

## ***Interactive comment on “Groundwater fauna in an urban area: natural or affected?” by Fabien Koch et al.***

**Fabien Koch et al.**

fabien.koch@kit.edu

Received and published: 30 July 2020

Dear referee,

we would like to thank you for your time and the constructive comments, which helped to improve the quality of the manuscript. Please find our detailed replies on the comments below. We hope that we answer all your remarks.

Referee #1:

This article is an application of an existing method to assess groundwater ecological condition. The article utilises a classification scheme based on a single threshold of proportion of crustaceans and oligochaetes within sample wells, with varying success. The manuscript acknowledges several limitations of using this single method suggest-

C1

ing that multiple methods should be used to fully understand impacts of humans on groundwater ecology. The research presented increase awareness of groundwater ecosystems and the threats facing them, however requires further analysis to justify some of the claims made. As such, I recommend major revision, purely because of the requirement for further statistical analysis.

Response: We partially agree. Thus, we performed a more profound statistical analysis (e.g. U-tests), which are presented below in our replies to the ‘specific comments’.

General comments

Generally, the sections flow well and it is easy to understand. The manuscript needs to be thoroughly edited as there are multiple issues with grammar, and the manuscript can be reduced in length particularly in the introduction. The figures and table are well presented. The methods and results section needs to have some aspects clarified. There is a lack of statistical analysis throughout the manuscript which detracts from the quality of the paper. The results show some interesting trends in the distribution of biota, however without the necessary statistical analysis of this data, it is difficult to establish if there are significant differences between landuses, or if these trends are just due to differences in sample size ( $n_8-n_{31}$ ) between the two landuses. This needs to be addressed, as currently there are speculations that differences in means indicates differences between landuses without any specific statistical analysis. A simple ANOVA or t-test would, in most cases, suffice and allow a more thorough analysis of this useful data.

Response: We agree on the grammar issues. Hence, the manuscript was carefully revised by a native speaker to ensure correct English.

We agree on the length of the introduction. Several sentences were shortened or even deleted (for more details see the marked manuscript attached).

We partially agree on the statistical analysis as already mentioned earlier, which is

C2

detailed below (Comment #12).

#### Specific comments

Comment #1: Introduction: In general the introduction is a little too long and can be made more concise. Eg paragraph starting line 45 and line 50 could be compressed and merged.

Response: We agree on the length of the introduction and the conciseness. Thus, we shortened or deleted sentences, e.g. we condensed the paragraph starting in line 47: "A closer look at the German federal state Baden-Württemberg is given in the study by Hahn and Fuchs (2009), which focuses on defining stygoregions based on different hydrogeological units. They conclude that the observed patterns of groundwater communities reflect a high spatial and temporal heterogeneity of aquifer types with respect to habitat structure, food, oxygen supply etc."

Comment #2: Line 35-37: Whilst these may be the usual temperatures for stygofauna within the region of this study, they exist in temperatures well over 14-16 deg on a global basis. This sentence needs to be rephrased.

Response: We agree on the reformulation of this sentence (lines 35-37). Nevertheless, it was not possible to find international values. Thus, we clarified the spatial reference of the values in this sentence: "Hence, they are assumed to be cold stenotherm, which means that they prefer cold temperatures and can hardly persist water temperatures over 16 °C (Brielmann et al., 2009) or rather 14 °C (Spengler, 2017) in Central Europe for a longer time period."

Comment #3: Line 38: Remove 'the' from "the German and European legislation"

Response: We agree. Done.

Comment #4: Line 44 remove 'data recorded by' in the brackets

Response: We agree. Done.

C3

Comment #5: Line 54: Typo error (Protocol for the Assessment: : ..

Response: We agree. Done.

Comment #6: Line 83: Korbel & Hose 2011 is correct reference, also consider Di Lorenzo et al. 2020 Ecological Indicators, 116, 106525.

Response: We agree and now consider this new study in the manuscript. Hence, we added the following paragraph (see lines 91-93 in the marked manuscript):

"This index is applied and tested by Di Lorenzo et al. (2020) in unconsolidated aquifers in Italy, which are located in nitrate vulnerable zones. They refined the index (wGHIN) and demonstrated its applicability on shallow and deep aquifers, yet also revealing that the new index has its limitations in terms of low correlations between the indicators."

