Aytac Guven and Neslihan Seckin

Estimation of Peak Flood Discharges at Ungauged Sites across Turkey.

Floods can cause loss of human life and damage to agricultural lands. Reliable estimates of maximum discharge are necessary to correctly size hydraulic structures, reduce the risk of their failures, and minimize downstream environmental damage. The reliable forecasting of the peak flood discharge at river basins is a common problem, which becomes more complicated when there is inadequate recorded data. The statistical methods commonly used for the estimation of peak flood discharges are generally considered to be inadequate because of the complexity of the problem. Hence, genetic programming (GP) has recently attracted the attention of the hydrologists. In this study, gene-expression programming (GEP) and linear genetic programming (LGP), which are extensions to GP, and also logistic regression (LR) were employed in order to forecast peak flood discharges at 543 ungauged sites across Turkey. Drainage area (A), elevation (E_s), latitude (L_a), longitude (L_o), and return period (T_r) were the inputs while the peak flood discharge was the output.

The proposed LGP and GEP models give a fast and practical way of estimating the peak flood discharges for all river basins across Turkey accurately and encourage the use of genetic programming in other aspects of hydrologic and hydraulic engineering studies. The LGP and GEP models presented in this study are simple explicit mathematical formulations, which are applicable to all river basins across Turkey. The proposed LGP and GEP models offer no restriction since they do not employ predefined functions unlike most regression-based models. GEP4 model based on 4 input parameters (A, E_s, L_o, L_a) was found to be better (MSE=0.765 and R=0.758) than LGP4 (MSE=0.97 and R=0.748) and LR4 (MSE=1.028 and R=0.651). GEP was observed to perform quite successful in estimating the peak flood discharges during flood events. Also, GEP4 (MDL=-23.289) was observed to outperform LGP4 (MDL=-6.422) and LR4 (MDL=19.783) based on MDL criteria, which determines the high generalization capacity of GEP in testing phase.

Besides the five input parameters (A, E_s, L_o, L_a, T) utilized in this study, the performance of the proposed models would be improved if more available independent variables on physiographic, soil and land use properties, and climate were added to the input data-set.

The results of this study suggest that both genetic programming techniques (LGP and GEP) predict the maximum flood discharge better than conventional regression technique, and that LGP and GEP can be successfully applied in estimating the peak discharges during flood events.