First of all we want to thank Jianting Zhu for the important and constructive points addressed in his review. The following lines will explain how we addressed his comments in the revised manuscript or, if we didn’t, why we didn’t.

Jianting: This study addresses important and challenging scaling issues related to hydrology. Specifically, it deals with upscaling of both hydrologic state variables and effective hydraulic properties at catchment/REW scales. The authors derived the time series of catchment-scale average soil saturation and hillslope effective hydraulic functions. They also found that the dominant patterns of soil heterogeneity and macroporosity (mainly macroporosity in my opinion after I read that paper) are enough to
represent the main aspects of the catchment scale hydrologic processes. Overall, it is a well written and organized paper, although there seem to be some repetitions of verbiage from time to time that can be reduced in the revised version. However, I have some concerns that I will elaborate below. Response: We will remove the repetitions from the revised manuscript.

Major Technical Comments I have four major comments about this paper.

Jianting: 1) The linkage between the two major objectives of this study: I found the linkage is not clear and needs to be better established. When generating times series of catchment-scale average soil saturations in the unsaturated zone by averaging the corresponding distributed model output, does the unsaturated zone have the same processes as the modeling that is used to generate the hillslope scale soil hydraulic functions. As I image the first part should be done in a more complete processes that also include the other hydrologic processes, such as the saturated zone and the concentrated overland flow zone etc. It is not very clear to me that if this modeling also uses the same set up as the one that is used to obtain the upscaled hydraulic functions. In other words, are the two main exercises in this study really tied? Is the process used in determining the effective hydraulic functions an integral part of the one used to generate times series of catchment-scale average soil saturations? If a different process is used to obtain the effective hydraulic functions, will the effective hydraulic function results be the same?

Response: There is in fact a close link between the two objectives, which will be brought out more clear in the revised manuscript. To generate the time serious of catchment scale saturation, we employ the complete model structure that has been build up and validated in the Weiherbach catchment within a simulation of app. 1.5 years: 169 hillslopes interconnected with a drainage network, the model is driven by initial and boundary conditions that have been observed during this period simulation period and the model accounts for all the processes (ET, soil moisture dynamics, runoff generation and concentration, discharge in the drainage network). As this model struc-
ture with this boundary conditions reproduced ET, runoff response and soil moisture dynamics observed during this period in the Weiherbach, we conclude that the average soil moisture (that is averaged over all 169 hillslopes in the catchment) is the best estimator and a physical consistent estimator how the “true” average catchment soil moisture developed in the Weiherbach catchment during this period. The important points to support this argumentation is that slight deviation from this model structure produce already different runoff responses and different time series of averaged catchment scale soil moisture, when used with the observed input. Within this context it is important to stress that different hillslope scale patterns of soils and macropores do produce clearly different time series of average catchment scale soil moisture, thus the differences do not average out! This fact is and the results from the catchment scale simulations with real boundary conditions is the justification to go for the second objective i.e. to the model to produce REW scale constitutive relations based on numerical experiments. In the first trial (not shown in the paper) we did this by employing the complete catchment model and natural boundary conditions, i.e. we just took time series from the simulations for objective one, plotted average catchment saturation against average capillary pressure and fitted a Brooks and Corey type of relation (Lee et al. 2006). However, this did not give could results, as the range of average saturations and capillary pressure values was very small. Hence, we employed the artificial boundary conditions specified in the manuscript, to cover a large range of values for average saturation and capillary pressure values. In a last step, we found out that simulation with just a single hillslope produce the same relations as simulation with the complete catchment model. This is because of the fact, that we just represent typical patterns of soils and macropores in the model, which are the same for each of the 169 hillslopes. The fact that is it sufficient to use a typical hillslope structure for deriving the constitutive relations will be further underpinned by new results that will be added to the revised manuscript. These will show that, even if we account for stochastical heterogeneity of ks-value and for macropores in an explicit manner, outflow from the hillslope structure is becomes already constant when averaged over one third of the
Jianting: 2) The average saturated hydraulic conductivity: The authors state that as expected, in homogeneous soils the average saturated hydraulic conductivity determined from a sufficiently large sample of point measurements is a good estimate for the hillslope scale/REW scale saturated hydraulic conductivity. I found this is quite strong statement and needs to be substantiated. I can only assume that the homogeneous soils here have to be texturally heterogeneous soils as opposed to the case study which is dominantly structurally heterogeneous. Otherwise it does not make sense to even mention the average saturated hydraulic conductivity determined from a sufficiently large sample of point measurements since the saturated hydraulic conductivity would be a constant for homogeneous soils. If my assumption does reflect what the authors meant, I would argue that the statement by the authors needs to be elaborated. For the drainage scenarios depicted in this study, the simple average is probably not a good effective parameter estimate, unless it is really a mildly heterogeneous case.

Response: This is of course not meant as general statement, because it would be nonsense. The statement just refers to this particular case of hillslope that consists of two homogeneous soil blocks (Calcaric Regosol upslope and Coluvisol downslope) and the layer interface is parallel to the direction of flow (that happens at the lower boundary in vertical direction). In case of a saturated soil it is easy to show that the effective hydraulic conductivity of this medium is the weighted arithmetic average of the conductivities of the two homogeneous soils (Jury and Horton, 2004, Roth, 1995). In other cases this is of course not true! We will clarify this in the revised manuscript.

