Reply to Anonymous Referee #2

We thank the referee for the valuable comments on our manuscript. These comments will certainly help to improve the paper. We will address the comments below and based on these comments we will change or clarify the manuscript.
The referee correctly states that although we use a simplified model to describe fluxes, we use a detailed groundwater model to derive the PDF for unsaturated zone thickness. In a follow-up paper with new measurement data we intend to show that we can derive this distribution and shape function from groundwater head measurements and do not need to use the groundwater model to derive them. The main idea is that a nested scale setup can reduce the huge amount of groundwater heads needed to derive accurate PDFs at catchment-scale.

Introduction:
Page 3754/line 44

- We agree to change "deltas" to "lowland landscapes".

Theory:
Page 3758/line 22

- Figure 1 introduces readers that are less familiar with lowland landscapes to the dimensions and groundwater head changes we are dealing with in this paper. We also hope that this figure illustrates what we mean with a decrease in variation of the groundwater depth with decreasing storage, when we refer to it later on in the paper. Therefore, we would like to keep this figure.

Page 3760/ line 2

- The referee is correct when stating that drained and un-drained parts of the catchment will have different characteristics of stream network density and more importantly, different groundwater depth distributions. We contemplated using two separate models for drained and un-drained areas with different shapes of the groundwater-depth PDF and the relation between mean and variation of groundwater depth. However, this would lead to four new parameters (describing the new shape of the groundwater table) without any means to assign a value to these parameters. We did not measure un-drained and drained groundwater levels and the detailed (5 by 5 meter) groundwater model is not detailed enough to grasp the groundwater head variations between tube drains (there are only two or three cells between two tube drains that are 15 m apart). Therefore, the choice of using one model for the entire catchment is a pragmatic solution that is based on the available data.
  In the current framework only relations between parameters and the groundwater depth can be included. If a discontinuity is that large that you expect it to have significant effects on the outflow and this discontinuity cannot be described by a function of the groundwater depth, then a second model should be added. However, when one does not have the information to fulfill the extra data demand to set up this second model (i.e., no information about the shape of the groundwater table in this second area or the discharge hydrograph does not

significantly show the effects of this second area), it is often better to use a single model structure.

Page 3760/ line 13

- We will clarify section 2.2. There is a clear distinction between l_{surf} and q_{ex} + q_{dr}: l_{surf} is the horizontal moving flux between locations, while q_{ex} (groundwater drainage), and q_{dr} (tube drainage) are the fluxes between the compartments of groundwater and surface water within a location. A part of q_{ex}, however, can be stored on the soil surface in ponds (S_{surf}).The change in soil surface storage therefore is the difference between q_{ex}+ q_{dr} and l_{surf}. This is also the reason that all three variables occur in the mass balance of the surface storage compartment.

Page 3762/line 18

- The delaying effect of the unsaturated zone the referee discusses is not included in our model although we agree that it exists. We assumed instantaneous equilibrium between all fluxes, and consequently any delays are ignored. Because we have shallow groundwater tables most of the time (especially during discharge events), this assumption seems justified. However, we have observed, just like Seibert et al (2003), that groundwater tables increase much more during wet conditions than during dry conditions in reaction to precipitation events of similar magnitude. We contribute this to a higher effective storage coefficient during dry conditions (close to the porosity) than during wet conditions. During wet conditions (high groundwater tables) the unsaturated zone is already very wet, with little storage space for extra water under unit gradient flow. A few drops of rainwater can then cause the groundwater level to increase many times more than it would do under dry conditions by effectively turning the capillary fringe into groundwater. For more examples of this effect see also Bierkens et al (1998) and Kim et al (1996).

Page 3770 line 8

- Surface storage in the model is always completely filled and in equilibrium with the saturated groundwater. The volume of surface storage is a function of the average groundwater depth and its variation. This volume is instantaneously subtracted from the saturated groundwater when the groundwater depth decreases (wetting). When the groundwater depth increases (drying) the excess surface storage is added to the saturated groundwater. Overland flow is caused by the head gradient between groundwater and surface storage: Surface storage is defined as a fraction of the volume of groundwater heads above surface elevation, $-m*u$ $u<0$. Therefore, the gradient between groundwater and surface ponding is given by $(m-1)*u$ $u<0$.

Page 3773/ line 10

We agree that a single value of *u,* separating zero and potential evapotranspirations, is a huge simplification, and at the local scale it may even be an over simplification of reality. See our discussion of a similar comment by referee 1 for a detailed rationale for this approach and a comparison with a more realistic approximation.

