Reply to Anonymous Referee #1

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.

In her/his first paragraph, referee 1 states as a major concern our treatment of the unsaturated storage as a variate without spatial pattern. We believe this simplification does not impose severe limitations on the catchment scale, and the model performance supports our view. Should the model fail for another catchment, it will of course have to be improved. In any case, starting with a simple model, testing it, and making improvements only when necessary, is a feasible approach. The model presented here offers new possibilities to do so, as the reviewer recognizes. For example, if a clear relation between the groundwater depth and the parameters that form the unsaturated zone storage function exists, these relations can easily be added to the model. Spatial patterns uncorrelated to the groundwater depth are assumed to be of minor importance within the pre-defined lowland catchment. If these patterns do exist and are expected to have significant influence on the discharge and groundwater depth, the catchment should be subdivided into separate models. We will clarify the assumptions and consequences in the paper.

Major comments:

Referee 1 offers nine major comments. The first deals with the lack of correlation between groundwater depth u or $\mathrm{d}u/\mathrm{d}t$ and other variables, such as the porosity and the drainage depth. The true test of these assumptions lies in the application of the model. Good model performance validates the assumptions that underlie it. The paper argues that the hydrograph of lowland catchments is mainly determined by the groundwater depth. The effects of spatial and temporal variations in groundwater depth outweigh the effects of spatial and temporal variations in all other variables/parameters. Therefore, we chose to create a water balance model around the distribution of groundwater depths and make all flux generating processes only dependent on groundwater depth, rainfall and evaporation. As a consequence, we were able to make many simplifications, for example, one soil type, no correlation between groundwater depth and tube drain location, flow resistance parameters and many more. In spite of these huge simplifications, the model is still able to describe discharge and groundwater depth well. This means that the model results do corroborate the chosen assumptions.

If the simplifications appear to be detrimental to model performance, that clearly signals a need for more detailed modeling, with its associated burden of heavier input requirements, more extensive calibration, etc. Though possible, this is unattractive, and our results so far have proved it unnecessary. This is related to major comment 7 discussed below.

The second major comment deals with the relation between the assumptions of instantaneous hydrostatic equilibrium and soil homogeneity. Here, the reviewer appears to have interpreted that instantaneous hydrostatic equilibrium implies a uniform soil. Upon re-reading our text we realized we need to improve our wording to avoid this. What we meant to convey was that hydrostatic equilibrium is a substantial simplification of the natural situation, even more so than the assumption of a single uniform soil. We considered it inconsistent to assume this extremely simplified hydrology of the unsaturated zone on one hand while on the other facilitating the spatial variation of the properties of this unsaturated zone. We therefore meant to argue (and will do so better in the revised text) that, if one is prepared to assume instantaneous hydrostatic equilibrium, one might just as well assume the vadose zone to be uniform, as a further but less consequential assumption with considerable benefits in terms of parameter requirements and measurement effort.

The third major comment addresses a potential inconsistency arising from the lacking correlation between the spatial variation of u and the variables that could explain this variation. We are aware of these relations and point out here that the variation in surface elevation (including the bottom of streams and ditches) is one of the main causes of groundwater depth variations. By assuming uniform soil properties and not considering the spatial distribution of the drain tubes, two factors that affect local groundwater depth cannot be explicitly taken into account when determining the spatial variation of u. Whereas the reviewer considers this a potential inconsistency, we adopt the view that our fixed distribution of groundwater depths pragmatically and elegantly incorporates the effects of these and other spatially variable features on the distribution of groundwater depths. At the same time, this allows us to maintain spatially and temporally constant parameters to quantify storage changes and fluxes. Thus, the fixed distribution of groundwater depths captures implicitly the combined effect of all heterogeneities of the soil, the vegetation, and the drain tube depths and locations.

The fourth major comment mentions the spatial correlation and connectivity issues. We share this concern. Particularly for overland flow and stream flow, a poor representation of connectivity can affect the response of discharge to rainfall. We will follow the reviewer's suggestion and address this aspect in the revised text.

