The work by Trambauer et al. if focused on the possibility of reliably simulating the flow field in the Holocene peat layer that outcrops in a large portion of The Netherlands. This task is accomplished by both experimental and modeling activities. The objective is clearly very interesting from the scientific point of view as the hydrologic response of peats is well known to be very complex. The goal is also very important from an applicative point of view because this peat unit constitutes the seal for the underlying sandy aquifer.

The major problems concerning the work described in the ms are related to the modeling effort. In fact, if the research focus is on the flow field in the upper 6-m thick peat layer, why the model domain is extended down to 300 m depth (p. 2075)?

Indeed we have also had the discussion whether to restrict the model to the CCL of 6 m thick or extend it to the bottom of the groundwater system. The reasons to select the deeper model: 1) especially for the steady state model, pressure heads in the first sandy aquifer to define the model's **bottom boundary** heads were not available at the experimental field. The closest wells with such data are more than a kilometer away from the site (see new figure 1). Estimation of the heads at the field (Kriging) would have led to errors which would have made model calibration more difficult, and 2) an existing regional model was already available to provide hydro-geological input data (top and bottoms of layers/conductivities) for the deeper layers. *A motivation for the boundary selection will be added in the article.*

Right bottom part of new Figure 1

As the experimental investigation revealed a significant lateral heterogeneity of the CCL also at the local scale (see Fig. 3), why and in which way the simulation domain is extended from the dimension of the test site (about 50 m long) to a 900 m long domain (p. 2076)?

Similar discussions we had whether to confine the model to the experimental field of 55 by 250 m where most of the data were collected or to the larger area of 700 by 900 m. The reasons to select the larger area: 1) having decided on a deeper model, pressure heads in the first sandy aquifer and deeper aquifers to accurately determine **lateral boundary** heads were not available at the field, and 2) although there are lateral heterogeneities in lithology in the peat area (total peat area is in the order of 5*5 km) these rather concern variations in thickness of the lithological layers, but the main picture remains intact: a 5 to 7 m thick CCL mainly consisting of peat with clayey layers at some depth causing a head difference (downward flow) ranging between 30 and 100 cm (see new Figure 3). There is no indication that a completely new lithology is present across the borders of the experimental field and from a hydro-geological point of view an extension is justified. *A motivation for the boundary selection will be added in the article.*

In my view, much more reliable results should be obtained by modeling only the Holocene unit in the experimental site with: - a much more refined grid able to represent accurately the peat heterogeneity, in both the horizontal and in the vertical directions to obtain much more accurate velocity fields and travel times; - using the pressure head in the first sandy aquifer as boundary condition on the domain bottom; the simulation of the confined aquifer system is absolutely useless; - using the water level in the 2 ditches limiting the experimental site as lateral boundary conditions.

At the experimental field we have only two wells into the first aquifer with an accurate lithological description of the CCL. The data are insufficient to represent accurately the peat heterogeneity by showing at many points in this field the variation in lithology. This also makes the use of a finer grid less useful since most of the cells are not supported by lithological evidence. *Grid size selection will be explained in the article*

Only very qualitative information can be obtained with the present model set-up. This is the main reason why I don't support the publication of the present ms in HESS.

We tend to disagree with the statement that 'only very qualitative information' can be obtained although this is also due to the fact that certain aspects have not been emphasized in the article (to keep it brief). First of all, there is the already mentioned point that although there are heterogeneities in the CCL there is also the continuation of the typical peat/clay layering and head differences in this layer across the boundaries of the experimental field into the model area.

Therefore the hydro-geological set up in the (larger) model area is realistic which leads to reliable results. Secondly, the article intends to compare a 4-layer set up for the CCL with the common 1-layer configuration. The obtained results between differences in computed heads, in flows (fluxes) and travel time are useful since the comparison is a comparison between two different concepts not requiring a 100% accurate representation of the reality. *Also, a statement on the usefulness of the model set up will be added in the article.*

There are also other weaknesses with the paper:

1) the conclusion about the possibility of using Darcy's law in peat soil is quite

ambiguous. What means "Based on this

information one is inclined to believe that Darcy's law with constant Kh may not be

valid ... but that its application in model computations is justified" (p. 2076)?

In fact the sentence said that ''may not be entirely valid'' meaning to say that errors introduced by applying the Darcy relation are very small and that computation (modeling) results are still very useful. *We will re-phrase this sentence in the article.*

2) slug tests were carried out to characterized Kh in the peat layers (p. 2070). But the

conductivities obtained by their interpretation are representative of which layers? It is

strongly evidenced just above that the CCL is very heterogeneous. Moreover, Fig. 5

shows that during the slug tests the water level remains well above the original position

(at time =500 s the water level displacement is stiff 50% of the initial value). Why?

