Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2015-545-RC1, 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Uncertainty assessment of a dominant-process catchment model of dissolved phosphorus transfer" *by* R. Dupas et al.

T. Krueger (Referee)

tobias.krueger@hu-berlin.de

Received and published: 4 February 2016

I enjoyed reading this paper. It is structured and clear and makes useful contributions to the GLUE limits of acceptability (LOA) approach and model evaluation in a water quality context more widely. The authors took great care to derive the LOA from independent evidence including a small number of repeat samples (which are hard to come by). I think the GLUE LOA approach shows its strength when the LOA are empirically grounded and not defined ad hoc as in some of the earlier 'proof of concept' applications. Hence this study will help greatly to operationalise the method. It does help that the authors come up with a model whose complexity is attuned to the data availability. This means the model actually manages to operate within the LOA without the authors having to unduly relax them post hoc. Where they relax the LOA this is done sensibly and well justified. Here again the paper goes beyond earlier GLUE LOA

studies (including my own) where the relaxation of LOA was a bit unsatisfactory.

One methodological inconsistency, however, is the translation of probability density functions (from regressions) into LOA and triangular weighting functions (sections 2.3.1-2.3.2, P11-13). Not only does translating pdfs into LOA by using the central 90% confidence bounds seem arbitrary. What is more, arguably, if you are happy with the regression (and the assumptions that come with it) then you have a probabilistic representation of uncertainty that you can use in GLUE. So why not use this information? If you're not happy with the regression (I wouldn't most of the time) then why not use a method more attuned to the assumptions you're happy to make? Examples in the GLUE LOA context are: Pappenberger et al. (2006), Krueger et al. (2010, 2012), Westerberg et al. (2011). I think this inconsistency weakens the method considerably when it is exactly the empirical grounding of the LOA and weighting functions that make it so valuable, as I argue above. Preferably, the method would be changed accordingly – or convincing arguments for the chosen approach presented in the paper.

Another aspect that is not entirely clear is the reasoning behind the 2-day aggregation of SRP loads (P12, L29-P13, L7). I see that you compare the model to total storm loads that way. But why not compare the model to daily loads (which it can simulate)? This would be a much stronger test of the model. Arguably, why need a daily resolution model at all if you are only interested in storm loads? And if you need the processes that the daily resolution covers, why not test these with data at the same resolution (which you have)? Preferably, the study would be adapted to take this into account – or we need convincing arguments in the paper why the present approach was chosen despite the concerns identified.

I'm also missing all the earlier GLUE LOA applications in the field of water quality modelling. Reviewing these in the introduction and discussing your own results against what they had to say would increase the scholarly impact of the paper I would say. Among those are: Page et al. (2003, 2004), Rankinen et al. (2006), Quinton et al. (2011), Krueger et al. (2012). Even if some of these don't call their approach LOA, they nevertheless, through the use of fuzzy performance measures with a finite support and multiplicative aggregation, effectively apply LOA within GLUE!

Specific points

P2, L5-7: It would be good to also refer to ecological impacts here.

P2, L13-15: The P fractionation in transit (e.g. resorption) would be important, too.

P3, L31-P4, L1: Here you could cite Krueger et al. (2012) where we dealt with evaluation data uncertainty (suspended solids, TP) explicitly in a GLUE limits of acceptability framework, albeit with even simpler models at finer time scales.

P4, L2-5: The grab sample uncertainty discussion could be usefully enhanced by referencing McMillan et al. (2012) where we discuss these issues at length by synthesising a large body of work.

P4, L3-5: Grab samples also represent a snapshot at a given point in the stream (e.g. Rode & Suhr, 2007).

P5, L12 and elsewhere: Please specify what +/- represents – one standard deviation?

P7, L13f: I think there should be "model" at the end of the title.

P10, L4-17: Here or elsewhere it would be good to note that no long-term depletion of soil P pools was modelled, i.e. effectively assuming steady state. This would also be an interesting point for discussion.

