Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-329-RC2, 2018

Interactive comment on "Field-based groundwater recharge and leakage estimations in a semi-arid Eastern Mediterranean karst catchment, Wadi Natuf, West Bank" *by* Clemens Messerschmid et al.

Anonymous Referee #1

Received and published: 4 July 2018

Answers by: C. Messerschmid, J. Lange and M. Sauter (21 October 2018)

This paper deals with groundwater recharge estimations for a semi-arid environment. The author Review of 'Field-based groundwater recharge and leakage estimations in a semi-arid Eastern Mediterranean karst catchment, Wadi Natuf, West Bank' by C. Messerschmid et al.'

General comments

Overall, the paper is remarkably chaotic. Points are being made based on material that is presented later,

We thank the reviewer for his/her extensive time and effort, put into the review of our manuscript. We diligently went through our manuscript based on all the comments of the three reviewers and understood that there might be some confusion regarding the methodological approach as well as the terminology, possibly attributable to our different backgrounds as consultants, hydrologists and hydrogeologists. As already addressed in our comments to review 3, we decided to restructure the manuscript to improve legibility and include a schematic diagram illustrating our understanding of the hydrogeological functioning of the system, the methodological approach as well as defining the individuals scientific terms employed in the study.

part of the mathematical relationships are presented verbally in the text instead of in equations,

Thank you for the comment; we shall express the relationships as equations

and the notation of units is at times inconsistent (see the detailed comments in the text to find many of these).

Thank you for the comment; we shall check units for consistency

Furthermore, text fragments appear in the wrong sections throughout the paper.

Thank you for the comment; we shall revise the structure and check for fluency of the text, grammar and semantics.

Introductory material is introduced in the conclusions, elements of the discussion appear in a figure caption, and methodological aspects are scattered throughout the paper.

Thank you for the comment; we shall restructure the manuscript (see our answer above), especially the introduction and methodology.

The only section of the paper that is well written is the geological description of the research area.

Thank you for the comment; we shall revise the manuscript thoroughly.

Many sections of the text are carelessly written:
- paragraphs are not well organized, evidenced by repetitions and discussions of related observations occurring at different locations within a paragraph

Thank you for the comment; we shall revise the structure of the manuscript thoroughly and avoid all unnecessary repetitions.

- terminology is used in way that make sense neither grammatically nor semantically, rendering the text incomprehensible

Thank you for the comment; we shall revise the manuscript and its language thoroughly.

- the tables are not numbered in order of reference

Thank you for the comment; we shall revise the format of the manuscript – also with regards to table captions and the order of tables

contain much too little information to allow the tables to be read independently of the text (or even after consulting the text)

Thank you for the comment; we shall revise the table captions thoroughly

- entire table columns are left unexplained

Thank you for the comment; we shall revise and expand longer captions thoroughly

The consequences of the poor structure of the paper are severe. The authors claim to have developed a new method to estimate groundwater recharge but even a very patient reader willing to go through the manuscript multiple times would be unlikely to be able to reproduce the approach.

Thank you very much for this important comment. As explained above, we shall restructure the manuscript; include explanatory diagrams and descriptions so that the importance and novelty of our approach and results become clear.

I suspect key elements were simply not reported because many technical details were omitted from the methodology section and could not be found elsewhere even though almost every section of the paper contains methodological elements.

Thank you for the comment; we shall revise the manuscript thoroughly and in particular expand the methodology.

This lack of completeness is aggravated by the liberal use of strong assumptions

Thank you for this helpful comment; we shall indeed reassess our assumptions thoroughly and revise them where necessary.

that are never made explicit, critically discussed, or tested (see the comments in the text for examples).

Thank you also for this comment. This point was also raised in review 2. We shall make our assumptions more explicit and/or explain them in more detail where necessary. We shall also statistical proof and critical discussion (see also review 2).

This hits at the heart of the science in this paper and cannot be remedied by a thorough revision. In setting up the project and the modelling strategy, here seems to have been a lack of critical thinking.

We thank the reviewer for still investing a major amount of time despite the shortcomings in the legibility. We believe that the criticism is a result of the lack of clear structure in the introduction and presentation of the methodology. We shall however reassess our general approach and modelling strategy and check it for gaps and errors.

Statistical jargon is used copiously but incorrectly.

Thank you for pointing this out. As mentioned already in our answers to review 2 and 3, we shall add statistical evidence. In addition we shall improve the relevant passages and we believe that it can be remedied so that our result will be presentable in HESS.

Any statistical analysis is missing, yet bold claims are being made about correlations, representativeness of a seven-year data set, and the inclusion of spatial variation within the research area even though only 8 monitoring sites were available.

Thank you for this comment, some of which was also raised by the other reviewers. Yes, you are correct. We somewhat sloppily dealt with statistics here. We shall take care of this in the revised version of our manuscript. Two sets of such statistical quantification were already added to our answer on review 2 – this concerns the representativeness of

a) the 5 daily read springs and b) the seven-year period and long-term rainfall variation. We shall add additional statistical material to this answer as detailed below and of course revise our manuscript accordingly. Some remarks on the spatial variation, representativeness and correlation are already included in our answers to review 2.

The contrast between boldness of the claims and the flimsiness of the evidence that supposedly backs them up casts serious doubt on the credibility of the science.

Please see our above answers. We shall hope that our revised manuscript and additional material will be satisfactory.

The HESS formatting guidelines are frequently ignored, for example by including footnotes, underlining parts of the text, placing equation numbers in front of the equations, etc.

Thank you for the comment. Yes, you are correct – we did not fully adhere to the HESS format in our draft version. We shall take care of this thoroughly for the resubmission of the revised manuscript.

HESS does not wish to publish regional studies.

Thank you for the comment. Yes, you are correct and we are certainly aware of this. We believe however that at least some of our results can be generalised. Nevertheless, please allow us to raise your attention to the fact that this article was written for a HESS special issue on the Dead Sea and environs, which addresses regional research problems (in most, if not all of the already submitted manuscripts for other articles).

The paper mentions that the method developed here is applicable to other areas, but does not indicate how, nor does it devote any section of the Methodology or the Discussion to the generalization of the results. The Conclusions repeat the claim of generalizability, but limit it to the Mediterranean region.

Thank you for this additional explanation. You are correct; and in our answers to review 3 we already stated that we shall include an additional chapter Discussion, which will shed light on this issue and in which we shall demonstrate that the results can be generalised to a larger area. Please allow us to already remark already in general: We are certain that our approach can be repeated in other areas of the Dead Sea area and its environs (relevant to the special issue of HESS, see above). Moreover that, we believe that if the results can be generalised to the Mediterranean climatic region, it is fairly valuable since there are a number of regions internationally that are comparable with respect to climate to the Mediterranean (e.g. parts of California, etc.).

This is followed by a list of conditions that need to be met before the method can be used. That list is so demanding that it serves as proof that the method can hardly be applied anywhere else.

As we said above, we shall address this issue. In our revised Discussion and Conclusions we shall discuss these issues in more detail and separately for each condition. However, please allow us to remark here already: It is not necessary that each and every condition is fully met in order to be useful for other researchers. Some of the incomplete or missing conditions can be addressed by additional field or other work. We are fully aware that in nature, no area resembles another one to 100%. And anyway, as reviewer 2 remarked: "Nobody is able to determine 100% correct values for each point in space". So, we shall aim at reaching a valuable compromise. Accordingly, future researchers may employ not the complete set of our methods but instead pick and choose certain ones that are applicable or deemed useful. To this end, we shall weigh the different conditions as to their importance. For example, if the catchment size cannot be determined with certainty and accuracy, a future researcher may choose not to follow our leakage estimation approach. Or if a greater variability of soil conditions is encountered in other terrains, then the need for additional field work or soil moisture stations may arise. Similar modifications of our approach may apply to the questions of representativeness of daily springs or of the measurement period, always depending on the exact individual problem encountered at the new study area…

Figure 4 seems to indicate measurement issues. When it rains after a long, dry period it stands to reason that the observed soil moisture storage in equivalent water layer cannot rise above the accumulated rainfall.

Thank you very much for this valuable and important comment. This touches on an important issue because some of our readings are indeed erroneous (most probably due to malfunctioning of the sensors/loggers installed). You are certainly correct and we are most grateful for your diligent, thorough reading: Soil moisture cannot rise above accumulated rainfall. We shall discuss this in the following.

If we look at Nov 2006, we see several showers before storage peaks at about 108 mm. It is not easy to see but I estimate the total amount of rain that generated that peak to be less than 50 mm. May 2008 is even more pronounced. This can only happen if recharge is captured from a large area and concentrated in the soil over a much smaller area.

Thank you for this comment. We checked our records and found the following:

On 4 Nov 2006, at station RK-W, soil moisture recordings raised to 69.2 mm, whereas accumulated rainfall had reached only 59.1 mm. And on 6 Nov 2006 a peak of 107.1 mm of apparent soil moisture was reached, although only 17 mm of additional rain had fallen meanwhile.

And on 17 March 2008, our apparent soil moisture raised from 55.3 mm to 87.7 mm (on 17 April 2008) with hardly any rain falling!

Now, as you correctly remark, this could only occur if rain (or other flows) were transferred from a wider area to the spot of measurements. However, this is most certainly not the case, as listed below:

This would require

1) considerable lateral flow before infiltration (overland flow)

1st mechanism: Overland flow wetting the soil. a) Station RK-W is situated on a naturally terraced hillside almost on the top of the summit. Hardly any water can flow down onto this site from above. b) The eroded rock banks form an irregular pattern natural “terraces” of 3-4 m width and between 4 and 11 m length. The particular terrace is rather small in area (3x4 m only) with pockets of relatively shallow soil (40 cm) surrounded by flat outcrops of rock. No inflow of runoff amasses from the surrounding rock outcrops within the terrace. c) Rainfall in early November 2006 was negligible, certainly way below the threshold that triggers runoff (evidenced also by the nearby instrumented runoff measurement station “Ein Ayoub”, which recorded no runoff event). The SM rise in March 2008 occurred during a dry spell (less than 1mm rain on 30/31 March 2008).

or 2) rapid infiltration through cracks to the groundwater, followed by lateral flow in the groundwater. For this lateral flow to converge towards a much smaller area there must be some kind of depression in the impermeable layer below the aquifer.

2nd mechanism: Lateral groundwater inflow accumulating over a depression in the underlying aquitard. a) The soil moisture station lies not only at (almost) the top of the hill but is also underlain by a 160 m thick series of karstified dolomitic limestone (Hebron Formation) of the regional Upper Aquifer with a thick unsaturated zone. The main section is massively banked; the lower section is more thinly layered. The bedding is sub-horizontal with a slight dip towards W and WSW. No folding is recorded laterally and in the subcrop (Walther’s law). Hence, there is no (topographic or structural) depression that could fill up with underground water. b) Groundwater, once percolated, moves rapidly through the karstified unsaturated zone down to the deep water table. So, even if there was a depression formed at the bottom of the aquifer, no filling up to the surface near soil horizons is imaginable.

The water moving toward that depression can only wet up the soil from below if the groundwater table is shallow enough.

Out of the above follows: Water tables in Hebron formation are very deep (usually more than 100 m below ground level). No shallow perched groundwater is found in this uniform and massively bedded karts formation of monotonous limestone/dolomite composition.

The authors do not mention any of the factors supporting mechanism 2, which is improbable anyway.

This is correct. Mechanism 2 does not apply.

But overland flows supporting hypothesis 1) were not mentioned either although the simple presence of the wadi indicates that overland flows do occur.

The Wadi at Ein Ayoub runs deep below this hillside. Ein Ayoub was equipped with a runoff measurement station. No runoff occurred in this period.

In the unlikely case that flow according to mechanism 1 or 2 wetted up the soil in a fraction of the area only, the normalization to mm should not have been applied, since the various variables presented in the graph and the overall analysis represent different areas. The dimension of choice should then be volume.

This is correct; normalization to mm is inappropriate, or would be if one of the mechanisms applied at least partially. We also checked the other two locations with misreadings – Kufr Fidiah (KF-W, low-UBK Formation) and Beitillu (BET, up-UBK Formation). KF-W is located in an open field, currently not cultivated and surrounded by walls. No signs of any runoff (throughout the entire observation period) were found here. BET is a private garden at the bottom of a hill, but again, surrounded by walls that inhibit overland flow reaching the SM-spot. In both cases, also mechanism 2, the bottom-up accumulation of groundwater into the soil column can be excluded with certainty.

In case that neither mechanism applies, the authors need to explain how the difference between soil storage change and rainfall can be positive. If they cannot, this must be a measurement error. The size of this error throughout the observation period and its effect on the recharge estimation error need to be examined then.

Thank you for this intensive investigation. Again, we certainly appreciate the diligence, with which you analysed our paper. Unfortunately, and as stated above, the misreadings are definitely a consequence of instrument failure (malfunctioning). We are grateful that this error has been pointed out and we shall certainly address it in our revised manuscript.

Thus, additional analysis on the size of the error shall be added to the paper. Attached to this answer we shall send a short summary of our findings, the result of which is:

Out of the 8 stations, three stations had measurement errors (WZ-uT, KF-W and RK-W mentioned above). In the three stations SM-readings above accumulated rainfall occurred during 6 days, 42 days and 55 days, respectively (in sum: 99 days). In addition, one station (RK-W in March/April 2008) had rising soil moisture in spring, during and despite a dry spell – during 44 days (together with the above, 147 days).

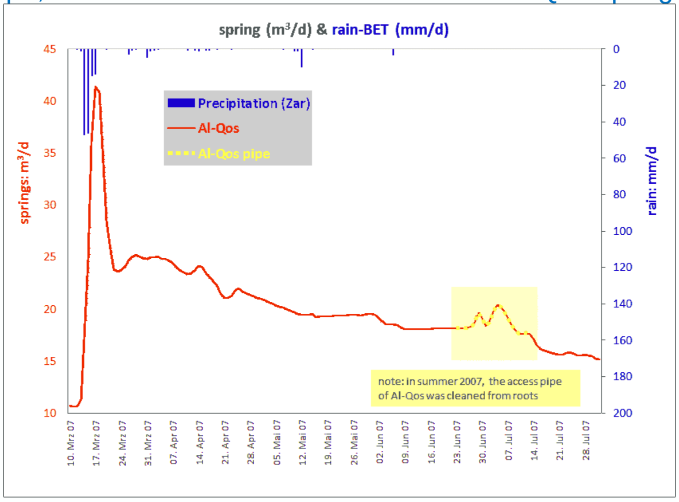
We recorded soil moisture during 1,818 days (spread over 8 stations). We modelled 8 stations over a period of 7 years (= 2,557 days), or for all stations together: 20,456 days.

The days with erroneous readings thus constitute 8.1% of the read-out days and 0.7% of the modelled days, respectively.

The caption of Figure 5 mentions the cleaning of one spring and the effect on its flow rate. The condition of the springs apparently affects their discharge. So, evidently this also affects groundwater recharge (see Eq. 5). Therefore, spring maintenance is an important factor for groundwater recharge, yet it is brought up only in a figure caption.

