In the following, comments by Referee E. Zehe are indicated with [Z] and replies by the author are indicated by [K].

[Z] *This study examines free energy conversions associated with subsurface water flows within numerical experiments at a synthetic hillslope. More specifically the author perturbed saturated hydraulic conductivity for a given set of soil hydraulic parameters in search for an optimum value maximizing entropy production/power in subsurface flows and elaborates on the role of the groundwater surface and on the role of purely random heterogeneity in this context. Furthermore, the author proposes an interesting approach how to use MEP for inferring/upscaling functional optimum, effective subsurface water characteristics within larger control volumes. Particularly, the last point is interesting and novel. The proposed study is, hence, without of high potential interest for HESS.*

Nevertheless, I got the imprecision that the study is compiled a little too hastily, which is partly reflected in superficial referencing (as explained below) and the too narrow range of properties that is investigated within the experiments. The study would greatly from a thorough revision a) to improve physical rigor in the proposed approach how to estimate "power" within the model domain and to broaden the type and range of terrestrial controls on subsurface water flows and related free energy dissipation within the numerical experiments.

[K] I would like to thank the reviewer for the constructive comments, questions and suggestions, which help to improve the manuscript. The points raised above also with respect to physical rigor and broadening the range of terrestrial controls will be addresses below and in detail in the revised manuscript.

[Z] *I encourage the author, for whom I have high respect, to further optimize this potentially very interesting study by addressing the following points:*
Major point concerning referencing:

- *I appreciate the authors effort to refer to recent studies dealing with thermodynamic optimality in hydrology (some of those I authored/co-authored myself). The way how these studies are discussed is, however, not appropriate. After reading the introduction, one is left with the imprecision that this is the first study using "physics-based" models instead of bucket models shed light on the role of MEP in hydrology. This is not quite the case, for instance: o Porada et al. (2011) used a 1d-implementation of the Richards equation combined with a SVAT-approach and showed for the 35 largest catchments in the world, that the MEP optimized configuration plotted in a short envelope around the Budyko curve.*
    - *Zehe et al. (2013) used a 2-d Richards-model (coupled with a SVAT and 1 d overland flow) within numerical experiments exploring free energy conversions and dissipation associated with subsurface water flows and different runoff components. More specifically we perturbed the flow resistance in hillslope scale models and found the those values which optimize steady state production of entropy (or equivalently the reduction of free energy), perform acceptable in predicting rainfall runoff behavior two different catchment.*
    - *Kleidon and Renner (2013) proposed a simple model (based on three parameters) to predict land surface energy exchange based on the idea of*

maximizing power in the virtual and sensible heat flux and that the sensible heat flux operates at the Carnot limit. The model performs strikingly well when being compared against flux tower data in different land use settings during convective conditions.

I very much agree that "physics based" models are superior for exploring to which extent thermodynamic optimality indeed applies; however, it is a matter of fairness to acknowledge that this has been addressed by the reported studies. Maybe the formulation PDE based or hyper resolution models is more appropriate as physics-based, as it does not imply the other models are not based on physics.

[K] I do not want to leave the impression of unfairness. I am very much aware of the aforementioned studies, which are of deep interest to me and contribute truly novel ideas to hydrology. By physics based and as a difference to previous studies I meant that Richards equation is applied at the support scale, and production of entropy is calculate at this scale as it is suggested by the theory in my opinion. To illustrate this, following Kondepudi and Prigogine 1998, the entropy balance equation is

$$\frac{\partial s}{\partial t} + \nabla \cdot J_s = \sigma \qquad \text{(i)}$$

where the first term on the left hand side is the temporal entropy change; the second term on the left hand side is the divergence of the entropy current (entropy that enters/leaves the macroscopic domain); and σ is internal entropy production due to sum of all internal, "microscopic" forces, F_i, and fluxes, J_i, defined as

$$\sigma = \sum_i F_i J_i \qquad \text{(ii)}$$

In the simulations, the entropy production was calculated directly at the support scale by summation of all local forces and fluxes. The term power was used because temperature does not change. However, this is inappropriate, because generation of power is simply neglected; it is the microscopic dissipation of the gradients by the fluxes in (ii), which produces the entropy as pointed out by the Referee. Additional thoughts on equation (i) and (ii) will be provided below, which will lead to a reanalysis of the simulation results and clarifications in the revised manuscript.