#### Methods

Comment #7: Line 116: replace 'with' to 'was'

Response: We agree. Done.

Comment #8: Line 116-117: improve sentence structure 'is mainly caused by' is incorrect, consider 'which' is mainly caused by: : : or rewrite sentence appropriately.

Response: We agree. Done. We replaced 'is mainly caused by' by 'which is mainly caused by'.

Comment #9: Line 120-124: condense and combine sentences

Response: We agree. Done. We condensed and combined sentences in this paragraph as follows: "A contaminant plume, which contains a polycyclic aromatic hydrocarbons concentration of up to 500 µg/l, of 200 m length over the entire aquifer thickness is located at a former gas plant in the east of Karlsruhe (Figure S1b) (Kühlers et al., 2012). Moreover, three parallel contamination plumes, of 2.5 km length each, can be found in the southeast of Karlsruhe (Figure S1b), where highly volatile chlorinated

C4

hydrocarbons ( $7 \mu\text{g/l} - 26 \mu\text{g/l}$ ) and their degradation products were detected (Wickert et al., 2006).”

Comment #10: Line 139-140: belongs in the results section not methods Section

Response: We agree. Done.

Comment #11: 2.3. I found this section hard to read, particularly due to some grammatical errors. This section is too long and verbose, it needs to be rewritten to make it clearer. The second paragraph starts well. I suggest removing the sentence start on line 160 “this requires to obtain also: : :”

Response: We agree on the grammar and the length of this section. Thus, we rewrote the paragraph and deleted some sentences (for more details see marked manuscript, lines 160-167).

Comment #12: Statistical analysis: You do not mention any of the statistical analysis completed in this paper. To be able to distinguish between forest and urban areas, you should at a minimum be completing some statistical analysis of the water quality data you have collected, even if this is simple ANOVA or t-test analysis. This is a major issue that detracts from the quality of this paper. I understand that you have used average values of the sampling wells, however determine whether there are statistical differences between (for instance) temperature at forested areas in comparison with urban areas, and look at the relationships between temperature, well depth and landuses. This analysis would greatly improve the scientific credibility of this study.

Response: We partially agree. We therefore performed U-tests instead of the suggested t-test, due to the possibility that the abiotic data does not follow a normal distribution. The results are as follows:

GWT forest vs. urban area:  $U = 248$ ,  $p\text{-value} = 3.3 \times 10^{-5} \rightarrow$  significant

Depth vs. GWT (forest):  $U = 64$ ,  $p\text{-value} = 1.6 \times 10^{-4} \rightarrow$  significant

C5

Depth vs. GWT (urban area):  $U = 203.5$ ,  $p\text{-value} = 5.5 \times 10^{-5} \rightarrow$  significant

Hence, we added the p-values of the U-tests to the manuscript (e.g. page 10, line 207). However, as various studies demonstrated (e.g. Amrhein et al., 2019) using p-values alone in a statistical analysis can lead to spurious interpretations, because p-values can exhibit wide sample-to-sample variability and therefore do not reliably indicate the strength of evidence against the null hypothesis (Halsey et al., 2015). Thus, we show the determined p-values, yet our focus remains on the presented spatial analysis using box-plots and other visual tools and comparisons.

Moreover, we added the following introductory sentence in Chapter 2.2. (line 147): “Mann-Whitney-tests (U-tests) were applied to detect potential impacts of groundwater characteristics (physical-chemical parameters), geology and well design on the groundwater quality as well as on groundwater fauna. Samples were regarded as significantly different if the p-value was  $< 5.0 \times 10^{-2}$ .”

Results/Discussion

Comment #13: Line 180: complete statistical analysis to indicate if there are significant differences in temperature between urban and forests areas- it appears that there are.

Response: We partially agree and added p-values of U-tests (see previous reply to comment #12).

Comment #14: Lines 192: while the box plots show that there are differences between forests and urban areas in DO and nitrate with landuses, these do not appear to be statistically different. I am not convinced that there are differences in DO and Nitrate between landuses this needs further discussion, as does the large differences in n values between the landuses.

Response: We partially agree (see previous reply to comment #12). The statistical analysis of the content of dissolved oxygen as well as of nitrate in the forest and urban area reveals significantly different distributions, which were added to the manuscript.

