Geophysical Research Abstracts
Vol. 14, EGU2012-11998, 2012
EGU General Assembly 2012
© Author(s) 2012

Flood quantile estimation at ungauged sites by Bayesian networks

L. Mediero, D. Santillán, and L. Garrote
Department of Hydraulic and Energy Engineering, Technical University of Madrid, Madrid, Spain (luis.mediero@upm.es)

Estimating flood quantiles at a site for which no observed measurements are available is essential for water resources planning and management. Ungauged sites have no observations about the magnitude of floods, but some site and basin characteristics are known. The most common technique used is the multiple regression analysis, which relates physical and climatic basin characteristic to flood quantiles. Regression equations are fitted from flood frequency data and basin characteristics at gauged sites. Regression equations are a rigid technique that assumes linear relationships between variables and cannot take the measurement errors into account. In addition, the prediction intervals are estimated in a very simplistic way from the variance of the residuals in the estimated model.
Bayesian networks are a probabilistic computational structure taken from the field of Artificial Intelligence, which have been widely and successfully applied to many scientific fields like medicine and informatics, but application to the field of hydrology is recent. Bayesian networks infer the joint probability distribution of several related variables from observations through nodes, which represent random variables, and links, which represent causal dependencies between them. A Bayesian network is more flexible than regression equations, as they capture non-linear relationships between variables. In addition, the probabilistic nature of Bayesian networks allows taking the different sources of estimation uncertainty into account, as they give a probability distribution as result.
A homogeneous region in the Tagus Basin was selected as case study. A regression equation was fitted taking the basin area, the annual maximum 24-hour rainfall for a given recurrence interval and the mean height as explanatory variables. Flood quantiles at ungauged sites were estimated by Bayesian networks. Bayesian networks need to be learnt from a huge enough data set. As observational data are reduced, a stochastic generator of synthetic data was developed. Synthetic basin characteristics were randomised, keeping the statistical properties of observed physical and climatic variables in the homogeneous region. The synthetic flood quantiles were stochastically generated taking the regression equation as basis. The learnt Bayesian network was validated by the reliability diagram, the Brier Score and the ROC diagram, which are common measures used in the validation of probabilistic forecasts.
Summarising, the flood quantile estimations through Bayesian networks supply information about the prediction uncertainty as a probability distribution function of discharges is given as result. Therefore, the Bayesian network model has application as a decision support for water resources and planning management.