Dear Editor, dear Hubert,

First of all let me thank you on behalf of all co –authors for giving me so much time to revise this study. Secondly, we want to thank Keith Beven and the two anonymous reviewers for their critical but extremely helpful comments and suggestions, which have been condensed into a completely restructured and stream lined manuscript, which is 6 pages shorter than the first ones. The major changes are in the paper are the following:

- We selected a new title (avoiding the term new) which reflects much better our scientific focus (the interplay of spatial organization and functioning of intermediate scale catchments) and the main idea of this paper (a re-interpretation of the HRU concept from a thermodynamic perspective).
- This combination offers, as explained in section 1, several useful and partly also novel perspectives:
    1. For defining functional similarity based on similar terrestrial controls on the pair of gradients and resistance terms controlling different land surface atmosphere energy exchange, rainfall runoff transformation and base flow production. We propose that this implies a scale hierarchy of "specialized" HRUs we call functional units (as further explained in section 2)…
    2. On alternative strategies to characterize intermediate scale catchments in replicate members of a functional units, focusing on the triple of flux, gradient and resistance (as explained in section 3)
    3. Requirements for structurally adequate models and options to reduce inherent equifinality based on information on controls on the pair of gradient and resistances (as explained in section 4.1 and 4.2)
    4. An energy centered view on hydrological dynamics (including rainfall runoff transformation), including the functional advantage of organized structures from a thermodynamic optimality perspective (as explained in section 4.3).

Sections 2-4 further elaborate these propositions, discussing pro and contra arguments. Each of those section refers, as recommended by the reviewers, to work and papers we regard as pioneering/ benchmarking in this respect, without providing an exhaustive literature review to stay as brief as possible.

Section 5 presents, as recommended by reviewer two, a much shortened outlook on our ongoing research that aims to test these four main propositions. Although reviewer 2 recommended skipping this outlook completely, we decided to share to our a–priory ideas and concepts how to advance characterization and modelling of catchments of organized complexity before publishing our findings in research papers. We do this on the risk of being proven to be wrong in the future, to vote for a publication culture that allows learning from failures. We think that opinion papers could exactly serve this purpose.

As a line to line response to the reviewer comments is, because of the new structure of the manuscript, not appropriate, we attach a brief list how we addressed the main reviewer comments, as well as our detailed responses to the reviewers from the open discussion phase.

Thanks again for your kind patience.

Best regards,

Erwin Zehe

Brief reply to Keith Beven

We sincerely thank Keith Beven for his critical but very helpful comments on our manuscript. We addressed most of them as summarized in the following:

- We selected a new title (avoiding the term new) which reflects much better our scientific focus – which is to contribute to better experimental characterization and modelling of intermediate scale catchments of organized complexity.
- Section 1 highlights that our intension was not to criticize the HRU concept, but to provide a re-interpret from a thermodynamic perspective. This combination offers, as explained in section 1, several useful and partly also novel perspectives:
    - For defining functional similarity based on similar terrestrial controls on the pair of gradients and resistance terms controlling different land surface atmosphere energy exchange, rainfall runoff transformation and base flow production. We propose that this implies a scale hierarchy of "specialized" HRUs we call functional units that function similarly with respect to one specific form of "water release", because the terrestrial controls on related gradients and resistances have different characteristic length scales (as further explained in section 2)…
    - On alternative strategies to characterize intermediated scale catchments by clustering observations in replicate members of functional units, focusing on the triple of flux, gradient, resistance (as explained in section 3).
    - Requirements for structurally adequate models, which should rely on thermodynamic consistent equations (flux= gradient 1/resistance) and disentangle matrix flow from vertical and lateral preferential flow. This offers options to reduce inherent equifinality, arising from the interaction of gradient and resistance terms in the governing equations, based on available information on controls on the pair of gradient and resistances (as explained in section 4.1 and 4.2)
    - An energy centered view on hydrological dynamics (including rainfall runoff transformation), including the functional advantage of organized structures from and thermodynamic optimality perspective. (as explained in section 4.3)
- In section 3 we stress that our proposed experimental strategy to cluster multiple methods in replicate members of functional units is in fact only feasible, if HRUs or the more specialized functional units do exist in the landscape. This is because this

strategy essentially relies on the idea of exemplary experimental learning. Thereby we acknowledge that an experimental test whether those functional units exist, can due to small samples neither operate at a high level of significance nor assure sufficient power to avoid a second kind error. We thus stress, as recommended by Keith Beven, that transferability of behavioral model parameter sets (as teams) is genuine test for the concept of functional units.

- In section 4.2 we refer to papers and models we regard as pioneering/ as benchmarks on the search for structurally adequate models for intermediate scale catchments (which balance complexity with parsimony).
    - We discuss pros and cons of the hillslope storage Boussinesq model and most importantly of the REW approach, thereby giving credit to the holy grail paper, as having pointed out the cardinal problem of deriving closure relations and in particular storage-discharge relations.
    - We also point out that thermodynamic consistency of the model equations, in the sense that they disentangle gradients and resistance terms, does not imply that a model needs to be based on partial differential equations. In this context we acknowledge WASA, mHm and Topmodel and Dynamic Topmodel.
- In section 4.3 we briefly discuss the present state of the arte with respect to thermodynamic optimality, summarize perspectives for their test within independent predictions. We then report on promising findings, thereby acknowledging that the value of these principles is still strongly debated. We also acknowledge that a test of concept based on successful uncalibrated predictions, relies implicitly on the strong assumption that the model is an acceptable representation of the system, accounting for its degrees of freedom and the feedbacks between processes that form structures and their impact on water and energy flows (which is beyond the scope of all environmental models that are currently available).
- Unfortunately, we did not deepen the discussion on similarity metrics (although we would love to do), simply to stay brief and not to provide a mixture of an opinion and research papers.

Please find much more details in our reply to your valuable review during the discussion stage (some of them might be a bit outdated). Again let me thank you very much for the valuable comments,

Erwin Zehe

Detailed response to Keith Bevens review, taken from the discussion section

We, (Erwin Zehe as EZ on behalf of all co-authors in the following) sincerely thank Keith Beven (KB in the following) for his critical assessment of our opinion paper (CAOS paper in the following). It is always a pleasure and a challenge to start an argument with KB because as he has, as a leading thinker, indeed tried and partly rejected many of the ideas we propose in the CAOS paper. Yet, we strongly disagree with KB that all ideas presented in the CAOS paper can be rejected outright (of course this is left to the Editor). Simply because also a major authority and leading thinker can fail in his assessment (think about Einsteins argument against the immanent stochastic nature of quantum mechanics). Before we intend to explain that this is also the case here, we shortly reflect on our understanding of an opinion paper and why opinion papers become fashionable: An opinion paper presents an opinion of several individuals for instance on possible innovations or critical issues of current scientific practice. As it is not a review, the referencing cannot be exhaustive, though of course important contributions should be cited. There might be many reasons to publish opinion papers: either to stimulate a debate or maybe also because it is faster to present and defend what we think than to present and defend what we have done.

