

EGU21-12439

<https://doi.org/10.5194/egusphere-egu21-12439>

EGU General Assembly 2021

© Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.



Supervised and unsupervised learning algorithms for extreme summer temperature prediction

Maria Pyrina, Eduardo Zorita, and Sebastian Wagner

Helmholtz Zentrum Geesthacht, Coastal Research, Geesthacht, Germany (maria.pyrina@hzg.de)

In order to predict the timing of extreme summer seasons in terms of 2 meters surface temperature (t2m), two algorithms are applied in the output of a global paleoclimate simulation. The model simulation is conducted with the Max Planck Institute Earth System Model configuration for paleoclimate (MPI-ESM-P). Global model output is provided for the period 0–1999 AD on a horizontal resolution of approximately 187 km (1.875°× 1.875°longitude by latitude grid). The t2m extremes are defined as the events with mean summer temperature higher than the 95th percentile of the training period (0-1970 AD) and are calculated separately for each grid point of the European region between 35°N-70°N, 10°W-30°E. The algorithms are trained only with data from the training period and are set to predict the summer t2m extremes of the test period 1971-1999 AD. The predictor data used for fitting the algorithms are chosen based on their known influence as boundary forcings of European summer climate. The predictor variables include springtime sea surface temperature (SST) from the NA region (0°-76°N, 85°W-30°E) and springtime European soil moisture (SM). The skill of the predictions is evaluated based on the extremal dependence index (EDI), which depends on the hit rate and false alarm rate. The EDI values vary between -1 and 1, where 1 is the skill of a perfect forecast. The first algorithm tested is a supervised learning algorithm, which is based on a random forest classifier (RF). RF predicts the highest EDI values over Scandinavia, Scotland and around the Mediterranean region (EDI>0.5), with the SST predictor being the main contributor to that skill. The second algorithm tested, is an autoencoder neural network (AE) that learns data codings in an unsupervised manner. AE surpasses the RF skill above most European regions and predicts the highest EDI values over the southeast Mediterranean, Central Europe, and the British Isles. The AE neural network is also trained to predict the absolute value of the extreme t2m events. The skill of reproducing the absolute value of the target t2m extremes is evaluated with the Mean Absolute Error (MAE), only for those extreme events that are reproduced by the AE prediction. The MAE values for the southeast Mediterranean region, Central Europe, and the British Isles are around 2 °C, 2.5 °C, and 1.5 °C, respectively. We have demonstrated the possibility of predicting a season in advance the occurrence of extreme summer t2m using an AE neural network. The AE neural network was tested in the virtual reality of a model simulation. The second step will be the application of the trained network on observational data.