Geophysical Research Abstracts
Vol. 17, EGU2015-11072, 2015
EGU General Assembly 2015

Data mining methods for predicting event runoff coefficients in ungauged basins using static and dynamic catchment characteristics

Ralf Loritz (1), Markus Weiler (2), and Simon Seibert (1)

(1) Karlsruhe Institute of Technology (KIT), Institute of water and river basin managment, Hydrology, Germany (ralf.loritz@kit.edu), (2) Faculty of Environment and Natural Resources, Chair of Hydrology, University of Freiburg, Germany

Transferring hydrological information into ungauged basin by regionalisation approaches is an ongoing field of research. Usually regionalisation techniques use physical landscape descriptors to transfer either model parameters or hydrological characteristics from a catchment to another. A common problem of these approaches is the high degree of uncertainty associated to their results. One reason is that often solely static (structural) catchment characteristics such as catchment area, physiographic properties or land use data are used for regionalisation. However, it is well known that the hydrological response of a 'natural' system is a complex and a non-linear interaction of its structure, state and forcing. Here it is important to note, that only structure is a static property. State and forcing are highly dynamic when considering the temporal and spatial scale of a rainfall-runoff event.
To overcome the limitations associated with 'static' regionalisation techniques we propose a regionalisation technique for event runoff coefficients combining static and dynamic catchment properties. The approach is based on the two data mining algorithms 'random forests' and 'quantile regression forests'. The static catchment characteristics include standard variables such as physiographic properties, land cover and soil data. The dynamic variables include event based properties of the forcing (i.e. rainfall amount, intensity,...) and proxies for the initial state of the catchment (i.e. initial soil moisture). Together with the runoff coefficient these quantities were extracted form hydro-meteorological time series (precipitation, discharge and soil moisture) using an automated rainfall-runoff event detection technique.
We tested our method using a set of 60 meso-scale catchments (3.1 to 205,6 km2, covering a range of different geologies and land uses) from Southwest Germany. We randomly separated the catchments in two groups. The first group (30 donor catchments) was used to train the data mining models . Based on the resulting relations we then predicted event runoff coefficients for the other half of the catchments (30 test catchments). With this regression method we are able to predict event runoff coefficient s in the test group with an overall root mean square error of about 5%. Furthermore our approach indicates that the dynamic characteristics (event precipitation and initial soil moisture) had a much higher importance for the prediction of event runoff coefficients than the static properties. In the next step, we applied random forest regressions to all 60 catchments individually based on the extracted event variables. The relative importance of the predictor variables of each of these regressions can be interpreted as indicators for the dominating rainfall-runoff controls within the basins (e.g. to identify initial storage or rainfall intensity controlled conditions).
We conclude that the ensemble regression tree methods provide insights into the 'functioning' of the individual catchments and that dynamic catchment properties (observed on meaningful spatial and temporal scales) have a very high potential for prediction hydrological functions in ungauged catchments.