Dear Reviewer. Thank you very much for your insightful and detailed comments and recommendations. Please allow us to answer as follows:

ANSWERS

Specific comments

a) Methodology. Wadi Natuf, prior to our work, was not "mostly poorly" but entirely ungauged. No rainfall, runoff, spring flow records, let alone soil moisture data were available.

b) We concur, thank you.

c) Novelty of our approach. Indeed, many, if not all of our individual tools and methods applied for Wadi Natuf are standard, well tested, used and published. Two things make our approach novel: First, a fully distributed approach was followed. Second, this is one of the many ungauged basins worldwide, which however are underreported in the literature. We specifically follow the recommendations of the IAH decade on PUB, which have described and prescribed in general terms but not been implemented before under such circumstances as found in wadi Natuf. (Most PUB applications deal with runoff, rather than recharge. Exceptions that tackle recharge, such as *Aish et al.* in Gaza, deal with gauged basins, where large amounts of subsurface and water level data are available and drawn upon in the study.) Our paper is a strong departure from this state of the art. And thirdly, maybe most importantly, it is the unique combination of individual tools, their application and synthesis in this one basin framework of ours, which is novel. Fourth, our results base on empirical field work, not calibrated models. So, yes, we agree with your comment that one novelty is the work in ungauged basins, but the other two factors should not be forgotten or underestimated either. We will try to make this more clear (abstract, introduction and conclusions). [If you are interested in a detailed discussion of each of these articles, on the *already proposed and formalized values of distributed recharge*, we can send you an additional file upon request.] At the end of our rewritten chapter Introduction we list seven different points that – taken together – characterise the novelty of our approach: The model is fully distributed, not lumped or semi-distributed [1]. It is a recharge model, not simply a rainfall-runoff model as in other studies before [2]. It concerns a hitherto ungauged basin [3] of deep karstic aquifers [4]. *It employs a new combination of* existing techniques that are based on observable processes, parameters and signatures with the help of a specific basin classification framework [5]. Our assessment adheres to the goal of parsimony [6] and integrates inductive and deductive steps. And, with the observation of processes deep underground being made impossible (by the occupation), it is grounded in surface and surface-near observations (so-called direct approaches) [7].

d) Length of the manuscript. Thank you for this comment and your recommendations. Partly the length and detail of the introduction is the result of the earlier reviews of the first manuscript, which then was split up into two different articles (This draft, reviewed by you, forms the second of the two articles.) We will try to shorten, condense and reduce the amount of information in the Introduction. (Real cuts cannot be done without also dismissing some of the content of this chapter.) So, a compromise will be necessary and attempted by us.

e) Ambiguity of the term “ungauged” – Dear reviewer, indeed, we also define and understand “ungauged” as lacking any flow or other hydrological information (incl. meteorology). We do not speak about special cases in karst, where some gauging data may be available. Where refer to “hundreds of wells” – deep Israeli wells – that are gauging the basin, we specifically refer to the downstream Coastal Plain area of the basin inside Israel (in the borders of 1949) only. By contrast, the upstream recharge areas, located in the occupied West Bank remain fully undeveloped, untapped by wells and thus ungauged. This is the study area of Wadi Natuf. All of Wadi Natuf was ungauged, prior to our study and field work. Such contrasts between gauged and ungauged regions with a basin are not only very common worldwide, but are especially relevant for the question of distributed recharge, which has to be investigated in the – typically ungauged – portions of otherwise already developed and used aquifer basins. In the case of Israel and the West Bank under prolonged military occupation, this contrast is particularly stark and it was specifically the underlying motivation of this study to try help close the gap, if not development and use, then at least in investigation, research and understanding.

f) “Classification and transposition into Recharge coefficients”. Thank you, yes, we understand your point and why there seems to be some inconsistency in the way we portrayed our methods and approach. We will clarify this, including changes in the tables presented… (see also, Specific Comments).

Specific comments

1) “*Abstract: Unless authors are directly focusing on ungauged AND karstic basins the sentence "relatively little attention has been paid to its spatial distribution" seems somewhat exaggerated.*” → As already discussed in the general comments, the novelty of our approach is the novel COMBINATION of otherwise well-established methods and procedures. We combine empirical field work with deductive methods. We work on an entirely ungauged basin – ungauged in any form, whether with regard to precipitation, runoff, spring flow, soil moisture or recharge. (The only exception is one single deep groundwater well under Israeli control, well Shibteen No. 4). Hence all our measurements were a first step into the unknown. But importantly and in addition, the novel character of our work stems from the fact that we attempted an assessment and quantification of GWR in its spatial distribution: In contrast to the literature suggested by you, our study does not only weigh and qualitatively evaluate certain “descriptors” or “physical factors” of recharge or estimate recharge distribution in relative terms or as “zones” of (weak, moderate or strong) GWR, but actually presents a fully distributed quantitative recharge model and assessment over the entire catchment area.