[Z] *Major scientific points:*
The author propose essentially to calculate power based on in- and outward soil water fluxes into the model grid elements and local gradients, to summarize these values over time and then compile the spatial integral. With this he adopts a macroscale formulation of power (usually employed at the system scale using fluxes across the boundaries and macroscale gradients) which applies for steady states in gradients and fluxes to the microscale. This is due to several reasons inappropriate because the grid scale gradients get depleted by the fluxes within a simulation time step means there is no steady state.

[K] In the simulations, conditions of dynamic equilibrium were produced i.e. the simulation year was run repeatedly with identical hourly atmospheric forcing until the water balance and entropy (ds/dt=0) over the simulation year did not change. This required more than 30 years (in some of the heterogeneous cases 60 years) of simulation! The only assumption in equation (i) is that the system is not far from equilibrium, which is the case in the simulations. Thus, applying equation (i) to instantaneous fluxes and gradients while including all required components in the divergence of the entropy current is appropriate in my opinion and results in ds/dt = 0 over one full cycle at dynamic equilibrium i.e. the statistical steady state. This is discussed again below.

[Z] *Secondly, soil water fluxes are essentially dissipative by their nature (pushing the system back to local thermodynamic equilibrium) and kinetic energy associated with soil water fluxes is marginal (this where power could be extracted from to perform work). The calculated term is thus not power (a source term in the free energy balance) but dissipation (the sink term in the free energy balance). As eq. 10 deals in fact dealing with dissipation, the proposed summation in is hence not appropriate, as positive and negative terms cancel out. (The latter wouldn't harm when dealing with a system conserving free energy, which does not perform work when traveling along a closed path). One can think about a case where all summands in Eq. 10 cancel, the equations suggests that nothing happened at all; although gradients have been depleted and re-established many, many times and which implies dissipation of energy. Maybe, the way to go is to multiply the net flux into a local grid element with temporal change in the local potential (see post discussion of Zehe et al. 2013 for the details) or to use the method proposed in Zehe et al. (2013).*

[K] The Referee is correct that the theory and analysis needs to be clarified. It was intended to integrate the equation (i) over the dynamic equilibrium of one cycle/year by calculating the total entropy production, which equals the divergence of the entropy current, because ds/dt = 0 over one full cycle i.e. one complete simulation year. Thus, it is correct that there is entropy production (dissipation) even if the net flux is zero. This and additional implications can ben illustrated in a simple thought experiment: removing the topographic slope from the synthetic hillslope experiment results in a bucket which is closed laterally and at the bottom, and is in contact with the atmosphere at the top. Performing the same dynamic equilibrium simulations as before will result in a net or mean flux of zero over one full cycle (because the bucket can not dry out infinitely). Because of the fluctuations in the flux there will be dissipation and entropy production. Thus, ds/dt = 0 over one full cycle; and the internal entropy production integrated over one full cycle equals the divergence of the entropy current integrated over one full cycle. However, the divergence of the entropy current now consists of the exchange with the outside via the top boundary (evaporation/infiltration) and the internal, local fluctuations close to equilibrium, which may be understood as local (entropy) sinks/sources. Thus, in the analysis these fluctuations need to be included. To check this I integrated

$$\nabla \cdot J_s = \frac{\mu}{T}\left(\nabla \cdot q\right) \qquad \text{(iii)}$$

and

$$\begin{aligned} \sigma &= \sum_i F_i q_i \\ F &= -\nabla \frac{\mu}{T} \end{aligned} \qquad \text{(iv)}$$

over one full cycle over the hillslope for one example of K_{sat} = 0.01 m/h (where μ is the chemical potential, in this case the hydraulic head; and T is temperature, in this case constant), which yields an absolute difference of -5.53x10^{-10} m^2/a, which is on the order of 10^{-15}% of the total entropy production. This tells me that the system is entropy conservative and equation (i) is applicable at dynamic equilibrium.

[Z] *The authors needs to show that his analysis reflects steady state behavior by compiling longer simulations using a longer time series reflecting the more than one year of forcing conditions.*

[K] Dynamic equilibrium (statistical steady state) over one full cycle/year is the prerequisite of the proposed method and was ensured with multiple decades of spinup simulation.