Our motivation to write the CAOS paper was to share the true a-priori research questions/ hypothesis driving the ongoing joint CAOS research. Why so? Some of us (especially EZ) suspect that sometimes the a posteriori synthesis of research is presented as hypothesis maybe because it is easier to publish success stories (this is of course personal opinion reflecting personal experience). Science history shows, however, that we learn even much more from our failures! The hypotheses, ideas and approaches presented in the CAOS paper have been accepted by an international jury of experts and we have interesting findings to be published in the forthcoming research papers from two years of research work (some match our expectations, some are truly surprising, at least to us). By presenting the initial ideas beforehand; the follow up research papers from CAOS will tell how much of these initial ideas will be corroborated or rejected. This is what we mean with none white washing scientific learning.

KB: missing reference to key papers dealing with these issues

EZ. In contrary to KB astonishing suspicion, we did not leave out references on purpose to fish for citations in an unfair manner. In fact we went back to the classics with our referencing with respect most contributions (maybe not all) which motivate or concept (organized complexity, HRU definition, the catena concept, the pattern process paradigm, predictive uncertainty and organizing principles). The reason why did not refer to the REW and other promising model concepts (hillslope storage Bousinesq model, Dynamic TOPMODEL), was not to claim our concept to be superior, but simply to be brief (maybe too brief). We admit that the REW approach is pioneering with respect to joint treatment of the mass, momentum, energy and entropy balance in larger control volume (so is the work of Troch on the hillslope storage Bousinesq model). With respect our hypothesis H2 we regard the REW approach, however, as too simple. Most of the applications of either REWASH, CREW, or THREW treat REWs and sub catchments as equal. This implies averaging over different ensembles (HRU!) and lumping of (soil) resistances and gradients into sub catchment scale averages. This is too simple at intermediate scales (Zehe and Sivapalan 2009) and leads to serious problems for instance when trying to close the equation for overland flow. Velocity depends on the square root of the gradient and spatially variable roughness along the flow path…. With such a kind of averaging we get simple equations. But we do not get rid of the complexity, it is just hidden in the closure relations. Boundary layer meteorology faces exactly the same problem: simple diffusion like equations for turbulent fluxes (in case of first or 1.5 order closure) and the entire complexity is in the turbulent exchange coefficients.

KB: nothing new about the EFU idea?

EZ. The CAOS paper suggests several innovative or at least useful approaches for better understanding intermediate scale catchments. Some of them are, as the EFU idea, not novel per se (we never claimed they are!); but the "just" reflect our ideas to add the "necessary" details to the most promising HRU concept. Others are, as thermodynamic optimality or the use of species distribution data as proxy for macroporosity, new in hydrology and thus naturally at a state of a hypothesis. In fact none of the ideas proposed in the CAOS paper is entirely new in science, but in combination they can very useful to characterize intermediate scale catchments and to link experiments and model. Our ideas might of course need a more rigorous explanation and definitions in fact they can be developed from a single theorem, which is again not new but well known in thermodynamics.

Theorem 1: Any kind of flux is equal to a "potential gradient" $\nabla\phi$ (temperature gradient, water level gradient, concentration gradient, soil water potential gradients) divided by a resistance R (inverses of either heat conductance, surface roughness, diffusion coefficient, hydraulic conductivity…). The former determines the (thermodynamic) force the latter determines dissipative energy losses along the flow path:

$$\vec{q} = {}^{1}/_{R}\,\overrightarrow{\nabla\phi} \quad \text{(Eq. 1)}$$

This theorem implies a hierarchy with the gradient on top as there is no dynamics without a force; (Thermodynamic forces are gradients of intensive state variables, which are continuous at interfaces, a soil moisture gradient is thus not a force).

Corollary 1: In larger control volumes of terrestrial systems both gradients and resistance are fields and depend on almost static controls and on system state variables (all this is well known). In this framework HRUs or functional units can be defined as classes of landscape entities/control volumes with similar terrestrial controls on the pair of $\nabla\phi$ and R (which directly leads to the hierarchy we propose in the CAOS paper as explained in definition 1). Relevant potential gradients in hydrology are:

- Soil water potentials, plant water potentials with respect to green/capillary soil water fluxes;
- Piezometric heads, surface and subsurface water level gradients with respect to blue water flows in the aquifer, in preferential pathways and the river network;
- The divergence in radiation fluxes causing near surface gradients in temperature and air humidity driving latent and sensible heat;
- The listed gradients are associated with differences in different free energy forms (capillary binding energy which is in fact chemical energy, potential energy…).

Corollary 2: Resistances in Eq. 1 are more than just a material/continuum property; 1/R is a tensor, reflecting the spatial heterogeneous and spatial organized arrangement of for instance soil material in the control volume. Subsurface flow resistance depends for instance on soil hydraulic conductivity $1/k(\theta)$, its covariance lengths and soil moisture. Connected structures (preferential pathways lateral pipes, vertical macropores) reduce the control volume resistance at a given driving gradient as they allow for advective flows, resulting in accelerated fluxes and shorter residence time distributions. In line with this idea vegetation is a preferential flow path for green water into the atmosphere.

Corollary 3 : Eq. 1 is immanently subject to equifinality because it is not an injective function (several elements of the start domain are mapped on the same element in the codomain):
- Several combinations of gradients and resistance compile the same flux. This might be frequently the case in hydrological systems as the (quasi static) controls especially on gradients driving lateral flows of blue /free water and on flow resistances (especially on vertical and lateral preferential flow paths) are partly independent in the landscape.
- The R term is none, as for instance preferential flow networks with different topologies and hydraulic properties may result in the same control volume resistance (e.g. Klaus and Zehe 2010). This source of equifinality cannot be eliminated and has to be accounted for.

In line with the argumentation of Bardossy (2007), the first source of equifinality can be reduced based on observations that characterize at least two out of the three variables (either q and R or q and $\nabla\phi$, or $\nabla\phi$ and R). Current observation technologies allow approximate characterization of static controls on gradients driving lateral flows of blue water (surface and bedrock topography). This is why we think a coupled treatment of the mass and momentum balance has more pros than cons and stands as the main supportive argument for hypothesis H2.

We also present two promising avenues to tackle the resistance problem.
1) Some preferential pathways (unfortunately the dead ended ones) are created by ecosystem engineers. There is a chance to get approximate information on their spatial pattern using species distribution models as explained in the CAOS paper.
2) Organizing principles allow for a priory optimization of the resistance term at a given gradient, either as a bulk resistance (Porada et al. 2011; Westhoff and Zehe 2013) or the

density of vertical and lateral macropores (Zehe et al. 2013; Kleidon et al. 2013). This implies the possibility of independent predictions. These can of course go wrong, but this is testable! And this is the main reason for hypothesis H3.

Definition 1: Based on corollary 1 we define the HRU/functional units as classes of landscape entities/control volumes with similar terrestrial controls on the pair of gradients $\nabla\phi$ and resistance R controlling either land surface energy exchange or rainfall runoff. (Note this this is necessary conditions for functional similarity, but not necessary and sufficient conditions as Eq. 1 is not an unique). According to Flügel (1996), 'Hydrological Response Units are distributed, heterogeneously structured entities having a common climate, land use and underlying pedo-topo-geological associations controlling their hydrological transport dynamics'. We think the definitions match well, our one has maybe a little bit more physical rigor and that it implies a) hierarchy of functional units instead of a one fits all HRU and b) that their dominance changes dynamically with prevailing boundary conditions.

