

General comments and evaluation: This manuscript introduces a coupled, cross compartment simulation of the water and energy cycles using the Neckar basin as a case study. I very much agree with the authors that coupled simulations of water and energy cycles are a key for a) advancing our fundamental understanding of environmental system dynamics and b) to identify and rectify deficiencies in data assimilation schemes. The scope of the manuscript is hence highly suited for the audience of HESS and I think that the proposed coupled model bears a huge scientific potential.

Unfortunately, the implementation of the coupled model study and its scientific presentation in the manuscript are far below the quality standard required for a publication in HESS. In the present form the paper has no clear scientific objectives. Page 3 of the introduction reads very much like a project proposal which lists all possible advantages of virtual realities – yet the manuscript does not address a single of these possible scientific objectives. This is a missed opportunity! Instead the authors provide hand waiving arguments, that plausibility of virtual simulation results is sufficient to use the virtual reality for scientific learning. I think this is a) wrong (see major point below) and b) implies that the manuscript is not reviewable, simply because plausibility of model results is nothing that can be falsified based on the provided model evidence (if the authors have a different opinion, they need to explain how to measure plausibility in an objective sense). In consequence the manuscript presents a large set of diverse and possibly very interesting simulation results in a manner, which does not support a targeted scientific learning process beyond the fact the model may provide those results in a form that is in accordance with the mind setting of the authors.

Given the huge potential of the coupled model I strongly encourage the authors to re-submit a much more focused study, particularly with clear scientific objectives. I hope that the points listed below will be helpful for this. I have doubts whether this can be achieved within the period usually granted for major revisions, particularly also because the revision requires additional sensitivity tests with the model system.

#### Major points.

- In contrary to the authors' statement, I think that virtual realities are only suitable for scientific learning, if they portray non-linear systems dynamics and its sensitivity to meaningful changes in environmental characteristics in an acceptable manner. This needs to be tested using predefined evaluation criteria and acceptance thresholds, thereby avoiding bias correction, to avoid that we find what we wish to find. Data assimilation procedures which work well in an error-prone virtual reality, must not necessarily do a good job in reality, particularly not if the model is biased! A revised study could hence focus on the question whether the proposed model system performs already good enough to act as virtual reality, thereby exploring related model sensitivities. Even if this will be not the case yet, the study would be extremely interesting and valuable. Computational expense is not really a bottle neck here, as there are suitable methods to assess sensitivity of also of computational very expensive models within less than 50 runs. Another possible objective could be to quantify how much skill in water balance simulations stems from the fact that we usually drive the SVAT part of hydrological models with observed dependent data of air temperature and air humidity. In the coupled model this equivalent to the case of perfect predictions of T and air humidity in the reference layers.
- The referencing is absolutely inappropriate. The authors should acknowledge past work of competing groups in the area of coupled, cross compartment modelling, of water in energy

cycles (e.g. in Hohenheim and in Sweden) and they need to explain how their approach compares to those. The author should also check the considerable body of literature on virtual landscape studies in hydrology (e.g. by Hopp, Gerrits, Ebel and Loague and many more). This might be helpful to focus the revised manuscript on clear objectives and science questions and a more targeted evaluation of the simulation results with respect to these questions.

- I truly miss a critical reflection of whether the proposed coupling of COSMO, CLM and PARFLOW at the selected grid resolution does compromise the physical basis of the concept used to parametrize shallow turbulence, of the Richards equation and of Mannings roughness as well. This is particularly astonishing because quite a few co-authors have a very strong physical background. I do not claim that there is a simple answer to questions raised below. But the investigation of land-surface atmosphere feedbacks implies, beyond coupling of models, also to enhance the theoretical fundament for this. In any case the author should be aware that a credibility of the “physics based” model paradigm stands and falls with the way how honestly we deal with the current limitation of our theories.
  - The Richards equation relies on the assumption of local equilibrium conditions. The latter is surely not fulfilled at a grid scale of 100 by 100 m (Or et al. 2015 WRR). So what is the physical meaning of a soil moisture and a matric potential defined at this scale? How does these effective quantities to observables, particularly the effective potential to binding energy density of water in the pore space? What would be the effect on simulated root water uptake and transpiration and hence latent heat release when running PARFLOW on a 1 by 1 m grid (as recommended by Or et al. 2015), using the same approaches implemented in CLM. Is it really the average binding energy needs to be represented at this grid scale (or also minimum and maximum)?
  - Though I am a layman in turbulence and boundary layer meteorology, I remember that Monin-Obukhov theory relies on a the assumption of horizontal homogeneity and quasi steady state in the Monin Obukhov layer, because it is essentially a diagnostic approach to determine wind, humidity and temperature profiles close to the land surface. Particularly also Businger Dyer stability functions imply horizontal homogeneity. I do not think that this assumption is justified at the selected grid scale?! Particularly not with respect to the fact the length scales of topography, landuse and soil heterogeneity is smaller than 100m. So how to deal with this in the future?
  - A proper accounting for river net geometries is as important for flood routing, as preferential flow paths for subsurface flows (which are neglected as well). An adjustment of Mannings n to compensate for errors is simply unphysical (as n is related to the size of roughness elements) and implies that a parameter with a physical meaning degenerates to a fudge factor and a physical approach degenerates to a conceptual approach.

#### Technical details

- The authors use three pedo transfer functions to estimate their soil hydraulic functions, a short note on the scale miss match would be appropriate. These functions provide an uncertainty range for all values, why not assessing the related model sensitivity?

- As far as I know there are more than 60 different approaches for stomata conductance in the market, which one was used and how do the results depend on the choice?
- It seems that ordinary kriging has been used to regionalize soil texture and soil organic content. Please provide data on the underlying variogram (nugget to sill ratio, effective ranges etc.). Secondly it is not clear how conditional simulations were used to account for sub-grid variability and on which data this has been based. Please provide details.
- The kinematic wave approximation can be rather inappropriate for open channel flow, as the water level gradient is during events not parallel to the slope of the river bed. Why not using the diffusion wave approximation?

Best regards,

Erwin Zehe

References:

- Or, D., Lehmann, P. and Assouline, S.: Natural length scales define the range of applicability of the Richards equation for capillary flows, *Water Resour. Res.*, 51(9), 7130–7144, doi:10.1002/2015WR017034, 2015.
- Damour, G., Simonneau, T., Cochard, H. and Urban, L.: An overview of models of stomatal conductance at the leaf level., *Plant. Cell Environ.*, 33(9), 1419–38, doi:10.1111/j.1365-3040.2010.02181.x, 2010.