ABSTRACT

Among several indicators for river engineering sustainability, the longitudinal dispersion coefficient (Kx) is the main parameter that defines the transport of pollutants in natural streams. Accurate esti mation of Kx has been challenging for hydrologists due to the high stochasticity and non-linearity of this hydraulic-environmental parameter. This study presents a new hybrid machine learning (ML) model integrating a Gaussian Process Regression (GPR) and an evolutionary feature selection (FS) approach (i.e. Covariance Matrix Adaptation Evolution Strategy (CMAES)) to estimate Kx in natural streams. The dataset consists of geometric and hydraulic river system parameters from 29 streams in the United States. The modeling results showed that the proposed model outperformed other mod els in the literature, producing more stable and accurate estimations. The FS approach evidenced the significance of the cross-sectional average flow velocity (U), channel width (B), and channel sin uosity σ to estimate the dispersion coefficient. In quantitative terms, the integrated GPR model with feature selection approach attained the minimum root mean square error (RMSE = 48.67) and max imum coefficient of determination (R^2 = 0.95). The proposed hybrid evolutionary ML model arises as robust, flexible and reliable alternative computer aid technology for predicting the longitudinal dispersion coefficient in natural streams.