Review #1

Field-scale leaching of a conservative tracer and an organic chemical was measured using a multi-compartment sampling device. Solute breakthrough was measured at a high spatial resolution for different flow regimes (snowmelt versus irrigation). The solute transport experiment was then calculated with a 2D variably-saturated water flow and solute transport model; random fields of a lognormally distributed scaling factor were used to mimic spatial variability in hydraulic properties. By trial and error the geostatistical parameters for hydraulic functions best describing the leaching data were identified. Additional flow boundary conditions were imposed during modeling to explore the effects of flow conditions on solute leaching.

This paper presents an interesting data set and associated calculations on a subject of broad interest to HESS readers. However, the experimental and numerical analysis are not really new as another paper on the same subject was recently published by the same authors (Schotanus et al. 2012) – also, similar work has been reported using wick samplers (eg Seuntjens et al. 2001 JContHydrol 51:13-39) with the main difference that the spatial resolution of their MCS is a bit higher than that of others (suction regulation with wick samplers is also possible). The main limitation of the current paper is its lack of detail in explanations and descriptions of assumptions, hypotheses, etc. Indeed, many statements are insufficiently underpinned: for example, comparison between flow heterogeneity for snowmelt and irrigation is hampered by incomplete description of the experimental setup, insufficient [basic] knowledge of soil hydraulic properties. I recommend to study solute leaching in each collector separately and try to calculate the pore-water velocity from moment analysis (Veffective); if this is compared with the application rate divided by an average water content (v0), the ratio Veffective/v0 is an indication of the degree of preferential flow (see eg. Mallants et al. 1996 – Hydrological Processes 10:55-70). The paper further lacks clarity and transparency in describing assumptions and hypotheses. Many of the conclusions are not supported by comprehensive data or model output. It is not clear how a difference in soil hydraulic properties can be the reason for a different flow regime when the same experimental setup is used. There is mentioning of uncomplete BTCs; what does that mean exactly? It needs quantification in terms of solute mass balance. When a single realization approach is used, inference of statistical moments of the distribution of relevant variables requires a flow domain that is sufficiently large compared with the correlation scales of the pertinent formation properties. In addition, in order to preserve details of the spatial structure of the formation properties, the size of the numerical cells must be small compared with the characteristic length scale of the heterogeneity of the relevant formation properties. Based on the criterion of at least four nodes per correlation length suggested by Ababou [1], the horizontal discretisation is too large compared to the value of the horizontal correlation length. Furthermore, the domain size should be at least 10 to 50 times the correlation length; the vertical domain size is too small (1.5 m versus a required 6 m if vertical correlation length is 0,6 m). There are many more comments on the annotated pdf-version of the manuscript.

We thank the reviewer for the effort he/she put in the review. Below, we will respond to all comments. As the above comments are also found in the supplement, our answer to these comments can be found below as well.

page 13452

r 17: Do you mean fraction of soil depth or fraction of pore space; was there any relation between infiltration rate and degree of non-equilibrium (see eg Nkedi-Kizza et al. 1983)

We mean fraction of pore space. We added this to the ms. With "The fraction of the soil that contributes to solute leaching… " we mean that in a cell or node no drainage occurs, or the concentration is 0, both throughout the experiment. We changed this in the text (see also comment 13459 line 24).

13453

r 21-23: A similar approach was developed by Seuntjens et al. (2001)- Grid lysimeter study of steady state chloride transport in two Spodosol types using TDR and wick samplers JContHydrol &

We included Seuntjens et al 2001 in the introduction. In Seuntjens et al a lysimeter experiment with two different soils was analysed. The parameters for CDE and SC types of transport were estimated for different soil layers using TDR with steady state flow. We performed an experiment with a multi-compartment sampler in the field. At the same location, two experiments were done, during snowmelt and with irrigation. Such an experiment has not been done before. The results of the experiment are discussed in a previous paper. In this paper, we extend the analysis by modelling. The results are used to estimate a random field for the soil hydraulic parameters. The experimental setup, the analysis of the experiment, and the modelling approach are considerably different from Seuntjens et al.

r 29: Can you confirm that the soil profile was unfrozen during the test because snowmelt often occurs on still frozen subsoil leading to low infiltration and ponding possibly followed by runoff.

No ice layer was observed under the snowcover, on top of the soil surface. At the time of solute application on top of the snowcover, the temperature in the soil was -0.4 C at 5 cm depth. As a high concentration of de-icing chemical was applied, any frozen soil water melted at this temperature. Thus, no ponding could occur on the soil surface of this coarse textured soil. Therefore, frozen soil does not contribute to heterogeneous flow. We added to the ms that no ice layer was observed.