[Z] *I d' like to encourage the author to better link the numerical experiment to natural systems and to explore a wider range in natural controls. The proposed simulation domain has, expect of a very small and homogeneous topographic gradient, not much in common with a natural hillslope. Also the soil should be better characterized than just dropping the van Genuchten-Mualem parameters (similarly the boundary conditions). Even when the author prefers not to deal with transpiration, it is straight forward to compare for instance fine and coarse porous soils (with different importance of capillary pressure in this concert), different topographic gradients (to come closer to a hillslope) and eventually forms.*

[K] The constructive suggestions by the Referee are very well received. A direct link to natural systems is the ultimately goal including the demonstration of the predictive capability of the methodology. Also major aspects raised by the Referee such geometry of the hillslope, soil characterization, boundary conditions, and additional processes such as transpiration need to be addressed. However, this is beyond the scope of a single study, because of the required compute time of the hyper resolution simulations for dynamic equilibrium and extremely large set of potential simulations. The scope of the study is to propose a new methodology supported by, at this point, synthetic hillslope experiments.

[Z] *The reported dependence that purely random heterogeneity in ksat yields to an elevated maximum in dissipation is by far too interesting, to treat it in such a brief manner. How does this depend on the variance of ksat (which I missed by the way in the manuscript)? How does this change when adding a spatial co-variance? Is it a steady decrease within increasing correlation length, or is there a maximum dependent on the correlation length? This is particularly interesting as correlation lengths in natural soils are short (several meters).*

[K] This appears to be indeed the case, but must be corroborated in the revised analysis. Following the discussion above, entropy production may increase with random heterogeneity, because of an increase in the fluctuations of the flux and perhaps also because of an increase in the mean flux. Again, there is an extremely large set of simulations required in order to interrogate the impact of e.g. spatial structure at different scales. In the revised manuscript, additional results will be provided for different variances of the random distribution.

[Z] *The last part dealing with MEP as an inference principle and the suggested approach to estimate a macroscopic gradient and conductance is for me the most interesting part of the study. However, the provide evidence does support the conclusion that MEP applies indeed to hydrological system in the sense that an optimized effective conductance performs well against observations. This implies a) to show that optimum hillslope structure have predictive power against data and that the up-scaled effective conductance works well when being used within a water balance simulations carried out at larger grid cells (as for instance shown by Lee et al. 2007). Westhoff and Zehe (2013) showed that MEP is not useful for conceptual modelling, as the optimum parameters had not predictive power for the water balance; Zehe et al. (2013) reports to 2 successful cases, which can be just a coincidence). So there is still room for more evidence here.*

[K] The Referee is correct that the demonstration of the predictive power against data and in water balance simulations at larger grid cells is lacking. At this point, it is not clear how to provide evidence, because the relationships provided so far are for a limited set of synthetic cases. More simulations and application to a natural hillslope are planned in future but are out of the scope of this study in my opinion.

[Z] *Please not that Eq. 1 is only valid, in case of mass flows, which are driven by chemical potentials and during steady states.*

[K] Please see my replies with regard to equation (i).

[Z] *Power is as a flux an intensive property and can thus not be imported/exported or balanced. Please refer to in/export of free energy which is an additive quantity.*

[K] This will be clarified in the revised manuscript following also the discussion above. The entropy balance including fluctuations will be revised.

[Z] *Hourly differences of P-ET are not equal to infiltration but equal to infiltration and surface runoff (Eq. 4). Why not using the influx into the model domain?*

[K] Correct, surface runoff will be added to the equation.

[Z] *How does the model deal with the saturated zone- in an iterative manner allowing for a free surface or by using storage coefficients?*

[K] Richards equation is applied in a continuum approach including the saturated and variably saturated zone allowing for a free surface (water table).

[Z] *Is it not really astonishing, that a saturated domain builds up as you use a no flow boundary?*

[K] The domain could be infinitely deep and there still would be a water table, which is purely the result of the interaction of non-linear, variably saturated flow and evaporation/infiltration at the top based on the atmospheric forcing time series.

[Z] *Page 5131 typo: toward by towards*

[K] Will be corrected in the revised manuscript.

[Z] *The gradient should point into the opposite direction of the flux (which reflects the second law of thermodynamics) not into direction of the flux*

[K] This will be corrected.

[Z] *Maybe replace small scale chaos with small scale disorder*

[K] This suggestion will be honored.