EGU21-2046
https://doi.org/10.5194/egusphere-egu21-2046
EGU General Assembly 2021

Oceanographic data reconstruction using machine learning techniques

Hrvoje Kalinić[1], Zvonimir Bilokapić[1], and Frano Matić[2]
[1]University of Split, Faculty of Science, Split, Croatia (hrvoje.kalinic@pmfst.hr)
[2]Institute of Oceanography and Fisheries, Split, Croatia

In certain measurement endeavours spatial resolution of the data is restricted, while in others data have poor temporal resolution. Typical example of these scenarios come from geoscience where measurement stations are fixed and scattered sparsely in space which results in poor spatial resolution of acquired data. Thus, we ask if it is possible to use a portion of data as a proxy to estimate the rest of the data using different machine learning techniques. In this study, four supervised machine learning methods are trained on the wind data from the Adriatic Sea and used to reconstruct the missing data. The vector wind data components at 10m height are taken from ERA5 reanalysis model in range from 1981 to 2017 and sampled every 6 hours. Data taken from the northern part of the Adriatic Sea was used to estimate the wind at the southern part of Adriatic. The machine learning models utilized for this task were linear regression, K-nearest neighbours, decision trees and a neural network. As a measure of quality of reconstruction the difference between the true and estimated values of wind data in the southern part of Adriatic was used. The result shows that all four models reconstruct the data few hundred kilometres away with average amplitude error below 1m/s. Linear regression, K-nearest neighbours, decision trees and a neural network show average amplitude reconstruction error of 0.52, 0.91, 0.76 and 0.73, and standard deviation of 1.00, 1.42, 1.23 and 1.17, respectively. This work has been supported by Croatian Science Foundation under the project UIP-2019-04-1737.