Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-314-RC1, 2017

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Data-based mechanistic model of catchment phosphorus load improves predictions of storm transfers and annual loads in surface waters" *by* Mary C. Ockenden et al.

S. Stoll (Referee)

sebastian.stoll@agroscope.admin.ch

Received and published: 13 June 2017

General remarks:

In the manuscript "Data-based mechanistic model of catchment phosphorus load improves prediction of storm transfer and annual loads in surface water“ the authors present different data-based mechanistic models describing the dynamics of discharge and phosphorus load in three catchments in the UK.

Generally, I find the manuscript to be very interesting, well written and suitable for HESS (after some revisions). While I agree with the authors that DBM models are very

Printer-friendly version

Discussion paper

helpful in detecting dominant transfer modes I think that some of the alleged benefits of the modelling approach are overstated. For example, I doubt that these models can "help in planning appropriate pollution mitigation measures" as stated in the abstract. The reason for that is the nature of these models. The only input driving the models is rainfall (and sometimes discharge) data. Many features which are known to influence the phosphorus dynamics (like soil type, soil phosphorus concentration, management practices, tile drainage density, etc.) and which would be the primary entry point for any mitigation measures are not directly considered. Accordingly, the effect of any changes in these features (e.g. management practices) cannot be evaluated (not saying that physically-based models are per se any better with regard to that given the parameter uncertainty). In my opinion, the presented DBM models are much better suited to analyze the effects in changes of the precipitation (as rainfall is the main input) under the condition that these future precipitation conditions are covered in the calibration period.

In addition, I would love to see some more analysis of the very nice data they collected. I would assume that the manuscript would greatly benefit if the model results would be discussed together with the data (for example detailed analyses of the hysteresis curves).

Specific remarks:

Title: Improvement compared to what, other models?

P1, L31-32: See comments above

P2, L7: The authors correctly point out the importance of the measurement uncertainty. However, in the whole manuscript no information is provided regarding the uncertainty of the rainfall, discharge and phosphorus measurements or how this uncertainty is handled in the modelling approach. Especially the stage-discharge relationship (regarding the discharge measurements of flood events) can be subject to considerable uncertainty which would directly translate into uncertainty of the phosphorus loads. One could argue that the measurement uncertainty is indirectly accounted for by the parameter uncertainty. However, given that the uncertainty bands are hardly detectable in the figures and measurements (without error bars) are not covered by it, it seems that either an important process is not captured by the model or that the measurement uncertainty is underestimated.

P2, L24-32: Here, the authors report the disadvantages and shortcomings of large, overparameterized process-based models (e.g. SWAT). I understand the motivation for that and even to a large degree agree with them. However, the authors should not only pick and describe the most extreme (or worst) process-based models. There are also parsimonious process-based models which can deliver reasonable results describing dynamics of phosphorus on hourly time steps (for example Hahn et al., 2013) or spatial herbicides losses (which have very similar transport pathways) (for example Frey et al., 2011) with few parameters.

P4, L14: What was the motivation to measure TP and not distinguish between or focus on particulate and/or dissolved phosphorus? Particulate (PP) and dissolved phosphorus (DP) can have different pathways and dynamics. While PP often shows a clockwise hysteresis (P peak before Q peak), DP often shows an anti-clockwise hysteresis (Q peak before P peak) (Dupas et al, 2015). By modelling them separately, it would be probably easier to identify a suitable transfer function and the corresponding pathways.

P5, L17: What is R in the equation, rainfall?

P6, L6-10: What is the motivation for setting up these short-term models for the Newby Beck catchment when the long-term model have similar performances and structures?

P6, L15: If I understand it correctly, the output which is used to identify and calibrate the model is also used as an input. I find this contra-intuitive and not really “proper”. Why not use a precipitation based antecedent wetness index?

P7, L7-10: Some scatter plots would be very helpful to illustrate the Q-P relationships.

Printer-friendly version

Discussion paper

P7, L16: Table number is missing.

P7, L19-20: Because Blackwater has the lowest specific discharge. It would be good to discuss and explain the differences in the specific discharges and P concentrations among the catchments.

P7, L28-30: So were model results actually used to fill data gaps for the longterm model? If yes, this should be clearly stated and discussed accordingly.

P7, L29-30: How can model results help in identifying problems in the extrapolation of the stage-discharge relationships, when the whole model itself is based and calibrated with data of exactly these stage-discharge relationships? In my opinion the model is only reliable for the conditions covered during the calibration period. If more extreme events would be included in the calibration period, the model and the parameters would very likely be different.