Thank you for this comment. You are correct. We created a misunderstanding, because we included a wrong formatting of Figure 5. This figure should only present the period from Oct-2006 until Aug-2009. (Before Dec-2006, we had large gaps in the reading of the 3 springs of Beitillu spring group. The formatting falsely indicated a continuous line with a linear trend of decreasing discharge for Al-Qos spring during summer 2006. This impression is an artefact of erroneous formatting.)

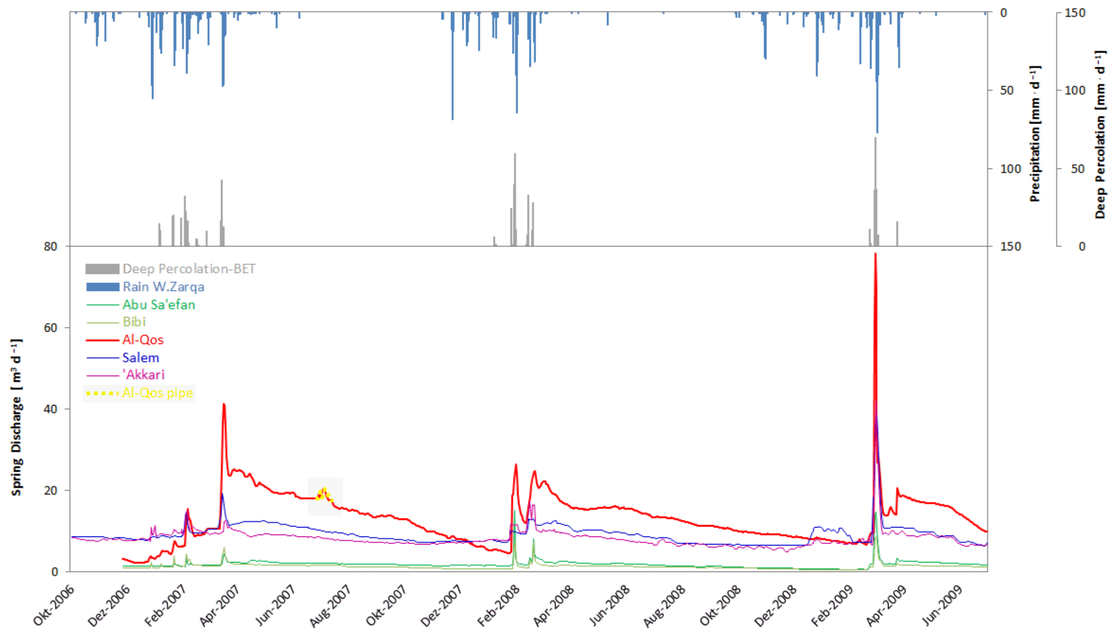
In addition, the short period of raised spring flow in Al-Qos spring (from 23 June until 12 July 2007) was not clearly discernible in this graph. We here add for your information a graph, zoomed-in on the short summer rise in Qos spring flow (see figure below):



We shall therefore correct the formatting of Figure 5 for the new manuscript.

The caption calls the effect temporary but is appears to double the well production for at least two years and the effect was still very strong at the end of the observation period so there is no clear picture of the persistence of the effect. It seems obvious that the net recharge to the aquifer feeding that particular well changed in 2007 with effects that last well over three years. I have the impression that the authors did not consider any of this in their calculations, or carried out a scenario study to examine the effect of different well maintenance practices.

Thank you for this comment. As stated above, this is a misunderstanding. We are sorry to have caused this confusion. A preliminary corrected layout of the corrected Figure 5 is shown hereunder.



It is obvious that only a brief rise for several weeks was observed in Al-Qos spring. We believe that this makes an answer to your interpretation obsolete.

Please also note the supplement to this comment:
https://www.hydrol-earth-syst-sci-discuss.net/hess-2018-329/hess-2018-329-RC1-supplement.pdf

Comments - hess-2018-329-RC1-supplement.pdf

Page 1

No.# 1 Please ident the first lines of the paragrpahs or separate them by blank lines for clarity.

Thank you for the comment. We shall see to that.

Page 2

No.# 1. Only if you know that the West Bank refers to the Jordan river do you know where the aquifer is.

Thank you for the comment. The West Bank actually refers to the occupied Palestinian territories (oPt) and is a well-defined political term (such as the State of Israel). (We trust that the readers of the HESS special issue on Dead Sea and environs will see the context). However, we shall revise our manuscript accordingly and point out that the area of the West Bank more or less overlaps with the recharge area of the Western Mountain Aquifer karst aquifer basin.

No.# 2. What does that mean?

Thank you for the comment. We shall see to another formulation: *Spatially, recharge varies strongly*… or: *Spatially, recharge is highly variable*…

No.# 3. One can pay attention but no emphasis, I believe.

Thank you for the comment. We shall see to that. You are of course correct.

No.# 4. Why does a leaky aquitard lead to a well-defined catchment area?

Thank you for the comment. We shall revise the manuscript and highlight the fact that the outcrops of the bottom aquitard surround the hill. This allows us to clearly define the groundwater recharge area.

No.# 5. What is a sub-aquifer?

Thank you for the comment.

"Sub-aquifer" is a term often used by Palestinian and Israeli scholars. Unlike the local perched aquifers, the regional aquifers actually comprise of 'aquifer systems' or 'aquifer complexes' that can be sub-divided into several "sub-aquifers". The term is usually attributed to individual formations within the (regional) aquifer complex and describes individual aquifer sections, according to their differences in lithology, recharge characteristics, conductivity, etc. For example, the Upper Aquifer in most regions is an aquifer system that comprises of the Hebron, Bethlehem and Jerusalem Formations as individual sub-aquifers…

As already noted in our answers to the other reviews, we shall properly introduce and define the terms we use. We shall see to that.

No.# 6. This does not belong in an abstract because it doe not help the reader determine if the paper is worthwile to read.

Thank you for the comment. We shall see to that.

No.# 7. These are not keywords, they are phrases.

Thank you for the comment. We shall see to that and break up the key words into smaller individual parts, such as karst aquifer, recharge assessment, Mediterranean climate, spatial distribution of recharge, aquitard leakage.

Page 3

No.# 1 I continuously struggle to find a coherent line of thought in the Introduction. It should be simple: what overall problem does the paper address? What work has been done in this field? What still needs to be done? What are we going to do? How does this close a gap in our knowledge or address a relevant problem? How does it contribute to the existing body of work?

None of these questions receives a clear answer, although I suspect that the authors know very well what they are doing and why. The last paragraph more or less manages to create an objective, but seems to be dis- connected from much of the Intro. In short: the Introduction needs to be thoroughly rewritten in order to

1) clearly establish the research problem in a larger (societal) context,
2) provide an overview of the current state of knowledge regarding the research problem,
3) identify from that a need for additional research
4) explain how the paper will address that need,
5) culminating in at least one objective of the paper.

Thank you for this helpful comment. As mentioned in the other reviews and above, we shall thoroughly restructure the entire manuscript and actually rewrite the Chapters 1. Introduction and 3. Methodology. Here we will stick to the context, state of the art, research question and objectives.

No.# 2 Around here you should give the location of the aquifer (by referring to a map for instance). If we know wehre it is it is easier to grasp its significance for the local populations.

Thank you for the comment. We shall refer to the location and set a reference to the figure of the location area.

No.# 3 Underlining is note permitted according to the HESS format I believe.

Thank you for the comment. We shall see to that.

No.# 4 differentiadedly

Thank you for the comment. We shall correct the wording.

No.# 5 This is very difficult to estimate for very deep unsaturated zones. And what about the travel time of the water in the unsaturated zone?

Thank you for this comment. We are not sure what you mean here: Do you mean soil moisture balances are difficult to establish over deep unsaturated zones? Well, this is the logic of direct procedures. Anyway, only the regional aquifers have deep unsaturated zones, the perched ones are very shallow and the thickness of the unsaturated zone is tens of metres maximum, often only metres deep.

Travel times are extremely brief (measured in Shibteen well in Lower Wadi Natuf, follow-up paper) and also stated by Schmidt et al. (2014) even for the deep, yet karstified regional aquifers (often hours, rather than days), even more so in the shallow perched aquifers: Here, spring response was observed well within a day's time.

No.# 7 If you want to estimate aquifer recharge you need both, so this would be 'and'.

Thank you for this comment. Ideally yes, both parameters, basin outflow and storage change are taken into account, independently. However, many integrated modelling studies, particularly in the WAB have worked with spring (and well) outflows only.

In this sentence we refer to Dörhöfer & Jesopait (1997), and they unequivocally stated '*or*', not '*and*', however not without adding that according to Bredenkamp et al. (1995) a reliable recharge estimation should always bank on employing several methods simultaneously in parallel and independently from each other… This is what we try to investigate in our introduction when weighing the options available in Wadi Natuf.

No.# 8 Before this point, the discussion is general. After it the text specifically targets the WAB. A separation into different paragraphs is needed here.

Thank you for the comment. Yes you are definitely right. We shall separate these paragraphs in our revised manuscript.

No.# 9 I presume this is infiltration from the wadi when it carries water.

Thank you for this question. Yes, indeed, transmission losses are sometimes also known as Wadi losses or Wadi bed infiltration. We shall add this remark

No.# 10 Is a wadi not by definition ephemeral?

Thank you for this comment. You are correct – so we shall reformulate: "*from ephemeral streams, such as wadis*"

No.# 11 If the springs are all above the wadi floor, the flow from the wadi to the springs must be zero. But your phrase suggests there is a flow, but one that is very small compared to the other flows feeding the springs.

Thank you for this comment. You have observed correctly. Contribution to perched spring flow by transmission loss in our study area is strictly zero, not 'negligible'. We shall separate the two issues: Recharge from TL in general is negligible. In the special case of the perched aquifers above the Wadis, such contribution by definition is impossible.

No.# 12 Please stick to English

Thank you for the comment. We shall see to that: We shall replace "talgrund" by valley floor"

No.# 13 Why do the aquitards and not the aquifers (possibly perched) feed the springs? Also, the leaky aquitards appear out of the blue in the text.

Thank you for the comment. We shall see to that. This was a typo! We meant leaky aquifers, not 'aquitards' of course.

No.# 14 This reference is apparently to a figure in another text. But what text?

Thank you for this comment. Please allow us to disagree: Our reference is correct – it is Fig.7. In the draft, all figures and tables appear at the end of the article. But in the final print version of the paper, this figure will appear within chapter 4!

No.# 15 But the wadis do carry water sometimes. Does none of the water in the wadis come from the catchment areas feeding the springs? Your statement here implies this.

Thank you for the comment. Allow us to repeat this point. In general terms, the amounts from Wadi flow (runoff) are a negligible portion (<1%) of the overall recharge budget. A separate point is the special case of the perched aquifers. Here, TL contribution to recharge is nil, since the Wadis run below the base of the aquifers.

By contrast, also the lands over the perched aquifer formations trigger runoff and thus contribute to overall wadi flow (although in negligible portions).

These mechanisms shall be explained in detail in Ch. 2. Area, and here we shall insert a note, pointing to this discussion.

No.# 16 Like the leaky aquitrads, this wadi appears from nowhere. I think you need to go back to the text and first provide us with a proper description of the local geohydrology, inlcuding cross-sections and maps.

Thank you for this comment. Well, not exactly from nowhere. Wadi Natuf (besides being the title of the paper) had already been mentioned twice before in the abstract (its size and the applicability of our methods beyond Wadi Natuf).

As already mentioned, we shall rewrite the entire chapters Introduction and Methodology. However, the correct place to introduce the area of Wadi Natuf is chapter 2. Area. Here, in the beginning of the introduction, we shall stick to the list of 5 points you kindly suggested in your comment No. 1 on page 3.

With respect to a general term like 'leaky aquitards', you are correct. We shall properly introduce such terms upon first mentioning, as already mentioned before and in our answers to the other reviews. We also plan to add a conceptual model graph, which will explain many (but of course not all) of the central terms of our methodology.

This implies that this material needs to be moved out of the Introduction. I believe that would be entirely appropriate. As it is, the Introduction is poorly organized and does not serve the purpose of an Introduction.

Thank you for this comment. Please see above.

Alternatively, you briefly introduce the WAB and the wadi and other features you need and then discuss past research done on them to makeit clear why your reserach has relevance and where it fits in. In that case you have to explain why this regional study has merits for the HESS readership. Anyway, you have to rewrite this - the text is simply too confusing in its present state.

Thank you for this very specific and helpful comment. We shall thoroughly check and test our options. So far, we tried to first introduce the general methods used worldwide and recapitulate what has been done in the region and which methods were applied in former research (and also present the research questions and objectives in the Introduction). Only then we wanted to talk about the specific conditions on the ground in Wadi Natuf (Ch. 2 Area) and finally apply these conditions to our specific approach (Ch. 3. Methodology).

With respect to the general applicability and merits of our research, please allow us to remind you that the paper is written for a special regional issue of HESS.

No.# 17 Yet the wadi is there. It can only erode itself into existence when it has runoff from time to time.

Thank you for this comment. We shall also revise this passage. Maybe it helps to simply quantify what we mean by negligible: Transmission loss form runoff does occur, but in orders of 0.6 mcm annually, when rain is over 58 mcm and recharge between 24 mcm and 28 mcm, annually (Annual net runoff leaving Wadi Natuf even stands at 0.11% of rainfall only). Therefore the contribution of runoff to recharge is negligible (2.3%). Runoff occurs regularly with 3 - 6 short pronounced events annually. It definitely is a stable feature in Wadi Natuf. It is no surprise at all that over millions of years, wadis were incised into the bedrock. However, this does not alter that fact that quantitatively spoken, and when compared to rainfall and recharge amounts "total surface runoff from Wadi Natuf was found to be negligible in an annual balance". We can only offer to modify the sentence in such a way that it becomes clear that we speak in relative terms, by comparison with rain and recharge.

Was the time scale of the study you quote too short, or has the climate changed and is the wadi a relict from a wetter era?

Thank you for this additional question. We shall make clear that this existing climate – despite an extremely low runoff coefficient – allows for the erosion of deep wadis over tens of millions of years (as is the case everywhere in the world).

No.# 18 This repeats itself, even using the same reference twice (Andreo et al.) to make the same point two times.

Thank you for the comment. You are correct. We shall see to that (i.e. remove the repetition).

No.# 19 utmost

Thank you for the correction. We shall see to that.

No.# 20 This is nearly incomprehensible. Do we need this level of detail here? If so, please explain it better.

Thank you for the comment. We shall revise the whole Introduction for legibility and flow of arguments. We already answered to the other reviewers that we must reconsider the use of the term 'ranking'.

But we do draw on the methodology of Radulović (2011). In his study - an assessment of the spatial distribution of recharge - he reduced the complexity by identifying the "most important natural factors, which influence recharge and enable the most reliable assessment"; he identified different (spatially distinguishable) recharge-related parameters such as topography, climate, runoff, vegetation, lithology and karstification etc. and 'categorized' them as classes of ascending weight (which are then used as factors in a formula).

As noted before, the Introduction will only provide an overview over existing methods. The Methodology will then go into more detail, further drawing on such works as Radulović's.

