Evapotranspiration Estimation Using Multimodel Ensemble

Evapotranspiration (ET) plays a crucial role in the Earth's hydrological cycle, influencing water resource management, agricultural productivity, and ecosystem dynamics. Accurate estimation of ET is essential for sustainable water resource planning and management. Multiple models based on different physical principles were developed. Past studies have reported varying accuracies, with performance of the models differing across sites, climate, biomes etc. and a single model could not be identified best under all conditions. This also limits the use of ET products in various applications. In recent years, ensemble modeling techniques have emerged as a promising approach to enhance the accuracy of ET predictions by combining the strengths of multiple individual models. Studies indicate that the ET estimates are improved even when basic methods of mean or weighted average techniques are used for the ensemble. Integrating machine learning techniques into ensemble framework utilizes the physical nature of model as well as the predictive power of machine learning algorithms.

Against this backdrop, this study aims to develop an ensemble ET over Indian region. The primary objective of the study is to improve the accuracy and reliability of ET predictions by utilizing the strengths of multiple modeling approaches. Three popular ET models for creating ensemble models namely: Priestley Taylor – Jet Propulsion Lab (PT-JPL), Soil Plant Atmosphere and Remote Sensing Evapotranspiration (SPARSE – Layer and Patch), and Surface Temperature Initiated Closure (STIC) belonging to different categories of ET models are used to create the ensemble. The study tests the ensemble based on mean, Bayesian Model Averaging (BMA), and k-Nearest Neighbor (kNN). The mean based ensemble in simple mean of all the four models. BMA is weighted average of the models where the weights are determined by the posterior probabilities of the models. kNN is a machine learning based ensemble where the model is trained using the available model output and the measured ET. The individual models are used to estimate the daily ET values initially and given as input to the ensemble models. ET at field scale and 1 km scales are used to estimate the ensembles. The models were tested at seven sites in India which is cropland and grassland.

The overall result indicates that the ensemble models have performed better compared to that of individual models. When all the sites are taken together, the RMSE of individual models are 63.02 Wm^{-2}, 66.39 Wm^{-2}, 63.31 Wm^{-2} and 86.88 Wm^{-2} for the models PTJPL, SPARSE Layer, SPARSE Patch and STIC simultaneously. Whereas the mean, BMA and kNN based ensembles have the RMSE of 54.06 Wm^{-2}, 52.92 Wm^{-2} and 36.52 Wm^{-2} simultaneously. KNN based ensemble have performed so well at field scale reducing the RMSE to almost half. To test the spatial extrapolation of the model, leave one out technique was used, where the training is done by data from all sites except one and testing on that site, which yielded a positive result that the ensembles are performing better compared to that of individual model. When tested at 1km scale the RMSE improved for ensemble models but there was no significant increase in the performance of kNN based ensemble compared to that of mean and BMA based ensembles. Also the leave one out method suggest that the training using the data from the same landcover can improve the ET predictions.