The fifth major comment is related to lateral flow within the catchment. Subsurface lateral flow within the porous medium (i.e., excluding drain tube flow) is assumed to occur in the saturated zone only (Eqs. 4 and 7). Integration over the catchment area only leaves the net lateral flux across the catchment boundary (Eqs. 8 and 9). Thus, lateral fluxes within the catchment are not ignored; they merely cancel out in the averaging operation. A catchment boundary is usually defined by the fact that water infiltrating on either side of it will be discharged by different streams, i.e., a no-flux boundary. As we explain below Eq. (9), we used this line of reasoning to set the integral of the lateral flux across the boundary to zero.

The sixth major comment criticizes our use of a groundwater model to provide information needed to drive our newly developed model. We are sympathetic to this viewpoint. Initially, we tried to fit the resistance parameters of the model and also fit the groundwater depth distribution parameters to a dataset of discharge and groundwater heads, but we ran into large equifinality problems. We realized that, in order to apply our model, we need to derive the shape of the groundwater depth PDF separately, by measurements or another model. Because the shape of the groundwater depth PDF is mainly determined by surface elevation, stream network locations and saturated hydraulic conductivity, a groundwater model is a good instrument to derive the PDF based on these spatially distributed datasets. In a follow-up study we intend to show that we can also describe the PDF from groundwater head measurements. By having a well-chosen nested-scale setup of discharge and groundwater head measurement we can bypass the problem of needing many groundwater head measurements to derive an appropriate PDF.

The seventh major comment raises the issue of the local starting point of the model development, and uses our simplified description of local evapotranspiration as an example. Reviewer 2 also comments on our approach to model evapotranspiration, so these comments merit a thoughtful reply. We originally implemented a Feddes-type function with two thresholds with a linear reduction of evapotranspiration between them, in line with the comments. It transpired that fitting the additional parameter was not possible using groundwater head and discharge data. Furthermore, we could not achieve an improvement in the model performance.
Stimulated by the comments of both reviewers, we revisited the problem and carried out a sensitivity analysis. Figure 1 shows that the two-parameter function gives nearly identical results as the single-parameter approach for the conditions we used to represent the catchment (standard deviation SD of groundwater depths between 0.2 and 0.4 m [0.3 m in Figure 1], and a linear reduction of evapotranspiration for groundwater depths increasing from 0.8 to 1.5 m). When the SD is reduced to 0.1 m, reflecting a smaller model area with less variation, the difference between the one- and two-parameter functions increases because the smoothing effect of the averaging operation is less effective. We will add to the revised text the possibility to use a more elaborate function to facilitate the modeling of smaller areas with limited variation in groundwater depth. Nonetheless, we respectfully disagree with referee 1's suggestion that it is necessary to start with a physically realistic process-description at the local scale. If the variation within the averaging area is large enough to dominate over local subtleties in the response to driving factors, such local variations will be drowned in the averaging process. In that case, the representation of the local process need only be unbiased to provide accurate results at the catchment scale. This improves computational efficiency of the large-scale model and makes it parsimonious in its parameter requirement if the representation of the local processes is prudently chosen. This line of argument goes back to the introduction of parallel column modeling by Dagan and Bresler (1979).

The eight major comment states that the effect of decreasing SD of the groundwater depth for wetter conditions has only a limited effect. For the follow-up study we reanalyzed the groundwater heads measurements. Now, we did not just calculate the standard deviation from the measurements, but first interpolated a groundwater table between the measurements and then calculated the groundwater depth PDF from every point in the field. Effectively, we weighted the contribution of individual measurements to the area they represent (we had a higher density of measurements close to the ditch). This resulted in a clearer tendency to smaller standard deviations under wet conditions (Figure 2). We will include this new analysis in the manuscript.

