

Author: Md. Abul Kalam Azad, Javed Mallick, Abu Reza Md. Towfiqul Islam, Kurratul Ayen & Md. Hasanuzzaman

Title: Estimation of solar radiation in data-scarce subtropical region using ensemble learning models based on a novel CART-based feature selection

Abstract:

Solar radiation estimation is essential with increasing energy demands for industrial and agricultural purposes to create a cleaner environment, negotiate climate change impacts, and attain sustainable development. However, the maintenance and operation of solar radiation measurements are costly due to the lack of pyranometers or their failure; hence, obtaining reliable solar radiation data is challenging in many subtropical regions. Despite its importance, a few studies use machine learning algorithms for solar radiation estimation in Bangladesh. To this end, this study contributes to filling the gap twofold. First, we presented the potentials of ensemble models, such as Bagging-REPT (reduced error pruning tree), random forest (RF), and Bagging-RF, which were compared to three standalone models, namely, Gaussian process regression (GPR), artificial neural network (ANN), and support vector machine (SVM), for estimating daily global solar radiation in three Bangladeshi regions. Second, we explore the optimal input parameters influencing solar radiation change at the regional scale using a classification and regression tree (CART)-based feature selection tool. Satellite-derived ERA5 reanalysis and NASA POWER project datasets were used as input parameters. The performance of the models was compared using performance evaluation metrics like correlation coefficient (r), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), index of agreement (IA), and Taylor diagram. Results suggested that the RF model performed 5.47–37.22% better than the standalone models in estimating daily solar radiation at Chuadanga in terms of RMSE. Besides, the other ensemble model Bagging-RF showed 14.9–25.03% and 11.46–30.97% greater performances in Dinajpur and Satkhira than the conventional models in RMSE metric. Besides, this study may provide knowledge to the policymakers to make critical judgments on future energy yield, efficiency, productivity, and operation, which are essential elements for investments and solar energy conversion applications in the subtropical areas of the world.

Keywords: Solar Radiation Estimation, Machine Learning, Ensemble Models, Bangladesh, Energy Efficiency, Climate Change, Feature Selection

DOI: <https://doi.org/10.1007/s00704-023-04638-3>