Reply to Anonymous Referee #2:

Ralf Loritz (RL): We would like to thank the anonymous reviewer #2 for his comments and the time he invested in assessing our manuscript. She/he has addressed many issues which will greatly improve the readability of our manuscript.

Reviewer: Summary and Recommendation:

Reviewer: *The manuscript presents 2D distributed physically-based modeling of water flow dynamics at two representative hillslopes differing in geology, soil, and vegetation characteristics. Authors proposed several scenarios with different conceptualizations of hillslope to analyze first-order controls on soil water dynamics and limitations in current modeling approaches. The manuscript conveys an interesting topic, potentially attracting readers of wide hydrological community. However, several points deserve further attention and need to be clarified/improved.*

General comments

Reviewer: *1. Rather than a rigorous study on hillslope modeling with detailed data-model comparison, authors set up representative hillslopes built on two perceptual models. As a result, no comparison of spatially dependent variables (although measured) was presented. Authors avoided intentionally the first step in modeling. The idea of using the state-of-the-art distributed model for analyzing soil water dynamics in a simplistic representative hillslope segment seems awkward as full potential of the model is not exploited.*

RL: We are thankful for this comment and will better explain our point of view in the revised manuscript. We are aware that distributed model studies often do a point to point comparison between observed and simulated fluxes and state variables in a catchment (Ebel et al., 2008; Scudeler et al., 2016; VanderKwaak and Loague, 2001). While this might be appropriate in less heterogeneous systems or in environmental system simulators as Biosphere 2 LEO, we regard this as rather difficult in more heterogeneous environments. In fact we think we can only be successful in doing point to point comparisons of simulated and observed state variables if the model is parameterized on highly resolved exhaustive observations of soil hydraulic parameters, rainfall and much more. As such observations are not at hand, we prefer comparison of statistical moments, based on the hypothesis that for instance observed and simulated soil moisture belong to the same ensemble and that they share similar dynamics. Also we started with what you claim the "*first step in modeling*" and did a comparison between soil moisture observations and soil moisture simulations at similar positions at a hillslope. Unfortunately this was not too conclusive. As can be seen in Figure 1 and 2 the variability of our measurements and observations cannot be grouped simply by their position in the landscape.

We do agree with reviewer #1 and #2 that we need to step beyond comparing medians of a single layer in our hillslope model to observations in the same depth. In the revised manuscript we will compare spatial variance of observed and simulated soil moisture observations using virtual observations. This virtual observations are located at different lateral positions on the hillslope in the same vertical depth (10 and 50cm) as the observations (For details see reply to reviewer 1).

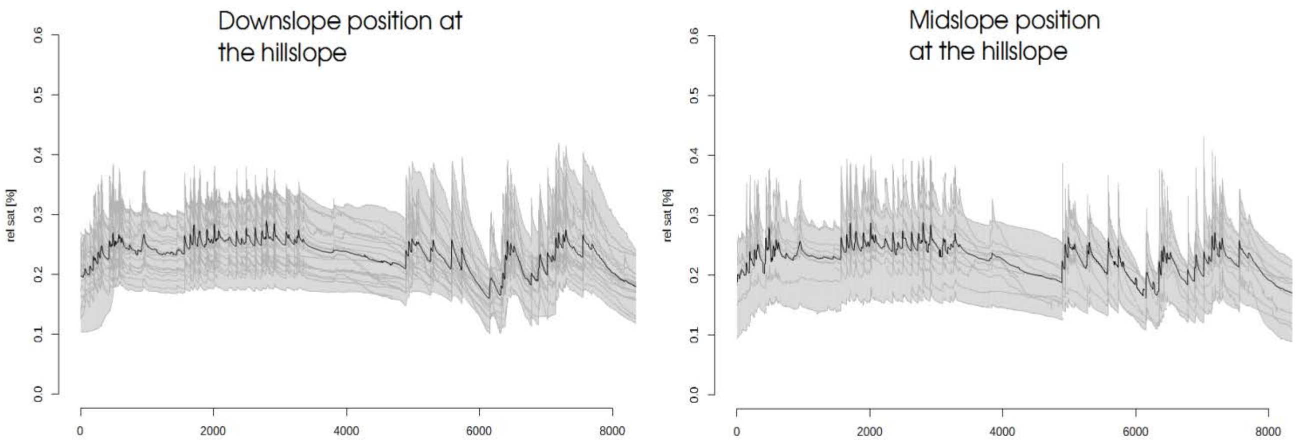

Figure 1 Soil mositure observation from downslope positions (left) and from midslope positions (right) at the hillslope.