C6

Furthermore, we agree that further discussion is needed with regard to the differences between both parameters and land use. Thus, the following paragraph was added in the manuscript (lines 217-221):

“In the urban area average nitrate concentrations are typically higher and correlate with the content of dissolved oxygen ( $U = 278$ ,  $p\text{-value} = 4.0 \times 10^{-3}$ ) showing the link between nitrate content and oxygen consumption. Wells with a content of dissolved oxygen below 1.5 mg/l have an average content of nitrate of 1.5 mg/l, caused by nitrate reduction under anoxic conditions. Groundwater with reducing conditions ( $< 5$  mg/l dissolved oxygen) has an average nitrate content of about 7 mg/l in contrast to groundwater with oxidising conditions with 9 mg/l, which is characterised by the oxidation of ammonium to nitrate.”

We agree that the results would be more meaningful, if more measurements wells were considered. However, this is beyond the scope of this study and therefore should be part of a future large-scale study.

Comment #15: Line 196: References in chronological order

Response: We agree. Done.

Comment #16: Line 201: ‘hold back’ should be ‘retained’ or ‘held back’ Line 201: suggest these sentences are combined and reduced eg ‘: : ..where atmospheric nitrogen in retained by forest soils and fertilization is prohibited due to water protection regulations’

Response: We agree. Done.

Comment #17: Line 207-209: Again you cannot claim ‘clear differences’ without adequate statistical analysis of these factors. You need to run further analysis of the data for this statement.

Response: We partially agree (see previous reply to comment #12). Thus, we added p-values from U-tests to corroborate these observations (see also comment #13 &

C7

#14).

Comment #18: Line 231: I would not say that amphipods ‘predominantly live within wells’ rather they have a habitat preference for open spaces such as wells.

Response: We agree. We reformulated the sentence: “Stygobiotic Amphipods, large-bodied invertebrates which due to their size have a habitat preference for open spaces such as wells, ...”

Comment #19: Paragraph starting line 231: It appears that amphipods are significantly higher in forested areas than urban areas, however without analysis this cannot be determined. This may be ecologically important and should be discussed. It is also worthwhile looking at the correlation between cyclopoida and amphipods as briefly mentioned in line 238.

Response: We partially agree (see previous reply to comment #12). Hence, we performed U-tests and added the corresponding p-values and the following sentences in the manuscript (line 264): “Although statistical analysis with U-tests showed no significant correlation between the abundance of Amphipods and land use ( $U = 92.5$ ,  $p\text{-value} = 1.5 \times 10^{-1}$ ), the higher number of individuals in the forest area can support the hypothesis that, as mentioned above, Amphipods indicate healthy groundwater ecosystems, as...”

“The lack of a statistically significant correlation might also be related to the low number of wells ( $n = 8$ ) and individuals ( $n = 46$ ).”

Moreover, we agree that it is worthwhile looking at the correlation between the abundance of Amphipods and the order Cyclopoida. The following correlations were found and added in the manuscript (line 275). The results of the total statistical analysis are as follows:

Abundance Amphipoda forest vs. urban area:  $U = 92.5$ ,  $p\text{-value} = 1.5 \times 10^{-1} \rightarrow$  not significant

C8

Abundance Cyclopoida forest vs. urban area:  $U = 46.5$ ,  $p\text{-value} = 5.0 \times 10^{-3}$  → significant

Abundance Amphipoda vs. Cyclopida (forest):  $U = 8$ ,  $p\text{-value} = 9.0 \times 10^{-3}$  → significant

Abundance Amphipoda vs. Cyclopida (urban area):  $U = 311$ ,  $p\text{-value} = 2.0 \times 10^{-3}$  → significant

Abundance Amphipoda vs. Cyclopida:  $U = 430$ ,  $p\text{-value} = 9.6 \times 10^{-5}$  → significant

Comment #20: Line 238-240: Incorrect grammar: : : remove 'be'

Response: We agree. Done.

Comment #21: Line 248: Incorrect grammar

Response: We agree. We rewrote the sentence as follows:

"In addition, quantities of Bathynellacea (371 individuals) were found in five monitoring wells, all located in the urban area, in a depth of 9.0 to 13.5 m at a GWT of 12-15 °C (Figure 4b)."

