press 'run' on R
The process should take few minutes.

Nothing happened ? this is normal !! Most of the outputs are commented (beginning by #), they are not launched, you need to uncomment it and launch them.

Explore a bit the code, try to understand its different sections. Try to uncomment some lines and generate plots, tables...to get your model outputs (we will see them in details in the following course).

Remarks :

All the references given in the previous course on model calibration and quoted in this exercise can be found in <https://github.com/SCAR/EGABICourse19> in the 'Literature' folder don't hesitate to also explore them to better understand the code.

EXTRA :

You can explore the following links if you have time

- 'Biomod2' R package content and vignettes

- 'dismo' R package tutorial

<https://cran.r-project.org/web/packages/dismo/vignettes/sdm.pdf>

- 'SDMPlay' R package and its vignette

<https://cran.r-project.org/web/packages/SDMPlay/vignettes/my-vignette.html>

- Browse « run species distribution models » on your computer, have a look at tutorial, explanations...etc