Geophysical Research Abstracts
Vol. 21, EGU2019-10349-1, 2019
EGU General Assembly 2019

Looking beyond spatial correlation in regional flood frequency analysis: exploring the potential of Generalized Least Squares and Top-kriging

Simone Persiano (1), Jose Luis Salinas (2), Jery Russell Stedinger (3), William H. Farmer (4), Alberto Viglione (2), Günter Blöschl (2), and Attilio Castellarin (1)

(1) University of Bologna, School of Civil Engineering, DICAM, Bologna, Italy (simone.persiano@unibo.it), (2) Vienna University of Technology, Institute of Hydraulic Engineering and Water Resources Management, Vienna, Austria, (3) School of Civil and Environmental Engineering, Cornell University, Ithaca, New York, USA, (4) United States Geological Survey, Denver, Colorado, USA

An important issue in regional flood frequency analysis is the impact of intersite dependence of flood series (i.e. spatial cross-correlation). Cross-correlation reduces the actual hydrological information content of regional datasets, thus impacting the accuracy of regional predictions. Our study investigates the impact of spatial correlation on the prediction accuracy of two regionalization procedures: Generalized Least Squares (GLS) regression, used by the Geological Survey in the USA, and Top-kriging (TK), a geostatistical approach developed in the mid 2000's, which is increasingly being used in Europe. GLS and TK have a completely different way of taking cross-correlation among flood series into account: GLS develops a regional regression model after describing the cross-correlation among concurrent flood series, and then the correlation among estimators of flood characteristics; TK estimators of flood characteristics exploit the spatial correlation structure of the region. Despite its importance, a detailed analysis of this aspect has never been performed in the literature, to the best of our knowledge. We refer to a dataset of 20 annual maximum flood series for a hydrologically homogeneous region in North-Eastern Italy, whose estimated regional Log-Pearson Type III (LP3) distribution is assumed to be the parent distribution for a Monte Carlo simulation framework. A total of 3,000 synthetic 20-site regions are generated under different cross-correlation scenarios (i.e. average regional cross-correlation equal to 0.2, 0.6, 0.8). For each region and cross-correlation scenario, GLS and TK are applied to obtain predictions of at-site flood quantiles (return periods equal to 10, 30, 50 and 100 years) in a leave-one-out cross-validation scheme. Finally, the performances of GLS and TK are evaluated relative to the prediction in ungauged conditions of the (known) theoretical flood quantiles and their sample estimates. Our analysis clearly shows that, for both methods, the higher the cross-correlation degree, the higher the performance in predicting sample estimates of flood quantiles and the lower the performance in predicting their theoretical values. That is, correlation between observed series produces a masking-effect on the true flooding potential at a given ungauged site, the higher the correlation the higher the effect. Also, TK outperforms GLS when they both assume flood quantiles to scale with drainage area alone (e.g. relative to the prediction of the 100-year flood in ungauged sites for a regional cross-correlation of 0.6, the median Relative Nash-Sutcliffe Efficiency, RNSE, is equal to 0.77 and 0.58 for TK and GLS, respectively). The prediction accuracy improves significantly and becomes similar for both methodologies when flood quantiles are regressed against several catchment descriptors (e.g. for the scenario described above, RNSE is equal to 0.86 and 0.88 for TK and GLS, respectively).