As to the quoted phrase “relatively little attention” – we would like to stress that the emphasis in this sentence lies on RELATIVE. Indeed, spatial distribution of GWR and its quantification has been much, much less studied than its temporal variability. This is something we fell, we cannot change (only clarify to avoid misunderstandings).

Concerning karst, and although most aquifers are indeed karstified, our method of ‘direct procedures’ tries to understand and quantify the processes of recharge, where they are accessible at the surface (without boreholes or other groundwater information in the underground). As a consequence, our SM-measurements and modelling takes place above the aquifer, at the soil layer and its interface with the underlying aquifer bedrock. Hence, the typical problems of describing the characteristics of karst (such as anisotropic flow) could be circumvented. Therefore, our model is applicable also in karstic rock, but as well generally in all kinds of aquifer types.

First, we will change the manuscript and state more clearly that the novelty of our approach lies in the combination of techniques, used in this study. Secondly, we will somewhat caution and specify our statement, that so far relatively little attention has been paid to spatial GWR assessment and quantification (and in a fully distributed model).

Last, regarding the literature suggested: Thank you very much! Some of the texts were indeed new to us (or were published after our work). However, they do not change our views expressed in the paper. We would argue that each of them falls short of such an approach as followed in our study (or does not even attempt to answer the same questions). We attach a paper with some discussion of the individual studies.

2) P. 1/26: Yes, thank you! We will indeed add a more comparable quantification of our results in mm/a (235 -274 mm/a).

3) P. 1/28: Following the above (1), we would still maintain that this COMBINATION of methods was used for the first time in such an environment and for such purposes (quantifying distributed GWR…). We would also like to stress that this second part of our series of articles on Wadi Natuf recharge cannot be read and understood without the first article. Much of the empirical field work is described in article 1. This second part on the regionalisation of the results of the study (in article 1) employs more deductive methods and conceptual work on the BCF. We were advised to split the originally submitted overall manuscript into two articles (both by reviewers and editors of HESS) – this 2nd part (here under review) is the result of that split.

4) P.2/5: Yes, thank you. We will remove "etc." and change the manuscript into: "their drivers, e.g. precipitation **and** evapotranspiration (Batelaan… "

5) P.2/6: We beg to differ: Throughout the Levant, the mostly Upper Cretaceous/Lower tertiary aquifers are all karstified. And they are all heavily developed, tapped, drilled and abstracted in all states – Lebanon, Syria, Israel and HK Jordan (ESCWA-BGR; 2013: 285) – with the notable exception of Palestine (the occupied West Bank of Wadi Natuf), where well development is rigorously restricted by the Israeli hydro-regime of its military occupation… As stated further below, in the same aquifer, the WAB, the Israeli downstream side is heavily developed and used, while the upstream occupied West bank zone (= Wadi Natuf) remains in an enforced virgin state of well development. So, in all karstic aquifers of the region, karstic aquifers wells are available in abundance – except for the West Bank. (Sentence remains unchanged.)

6) P. 2/36: We have found the term bedrock in the literature not only describing crystalline bedrock formations, but also as a more general term (Nature, USGS, etc.). However, we can change this term if required.

Second, yes, thank you. We change the sentence. Scanlon et al. report rather than recommend such methods (unsat. zone observations).

7) P. 3/7: Geology in our view encompasses more than only lithology. Also, the quoted authors (Sanz et al., 2011) differentiate somewhat between the two. The second mentioning of "lithology" refers to another set of studies (Batelaan, Aish, etc.). It is not our repetition but theirs. (Please note, as already mentioned, we here not only list the three groups but also forms of combination and interaction with each other. Repetitions therefore are unavoidable.)

8) P. 3/9: Again, land use is not simply repeated as a mistake. Rather, the different authors use different approaches, i.e. different combinations of factors (like land use together with other

factors). We wrote: Authors use three principle groups (and use one or two of them). We can clarify the distinction between 3 principle groups and their combination by different authors.