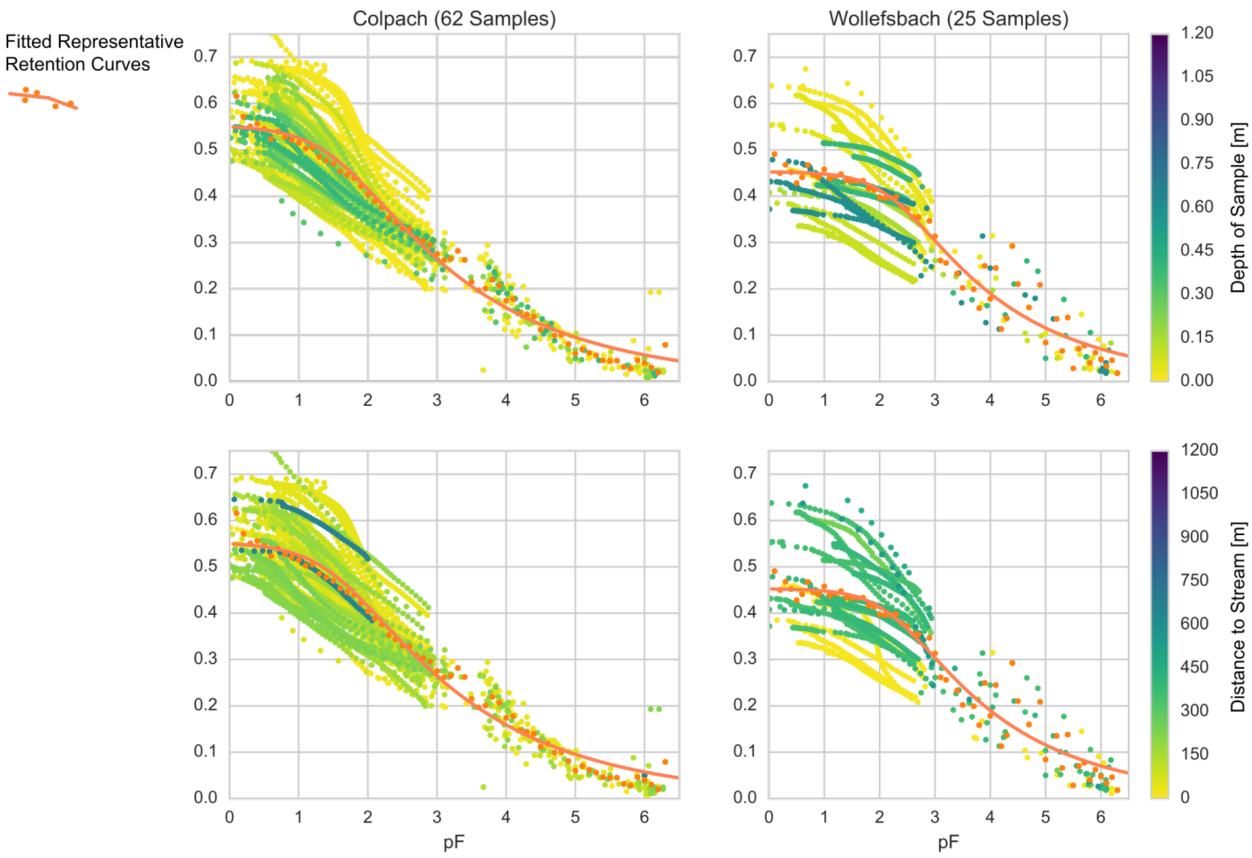

Figure 2 Fitted soil water retention curves (orange) and measured soil water retention relationships with a color key for their measurement depth (two upper plots) and for their distance to the stream (to lower plots) for the Colpach and Wollefsbach catchment.