The tests were mainly done in the first and or second layer below surface in the depth range of 0.5 to 2.2 m. The tested material was (muddy) peat, mud or a mix of peat, clay and silt (see Figure 4). Although being heterogeneous the K values could be attributed to a specific lithology. The reason that after 500 s the water level is still at 50% of the original level is a result of the relatively low conductivity of the peat (about 1 m/day). Further monitoring would have seen heads approaching zero displacement. *Some explanation will be provided in the article.*

3) data about the result of the monitoring activity (e.g., water level vs time at the different

monitoring points) must be shown;

True*, we will add a graph showing the monitoring results*

4) what are the "standard rules for model calibration" (p. 2075)? What are the "GxG

maps" (p. 2078)? Are the resolution of these maps comparable with that of the investi-

gation carried out to characterize the experimental site?

The models have only been calibrated with respect to heads and Konikow and Bredehoeft (1992) have pointed out that such calibrations have limitations in the sense that errors cannot be completely prevented. This is a result of the non-uniqueness of the solutions when calibrating (and validating) against heads only (we wanted to calibrate/validate against flows as well, but flow data could not be provided at the last moment). *The sentence will be rephrased and the article will be added to the list (Advances in Water Resources 15 (1992) 75-83: Ground-water models cannot be validated, Leonard F. Konikow & John D. Bredehoeft)*

The G*G maps provide phreatic groundwater levels on a 25 * 25 m basis and are based on data collected at (scarce) observation wells. Values for intermediate cells have been calculated with formula taking into account typical Dutch terrain conditions. It is true that the resolution of the G*G maps is not the same as the model grid resolution (5 * 5 m), but we have looked at aerial averaged computed values when comparing with the G * G maps. Also, at the experimental field computed phreatic levels have been compared directly with observed values. *We will provide some explanation in the article.*

5) Section 5.1.3 is useless and can be deleted;

We assume that the section which is meant is 5.1 and then the part on the Groundwater balance. This part is necessary as the balance is usually one of the main considerations for model design. For example, this part indicates that both recharge and capillary flow have to be considered in the model. Also it is shown that the interaction with the ditches has to be taken into account. But most of all the section contributes to the general understanding of the flows taking place in the CCL

6) do the simulations account for the water flow in the shallowest unsaturated CCL? Or

is the CCL completely saturated?

No, the simulations are only for the saturated part of the CCL. The CCL is almost completely saturated; only the upper 30 to 70 cm is unsaturated. *We will add some explanation in the article*.

7) in p. 2078 it is written that the "transient model was initially built ... for the winter

period 2009 to 2010" and then "extended ... to simulate 4 and 8 yr simulation periods".

But in which way this task is carried out? What is the reliability of extending data

recorded over a winter season to a 10 yr long period? Some discussion is needed;

''Extension'' may be a misleading word. The transient model was calibrated for the winter period 2009 to 2010 and then the model period was lengthened to simulate from 2006 to 2010 (4 yrs) or

2002 to 2010 (8 yrs) using historical data on recharge, open water levels, etc. We still think that reliability is in order since the historical data are of good quality. The model storage parameters that are also needed could be calibrated using the winter period (although we would have liked to do a ''summer check'' as well). *We will explain this issue a bit more in the article.*

8) it is a nonsense to compare the model sensitivity to hydraulic conductivity and

groundwater recharge (p. 2079-2080). The former is a factor forcing the system, the

latter an intrinsic property of the system;

We are not sure what this comment means. Certainly we are not trying to compare conductivity with recharge, but only the effects that (changes in) these parameters have on the computed heads in the model. This is one of the items (in addition to parameter uncertainty and realistic parameter value ranges) to be considered, before one starts on model calibration.

9) model results in the heterogeneous simulations (Figs. 8, 10, and 12) need to be

clarified. Due to the CCL heterogeneity (Kv in the 3rd layer is 3 to 4 order of magnitude

less than in layers 2 and 4 - see Table 2) it should be expected a much longer travel

time to cross unit 3 instead of the other layers (i.e., much more bullets within the 3rd

FD layer). But this is not what the pictures show. Why?

Yes, an increase in travel time as a result of a lower velocity that is indeed what one would think first, but in reality velocities across the confining layer do not change much. What actually happens in the confining clay layer (layer 3) is a drastic increase in (head) gradient. This was also clearly observed in the field (large differences in groundwater levels in the intermediate boreholes, just above the clay, and in the hole in the first aquifer). Also the effective porosity in the clay layer is lower. Both parameters increase the velocity and ''off set'' the much lower value for the conductivity: v = (Kv (much lower)/n (lower))* i (higher).

10) Fig. 1: all the borehole locations and the modeling section must be shown in the

aerial photograph;

Correct, this will be done

11) Figs 8 to 13: add proper vertical and horizontal scales. Why the domain of Figs.

10-13 is different from that used in Figs. 8-9?

Yes, we will add. There is no specific reason why the domain of Fig 10-13 is different from Fig 8-9 (Fig 10-13 is enlarged a bit to obtain a better view of the complex flow paths).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 2065, 2011.