9) P. 3/6: Yes, but do we understand you correctly that you want "fractured density" to be mentioned (as was used by Radulovic)? However, our aim here was not to be as exhaustive as possible in a list of criteria. Rather to the contrary, we aim for simplification and here want to show the three principle groups that are used most frequently. → Is it necessary to add a sentence that lists all other possible criteria? For example, to determine "fracture density", you usually need very extensive ground work, mapping and even underground information. This can never be done for an entire area or catchment, but only and at the best, at some isolated spots in the field... – and such preconditions are very far from Predictions in Ungauged Basins (PUB)!

10) P.3/20: Yes, thank you. We could have specified the type of model: "*In his* ***rainfall-runoff*** *model*, Seibert (1999) developed relationships between the calibrated model parameters and the physical catchment characteristics of landscape found in the field". However, since you asked us to shorten the Introduction, we removed this sentence from the manuscript.

11) P.3/42: Thank you. You are right, where basins remain completely ungauged, it is often difficult to find and identify any such signatures. So, signatures require some, if even poor gauging beforehand (or of course field observations and measurement campaigns during a study). Yet, despite this fact, this is what the quoted PUB-authors recommended. We could alter the sentence slightly and add: *„Although, such signatures are often lacking due to the state of remaining ungauged…*". Would that solve the problem? In our case of Wadi Natuf, after and due to the measurement campaign of our study, we were able to quantify such signatures. The patterns of such signatures were already detected before measurements began, during the stage of field observations and therefore, measurements were designed and carried out accordingly (see also part 1 of our article series). Your comment was considered and has changed in our new, shortened version of the Introduction.

12) P. 4/9: Thank you, we will amend this sentence into: The WAB is **a complex of** up to 1000 m thick Upper Cretaceous carbonatic karst aquifer (SUSMAQ, 2002) and conventionally divided into two regional aquifer layers (Fig. 1) – an Upper Aquifer (UA) of Turonian to Cenomanian age and a Lower Aquifer (LA) of Upper Albian age, (see Fig. 2a in Messerschmid *et al.*, 2019) **– and separated by poorly or non-permeable layers of Lower Cretaceous age (Yatta formation)**.

13) P. 4/19: Ok, we will change into: …monitored and gauged through **a network of hundreds of Israeli groundwater abstraction and monitoring wells (tapping the deep aquifers). By contrast, the slopes** , the WAB recharge and accumulation zones **(inside the occupied West Bank)** remain …

See also general comments

About your comment: "*In my understanding, "ungauged basins" refer to watersheds where stream (or spring)* ***discharge is poorly known*** *or even not known at all. Here, the ambiguity is left because* ***in karstic*** *regions springs can be ungauged, the Wadi Natuf can be ungauged as well but the authors mention ""gauged though hundreds of Israeli deep wells" suggesting that gauged/ungauged refers to aquifer monitoring by wells.*"

We shall explain here that the WAB is well gauged in its **discharge** area, but that almost the entire **recharge** area (and Wadi Natuf) remains completely ungauged (groundwater flow and discharge of springs is not known at all) = PUB.

14) P. 4/End of section 1.5: Thank you for this comment. Indeed, we did not intend to already introduce results of other studies, not least in order to avoid unnecessary repetition (which you criticised on other occasions) and since later, under results and discussion we will come back to this point. It should also be mentioned that in part 1 of the series, we presented findings from other studies extensively (see Table H1 in the Appendix). Maybe, it would suffice here to make reference to this article and Table of comparison (see: https://hess.copernicus.org/articles/24/887/2020/)? As for the results of THIS study ("values for recharge"), we do not think it is proper and necessary to already pre-empt the final results of our study in the middle of the chapter Introduction. So, we only briefly added here a RC range of 30-50% (together with a reference to Annex H of the first paper, Messerschmid *et al.*, 2020).

15) P. 4/51: We would like to state that we cannot mention our BCF criteria in this sentence about recommendations by Hrachowitz. Hrachowitz, in his study did not specifically mention these criteria. To the contrary, the combination in our BCF is our novel combination of such criteria, as will be presented at the appropriate time (methods and results of the study), not already here. Therefore, in order to follow your recommendation, we can only mention again the three groups (but not as groups already identified by Hrachowitz).

16) P. 5/6: Yes, you are right! This is now Messerschmid et al. 2020! Thank you.