No.# 21 This paragraph is very messy. Elaborate discussions are inserted in what is essentially a ranking of three factors. The grammar is inconsistent becasue the perspective of the narrative changes fluidly. This creates an intertwined set of lines of thought, leaving the reader wondering what points are being made and for what reasons.

Thank you for the comment. You are correct. We shall see to that (i.e. work on grammar & language).

No.# 22 Which processes? You were discussing negligible runoff in Wadi Natuf, and various methods to estimate spatially variable annual recharge, but not the recharge process itself.

Thank you for the comment. You are correct. We should not have spoken of processes here, but of the role of the above factors that rule recharge. In general and as already noted, we shall add a diagram of the conceptual recharge model that will introduce the concepts we draw on. And we shall also refer to Ch. 3. Methodology, where more details on our methods will be given.

No.# 23 Field capacity as a concept has had a questionable reputation. You have to at least specify how it is defined here. Here you can replace it by 'soil hydraulic properties'. Later in the text that is more complicated - all the more reason to provide a proper definition.

Thank you for the comment. You are correct. We shall see to that and introduce and define our terms properly. In Methodology, we shall reconsider the use of Field Capacity concepts. Here in the Introduction, as you suggested, it will suffice to speak of soil characteristics in general.

No.# 24 Please rephrase.

Thank you for the comment. You are correct. We shall see to that and improve clarity.

No.# 25 Apparently the soil and the unsaturated zone are not the same in your view. I suppose in a desert this can be a valid distinction. But the soil must al least be part of the unsaturated zone, yet you seperate the two.

Thank you for your attention to detail. You are correct. We shall make clear that we speak of the bedrock portion of the unsaturated zone (as opposed to its soil portion).

No.# 26 I am not familiar with this term. How is it defined and quantified?
Do you not simply mean its hydraulic conductivity, or perhaps its sorptivity? Both terms have well defined physical meanings, unambiguous dimensions, and defining equations.

Thank you for the question. You are right. Receptiveness or rather 'receptivity' is a term rarely used. We should probably better speak of general recharge capacity (which includes absorption but also storage, permeability and conductivity etc.). Here we simply wanted to state that recharge is also a function of (and limited or enabled by) the underlying bedrock material.

Page 4

No.# 1 Explain abbreviations on first use.

Thank you for the comment. You are correct. We shall see to that.

No.# 2 Youi either adopt parameter values of calibrate them, not both.

Thank you for the comment. We shall work on the language here as well to point out the difference between parameters taken from the literature and parameters measured through empirical field work in situ.

No.# 3 I do not think it is as black and white as you state it here. I can envision many scenarios where you combine the two approaches. Even if you calibrate every parameter on field data, the underlying equations were taken from the literature. Or you you develop your own theory here?

Thank you for the comment. You are correct. This sentence was not meant to claim that the two approaches were absolutely mutually exclusive. You are right in your remark that one can "combine the two approaches"; which goes to show that they are indeed distinct. We do not try to stretch the point any further than that (we of course do not claim that we directly "control the spatial variability of … factors determining recharge"). We shall revise our manuscript accordingly.

No.# 4 Why would you repeat a calibration. There are many techniques for multiparameter calibration available.

Thank you for your comment. You are correct; there are many ways of calibration available and we don't want to delve deeper into these issues here. Please refer to our answers to the other reviews on over-parameterization, equifinality problems (RC2) and the general benefit of empirically grounded field observations (RC3).

No.# 5 This sentence for the first time explains what this study adds to exisitng work. But the preceding text does not really lead up to it, although some aspects could point in this direction if they were better phrased to fit the purpose.

Thank you very much for this comment. We shall honour the KIS-approach (*Keep It Simple!*) in our revised manuscript.

No.# 6 But you do not provide references to support this claim.

Thank you for the comment. We believe the next sentence delivers such support (reference to Cheng et al., 2017). But you are correct – we shall point out that also the finding of underestimation in annual approaches refers to Cheng.

No.# 7 In this paragraph you again jump from the general to the specific without warning. This creates considerable incoherence in the line of argument.

Thank you for the comment. You are correct. We shall work on the flow of arguments – also by introducing more sub-chapters with titles.

No.# 8 Which claim?

Thank you for the comment. You are correct. We mean 'finding' not 'claim'. We shall change that. We actually meant the observation (among others, by Cheng) that recharge in (semi-)arid areas is often underestimated when the estimation interval is too long.

No.# 9 Nearby to what?

Thank you for the comment. You are correct. We shall see to that and be more precise, i.e. "nearby Wadi Natuf" (and add the distance: "8 km to the East").

No.# 10 You can have infiltration excess overflow and saturation excess overflow. Some time into prolonged rainfall the difference between the two becomes moot. Please consult any hydrological or soil physics textbook for a more insightful discussion of infiltration and runoff generation.

Thank you for the comment. Please refer also to our answers to review 2. We do not believe this difference is moot: "*Other field studies … pointed to the importance of soil saturation*" before "*surface runoff and recharge*" are triggered. It is an essential part of the discussion of countless papers not only but especially in semi-arid climates and also in the region (as quoted by us, and also by Ries et al., 2015, etc.).

We would not state that any infiltration excess can be excluded to 100% in Wadi Natuf. What interests us is rather the question whether we can determine a dominant, most typical process – and this is saturation excess for both runoff and recharge in this area of the West Bank! With all due respect, we have studied, published and co-authored hydrological textbooks on saturation excess runoff.

By the way – we did not mention it in our paper because the results were not conclusive (not reliably quantifiable), but we also performed a large series of double-ring infiltrometers tests (as a first attempt to approach infiltration). And all the tests showed a clear pattern, namely that the observed apparent infiltration rates of the soil were unreasonably high! This confirms an observation of other scholars in the region (T. Grodek from Hebrew University), that apparent soil infiltration capacity in such tests is too high to be true, particularly, but not only in summer, when desiccation cracks are frequent. This goes to show that, at least in our area, such (apparent) infiltration excess is hard to reach and thus hardly poses a limit (threshold), beyond which runoff is triggered. In other words – reliance on such high (apparent) infiltration rates and excess underestimates runoff generation (compare also with Lange et al., 2003 and his findings during sprinkler test with extremely high artificial 'rainfall' intensities…)

No.# 11 Where does this come from and why do we need to know this? One of the cited papers is about paleoclimatic research but you seem to be interested in current recharge.

Thank you for the comment. We shall revise the manuscript, also in structure. Our point was that it was clearly discernible that the quoted studies (like cave drip, tracers, etc.) have the problem that they refer to a plot-scale and do not regionalize.

The work by Bar-Matthews is seen – not only by us – as widely relevant not only for "paleoclimate research" but also for current recharge... We shall restructure the paragraphs, so that it becomes clear why we quote these studies here.

No.# 12

Thank you for the comment. We shall see to that.

No.# 13 How is this relevant for you, you are working in the Negev desert, are you not?

Thank you for the comment. Clearly not. We hope we have made clear that we do not work in the Negev (actually, the Negev is only mentioned once, in a table in the Appendix, quoting other areas' RC).

A cursory look at our paper would reveal that we work in the wider Jerusalem area (Line 293) and more specifically near Ramallah (L624, 678) in the West Bank (see Lines 24, 189, 190, 220, 339, 559, 682, 685, 792), see also L180: "Wadi Natuf is a 103 km^2 catchment stretching from the mountain plateau at the crest of the West Bank" and L26/27 "Wadi Natuf … in the West Bank slopes and mountain region,". Refer also to our location map Figure 1a (L779) and to your own comment No.1 (p.2). Last not least – refer to the title (line 4)…

No.# 14 Is this not a massive problem if you are studying recharge?

Thank you for your comment. Yes, it may seem so. But we look at it from a different angle - it simply means that indirect procedures (drawing on spring discharge) are excluded, which is a fairly typical situation worldwide. We believe that this is why our approach is relevant (and this is what we continue to discuss in the following paragraph).

Page 5

No.# 1 Without more information about the model set-up this information has little value.

Thank you for your comment. Please also refer to the details of our answer on comment No. #4 (p.5) in review 3. You may be right – our description was too brief to be understandable. We do not want to delve into the details of their modelling codes. All we wanted to show was that Weiss & Gvirtzman used a rather feeble crutch, which is based on a voluntary conceptual misunderstanding when they deliberately treated (or had to treat) the unsaturated zone as fully saturated – which means, they simply inserted an artificial condition by which all leakage is eliminated and thus did not have to worry about this crucial source of error and uncertainty (compare also to your comment #14, p.4 above). So, in this sense this attests to their conceptual weakness and in our view therefore to the conceptual strength of our approach. We definitely should make this point more clearly and explicitly.

No.# 2 What does this mean?

Thank you for your comment. We shall remove the page number in the quote "2010: 23)".

No.# 3 But are not the aquifers being recharged instead of the formations they are in?

Thank you for your comment. As already explained in our above answer on the sub-aquifers (No. # 5, p.2), in our region conventional nomenclature assigns different 'aquifers' or 'sub-aquifers' to the different individual formations. In this sense, the two terms become stratigraphically somewhat synonymous. Strictly speaking, we did not exactly say that, which you understood: We did not say that the formations were recharged. We want to spatially differentiate recharge not INTO but FOR the different formations at hand: "This study aims at the estimation of lithology-, soil- and landform-specific recharge **for** the formations in the carbonate aquifers of Wadi Natuf, located in the central WAB". This is based on the central fining that recharge as a surface process differs at different locations – different aquifer outcrops. We try to assign specific recharge rates (and RCs) for the different formation outcrops. In addition, it is not always the case that the recharging aquifer is identical to the formation at which recharge as a surface process takes place. In the deep regional aquifers ('aquifer complexes'!), one or even two formations can belong entirely to the unsaturated zone ABOVE the water table. (We are aware that from a different perspective of, say, indirect procedures based on storage change, the wording would be different, but not in our case). Also note that we used the plural for 'formations' because in some aquifers, several formations belong to one and the same aquifer (or rather 'aquifer complex'). But not the other way round! In our area, one formation never houses tow independent aquifers…

We need to stress here again, that in our study, deep percolation is equated to recharge. This definition is essential and basic to our approach. Once the infiltrated soil water reaches beyond the soil and enters the unsaturated bedrock, we speak of and consider it recharge.

No.# 4 This section is very well written.

Thank you very much - finally one good chapter.

No.# 5 This repeates the Introduction. I think this information fits better here.

Thank you for this comment. Yes, according to your above list of 5 points for the Introduction, such area-specific information should be placed in Ch. 2. This shall be dealt with within the general restructuring of the manuscript. However, already in the Introduction we already have to announce – at least in very general terms – that we shall and can neglect runoff.

Page 6

No.# 1 How old is that (for those of us less familiar with geological time tables)? I have the impession the nomenclature for the time periods is local. Is it possilble to provide a table with names and times? the legend in Fig. 1 seems to contain some of that information but the resolution is too low.

Thank you for the comment. We shall include the age information in Fig 1b.

No.# 2 There are several elements in here that appear to be Results. Parts of the text and the tables are so incomprehensible that I cannot properly review them.

Thank you for the comment. Please allow us to note that we feel that here, we must already report the important finding that aquitard outcrops beneath the perched aquifers surround the hills and thus fully isolate these local aquifers. The repercussions of this belong to and should be part of the Methodology. Other elements indeed rather belong to Results. This shall be dealt with within the general restructuring of the manuscript.

Page 7

No.# 1 Convincing point. Earlier on, you mentioned potential applications of your approach to aquifer/wadi systems elsewhere in the world.

1) How important is the complete extent of the aquitards below the perched aquifers for applying your approach?

Thank you for your comment. Please allow us to note not only that all of the perched aquifer formation is uniformly underlain by the aquitard, but to stress in addition that the outcrops of this aquitard surround the hill – which allows a catchment size determination with certainty. This is indeed very important but only for one aspect off our work; the leakage estimation through a water budget approach rests on this finding. The other parts of our study do not depend on this special formation and outcrop geometry.

2) Is anything known about the occurrence of this configuration (perched aquifers completely underlain by aquitards) elsewhere?

Thank you for this question. Regionally, there are many such areas in other parts of the West Bank Mountains and in other hillsides of the 'Mountain aquifers' in all of

Israel/Historical Palestine (Lower and Upper Galilee, Mt Carmel) and also in parts of the Northern Highlands on the Jordanian side of the Jordan Rift Valley – hence also the studies by Gvirtzman and others.

In general, your question is not easy to answer and to determine with precision. But we can approach this question conceptually. You need two ingredients for such a geometry: a) A relatively diverse lithostratigraphy with alternating aquifers/aquitards and b) a certain type of undulating landscape with hills of the right size (not flat plains and not very steep, high mountains; such conditions are found in many places around the Mediterranean Sea and in many of the 'central highlands' and 'low mountain ranges' worldwide…)

But to the best of our knowledge, this not so rare instance has never been used systematically in the hydrogeological research…

No.# 2 How exactly?

Thank you for the comment. You are correct. This is an omission that shall definitely be addressed in our revised manuscript.

No.# 3 Soil moisture measurments appear at the start and the end of this paragraph. How are these related? I think restructuring this paragraph might improve its clarity. Also, you can refer to some of the tables to provide part of the information I ask for below.

Thank you for the comment. You are correct. The soil sampling (at the end of the section) complemented the 8 SM stations (at its beginning). We shall bring these issues together and refer to the tables. The explanation of the relation between the two passages is already stated, we believe: "In order **to increase the representativeness** of point-scale measurements, parallel soil moisture measurements were carried out … (Table 1a)."

No.# 4 At what depth(s) were the sensors installed (refer to Table 1a here),

Thank you for your comment. Yes, we shall refer to Table 1a here.

what were the sensors, and what depth interval did their measurements represent?

Thank you for your comment. But the sensor types are described already (both, Theta and ECH2O), just three sentences down your question.

No.# 5 That's a bit unspecific, although not essential. Bit I suppose you know the exact number, so why not report it?

Thank you for your comment. We are still searching for this information (from the year 2003). Actually, it was over hundred samples – but to stay on the safe side, we only mentioned 'dozens'. If we find the exact number, we shall surely include it.)

No.# 6 From what depth intervals did you take the samples?

Thank you for your comment. This can be answered with certainty: We sampled at all depths – i.e. from the surface down to the bottom of the soil (just above the bedrock).

No.# 7 I do not fully understand how. Please explain in section 3.

Thank you for your comment. You are correct. We shall see to that and explain the concept of the regionalisation step of spring measurements in more detail in section 3.3 hereunder (and as already stated in note # 2 above). And we shall in particular explain the role and use of key date measurements for the spring groups.

No.# 8

Thank you for the comment. We shall see to that (i.e. delete "parallel").

No.# 9 Explain abbreviations on first use.

Thank you for the comment. You are correct. We shall see to that

No.# 10 A bit more deail please: depth intervals, sample volume, etc.

Thank you for the comment. You are correct. We shall see to that – for depths, see also comment No. # 6. But many sampling depths were chosen randomly by the field team (of SUSMAQ project).