Reviewer: *2. The hydraulic functioning of the simulated hillslope segment is questioned. Given the vertical height of the 2D flow domain (2 m), maximum soil depth of 1.8 m, and thickness of the soil-bedrock interface (0.2 m), bedrock was excluded at some locations. This hydraulic setting affects deep percolation fluxes across the interface. Furthermore, saturated hydraulic conductivities are too high for* *soil layer, soil-bedrock interface, and drainage system and too low for bedrock. Can still be laminar flow assumed for near saturated conditions (Richards eq.)? Is there any measurement indicating such high Ks value of the soil-bedrock interface? Can Ks value of the soil-bedrock interface be higher than upper soil? Similarly, bedrock porosity values seem too high, is there any (experimental) justification? Grid size of 1 m used for macropores seems unrealistic.*

RL: Thank you for this comment. In the following section we will explain the unclear parts of our parametrization in more detail. In a revised manuscript we will rephrase the parts which reviewer #1 and #2 found not conclusive. It is our goal to make our modelling as transparent as possible.

Bedrock parametrization:

Maximum soil depth with the bedrock interface was 1.8 m. Hence bedrock was not excluded at any location in our model domain. We apologize and we will stress this in a revised manuscript.

In both catchments we expect no major groundwater body beneath the hillslopes (Wrede et al., 2015). In the Wollefsbach catchment you can find redoximorphic features in deeper soil layers indicating a local episodic groundwater body above an impermeable bedrock. Also the schist bedrock of the Colpach catchment is besides of cracks and fractures described as quasi impermeable. We therefore think a Ksat of 10^{-9} m/s can be justified in both catchments.

Hydraulic conductivity of the soils:

Yes, the hydraulic conductivities of the young silty soils are high. These values are corroborated based on a large set of observation using both undisturbed soil samples and constant head permeameter measurements (Jackisch et al., 2016). Furthermore Wienhöfer et al. (2009) reports partly similar high values for a clay loam in a forested slope in the Austrian Alps. These high values of Ks and porosity might arise in case the fine silty material is aggregated with large inter-aggregate pores.

Bedrock interface velocities:

Yes, subsurface lateral flow can exceed Ksat values of the topsoil. Wienhöfer et al. (2009) derived these conductance estimates particularly for the lateral pipes based on a series for two tracer experiments in a forested site. He observed breakthrough velocities of $2x10^{-2}$ m/s and average travel velocities of 10^{-3}m/s for a lateral transport distance of order of 30 m. These high velocities coincided with a Peclet number of 2, which indicates that these fast reactions were mainly due to lateral preferential flow. Top soil Ks values in this area where on average 10^{-5} m/s. Furthermore Angermann et al. (2016) report breakthrough velocities of 10^{-3}m/s in the periglacial deposits which are on top of the schist bedrock in the Colpach.

Laminar flow assumption:

We agree with the reviewer that, with a Ks value of $5x10^{-3}$ m/s, a vertical grid size of 0.1 and a porosity of 0.25 we have a Reynolds number of 50. This is slightly above the threshold for strict laminar flow, which is 10 according to Bear (1972). This implies that strong fluctuations might cause turbulence, which, implies that we need a different flux law. However, we are still below the threshold of 1000, were flow is entirely turbulent. Nevertheless in a revised manuscript we will reduce Ks to $1x10^{-3}$m/s, to stay below the critical Reynolds number of 10.

Macropores:

We agree that a grid size of 1 m seem too coarse for the macropores. For a grid size of 1 m the Poisson process does often generate several macropores in the same grid cell, the generation process is carried out at a much smaller grid (1 -2cm). All contribute to the enhanced conductance of the model element, the enlarged conductance is hence an effective representation of the subscale macropores. This infiltrability is sufficient to allow a fast flow to the bedrock (in the Colpach) anddrainage (in the Wollefsbach) and then subsequent fast lateral flow. This is shown to be sufficient for runoff generation, but it has certainly deficiencies when dealing with solute transport. To improve this we usually work with an adaptive grid size of a few centimeters, generate the macropores, and reduce the grid resolution in areas without macropores. This was done in the Wollefsbach. We already employed the same approach in the Colpach, and simulated the hydrological year of 2014, the result is the same as with the coarser grid, as shown in Figure 3 and 4 in the reply to the first reviewer. In a revised manuscript we will reduce the grid size of the macropores in the Colpach to 0.1 m similar to the grid size in the Wollefsbach catchment.

