Hydrol. Earth Syst. Sci. Discuss., 11, C2160–C2163, 2014
www.hydrol-earth-syst-sci-discuss.net/11/C2160/2014/

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Development of a large-sample watershed-scale hydrometeorological dataset for the contiguous USA: dataset characteristics and assessment of regional variability in hydrologic model performance" *by* A. J. Newman et al.

H. Gupta (Referee)

hoshin.gupta@hwr.arizona.edu

Received and published: 1 July 2014

See also attached pdf document:

Review by Hoshin Gupta of: Newman AJ, MP Clark, K Sampson, A Wood, LE Hay, A Bock, R Viger, D Blodgett, L Brekke, JR Arnold, T Hopson and Q Duan (2014, Development of a large-sample watershed-scale hydrometeorological dataset for the contigu-

ous USA: dataset characteristics and assessment of regional variability in hydrologic model performance, Hydrology and Earth Systems Sciences, Discussion Paper, 11, 5599–5631, www.hydrol-earth-syst-sci-discuss.net/11/5599/2014/ doi:10.5194/hessd-11-5599-2014.

Summary of the Paper: The paper presents a thirty-year (1980–2010) "large-sample watershed-scale data set" (671 small- to medium-sized basins) of daily forcing (max & min temperature, precipitation, shortwave downward radiation, day length, humidity, potential evapotraanspiration and snow water equivalent) and hydrologic response (basin-outlet streamflow) data for basins across the contiguous United States, spanning a wide range of hydroclimatic conditions.

1) The original source of the data is daily, high spatial resolution (1 km) gridded Daymet (T & P) meteorological data (necessary to represent spatial heterogeneity in complex topography) and USGS streamflow observations. 2) The data are provided (via areal averaging) for three spatial configurations: a) watershed-lumped, b) HRU and, c) elevation band. 3) The selected basins are almost exclusively smaller, headwater-type basins having minimal human influence (having been subset from the USGS's GAGES-II/reference/HCDN-2009 dataset), and so can be expected to represent natural flow conditions reasonably well.

In addition to presenting the data, the paper provides a model performance benchmark using the coupled SACSMA-Snow-17 model, calibrated to minimize the RMSE to 15 years of daily streamflow via SCE global optimization.

Overall Review Comments: I congratulate and commend the authors for an excellent job in compiling and reporting this "large-sample" data set of watersheds across the CONUS that can be used for hydrological and modeling investigations emphasizing spatial extensiveness (breadth) and generality of hydrological understanding (and therefore associated model performance). The paper is an excellent example of a relatively comprehensive study and report, and should function as a "model" or "template" for similar studies to be conducted for the other continents on our planet. In addition to reporting a carefully considered and reasoned approach to basin selection, they authors provide a benchmark assessment of simulation performance based on a "standard" lumped catchment model and calibration approach. I have only a very few suggestions for how the paper might be improved.

1) It might be nice to see (in Section 2) some more analysis of "basin characteristics" that would facilitate comparison/contrasting of the CONUS basins with ones from other continents. This analysis should probably include a) physical descriptors such as size, mean elevation distribution, shape (length to width ratio), and river characteristics such as distribution of stream order, dendritic pattern etc., and b) climatological descriptors including both annual values and monthly climatology. 2) While NSE = 0.55 is "OK" as a benchmark, it does reflect a relatively low level of model performance. Perhaps the authors could also slightly expand the discussion in the text to mention also the fraction of catchments exceeding (say) NSE = 0.8. This would help to set the tone for future studies by setting an "OK" level and a "Good level". 3) I think it would be good to more strongly emphasize the role of large-sample studies to help identify "outlier" catchments (and regions), along with the important function of "characterizing" the nature of the "outlier" (e.g., as being likely due to data errors, model inadequacy, calibration failure, etc). 4) The issue of performance on basins with strong annual climatology does not come through very strongly in the discussion. I wonder if it would help to include a map showing where the "climatology" is strong and where it is relatively weak (e.g., strength of flow correlation (?))". 5) The authors report MSE decomposition components for bias and variability. For completeness, perhaps they could also report the obs-sim cross-correlation coefficient.

Additional Suggestion: As the basis for a complimentary study, I think it would be interesting to repeat the calibration-evaluation study while interactively removing the (say) 5% or 10% of the time-steps (used to compute the performance measure) that correspond to the largest simulation errors (and therefore strongly influence the selected

"best" parameter set). While this might lead to interesting response surface artifacts during calibration (but none that SCE should not be able to handle), I wonder if this would lead to more stable calibration results, when viewed for the evaluation period (functioning as a sort of fault-detection strategy)?

Minor Comments: 1) The sentence "Gupta et al. (2014) emphasize . . ." beginning on page 5602 line 4 cites the paper twice (at beginning and end). 2) The phrase "well know" on page 5605 line 20 should be "well known" 3) On page 5606 line 3, the phrase "necessitating a snow model is required," either the word "necessitating" or the phrase "is required" should be removed. 4) On page 5607 line 24 "is shown by" should be "as shown by". 5) On page 5609 line 23, the term "poorlyfollowing" should be "poorly following". 6) In Section 4.3 it might be interesting to compare spatial variability results with those reported by Martinez & Gupta (2010, 2011 WRR). 7) On page 5615 line 8 should this be "which utilized 425 of the basins . . ."?

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/11/C2160/2014/hessd-11-C2160-2014-supplement.pdf

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 11, 5599, 2014.