KB: Why a hierarchy and where is the innovation?

Definition 1 and Theorem 1 imply that there could be a scale hierarchy of functional units, due to a scale hierarchy in the terrestrial properties which control the gradients driving land surface energy exchange (during radiation driven conditions) and lateral flows of free water (during rainfall driven conditions) and also because the terrestrial properties controlling the related resistances have different characteristic extends. In line with this we think that the fluxes dominating either the energy balance or rainfall runoff transformation operate at a hierarchy of different characteristic spatial extends (or REVs) are driven by a hierarchy of different gradients (with respect to type, strength and direction), feed from different water sources (capillary bounded 'green' water or free 'blue water) and are facilitated by different types of network like structures (including vegetation).

We thus step beyond the idea of a "one-fits-all-HRU" and postulate a spatial hierarchy of functional units named lead topologies and embedded elementary functional units EFUs which act similar with respect to either rainfall runoff production or the energy balance. As sketched in Figure 3 EFU's are deemed to act in parallel during radiation driven conditions, controlling the radiation balance, the Bowen ratio, soil heat flows and upward vertical flows of capillary water in the soil matrix. The key terrestrial determinants are slope and aspect (determining exposure to global radiation) plant and soil albedo (determining the net radiation), soil type and depth to bedrock (determining partly the retention properties and thus how soil water potential evolve). The big unknowns are macropores and the functioning of vegetation itself, which are for us the key determinants for the resistance terms. As EFU's control the energy balance, they control also the water balance i.e. partitioning of rainfall into ET and runoff. This might explain why land surface models in meteorology do a good job in reproducing the energy balance, though they do not account for lateral flows.

The next higher scale level is the hillslope scale, which determines terrestrial controls on lateral potential gradients driving lateral flows of free blue water. Gradients build up at inclined material interfaces i.e. the land surface, the bedrock surface and the groundwater surface. Expect for groundwater-dominated systems these potential gradients are thus largely determined by the morphology and topography and permeability of these interfaces as explained in the CAOS paper. EFUs act in a series during rainfall driven conditions and get interconnected by lateral flows either at the surface, in subsurface lateral drainage networks or at the bedrock interface (all big unknowns it admit). The music concerning similarity plays no longer at the EFU scale but at the hillslope scale. So why not simply speaking of functionally similar hillslope classes? This is because riparian zones function in a different form but also controls lateral blue water flows. Hillslopes are not permanently hydrologically connected to the stream, riparian zones are. But both hillslopes and riparian zones control lateral blue water flows and both may consist of several different EFUs. Maybe PCU potentially connected units is the better word than lead topologies.

We also suggest that dominance of functional units is nothing static, but depending on the situation similar EFU will act similar (in case they receive similar forcing) or similar (lead) topologies. Based on this perception, we developed a hierarchical combination of EFU and lead topology objects, which are based on simplified but physically consistent process descriptions, with the river network and a groundwater domain as catchment scale objects. As the CAOS model consequently disentangles matrix flow and preferential flow into separate process domains, we hypothesized that a) model parameters that control the energy balance at the EFU level can be estimated independently from those that control rainfall runoff at the lead topology scale and b) acceptable parameter sets should be transferable among class members of the same functional unit (compare supplement 'CAOS model and verification' that will be attached to the revised manuscript). There are at least a couple of examples that functional and structural model parameters are transferable at the hillslope scale (Weiherbach catchment; Zehe and Blöschl, 2004) or where the entire Mallacahuello catchment can be presented by a single hillslope (Zehe et al. 2013). As second test of concept is whether this concept allows to avoid redundant calculations by dynamic grouping.

We, furthermore, propose and established a stratified observation network drawing from process hydrology, soil physics, geophysics, ecology and remote sensing <u>in replicates</u> of candidate functional units in the Attert River basin (Luxembourg), to search for typical and similar functional and structural characteristics. Expect of the B2-LEO, we do not know other experimental studies which conduct identical experiments and monitoring in replicate control volumes, (at the field and slope scale) to experimentally check the idea whether functionally similar control volumes can be detected. Of course we are still at a stage where the experimental design is a hypothesis itself - in the future this will for sure to be refined or even rejected. We are aware, anticipating KBs comment, that inversion of geophysical proxies are non-unique and yield at best site specific petro-physical relations. In fact we do not invert these data, but compare them for instance with augers to pick horizons and derive estimates for subsurface structures to constrain the model (of course in an hypothesis based manner) our we make use of time lapse GPR to detect water flow subsurface structures (this works, but of course we cannot quantify the flow rate exclusively with it).

KB has doubts on the idea of co-evolution, because this is contaminated by management

EZ: Agreed! Our first guess predictors for detecting EFU in a given geological setting are the topographical positions, landuse, hillslope aspect, soil type and of course management practice. The latter plays a key role as it controls either the age spectrum and species composition of trees in forest areas or surface preparation (roads, etc), optionally cutoff of macropores, and selection of crops in agricultural areas. However, also the different forms of landuse might have well adapted /co-evolved to what makes sense/ brings best profits in different landscape compartments (Savenije 2009) and is thus not totally independent from the landscape evolution. In this sense there might be much to be discovered with respect to the question why landscape are managed the way they are. But there is also much to be discovered with respect to the questions whether co-evolution exists and optimality principles are helpful to describe steady state configuration of a geo-ecosystem (or potential natural state if you wish) or not.

KB: The HRU concept does not neglect exchange

True, the concept does not neglect and we will correct for this in the revised manuscript. But its model implementation does pretty often neglect exchange (at least at the hillslope scale). Due to my experience with PRMS, in contrary to the picture you mention, HRU are represented by the same parameter sets, but their interaction is not treated in a spatially distributed manner. Water does simply not flow downslope driven by potential gradients. We regard this as being too simple in intermediate scale catchments, and a missed opportunity with respect to reduce degrees of freedom in the model (as already explained)

KB: Organizing principles are too speculative

EZ: We respect and share KB concerns and fully agree that there is more to be done to explore their practical value either for uncalibrated predictions or for estimating hydro-pedological characteristics. This is exactly one major objective of the research outlined in the CAOS paper. In contrary to Keith we think there are promising results which justify a high expectation and the search for clear tests. Westhoff and Zehe (2013) clearly showed that thermodynamic optimality (TO) is of little use for constraining model parameter sets of conceptual HBV type models. The intersection between parameter sets matching the water balance and those that maximize entropy production was small in case of varying a single parameter (this is what we wish), but unfortunately zero in case several parameters are varied. Water flows in these models are not driven by gradients in intensive state variable /potentials and our effort to define proxies didn’t succeed.

