EGU21-218
https://doi.org/10.5194/egusphere-egu21-218
EGU General Assembly 2021

Comparison of wavelet and machine learning methods for regional drought prediction

Gilbert Hinge[1] and Ashutosh Sharma[2]
[1]Assam Downtown University, Guwahati, India (gilberthinge@gmail.com)
[2]Indian Institute of Technology Roorkee (ashutoshsharma93@outlook.com)

Droughts are considered as one of the most catastrophic natural disasters that affect humans and their surroundings at a larger spatial scale compared to other disasters. Rajasthan, one of India's semiarid states, is drought inclined and has experienced many drought events in the past. In this study, we evaluated different preprocessing and Machine Learning (ML) approaches for drought predictions in Rajasthan for a lead-time of up to 6 months. The Standardized Precipitation Index (SPI) was used as the drought quantifying measure to identify the drought events. SPI was calculated for 3, 6, and 12-month timescales over the last 115-year using monthly rainfall data at 119 grid stations. ML techniques, namely Artificial Neural Network (ANN), Support Vector Regression (SVR), and Linear Regression (LR), were used to evaluate their accuracy in drought forecasting over different lead times. Furthermore, two data processing methods, namely the Wavelet Packet Transform (WPT) and Discrete Wavelet Transform (DWT), have also been used to enhance the aforementioned ML models' predictability. At the outset, the preprocessed SPI data from both the methods were used as inputs for LR, SVR, and ANN to form a hybrid model. The hybrid models' drought predictability for a different lead-time was evaluated and compared with the standalone ML models. The forecasting performance of all the models for all 119 grid points was assessed with three statistical indices: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Nash-Sutcliffe Efficiency (NSE). RMSE was used to select the optimal model parameters, such as the number of hidden neurons and the number of inputs in ANN, and the level of decomposition and mother wavelet in wavelet analysis. Based on these measures, the coupled model showed better forecasting performance than the standalone ML models. The coupled WPT-ANN model shows superior predictability for most of the grid points than other coupled models and standalone models. All models' performance improved as the timescale increased from 3 to 12 months for all the lead times. However, the model performance decreased as the lead time increased. These findings indicate the necessity of processing the data before the application of any machine learning technique. The hybrid model's prediction performance also shows that it can be used for drought early warning systems in the state.