P8, L4: Should be "Table 2"

P8, L4-18: I find the discussion and evaluation of storm Desmond a bit constructed and unnecessary. You don't need a DBM model to realize that discharge and P load was underestimated when there are reports of out-of-bank discharge bypassing the gauging station. The model also doesn't help in the quantification of the missed P and Q. As mentioned before, the model was trained under different conditions and is therefore in my opinion not really valid for very extreme cases not being part of the calibration period (again not saying that physically based models are any better).

Table 2: According to the time constants and order of the Q- and TP models, there are two pathways contributing to the discharge generation with only the fast pathway contributing to the TP generation. If I understand the concept of the TC correctly, TP reacts before the discharge rises. Is this in agreement with the measured data?

Table 3: What is the meaning of the term "using Qsim". If model outputs instead of actual measurements were used, this should be clearly stated, justified and discussed

(for example why is the performance worse for "using Qsim" that "using Qobs"?) In relation to that, how was TPLoad calculated? Did the authors used the modeled Q to calculate TPLoad or did they use the measured Q? Again, if modeled Q was used, this should be stated, justified and the consequences discussed.

Table 3 cont'd: For the Newby and Wylye TPLoad models effective rainfall was used as input, while regular rainfall was used for the discharge model. What is the meaning of that? Does it mean that for TP dynamics antecedent conditions are important, while they are not important for the discharge dynamics? Again, I would advise to discuss these findings as well as the different time constants and their percentages together with the actual measured data.

P9, L26: What does "effective rainfall (from the runoff model)" mean?

P11, L7: Same point again. What does "effective rainfall simulated by the rainfall-runoff model mean?

P12, L1: It's nice to have models with a low parameter uncertainty. However, when the uncertainty bands do not encompass the measurements, it's not really better situation than having a large parameter uncertainty. The model is either missing an important process or measurement uncertainty is not accounted for. A third reason could be a too narrow parameter sampling space in the MC method.

P12, L28-33: Although, the authors openly discuss the limitations of their modelling approach, there is one point I miss. They argue that understanding the rainfall-Q/TPLoad relationship through DBM models can help to identify dominant modes of the catchment and can therefore be used to target management interventions. I would argue that this is only possible if the identified dominant modes or pathways can be related to specific areas in the catchment. In my opinion it is not enough to know that 70% of the TPLoad was activated via a fast pathway. It is necessary to know which areas in the catchments are connected to the stream via this pathway, how these areas are managed and what their soil P status is. To actually plan and implement intervention strategy, you need to know where (on which fields) and how to intervene. The "how" is strongly dependent on the "where". If you identified some fields with subsurface tile drainage as the contributing areas you would need a different intervention strategy as for example on a field with a tendency for surface runoff due to soil compaction. Knowing the temporal dynamics is not good enough, you would also need information about the spatial patterns.

Authorship:

I thought long about including this very last comment in the review. However, given the many discussions I had with colleagues about this very issue in the past, I feel somewhat obliged to mention that I find the number of authors contributing to this manuscript too excessive, given the nature of the article (a regular modelling study). I am very much in favor in acknowledging significant contributions (for example with respect to data gathering) with a co-authorship, however this seems not to be the case here. The authors state themselves that while two persons were responsible for the modelling, three persons did project management and the remaining fourteen (!) basically discussed the results and did some editing. I certainly don't want to offend any of the authors and obviously have no insights in the preparation process of the manuscript. Nonetheless, I would encourage each co-author to reflect if in their opinion they really contributed significantly to this manuscript.

References

Dupas, R., Gascuel-Odoux, C., Gilliet, N., Grimaldi, C., Gruau, G., 2015. Distinct export dynamics for dissolved and particulate phosphorus reveal independent transport mechanisms in an arable headwater catchment. Hydrol. Process. 29 (14), 3162e3178. http://dx.doi.org/10.1002/hyp.10432.

Frey,M.P.; Stamm,C.; Schneider,M.K.; Reichert,P. (2011) Using discharge data to reduce structural deficits in a hydrological model with a Bayesian inference approach and the implications for the prediction of critical source areas, Water Resources Research, 47(12), W12529 (18 pp).

Hahn,C.; Prasuhn,V.; Stamm,C.; Lazzarotto,P.; Evangelou,M.W.H.; Schulin,R. (2013) Prediction of dissolved reactive phosphorus losses from small agricultural catchments: Calibration and validation of a parsimonious model, Hydrology and Earth System Sciences, 17(10), 3679-3693.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-314, 2017.