Hydrology and
Earth System
Sciences
Discussions

# *Interactive comment on* "Evaluating Hydrological Model Performance using Information Theory-based Metrics" *by* Y. A. Pachepsky et al.

**Anonymous Referee #2**

Received and published: 14 March 2016

The paper basically investigates a model-data comparison with three quantifiers from Information Theory in addition to the conventional Nash-Sutcliffe model efficiency: MIG, EMC and FC. The time series have a length of exactly 11 years (why?) and are from the period 1960-1970 (why? Could you not find any more recent measurements?), at daily resolution, which means around 4000 data points per time series. The observations (streamflow) from five different watersheds are compared to 8 different hydrologic models. The latter are not characterized to any detail, there is only a Table with a rather short characterization of each of these models (they do not have names?). What is worse, the calibration process is not even mentioned. Nobody can judge whether there has been an automatic parameter optimization performed, whether the full parameter space has been exploited, what was the objective function (presumably NSE, or RMS, or something else?). Whether the statement "more com-

plex models perform better than simpler ones" are supported by these experiments remains unclear; in a more general setting (models in the geosciences), it is usually wrong. The NSE results point to severe mismatches between some of the models and the observations. Models S1 and M1 seem to produce perfectly constant output (runoff values) for some of the watersheds, since their location in the MIG-FC and MIG-EMC planes are very close to the lower left corner. However, no simulated time series are shown. The measures from Information Theory are described in more detail than necessary, since the methodology has been published a long time ago; in addition to Lange (1999) and Wolf (1999), the seminal paper by Wackerbauer et al. (1994) has to be mentioned in this context. The current authors underexploit the method seriously by restricting themselves to word length = 2, which is unnecessary given the length of their time series. More concerning is that they do not seem to be aware of the upper limit curve given by the binary Bernoulli process, as shown and discussed extensively in the relevant references. This couples MIG and FC together in both the very random and the very regular regimes of the complexity plane, and that is what the authors necessarily observe in their results; there is no room for interpretations like they attempt, since they simply (re-)discover a mathematical property of the measures. Another issue is the calculation of differences as Euclidean distances, which is not justified given that the manifold is not that simple as the authors seem to assume (there is an almost parabola-shaped limit curve). My suggestion would be to use the Hellinger distance in pattern space – although this is of limited value at L=2, but makes sense for higher word lengths. The MIG values for the runoff time series are surprisingly small compared to what others have found (e.g. Hauhs and Lange 2008). This might be due to the size of the watersheds (large), but also renders the modelling task less demanding (as compared to, say, first order watersheds). It is thus surprising that the models chosen are not able to cope with the modeling problem better than (most of them) actually do. The paper is full of grammar errors, typos and other language problems, and simply not written well. A larger set of rather detailed comments is contained in the uploaded file. It might serve as guidance for necessary

improvements of the paper. Given these limitations, the reviewer suggests to reject the paper for publication in HESS.

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/hess-2016-46/hess-2016-46-RC2-supplement.pdf

---

C3