13454

r 2: Please quantify, e.g. use coefficient of variation or standard deviation of concentrations

We added to the ms: "In the irrigation experiment, more isolated peaks in the concentration were found (spatial autocorrelation of 0.26 versus 0.61 during snowmelt),"

r 3: What was the reason for the lower exchange? Was this due to the higher saturation degree of the soil resulting in activation of preferential flow paths and reduced solute mixing. The irrigated soil was probably drier with more matrix transport providing more opportunity for mixing.

The main reason for less exchange, is the shorter residence time in the soil during the snowmelt experiment. Due to the shorter available time, less exchange can occur. We added this to the ms.

r 6: From this I deduce the experiments were on different soils, right? [Although after reading Schotanus et al 2012 it seems it's still the same setup - so how can there be differences in soil hydraulic properties?] If so it is needs mentioning what the major differences in hydraulic properties were. The level of detail provided is insufficient for the reader to draw some useful conclusion and take a message home.

Apparently, our text was not clear, as the reviewer got the wrong impression. The experiments were on the same location, with the sampler left in place in between the experiments. In this way, the effect of the infiltration rate on heterogeneous leaching can be studied while all other experimental conditions remain the same. The soil hydraulic properties only differ on a small scale. Roth (1995) analysed this type of soil heterogeneity with a model, with different infiltration rates.

We did not take many samples to determine the soil hydraulic properties on many samples, as is done by others. Instead, we performed a leaching experiment with a very high resolution (0.001 m2). By performing the experiment twice on exactly the same location, but with different infiltration rates, and analysing the type of heterogeneous flow, we deduced that the soil hydraulic properties vary on a small scale.

We have changed the text and added to the ms: "Two experiments were done, while the MCS remained in the same location"

r 20: please do a more complete literature review and provide a more comprehensive overview - there is much more similar work out there than this paper suggests

Indeed, an extensive literature exists on heterogeneous flow, which we could not include all in our paper. There are some excellent review papers. We included these review papers in our literature review to give a broad overview of the literature available.

13457

r 2-4: This is a very general statement which needs further details; how were the "effective" results defined? Mean field-scale BTC; variance of mean BTC?
With effective results, we mean the same spatial variability of drainage, throughout the entire experiment (such as Fig. 6). We added to the ms: "Instead, we were looking for a distribution of the soil hydraulic properties in a vertical plane, which will give the same spatial variability of drainage, throughout the entire experiment. The simulated drainage was sorted and cumulated, and then compared to the measured sorted and cumulated drainage. The degree of correspondence was determined with the mean absolute error.".

r 6: Explain; what was the mass recovery?
With uncomplete BTCs, we mean truncated BTCs, i.e. that the concentration did not decrease to a non-detectable level before the experiment was stopped. We changed this in the ms.
The mass recovery of the tracer was 42%. Similar values were found by Lissner et al, who used closed lysimeters from the same area. Probably, the tracer is stored in the profile, which leads to low mass recovery. We added the mass recovery to the ms.

r 10: how?
We added to the ms:
"For the saturated hydraulic conductivity and the pressure head, the standard deviation, and the correlation lengths in depth and width direction were determined with Hydrus-2D, using the spatial distribution of the drainage. The parameter set resulting in the smallest deviation between observed and calculated drainage was selected. As a criterion for the selection of the most appropriate random distribution for this soil, the total drainage per cell was used, sorted in a decreasing order, and then cumulated. Random fields with Miller-similarity were generated in Hydrus-2D. For details on the generation of these fields, we refer to the Hydrus manual (simunek,1999)."

r 16: what is the source?
We added to the ms: "Degradation was modelled as a first-order process, with a half-life time of 10 d. In the field, the half-life time of PG can be up to 17 d (french,2001), however, the degradation rate can vary in time, for instance as soil temperature or wetness vary."

r 23: remove 3
changed in ms

r 24: What was the source of this data? Different geographic regions? Or is it synthetic data?
The weather data is measured at the airport, where the experimental site is situated. For the year 1997, the snowmelt is removed from the measured data, thus only 1997 without snowmelt is synthetic data, and only during the snowmelt. This is added to the ms. We added to the ms: "To investigate the effect of the atmospheric flux on solute leaching, three different atmospheric input fluxes were used, measured at the weather station at the airport where the field station is situated."