17)P. 5/15: Question: Differentiation of recharge classes. – based on which criteria? Answer: The classes were differentiated in two steps: A) Qualitatively, based on observable features and well-established physical principles of recharge processes (for example difference between karstic carbonates and clay; or between forests and sparsely vegetated land forms, and of course between thin and thick soils). B) Quantitative differentiation was based directly on the results of the SM-percolation models (in Messerschmid et al., 2020). See new version of the Introduction.

18) P. 5/34: Impermeable chalk. Thank you – we understand your concern. Indeed, in other areas of the world, "Chalk" is a term for extremely powerful, productive aquifer formations, such as famously in the case of the "British Chalk" – a microporous white limestone formation, which however exhibits rather highly competent rock, in which fractures and fissures remain open and therefore act as excellent flow paths! (see Introduction by BGS - The Chalk). However, in the Levant, the Senonian chalks are almost invariably known as powerful aquitards (even dubbed "aquicludes" here). This is because this chalk is very soft and incompetent. Hence, it does indeed act as a powerful separation layer between over- und underlying aquifers (Eocene and Turonian aquifers), both on a regional and local scale. Few exceptions of some minor aquitardal limestone series within the Senonian formations actually prove this rule. Therefore, in general, as well specifically for Wadi Natuf, we would find it rather misleading to attest a "low permeability" to the Senonian chalks, rather than stressing the completely "impermeable" and separating nature of these aquitards ("aquicludes")…

- "The thick **chalk-limestone** Senonian to Eocene sequence is generally an **effective aquiclude**" writes **Carmi** (1989): https://www.sciencedirect.com/science/article/abs/pii/0022169489901698
- see also: **Avisar** et al (2004), Fig. 2 (modified after Rosenthal et al. 1999) - https://link.springer.com/article/10.1007%2Fs10040-004-0322-8
- see also works of Akiva **Flexer** (TAU): https://link.springer.com/article/10.1007/s100640050030
- and **Qannam** (2003) in Al-Arroub: https://tu-freiberg.de/sites/default/files/media/institut-fuer-geologie-718/pdf/fog_vol_9.pdf
- and **Sauter** et al. (2005), Final Report of GIJP: "*The formations of the Mount Scopus group can **generally** be considered as an **aquiclude**.*" https://www.cleaner-production.de/fileadmin/assets/bilder/BMBF-Projekte/02WT0161_-_Abschlussbericht.pdf

19) P. 5/36: One has to read this statement carefully and correctly (see below). First, we refer to Fig.2b in the previous study, not in this manuscript under review (we do not have a Figure 2b in this manuscript). By contrast, Fig. 2b in the previous article clearly indicates the entire list of existing lithostratigraphy (top left corner with legend in the map). Figure 2 of this manuscript has nothing to with it, as it is meant to present the correlation between LU/LC and geology… Second, we are speaking of the WAB portion of Wadi Natuf only. Thirdly, we present the "conventional view", i.e. only two large regional aquifer complexes, UA + LA. (UA = *Heb, l-Bet, u-Bet & Jerus.*; LA = both Beit Kahil formations, or *l-LBK, u-LBK, l-UBK & u-UBK*). This is NOT the more refined hydrostratigraphy as applied in this study. It is only a brief summary introductive statement about the geology and hydrogeology of the WAB in Wadi Natuf. Roughly two thirds of outcrops of this WAB portion are made up by the conventional LA and UA formations (64,4%). In other words: This does explicitly NOT include the aquiferal portions of middle Yatta formation (as it is considered *The Middle Aquiclude* in the conventional view). Also excluded here is the deep Ein Qinya formation aquifer outcrop portion as well as the shallow alluvial deposits and their share of total WAB are in Wadi Natuf (not part of the "Mt. aquifer" stratigraphic column). For more details, see further explanations in part 1 of the article series. The exact outcrop sizes of each formation are detailed in addition further down in Table 4 (km^2 outcrop of each formation) for the entire Natuf catchment (incl. the Eastern EAB portion). Now, the text is slightly adjusted into:

All formations of the WAB are covered in this study (Fig. 2b in Messerschmid *et al.*, 2020). Together, the aquifer formations cover around two thirds (64.4 %) of the outcrop areas in Wadi Natuf; they are entirely carbonatic and in most parts strongly karstified. The other third of the area consists of outcrops of less permeable and fully impermeable formations.