When using a process based model (Zehe et al. 2013) the results were however promising for two distinctly different landscapes. In the Weiherbach catchment, where capillary binding energy dominates free energy dynamics of soil water, we found that the thermodynamic optimum surface density in macroporosity yielded acceptable uncalibrated rainfall-runoff predictions. In the Mallalcahuello catchment in Chile, where free energy dynamics of soil water is dominated by it potential energy, we were able to estimate the annual runoff coefficient based on assuming long term steady state with respect to potential energies of soil water during rainfall runoff. Furthermore, Kleidon and Renner (2013) developed a very simple model for the land surface energy exchange based on TO which performed without calibration not too bad against flux tower data at three different sites. Still all these promising findings might be just by the matter of coincidence (the atom model of Bohr worked nicely for hydrogen but failed for all other elements, just by chance …). We thus need a more rigorous testing here that allows rejection of these principles. The challenge is to find a closed model experiment or real experiment that accounts for all the necessary positive and negative feedbacks between structure formation and the processes which are controlled by these structures…

KB: The holy grail paper already discussed all the necessary aspects of functional units.
EZ: I was as an eponymous reviewer of this paper as critical as KB with the CAOS paper. I admit that storage discharge relations are certainly important for one (storage and release of free/blue water), but not for all catchment functions. This includes land surface energy exchange, as we need capillarity here, but also forms of runoff generation that are not a strictly monotonous function of storage (for instance bypassing, Horton overland flow). As far as I remember the paper refers strongly to the REW concept and thus to the sub catchment scale, which is not the entire story; as we think the hillslope is equally important for understanding intermediate scale catchments. A reference to Jim Kirchner (his work on catchments as simple dynamic systems) would be even more appropriate as he developed the idea of storage discharge relations to something of practical use. In fact we tried his ideas, in the Attert but not with too much success up to now.

Keith: mentioning of the review on preferential flow
EZ: We define preferential flow in general rapid advective movement of water and solutes. The clue is that we have large fluxes even in case of small driving gradients, as specific dissipative frictional losses are small. In this sense flow in pipes, macoropores, cracks, surface rills network and the river are preferential flow.

Keith: The use of entropy is a little superficial
We agree that this passage is not exhaustive and relates to one interpretation of entropy An exhaustive explanation of how entropy production and export relates to hydrological processes based on the Clausius definition can be found in Kleidon et al. (2013) or Zehe et al. (2013). By the way absence of gradients as state of maximum entropy follows directly from the Clausius definition and related definitions of thermodynamic potentials (such as Gibbs free energy). The link of entropy and information and distribution of possible mircostates that belong to the same observed macro state is furthermore well established in statistical mechanics and standard part of lectures on thermodynamics. We will add a reference to Kondepudi and Prigogine (1999), they further explain why a strong spatial covariance is organized compared to a white noise (although both have the same variance). We explicitly referred to the river and catchment because covariance is not a good measure for connectivity (as the invariance properties are different).

Keith: Mentioning Imbeaux (1897)
EZ: Thanks for this, we will change the phrasing. Still we regard Jims work on organized complexity (Looking for hydrologic laws) as key paper in this area.

Keith: partly poor/strange wording
EZ: Sorry for the bad English. With “superordinate” we mean gradients that dominate flow at the next higher scale and thus hierarchy level. We selected the name lead topology to reflect our perception that the arrangement of EFU with its (textural properties) along the gradient controlling lateral flow/free water is of key importance as it controls hydrological connectivity for stream flow generation during rainfall events. Data gridded conceptual models are insensitive for “flipping” of the soil catena as most of them (HBV, LARSIM, WASIM), PREVAH) assume that the gridded elements contribute in a parallel manner to runoff production. Water does not flow downslope, yet these models work well in the input output paradigm.

Keith: what means unique range of settings
EZ: What we meant is that the Attert observatory covers, in a nested design, 9 sub catchments ranging up to 250 km^2 with 4 different geologies, different land uses and a large climate gradient. This is not unique in the world but very rare to have such an observatory with catchments of mixed and clean geologies so closely co-located.

KB What is the role of Non-Gaussian transport for the hydrograph
EZ: Agreed! Displacement of old (partly well mixed) water contributes largely to the slow branch of residence times and the hydrograph. But predicting the hydrograph is not sufficient if we go for water driven transport (which we want), where preferential flow operates at the fast part of the residence time distribution. Macropores affect, as Keith knows, the hydrograph in several ways. For instance by enhancing infiltration and reducing connectivity of overland flow paths. (. This of course also affects groundwater recharges (and the mixed part on the residence time distribution. But there is also experimental evidence that vertical and lateral preferential flow strongly contribute to spring hydrographs, as shown in a series of two tracer experiments performed by Wienhöfer et al. (2009) in Austria (tracer breakthrough curve yielding Peclet number around 2 after nearly 30 m lateral transport into the spring). This implies preferential transport in the near field, which is nicely corroborated by a related model study (Wienhöfer and Zehe, 2014). Jim Kirchner provided similar evidence for this in a beautiful talk (Kirchner, J.W., Reflections of preferential flow at the hillslope and catchment scale, Monte Verita International Workshop on Preferential Flow and Transport Processes in Soil, Ascona, Switzerland, November 2006.)

KB: There is earlier work explaining mobilization of prevent water by pressure transduction as being related to the momentum balance
EZ: We are sure it is, because it is straight forward from a physical point of view. We put this citation in as an example. It is not that we claim this insight to be new. But it is, according to

Corollary 3, a relevant argument that spatially explicit models, which can at least in a simplified form account for the momentum balance, are structurally more adequate and thus potentially less uncertain (see reply to next comment).

KB: Equifinatility due to lumped treatment of gradients and resistances

EZ: We apologize for putting the wrong reference and correct it. This was meant to give the merits to you and Andy, as you put the finger on the wound that we immanently solve ill-posed problems with hydrological models and that our solutions space is therefore infinite (this contribution will last). With respect to state of the art in hydrological process research, near surface geophysics and scientific computing we suggest that explicit treatment of the momentum balance in hydrological models offers more advantages than drawbacks (as already explained).

KB: on transpiration

EZ: Agreed, partitioning of energy into latent and sensible heat is controlled by transpiration and thus by vegetation. This is also explained later in the same section of the CAOS paper. Due to theorem 1 we always address the gradient first (because this is the driving force related to the radiation balance) and discuss then the controls on the resistance terms (this is where vegetation plays, of course again with internal potential gradients)…

KB: Richards equation is falsified per se

EZ: Bashing the Richards equation as being wrong per se (in fact Darcy Buckingham concept) is a little bit too general for my taste. Darcy Buckingham is in fact a diffusion concept and can as such not deal with fast advective processes. Agreed, but it is our problem that we expect it to work there! We also agree that potentials, relying on local equilibrium, are not well defined when flows get fast.

We think soil water potentials are well defined during radiation driven conditions and we need capillarity to describe what we see: rising water performing work against gravity! We do not know any better concept than the matric potential to account for this. Maybe Keith knows, and we are happy to use that. Capillarity is present at many scales, reflecting that water acts as a wetting fluid and the fact that soil are porous media. …. One cannot like capillarity (I do without this there wouldn't be any water storage against gravity) but one cannot ignore it. (Note Newtons mechanics is wrong, as it fails at the quantum scale, still we use Newtons law in classical mechanics)

Keith: Annoying similarity to Beven JH 1989 (no reference given at the end)

EZ: I hope KB forgives me that I do not know all his papers by heart (simply too many). As explained I did not leave out references on purpose, possibly I never read this work. I am happy to refer to this when you provide me the reference. However, the references in this passage are pretty new which show that the entire issue is still under debate.

