

## ***Interactive comment on “Disentangling timing and amplitude errors in streamflow simulations” by S. P. Seibert et al.***

### **Anonymous Referee #1**

Received and published: 9 May 2016

Overall, I like this paper very much.

But I am a little disappointed: the metric proposed here is not new. The authors already published it. They have a new improved version, this is great: but in this case, you should show us on some “large enough” dataset that the new version of the metric is better than the previous version.

Naturally, this is not easy: on which basis would you show us that the new version is “better”? The fact that we have two different numerical values proves nothing of course. You could do again what you have done with the survey of experts in Crochemore et al. (2015): is one of the versions better correlated with the opinions of the experts? An alternative would be to use the new metric to calibrate and multiple metrics to validate. If you can get better numerical results in validation on a range of numerical criteria by

[Printer-friendly version](#)

[Discussion paper](#)



“extracting” information from data during the calibration process with the new version, this would indicate a “better” criterion.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-145, 2016.

**HESD**

---

Interactive  
comment

[Printer-friendly version](#)

[Discussion paper](#)