Materials and methods.

Page 3776 / line 21

- We agree to remove figures 5 and 6.

Page 3776/line 22

- These are surface water levels. We estimated these surface water levels for all ditches and brooks from a detailed airborne radar image. The water levels are of course the water levels at the time of flight, but we used them in the groundwater model year-round. We do not calculate surface water levels dynamically because there are just too many surface water levels in the 65 km of streams and ditches within 6 km^2. We only use the draining function of each water course with a year-round fixed water level, but do not route the water downstream. We will clarify this in the manuscript

Results and discussion

Page 3777/ line 20

- We did not calibrate the model, as we explained in the materials and methods section 3.2. We checked the model for inconsistencies with measured data and performed a sensitivity analysis. It turned out that groundwater heads and discharges were most sensitive to the phreatic storage coefficient. The phreatic storage coefficient expresses the units of rainfall needed to obtain a unit groundwater level increase (a value between 0.05 and 0.25). The representation of the unsaturated zone by a single storage coefficient is just not good enough to get a good fit during both dry and wet conditions (under wet condition it will be around 0.05 and under dry conditions around 0.25). This again stresses the importance of the unsaturated-saturated groundwater interaction as implemented in our model. We chose not to calibrate the model, because by calibration we would tend to underestimate the storage coefficient, in order to represent peak discharges and peak groundwater heads (during peaks errors in groundwater head and discharge are largest). However, the shape of the groundwater table is only moderately influenced by the phreatic storage coefficient. This shape (the curvature of the groundwater table between draining ditches/tube drains) is mainly determined by the positions of ditches, tube drains, and streams, the hydraulic conductivities and the surface elevation. For this study we were mainly interested in the spatial mean and corresponding spatial variation in groundwater depths.

  We pointed to an under estimation of overland flow by the groundwater model as a possible source of error. This is strongly related to the phreatic storage coefficient. A low storage coefficient gives high groundwater tables during wet periods and consequently more overland flow.

  In our simple model we cannot easily distinguish between groundwater flow to streams and ditches and overland flow, because at point scale there is not much difference between both routes (both routes occur when groundwater heads are above surface elevation). Therefore, in Figs. 15 and 17 we do not distinguish overland flow separately.

Page 3780/ line 22

- Yes, we mean saturation excess overland flow and will change this.

Page 3781 /line2

- We like to keep these first parts of figure 10, because they put the distributions of figures c and d into context. The extensive drainage area during wet periods translates into a relatively large area with negative groundwater depths.

Page 3784 line 21.

- The hysteresis remains difficult to see with the presented dataset. We also agree that there are many more sources of error that can cause deviations between measurements and model. Measurement errors in rainfall, calculations of evapotranspiration by Makkink-method are major sources of error. Also the equilibrium assumption for the unsaturated zone can cause errors: During infiltration the unsaturated zone contains more water than under unit gradient flow, while during evapotranspiration the unsaturated zone contains less. We will try to expand the discussion on the sources of error and their effects, and we will remove the focus on hysteresis.

Conclusions

Page 3786/ line2

- The PDFs of the unsaturated zone thickness have to be determined via other sources of information than discharge in order to prevent large equifinality problems. The PDF's can be constrained by groundwater head measurements or a groundwater head simulation or possibly remote sensing of the catchments wet fraction.

  A huge advantage of the groundwater depth PDF is that, although you need measurements on many locations and under different discharge regimes, the groundwater depth itself is easy and cheap to measure. Far more easily than for example the soil parameters that are needed to build a groundwater model. By measuring the groundwater depth itself, all effects caused by heterogeneity of soil and vegetation are indirectly measured too. A well chosen nested scale setup could circumvent the large amount of measurements needed to characterize catchment-scale groundwater depth PDFs.

References

Bierkens, M.F.P.: Modeling water table fluctuations by means of a stochastic differential equation, Water Resour. Res. 34, 2485-2499, 1998.

Kim, C. P., Stricker, J. N.M. and Torfs, P. J. J. F: An analytical framework for the water budget of the unsaturated zone, Water resour. Res. 32, 3475-3484, 1996.

Seibert, J., Rodhe, A, and Bishop, K.: Simulating interactions between saturated and unsaturated storage in a conceptual runoff model, Hydrological Processes 17, 379–390, 2003.