13458
r 7-8: Not clear. Porosity of what? Snowcover?
Indeed, we mean porosity of the snowcover, we added this in the manuscript: "(i.e. porosity of the snowcover x depth of the snowcover, m)"

r 13-15: Does this mean it's synthetic data which you make up by subtracting infiltration from snowmelt using real data? It is difficult to understand how such data set with less infiltration (snowmelt is removed) can have higher soil moisture
r 17: how was this quantified? did you run hydrus with a warm-up period? If so what data was used for both cases? Show data on soil moisture content.
Yes, a warm-up period was used in Hydrus. For 1997 with snowmelt and 1997 without snowmelt, the same duration of the warm-up period was used (87 days). For 1997 with snowmelt, the solute application started at the beginning of the snowmelt. Just before snowmelt, the soil is dry, because

during winter there is no infiltration, as all precipitation is stored in the snowcover. During 1997 without snowmelt, there is infiltration, as precipitation is not stored in the snowcover. As a result, the soil moisture content is higher during 1997 without snowmelt. We added this explanation to the ms. Moreover, we changed Fig. 2. Now, different colours are used for the different parts of which the infiltration series consists (eg for the snowmelt part). Also this was added to the ms.

13459

r 3: the leached mass can be obtained from the zeroth time moment; why did you not use that? As a more general comment, mass balance data must be provided for each individual BTC and the mean BTC

We talk about mass passing a control plane, i.e. a flux, and this can be determined using Kreft and Zuber. We believe this mass that arrived is equal to the zeroth moment of the breakthrough curve. We added a reference to Kreft and Zuber (1978).

To include a mass balance for each individual BTC, the applied mass should be known for each cell. This is not known as the flow paths are not known exactly. Only the applied mass for all cells together is approximately known. Therefore, the mass balances for individual cells could not be included.

r 13: How does the total leaching of all cells relate to the applied solute mass M0 = C0 x t0 (zeroth time moment)? In other words, what was the mass balance.

We added to the ms: "In the snowmelt experiment, the bromide recovery was 43%. In the irrigation experiment, the bromide recovery was 42%, and the PG recovery was 32% (schotanus,2012)."

r 23, instant: i.e. here the cumulative drainage until the first detectable concentration

Yes, we added this to the ms.

r23, height: do you mean: magnitude?

Yes, we replaced height with magnitude in the ms

r 24, tailing: i.e. the degree of non-equilibrium (deviation from the Local Equilibrium Assumption)

Correct, this tailing is indicative of deviation of LEA.

r 24, contribute: when is a cell considered not to contribute to leaching? When zero drainage measured at a particular time (here: cumulative drainage)? Or another criterion?

We added to the ms: "In a cell or node that does not contribute to leaching, no drainage occurs, or the concentration is 0, both throughout the experiment."

13460

r 8:It would be useful to indicate the cumulative leaching associated with the applied tracer velocity v0 = infiltration/water content;

this way the cells exhibiting faster than v0 transport can be distinguished from those with slower than v0 transport

We added to the ms:" (in the snowmelt experiment)The CV in the drainage of the cells was 0.9. (…)This is supported by the CV in the drainage of the cells, which was 1.1 (in the irrigation experiment), higher than in the snowmelt experiment."

We did not include an analysis with v0 in the ms, because it is not possible to define v0, as water flow in our field experiments was transient. We also refer to our reply for 13461, line 4.

r 14, same time: Define - after how much cumulative drainage?

r 15, larger amount: define

r 16: define - how much wetter.

r 17, highest solute flux density: observed in any of the drainage celss (ok?)

r 18, lower: how much lower?

These sentences were made more clear, by making the qualitative statements quantitative.

r 19: discuss the possible effect of differences in applied cumulative drainage
r 28: Could sorption be an important process? If so, this could have retarded the arrival time which is then interpreted as cells not contributing to leaching, although it is more a matter of delayed response.
No, adsorption was not an important process. PG does not adsorb, adsorption was not included in the modeling either. We added this to the ms: "Adsorption was not considered to occur in the model (french,2001).".