Reviewer: *3. Runoff processes (subsurface lateral flow, overland flow, and deep percolation) and their partitioning mostly depend on hydraulic setup of the hillslope (see point 2). Observed catchment streamflow was compared with the sum of the three runoff component. It would be interesting to see individual contribution of each process to streamflow.*

RL: We agree with the reviewer and will analyze and display the three components separately in a revised manuscript.

Reviewer: *4. Some details of the model were not unveiled in the manuscript. For instance, algorithm of the root water uptake module remains unclear as well as the parameters (not sufficient to refer to previous study). As one hydrological year was considered in the simulations and most of runoff occurred during the winter season (Figure 8), no information was given on snowmelt runoff. More comments are appended below.*

RL: We apologize that this was not clearly written in our discussion paper. We will make our ET simulations in a revised manuscript as transparent as possible and rework the description. For further details we refer to our reply to the first review.

Reviewer: *5. The manuscript is too long. Some parts can be shorten and condensed (see Detailed comments). Please focus more on hydrology and less on philosophy. Conclusions part resembles Discussion.*

6. Discussion should be condensed, there are too many points which are discussed. The results obtained in this study are not related to previous literature, i.e. no deep discussion is provided. Focus on the main aspects of the results. Some discussion parts are too vague and superficial, adding limited value to *objectives ("to identify limits in our theories and related physically-based models") and overall knowledge.*

RL: We will streamline our revised manuscript by removing most of the virtual experiments, shorten our introduction and be more precise in our discussion. Our goal was to show that perceptual models can be used to constrain physically based hillslope models in a qualitative manner. Moreover, it was our objective to show that such a modeling approach can be based primarily on field measurements and literature values and therefore we intentionally avoided tuning of our model parameters against a single hydrological response. Finally we wanted to show which is the most important information or data source that is needed for setting up our hillslope models.

We apologize that we lost track at some of the sections of our discussion paper and think that with the comments and references of reviewer #1 and reviewer #2 we are able to be more concise and improve our manuscript.

Detailed comments

Reviewer: *Lines 14-27: This is too long introduction in the abstract. Please shorten.*

RL: We will consider shortening the abstract.

Reviewer: *L16-20: Conceptual models can also be physically-based. The use of conceptual models is incorrect.*

RL: We agree with the reviewer. In fact we adopted the term physically-based, as it is widely used for models based on the Richards equation. We will add a note to the introduction to clarify this.

Reviewer: *L34-6: Not true, internal water storage was not simulated well (see Figure 9).*

RL: We did not use the term “well” in line 34 but “some success” and we think this is true. We agree that the model performance could be improved by further tuning, but we do not think this is a real surprise neither was this our goal. Our objective was to set up our model based on field measurements where it is possible and on literature where it isn’t. Relating to our issue we think it is legitimate to write we had some success in simulating soil moisture observations as well as sap flow dynamics.

Reviewer: *L79-85: The literature body of recent hillslope 1D&2D modeling applications is somehow limited to a narrow window (mostly of the author’s group). For instance, predictions of 2D Richards-based hillslope model with a provision for preferential flow were compared with field data (hillslope discharge and spatially distributed pressure heads) in recent studies.*

RL: Thank you for that advice, we did not leave out studies on purpose. We will carefully check the literature once more to refer to other relevant studies.

Reviewer: *L110-3: Yes, hillslopes are indeed important in some headwater catchments. However, wetlands/riparian zones may control the runoff generation in other headwater catchments.*

RL: You are right, with our model we are not able to check how important wetlands are for the runoff generation. But we would like to highlight that there is also a saturated area at hillslope foot as a result of our bedrock topography. We further think that the riparian zone is still is a part of a hillslope.

