Supplement of Geosci. Model Dev., 12, 613–628, 2019
https://doi.org/10.5194/gmd-12-613-2019-supplement

Geoscientific Model Development
Open Access
EGU

Supplement of

Topological data analysis and machine learning for recognizing atmospheric river patterns in large climate datasets

Grzegorz Muszynski et al.

Correspondence to: Grzegorz Muszynski (gmuszynski@lbl.gov)

Supplement Text

1 Implementation details

This supplement describes technical details of a key building blocks that have been used in the implementation of our atmospheric river pattern detection method. This method consists of two stages: feature extraction—topological data analysis (TDA) algorithm and binary classification—support vector machine (SVM) classifier. Below we provide details about the actual implementation—*e.g.*, the data structures and programming languages used, the external software and packages/libraries used for each stage of the method.

1.1 TDA Algorithm

The TDA algorithm employs the union-find (often called disjoint-set) data structure to extract topological feature descriptors (connected components)—described in the main manuscript. The data structure and algorithm are implemented in C++ programming language. However, this algorithm uses the very portable language, it is dependent on the custom structures and packages/libraries in The Toolkit for Extreme Climate Analysis (TECA). The TECA is a collection of climate analysis algorithms for extreme event detection and tracking implemented in a scalable parallel framework. More details about the TECA installation can be found in the paper of (Prabhat, et al., 2015) and user manual (https://github.com/LBL-EESA/TECA/blob/master/doc/teca_users_guide.pdf). The code of our algorithm needs to merged with the TECA code and complied all together with it.

1.2 SVM Classifier

The binary classification task utilizes SVM classifier (here, based on kernel functions). It employs the C-Support Vector Classifier (C-SVC) from Python scikit-learn that the implementation is based on libsvm. More details about the installation and the use can be found on the sckit-learn and the libsvm websites (https://scikit-learn.org/stable/modules/svm.html, https://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html).