Analysis of Sediment Transport Formulas using a Data Mining Technique with Applications to South Korean Rivers

Eun-kyung Jang[1], Un Ji[2], Woonkwang Yeo[3] and Pierre Y. Julien[1]
[1]Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, Colorado
[2]Hydro Science and Engineering Research Institute, Korea Institute of Civil Engineering and Building Technology, Goyang, South Korea
[3]Department of Civil and Environmental Engineering, Myongji University, Yongin, South Korea

Abstract. This research explores new ways to improve the applicability of sediment transport formulas to field conditions. Several formulas derived from laboratory measurements are often applied to rivers with very high annual maximum to minimum discharge ratios. For instance, it is particularly difficult to find appropriate sediment transport formulas for South Korean rivers where very low flow periods are frequently followed with massive floods from typhoons during the summer months. This study uses data mining techniques to derive the statistical pattern for the development of site-specific sediment discharge formulas based on a combination of 14 dimensionless and dimensional variables. For each case, the mean discrepancy ratio, the correlation coefficient, the root mean square error and the mean absolute percent error were calculated and used to review the accuracy of the predictions. As a result, the best fit was obtained after data mining from a sediment discharge formula using five-dimensional variables: flow velocity, flow depth, bed slope, river width, and the median diameter of the bed material. Also, the results of the data mining formula were compared with several original formulas from: Engelund and Hansen, Ackers and White, Yang, Brownlie and van Rijn. The evaluation has been made through comparisons with field measurements of discharge and sediment transport at 28 rivers stations in South Korea. Among the accuracy predictors for the comparison, the mean discrepancy ratio by data mining was 1.42. This shows improved discrepancy ratios compared to the original formulas of Engelund and Hansen (2.72), Ackers and White (1.44), Yang (1.65), Brownlie (1.55) and van Rijn (1.64), respectively.