Reviewer: *L181-4: A reference is needed here to support the statement on water balance. Introduction section seems to be quite long, it can be effectively shorten without losing the central messages.*

RL: Thank you, we will shorten the introduction. Especially the first passage and the part dealing with ET.

Reviewer: *L311-4: Double mass curve is not well suited for studying annual water balance since it relates runoff to precipitation. There is no provision for e.g. storage changes.*

RL: We think it is a suitable measure for detecting difference in seasonal runoff behavior. This is corroborated in the work of Seibert et al. (2016) and Jackisch (2015). We agree that strong inter-annual variations of runoff coefficient can be either a result of storage changes or due to changes in ET. The double mass curve cannot be used to discriminate these reasons. However, it is well suited to detect the shift between the summer and the winter regime. Furthermore do we not expect large groundwater bodies in both catchments. Hence inter annual storage changes shouldn't be of high relevance. We will explain this in the revised manuscript.

Reviewer: *L439-42: Any experimental justification for arrangement of the structures? Taken from previous studies on different hillslopes?*

RL: Several field experiments related to macropore flow took place in the Colpach as well as in the Wollefsbach (Angermann et al., 2016; Jackisch et al., 2016). We apologize as we did not do a good job in describing how we parametrized our model with respect to the macropore setup. We will make the parametrization of the macropores simpler and more transparent in a revised manuscript (for details see reply to reviewer 1).

Reviewer: *L453-4: Grid size of 1 m seems as a crude approximation for macropores.*

RL: We agree and will change this to 10 cm (See reply to reviewer 1 and above).

Reviewer: *L472-3: Ohm*m. In Figure 6B, contour line of 1500 ohm*m is situated in depths ranging from 1.0 m to 3.2 m for 100 m hillslope length. How such spatial configuration could be simulated in 2 m high hillslope segment? Please use the same hillslope segment in Figure 3 as ERT cross-section shown in Figure 6B.*

RL: Good point this was not explained well. Rather than directly transferring the ERT bedrock topography into our hillslope model we divided the deepest bedrock position and transferred it relatively to our 2 m deep and 350 m long hillslope. We chose a 2 m deep hillslope after analyzing the mean depth distribution of the 1500 ohm*m bedrock surface of 7 ERT measurements in the catchment. Our goal was not to have an exact image of the ERT measurement we show in Figure 6B but to generate a bedrock topography which follows the image of the perceptual model. We will better explain this in the revised manuscript.

Reviewer: *L474-7: Model of van Genuchten assumes zero air-entry value. Thus, alpha parameter is not reciprocal to air entry value. Not clear explanation of the soil hydraulic parameters - list also the parameters of macropores in Table 1. Furthermore, show the (variable) depths of the soil structures in Table 1.*

RL: We apologize for this. We will correct this in the revised manuscript.

Reviewer: *L501-3: Be more specific, what kind of data?*

RL: We will add a table with the vegetation parameters in a revised manuscript.

Reviewer: *L508-9: Pressure-water content relationship was measured in 0-3 pF range in detail, resulting in large spread of data points. This is not true for smaller pressure head values. These aspects may invalidate the fitted representative curves used in modeling.*

RL: It is correct that the soil water retention was measured with two different methods for 0-3 pF and above respectively. The Hyprop apparatus (for the low tension range) allows for a continuous measurement resulting in a large number of data points. The WP4C apparatus can only measure matric potential at discrete states and resulted in 2-5 points for each sample. For the calculation of the representative curve (as also for the fitting of individual retention curves to the data), the pF range is binned into steps of 0.05 pF. The resulting curve is fitted to the respective mean of each bin (when data is available).

This method reduces the effect of different data point densities. As we use all records of the grouped soil samples, there is a very strong basis for such a fitting also in the higher tension range. Hence the concern of the reviewer may not justify. Notwithstanding, there may be alternative means to derive such a representative retention curve.

