

## ***Interactive comment on “An integrated water system model considering hydrological and biogeochemical processes at basin scale: model construction and application” by Y. Y. Zhang et al.***

**Anonymous Referee #2**

Received and published: 2 September 2014

The manuscript describes a novel attempt for integrated water modelling with relying on existing submodels. The model description is coupled with a case study in the Sha Ying catchment (East China), where a partial comparison is made to the SWAT model based on a previous study. The topic is timely and would nicely fit into the scope of HESS. I think, however, that the manuscript needs significant changes to be acceptable.

Certain parts (Introduction and conclusion) of the paper promise more than the study actually does/did. While this is mainly a stale issue, there are more severe problems with the general presentation of the work. The introduction is built around the idea

C3563

that the (in my opinion old yet persisting) challenges of integrated modelling need to be addressed by new models because old models have intrinsic problems (being limited in scope or too simplistic) that prevent their meaningful usage in integrated assessments. While this may be true, the reader gets surprised in the following section that the new model developed to solve these problems is mostly a different mix of the old ingredients of the same conceptual models (for example significant bits of QUAL2K and SWAT are re-used). For this reason I think that the presentation of this indeed interesting model and case study should be done in a different way. As SWAT is really encompassing many of the proven thematic submodels, it is not a shame that HEXM has a significant structural similarity to it. Additionally, SWAT is very far from being perfect from both practical and theoretical points of view, so alternative models have their *raison d'être* as well. I would less emphasise the general problems of SWAT because - as a close relative - HEXM shares most of them. The focus could be put on the advantages of HEXM stemming from the different scope (for example dams) and different submodels (hydrology, soil, etc).

A structural problem is the lack of sufficient discussion. While the model itself and the results of the case study are presented in acceptable detail, the theoretical implications of putting these submodels together and a structural comparison of the new model to the criticised predecessors is completely missing. It would be interesting to learn about the problems when conceptual models operating on different spatial and temporal scales (for example hydrology operates on subcatchments, soil biogeochemistry and erosion are on the site scale) are connected in an integrated framework. These issues have to be addressed if HEXM is considered as a new integrated model and not a modification/extension of SWAT.

---

Specific comments

P 9222 L 5: I think that this hindrance will actually continue in the future too.

C3564

P 9222 L 7: This is difficult to understand: What do you consider as “the traditional hydrological method” for solving water pollution and ecological degradation?

P 9222 L 9-11: Process-oriented modelling may indeed be the most efficient tool, but here no justification is given. What are the arguments against other methods?

P 9222 L 14: I think that “energy process” is rather “energy fluxes”.

P 9222 L 15: These sentences can be made much shorter without any loss of information: “For example, the physiological and ecological processes of vegetation affect evapotranspiration, soil moisture distribution and infiltration, and nutrient sorption and movement. On the contrary, soil moisture and nutrient content directly affect crop growth. Overland flow affects the pollutant loads to water bodies.”

P 9222 L 27: What is the basis for this statement?

P 9223 L 4: “Darcy’s”

P 9223 L 8: Here I think “remote sensing” is more appropriate than GIS. Moreover I don’t see the relevance of “GPS”.

P 9223 L 11: Here you state that the combination of these rather old yet practical knowledge (in the previous paragraph) into an integrated model system stems from the 1980s. Then I think it is unnecessary to have so detailed description of these relationships in the previous paragraph. The following sentence essentially summarises the whole story: “Several models have been developed based on the mature models of different disciplines (hydrology, environment and ecology).”

P 9223 L 14: “based” is probably not the best term here. I would suggest to write that different models put the emphasis on different processes.