20) P. 6/5: LBK & UBK comment: "*provide the **significance** of LBK UBK as given in Table 1.*" It is not entirely clear what you mean by that. Are you asking, a), why LBK and UBK form a *major "regional aquifer*" in Tab.1? Or do you ask, b), why they are differentiated as in Tab 1? Or c), is it the word "must"?

Answer to a), LBK (+ UBK) are well known and used as main target aquifer in all deep wells (most productive ftn.) in the West Bank.

Answer to b), l-UBK differs strongly from u-UBK or l-LBK (this is also a practical problem, felt in many deep wells, where these differences were described in terms of different lithology, fracturedness, etc.) And to c), of course, this can be changed: On a local scale, the aquifers CAN be differentiated into different (more or less permeable) parts.

21) P. 6/21: Yes, most of the springs are perennial. And a few (some 10-15%) of them are very small springs that dry out seasonally each summer). More details on springs are found in part 1 of the article series.

22) P. 8/44: Yes, indeed, this correlation is repeated several times throughout the article (like before on page 8, line 17,18). We thought this is a central fact worth emphasizing. But we can also avoid such repetition…

23) P. 9/10: yes, fine. We will change into "mineralogical composition".

24) P. 9/12: Fine, no problem. If that helps, we can list these karstic features studied in the field: The text can be amended as follows: "First, existing geological maps in the scale 1:50,000 (GSI, 2000; 2008; Rofe & Raffety, 1963) were corrected, complemented and refined by extensive field mapping and

remote sensing (stereoscopic aerial photographs) with the target to detect, describe and interpret the lithological rock content, (chemism, texture, grain distribution), the degree of crystallisation and structural features like folding, faulting, cleavage, jointing, as well as primary porosity and karstic features. These karstic features encompass **epi-karst surface features** (karren and schratten landscape), **karstic solution holes** (especially in the l-LBK formation, aka "Swiss cheese" formation and in parts of the Hebron formation), well-known **caves and karstic channels** in the underground; wide, **karstified fractures** (incl. their width and prevalence). In addition, there are **indirect indicators** of sub-surface karst (such as rapid interflow emerging at "bleeding hills" and travertine crusts)."

25) P. 10/2: Ok we will remove the line with "etc." from Table 2 and will amend the text in line 18 as follows: "were inserted 17 in Table 2 **as a generic example, in order** to obtain…"

26) P. 10/15: Thank you for this comment. We can see now that our statement was confusing. What we wanted to say is that both LU/LC and geology are not quantifiable parameters in the sense that forest and grassland or limestone and chalk are not quantitative categories. But of course, the area size of their occurrence can be and was quantified. And in addition, quantitative values can be attributed to these qualitatively different features if needed. In fact, we did attribute different RC-values to these land use features and geological formations as a result of our BCF. We should amend our text here and rather state that **LU/LC & geology features were ranked according to their recharge potential and correlated with soil thickness** (see: ranked classes of recharge potential from low to high, as in Table 2). **This was based on a soil thickness correlation matrix in which representative typical soil depths could be attributed to the different formations and land forms** (this soil matrix is shown in Appendix D, Messerschmid et al., 2020). Regarding your second question, yes, a LU/LC map is available and presented in Figure 2 (p. 7).

27) P. 10/21: Thank you. Yes, indeed - maybe we should formulate it more simply and straight forward without unnecessary details: "Five different formation-specific RC-values were obtained from the monitored and modelled SM-station data (Messerschmid et al., 2020), representing 5 different classes of recharge potential (between 57% and 42%). In addition, impermeable formations were not monitored at SM stations. For them a sixth class of zero recharge (RC = 0%) is added. According to the correlation and grouping in the BCF, the different specific RC-values of the modelled formations could be attributed to other formations as well (see Table 4). By this step, all existing formations in Wadi Natuf were assigned specific RC-values of annual recharge (ranging from 57% to 42% in aquifers and down to zero recharge in the aquitards). It should be added here that for three formations, an additional step was needed. These formations were found not to be uniform but either consist of different sub-facies at different locations of the catchment (e.g. u-Bet and l-Bet) or represent different lithologies in the stratigraphic column (within l-UBK formation, see photo, Fig. 3c, above). In these instances, an additional, intermediate RTC value was introduced, based on the arithmetic mean of classes II and IV (49.4% - as average between 44.7% and 54.1%)."

