Supporting Information

Advancing Micro-Nano Supramolecular Assembly Mechanisms of Natural Organic Matter by Machine Learning for Unveiling Environmental Geochemical Processes

Ming Zhang [1], Yihui Deng [2], Jiawei Zhou [3], Jing Gao [2], Daoyong Zhang [1,*], Xiangliang Pan [2,*]

[1] College of Geoinformation, Zhejiang University of Technology, Hangzhou, 310014, P. R. China

[2] College of Environment, Zhejiang University of Technology, Hangzhou, 310014, P. R. China

[3] College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, P. R. China

Corresponding authors:

Daoyong Zhang

E-mail address: zhangdaoyong@zjut.edu.cn

Xiangliang Pan

E-mail address: panxl@zjut.edu.cn

Text S1. Additional descriptions about data acquisition

The literature search was conducted based on the Web of Science Core Collection database. The literature search mainly focused on commonly used machine learning algorithms. The search formula is TI = ("machine learning" OR "deep learning" OR "supervised learning" OR "deep neural networks" (DNN) OR "xgboost" OR "long short term memory" (LSTM) OR "transformer" OR "random forest" OR "artificial neural network" (ANN) OR "support vector machine" OR "decision tree" OR "convolutional neural network" (CNN) OR "recurrent neural network" (RNN) OR "graph neural network" (GNN) OR "DNN" OR "CNN" OR "RNN" OR "GNN" OR "LSTM" OR "ANN") AND TI = ("supramolecular" OR "protein*" OR "peptide*" OR "lipid*" OR "polysaccharide*" OR "DNA" OR "RNA" OR "viral*" OR " dissolved organic matter" OR "natural organic matter" OR "solid organic matter" OR "humic*" OR "EPS" OR "biopolymer"). Document type was limited to article. The retrieval date was September 30, 2024. A total of 3482 articles were obtained. The software VOSviewer v.1.6.20 and Pajek v.5.19 were used to visualize and analyze the keyword co-occurrence in the literature.