Sample volumes (between 200 and 400 g in most cases – in rare cases, up to 1.5kg) shall be added.

No.# 11 wilting point for agricultural crops is pF 4.2. For desert plants it can be much higher. Please define both field capacity and wilting point. From this text I suspect you used the water content at a fixed matric potential as field capacity. If so, what was that matric potential value, and why did you select it?

Thank you for your comment. If you please refer to reviews 2 (Line 297) and 3 (No.2, p.18), where this matter was already discussed. We followed their recommendations and agreed that we shall refrain from using the term WP and use 'minimum soil moisture' (SM_{min}) instead.

No.# 12 i

Thank you for the comment. We shall see to that correction accordingly (see comment above).

Page 8

No.# 1 The unit % is not one I am familiar with for either field capacity or wilting point. They should be volumetric water contents, possibly converted to mm by integrating over the soil depth.

Thank you for your comment. Yes, we shall convert this to mm/mm, and when possible also normalize it into mm.

No.# 2 OK, here you define FC. What is a soil moisture plateau? Is it uniform across the area or specific for the various formations?

Thank you for your comment. We shall properly define our terminology, as stated already. In the graph in Fig. 4 you can see that observed soil moisture rises repeatedly over several seasons to a 'maximum' of 112.5 mm (quoted in Table 1.a). This recurrent level can therefore be called a plateau, as visible in the graph. This is equated with Field Capacity. This is in our terminology distinguished from 'absolute maxima' of so-called 'peak SM', which only occur occasionally and always only for very brief periods. Field capacity is the moisture content a soil can carry against gravity. The brief peak SM periods therefore show an unstable situation of water content above field capacity (often including water in open cracks and always above the level the soil can hold against gravity), in other words: brief over-saturation.

To your second question: Importantly, the values of FC (and also peak SM) are different at every station. They are location specific physical soil characteristics (but stable over time). We shall add some of these explanations discussion to the manuscript.

From Fig. 4, the plateau is relatively clear. What remains unlcear is what causes this plateau to exist. What worries me is that the observation data do not show the plateau as clearly. The show a wet range between 100 and 110 mm without clear leveling off.

Thank you for this comment. If we equate the 'plateau value' with FC, then it is quite understandable how and why it exists - as a typical physical soil property, i.e. as a function of soil chemistry, granulometry, porosity, texture, etc.

The graph in figure 4 and also in the other soil stations is rather clear. The deviation from the value indicated in Table 1.a. (112.5 mm) is actually quite small, not large. And at least during the first three seasons, the graph shows a rather remarkable levelling-off at exactly the same height during peak rainfall periods – i.e. full soil saturation and thus recharge.

(Please note that peak SM is also discussed in the text, lines 301 and 305. And please also refer to our answers to your comment No. 1 on page 28)

The plateau may well be an artefact of your model, in which case I do not know how to interpret it.

Thank you for this question. Yes, sure - the ‘plateau’ of the modelled graph is an “*artefact*” and not by coincidence. It was produced by us. This is at the centre of the model. Deep percolation occurs always when full saturation of the (field capacity – ‘plateau’ value) is reached and additional net rain (HEP, after deducting ET) is added and infiltrates. All this additional rain then percolates. This is exactly the logic of our parsimonious model. In other words: It is the opposite of an ‘*artefact*’ and instead the intended core condition of the model.

No.# 3 The readabiltiy of the paper would be improved if you would give these relations as equations instead of hiding them in the text. I find myself using the 'search' tool much more often than I should to read a paper.

Thank you for the comment. Fine, we shall see to that

No.# 4 I understand what you mean but the term is unfortunate, as a soil cannot be oversaturated. 'Excessive soil wetness' perhaps?

Thank you for the comment. Oh yes, sure; soils can indeed be over-saturated – i.e. always when field capacity is exceeded. Such over-saturation (peak SM) in our conceptual model is a passing feature and indeed was never observed to last longer than a few hours during peak rainfall events (because it is the water content beyond what the soil can hold against gravity). One may also call it *excess water content*… There is a lot of literature that explicitly refers to this concept (not least Ries et al. in HESS, 2015, from our region, which we quote in the text).

No.# 5 Grammar is off. What do you mean?

Thank you for the comment. We shall see to correcting that. What we meant was: we subtracted minimum SM values from the measured soil moisture values and obtained values of “available SM”. (Equally, by subtracting SM_{min} from observed maximum SM values results in maximum available SM, which represents FC_e – see Table 1.a.).

No.# 6 Not the proper HESS format

Thank you for the comment. We shall see to that…

$$AW_{i+1} = AW_i + P_{i+1} - ETa_{i+1} - DP_{i+1} \qquad (3)$$

No.# 7 This section is very difficult to comprehend, in a large part because terms like ranking, correlation, criteria et. are used in ways that are not semantically sound. A through rewrite is required.

Thank you for the comment. We shall revise our terminology and introduce it properly. As already discussed above and in the other reviews, we shall rework the entire concept and wording of the ‘ranking’ procedure.

No.# 8 You do not seem to rank anything, but rather estimate recharge rates.

Thank you. See our comment above.

No.# 9 Are the travel times in the unsaturated zone short enough for percolation to reach the perched aquifer within a year?

Thank you for this question. In the strongly karstified Mt. Aquifer, despite the very thick unsaturated zone, travel times lie within days (often hours! – Schmidt et al., 2014). Piezometric aquifer response is almost instantaneous. Extremely long travel times will take weeks at maximum.

Page 9

No.# 1 1b?

Thank you for the comment. You are correct. Table 1b and Table 3 are more appropriate.

No.# 2 What does this mean?

Thank you for the comment. You are correct. We shall see to that and explain the evaluation procedure.

No.# 3 What does that mean in this context?

Thank you for the comment. You are correct. We shall see to that and explain the evaluation procedure. See comments No. # 7 and 8 (on p. 8).

No.# 4 What does this mean in this context?

Thank you for the comment. See comments above.

No.# 5 Why is this not reflected in the units then?

Thank you for the comment. We shall revise this and be more specific. Tables 4, 5 and 6 show values, normalized to mm/a.

No.# 6 The calculation of Q is an essential step, yet you do not provide equations for it. The reference 'as described above' without even a section number is inadequate.

Thank you for the comment. We shall see to that.

Because of the generally chaotic flow of the paper it is cumbersome to look this up in the main text. You should not force the reader to do the work, but simply provide the equations instead.

Thank you for the comment. You are correct. We shall see to that. The method of calculation of spring group discharge shall be explained in more detail here (in a new sub-chapter Spring Grouping and Budgeting Methodology) and equations shall be provided.

No.# 7 This too should be formalized in an equation.

Thank you for the comment. We shall see to that (i.e. providing an equation on how we calculated spring group discharge in other years, based on reference year 2003/04).

It is important to note that some rather severe assumptions are implicit in this approach that you do not acknowledge at all in the text. I would prefer to have these made explicit and critically discussed. You can then also explore what happens to the accuray of your simulations (e.g. systematic over- or underprediction of groundwater recharge) if these assumptions are wrong.

Thank you for the comment. We shall see to that and lay bare our assumption, which was that the five daily springs are representative for the entire respective spring groups and that reference year 2003/04 is representative for the other years. Other than that, our spring budget calculations were quite straight forward.

By ignoring the underlying assumptions you missed a chance to develop a more penetrating analysis I believe.

Thank you for the comment. You are correct. Please refer to comment No. # 6.

No.# 8 And yet, 30 lines below you discuss the termporal variation in terms of unquantified correlations and claim that annual rainfall is insufficient to estimate recharge.

Thank you for the comment. You are correct. Our description has been somewhat confusing. Although this paper does not focus on temporal variation as much as on spatial, we shall revise this section and eliminate the contradictions. Please also refer to review 2 and 3 and our answers above.

Again, Cheng is right that in order to establish recharge, the analysis should be performed at the event, not annual level. But once we have modelled a series of annual recharge rates, we can then work with them as annual values. This is a quite common approach.

We shall revise our manuscript accordingly.

I do not think you can maintain the viewpoint that you are only looking at spatial aspects and still can write a meaningful paper.

Thank you for this comment. Please refer to the above. If we crated anywhere the impression that we were exclusively looking at spatial variation and entirely disregarding temporal variation, then we certainly have to correct this misunderstanding. We shall revise the manuscript accordingly and point out and distinguish what is the focus of our paper and what are other considerations.

No.# 9 That seems to be wildly optimistic and you provide absolutely no proof. A sample size of 7 individuals may give you a more or less realistic estimate of the average of a variate, but only if variation is minimal can it be expected to give you a somewhat reliable estimate of the second and higher moments of the variate's distribution. But you have not carried out even the most basic statistical analysis and yet you write down this claim, which you cannot substantiate at the slightest.

Thank you for this comment. You are correct. We shall provide statistical analysis on that claim (that our seven years of observation happened to lie well within the range of long-term rainfall variation).

No.# 10 A statement like this belongs in the main text.

Thank you for the comment. You are correct. We shall eliminate footnotes.

Page 10

No.# 1 This only applies to the high moisture levels, but even there, the fluctuations are sometimes out of phase.

Thank you for this comment. You touch upon an important point. This also pertains to your comment No. #1 on page 28 (Fig. 4) and to your initial remarks on the Figure. Please also refer to our answers there on the mechanisms you proposed.

.

The model performance around December is very poor. May 2008 is also a complete miss. The readers expect a critical evaluation of the model. With appraisals like this you are not delivering that.

Thank you for this comment. As you will see in our attached report, the brief periods of misreadings actually were in November and April. You are correct – the readers expect and deserve a thorough investigation and explanation on this point, which we will provide for the revised manuscript (as shown in the attached file on instrument failure in SM readings).

No.# 2 In 2007-2008, yes. In other years, not so much.

Thank you for this comment. Apart from the November problems, the winter readings and the match is excellent also in 2005/06 and 2006/07. But surely, we shall provide statistical evidence for this as well.

No.# 3 This analysis is underwhelming. Can you not derive anything more quantitative? In the title you state you are doing recharge estimates, but from this it appears as if this is not doable.

Thank you for this comment. Please allow us to again disagree to some extent. It of course depends on what one understands by recharge estimates. It is certainly true that we do not focus on a micro-scale event-by-event analysis of the temporal patterns of recharge triggering and on details such as thresholds. This is because our research question clearly states (shall state) that we focus on the spatial variability of recharge.

However, we can indeed discuss this point you raise a little bit, if you deem it worthy and necessary. Because our model certainly does deliver certain results on recharge thresholds and the like. Only, and this is important, we have to caution already that we will not find a uniform pattern of threshold depending merely on the amount of accumulated rainfall. This of course is just what one should expect, if we follow Cheng's observation. The individual pattern of event distribution, its size, intensity, number, concentration, dry spells etc. are what can lead to recharge in an apparently dry winter and to no recharge in an apparently wet winter – at least and especially for the onset of recharge and for the phase of wetting-up of dry summer soils during the early winter months. Since the event distribution changes from year of year, so too *should* the recharge alter from season to season. We shall not attempt to make a major point out of the discussion of recharge patterns in individual years, let alone, events. However, we shall give some detail and discussion on some of our findings and model results here in the text.

No.# 4 I think you mean FC (also in the line below).

Thank you for this comment. You are partially correct. But we should remark that as we here refer to area normalised FC (in mm), we wrote it as FC_e. However, the sentence here would also make sense if we spoke about FC instead. If we talk about reality on the ground, FC and FC_e mark the same physical condition – namely, "maximum SM" – whether expressed as total water content or only as maximum effective water content. In both cases, this marks the limit beyond which recharge is triggered according to our saturation excess model.

We could also emphasise and add to the sentence that we speak about maximum available SM by saying: "*usually the FC_e threshold for deep percolation was reached in*"…

No.# 5 -

See comment above.

No.# 6 The reasong in this paragrpah is incomplete, although I believe it is valid. If you want to show the effect of different flow domains in a karstic system you have to contrast the response of the wells to a recharge event under different conditions. In this case: a dry matrix domain versusa moist matrix domain. Here, you only show the effect of the dry matrix domain and let the reader figure out the reponse fo the system when the matrix domain is wet, but that is your job.

Thank you for this comment. You are partially correct. And we shall revise some of our wording here. However, we believe that at least implicitly, we covered the wet matrix too in this sentence: This is because the strong events mentioned in this sentence certainly refer to an already wetted system. And actually, we rather thought less about a wet vs. dry MATRIX, but rather the effects of matrix vs. main conduits.

But we shall make these ideas more explicit in our revised manuscript.

No.# 7 Why did you not metnion this type of models in trhe Introduction, if you are using it here to explain your results?

Thank you for your comment. You are correct; we shall introduce double porosity earlier, either in the introduction or in chapter 2. Area. Btw, we found the existence of a double porosity not all too surprising in such a karstic environment as ours.

No.# 8 What is the point of this number? It is not a correlation coefficient,...

Thank you for your comment. You are correct. The wording of this sentence is somewhat confusing. We shall revise the language and omit the number or discuss it properly.

Please allow us to clarify: 742 mm/a was the maximum rainfall measured as average area precipitation of the entire Natuf catchment (the maximum value of all annual area averages). We did not want to express more than the following finding: Consistently, in all sub-catchments, maximum rainfall occurred in 2004/05. And also consistently, during that same year, all individual SM stations showed the highest DP rates of all years modelled. The year of maximum rainfall matches the year of maximum recharge. We termed this match a "correlation" (defined as "*a mutual relationship or connection between two or more things*"), which are here the maximum rainfall and the maximum recharge. It is of course not a "correlation coefficient". In our view however, this is an important statistical finding with respect to the quality of our model that deserves discussion (and could be discussed in our new chapter Discussion).

However, you are correct: Since we did not model with the overall area rainfall but with the daily precipitation heights of the respective sub-catchments, the mentioning of this number is confusing and shall be dropped.

… and it is unclear whether it is rainfall or perhaps recharge per area.

Thank you for your comment. You are correct – our mistake.

Furthermore if you state that there is a correlation, a single pair of rainfall-recharge numbers proves absolutely nothing. You have to plot the one against the other and analyse the trend.

Thank you for this additional explanation. As mentioned before, you are correct. If it was only one station where the years of P_{max} and DP_{max} match, this would "prove absolutely nothing". We failed to stress that this was a consistent pattern throughout all of our stations (models), and we shall correct this.

As to your suggestion, we could indeed replace Table 2 by a graph, which plots the individual recharge rates against each other and against rainfall and then discuss the trends. (Please note that we would then send Table 2 into the Appendix. This is because, being an article for a regional issue of HESS, the absolute values of area recharge for individual years and stations are of interest to regional researches and should be documented).

No.# 9 Variables can be correlated or uncorrelated, but there is no such thing as a systematic correlation. Also, you fail to provide any quantitative measure of the correlation you claim to have found. This is another example fo sloppy use of terminology that weakens the paper.

Thank you for your comment. See above. We shall revise this section accordingly.

No.# 10 Either they do or they do not. Again, you do not provide quantitative information.

Thank you for your comment. You are correct. We shall revise this as well.

No.# 11 This sounds terribly specific. You claim that only the soil moisture balance approach with daily time steps can work.