P 9223 L 17: I think that the difference between true empirical and process-based conceptual models is overemphasised. The attributed real-world meaning of individual parameters and processes of conceptual models is not guaranteed to persist when the

C3565

model is calibrated (see e.g. Mantovan and Todini, 2006; Mantovan et al. 2007). At the same time, calibration-free models are rather rare in environmental modelling. This means that most of our (sub)models are actually empirical to a certain degree. That’s why I think one should not overemphasise the distinction between “process-based” and “empirical” models. The main message of this part seems to be that due to the complexity of the integrated system it is rare that each process is simulated with the same detail. Thus one gets best performance for processes that are described with the most detail and only approximate results for others outside of the model’s focus.

P 9224 L 7: This highlights the tradeoff of integrated models. SWAT covers enough processes to be called an integrated model, but from the perspective of specialised models SWAT’s processes are overly simplified. However, the entire integrated system is so complex that it is one of the largest environmental models and there is practically no chance to calibrate everything.

P 9224 L 7: “described in its model setting” can be deleted.

P 9224 L 12: As mentioned before, “over-simplified” is only true at the scale of single processes.

P 9224 L 18: It’s not a “new challenge”, this is the original challenge of integrated modelling from the very first moment. Just like the tradeoff of resource allocation in integrated water management. Nevertheless, the statement is right in the sense that integrated management has just been adopted in the last decades and these issues are now becoming obvious for the practitioners.

P 9224 L 26: And most importantly, there is MUCH more data than before, which allows for more detailed analyses.

P 9225 L 1: “non point source pools” of which pollutants/nutrients?

P 9225 L 9: “to lay the scientific foundation to promote the implement of integrated river basin management all over the world” is a bit overstated, so I suggest to remove.

C3566

P 9225 L 10: "as follows."

P 9225 L 21: TVGM is a nice member of the class of saturated path models where the proportion of runoff (=the proportion of saturated area in the (sub)catchment) is the function of the areal mean soil moisture. However, just like the other members of this class, TVGM is a conceptual model. While using TVGM may be a step forward from totally empirical (just like the SCS curve number method) models, SWAT's other built-in runoff function (Green-Ampt, a discrete simplification of the Richard's equation) is even closer to the physical description of runoff formation on small scale. So (i) SWAT's empirical nature should not be emphasised so much and (ii) should not be used as the main argument for developing HExM.

P 9225 L 25: A reference should be provided for the DNDC model.

P 9226 L 3-5: These can be deleted: "based on hydrology", "based on ecology", and "for environment".

P 9226 L 7-11: It is a pity that there is no Discussion section, because this sentence could be used to generate valuable content there. The coupling of (sub)models working on such different spatial and temporal scales poses severe theoretical difficulties. The outputs of submodels are only meaningful on their spatial and temporal scale. This means that a surface runoff calculated on a HRU or subcatchment scale cannot be directly used in a field-scale model of for example diffuse pollution. The multi-decade struggle of hydrologists to transfer knowledge between hillslope and catchment scales provides a nice illustration of these theoretical problems (see a summary in: Kirchner 2006).

P 9227 L 10: SWAT was criticised to be overly simplistic, but the authors chose a simplified PET estimation method instead SWAT's optional more accurate and robust Penman-Monteith due to limited data availability. On one hand this is a perfect illustration of the unavoidable tradeoff of modelling between detail and usability. On the other, it somewhat devaluates the arguments against SWAT.

C3567

P 9227 L 13: Same remarks as for P 9225 L 21

P 9231 L 1: Is eroded material anywhere retained along the overland transport paths? This is a significant factor that actually determines the particulate diffuse loads of streams. Usually a tiny portion of eroded particles manage to reach the nearest stream.

P 9231 L 11: Same problem as with SWAT: QUAL2 was blamed to be "subject to computational instability and time consuming due to its complexity" on P 9223 L 26. Then why is it used here? (I mean it's OK to use but then it should not be presented as something inappropriate for integrated modeling)

P 9233 L 2: It would be great to learn more about the uncertainty analysis method (Bayesian approach). According to section 3.2 it must be a kind of informal method because the evaluation functions presented there are no formal likelihood functions. But that's the most the reader can guess about the uncertainty assessment.

