Responses to the Comments by Reviewers

We thank Professor Graham Fogg and an anonymous reviewer for their constructive comments. The manuscript has been significantly improved by addressing the comments. The following is our point-to-point responses to their comments.

Responses to the Comments from Reviewer #1

General comments

This work proposes an extensive review on the evaluation of Deep Soil Recharge (DSR) in arid and semi-arid regions. The validity of Annual Recharge Coefficient is questioned based on the test data of DSR measured by the improved lysimeter. The issue raised by authors is of great significance, but the section of data analysis (part 3.2) is not well structured. There is a lack of explanations on the new apparatus and details on the field test are not clear enough as well. Experiment data presented by the authors is limited to support the conclusions. Thus, the manuscript requires significant improvement before it is accepted for publication.

Response: Thank you for the positive comment. We give a detail explain of the new lysimeter at part 2.2, page 9, and descript field testing of the new apparatus at part 3.1, page 11. And we have reorganized the section of data analysis (part 3.2) by improving the explanations on the new apparatus and details on the field test (line 258-270). Experimental data presented are as complete as possible. The conclusions drawn are based on a careful analysis of the complete set of data. We believe this revised manuscript meets the requirement of HESS and is now ready for publication.

Specific comments:

1. Some of the conclusions are not supported by the test data. For example, the authors claim that "The temperature influences the DSR rate" (in line 343). The evaporation intensity varies with temperature and affects the quantity of DSR indeed. However, if temperature influence is considered, the delay time is an issue remains to be discussed.

Response: Thank you for your careful observation, we compare two precipitation events which are similar in strength (17.2 mm for April 4 and 16.8 mm for October 5) but different in the DSR delay time (36 days for April 4 and 16 days for October 5) in line 354. Temperature is the most likely factor for such a delay time. This is the primary reason for above sentences. So we draw a conclusion that temperature affects the DSR rate. However, finding out exactly how the temperature affects the DSR rate will require additional field experiments that should be pursued in the future. See page 20 line 354-362 for details.

2. Authors claim that "recharge is somewhat positively correlated with a few strong precipitation events (greater than 10mm), and very closely correlated with the strongest precipitation event" (line 426-428). However, the data in table 3 seems do not support this conclusion.

Response: In Table 3 as showed below, we showed inter-annual statistics of strong precipitation and its percentage in total annual precipitation amount. Comparing the maximum precipitation events in 2013 to 2015, we can conclude that recharge is somewhat positively correlated with a few strong precipitation events, such as the 32mm, 15mm, and 17.2mm maximum precipitation events in 2013, 2014, and 2015, respectively. Such a positive correlation is particularly strong for 2013 which has the largest maximum precipitation event of 32mm. This positive correlation is weaker for 2014 and 2015 which have moderate and somewhat similar maximum precipitation events (15mm and 172.2mm, respectively). For these two years, other factors such as rainfall temporal distribution may also be a concern. Because of this, we stated on page 24, line412-416 that "but the determination of the threshold for strong precipitation events that directly contribute to DSR is still unclear and requires further investigation."

Year	Number of strong precipitation	Maximum precipitation event (mm)	Annual DSR (mm)	Annual DSR /annual precipitation (%)
2013	2	32	20.2	24.33
2014	4	15	20.6	10
2015	6	17.2	9.2	4.94

3. The quantity of DSR is actually given by the mass balance of surface layer. Surface runoff, evaporation and transpiration are critical components of water balance besides precipitation. It is necessary to present more monitored data, especially about evaporation and surface runoff, to support the conclusions in the paper.

Response: This is a nice comment. Firstly, there is no runoff at the studied area which is essential desert and easy to penetrate. Secondly, as stated correctly by this reviewer, the basic idea of lysimeter is water balance. So if the point of measurement is relatively shallow, one must consider evaporation and transpiration process. However, the DSR measurement reported in this study is NOT at relatively shallow depth, instead, it is specifically at a sufficiently deep location (2 m) is to make sure that evaporation and transpiration are both negligible. In another word, the downward DSR measured at such a deep depth is regarded as completely recharging the underneath groundwater aquifer. Please see page 7 line 150-161 for details.

4. Previous studies have shown that Annual Recharge Coefficient varies with the water table depth. To avoid the influence of water table depth, the dynamic of phreatic water table from 2013 to 2015 in the study area is suggested to be presented in the paper.