Thank you for your comment. You are correct and this belongs in the Methodology. We shall revise this accordingly.

Why are you excluding numerical modeling with time steps that vary depending on rainfall and evapotranspiration rates, for instance?

Thank you for your comment. You are correct. But we did not intend to exclude any models. Anyway, such arguments belong to a new chapter 'Discussion'. We shall revise this accordingly.

In cases of surface runoff, daily time steps seem to be too long. This claim seems to be too specific and too definite based on the extent of our analysis and the regional nature of your study.

Thank you for your comment. You are correct. In surface runoff, much smaller time steps were used (measured in 5-minute steps). Anyway, we are only discussing recharge here. We shall revise this as outlined above.

No.# 12 There are no soil properties in Table 2.

Thank you for your comment. You are correct. This is a typo. The soil features are presented in Table 1a. We shall change this.

No.# 13 You have no idea if its is unless you have a much longer weather record available and carry out a proper statistical analysis. This will also require you to define quantitative criteria by which to judge whether this 7-year record is representative of the record of x years (x >>7). As itis, this is just unsubstantiated speculation.

Thank you for your comment. You are correct. Please also refer to reviews 2 and 3. We shall add a statistical analysis of long-term precipitation, as stated above.

No.# 14 the rainfall converted to

Thank you for your comment. You are correct - we shall correct this into RC (of >57%).

No.# 15

Thank you for your comment. You are correct. We shall correct the typo ("185 and 233 mm, respectively").

Page 11

No.# 1 Table 2 does not have a column with recharge coefficients.

Thank you for your comment. We shall explain better: Recharge coefficients are listed there (but as a line, not as a column); they are not marked as "RC", but expressed as "avg. DP/avg. P". This shall be replaced and explained in the caption.

No.# 2 How do you define this?

Thank you for your question. The term 'Recharge potential' in the literature is usually used as a general term to describe spatial differences in recharge. It is a stable physical condition of the ground, regardless the temporal variations of rainfall in individual years – hence '*potential*'. Recharge potential is often used in maps that usually indicate areas of low, medium and high 'recharge potential' (compare e.g. Senanayake et al., 2016).

No.# 3 How can that be? You only have one estimate per site, right? So you simply have an estimate on each location.

Thank you for your comment. You are correct. We had "only" 8 locations with long-term series of intensive soil moisture measurements and modelling. This however is a remarkably high number, especially in this region, where nobody has performed such a study (duration and spatial density). However, the expression is somewhat unfortunate and we shall reformulate this (since the 8 sites are simply 8 sites and thus the eight different recharge rates were not derived with a particularly "high" or "low" resolution, but simply in 8 key locations and formations.) Discussion of the spatial resolution of our approach rather belongs into the new chapter Discussion (or Conclusions).

No.# 5 Repetitive.

Thank you for your comment. You are correct. We shall address this when we restructure the manuscript.

No.# 6 This is Methodology and should not be here. That being said, the methodology does not apear to be sound. How can you assign a value to a continuous variable (RC) based on a rank number?

Thank you for your comment. You are correct. We shall discuss this in Ch. 3. Methodology and there also in greater detail on how we assign values derived from modelling to other formations.

I am almost through the paper and have not encountered a clear explanation of the three ranking methods (not scenarios). Also, I have not seen a table with the three estimated recharge coefficients and amounts for the soil moisture sites, only for the aggregated scale in Table 4, to which you have not referred at this point.

Thank you for your comment. You are correct. We shall address this as stated above several times.

Given the level of uncertainty that you for some reason do not discuss but must be sizeable, I think the values in Table 4 are so similar the emphasis you place on these three approaches is overblown. You could simply have stated that the recharge estimates based on these three approaches varies by x % (less than 20 %), and possibly that you recommmend one of them because your analysis suggested its reliability was superior. Note that this kind of analysis is completely missing from the paper.

Thank you for your comment. You are correct. We shall address this in our revised manuscript. Indeed the differences between the three alternatives are so small ($\pm$10%) that they pale against the uncertainties created by the three work steps (measurements, modelling and assignment). We could therefore simply stick to one alternative. We shall also discuss, as you suggest, the pros and cons of different alternatives using different factors – but again, better under the chapter Discussion than under Results. However, one point should be made here: The fact that all three independent alternatives result in such similar overall recharge is an important result and finding in itself and should be emphasised.

No.# 7 The grammar is incorrect, obscuring the meaning.

Thank you for your comment. You are correct. We shall correct and reformulate this (…*lies in the range of values, reported for WAB and other basins such as*…).

No.# 8 This material belongs in the Introduction.

Thank you for the comment. We shall see to that, i.e. we will mention other aquifers summarized as "Mt. Aq." in the Introduction already. However, we are not sure that we can present such level of detail in the introduction. This remains to be seen.

No.# 9 The < and > signs make absolutely no sense here.

Thank you for your comment. We shall correct this and express it in words, i.e. "*range from ca. 40% to 47%*".

No.# 10 Why is this relevant?

Thank you for the question. The relevance stems not only from the similarity of the values, but also from the fact, that Ein Harrasheh (gw catchment) overlaps with Wadi Natuf (surface catchment) – in other words: lies in the immediate vicinity.

No.# 11 Where is the other parenthesis?

Thank you for the question. You are correct - the other bracket should have stood at the end of the sentence – after "to minimum recharge rates". But we could also remove the entire parenthesis in the revised version.

No.# 12 What does this mean?

Thank you for the question. We shall explain this in our revised version, as announced before.

No.# 13 You did not emprirically test them. At least I did not read anything about a test in the Methodology or in the Results. I believe you mean you estimeted them based on 7 years of data. Be that as it may, what point are you making here?

Thank you for the question. You are correct, the expression is wrong – they were not 'tested' but are based on empirical work (measurements) and our model. We shall express this more clearly in our rewritten section (here, namely that we interpolated within the modelled range of RCs, we did not extrapolate).

No.# 14 This phrase is nonsensical. You cannot rank apples according to pears, you can only rank a variable according to its onw value.

Thank you for the remark. You are somewhat correct. As already stated, we shall rewrite this section with new terminology, arguments and structure. (We might even refrain altogether from the term 'ranking'.)

In our understanding, to give an example, one would rank children according to their size or their age, or by age… Both seem equally possible in English: "A ranking is a relationship between a set of items such that, for any two items, the first is either '*ranked* higher than', 'ranked lower than' or 'ranked equal to' the second." So, one would equally rank formations according to their recharge potential. Yet, you are correct that "*ranking RCs according to soil thickness*" makes no sense. Formations and soils are items, coefficients and thicknesses are parameters (that indicate their relationship). One cannot rank parameters by parameters.

Page 12

No.# 1 Introduction material.

??? (You marked the header, inserted by HESS)

No.# 2 Why this hyperbolic language? These formatioins have the smallest recharge coefficents - simply say so.

Thank you for this comment. We shall revise this expression, e.g. they have *a smaller* RCs (but not '*the smallest*', because the marls have the smallest, with an RC of 0%).

No.# 3 This is Methodology. The spring groups should have been introduced much earlier.
Thank you for your comment. You are correct. We shall restructure this.

No.# 4 Partially methodology. Why is 2003/04 a reference year, and why did you not apply this estimate to all years in the data set?
Thank you for your comment. You are correct. We shall restructure this and also explain the methodology of spring grouping and budgeting (e.g. why 2003/04 was chosen as reference year - because we had two key date measurement campaigns in this year). However, we did of course apply this pattern found in 2003/04, to all the other years.

Earlier on you implied that the 7-year obsrvation record had a wide range of weather conditions, so it would be reasonable to test the validity of this approach for different conditions.
Thank you for your comment. You are correct. We shall add statistical analysis on this, as announced before.

It also would perhaps have allowed you to see if there might be an effect of the weather on recharge that lasts more than one year, for instance after a prolonged drought or an unusually wet period.
Thank you for your comment. As already mentioned before (see travel times), recharge in Wadi Natuf and almost the entire WAB is *far* too fast for such a possibility!!

No.# 5 One line above you claim you did this for one year only!   This too belongs in the Methodology section.
Thank you for your comment. You are correct. We shall explain our spring methodology in detail and under Ch. 3.

No.# 6 This is methodology again. Also you have to explain this better.
Thank you for your comment. You are correct. This belongs to Ch. 3 and is a repetition. We shall revise this accordingly and explain the routine of 'reference years'.

No.# 7 This belongs in the Introduction.
Thank you for your comment. You are correct. We shall restructure this. Probably part of it will go into Ch. 2. Area. Another part will be placed in the Introduction.

No.# 8 repetitive
Thank you for your comment. We shall revise this.

No.# 9 repetitive
Thank you for your comment. We shall revise this.

No.# 10 If you assume that the time scales of all subsurface flow processes are well below one year, but you never discussed this.
Thank you for your comment. We shall make this clear.

Page 13

No.# 2 How exactly? Where are the equations, and what assumptions do you need?
Thank you for your comment. No further assumptions or equations are needed. This step is straightforward and can be calculated according to Formula 5 (which we shall refer to in the revised manuscript). Please note, that since the catchment area is the same for all three factors (Q, R and L), we can calculate this in m^3/a as well as in normalized mm/a. (Please note also that here, we shall mention the detail of the captions to Fig. 6 - which will then be a repetition!).

No.# 3 These conditions are so strict that in fact you prove with it that you have carried out a regional study.

Thank you for your comment. You are partly correct. Please also refer to our answers to your comment #1 (part 1 and 2), p.7 and to your initial (free text) comments with the '*list of conditions*'.

First of all, our paper should be considered a regional study (for this regional special issue of HESS), by all means. And it compares with existing studies in the region (the WAB, as well as the Mt. Aquifer, or the Cretaceous aquifers in the Levant, such as Jordan, Syria, Lebanon and of course Egypt, which is a riparian to the WAB in the Sinai).

However, we believe that indeed this approach can be applied everywhere, where conditions are similar. We do respectfully disagree that the conditions are so terribly strict! Please allow us to illustrate this by the following questions.

A - **Can we determine total spring outflows of an aquifer with sufficient precision?** This often is possible, and not only when we have an aquitard outcrop that surrounds the hillside.

B – **Can we identify typical SM locations that stand for different groups of recharging units?** (Please note that only one such possibility is that these units follow the lithostratigraphy of formations.) And can we determine FC and thickness of the typical soil locations? This most certainly is the case in many areas all around the world.

C – **Can we build an average over a certain period that is representative?** If we compare with many, if not most hydrological studies, that often only measure one or two seasons, we definitely lie at the upper end of the range! Otherwise, the new research would have and could find ways around this problem, or has to state her or his results in a more modest fashion…

D – **Can we identify typical spatially distributed physical features such as rock lithology, LU/LC etc. that rule recharge; and can we then arrange these items into a ranking order according to the controlling (parameterized) factors?** This can but most not be done along formations; also other methods could be chosen…

E – **Can we determine with accuracy the groundwater catchment area of a formation?** In this – largely ungauged – sub-basin, we used the particular geometry of the bottom aquitard outcrops. But also other methods are applicable of course, and in many cases of previously studied basins, catchment areas were already established by previous research (modelling, tracer studies, field mapping, etc.).

In addition, we should state that if in other areas one or two conditions are not fully met, a deviation from our approach is most likely possible… So, definitely, there is no reason in principle, and also little reason in practice, why our approach should not work elsewhere…

Page 20

No.# 1 This is not the right place of the units. Also you are not consistent in the designation of year. Elswhere you use a. Consult the HESS guidelinesplease.

Thank you for the comment. You are correct. We shall replace all 'yr' by 'a' and change the position of the unit indication in the table.

Page 21

No.# 1 Way too little information here. The table cannot be read independently of the text. This table appears to be important but its compostion is messy.

Thank you for the comment. We shall see to that and expand our caption texts (so that figures and tables can be read independently) since you explicitly formulate this as a requirement. (We would like to note that other articles in HESS by no means employ exhaustive captions that explain every detail in the table or figure independently to the text of the article). However, this will make repetitions unavoidable.

No.# 2 The 'potential rank' column is utterly unclear. Neither the table nor the text explains what the ranking means (The text disusses ranking of formations). The symbols do not convey any information because they are not explained.

Thank you for the comment. You are correct. The table shall be revised for content and legibility.

The caption will explain that the twelve formations were “ranked” (or were *listed in descending order*) by their recharge potential (in numbers: 1 to 12). Hereby, the lower end (“ranking position” # 1, 2, 3) stands for the highest potential (RC = 57.3%) and the highest number (“ranking position” # 11, 12) stands for the non-recharging impermeable formations with zero recharge potential (RC = 0%).

No.# 3 So there are no ages at all, just a ranking of ages. Thus, the first column has a misleading header.

Thank you for the comment. We shall clarify this issue in the revised manuscript version (the first column indicates the ‘*ranking order*’ according to the age of the respective formations).

Page 25

No.# 1 Scale is not readable and orientation (compass arrow) is missing,. Symbols not explained. Concept inlet is not helpfull - the definition of leaching it appears to provide is a water level of some sort. Colored lines eihter denote the WAB or the isohyets, but are not properly definedand the isohyets lack numerical values and a reference height

Thank you for the comment. We shall see to that. The figure has been thoroughly restructured and revised.

No.# 2 The compass arrow is missing, and the legend is unreadable. Symbols are not explained. Are they main springs or soil moisture monitoring sites?
I need this legend for the names of the geological time periods, but the pixelation is too coarse to allow me to read them.

Thank you for the comment. The figures have been reprocessed for clarity, North was indicated, the legend added and names were made legible. Symbols of main springs shall be mentioned (red dots). The red boxes indeed indicate soil moisture measurement sites.

Page 28

No.# 1 This figure seems to indicate some measurement issues, or point at an inappropriate choice of units.

Thank you very much for this important comment.

We have already answered on this issue before (and we shall attach a file with additional material on SM misreadings to this answer).

We already explained that it is indeed a measurement issue. Soil moisture above rainfall is simply impossible. In Wadi Natuf, at our specific sites, we do not have redistribution of water. A karstic terrain at the top of a hill, with small terraces and a few metres wide soil pockets surrounded by karstic rock simply preclude this possibility of redistribution at or near the surface.

We are not sure about the reasons for the misreadings. Maybe, partially, they could be a systematic problem that has to do with the desiccation cracks after summer, where rainwater can infiltrate rapidly. But what we can see as a pattern is that such misreadings occur in late autumn/early winter, when soils are not yet (fully) wetted. By contrast, the sensors and loggers worked fine during the crucial period of heavy storms, full soil saturation and thus recharge! Therefore, during the important times of the model, during periods of deep percolation (=recharge) the field observations are in sync with the model graph (see Fig.4).

When it rains after a long, dry period it stands to reason that the observed soil moisture cannot rise above the accumulated rainfall. If we look at Nov 2006, we see several showers before storage peaks at about 108 mm. It is not easy to see but I estimate the total amount of rain to be less than 50 mm. May 2008 is even more pronounced. This can only happen if recharge is captured from a large area and concentrated in an area of at most 50 % of the catchment area. If that is indeed the case, the normalization to mm should not have been applied, since the various variables presented in the graph and the overall analysis represent different areas. The dimension of choice should then be volume.

