Title: Hybrid ensemble machine learning for groundwater level prediction in data-scarce northern Bangladesh

Author names: Md. Abdullah-Al Mamun, Mst. Nazneen Aktarb, Javed Mallickc, Saeed Alqadhic , Abu Reza Md. Towfiqul Islam

Abstract: Groundwater depletion threatens sustainable urban water supply in northern Bangladesh, where overextraction, land-use change, and climate variability accelerate groundwater level decline. This study proposes a hybrid ensemble machine learning framework to improve short-term groundwater-level prediction by incorporating hydro-meteorological and anthropogenic variables. Ten individual algorithms and six hybrid ensemble models were evaluated across eight observation wells in Rangpur and Dinajpur under contrasting hydrological conditions. Model performance was assessed using mean absolute error, root mean square error, coefficient of determination, and correlation metrics. The best-performing hybrid ensembles reliably offered the highest predictive accuracy and stability across all wells. In the test stage, the hybrid models achieved root mean square errors of 0.01–0.24 m and mean absolute errors of 0.01–0.21 m, with coefficients of determination ranging from 0.99 to 1.00. The proposed framework provides a robust data-driven tool for groundwater forecasting and supports sustainable urban water resource planning in data-scarce regions.

Key words: Groundwater level predictionensemble learningurban groundwaterpermutation importance, SDG 6data-scarce region

https://doi.org/10.1080/1573062X.2026.2649885