KB H1:

EZ: According to Flügel (1996), 'Hydrological Response Units are distributed, heterogeneously structured entities having a common climate, land use and underlying pedo-topo-geological associations controlling their hydrological transport dynamics'. The underlying assumption is that a similar structure is a sufficient proxy for predicting a similar hydrological functioning. We follow this idea but further refine as we regard a one fits all HRU as inappropriate, as the processes governing landsurface energy exchange and rainfall runoff process operate at different characteristic scales and are controlled by partly different landscape properties (see above). We already stressed in accordance with KB that transferability of the parameter sets among different units at the different hierarchy levels is a key benchmark.

KB: H2: There is no simple solution for the closure problem

EZ: Agreed energy closure is not simple: This is exactly what we point out with the discussion of land energy feedbacks as fundamental challenge. I do not think that coupled modelling of water and heat is a problem in soil (this goes back to de Vries work in the 50ties, despite that Darcy Buckingham has problems with fast, gravity driven flows)

Our momentum closure is adaptive to the context:

- For capillarity driven conditions we take Richards (coupled with head balance and A-D equation) unless KB names something better.
- For rainfall driven conditions we assume that flow preferential in drainage structures dominates and assume quasi steady state (as many do). The potential gradient is either one (in case of vertical flow) or equal to the water level gradient. The key problem is proper accounting for frictional losses, we do this with Darcy-Weißbach in case of lateral preferential flow (compare also supplement 'CAOS model and verification').

KB: H3

EZ: We already explained this above.

KB: 3268 the issue of fast growing roots

EZ: agreed, in case roots can grow that fast.

KB: EFU defined how 3271

EZ: All members of a EFU class that belong to the same ensemble with respect to time invariant controls/steady state controls on the gradients and the resistances that determine the energy balance (interception of radiation, slope, aspect and albedo) and of green water and heat fluxes (retention properties and depth to bedrock, thermal properties). Candidate EFU are defined as being homogeneous with respect to soil, aspect, hillslope/ catena position This implies homogeneous habitat conditions for key ecosystem engineers as explained in the CAOS paper This is to be corroborated by similar dynamics of sap flow, average soil moisture and potential dynamics and temperature dynamics (as H0) and transferability of the related model parameters.

KB REV, EFU Size

EZ

- In general we define upper and lower boundaries of EFUs classes by (inclined) material interfaces i.e. the landsurface and the bedrock.
- Lateral boundaries are also marked either by a significant change in textural properties or a change in the density of vertical preferential pathways as both determine vertical control volume resistance R˙. As an EFU is a means of separating a scale of functional homogeneity (Reggiani and Rientjes, 2005) it has to be much larger than the covariance lengths of porosity, hydraulic conductivity and of the vertical and lateral length scale of the macropore network to assure ergodic conditions for gradients and flow resistances. At grassland sites these correlation lengths appear to be in the order of a few meters as explained in the CAOS paper. Changes in the surface density of vertical macropores are controlled by ecosystem engineers (earthworms, rodents) and vegetation and their specific manner to build these flow structures. As both find different habitats when changing the aspect (as radiation input is different), we expect aspect to be an important discriminator for evolution of different pedological EFU classes.
- On the other hand an EFU must be small enough to assure local thermodynamic equilibrium within a soil layer, which implies that a laterally uniform matric potential is well defined at this scale (Zehe *et al.* 2006, Vogel and Ippisch 2008). During rainfall driven conditions this extent can be deemed as very small, as rapid flows in connected structures disturb local equilibrium. Field studies of Brocca *et al.* (2007), Blume et al. (2008) and of Western *et al.* (1998) report however that soil moisture patterns are temporally stable not in the sense that values are themselves stable but their ranks within the probability distribution do not change over time. Zehe *et al.* (2010) found consistent results for two sites of 20 by 20 m that were instrumented with 40 TDR sensors respectively. Ranks of the distributed soil moisture time series were stable in time, especially during energy driven conditions. We thus suggest that matric potential at a constant depth is during energy driven conditions rather homogeneous at this extent, even if the pore space and soil moisture are heterogeneous. Otherwise small scale soil moisture variability wouldn't persist in time but be smoothed out by lateral flows. We expect thus that an EFU has typically a lateral extent of approximately 25 by 25 m.
- Note: This is why we go for Richards, heat balance etc. during fair weather but alternative approaches h during rainfall driven conditions.

KB: How to experimentally infer new understanding on storage, mixture and release

EZ. Past investigations in the Alzette and Attert River basins in Luxembourg have demonstrated the first order control of geology on winter stormflow coefficients (e.g. Pfister et al., 2002). Recent investigations in our nested catchment set-up have further suggested geological controls on: isotopic signatures in baseflow and catchment dynamic storage (as per Sayama et al., 2011). We have been able to document that isotopic and geochemical signatures in streamwater exhibit a large variability between catchments, but also along individual catchment flow duration curves (FDC) – e.g. concentrations increasing/descreasing in the lower part of the FDC and stabilizing in the upper part of the FDC (Pfister et al., in prep.). Further investigations are needed to understand what implications this variability has on basic assumptions related to end-member mixing analysis, as well as time variant transit time and transit time distributions. Building on data from our nested catchment set-up, we also see potential here for a storage-based catchment classification scheme.

KB: 3274 Importance of drainage via fractures/dip layers into the deep subsurface supplying the energy balance

EZ: This is an important point in the presence of deep rooting plants that may feed on this water source. In case of a pristine landscape we believe that vegetation will adapt, because some of them might take advantage of these niche.

KB: Persistence of soil moisture patterns

EZ: Indeed this reflect different in retention properties, which implies that soil moisture variability is important for storage, it does not reflect the difference capillary binding energy during radiation driven conditions (see comments on EFU extent).

KB: Thermodynamic consistency

EZ: Let us explain what we mean. Different forms of free energy (note this is not conserved as dissipation is a sink), are products by conjugated pairs of an intensive (continuous at interfaces, non-additive, such chemical potential (related to soil water potential), velocity, pressure) and extensive state variable (discontinuous, additive, mass, momentum, volume). Gradients of intensive variable drive fluxes. As reductionist models account for this they allow to trace these free energy conversions. This is why the CAOS model (note this without H ☺) is based on PDE, however, with weak coupling.

KB: Lack of convincing evidence that optimality principles hold

EZ: Agreed, for instance as optimality refers to steady states during steady state. This is not too easy to be detected. Note we do not claim they are true but we search for scientific tests in an open and unbiased manner (compare details above).