Thank you for your comment. We checked this issue thoroughly and can exclude the possibility of rainfall captured from large areas.

No.# 2 Only the contents of the first sentence pertains to a figure caption. Not enough information is provided to read the figure independently of the text.

Thank you for the comment. We shall see to that (i.e. only make the figure intelligible, not discuss its results).

No.# 3 In the text you write that FCe was calculated, yet here you state you simply set it to some value. These two statements cannot both be true.

Thank you for the comment. Please allow us to explain: Indeed, we calculated the physical soil parameters (like FC-values) from our field observations and then used them in the model. So we set our model up with empirically derived values, calculated from real-time in-situ field observations; whereby '*calculation*' here means simply multiplying observed soil moisture values (m^3/m^3) by observed soil thickness (mm or m). In other words, we simply transformed the field readings into SM-values (also SM_{max} follows directly from the observed plateau values. There was no manipulation involved…

So, "FC_e was set to 112.5 mm" by calculation, i.e. measured available SM (%) by soil thickness. In our view, this is no contradiction. However we shall reformulate this explanation to avoid misunderstandings.

No.# 4 This material belongs in the discussion, not in the figure caption.

Yes, thank you – we shall avoid such discussion of results in the captions.

Page 29

No.# 1 This type of info belongs in Materials and Methods, not in a figure caption. The condition of the springs apparently affects their discharge. So, evidently this also affects groundwater recharge (see Eq. 5). Therefore, spring maintenance is an important factor for groundwater recharge, yet you bring it up only in a figure caption. You call the effect temporary but is appears to double the well production for at least two years and the effect was still very strong at the end of the observation period so you have no clear picture of the persistence of the effect. It seems obvious that the net recharge to the aquifer feeding that particular well changed in 2007 with effects that last well over three years. Did you consider any of this in your calculations, or did you carry out a scenario study to examine the effect of different well maintenance practices?

Thank you for the comment. We shall see to that (i.e. discuss the state of the spring's access pipes in chapters 2. Area or 3. Methodology). We shall clarify this discrepancy in the text. Please also refer to our previous answers on Al-Qos spring readings.

Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-329-RC2, 2018

Hydrology and Earth System Sciences Discussions
Open Access
EGU

Interactive comment on "Field-based groundwater recharge and leakage estimations in a semi-arid Eastern Mediterranean karst catchment, Wadi Natuf, West Bank" *by* Clemens Messerschmid et al.

Anonymous Referee #2

Received and published: 23 July 2018

Answers by: C. Messerschmid, J. Lange and M. Sauter (21 October 2018)

This paper deals with groundwater recharge estimations for a semi-arid environment. The author distinguishes between deep percolation into the major aquifer system and quick spring outflow.

Thank you very much for this comment. There are obviously some misunderstandings about the concept of infiltration processes and the general hydrogeological set-up.

Therefore, we designed a schematic diagram illustrating our understanding of the hydrological / hydrogeological situation that also includes the terms employed during the study. As also found by Schmidt, Sauter et al. (2014) in the neighbouring Eastern Aquifer Basin (EAB), deep percolation (DP) has to be distinguished from the surface-near process of rainwater infiltration into the soil. DP (or recharge) here is understood as the transition of water from the soil column into the underlying bedrock, the unsaturated zone, regardless whether it happens on outcrops of the two regional aquifers or on the local perched aquifer formations. (In areas with deep unsaturated zone, we further distinguish between potential recharge and actual recharge down to the water table.)

In Wadi Natuf, deep percolation (recharge) occurs in both types of aquifers, but only the shallow perched aquifers issue groundwater locally as spring flow. The deep regional aquifers continue to flow into the Coastal Plain near the Mediterranean. We therefore did not distinguish between DP and spring flow (Q).

We shall include a schematic diagram that presents the conceptual model and main processes at work and shall improve the understanding of the terminology.

Especially for water scarce environments, those estimations are very important and the development of new ideas and methods would be very welcome. Unfortunately, the presented study just applies quite simple (and from my point of view, not really adequate) methods,

Thank you for this comment.

Unfortunately, indirect procedures that are based on basin outflows or storage change observations are not applicable in Wadi Natuf (as the entire recharge zone of the WAB). We therefore had to resort to direct procedures, i.e. a quite standard classical soil moisture balance model that has been successfully applied in the region. However – and this is also what reviewer 3 valued – we invested a lot of effort on field measurements to make our robust parsimonious model reliable. We used empirical field-measured soil parameters. And we corroborated the model's findings by the high resolution SM records. In addition we compared the modelled periods of recharge through comparison with spring flow in the perched local systems. These direct procedures apply to both, local perched and regional deep aquifers.

What makes this study special is in addition that we selected well controlled perched aquifer systems where the crucial factor of catchment size is observable in the field. This allows us to budget these systems and therefore calculate groundwater leakage by comparison with spring output.

which are lagging behind modern and creative techniques.

Thank you for this comment. Without pre-empting the detailed answers below, please allow us to state in a summary fashion: We cannot see how we lag behind modern methods and in fact the most recent models that are used in the study area (like Ries, 2015 and Schmidt, 2014). It should be noted that this article will be published in a special issue of HESS on the Dead Sea environs. With respect to the Western Aquifer Basin (WAB) or so-called „Mountain Aquifer“ in general, we do believe that our article and the methods employed are state of the art. And by comparison, we consider our technique for leakage estimations in the perched systems quite 'creative'.

Moreover, the paper is not very well structured and scientific terms are not properly used.

Thank you for the comment. You are correct and we take this comment seriously. We shall revise the manuscript thoroughly. We will work on the language, style, structure and flow of arguments.

Specific concerns:

Line 18-21: There are tons of studies applying spatially distributed hydrological model for groundwater recharge estimation. It is very common that they base on "a) soil type and soil condition, b) land forms such as relief, vegetation and land use, and c) lithology and hydrogeological characteristics of the subsurface rock formation".

Thank you very much for this comment. We appreciate that you consider our methods of spatial differentiation to be modern and adequate. But you are correct. The sentence is misleading as we should have pointed out that our remark about "previous research" was referring to publications in the region (the Dead Sea environs, which are the topic of this special issue of HESS). Or as Ries (2015) pointed out: „Only limited knowledge on recharge dynamics is available for the carbonate Mountain Aquifer system shared between the West Bank and Israel, although it is of regional importance.“ We should indeed stress that previous studies in the region sometimes used one or the other of our approaches, but never this combination. We shall make this clearer in our reworked introduction and discussion of the state of the arts.

Line 29: Evapotranspiration records, really?

Thank you very much for this clarifying comment. You are correct: Potential evapotranspiration (ET_p) was not 'recorded', not measured but was calculated, however based on local real-time recorded weather data…

Line 30: What do you mean by "solid"?

Thank you very much for this comment. You are correct: We shall choose a better description. What we meant was a "robust" model (*a simple but robust parsimonious model of soil moisture and percolation*).

Line 31-32: Model fitting also bases on quantitative field observations.

Thank you very much for your comment regarding this misleading term in this context. We should point out that we did not perform any 'model fitting'. We calculated our soil moisture balance model with real field data directly entering the equations. This is a difference to models that are based on assumed data, however derived from previous field observations (or text books). Notably, the existing WAB models and even the more detailed soil moisture models like Sheffer et al., 2010, used in our region are based on assumed soil thickness and properties from textbooks such as the Committee on Soil Classification [1979] and were then calibrated through a million runs. In other words: You are right – also these textbook assumptions are based on field observations. Only in that sense did Sheffer 'translate' existing *quantitative field observations* into assumed parameters. We shall make this point more clearly in order to avoid misunderstandings.

Line 37: Eight soil moisture measurement stations for six litho-facies groups (with different topology *[?]* and soil characteristics) are not really much. I would not consider this as "high spatial resolution".

Thank you for your valuable comment. You are right. Soil moisture studies in general cover a wide spectrum of field investigation – from studies with hundreds of sites to studies focussing on only one hilltop – of course always depending on the exact research question.

In general, for a highly heterogeneous area, of course we always need better and higher resolution with respect to time and space. Regionalisation to catchment scale is always a critical and probably the most important step. But in our case, we believe that the number of sites investigated by ourselves does represent the spectrum of material and site characteristics encountered in the area; and therefore, we believe that we are able to cover the spatial variability (and general pattern of temporal responses) of the different of the sub-catchments to recharge events.

Finally we would like to point out that inn our region (the WAB) our study does stand out with its emphasis on dense and long-term field observations (as Reviewer 3 also acknowledged).

Line 79: Indirect recharge originates not only from transmission loss in wadis. Also, the infiltration in cracks and shafts of outcropping karst is indirect recharge. As soon as karst is outcropping, this cannot be neglected. In principle, your idea is ok, but I would not phrase it like this.

Thank you for your comment. We shall revise our manuscript, particularly with the use and definition of terminology. In addition we shall introduce a schematic conceptual diagram explaining and defining the main processes in this study.

You are right. There are many definitions of direct and indirect recharge. And their suitability always depends on the research question.

In our study, we apply the definition of Scanlon (2002) for infiltration and percolation. And with Schmidt et al. (2014) we define deep percolation as water movement below the root zone into the bedrock, equalling recharge - whether into local shallow perched or regional deep aquifers. Direct or diffuse recharge is what we measure at the soil stations. This includes local redistribution from surrounding rock outcrops (in a scale of merely a few metres radius). Indirect, redistributed recharge (Scanlon, 2002), by contrast, are such forms as transmission loss.

Line 92: What about the studies you are referring to in the section line 133ff? Some of them apply spatially distributed models for the WAB.

Thank you for your comment. We shall clarify this statement in our new manuscript. But we should distinguish two different points here. In our line 92 we referred to models that estimate recharge based on surface characteristics. While spatially distributed models of WAB recharge amass, none is based on surface characteristics, processes or properties observed in the field in the framework of the respective research (most of them anyway apply indirect procedures, based on basin outflows and storage change).

The second point is the scale of spatial differentiation: The studies available approach the problem mostly from a regional, catchment scale type of approach, i.e. none of them differentiated into small scale sub-catchment investigations that provide insights into processes within individual formations (their individual soil, lithological, vegetation etc. characteristics, or the actual flux, i.e. spring discharge measurements of individual formations). The spatially distributed models you mention employ regional topographic, geological and soil maps and land use information on an entirely different scale and hence properties and processes (the probably most detailed study, which also employs direct procedures, Sheffer et al. (2010) uses "*fitted recharge coefficients for two classes of lithology (lst & marl)*" over the entire basin are of over 9000km^2 (that is: **two** classes: limestone/dolomite & chalk/marl) - sic!

As already mentioned, we shall revise the manuscript to highlight these differences more clearly.

Line 98: I disagree with the state about the controlling factors/processes. The listed factors are the relevant ones for a saturation excess situation. In semi-arid areas an infiltration excess situation seems to me more plausible (or better both in combination). This means, additionally, you should consider the infiltration capacity. (Consequently, models need to have a high temporal resolution.)

Thank you for the comment. Please allow us to point out two points here.

You are right; conventional wisdom usually attributes infiltration excess as main recharge process to semi-arid climates. However, recent regional studies (in Wadi Natuf and adjacent catchments) have shown that saturation excess is the typical and predominant process at work (Ries, 2014, 2017, Lange et al., 2003).

This was emphasised by Ries in his articles: "*Our results suggest that groundwater recharge is most prominent when single rainfall events are strong enough to exceed field capacity of soil pockets over a wide range of soil depths*" (Ries et al., 2015). This finding was the basis for recharge model by Schmidt (2015) in neighbouring Wadi Aujah, which we modified and employed in our study.

This finding on saturation excess also applies to runoff, even in the semi-arid areas of the adjacent Eastern Aquifer Basin: "*Our results show that runoff in the semi-arid headwater area is strongly related to long lasting rainfall events of high amounts and is predominantly generated by saturation excess overland flow (SOF). Observations from the arid runoff plot indicated a strongly contrasting behaviour with dominating Hortonian over-land flow (HOF)*" and "*temporal runoff patterns indicate that runoff was predominantly generated by saturation excess overland flow during events with high rainfall depths*" (Ries et al., 2017).

Secondly, we would like to emphasise that while Wadi Natuf covers both climates, the main recharge area in the mountains belong to a Mediterranean sub-humid rather than semi-arid climate (compare with Lines 24 and 185).

Since this is an important and somewhat astonishing finding, we shall point this out more clearly in the conclusions and methodology of our revised manuscript.

Line 114-118: Fitted conceptual hydrological also base on reliable field data (to define the parameter range) and not "indirect theoretical values" - whatever this means.

Nobody is able the determine 100% correct values for each point in space. Moreover, none of these models is 100% physical. They all have empirical / conceptual parts, i.e. parameter fitting makes sense.

Thank you for your comment. Generally speaking, you are of course correct and with you kind permission, we shall actually include this sentence in our discussion: "*Nobody is able to determine 100% correct values for each point in space*".

Of course, we are not able to cover the whole catchment with a detailed high-resolution measuring network and therefore considerable uncertainty still remains, typical for natural systems.

But again, we would like to stress that it makes a difference whether one basses his model on real-time field measurements or whether one retro-fits assumed model parameters, thus creating additional sources of insecurity (inaccuracy). Such insecurities only rise with the number of parameters in the model, especially when several parameters are fitted together at once and thus create the typical problems of equifinality. This is why we were careful to choose a simple, robust parsimonious model and abstained from retro-fitting. Of course this does not mean that our results are 1000% correct and reliable, but at least we are confident that in our approach the benefits outweigh the shortcomings. This way we can obtain temporally and especially spatially distributed information that goes beyond what is generally available to date. This has not been achieved in the region so far.

Line 119-121: This statement requires further explanation. For me, it is too general and actually, just depends on the basis of the recharge coefficient determination.

Thank you very much for pointing this out. You are correct and we shall revise our manuscript here. We shall point out that with Cheng et al. (2017), we were careful to model recharge not on an annual but on the event scale (day steps). However, you are correct, annual recharge values and coefficients as a result of such more refined analysis certainly do have their place and indeed were used by us in our analysis and in our comparison with reported results existing in the literature (Appendix).

As suggested by Reviewer 3 (p.4, Line 9), we shall replace "annual values" by '*annual estimation methods*'.

Line 149: In line 74 you have classified this (correctly) as an indirect method.

Thank you for pointing out this misleading formulation. In our new Introduction and Methodology we shall rework our terminology. And we shall insert a new conceptual diagram that demonstrates our system understanding.

As for line 149, you are indeed right – we here describe indirect procedures (based on spring and other basin outflow) and our unfortunate wording, mentioning "direct measurements" may have mislead the reader to believe we speak about direct procedures here. This was not intended. We should have clarified that such indirect, integral procedures are not applicable in strongly leaky aquifers, such as the local perched aquifers in Wadi Natuf because "the problem … is that **leakage** cannot be quantified easily" and independently from both, spring flow and recharge (Line 153). This is also the problem of Gvirtzman's studies quoted directly hereunder. And therefore, this was precisely the reason why we had to resort to "*direct procedures*". We shall reformulate this and use less confusing terms.