Jianting: 3) The inconsistent use of hydraulic functions: The soil hydraulic properties after van Genuchten and Mualem were measured in the laboratory using undisturbed soil samples. But for REW-scale soil effective parameters, the parametric relationships of power-law type as shown in Eq. (5), similar to the Brooks-Corey model [Brooks and Corey, 1964], were assumed. I found that it will make more sense to use a consistent model. It is even more desirable to relate input hydraulic parameter structure to the total hillslope length!
Response: We agree that using the van Genuchten-Mualem approach to parametrized the REW scale constitutive relations would be more consistent, as the REV scale model CATLFOw uses this approach. In this case we could compare REV and REW scale parameters. However, the clear advantage of the Brooks and Corey approach is, (the bubbling pressure and the exponent) compared to the 5 parameters of the van Genuchten approach, which would make the parameter estimation much more involved (including multiple optima). As we want to minimize the number of model parameters, and as the REW scale curves appear to be much simpler than the REV scale curves, we still prefer the Brooks and Corey approach. However, in the revised manuscript we will compare the bubbling pressure to the inverse values of alpha from the van Genuchten model.

Jianting: 4) The dominant heterogeneities that dictate the hydrologic processes. After reading this paper, the take home message seems to be that the macroporosity heterogeneities and patterns dominate both the hydrological processes and the effective hydraulic properties for the catchment (REW). In other words, the critical subscale soil heterogeneities that actually impact hydrologic processes at next higher scale level are related to the macropores. At the end of this paper, the authors postulate that a set of typical closure relations exists for each landscape. I feel this hypothesis might be a natural extension (step) of what being studied in this paper, i.e., structural heterogeneities. For textural type of heterogeneities, the degree of heterogeneities and spatial correlations rather landscape might be more dominant in determining parameterisations of hydrological processes at the next higher scale. My question is: will a set of typical closure relations always exist for various heterogeneous scenarios, such as a mixture of both structural and statistical (textural) heterogeneities?

Response: This is indeed a very important issue. We also had the feeling that the revises manuscript has to compare the effect of textural and structure heterogeneity. In the new manuscript we additionally investigate the effect of heterogeneity of soil hy-
Hydraulic conductivity and macroporosity. To this end we compared 4 cases - The normal hillslope catena with homogeneous soils - The normal hillslope catena + a heterogeneous pattern of ks generated with turning bands (total variance was estimated based in field data to 1 order of magnitude) - The normal hillslope catena + a heterogeneous pattern of ks + a typical pattern of discrete macropores (generated with a Poission process, data taken from field investigations). As observed in the Weiherbach the macropores penetrate in average 0.8 m into the depth and end in the soil matrix. - The normal hillslope catena + a heterogeneous pattern of ks + a deep pattern of discrete macropores (generated with a Poission process, data taken from field investigations). In this case the macropores penetrate the lower boundary of the modelling domain, hence drainage from those macropores is not controlled by the soil matrix at their end.

As will be shown in the revised manuscript, the first three model structures yield the same spatially average outflow after a total averaging length (REL) of 50 m, which is app. 0.25 of a typical Weiherbach slope. Only the forth case gives, as expected different spatially averaged outflow values. This shows: there is an REA/REL for the Weiherbach, were drainage into the s-zone can be treated homogeneous. This shows furthermore, that the average value embeds information on the spatial connectivity of macropores /structures with them at the next higher scale, as the 4th case yields a clearly higher average outflow. However, case 4 is not the relevant for the Weiherbach catchment as earthworm burrows there never penetrate down to the saturation zone (which is located between 3 and 30 m. However, to address the always in the question, I would not go that far that we always find representative closure relations. One possible case would textural heterogeneities with correlation lengths much larger than the hillslope (I don’t know whether something like this exists).

Specific Minor Comments Jianting: Page 1630, line 26. The word “hillslope” appeared suddenly. Does it imply that what the authors summarized earlier is not related to “hillslope”?

Response: As stated in the response to major point 1, the time series are derived at the
catchment scale. For the constitutive relations it was sufficient to work on the hillslope scale.

Jianting: Page 1633, line 22. “is” should be “are”. Page 1635, line 19. The authors stated that the approach is similar to the perturbation methods. In what sense is the approach similar to the methods? Does the approach have the same limitation of being applicable to mildly heterogeneous media? If this is the case, I think the authors should be more specific about it. Response: No, the approach has not the same restriction, as will be shown in the new manuscript.

Jianting: Page 1638, section 2.3. We focus on the exchange term $e_{us}$, which denotes groundwater recharge or capillary rise. My question is: why was only this exchange term used in deriving the catchment-scale hydraulic functions? While this term is certainly influenced by hydraulic properties mostly, all other terms will also carry fingerprints of hydraulic properties in the catchment, in my opinion. This comment is somehow related to my major comment 1).

Response: We focus on the exchange between the unsaturated and saturated zone, this is the only exchange term in the CREW model, that contains REW scale hydraulic functions. (see Lee, E. Zehe, M. Sivapalan Page(s) 1667-1743. SRef-ID: 1812-2116/hessd/2006-3-1667). E.G. the closure relation for infiltration makes only use of the saturated hydraulic conductivity.

Jianting: Page 1638, line 18. “?” needs to be deleted. Response: Will be deleted in the revised manuscript.

Jianting: Page 1650, line 16 - 18. It should be noted here that the time series of the upscaled catchment scale soil moisture are not just simply the arithmetic averages of the observations. Then what is it? I also have some comments on the bottom panel of Figure 4. From the way it is presented right now, it does not seem to tell us anything. While the authors declare that the average soil moisture simulated by the landscape and process compatible model structure falls in the range of the observed soil values,
I feel it might be more appropriate to show comparison with the average observations, such as some type of area weighted average if the observation points are not evenly distributed spatially. Response: This is a very good idea and will be done in the revised manuscript.


Jianting: Page 1653, line 26. Delete “not”. Response: Will be deleted in the revised manuscript.


Erwin Zehe


Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 3, 1629, 2006.