Response: A nice suggestion. We have added the dynamic phreatic water table

position in 2013-2015 in the revised paper (see page 7 line 154). It is found that the water table depth is greater than 4m in 2013-2015, so its influence to DSR is negligible.

5. In Figure 1(B), how to measure the flux at depth A? More details about the new lysimeter are required.

Response: Thank you for your careful observation, there is a rain gauge at depth B and the column between depth A and depth B is at a balance stage so flux at depth B is the same as at depth A. We have revised the caption for this figure.

6. Precipitation events are suggested to be presented by using columns (or vertical lines) in Figures 3, 4 and 5. 7. English should be improved because the text is somewhat difficult to follow.

Response: Implemented. The English has been thoroughly checked by a native English speaker.

Responses to the Comments from Professor Neuman (Reviewer #2)

I generally like this article as it is based on three years of observations and in groundwater hydrology, specifically recharge, there is no substitute for observations are these are far and few.

Response: Thank you for the positive comment.

Specific comments:
(1) The authors seem to knock on annual recharge coefficient as well as models and state neither would work. I concur with the first one (annual recharge coefficient) but I am not so sure that you can make the same statement on models. Most models are complete depictions of hydrological cycle and if done correctly (implying that all the components of the water balance are correct), then recharge should be accurate

Response: This is a nice comment. We have revised the text to only concentrate on questioning the annual recharge coefficient, but the models. As corrected stated by this author, if the model is established properly, recharge should be accurately estimated. See page 4 line 80-94.

(2) I find the figures 3, 4 and 5 very interesting. However, there is a large component of the infiltrating water that evaporates and if that is not subtracted from the rainfall you cannot estimate the recharge. In fact you cannot just compare 2013 to 2014 to 2015 without accounting for evaporation of the infiltrating water in the inter-storm periods. I think your

observation that recharge is dictated by high intensity rainfall is correct; during high intensity (and long duration rainfall) the saturation of the soil profile hastens recharge and decreases evaporation (due to lesser atmospheric demand especially if it is raining!).

Response: This is a nice comment. Firstly, there is no runoff at the studied area which is essential desert and easy to penetrate. Secondly, as stated correctly by this reviewer, the basic idea of lysimeter is water balance. So if the point of measurement is relatively shallow, one must consider evaporation and transpiration process. However, the DSR measurement reported in this study is NOT at relatively shallow depth, instead, it is specifically at a sufficiently deep location (2 m) is to make sure that evaporation and transpiration are both negligible. In another word, the downward DSR measured at such a deep depth is regarded as completely recharging the underneath groundwater aquifer.

(3) A better analysis of length of the storm, atmospheric evaporation demand (should be very easy to calculate) should help in estimating recharge (with a simple model as compared to SWAT or HYDRUS). This will in fact justify your hypothesis that recharge is dependent on a few high intensity events.

Response: This is an interesting suggestion and certainly will be pursued in a future study to justify the hypothesis that recharge is dependent on a few high intensity events. The purpose of this study, which represents a first step in such an endeavor, is to provide direct field evidences to question the concept of annual recharge coefficient. A complete modeling of the storm, atmospheric evaporation demand will be pursued elsewhere.

We revised all the figures.

• Line 80-82: Modeling is an efficiency way to test different hypothetical scenarios and it may be used to predict DSR in the future if the model is calibrated carefully

Response: Implemented. Revised.

• Line 178-179: At the soil surface there is a device to measure the amount of the precipitation and

Response: Implemented. Revised.

• Line 178-179: At the soil surface there is a device to measure the amount of the precipitation and

Response: Implemented. Revised.

• Line 259-271: In September 1, 2012, mobile sand dune within the study site was set as the monitoring plot….

Response: Implemented. Revised.

• Line 333: That is one reason

Response: Implemented. Revised.

• Line 339: This is the other reason why precipitation in 2014 (205.6 mm) is greater than 2013 (83 mm) but the overall DSR in 2014 is less than that in 2013.

Response: Implemented. Revised.

• Line 361-362: Comparing two precipitation events which are similar in strength but different in the DSR delay time, temperature is the most likely factor for such a delay time

Response: Implemented. Revised.

• Line 407-415: Such a positive correlation is particularly strong for 2013 which has the largest maximum precipitation…..

Response: Implemented. Revised.