Line 187: You are criticising annual recharge coefficients, but using annual runoff coefficients. For my taste, exactly the same problems with recharge coefficients are valid for runoff coefficients.

Thank you for this valuable reminder. As in your comment on Line 119-121, you are right: Annual runoff coefficients as well as recharge coefficients are a valuable tool. And we should and indeed do use them! But as for the process of estimating runoff, we must (analogue to Cheng's demand for recharge) work on the event scale. This was of course done in the previous paper on runoff and transmission loss (recording in 5-minute time steps). Only after we derived and applied event-based values annually, can and should we work with such annual coefficients. We shall revise the manuscript and point out this difference.

Line 250: What does extensive field mapping mean?

Thank you for this comment. We shall briefly indicate the scope of field mapping. We conducted several hundred field visits (many field reports) to refine the general geology, to validate and correct the local lithologies (indicated in the regional maps 1:50,000) and to map the hydrogeologically relevant physical characteristics. These include jointage, primary and secondary rock porosity, degree of karstification, fractures and caves, structural features, bedding geometry and surface crusts, etc. Furthermore we carried out a detailed survey with hundreds of soil samples to determine the prevalent soil type, thickness and texture distribution for each formation and we dug up the soil column down to the bedrock to correlate soil depth to geology and LU/LC (landforms) – see Table A1, Appendix. Last not least, we extensively mapped and measured additional, hitherto undocumented springs, their issuing formation, bedding geometry, spring type, soil and land use/land cover characteristics, etc.

Line 283-284: What does "successfully" mean? The model was already applied in 1992 – lot of development has happened since then.

Thank you for this comment. You are correct; a lot has been published since 1992. But this does not automatically devaluate the article and its findings. In our view, the general type of soil moisture balance approach has not radically changed over the years. Furthermore, the general advantage of the 1992 study is that here, we deal with a karst catchment that is well delineated, relative simple with respect to geometry, continuously monitored since 1950s and well instrumented (for a karst catchment) with groundwater potential monitoring. Therefore it provides a base study for a well-studied and well-controlled karst aquifer flow system, which still is rarely the case…

Footage 2 on page 7: Potential ET is only energy limited. Why you are referring to precipitation characteristics?

Thank you for this comment. Indeed, the Hargreaves-equation only uses temperature as input. We shall clarify this in our revised manuscript. The reference to precipitation patterns was only meant to illustrate the fact that the semi-arid conditions, for which the equation was developed strongly resembles Mediterranean climate, not only in temperatures, but also in rainfall heights and distribution). We shall therefore revise the statement (and remove the footnote as suggested by the other reviewers).

Line 297: You considering the minimum recorded soil moisture as the wilting point. What is the basis for that assumption?

Thank you for this comment. This is yet another example of terminology that shall be addressed. We shall change the manuscript and use the term "minimum SM" instead of "WP" (because, indeed, we did not measure the matrix potential of the soil).

Such fully dried soil after the long summer season serves as the starting point (set as zero) of soil moisture accumulation during the winter season. It should be noted that the empirically measured SM_{min} was a stable value throughout the seasons and can thus be taken as a physical in-situ soil parameter, rather than a passing state in our time-series data.

Line 303: You are talking about "bypass flow through preferential pathways". Earlier you talked about directly exposed karst (line 183-184) and soils with high clay content and cracks during dry season (line 225). These are clear signs for indirect recharge processes and preferential flow. However, your model does not account for this.

Thank you for this comment. You are correct; in Lines 225 and 303 we address desiccation cracks forming in the soil over the summer. However, we did not refer to karts alone in lines 183/184, but simply informed about the diversity of Wadi Natuf when stating that here all formations (of the WAB) have outcrops and therefore the Wadi offers a great condition for the study of the spatial variability of recharge.

As already stated before, we shall revise the manuscript and detail the definitions of terms used in this study. We would define direct recharge as an in-situ process that allows the infiltration and percolation of <u>locally available</u> water (rain, then soil water). If such locally precipitated rainwater infiltrates the soil in situ through desiccation cracks, we do not see the condition for transfer or concentration of water from other places. This of course is a matter of scales. Indeed, as has been studied and published in the region (Sohrt et al., 2014; Lange et al., 2003) on a very small scale, in a radius of a few metres, within a single terrace, on one single rock bank or within one single soil pocket we do find concentration and transfer of rainwater. However, according to our definition and to the scale of this study – the reference unit of which is the formation outcrop – this does not account for indirect recharge, as opposed to transmission losses of runoff that is transported and concentrated over kilometres of length and many square-kilometres.

In addition, we should pint out again, that the in-situ bypass flow through desiccation cracks is fully accounted for by our model (see results).

Hence we shall rework and reword our manuscript to point out this difference in definition.

Line 313: So far as I understand, this is a simple one bucket model. If the bucket is full, you get recharge. It seems to me that you not distinguish whether soil moisture is stored at the surface or in e.g. 50 cm. As long as there is enough water (it doesn't matter in which depth) it will be evaporated by potential evaporation. If so, this is far away from real hydrological processes and also not state-of-the-art.

Thank you for this interesting remark. Indeed we use the standard one-layer ("one bucket") model, also applied by comparable studies in the region (Sheffer, Schmidt, etc.).

However, we did measure at two or three different depths simultaneously. The comparison of the different sensors shows that during the first weeks, up to one or two months of the early winter, a sort of 'wetting front' slowly progresses downward into the deeper soil layers. Later in the winter and in spring, as expected, deeper layers indeed show soil moisture contents higher than those at the surface. Also, the green season is winter and spring in this climate. Thus transpiration becomes more important, whereas in summer and autumn, soil evaporation dominates among the dried up dormant plants. Through the roots, plants can mobilise water from deeper horizons and bring it up to transpiration. Our combined graph then refers to both, deeper and more shallow soil layers, or in other words, both, transpiration from deeper roots as well as evaporation from the surface.

You are of course correct that this is a simplification of the real processes, as intended by our parsimonious model. Ours was a conscious decision not to over-parameterise our model, because the advantages are bought with problems of equifinality etc. (see above)

As Ries et al. (2015) noted: "The calibrated soil hydraulic parameters of our model should be treated as effective parameters that represent both preferential and matrix flow components. Despite increasing work on (preferential) water transport in heterogeneous porous media, there is still no convincing integrated physical theory about non-Darcian flow at the scale of interest (Beven and German, 2013). And even if such a theory existed, measurement problems in natural clay soils would restrict its application to laboratory monoliths. From this perspective, the **use of a simple model** with a minimum number of calibrated parameters seemed to be a **valid compromise** to infer statements on groundwater recharge from a limited number of measurements in the unsaturated zone."

In other words, it would have been impossible to employ our advantage – the strictly field-based approach that Reviewer 3 appreciated – and at the same time introduce such a degree of parameterisation as HYDRUS requires. Admittedly, this comes with strings attached and brings a certain degree of inexactitude and inaccuracy into play, but as you mentioned before: "Nobody is able the determine 100% correct values for each point in space"

Line 371-372: From my point of view, the temporal pattern, i.e. interannual variation, of precipitation is one of the very crucial factors for recharge processes especially in semi-arid environments. I would not separate it.

Thank you for your valuable comment. We of course respect your esteemed expert opinion, but again, we must point to the fact that we are in a trade-off. On the one hand we definitely agree that inter-annual rainfall variations play a pivotal role in recharge, besides the factor that stands at the centre of our attention: spatial distribution and variation of recharge! On the other hand however, the strength of our paper – at least in our view – stems from the fact that it is solidly anchored in real time field observations (over a period of seven years). It would of course be desirable to have 30 years of real-time field records at our disposal; and it is by no coincidence that reviewer 1 and you have pointed out that for a long-term observation, a period of 30 years is standard and required (see your comment to Line 372-373 below). We do not believe that we can do justice to a reliable analysis of inter-annual variation with data spanning only over 7 years. This is why we point to a follow-up paper, which will, by contrast to this paper, not be banked on empirical field data.

Line 372-373: There is a clear definition that min. 30 years are needed to represent climatic condition.

Thank you for this follow-up comment. Of course, you are correct about the international 30-year standard, especially when one is talking about climate variation and climate change. Again, we have the impression that maybe we formulated in an ambiguous way that invited misunderstandings (and shall revise the manuscript accordingly). By no means did we want to claim that our seven-year record of recharge variation can be considered a long-term reliable average. All we wanted to remark in this line was that the rainfall during these seven years happened to cover the span of long-term variation – in rainfall, not in recharge! We shall address this point in our revised manuscript and allow us to add a separate analysis on this point attached to this file. An assessment of the long-term average recharge in Wadi Natuf (or the WAB) is not the purpose of this paper.

Line 380: What about an objective measure for the goodness of fit, like NSE, KGE etc.?

Thank you for the comment. You are correct and we shall include such statistical measures comparing modelled with observed soil moisture.

Line 392: Can you quantify this correlation? Btw, from my point of view, the correlation is not an ideal measure. Again, I would propose to use the common ways to compare data series in hydrology: RMSE, NSE, KGE etc...

Thank you for the comment. We shall also include a correlation measure on the periods of DP vs. Q in the revised manuscript.

Line 401: I am not so sure what you mean by "correlation". Your visual impression that two data series fit to each other?

Thank you for this comment. You are correct. We did not correlate the time series statistically but only visually compared the time series with each other. We shall include such statistical analysis in the revised manuscript.

Line 440-441: I not really understand this sentence. Earlier, you stated that you do not calibrate. However, "this was done in order to minimize the bias" sounds like fitting for me. And which bias do you mean?

Thank you very much – you touch upon an important aspect of the review that also was mentioned by the other reviewers. With respect to the immediate wording here; the reviewer is correct. The term bias is misleading. We shall revise the entire section of the "ranking", not only with regards to the problematic terminology and structure, but also to some aspects of the content (see also Review # 1). We produced three different sets of RC-values for the un-modelled formations, independently from each other and each according to one of the three main aspects – geology, soil thickness and LU/LC ('landform'). This has nothing to do with calibration, nor with fitting. We simply wanted to state that we reduced the arbitrariness of RC attribution by testing several alternatives. The aim was to select the most realistic alternative and also show the margin of error between the three alternatives. But as already mentioned, we shall revise the whole section thoroughly.

Line 488-489: What does representative means? You should provide numbers that the reader can follow your ideas and conclusions.

Thank you for this comment. The reviewer is correct – the term should have been explained and illustrated. We shall add calculations that present how representative these springs really are, quantitatively speaking. We attach to this file an overview over some statistical analysis of key date measurements, showing that we found a standard deviation of 1.15% and 0.72% for the two spring groups, respectively.

Line 545: Why you consider the soil moisture measurements, FC and thickness as representative? What is the basis for this assumption?

Thank you for this comment. Please allow us to explain. We documented our findings in Table 1.a. and Table A1 (Appendix).

From Table 1.a. (calculating field capacity/soil thickness) follows a range of FC-values between 19% and 39% as minimum and maximum values, respectively. The average and median values for all stations lie at 29.8% & and 32.2% respectively. This is in congruence with findings in the literature for the region. For example, Sheffer et al. (2010) base their assumed soil values on Dingman (1994) and quote values of 30% and 32%, for terra rossa and rendzina soils, respectively. By contrast, Ries et al. (2015) did not assume but measured soil water contents in the field and show values between 22% and 30% in their soil moisture plots (Fig. 4 in Ries, 2015). As we explained in the text, we not only measured 8 locations with 2-3 sensors each as time series, but also collected hundreds of soil samples for several seasons. (The sampling shall be explained in more detail in the revised manuscript – compare our answer to comment on line L 250).

With regards to the sensitive factor of soil thickness, we would like to point to our Table A1. Soil depth matrix. After mapping and weighing the distribution and prevalence of different landforms (both, land use and land cover), we designed our field probing campaign in a fashion to probe all important types of landforms for each of the different formation and hydrostratigraphic groups, such as Lower and Upper Aquifer, etc. The table indicates our SM-locations as well as typical landforms (shaded in grey). We believe that this tedious effort gave us good control over typical distributions of soil thickness under different field conditions and not based on literature but on first hand empirical field work.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-329, 2018.

Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-329-RC2, 2018

Interactive comment on "Field-based groundwater recharge and leakage estimations in a semi-arid Eastern Mediterranean karst catchment, Wadi Natuf, West Bank" *by* Clemens Messerschmid et al.

O. Batelaan (Referee)
okke.batelaan@flinders.edu.au

Received and published 3 August 2018

Answers by: C. Messerschmid, J. Lange and M. Sauter (18 October 2018)

Review:

Field-based groundwater recharge and leakage estimations in a semi-arid Eastern Mediterranean karst catchment, Wadi Natuf, West Bank; Messerschmid et al.

I have mixed feelings about this paper. In principle the broader background problem/issue (spatial/temporal recharge estimation of semi-arid mountainous areas) of this paper is interesting and important. Water management in the study area is also a big issue, hence the topic of the paper is good. Further, it is commendable that this paper aims and is based on extensively collected field data.

However, the introduction, presentation of the data, methods, results is far from mature. The paper needs a big overhaul of its structure (i.e. where in the paper is presented what), the logic and analysis support, and finally a detailed and broader discussion linking it to state of research in this field of research.

Thank you very much, estimated Prof. Batelaan, for the time you invested and for your constructive review and comments that are very helpful to process the manuscript.

We do understand that some of the passages are difficult to follow. This is however also a matter of the subject, i.e. the complex recharge mechanisms in a spatially distributed parameter field, in a semi-arid climate, with rapid recharge and complex runoff processes.

Here we shall address your comments and restructure our manuscript for better legibility. We shall thoroughly revise the state of the art section; provide a new paragraph and graphics illustrating our methodological approach and concept of data analysis.

I got several times confused or could not follow the argumentation, a lot of work will have to go into making the text a lot more clearer.

We shall provide a conceptual diagram illustrating our understanding of the recharge mechanism in the area and a flow diagram of the different work steps, data analysis, modelling procedures and calculations.

Moreover, the used terminology is often not very precise. It is difficult from the many tables to grasp the significant points. The figures and the captions could be improved, potentially more figures instead of tables could help to clarify the paper.

We shall follow your advice. Some of the terminology is a result of the complex recharge mechanism, the different disciplines involved (hydrology, hydrogeology, geology) and the specifics of the region.

I have provided in an attached marked up version of the paper many more general and detailed comments.

Thank you very much for your efforts. This is very helpful.

Please also note the supplement to this comment:
https://www.hydrol-earth-syst-sci-discuss.net/hess-2018-329/hess-2018-329-RC3-supplement.pdf

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-329, 2018.

Comments hess-2018-329-RC3-supplement.pdf

Page: 1

No.#1 not clear from the title what this is, i.e. vague

We shall reformulate the title and unambiguously. The new title is now:

Spatial variations of groundwater recharge in a Mediterranean leaky karst aquifer system – parsimonious modelling and field-based estimations in Wadi Natuf, West Bank.

Page: 2

No.# 1 unclear

This passage will be reformulated. (We wanted to express that simulated recharge periods were compared to the temporal patterns of spring hydrographs.)