KB criticizing the summary of our ideas at beginning of chapter 4

EZ: There is an imbalance concerning the amount of precision KB is asking for in our statements and his often imprecise formulations. This section is by far presented as a conclusion based on findings but a summary of our propositions that need to be tested. This is clearly stated and we never claimed that we have an easy answer. We also never criticized the HRU idea per se (we admire it and believe in it and try to advance it) nor the suit of methods that haven been proposed to identify HRU's. We criticized the implementation in models (this will be stressed more clearly) which often, not always, ignores exchange and that a test of concept is missing.

Such a test of concept cannot exclusively rely on observations (even if they are conducted in replica) as we have to test a null hypothesis e.g. members of candidate EFU belong to the same ensemble with respect to the energy balance and related green water and heat fluxes. The level of confidence will be low as the sample of observations of sapflow or soil moisture is small within the EFU even if we employ time or space substitution. This exercise must thus be essentially combined with a test, whether the model structural and functional parameter are transferable within class members at the same hierarchy level.

We would be very happy to provide details on the model and the model falsification concept (also dealing with the inverse problem) as well as on the metrics issues. This could be addressed within two separate supplements of the revised manuscript or course again in a manner that remains hypothetical unless we add results (which would blow the entire story).We think that reviewer 2 is pretty right that we might have overdone it a little with our wish to make all or initial ideas transparent beforehand; as such a level of detail too much for an opinion paper. A little bit less is maybe more.

Thanks again very much for the good and constructive points. Allow us to assure you that we would have taken them equally serious, if they had been communicated with a little less sarcasm.

Erwin Zehe

References

Blume, T., Zehe, E., and Bronstert, A.: Investigation of runoff generation in a pristine, poorly gauged catchment in the chilean andes ii: Qualitative and quantitative use of tracers at three spatial scales, Hydrological Processes, 22, 3676-3688, 10.1002/hyp.6970, 2008.

Blume, T., Zehe, E., and Bronstert, A.: Use of soil moisture dynamics and soil moisture patterns for the investigation of runoff generation processes with special emphasis on preferential flow. , Hydrology and Earth System Sciences, 13, 1215 -1233, 2009.

Brocca, L., Morbidelli, R., Melone, F., and Moramarco, T.: Soil moisture spatial variability in experimental areas of central Italy, Journal of Hydrology333 356-373, 2007.

Flügel, W-A.: Hydrological Response Units (HRUs) as modeling entities for hydrological river basin simulation and their methodological potential for modeling complex environmental process systems. Die Erde 127,42–62, 1996.

Klaus, J., and Zehe, E.: Modelling rapid flow response of a tile drained field site using a 2d-physically based model: Assessment of "equifinal" model setups, Hydrological Processes, 24, 1595 – 1609, DOI: 10.1002/hyp.7687., 2010.

Kleidon, A. and Renner, M.: A simple explanation for the sensitivity of the hydrologic cycle to surface temperature and solar radiation and its implications for global climate change, Earth Syst. Dynam., 4, 455-465, doi:10.5194/esd-4-455-2013, 2013.

Kleidon, A., Zehe, E., Ehret, U., and Scherer, U.: Thermodynamics, maximum power, and the dynamics of preferential river flow structures at the continental scale, Hydrology And Earth System Sciences, 17, 225-251, 10.5194/hess-17-225-2013, 2013.

Kondepudi, D., and Prigogine, I.: Modern thermodynamics: From heat engines to dissipative structures, John Wiley Chichester, U. K., 1998.

Mellor, G. L., and Yamada, T.: Hierarchy of turbulence closure models for planetary boundary-layers, Journal of the Atmospheric Sciences, 31, 1791-1806, 10.1175/1520 0469(1974)031<1791:ahotcm>2.0.co;2, 1974.

Pfister, L., Humbert, J., Iffly, J.F., Hoffmann, L.: Use of regionalized stormflow coefficients in view of hydro-climatological hazard mapping. HSJ 47: 479-491, 2002.

Porada, P., Kleidon, A., and Schymanski, S. J.: Entropy production of soil hydrological processes and its maximisation, Earth Syst. Dynam., 2, 179-190, 10.5194/esd-2-179-2011, 2011.

Reggiani, P., and Rientjes, T. H. M.: Flux parameterization in the representative elementary watershed approach: Application to a natural basin, Water Resources Research, 41, W04013, 2005.

Sayama, T., McDonnell, J. J., Dhakal, A., Sullivan, K.: How much water can a watershed store ? Hydrological Processes, 25, 3899-3908, 2011.

Vogel, H. J., and Ippisch, O.: Estimation of a critical spatial discretization limit for solving richards' equation at large scales, Vadose Zone Journal, 7, 112-114, 10.2136/vzj2006.0182, 2008.

Westhoff, M. C., and Zehe, E.: Maximum entropy production: Can it be used to constrain conceptual hydrological models?, Hydrology And Earth System Sciences, 17, 3141-3157, 10.5194/hess-17-3141-2013, 2013.

Westhoff, M. C., and Zehe, E.: Maximum entropy production: Can it be used to constrain conceptual hydrological models?, Hydrology And Earth System Sciences, 17, 3141-3157, 10.5194/hess-17-3141-2013, 2013.

Wienhofer, J., and Zehe, E.: Predicting subsurface stormflow response of a forested hillslope - the role of connected flow paths, Hydrology And Earth System Sciences, 18, 121-138, 10.5194/hess-18-121-2014, 2014.

Wienhöfer, J., Germer, K., Lindenmaier, F., Färber, A., and Zehe, E.: Applied tracers for the observation of subsurface stormflow on the hillslope scale, Hydrology and Earth System Sciences, 13, 2009.

Zehe, E., and Bloeschl, G.: Predictability of hydrologic response at the plot and catchment scales – the role of initial conditions, Water Resources Research, 40, W10202, doi:10.1029/2003WR002869, 2004.

Zehe, E., and Sivapalan, M.: Threshold behavior in hydrological systems as (human) geo-ecosystems: Manifestations, controls and implications Hydrology and Earth System Sciences, 13, 1273 - 1297, 2009.

Zehe, E., Ehret, U., Blume, T., Kleidon, A., Scherer, U., and Westhoff, M.: A thermodynamic approach to link self-organization, preferential flow and rainfall-runoff behaviour, Hydrology And Earth System Sciences, 17, 4297-4322, 10.5194/hess-17-4297-2013, 2013.

Zehe, E., Lee, H., and Sivapalan, M.: Dynamical process upscaling for deriving catchment scale state variables and constitutive relations for meso-scale process models, Hydrology And Earth System Sciences, 10, 981-996, 2006.

Brief reply to reviewer II

We sincerely thank reviewer II for his critical but very helpful comments on our manuscript. We addressed most of them as partly already explained in our brief reply to Keith Beven and partly summarized in the following:

- We removed the "proposal like" language and avoided referencing to the CAOS project in the revised manuscript (except of section 5), as recommended by the reviewer.
- Section 5 presents now a very much streamlined outlook on our ongoing research that aims to test our 4 main propositions. Although reviewer 2 recommended skipping this outlook completely, we think this is valuable as we share our a–priory ideas and concepts how to advance experimental characterization and modelling of intermediate scale catchments of organized complexity before publishing our related findings in research papers (which implies the risk of being proven to be wrong). We take this risk to vote for a publication culture that allows learning from failures.
- We did not add preliminary results to our outlook to the ongoing test of concept. This is simply because to stay within the scope of an opinion paper and avoid some kind of mixture of opinion and scientific article – which might be very difficult to review.