To clarify this, we will change the used formulation L364ff:

"For both geological settings we estimated a mean soil retention curve by fitting a van Genuchten-Mualem model to all recorded retention data points of all soil samples in each group (51 and 28, respectively). This was done as maximum likelihood method accounting for the different data point density coming from the two measurement techniques. The tension axis was binned in 0.05 pF increments. The resulting curve is fitted to the respective mean soil water value of each bin (Table 1 and Figure 7)."

Reviewer: *L585-7: Does this mean that value of the saturated hydraulic conductivity was increased 75 times compared to reference scenario? Reference scenario used already high Ks value.*

RL: The point behind this virtual experiment was to show that emergent structures observed in the catchment have a strong influence on the runoff generation. With our static parametrization we unnecessarily constrain the agility of our model (Mendoza et al., 2015) and are not able to simulate the runoff generation in the Wollefsbach in summer. Instead of increasing the Ks value we should have added extra vertical macropores to our hillslope model. But as you point out above the strong increase of Ks actually violates the laminar flow assumption of the Darcy Richards equation. We admit that this virtual experiment was not done in a smart way and we apologize. We will remove this from the revised manuscript.

Reviewer: *L602-5: This can't be concluded, so far only runoff component of hillslope balance was shown. No comparison on hillslope storage was made.*

RL: We will revise this and state that the hillslope models closely portray the seasonal pattern of the catchment's runoff production, also because simulated and observed annual runoff coefficients match very well.

Reviewer: *L607-9: I would not say that in case of Wollefsbach, please see the winter period in Figure 8D. To support this statement, provide efficiency coefficients for both winter and summer period.*

RL: Thank you for this good comment we will do this in the revised manuscript.

Reviewer: *L610-2: Is hydrological year 2013 meant here? I miss the point. A1 scenario for Colpach has NSE = 0.84, so why discuss smaller NSE? Instead, scenario A2 for Wollefsbach requires discussion (NSE = 0.26).*

RL: Yes, we also tested the model for the hydrological year 2013. Since we only have runoff and rainfall for the year 2013 a more detailed test is not possible. Our goal was to show that our model also work in a different time period (Split sample test (Klemeš, 1986)).

Reviewer: *L613-5: Not clear what is discussed, Weirbach simulation?*

RL: We are sorry. We meant the Weierbach catchment which is a headwater catchment in the Attert basin. Our goal was to show that our model also works in a different catchment in the same hydrological landscape (Proxy-basin test (Klemeš, 1986)). We will clarify this in the revised manuscript.

Reviewer: *L624-5: Infiltration-excess overland flow was most likely not simulated due to extremely high Ks value of (top)soil. Was the extent of overland flow decreased by further increasing Ks value of bulk soil (emergent structures)? Such short high flow events also could not produce saturation-excess overland flow. Please check.*

RL: It is not the Ks value that is of relevance here but Ku(theta). The latter is in the order of 10^{-11} m/s in summer as specified in our reply to the reviewer 1. In Catflow we usually treat infiltration and runoff generation in a mixed or Cauchy boundary condition in combination with a fine discretization of the top soil layer (< 5cm). Rainfall is treated as flux boundary condition, until the upper element gets saturated within a time step. We then switch to a Dirichlet boundary condition while using the overland flow depth as pressure boundary. Infiltrating water flux is calculated according to the Darcy law: the product of the finite difference approximation of the potential gradient and the averaged hydraulic conductivity $k_{1+1/2}$ between the two upper nodes. When using the geometric mean sqrt(ksat*k(θ_2)) and a ksat of 1e-4 m/s and k(θ_2)= 1e-11 m/s the averaged conductivity is $k_{1+1/2}$ = 0.5 e-7 m/s. In fact it is the conductivity of the second node which acts as bottle neck and may produce infiltration excess.

Reviewer: *L666-8: This suggests misrepresentation of the soil profile (e.g., soil layering) as well as misparametrization of the soil hydraulic properties.*

RL: Since the model works acceptable in winter we think it is because we do not account for emergent structures like worm borrows and cracks. If we would have misparametrized our model why should it work in winter for soil moisture?