No.# 3 vague, unclear. Moreover, an aquitard is by definition 'leaky'

Thank you, we shall work on the terminology and introduce the concepts early in the article – such as here: aquitards and aquicludes.

(Strictly speaking, there is no such thing as pure "aquicludes" in nature. However, in this region, the term is often used and attributed to some of the formations we study. Therefore we shall define "aquicludes" here as formations with no or negligible leakage, whereas "aquitards" shall be defined as formations that allow significant and measurable amounts of downward leakage from the aquifer formations above.)

Page: 3

No.# 1 I strongly have the feel that the Introduction needs some restructuring. It does not have a logical structure. It starts too specific - detailed.

Thank you. Yes, we shall rewrite the Introduction (and Methodology) with special emphasis on the flow of arguments. We have been working on this issue already, i.e. to illustrate the objectives, methodological approach, as well as the data evaluation in a more concise way

-What is the bigger picture problem here? Who has worked on it and what is not solved? The literature has not been optimally used for this.

cm: Thank you for the comment; we shall revise the manuscript thoroughly.

-You are going to work on that scientific gap/problem in more detail in a specific area. What is the general characteristics of this area? You already need a figure to support this.

Thank you for the comment; we shall revise the manuscript thoroughly. We shall try to stick to the general features here and otherwise refer to the appropriate section under chapter 2. Area ("as will be outlined in the description of the study area, …").

-What is problem in this area; this leads to a specific hypothesis with a scientific question. A schematic figure could help in making this clear (I imagine something with a cross-section with the deeper aquifers in it but also the hill-slopes with there perched aquifers, underlying aquitards; arrows (?) for recharge and leakage processes.

Thank you for the comment. You are correct. We shall insert a conceptual diagram that explains some of the processes (and terminology) used in the article. Otherwise, more detail will be given in Ch. 3. Methodology.

In your writing, especially terminology, you should also try to be more accurate, many sentences seem vague or use a series of words which confuse.

Thank you for the comment; we shall revise the manuscript in this respect (introduce a conceptual diagram and references to the literature).

No.# 2 Why a range?

Thank you for your comment. We shall clarify that different definitions of the boundary conditions and area size of the basin exist in the literature (especially with respect to the area south of the Afiq channel near Gaza, and with the Sinai portion of the WAB…)

No.# 4 vague, please be more specific

Thank you for this comment. Yes, indeed, we can be more specific: The WAB with a current average use of 389 mcm/a provides 18.261% of total Israeli bluewater production (from wells, springs and surface water Palestinians enjoy only a 6 %-share in basin abstractions.

No.# 5 what do you mean by 'intermediate local'? Better not to use such descriptions.

Thank you. We will introduce the concept of the local perched aquifers more thoroughly. Here, in the introduction we will give a general characterisation and otherwise refer to the details under Ch. 2 Area and 3. Methodology. (We will avoid the misleading term ‘intermediate’ for the aquifers positioned between the Upper and Lower regional aquifers.

No.# 7 aquitard are leaky by definition. 'leaky aquifers'?

The aquifers overlying aquitards are defined as ‘leaky’. The springs in the perched aquifers are contact springs. However, you are right, they are not fed by the aquitards but by the leaky perched aquifers above.

No.# 8 Normally the first fig should be Fig. 1. What is Ch. 4 referring to?

Thank you for your comment. We shall take out the reference to this figure. Thus all figures will be first mentioned in order of their numbers.

No.# 11 why 'intermediate'

Thank you for your comment. We shall produce a conceptual diagram that demonstrates the system characteristics, and properly describe the in the text. More details will be found in Ch. 2 Area.

No.# 13 Not clear what that is without a proper geological description

Thank you. We shall try to only refer to general features here and discuss the details – like regional aquitards according to the nomenclature – in Ch. 2.

No.# 15 not clear what that is, needs explanation.

Thank you we, will duly introduce and reference this term (“effective precipitation” or HEP, “hydraulically effective precipitation”).

No.# 16 Where is this? The area is not introduced, refer to a map?

Thank you for this comment. This was helpful. We shall refer to Fig.1 (under the current numbering system).

No.# 19 water

Thank you for this comment, we shall insert the word ‘groundwater’.

No.# 21 vague

Yes, thank you for this comment. We shall give a more detailed description of Radulovic’s methods.

No.# 22 not clear what this refers to

Thank you. We shall clarify: “*to understand these recharge processes conceptually*”

No.# 25 'Recharge infiltration' is not a good term. 'recharge' and 'infiltration' are different processes.

Thank you. You are right. We shall work on the terminology, as stated above. Introduce our terms more properly in the beginning and stick to them throughout the article.

(We define infiltration as transition of water into the soil, whereas deep percolation stands for the movement of soil water beyond the root zone and soil layers into the bedrock (usually the unsaturated zone. This deep percolation constitutes ‘potential recharge’ as the water has not yet entered the water table in the aquifer, which would be ‘actual recharge’.)

Page: 4

No.# 1 not explained before

Thank you. We shall explain RC (Runoff coefficient) upon first mentioning.

No.# 3 'landform' is not appropriate as a description for LU/LC change

Thank you for the comment. You are correct. We will refer to LU/LC instead of 'landforms'. (Much of the area is not under use, so the term will stress more on natural topographic features like relief and natural vegetation than on anthropogenic features such as agriculture and other land use.)

No.# 7 in this area? or in general? The last would not be correct.

Thank you for this comment. Indeed we were referring to the study area, or wider area (like the WAB and Dead Sea environs).

No.# 9 you mean probably something like 'annual estimation methods'.

Thank you, yes, we shall replace the term.

No.# 10 in the case of

Thank you. We shall reformulate this.

No.# 11 this sentence needs rephrasing, too complex. What is 'temporal distribution pulse', i.e. not well defined.

Thank you for the comment. You are correct. We shall see to that.

No.# 13 s

Thank you, we shall correct this.

No.# 14 do you really mean accuracy? From what follows it seems not.

Thank you for the comment. You are correct. Abusaadah probably also meant 'reliability'. (However, when he stresses the need for data refinement, he also refers to accuracy.)

Page: 5

No.# 1 [he deleted: "*and aquitards*"]

Yes, thank you. Of course, Peleg & Gvirtzman try to understand processes within the aquifer. However, in their text, they also stress the pivotal role of the aquitards beneath: "*Perched springs in nature emerge from aquifers laying on aquitards within the unsaturated zone*"

No.# 2 of

Thank you, we shall correct this.

No.# 3 by

Thank you, we shall correct this.

No.# 4 what do you want to say with this?

Thank you for the comment; we shall explain this quote more thoroughly: Two exceptions stand out. Peleg & Gvirtzman (2010) and Weiss & Gvirtzman (2007) studied recharge in perched local aquifers and aquitards on small-scale erosionally isolated groundwater catchments, using spring discharge[4] from these units and applying it to numerical models, however, without accounting for downward leakage in the multi-layer systems or explicitly excluding such leakage "*by treating the bottom unsaturated layer as if it is saturated*"(Weiss & Gvirtzman, 2007).

No.# 6 aquitard

Thank you for the comment; Yes, into the underlying aquitard (*through the bottom of the perched aquifer…*)

No.# 6 [deleted: "*generation*"]

Thank you for the comment; we shall revise the manuscript thoroughly.

No.# 8 How much? Which station, years?

Thank you for the comment; we shall revise the manuscript thoroughly.

No.# 10 Start your study area description with this type of more general description

Yes, thank you. We shall see to that.

Page: 6

No.# 1 F

Thank you for the comment; you are correct. We shall revise the manuscript accordingly.

No.# 2 You jump in the order of your Tables. Provide the tables in the order that you reference them.

Thank you very much for the comment. In the final paper, the tables will be inserted IN the text at the appropriate location. We shall make sure that we adhere to the appropriate order of the tables.

No.# 3 exceptional: I think you mean something else

Thank you for the comment; we want to describe the sparse occurrence of coniferous forested areas

No.# 6 You only very roughly, descriptive, show that there is a correlation; too base on that a conclusion that you can regionalize the recharge is too strong.

Thank you. We will revise the manuscript thoroughly and explain this in more detail. In addition, we shall add references to already published papers…

Page: 7

No.# 1 better 'recharge area'

Thank you, we shall change this into “recharge catchment areas”.

No.# 4 what do you mean?

Yes, thank you; the procedure of key date measurements at the springs shall be explained more thoroughly, together with visual material (graphs or tables).

No.# 7 unclear

Thank you for the comment; we shall revise the manuscript thoroughly and explain the procedure in more detail.

No.# 9 rephrase

See above…

No.# 12 why?

See above...

No.# 15 not explained before

Thank you for the comment. You are correct. We shall explain abbreviations upon first mentioning.

No.# 16 a data record of 13 years of ....

Thank you for the comment. We shall revise this and make clear that we do not speak of 13 years of measurements, but of a set of 13 SM-measurements.

No.# 19 a time resolution of half-hour.

Thank you. We shall revise this.

No.# 23 no footnotes, include or do not include it in the text

Thank you for the comment; we shall revise the manuscript and insert this list into the main text.

Page: 8

No.# 2 The ranking procedure is not clear

Thank you for the comment; we shall revise the manuscript thoroughly over this point. We will present our work differently and also using a different terminology.

No.# 3 have

Thank you; we shall revise the script.

Page: 9

No.# 1 I lose you in this paragraph, I cannot follow well the methodology, correlation, scenarios, etc.

Thank you for the comment; we shall revise the manuscript accordingly and shall change the terminology (‘ranking’, ‘correlation’, ‘scenarios’, etc.). We shall also revise our tables and graphs accordingly.

Page: 10

No.# 1 This is really a results and discussion section. I would strongly suggest to separate the results and discussion in order to get a clearer presentation and better discussion (strength-weaknesses of approach, comparison with other studies, upscaling, etc)

Thank you for the comment; we shall revise the manuscript accordingly. We might actually change the entire structure and add a chapter on Discussion.

No.# 4 should be quantified

Thank you for the comment. You are correct. We shall see to that and discuss the quantitative findings of Figure 4 in the text.

No.# 7 should be quantified

Thank you for the comment. You are correct. Again, we shall quantify the findings in the text.

No.# 9 quantify

Thank you for the comment. Again, we shall present these quantifications in the text as well.

Page: 11

No.# 2 This is a method description and does not belong in the results section.

Thank you for the comment; we shall revise the manuscript and structure thoroughly. The description will move to the Method section.

Page: 12

No.# 2 This is more a method description and is partly a repetition of earlier explanations.

Thank you for the comment; as above, we shall revise the manuscript and structure thoroughly. The description will move to the Method section.

Page: 13

No.# 1 You should discuss this

Thank you for the comment; we shall revise the manuscript and structure thoroughly. We shall already discuss some the concept of the leaky aquifers in Ch. 1. Introduction (together with a conceptional model graph) and in Ch. 3. Methodology.

No.# 3 rephrase sentence

Thank you; we shall rephrase the sentence: "The study demonstrated that it is possible to apply empirical approaches based on field-measurements in order to estimate spatially differentiated recharge in the Western Aquifer Basin."

Page: 18

No.# 1 What is the difference between the first two lines?

Thank you. We shall explain that the 1st line is number of years (#), while the 2nd line refers to the period of recording

No.# 2 abbreviations should be explained in the caption. Are the values averages for the years measured?

Thank you. You are correct. We shall write extended captions (and we shall keep the abbreviations to a minimum and explain them properly). We shall also modify the table as follows:

- No, the values do not show the averages; instead they present stable physical soil properties that remained unchanged throughout the measurement period (such as for example installation depth of sensors).
- In the case of "SM peak", we hereby present the absolute maximum recorded SM.
- As suggested by the other reviewers, we will change the term "WP" and instead use "minimum SM". This is because they rightly remarked that SM_{min} is not always equal to WP. It should be noted that both, FC and SM_{min}, are absolute values, which remained stable over the years.

Page: 19

No.# 2 This should be table 2

Thank you. You are correct. We shall revise the table numbering.

Page: 20

No.# 2 not clear what these are

Thank you. We shall explain in the captions (and in Ch. 3. Methodology) that these abbreviations refer to rainfall sub-catchments.

No.# 4 avg?

Thank you. As already noted, we shall revise and keep abbreviations to a minimum – and replace them by words (such as this symbol for average)

Page: 21

No.# 1 unclear

Thank you for your comment. As already noted above, we shall modify our procedure and terminology such as for “ranking”, “scenarios”, etc. This will be explained in both, the manuscript and the Table captions.

The captions (and Methodology) will also note and explain that Wadi Natuf surface catchment extends over two groundwater basins and that this table refers to all of the Natuf surface catchment, therefore covers parts of both basins (WAB and EAB) - see also map Figure 1b.

Page: 22

No.# 1 Why this order of alternatives? Why vertical line between Alt.-2 and -1

Thank you for this comment. We will modify this table as well and explain some details of the table in the captions (the procedure in Ch. 3. Methodology).

Page: 25

No.#1 See my remark on the Introduction

Thank you for this comment. As already noted, we shall modify the entire figure. And we shall modify the small inlet figure at the top right corner and use it in the Introduction as a schematic conceptional figure to present our overall approach.

In addition we shall separate the figures and thus resize it and make them better readable.

No.# 2 Legend is probably not readable

Thank you; you are correct. The figures shall be changed accordingly

No.# 3 It would be better to have a legend

Thank you; you are correct. The new version will have a legend

No.# 4 not readable

Thank you; you are correct. We shall change this.

No.# 5 I think this figure wants to show too much. I would split it.

Thank you; you are correct. The figures shall be split.

No.# 6 s

Thank you; we shall revise the caption accordingly.

No.# 7 There is no a and b in the figure indicated

Yes, you are correct; thank you. The references for a and b shall be indicated.

Page: 26

No.# 1 a and b is missing

Thank you for this comment. You are correct; a and b shall be included here as well.

No.# 2 What is UBK?

Thank you for your comment. We shall explain the abbreviation UBK in the chapter Methodology (and briefly in the captions here).

No.# 3 F Change this everywhere

Yes, thank you. As already noted, we shall change all the word ‘formation’, when using it as specific name, not only as a general term in to ‘Formation’ with capital F. (The existing literature is divided on this point.)

No.# 4 s

Thank you for this comment. However, we here show only one inlet photo (so, a plural ‘s’ would be misleading).

No.# 5 o

Thank you for this comment. We shall modify accordingly.

Page: 30

No.# 1 Scale of map, north arrow is missing

Thank you. You are correct, the scale & north arrow shall be added.

No.# 2 I am unclear what this sub figure is showing

Thank you for the comment. You are correct. We shall elaborate on this in the figure captions: the inlet figure shows the hydraulic boundary conditions for the spring group Beitillu. We shall explain the captions (and in the text) that most hydraulic boundaries are no-flow boundaries that run along the line of lithostratigraphic changes from permeable to impermeable; only a very small portion of the hydraulic boundaries (and here the small part indicated in blue) represent flow boundaries. We here set it as a stream-flow boundary according to the prevalent local groundwater flow direction.