Again let me thank you very much for the valuable comments,

Erwin Zehe

Detailed response to review II, taken from the discussion section

We, (Erwin Zehe as EZ on behalf of all co-authors in the following) sincerely thank the anonymous reviewer for his helpful assessment of our opinion paper (CAOS paper in the following).

Reviewer: Is this meant an opinion paper?

EZ: We think this is indeed an opinion paper as it reflects our opinion at two levels:

- We need a publication culture that allows sharing of our scientific failures, because there is much to learn from it. I think an opinion paper is the right tool for this.

- It reflects our opinion on how to advance our predictive understanding of how spatial organization controls intermediate catchment functioning (this is not a novel problem but not solved (we explain this after explaining the first point).

What was our idea on how to address the first point? The hypotheses, ideas and approaches presented in the CAOS paper have been accepted by an international jury of experts. Before we will publish our findings from two years of research work in the forthcoming research papers, we wanted to share the true initial ideas and a priory hypothesis beforehand. The underlying idea this is to share not only the success stories but also what we learned from our failure with the community, because some the findings are line and some are pretty surprising. We did this certainly not to advertise our project and will reduce our referencing to it to the necessary absolute minimum.

We admit that we have maybe overdone it a little with this idea, especially when it comes to section 4. Again this has a good reason. EZ was editor of one of the very first opinion papers by Savenije (2010) introducing his idea of the flex- topo modelling in HESS. A major critique in the first reviewer round was that the author should point out ways to show that his proposed model concept is indeed structurally more adequate than others. Section 4 is part of this paper, because we did not want to stop at the stage of just arguing that the presented EFU concept is (partly) experimentally testable but share our ideas how we are going to do this. This section has partly the character of a proposal, simply because it is a proposal how to tackle this problem. Adding more results to this section is not possible, because it is the privilege of the PhD students and Post Doc to publish their main results with their names at first place. Hence, we will reduce the length of this section to the minimum necessary amount.

Still we would like to stress that the proposed experimental design is at least pretty rare, if not unique with its effort to conduct replicate experiments and monitoring at members of EFU and hillslope that are expected to function similarly (of course there is B2-LEO). We are well aware of Terreno as the two authors Theresa Blume and Peter Dietrich coordinate experimental activities in the Terreno Müritz/Ücker and Terreno Bode observatories.

Reviewer: Past work on similar issues should be acknowledged
EZ: With respect to the second point - the old problem of organized complexity – we present several novel ideas that could due to our opinion bring new momentum to our understanding. These are motivated by the HRU idea and earlier work that proposed simplified but physically consistent model approaches for larger control volumes (REW approach, Bousinesq model) and our working believe that the catchment is an organized fingerprint of past processes. In fact we went back to the classics with our referencing with respect most contributions (maybe not all) which motivate or concept (organized complexity, HRU definition, the catena concept, the pattern process paradigm, predictive uncertainty and organizing principles). We will add references to the REW and other promising model concepts (hillslope storage Bousinesq model) in the revised manuscript and discuss their pros and cons with respect to the balance of model complexity and simplicity.

Our main points are quickly summarized in the following:
1) A thermodynamic perspective on hydrology is useful (while not novel per se) as it implies/reminds us that a flux is equal to a "potential gradient" $\nabla\phi$ (temperature gradient, water level gradient, concentration gradient, soil water potential gradients) divided by a resistance R (inverses of either heat conductance, surface roughness, diffusion coefficient, hydraulic conductivity…). The former determines the (thermodynamic) force the latter determines dissipative energy losses along the flow path:

$$\vec{q} = {}^{1}/_{R}\,\overrightarrow{\nabla\phi} \text{ (Eq. 1)}$$

In larger control volumes of terrestrial systems both gradients and resistance are fields and depend on almost static controls and on system state variables (all this is well known).

In this framework HRUs or functional units can be defined as classes of landscape entities/control volumes with similar terrestrial controls on the pair of $\nabla\phi$ and R. on the pair of gradients $\nabla\phi$ and resistance R controlling either land surface energy exchange or rainfall runoff. (Note this this is necessary conditions for functional similarity, but not necessary and sufficient conditions as Eq. 1 is not an unique equation). According to Flügel (1996), 'Hydrological Response Units are distributed, heterogeneously structured entities having a common climate, land use and underlying pedo-topo-geological associations controlling their hydrological transport dynamics'. We think the definitions match well, our one has maybe a little bit more physical rigor <u>and</u> that it implies the proposed hierarchy of functional units (EFU and lead topologies) instead of a one fits all HRU and that their dominance changes dynamically with prevailing boundary conditions (this is thoroughly explained in our response to Keith Beven).

2). Eq. 1 is immanently subject to equifinality because it is not an injective function (several elements of the start domain are mapped on the same element in the codomain):

a) Several combinations of gradients and resistance compile the same flux. This might be frequently the case in hydrological systems as the (quasi static) controls especially on gradients driving lateral flows of blue /free water are independent from the properties controlling flow resistances (especially when it comes to vertical and lateral preferential flow paths). In line with the argumentation of Bardossy (2007) the first source of equifinality can be reduced based on observations that characterize at least two out of the three variables (either q and R or q and $\nabla\phi$, or $\nabla\phi$ and R). Current observation technologies allow approximate characterization of the static controls on gradients driving lateral flows of blue water (surface and bedrock topography). This is why we think a coupled treatment of the mass and momentum balance has more pros than cons and stands as the main supportive argument for hypothesis H2.

b) The R term is no unique as resistances in Eq. 1 reflect the spatial heterogeneous and spatial organized arrangement of for instance soil material in the control volume.

Subsurface flow resistance depends for instance on soil hydraulic conductivity $1/k(\theta)$, its covariance lengths and soil moisture. Connected structures (preferential pathways lateral pipes, vertical macropores) reduce the control volume resistance at a given driving gradient as they allow for advective flows, resulting in accelerated fluxes and shorter residence time distributions. Preferential flow networks with different topologies and hydraulic properties may result in the same control volume resistance (e.g. Klaus and Zehe 2010). This source of equifinality cannot be eliminated and has to be accounted for. To our opinion there are 2 promising avenues to tackle the resistance problem. As some preferential pathways are created by ecosystem engineers, there is a chance to get approximate information on their spatial pattern using species distribution models as explained in the CAOS paper. Organizing principles allow for a priory optimization of the resistance term at a given gradient, either as a bulk resistance (Porada et al. 2011; Westhoff and Zehe 2013) or the density of vertical and lateral macropores (Zehe et al. 2013; Kleidon et al. 2013). This implies the possibility of independent predictions. These can of course go wrong, but this is testable! And this is the main reason for hypothesis H3.