Reviewer: *L670: This paragraph needs a reference to Figure 10A. Evapotranspiration module is treated as a black-box for readers. We are left unaware what parameters were changed. Was there a difference between potential and actual ET fluxes? Did any water stress occur? Add information on Wollefsbach hillslope.*

RL: We are again sorry. We will make our ET simulation more transparent.

Reviewer: *L690-2: Do not write "it may be", this must be exhaustively explained by the model. Less storage above the interface for steeper hillslope setting would lead to increased storage in bedrock. Note that deep percolation was considered when comparison with observed streamflow was performed. This would also cause a delay in simulated runoff compared to reference scenario.*

RL: Sorry for using the term "it may be". We will rephrase this. The bedrock is quasi impermeable and why should a steeper hillslope increase the storage in the bedrock? Deep percolation is close to zero.

Reviewer: *L788-90: Satisfying match was obtained due to large measurement variability. Many scenarios with different parameter sets would fall within measured soil water content range (even different modeling approaches can provide similar match).*

RL: We refer here to the 12-hour rolling median not to the ensemble of all observations.

Reviewer: *L802-4: Given the comments above, I would hesitate to make such statements.*

RL: We think that the statement is well justified in respect to the runoff simulation. But not in the context of the simulation of the soil moisture.

Reviewer: *L810-2: The values contradict the previous statement.*

RL: You are right. We are sorry we meant 51.

Reviewer: *L812-5: I do not agree with this statement, see Figure 9.*

RL: Our model uses a representative soil water retention curve and does well with respect to discharge simulation in winter. If we now combine this with the fact that 90% of the overall runoff within a hydrological year is produced in winter one might come to the conclusion that heterogeneity of soil-hydraulic properties is not the major control on runoff generation in a hydrological year. Please note that we do not consider spatial variability of soil properties as unimportant. But since three to four parameter models able to simulate the runoff of catchments we think the heterogeneity of the soil-hydraulic properties cannot be of such importance for the runoff generation in a catchment.

Reviewer: *L886-9: It may be also due to location of a large depression (considered at the hillfoot region in this scenario). Therefore, these statements are not fully justified by the results.*

RL: You are right. We were too fast with our conclusion. Since this belongs to a virtual experiment we will remove this from the revised manuscript.

Reviewer: *L911: The message of this section remains unclear. What is suggested here? The need to use distributed model of representative hillslope with spatially uniform rainfall and pET fluxes?*

RL: We are sorry. The message was that a single hillslope won't be able to explain patterns which are a result of phenomena which have a larger spatial extent than the hillslope itself. To account for them you need a fully spatially distributed model.

Reviewer: *L936: This section is too vague.*

RL: This section will be removed from the revised manuscript.

Reviewer: *L995-7: This can't be concluded in such a general way. Beside hillslopes, riparian zone may play an important role in runoff generation in some catchments.*

RL: We think a riparian zone is still part of the hillslope. You could easily add a riparian zone to a perceptual model if you think that the riparian zone is a major control in your catchment.

The following comments are rephrased or corrected. We apologize for these mistakes.

L166-8: Is there any source to justify this statement?

L172-4: Add a few references here.

L223-5: Delete "concentration".

L268-70: ". . .perceptual. . .".

L421-4: Delete "a function of".

L424-6: This is not clear, please explain.

L428-30: The algorithm of RWU is not clear.

L450-3: Define slope angle. Was variable hillslope width considered (total area remains unclear)? The same pertains to Wollefsbach hillslope.

L464-6: Boundary conditions are not clearly explained. Does free outflow refer to free drainage BC (with unit hydraulic gradient condition)? Is gravitational flow boundary condition seepage face BC?

L486-9: I did not find any band generator in Zehe et al. (2010a).

L518-22: Boundary and initial conditions are identical for the two hillslopes. Do not repeat the information.

L523: "Model scenarios"

L527-31: Not left boundary (Figure 3CD)? What is the added value of log NSE compared to NSE criterion?

L551: "In VE2.1 scenario, . . ."

L558-9: Instead of "Last not least" use "Finally".

L565-8: Need to say what parameters were changed and what were kept unchanged. Otherwise it is a black-box.

L581-2: Delete "until the onset of the summer period".

