Review of "Does drought alter hydrological functions in forest soils? An infiltration experiment" by Gimbel, K.F., Puhlmann, H., and Weiler, M.

This manuscript addresses an interesting and important premise, which is what happens to the infiltration capacity of soils under prolonged drought. This requires monitoring soil properties over longer periods compared to the time frame of more common experiments imposing short but intense dryness. While the experiments seem well executed, I found the presentation to be lacking and the interpretation of the results to be problematic. In some places the conclusions do not follow directly from the results presented. I outline some major issues below.

Major comments:

1. Description of the dye pattern analysis (section 2.4): this section is tailored for those already familiar with dye pattern analysis. Otherwise, it is difficult to understand the reason the 3 metrics (volume density, surface density, and stained path width) are selected for characterizing flow patterns within the soil column. In addition to referring to previous literature that adopts these metrics, I think the authors should include more descriptions for the advantages of using these metrics and how they relate to physical processes in the soil. In addition, results pertaining to surface density is not presented anywhere in the results section. How does information from surface density complement that from volume density?

2. Soil moisture changes (Section 3.1): There are 2 lines delineating the dates of experiments in 2011 and 2013, but the period over which the simulation has been conducted is never indicated. Which year was this? In Figure 4, what does the green line signify? This needs to be explained in the legends. Throughout the manuscript, the authors use qualitative words to describe quantifiable results, such as in page 7698, line 5, "before the experiment in 2013, the soil moisture status of the drought treated plots and the control plots are *very similar*." I found this to be vague and misleading, and furthermore inadequate to support the main result from this section, which is that "the observed infiltration patterns and changes among the sites are mainly a result of change in soil properties [due to drought]." In fact, differences in trajectories between drought plots and control plots can be observed even *prior* to the start of experimentation (e.g., coniferous forest in Schwabische Alb, deciduous forest in Schorfheide-Chorin), sometime to the same extent observed *after* rainfall-exclusion. The authors need to address these differences, using statistical evidence if possible. In general, this section needs to be overhauled and written to highlight the connection between the main points. Also, if soil measurements have been taken, why not show them on the plots?

3. Interpretation of results: As mentioned before, the authors have a tendency to make broad stroke generalizations on the results that should otherwise be addressed with more nuance to accommodate for other explanations. On page 7698, line 20, "In general, coniferous plots under drought had higher WDPTs than deciduous plots." By looking at Figure 5 it is clear that this is the case for 2 out of 3 sites and far from a general observation. This tendency continues throughout the manuscript: "in general, the patterns of the control profiles are *similar* in vd, SPW… (page 7699, line 24)" and "The comparison… showed no differences which can be addressed to other reasons than small scale heterogeneities of soil properties… All control plot profiles can be assumed to be comparable to the pre-drought plot profiles (page 7703, line

23).” I would dispute the accuracy of those statements. This becomes extremely disconcerting in Section 3.3, when the authors dismiss “time dependent changes of the soil characteristics” by equating control profiles to pre-drought profiles and attributes observed differences between pre- and post-drought profiles to the effects of rainfall-exclusion. However, by looking at Figure 6 and 7, it is not apparent to me the degree that the differences can be attributed to either the pre-drought and control pair or pre- and post-drought pair. In some cases, the difference between post-drought and control profiles seem much less than pre-drought and control profiles (Ashwabische Alb, coniferous, 60mm), which would invalidate the authors' premise. These differences are brushed aside, which to me raises red flags about the validity of the ensuing arguments. The authors should strive to clarify this section a bit more. It would help, for example, to reorganize Figures 6 and 7 to highlight the similarity and differences between the 3 classes of observations (control, pre, post) and include the flow processes bands in Figure 6.

Clarity issues and other comments:

1. Page 7690, line 15: “WDPT tests” This is the first time this acronym appears in the paper and needs to be written out.
2. Page 7691, line 15: “these shrinkage cracks foster bypassing of the soil matrix” This is done through preferential flows? The sentence as it stands now does not make much sense and needs to be expanded.
3. Page 7692, line 13: “in respect to the expectable behavior” needs to be changed to “expected behavior”
4. Page 7692, line 15, “avoiding tentativeness due to an overreaction to…” needs to be rephrased.
5. Page 7692, line 19, “because they reflect integrally…” needs to be changed to something like “they reflect the integrated changes in soil hydrological functions…”
6. Page 7694, line 8, “The incoming precipitation was reduced… to the level equivalent to an annual drought with a return period of 40 years” This would imply different levels of reduction for each of the sites. The basis for this choice was puzzling to me. The authors clearly points to projected climate change with increasing dryness in Europe (Page 7691, line 10) and thus a drought level with a return period of 40 years calculated using historical data would contain little meaning when applied to future, nonstationary conditions. In theory 40-year droughts would become increasingly likely in the future, but the frequency with which it happens would depend on each site. What is the advantage of using this instead of a uniform reduction across each site? Additionally, the actual amount that was reduced should have been listed somewhere in the paper.
7. Page 7694, line 17: “experimental area was kept shaded and sheltered” how does shaded differ from sheltered?
8. Page 7699, line 12: “By comparing the pre-drought pattern and the pattern for the control plots…” This sentence is convoluted and needs to be rephrased for clarity.
9. In general the paper needs to be rewritten with an eye on clarity of the sentences and the organization of the paragraphs (to emphasize a few main points).
10. Figure 6: The black and grey regions are not properly explained. They indicate the vd of stones but what differentiates between them?