3) We think that hydrology lacks realistic and falsifiable models at intermediate scales and we share or opinion why this is the case and how to fill this gap by formulating clear criteria. The first is an explicit at least one dimensional accounting for the momentum balance, the second is spatially explicit accounting for preferential flow paths (especially the lateral ones which cause advective exchange between EFU during rainfall driven conditions). In the context we do not think that the problem is in the equations themselves. We have either diffusion like equations for diffusion like problems, advection equations for advection problems and advection dispersion equations. There is a problem when using the inappropriate equation for a certain type of flow problems. In this sense the Darcy equation is inappropriate when flows get preferential (advective) or the use of the advection dispersion equation is in appropriate at low Peclet numbers. So where what is the chicken and what is the egg?

Advection implies high velocities even at small gradients which essentially requires low dissipative losses. We thus regard the presence of connected preferential pathways as the chicken, which has to be well represented in the model, and fast flow as the egg. In line with this we regard the topology of the flow network as the key for a structurally adequate model structure. Otherwise hydrologic routing schemes couldn't work at all, because they preserve the network topology but violate the "flow law". Similarly preferential flow and transport in soil can be well predicted based on assuming Darcy flow, which is not correct, when the topology of the flow network is well represented. What is thus falsifiable?

- The spatial structure of the model, (it representation of the covariance structure of textural properties and the topology of preferential pathways),
- How processes interact among different domains (bidirectional, unidirectional)…

Our benchmark for this is spatial transferability of the structural and functional model sets among class members of the same functional unit.

We will revise the manuscript to better explain these main points. We again thank the reviewer for his effort and the helpful comments,

Erwin Zehe

References

Bardossy, A.: Calibration of hydrological model parameters for ungauged catchments, Hydrology And Earth System Sciences, 11, 703-710, 2007.

Flügel, W-A.: Hydrological Response Units (HRUs) as modeling entities for hydrological river basin simulation and their methodological potential for modeling complex environmental process systems. Die Erde 127,42–62, 1996.

Klaus, J., and Zehe, E.: Modelling rapid flow response of a tile drained field site using a 2d-physically based model: Assessment of "equifinal" model setups, Hydrological Processes, 24, 1595 – 1609, DOI: 10.1002/hyp.7687., 2010.

Kleidon, A. and Renner, M.: A simple explanation for the sensitivity of the hydrologic cycle to surface temperature and solar radiation and its implications for global climate change, Earth Syst. Dynam., 4, 455-465, doi:10.5194/esd-4-455-2013, 2013.

Kleidon, A., Zehe, E., Ehret, U., and Scherer, U.: Thermodynamics, maximum power, and the dynamics of preferential river flow structures at the continental scale, Hydrology And Earth System

Savenije, H. H. G.: HESS Opinions "Topography driven conceptual modeling (FLEX-topo)", Hydrol. Earth Syst. Sci., 14 (12), 2681–2692, doi:10.5194/hess-14-2681- 2010, http://www.hydrol-earth-syst-sci.net/14/2681/2010/, 2010.

Westhoff, M. C., and Zehe, E.: Maximum entropy production: Can it be used to constrain conceptual hydrological models?, Hydrology And Earth System Sciences, 17, 3141-3157, 10.5194/hess-17-3141-2013, 2013.

Zehe, E., Ehret, U., Blume, T., Kleidon, A., Scherer, U., and Westhoff, M.: A thermodynamic approach to link self-organization, preferential flow and rainfall-runoff behaviour, Hydrology And Earth System Sciences, 17, 4297-4322, 10.5194/hess-17-4297-2013, 2013.

Reply to reviewer III

We sincerely thank reviewer III for his helpful comments.

We agree that the title of the first version left many open questions. We thus selected a new title (avoiding the term new) which reflects much better our scientific focus – which is to contribute to better experimental characterization and modelling of intermediate scale catchments of organized complexity.

- Section 1 highlights that our intension was not to criticize the HRU concept, but to provide a re-interpret from a thermodynamic perspective. This combination offers, as now explained in section 1, several useful and partly also novel perspectives:
    - For defining functional similarity based on similar terrestrial controls on the pair of gradients and resistance terms controlling different land surface atmosphere energy exchange, rainfall runoff transformation and base flow production. We propose that this implies a scale hierarchy of "specialized" HRUs we call functional units that function similarly with respect to one specific form of "water release", because the terrestrial controls on related gradients and resistances have different characteristic length scales (as further explained in section 2)…
    - On alternative strategies to characterize intermediated scale catchments by clustering observations in replicate members of functional units, focusing on the triple of flux, gradient, resistance (as explained in section 3).
    - Requirements for structurally adequate models, which should rely on thermodynamic consistent equations (flux= gradient 1/resistance) and disentangle matrix flow from vertical and lateral preferential flow. This offers options to reduce inherent equifinality, arising from the interaction of gradient and resistance terms in the governing equations, based on available information on controls on the pair of gradient and resistances (as explained in section 4.1 and 4.2)
    - An energy centered view on hydrological dynamics (including rainfall runoff transformation), including the functional advantage of organized structures from and thermodynamic optimality perspective. (as explained in section 4.3)

S

In section 4.1 and 4.2 we suggest that structurally adequate models for intermediate catchments

- should rely on thermodynamic consistent equations (any flux is equal to gradient times 1/resistance), because this offers options to reduce inherent equifinality as stated above;
- disentangle matrix flow from vertical and lateral preferential flow, because they are independent sources of equifinality (compare section 4.2) and reflect different forms of spatial organization (compare section as explained in section 2.1)

As recommended by the reviewer to papers and models we regard as pioneering/ as benchmarks on the search for structurally adequate models for intermediate scale catchments (which meet these requirements balance complexity with parsimony).

- In section 4.2.1 we give credit to the hillslope storage Boussinesq model and most importantly of the REW approach. But we point also out that zero dimensional treatment of process domains (as done in the REW approach), h is not appropriate for representing spatial organization at the hillslope and smaller scales (compare section 4.2.1).
- Section 4.2.1 also points out that thermodynamic consistency of the model equations, in the sense that they disentangle gradients and resistance terms, does not imply that a model needs to be based on partial differential equations. In this context we acknowledge WASA, mHm and Topmodel and Dynamic Topmodel.
- In this context we do not refer to models, which do not spatially resolve terrestrial controls on gradients driving lateral and vertical flows, because they do not allow constraining inherent equifinality in the proposed form.

The reviewers comment on treatment of lateral flows:

- We agree that our formulation was misleading here. The idea of the CAOS model is to subsequent treat/ add processes at those scale levels where they “kick in”
- We thus neglect lateral flows within EFU objects but account for all vertical processes sustaining the energy balance and infiltration.
- As EFU are arranged along topographic gradient, there is no way of neglecting lateral flow at the next higher level (which of course can happen in different forms)
- We treat lateral flow thus in separated objects (a fast flow object and a saturated flow object) which together with the EFUs form a Lead topology (called lateral topological units in the second manuscript to please Keith Beven).

- Lateral topological unit are basicly hillslope with similar a driving topographic gradient and a similar topologically connected lateral flow path (surface rills, or pipes or the bedrock micro-topography) facilitating lateral exchange (as explained in section 3.2)

We fixed all technical issues as recommended.

Thanks you very much again for your valuable comments

Erwin Zehe