28) P. 10/19: transition from Table 3 to Table 4: Thank you for your comments. You are right, the text and the tables are somewhat confusing. We added information and restructured Table 3. An additional "class" of recharge coefficients was inserted, not based on direct results of SM-percolation modelling, but as an average of two of the classes calculated by this modelling. (classes II and IV → new average: 49.4%). This applies only to the group (column) "geology" and it is based on the fact that the three formations concerned all display a mix of different facies (either vertically in the rock column or laterally in the outcrops of the catchment). We therefore have to also adapt the text of the manuscript accordingly (see answer to comment 27), above).

Table 3. Conceptual basin classification framework, specific for Wadi Natuf

	group 1 - LU/LC		group 2 - soil		group 3 - geology		
	formations	features	formations	feat.	formations	features	
I			All, **Jer**, l-LBK	--	All, Jer, l-LBK	well dev. karst (& gravel)	Increasing recharge potential ↑
II	u-UBK	cliff, mostly rock outcrops	**u-UBK**, u-LBK	-	Heb, u-UBK, u-LBK	karstified lst / dol	
*				(…)	l-Bet, u-Bet, l-UBK	lst / dol (some marl / chalk) (*Nari for u-Bet*)	
III	All, Jer, Heb, u-LBK, l-LBK	olive terraces, rock outcrops	u-Bet, **Heb**, EQ	-/+			
IV	u-Bet, l-UBK	arable but uncultivated, grass- & shrublands	**l-UBK**	+/-			
V	l-Bet, l-Yat, EQ	mixed, transit. woodlands	**l-Yat**, l-Bet	+	l-Yat, EQ	mixed lst + marl	
-	*(as Gr.2)*	agric. plains, forests	Sen, u-Yat, Apt	++	*(as Gr.2)*	marl (chalk)	

Note: Left column: classes of measured recharge potential (I – V); middle columns: groups of phys. features (1-3); formation names as in Table 1; soil thickness increases from thin (--) to thick (++). The formations shown in bold type were the ones monitored, measured and modelled. The grouping and class distribution was based on field work and literature, e.g. SUSMAQ (2002), LRC (2004), GSI (2001), Keshet and Mimran (1993), Messerschmid (2014) and Messerschmid *et al.* (2018). Aquitards, i.e. impermeable formations, where recharge is assumed zero, were not measured in SM-stations (bottom line of Tab. 3).
Regarding class *: The class marked with asterisk * was not measured in SM-stations. Instead, for group 3 (Geology), the average of RC for classes II and IV was taken (as 49.4%), because these formations appear in two facies types, which are more and less permeable, respectively.

29) P. 10/24: Ok, thank you – we will insert reference to Table 3.

30) SECTION RESULTS: We are not sure what precisely you suggest. Do you want us to transfer some of the information from Section 3. Methods to section 4. Results? (*NB: We re-formatted the numbering in RESULTS to 4.1, 4.2, etc.*) otherwise, fine, we amended Figure 5 and added a map for LU/LC and for Geology; it now shows all three different runs for all three groups.

31) P. 12/6: Fine, we can change into "aquifer" here.

32) P. 13/9: Discussion often starts be recounting the state of art (referring back to the introduction and thus introducing the relevant questions for discussion). Admittedly, this is a matter of personal style and taste. But many articles follow and many tutorials on scientific writing suggest such a procedure. We can of course also simply cut the entire paragraph (line 9-22), if this is really what is demanded by you and the other reviewers and editors.

33) P. 13/21: Oh, yes, of course – thank you! "temperate" not "moderate"…

34) P. 13/24: Yes, thank you. We will reformulate, but what we intended to say here, and referring to the recount of existing works in lines 9-22 before, we wanted to state that our study went a step further *than* **the existing literature**, a step beyond the methods employed already in the quoted literature. This is part of the novelty of our approach. Again, the novelty here is– to the best of our knowledge – the unique COMBINATION of otherwise well-established methods. But this, by the way, is not the only novelty. Other new features are the application in hitherto completely ungauged basin, the strictly empirical nature of our approach (and its depth), as well as the truly distributed recharge analysis, which hitherto has not been performed under such circumstances (karstic, PUB, recharge, not runoff calculations, forward calculating models, not calibrated and retro-fitted modelling, etc.)

35) P. 13/33 ff.: Thank you. We will shorten this paragraph, as follows: "Our three independent runs of conceptual analysis and attribution resulted in a close range of total WAB recharge (24, 26 and 28 mcm/a). This suggests that our transfer procedures delivered a robust and realistic representation of the processes at hand (which we prioritized over allegedly "exact" but less reliable results)."

