Geophysical Research Abstracts
Vol. 14, EGU2012-8012-3, 2012
EGU General Assembly 2012

How much complexity is warranted in a rainfall-runoff model? Findings obtained from symbolic regression, using Eureqa

R.J. Abrahart and D.J. Beriro

School of Geography, University of Nottingham, Nottingham, United Kingdom (bob.abrahart@nottingham.ac.uk)

The information content in a rainfall-runoff record is sufficient to support models of only very limited complexity (Jakeman and Hornberger, 1993). This begs the question of what limits should observed data place on the allowable complexity of rainfall-runoff models? Eureqa[1] (Schmidt and Lipson, 2009) - pronounced "eureka" - is a software tool for finding equations and detecting mathematical relationships in a dataset. The challenge, for both software and modeller, is to identify, by means of symbolic regression, the simplest mathematical formulas which describe the underlying mechanisms that produced the data. It actually delivers, however, a series of preferred modelling solutions comprising one champion for each specific level of complexity i.e. related to solution enlargement involving the progressive incorporation of additional permitted factors (internal operators/ external drivers). The potential benefit of increased complexity can as a result be assessed in a rational manner. Eureqa is free to download and use; and, in the current study, has been employed to construct a set of rainfall-runoff transfer function models for the Annapolis River at Wilmot, in north-western Nova Scotia, Canada. The climatic conditions in this catchment present an interesting set of modelling challenges; daily variations and seasonal changes in temperature, snowfall and retention result in great difficulty for runoff prediction by means of a data-driven approach. Data from 10 years of daily observations are used in the present study (01/01/2000–31/12/2009): comprising [i] discharge, [ii] total rainfall (excluding snowfall), [iii] total snowfall, [iv] thickness of snow cover, and [v] maximum and [vi] minimum temperature. Precipitation occurs throughout the whole year being slightly lower during summer. Snowfall is common from November until April and rare hurricane weather may occur in autumn. The average maximum temperature is below 0 ^{0}C in January and February, but significant variation may result, producing milder weather and snowmelt throughout the winter. The average minimum temperature is below 0 ^{0}C during half of the year, such that freezing and melting occur frequently. The principal rainfall-runoff drivers are found to be lagged discharge and lagged precipitation, as expected. The complexity-accuracy trade-off, is nevertheless found to exhibit threshold behaviour, in which snow cover is eventually included at higher levels of complexity to account for multifaceted cold season processes.

REFERENCES

Jakeman, A.J. and Hornberger, G.M. (1993) How much complexity is warranted in a rainfall-runoff model? *Water Resources Research* 29(8): 2637-2649.

Schmidt, M. and Lipson, H. (2009) Distilling Free-Form Natural Laws from Experimental Data. *Science* 324(5923): 81-85

[1]http://creativemachines.cornell.edu/eureqa