EGU22-13251, updated on 18 Aug 2022
https://doi.org/10.5194/egusphere-egu22-13251
EGU General Assembly 2022

The ‘rtry’ R package for preprocessing plant trait data

Olee Hoi Ying Lam[1,2], Susanne Tautenhahn[1], Gabriel Walther[3,4], Gerhard Boenisch[1], Pramod Baddam[1], and Jens Kattge[1,4]

[1]Max Planck Institute for Biogeochemistry, Jena, Germany
[2]University of Wisconsin-Madison, USA
[3]University of Jena, Germany
[4]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Germany

From evolutionary biology, functional ecology, earth system modelling to landscape management, plant trait data are used to determine how the plants respond to the environmental factors and can act as indicators of ecosystem functions. In 2007, the TRY initiative was founded as an integrated database of trait data and all additional attributes relevant to understanding and interpreting a given trait value. Since then, the TRY database has integrated more than 400 datasets, including both original datasets and collective databases.

Due to the unique long table structure, the relevant information (e.g. trait names, species names, ancillary data representing context information, units of trait data, and identifiers for duplicates and outliers) for trait data filtering is stored at different places of the released TRY data. This makes the process to find all relevant information to select or remove trait data not straightforward without knowledge of the inherent data structure.

The ‘rtry’ package is an R package that provides a set of easily applicable functions for the basic steps of data preprocessing and is designed in particular to support the data exploration and removal of the plant trait data, taking advantage of the features of trait data released from the TRY database. This package is supposed to be applicable without advanced knowledge of the data structure released from TRY or the R software. Most importantly, despite the ‘rtry’ package being developed to support the application of plant trait data received via the TRY database, it is also applicable to other trait data.