36) P. 14/5: And hereby, we will focus on the point that and why three different runs were carried out separately for each group of physical features.

37) P. 14/26: Yes, you are right. It should read: "*The results **of the measurements and analysis in part 1 of the series confirm** the well-known fact that the temporal distribution of events (precipitation) strongly affects the percolation rates.*" We also wanted to remark on the side that our daily steps of analysis were the appropriate time scale; we thought this a relevant

issue under the topic of "process representation" (section title 5.1) – but if you deem this unnecessary in this paper, we can also simply delete this statement…

38) End of Section 5.1: Yes, indeed, limitations and caveats are discussed at the end of the Section Discussion, not at its introduction. → We added a new section No. **5.2 Limitations & caveats** (followed now by a new section No. "**5.3**" for "Annual RC – overall basin RC …")

- To begin with, the results of the SM-models, the RC-values of the different formations as input data for our basin classification framework (BCR), are taken as correct and reliable. A discussion of the limitations and caveats of the results and methods can be found under Messerschmid et al. (2020).
- However, the process of setting up a BCF and attributing different classes of recharge potential (RP) to the different physical features (under the 3 groups selected) is a deductive step, which relies on the translation of qualitative observations in the field into quantitative classes of RP. Therefore, the exact classes under the here developed BCF, although based on and rooted in universal physical laws and well-established evaluations, could be somewhat imprecise and incorrect. Some classes could have been selected wrongly and could under- or overestimate certain factors (features) for the decision. This is why we found it imperative to establish three independent runs of classification for the three different groups of indicators, which allows us to weigh and compare and thus evaluate the reliability of the BCF.
- Another theoretical possibility is that some processes, although present in the field, were not detected and included in the set-up of the BCF. However, the approach of this study used the most commonly known principle groups of physical landscape criteria quoted in most of the literature (see Ch. 1, Introduction). Therefore it can be stated with confidence, that the processes covered by our selection of classification criteria belong to the most important, principle processes of GWR and it is rather unlikely that a major process was overlooked.
- The possibility of overlooking a minor process is always and necessarily a by-product of such simplification. Hence, such simplification, a major characteristic of our approach is not only a strength but also a (relative) weakness. However, it should be repeated here once again that the need for simplification of the host of processes at work in groundwater recharge is strongly recommended and explicitly highlighted by the existing PUB-literature (see Ch.1, Introduction).
- In addition, the overall results were also weighed against and compared with similar results from other catchments, especially such in in the WAB and its environs.
- Lastly and although the correlation between the three groups of observable features was clearly observed and investigated in depth within Wadi Natuf, it may be absent in other catchments. This then would pose a limitation to the applicability of the approach chosen. However, in such a case, other correlations can and should exist; they should be studied and detected individually for each other basin, but otherwise following the same approach as designed for this study.

39) Section 5.2: "Literature on RC values in other areas…": Thank you. The studies of Allocca und Marto Rosillo are useful. They **shall be briefly mentioned and their avg. RC values quoted**. But in the other studies we cannot see the applicability for our purposes [further details and discussion are available upon request]. Hence, we **added** to the text and references:

Allocca et al. (2014) quoted average annual RC-values ("effective infiltration") from other countries (Hungary, Greece, Spain, France and Croatia) to range from **35% to 76 %** and of 27 % for Tennessee (dolomites) and **found** recharge coefficients of **50 – 79 %** in their own study in the southern Apennines. **Martos-Rosillo** et al. (2015) present a review of gw recharge studies in Spain. They found spatial variations due to: "*the degree of surface karstification and the development of the vegetal cover–soil–epikarst system in the carbonate aquifers*". "*The recharge may range anywhere **from 7 to 720 mm/year**. The mean coefficient infiltration or recharge rate is 38 % of the rainfall, ranging **between 4 and 62 %**.*"

[**Added to References:** Martos-Rosillo, S., González-Ramón, A., Jiménez-Gavilán, P., Andreo, B., Durán, J. J., & Mancera, E.: Review on groundwater recharge in carbonate aquifers from SW Mediterranean (Betic Cordillera, S Spain). Environ Earth Sci, 74(12), 7571-7581, 2015. https://doi.org/10.1007/s12665-015-4673-3]

40) P. 14/16: yes, thank you. We will write "geology", not "lithology"…

Clemens Messerschmid, Amjad Aliewi (Dec 2021)