With the groundwater model it would not be possible to realistically simulate a wetter catchment, because the model does not have dynamic surface water levels or the capability to pond. The groundwater level will just follow the surface elevation under very wet conditions and tend to a standard deviation of 0 and a mean of 0.

The ninth major comment recommends to elaborate on the difference between our model and existing models. Other models such at TOPMODEL and the soil routine in HBV also use spatial distributions. The primary source of variation in these models is either slope type or soil type, which can be derived from topographic maps and remain constant in time. The main difference with the model presented here is that we chose groundwater depth as the main source of spatial variation that influences discharge generation. This distribution is not constant in time but a function of storage. We therefore defined relations between the distribution parameters and the storage. This resulted in a much more dynamical model driven by continuously changing groundwater head gradients.

Minor comments:
We indeed used changes in unsaturated zone thickness to describe changes in saturated storage. We will try to clarify this.

Page 3755, line 3
We meant to say that these models have been originally developed for a sloped catchment. Consequently, they do not focus on the processes that are important in lowland catchments, such as unsaturated-saturated zone interactions, surface ponding and a strongly changing active drainage area.
We admit that some of these models can be applied successfully to lowland catchments, but in our paper we show that a continuum of point-scale linear reservoirs, including unsaturated-saturated zone interactions and surface ponding, is much more appropriate to model lowland catchments

Page 3776
We will clarify the manuscript by combining sections 3.1.4 and 3.1.3.

P 3778
The 5 cm standard deviation was an intuitive estimate for 5 by 5 m cells. We felt that we needed slightly higher standard deviation for a continuous surface elevation. However, we have to take into account that also the surface elevation measurements from radar will have an error with a standard deviation of a few centimeters. We will describe this assumption more clearly in the manuscript.

P3780
The weighting factors were determined by an iterative procedure. We wanted all the different components of the objective function to contribute in the same order of magnitude to the objective around the optimal solution. First, we ran the optimation with a first estimate of the weighting factors and changed their values based on these results to ensure an equal contribution around the optimal solution. This will be clarified in the manuscript.

P3780

We like to keep the discussion and results section together. We think this improves the context of the discussion, by being able to refer directly to the results. One can argue either way, it is to a high degree a matter of personal preference.

P3782

First, we found good solutions with unrealistic soil parameters, but we saw thatmany combinations of soil parameters would give acceptable results. Therefore, we added the values reported by Wösten et al. (2001) as prior information to the objective function to ensure that the solution would be close to the values found by Wösten et al.(2001). We will describe this in more detail in the paper.

We had to use the estimated tube drain contribution in the calibration. There is not enough information in the hydrograph and groundwater levels to make this distinction between flow routes. Therefore, it is a good thing that the calculated contributions are close to the estimated contribution, but it is certainly not an independent validation.

P 3785

The hysteresis remains difficult to see with the presented dataset. We will change the text, so that is does not focus on possible sources for hysteresis, but focuses more general on the most important sources of errors.

References

Dagan, G. and Bresler, E.: Solute dispersion in unsaturated heterogeneous soil at field scale: I. Theory, Soil Sci. Am. J., 43, 461-467, 1979.

Wösten, J.H.M., Veerman,G.J. De Groot, W.J.M. and Stolte, J.: Waterretentie- en doorlatendheidskarakteristieken van boven- en ondergronden in Nederland: de Staringreeks. Vernieuwde uitgave 2001.; 86 pp http://www2.alterra.wur.nl/Webdocs/PDFFiles/Alterrarapporten/AlterraRapport153.pdf, 2001.

