**Title: Vignette for running N-SDM**

**Author: Antoine Adde (**[**aadde@unil.ch**](mailto:aadde@unil.ch)**)**

**Date: 04.05.2022**

1. Welcome

The aim of this vignette is to help you getting started with running N-SDM on HPC environments. Before running N-SDM, you should make sure you have an appropriate [set up with](https://prioritizr.github.io/cibio-workshop/introduction.html#setup) the following requirements:

* A cluster with Slurm workload manager
* Available modules: gcc; r; proj; perl; curl; geos; gdal
* installed R packages: c('data.table', 'stringi', 'stringr', 'plyr', 'readxl','writexl', 'parallel', 'sp', 'raster', 'rgdal', 'zoo','fst','tools', 'glmnet', 'gam', 'mgcv', 'randomForest', 'RRF', 'lightgbm', 'ranger', 'maxnet','caret', 'ROCR','ecospat', 'chron','ggpubr')

The nsdm.zip file downloadable from XXX should be unzipped in you work directory. It contains four main folders: *Data* (species and covariate data), *Scripts* (codes and functions), *Scratch* (where intermediate N-SDM outputs will be written), *Save* (where final N-SDM outputs will be saved).

1. Getting started
   1. Data

In this example run we will model the current and future distributions of three test species (Larix decidua, Capra Ibex and Cantharellus cibarius) in Switzerland using a suite of more than 100 candidate covariates. Following the spatially nested modelling framework of N-SDM, we will use European wide GBIF occurrence data at the global level and Swiss wide InfoSpecies occurrence data at the local level. Only 1km resolution bioclimatic (CHELSA) covariates will be used at global level whereas 8 environmental categories of 100-m resolution covariates will be considered at the local level. Meta information on these data can be found in the supplementary material of the main manuscript. Some of these covariates are equipped with focal windows (e.g., land use and cover) and others (e.g., bioclimate) are temporally dynamics.

* 1. N-SDM settings

N-SDM setting edit can be done by modifying the .csv file *workdirectory/scripts/nsdm-workshop/main/settings/settings.csv*. Be careful when saving *settings.csv* to use “;” as the delimiter. In the same directory *workdirectory/scripts/nsdm-workshop/main/settings/* the .xlsx file *param-grid.xslx* allows specifying the grid for hyperparameter tunning. The .xlsx file *expert-table.xslx* allows for expert-based prefiltering of taxon-specific candidate covariates.

1. Running N-SDM

Position yourself at *workdirectory/scripts/nsdm-workshop/main*, which is the location of the main N-SDM bash file (nsdm.sh). We encourage you running it in a background no hangup mode to prevent the command from being aborted automatically if logging out or exiting the shell, such as: *nohup bash nsdm.sh &*.

1. Outputs

Explore saved outputs *workdirectory/save/nsdm-workshop*. Maps, sacct statistics, ODMAP protocol etc…