P11, L4-16: Here especially it would be good to cite other GLUE LOA studies in water quality modelling, see above.

P12, L20-22: Does this imply that no intercept was fitted in the regression equation? Would be good to clarify either choice.

P14, L7-9: Why were the weights summed (average) and not multiplied in keeping with the LOA concept? Krueger et al. (2012) discuss this.

P16, L23-25: No. What you must say is that you cannot reject this set of processes as a hypothesis of dominant control given the available evidence! There is no confirmation here, only a failed rejection.

Section 4.2, P17-19: Here I'm missing a discussion of the neglect of farm management practices in the model – which are vitally important if the model is to eventually have any bearing on catchment management.

P18, L5-8: If you want to make this point then you should also discuss what benefit the finer resolution of the SW-GW interactions brings given that the subsequent P processes are much coarser (e.g. there are no hyporheic zone P reactions).

Fig. 7: The 1st storm in (a) is not easy to see (lines too close together) – consider different x-scale or 2 panels or else. In (b) the mix of lines and vertical lines with triangles to represent the LOA at the different resolutions (storms vs. baseflow) is confusing. Best would be to evaluate the model during the storms at the same resolution as during baseflow (daily, see above). If you can convince that this is not necessary then think of a different representation, maybe only the vertical lines but making the triangles smaller.

Fig. 8: Can't distinguish the lines and the LOA to get a sense of model fit. Consider scaling x differently, see above, and making model lines smaller and in same colour. Emphasise LOA lines (perhaps move in front of model lines).

References

Krueger, T., J. Freer, J. N. Quinton, C. J. A. Macleod, G. S. Bilotta, R. E. Brazier, P. Butler, and P. M. Haygarth (2010), Ensemble evaluation of hydrological model hypotheses, Water Resour. Res., 46, W07516, doi: 07510.01029/02009WR007845.

Krueger T, Quinton JN, Freer J, Macleod CJA, Bilotta GS, Brazier RE, Hawkins JMB, Haygarth PM. 2012. Comparing empirical models for sediment and phosphorus transfer fromsoils to water at field and catchment scale under data uncertainty. European

Journal of Soil Science 63(2): 211–223.

Page, T., K. J. Beven, J. Freer, and A. Jenkins (2003), Investigating the uncertainty in predicting responses to atmospheric deposition using the model of acidification of groundwater in catchments (MAGIC) within a generalised likelihood uncertainty estimation (GLUE) framework, Water Air Soil Pollut., 142(1-4), 71-94.

Page, T., K. J. Beven, and D. Whyatt (2004), Predictive capability in estimating changes in water quality: Long-term responses to atmospheric deposition, Water Air Soil Pollut., 151(1-4), 215-244.

Pappenberger, F., P. Matgen, K. J. Beven, J. B. Henry, L. Pfister, and P. Fraipont de (2006), Influence of uncertain boundary conditions and model structure on flood inundation predictions, Adv. Water Resour., 29(10), 1430-1449.

Quinton, J. N., T. Krueger, J. Freer, G. S. Bilotta, and R. E. Brazier (2011), A case study of uncertainty: Applying GLUE to EUROSEM, in Handbook of Erosion Modelling, edited by R. P. C. Morgan and M. A. Nearing, pp. 80-97, Blackwell Publishing Ltd, Chichester.

Rankinen, K., T. Karvonen, and D. Butterfield (2006), An application of the GLUE methodology for estimating the parameters of the INCA-N model, Sci. Total Environ., 365(1-3), 123-139.

Rode M, Suhr U. 2007. Uncertainties in selected river water quality data. Hydrology and Earth System Sciences 11(2): 863–874.

Westerberg, I., Guerrero, J.-L., Seibert, J., Beven, K. J., and Halldin, S.: Stage-discharge uncertainty derived with a non-stationary rating curve in the Choluteca River, Honduras, Hydrol. Process., 25, 603–613, doi:10.1002/hyp.7848, 2011.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2015-545, 2016.