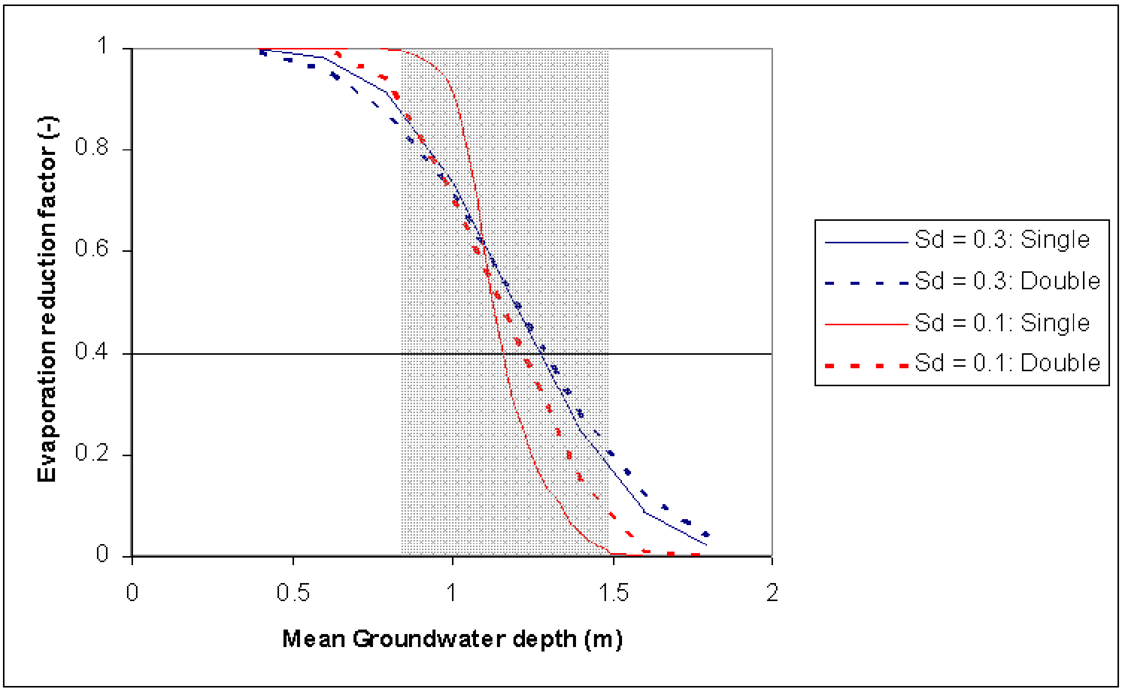

Figure 1. Sensitivity analysis of the evapotranspiration reduction function. The solid lines give the reduction for a single switch point representation, while dotted lines give reduction for the double switch point representation with linear decline (Feddes function). The shaded area gives the values between which there is a linear decline for the double switch point representation. The single switch point is located exactly in the middle of the shaded area.

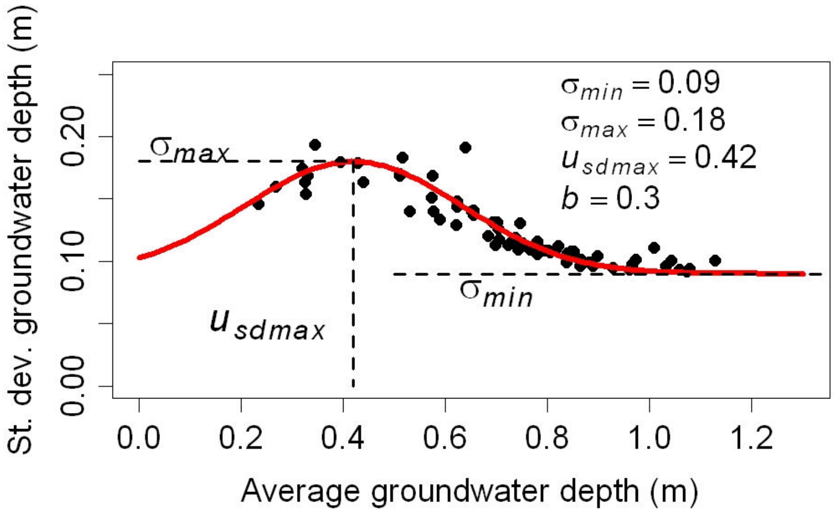

Figure 2 Relation between the average groundwater depth and the standard deviation of measured groundwater depth at the field site.