READ ME

Study: “Predicting predator-prey interactions in terrestrial endotherms using random forest”

The files include a folder with data files (.rds format) and a folder with R scripts. File paths will need to be adjusted once downloaded, and can be found in the scripts by searching for “###”.

DATA FOLDER

The data folder contains 4 files.:

* GloBIplus\_Int20EVs.RDS contains the enhanced global interaction records,
* allNon\_sameCont\_1.RDS and allNon\_sameCont\_2.RDS contain the enhanced global non-interaction records – these two datasets need to be combined (rbind) and renamed allNon\_sameCont.RDS or Non for running scripts,
* allperms\_cut2\_20EVs.RDS contains the interaction and non-interaction records for the seven focal predators from the Simpson Desert.

Each of these files include all ecomorphological traits and phylogenetic eigenvectors used in analyses.

FUNCTIONS FOLDER

The functions folder contains a file (all\_functions\_ranger.R) with functions required by the script files.

SCRIPT FOLDER

The script folder contains three more folders.

1. OPTIMIZATION OF RANDOM FOREST MODELS

contains scripts that identify optimal parameter values for 6 different random forest models (that differ in terms of the trait data used).

1. APPLIED TO GLOBAL AND SIMPSON DESERT DATASETS

contains scripts that apply the 6 optimised model (as determined in step 1)) to the enhanced global interaction/non-interaction data and the Simpson Desert data (for the seven focal predators).

1. DATA QUALITY MANIPULATION AND MODEL PERFORMANCE

contains two more folders of scripts that test the effect of modifying training data quality on model performance (when training on the enhanced global data and applied to the Simpson Desert data).

1. RECORDREMOVAL&REPLACE\_MODELPERFORMANCE

contains scripts that test the effect of removing records or switching interaction records to non-interactions (false negatives) on model performance. These modification to training data quality were made to different subsets of the data including: the whole dataset, focal prey species only, focal predator species only, and non-focal species (non-Simpson Desert) only.

1. CORRELATION&CHANGE\_PROBABILITY

contains scripts testing the effects of modifying the focal-predator component of training data (removing records or switching interactions to noninteractions) on (a) relative suitability of different prey for each predator and (b) the mean probability assigned to potential prey for each predator.