L582-3: "Summer period was started with ...".

L593-4: This a harsh break from runoff to 2D saturation distribution, I suggest discussing runoff first and then move to saturation.

L622-4: Please provide the volume proportions (overland and subsurface flow, deep percolation) of runoff for both hillslopes.

L631-5: Not clear ". . . terrestrial filter properties . . .".

L708-10: Please make a reference to VE2.3.

L732-5: Please reword.

L737-41: Delete “revealed and” and “matching”.

L773-6: Instead of “parallel” use “lateral”.

L778-80: Instead of “in concert with” use “and”.

L796-9: Instead of “benchmarking” use “comparison”.

L828-30: Check KGE values (see Table 2 and Figure 11).

L884-6: Is VE2.3 scenario discussed here?

L908-9: Not clear what are SVAT modules.

L912-3: Not clear “. . . models encountered to capture flashy . . .”.

L1171-3: Cite this study in HESS.

L1272-5: Improve the reference.

Figure 3: Add dimensions instead of “small section”.

Figure 4: Dimensions are necessary.

Figure 9: No green color found. Stick to Colpach and Wollefsbach.

Figure 10: Relative saturation > 1?

Figure 12: Add hillslope location.

Table 2: Check KGE values shown in Table 2 and Figure 11.

Angermann, L., Jackisch, C., Allroggen, N., Sprenger, M., Zehe, E., Tronicke, J., Weiler, M., Blume, T., 2016. In situ investigation of rapid subsurface flow: Temporal dynamics and catchment-scale implication. Hydrol. Earth Syst. Sci. Discuss. 2016, 1–34. doi:10.5194/hess-2016-189

Bear, J., 1972. Dynamics of Fluids in Porous Media.

Ebel, B. a., Loague, K., Montgomery, D.R., Dietrich, W.E., 2008. Physics-based continuous simulation of long-term near-surface hydrologic response for the Coos Bay experimental catchment. Water Resour. Res. 44, 1–23. doi:10.1029/2007WR006442

Jackisch, C., Angermann, L., Allroggen, N., Sprenger, M., Blume, T., Weiler, M., Tronicke, J., Zehe, E., 2016. In situ investigation of rapid subsurface flow: Identification of relevant spatial structures beyond heterogeneity. Hydrol. Earth Syst. Sci. Discuss. 2016, 1–32. doi:10.5194/hess-2016-190

Klemeš, V., 1986. Operational testing of hydrological simulation models. Hydrol. Sci. J. 31, 13–24. doi:10.1080/02626668609491024

Mendoza, P.A., Clark, M.P., Barlage, M., Rajagopalan, B., Samaniego, L., Abramowitz, G., Gupta, H., 2015. Are we unnecessarily constraining the agility of complex process-based models? Water Resour. Res. 51, 716–728. doi:10.1002/2014WR015820

Scudeler, C., Pangle, L., Pasetto, D., Niu, G.-Y., Volkmann, T., Paniconi, C., Putti, M., Troch, P., 2016. Multiresponse modeling of an unsaturated zone isotope tracer experiment at the Landscape Evolution Observatory. Hydrol. Earth Syst. Sci. Discuss. 1–29. doi:10.5194/hess-2016-228

VanderKwaak, J.E., Loague, K., 2001. Hydrologic-Response simulations for the R-5 catchment with a comprehensive physics-based model. Water Resour. Res. 37, 999–1013. doi:10.1029/2000WR900272

Wienhöfer, J., Germer, K., Lindenmaier, F., Färber, A., Zehe, E., 2009. Applied tracers for the observation of subsurface stormflow at the hillslope scale. Hydrol. Earth Syst. Sci. 13, 1145–1161. doi:10.5194/hess-13-1145-2009

Wrede, S., Fenicia, F., Martínez-Carreras, N., Juilleret, J., Hissler, C., Krein, A., Savenije, H.H.G., Uhlenbrook, S., Kavetski, D., Pfister, L., 2015. Towards more systematic perceptual model development: a case study using 3 Luxembourgish catchments. Hydrol. Process. 29, 2731–2750. doi:10.1002/hyp.10393