

## ***Interactive comment on “Comparative assessment of predictions in ungauged basins – Part 1: Runoff hydrograph studies” by J. Parajka et al.***

**Anonymous Referee #1**

Received and published: 26 February 2013

This is a well-written paper which makes a significant contribution to the literature. The only concern I have is with the intercomparison of Nash-Sutcliffe Efficiency scores across widely differing environments. I think this topic needs to be mentioned by the authors in their paper. The paper contains a wealth of valuable information, which the authors have made accessible and digestible. I would expect it to lead to many follow-up studies, and to provide useful guidance for hydrological modelling studies and applications.

### Specific Comments

P379 Line 24 "The predictive accuracy was then described by the Nash-Sutcliffe efficiency (NSE, Nash and Sutcliffe, 1970) of daily runoff." This is a limited method of assessing model performance. It would be useful to include some commentary on

C142

strengths and weaknesses of using NSE to evaluate hydrological models.

For example, I note these two quotes from the literature:

"the Nash efficiency value, while a convenient and normalized (–inf to 1.0) measure of model performance does not provide a reliable basis for comparing the results of different case studies" Schaeffli, B. and Gupta, H. V. (2007), Do Nash values have value?. Hydrol. Process., 21: 2075–2080. doi: 10.1002/hyp.6825

"A comparison of NSE across basins with different seasonality (as is often reported in the literature) should therefore be interpreted with caution." Gupta H.V., Kling H., Yilmaz K.K., Martinez G.F., 2009: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modeling, Journal of Hydrology, 377(1-2), pp. 80-91. doi:10.1016/j.jhydrol.2009.08.003.

The main reason for raising this issue is that the paper analyses catchments from a wide range of environments, including some which are strongly seasonal and some which are not. It is easy to get a good NSE score in a strongly seasonal catchment, even with a rather crude model. Is a high NSE score in a strongly seasonal environment a reflection of high accuracy of modelling in this environment, or an artefact of the scoring?

I see no problem in using NSE to compare different methods on the same catchment, but there may be a problem of interpretation when comparing the same method on markedly different environments (e.g. section 4.5).

P380 Line 24 "Four catchment characteristics are analysed: aridity index, mean elevation and catchment area" The list contains three characteristics, rather than four.

P385 Line 22. The authors use number of flow recorders as a measure of data availability. It is true that more flow recorders will enable better predictions. However, the number of gauges is a very crude measure of information availability. A set of ten gauges within a study area of 100 km<sup>2</sup> is very different to ten gauges within a study

C143

area of 100,000 km<sup>2</sup>. Do the authors consider that the number of gauges is more useful in characterising information availability than the gauge density? (density = number of gauges per sq km). In addition, the quality of flow estimates often depends on the quality of rainfall information. If the rainfall data is very sparse, then one might also expect poor simulations. I think this is worthy of a comment in the paper, even if no information is available on rainfall data density.

P388 line 20 "Overall, this very clear pattern of an increase of the performance with catchment scale may be due to two reasons. The first is a trend for an increasing number of raingauges within a catchment as the catchment size increases." Why does more raingauges lead to better models? Surely it depends on the density of raingauges relative to the correlation length scale of the rainfall? It might be useful to cite Schaake (1981, see <http://www.nws.noaa.gov/oh/mopex/raingage%20density%20requirement.htm>)

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 375, 2013.