13461
r 1, more homogeneous: This is not easy deduced from the leaching surfaces; a much simpler and quantitatively more appealing approach is to calculate the standard devation or CV of concentrations across the cells.
Fig 3 shows that 90% of the soil leaches bromide during the snowmelt experiment. Fig 4a show that about 75% of the soil leaches bromide during the irrigation experiment. Moreover, it is shown that in the irrigation experiment a smaller fraction leaches large amounts than in the snowmelt experiment. We added the CV in drainage and leached mass to the ms.

r 1, increasing: which cases do you compare and what is the magnitude of the increase?
We compare Fig 3 and Fig 4a. We added this to the ms.

r 4: Adding some basic statistics of the leaching behaviour would be very useful. eg add a frequency distribution of the effective velocities based on temporal moments; this would give a first indication if the heterogeneity is caused by (small-scale) textural differences or structural phenomena
In the previous paper we fitted transport parameters v and D to BTCs. The cells were sorted and divided in three groups: fast, average and slow responding cells. For each group, and the total, v and D were fitted using CXTFIT. For this new manuscript, we tried to fit a v and D for each cell separately to obtain a distribution for v and D. However, this was not successful. One of the reasons for this is that the applied mass should be known. This is approximately known for all cells together, but not for individual cells, as the flow paths are not known. Therefore, we were not able to include a distribution for v and D.

r 10: remove van
Ok

r 11: Why wasn't the drainage data used also in the fitting? Or alternatively, use the drainage data to validate the fitted soil hydraulic properties. This can be easily done.
We used the pressure head to fit the soil hydraulic properties, because water flow is driven by differences in potential. Thus, the drainage is a result of the potential, which we used. Moreover, this pressure was applied to the porous plates in the MCS, thus water collection in the MCS is also driven by this pressure head.

r 11: pressure heads
Ok

r 13: what is the 95% confidence limit?
We added to the ms: "The optimal value for alpha was 14.85 m-1 (95% confidence interval: 14.38--15.30) and for n 3.165 (-) (95% confidence interval: 3.09--3.24)."

r 16: independently? how?
We added to the ms: "The values of the other soil hydraulic parameters were the residual water content theta_r=0.045 m3 m-3 (from the category ``sand'', Carsel and Parish 1988),"

r 20: You probably mean: different sets of parameter values for {sigma, lambda} were selected in a trail-and-error way and applied manually in hydrus; the parameter set resulting in the smallest deviation between observed and calculated total drainage per cell was selected.

Indeed, selected is more appropriate than fitted in this case. We changes this in the ms into: "After fitting the soil physical parameters, the standard deviation and the correlation lengths in width x and depth z direction were varied manually for K_sat and the pressure head h. The parameter set resulting in the smallest deviation between observed and calculated drainage was selected."

r 23: This needs additional mathematical explanation: what was generated was a single realisation of Y = ln(scaling factor f);
h = h* exp (-sigma^2 Y)/ (f)
K= K* exp(-sigma^2)*
f^2)
We generated the fields for the scaling factors, with Miller similarity. We added to the ms:" Random fields with Miller-similarity were generated in Hydrus-2D. For details on the generation of these fields, we refer to the Hydrus manual (simunek,1999)."

r 26: how does this value relate to the statistics of the effective pore-water velocity? See comment above on determining basic statistics for Veff
We refer to our reply for 13461, line 4.

r 26: why this large difference in correlation length? Usually the spatial correlation is larger in the horizontal than in the vertical direction owing to stratification or soil layering. The high value for lambda_z suggest the soil is more similar in the vertical compared to the horizontal direction. What arguments do you have to substantiate this?
The correlation length of 0.6 m is only used to test the discretisation. For the other simulations we use a vertical correlation length of 0.15 m. We used the correlation length of 0.6 m to test the discretisation, because this is an extreme value.

13462
r 7: Your assumptions violates the the criterion of at least four nodes per correlation length suggested by Ababou [1988]. Furthermore, with a single realization approach, inference of statistical moments of the distribution of relevant variables requires a flow domain that is sufficiently large compared with the correlation scales of the pertinent formation properties (say at least 10 to 50 correlation lengths) . With a vertical correlation length of 0.6m, the vertical domain size should be at least 6 m, while it is only 1.5m . For the horizontal domain size at least 0.5 to 2.5 m would suffice. This is ok .
We are aware that the spatial discretisation should be smaller than the correlation length. Therefore, we investigated the effect of the discretisation. This is explained on page 13461 line 23 – page 13462 line 11. We found that the results were similar with a discretisation of 0.0125 and 0.025 m (see included Figure 1, below). We use this figure to demonstrate the results are similar, but we do not include this figure in the ms, as we think it is not informative enough for the ms. We include a comparison with Ababou 1989 in the ms.
As the MCS is located at a depth of 0.51 m, we are interested in solute leaching at this depth, also in the modelling. Therefore, we do not use a much larger domain. Moreover, we would like to emphasize again that the vertical correlation length was not 0.6 m, but 0.15 m. In this case, the domain is 10 times as large as the correlation length.
We checked whether a groundwater table at 1.5 m deep would influence leaching at 0.5 m, this was not the case.