P 9234 L 9-18: Would be interesting to know the number of people living on the catchment.

P 9235 L 7: This sentence about LH-OAT is just confusing the following description of calibration by SCE-UA, so it could be moved to section 3.3.

P 9235 L 18-24: This paragraph could be shortened because most of these performance measures are well known.

P 9235 L 21: The optimal value of NS is not close to 1, it IS 1.

P 9235 L 24: What are the reasons for not using NS for NH<sub>4</sub>? NS does not have any criterium on the amount or frequency of measurements.

P 9236 L 1: The references here do not fully explain why the different objective functions have to have these weights. It seems to me that "Madsen 2003" and "Efstratiadis and Koutsoyiannis, 2010" both describe that different objectives can be merged into one by weighting, but I haven't found any explanation for these actual values. Are

C3568

these weights specially tuned for the Sha Ying catchment?

P 9236 L 5-9: This description is a bit confusing. I would rephrase this as follows: "Over 200 parameters (93 lumped, 112 distributed) control the hydrological, ecological and environmental processes of HEXM. To reduce the dimensions of the calibration problem we restricted SCE-UA to calibrate only the sensitive parameters as defined by LH-OAT."

P 9236 L 10: These parameter abbreviations (WMC, WM, ...) do not make too much sense when one has to look them up in the appendix. I would keep the textual definitions only.

P 9238 L 18: Calling an NH4 model as "environmental simulation" is a bit overstated again. Rather "Water quality" or even "Modelled Ammonium concentrations".

P 9240 L 10: Due to the differing sizes of subcatchments it does not make too much sense to report absolute annual yields. I would suggest to make them specific to catchment or cultivation area.

P 9243 Appendices: These could be significantly shortened if they would concentrate on those parts which are not published elsewhere. Or an even more complete description (with units!) could be provided as supplementary material.

P 9243 L 5: I would recommend to use the notation of surface runoff (for Rs), fast flow/interflow (for Rss), and baseflow (for Rg) instead of having 3 types of runoffs.

P 9255 L 3: "improved USLE"

P 9263 Table 1: Qual2K has a simplified 1D channel hydraulic model inside, but it doesn't do anything with hydrology.

P 9266 Table 4: What is the reason for the high bias of SWAT in certain subcatchments (Yingshang, Zhoukou) during calibration? One would assume that the calibration mechanism tries to eliminate bias as much as possible when calibrating.

C3569

P 9267 Table 5: Interesting to see NS values like -9 and -81. These must be as wrong as possible. What may be the reason?

P 9268 Table 6: For some stations (Zhoukou, Huaidian, Fuyang, Fantaizi) SWAT seems to be less biased, although HEXM has less variance in bias among the stations. Why is HEXM mostly wrong by about 30%?

P 9272 Figure 4: A part of the WQM figure was taken from the QUAL2 manual. This has to be indicated!

P 9276 Figure 8: There seems to be an issue with the dam regulation module because 4 out of 7 stations frequently have tiny simulated values when the observed flow was between 10 and 100. + How can the red trend line belonging to the grey dots be above the black when most grey points are below the black ones? (cf. Fuyang, low flow)

P 9277 Figure 9: The simulated outbursts of NH4 seem to be attached to very low simulated flow. Can these be identified as the seriously underestimated points on fig 8?

---

## REFERENCES

Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resour. Res.*, 42, W03S04, doi:10.1029/2005WR004362.

Mantovan, P., and E. Todini (2006), Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology, *J. Hydrol.*, 330, 368–381, doi:10.1016/j.jhydrol.2006.04.046.

Mantovan, P., E. Todini, and M. L. V. Martina (2007), Reply to comment by Keith Beven, Paul Smith, and Jim Freer on "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology", *J. Hydrol.*, 338, 319–324,

C3570

[doi:10.1016/j.jhydrol.2007.02.029](https://doi.org/10.1016/j.jhydrol.2007.02.029).

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 11, 9219, 2014.

C3571