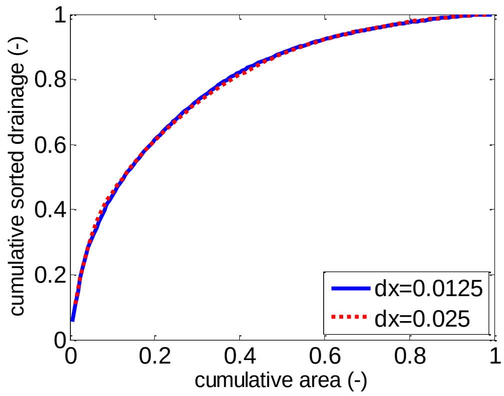

Figure 1. Sorted and cumulated drainage, with spatial discretisations of 0.0125 m and 0.025 m.

r 13-14: was there any combination for which the horizontal correlation was larger than the vertical? and if not, why?
Yes, such combinations were used too, but it appeared that these could not reproduce the spatial distribution as we measured in the field.

13463
r 22: The simulated maximum is 0.02 compared to 0.01 for the measurements. This is a factor of 2 difference? This seems considerable - and does seem to contradict your statement of "sufficiently well"
We included in the ms: "To compare the magnitudes of the solute flux densities, the solute flux density of the simulation should be corrected with 59/100, as the simulation consists of 59 cells, and the experiment of 100 cells. The maximum solute flux density is slightly larger for the simulation (0.010 and 0.012 d^{-1})."

13464
r 4: It would be useful to plot the scaling factors versus amount leached, and relate the scaling factor to texture (eg high values for coarse and low values for fine texture - right?)
Such a figure seems a good idea. For our analysis we actually made such a figure. However, we did not include it in the manuscript, as the amount of leaching not only depends on the scaling factor at a particular depth, but reflects an integrative scaling factor from that particular depth to the soil surface. Furthermore, the exact pathways of the solute or water is not known, and thus the integrative scaling factor is not known exactly either. Therefore, we did not include this figure, but only a sentence instead: "When the total leached amount of the cells in Fig.7 is related to the scaling factor in the particular cells at 0.5 m depth (Fig. 8), the leached amount generally decreases with an increasing scaling factor.".

r 5: calculate the ratio of flux to Ks; that is a useful criterion to evaluate if flow will be close to saturation (then coarse textured material governs flow) or rather highly unsaturated (then flow is governed by fine textured material).
We added to the ms: "The saturated hydraulic conductivity (French et al, 2001) is high compared to the water flux (respectively 57 m/d versus 0.016 m/d)."

13467
r 20-22: The degree of tailing is difficult to discern from fig 12 ; there are better ways to display the tailing of BTCs;
The maximum value in the legend is different for the tracer and degradable solute. Furthermore, in Fig 12, the infiltration rate is high (compared to fig 11). As a result, the differences in solute leaching for the tracer and degradable solute are much smaller, which is actually the purpose of Fig 12.

13468
r 2: Do you mean: solute mixing or dispersion in a medium composed of three types of sand is enhanced at higher flow rates?
Yes, we changed this sentence in the ms.

r 3: Do you mean as opposed to the spatially correlated approach - and why does a stream tube approach produce such results?
Opposed to a spatially correlated approach, when independent stream tubes are used, the effect of the infiltration rate on the spatial distribution of the solute leaching is ignored, as solutes cannot move laterally.

13475

0.5 0.05 0.15 0.054 0.020: This would have been a better choice as it does not invalidate the criterion of at least 10 correlation lengths.
We agree, therefore we used these parameters for the simulations.

13477
Improve comprehensiveness of graph by adding eg a horizontal arrow above data removed from the time series to produce case b.
We made the figure more clearly by adding colour to the figure. Now, the building blocks of the precipitation series can be distinguished in part a and b, because the blocks are coloured and a line was added.

13478
grey-scale does not have a legend
We changed the caption of this figure, now, the caption includes the dimensions.

This is basically a spatial picture of the mass recovery if the rescaling per cell is done with respect to the applied solute mass M0; the advantage of rescaling to M0 is that one gets an idea of the spatial variation in mass recovery.
We used the leached mass for scaling, the only difference is the maximum value in the colour bar. For the degradable solute, we scale with the leached mass of the tracer, which gives the mass recovery of the degradable solute compared with the tracer. See also our earlier response that the applied mass per cell is not known.