Comment #22: Line 274- 280: The issue of purging wells needs further discussion as this is a limitation of your study. If you are looking at proportions of crustaceans to oligochaetes this is almost certainly affected by sampling method. The sentence on line 277 needs to indicate that relative abundances and proportions of crustaceans is likely to be impacted by the sampling methods, thus caution must be taken when interpreting the results.

Response: We partially agree that the sampling method is a limitation of our study. The standing water in the monitoring wells can host a larger number of individuals, caused by filtration effects. Yet, the proportional differences between the two groups are similar between wells and aquifers, as already demonstrated by various previous studies (Hahn and Gutjahr, 2014; Hahn and Matzke, 2005; Korb et al., 2017). Moreover, only large Amphipods, which were found in three wells, prefer living in open space (well

C9

water) (Hahn and Matzke, 2005; Korb et al., 2017). Other Crustaceans are smaller, like the order Cyclopoida, and are not influenced by filtration effects.

We do not agree that the sentence in line 277 needs to indicate that relative abundances and proportions of crustaceans is likely to be impacted by the sampling methods, as justified by the explanation above.

However, we added the following sentence (line 329): "Nevertheless, pumping can result in the selection of the taxa, especially in the presence of very fine sediments, and can result in changes of the sediment composition in the surrounding of wells and therefore in changes of habitat conditions."

Thus, we are confident about the appropriateness of the used sampling method, yet also mention that caution must be taken when interpreting the results (line 323).

Comment #23: Lines 295 -300: Could this also be due to organic carbon supply? Would level 2 assessment clarify these issues if it were undertaken?

Response: We agree that the food supply is one of the most limiting parameters for the survival of groundwater fauna. Thus, we added the following sentence in the manuscript (lines 355-357): "If the organic carbon supply varies on a small scale, this can influence microbiology and therefore groundwater fauna as well, although short-term changes in nutrient supply can be compensated by groundwater fauna."

We agree that the application of Level 2 might help to get a better understanding of the living conditions of groundwater fauna and might explain why some measurement wells are not populated. The assimilable organic carbon is one indicator which can be chosen as criteria (of the category microbiology) for the evaluation according to Level 2. However, the application of Level 2 is time-consuming and cost-intensive.

Comment #24: Line 305: Could the high (35%) of urban areas displaying natural sites be due to the sample methodology; ie were they classified as good incorrectly due to high proportions of crustaceans that may be influenced by the lack of purging of the

C10

wells?

Response: We partially agree. The standing water of the monitoring wells can contain a larger numbers of individuals than the surrounding aquifer, because wells serve as traps for the groundwater fauna and filtration effects can occur. Yet, as mentioned above, the proportional difference between the two groups will be similar, which is the main criterion for good ecological conditions in this assessment. In addition, some smaller Crustacean (e.g. of the order Cylopoida), which were found in larger numbers in most wells, are not influenced by such effects.

Conclusion

Comment #25: The conclusions of this study to me indicate that the method you have adopted (ie net sample wells and use the proportions of oligo/crustacean populations to determine ecosystem condition) need to be investigated further. The disproportionate number of crustaceans in wells due to sampling methods may be impacting the assigning of "OK" condition to sites that are actually impacted. Potentially a wider range of indicators need to be used including expanding on the use of only oxygen concentration in the classification scheme. The Level 2 assessment (Figure 1) also needs to be discussed in the conclusion.

Response: We disagree that measurement wells are incorrectly classified as good due to the sampling method (see comment #24). This study is focusing on existing approaches to obtain an initial impression on this complex topic. However, we fully agree that there has to be a defined sampling method to achieve representative sampling and comparable assessment of groundwater fauna in the future (see line 336).

We agree that a wider range of indicators has to be used in such classification schemes. Thus, we added the following sentences to the conclusion (line 386): "Level 2 assessment of Griebler et al. (2014) can help to achieve a more reliable and quantitative ecological assessment of urban aquifers, as it divides groundwater ecosystems in ecological grades according to the intensity of anthropogenic disturbance. It is

C11

based on the use of local reference values and the collaboration with experts, which is however challenging to apply. Therefore, further studies with large-scale and repeated measurement campaigns are needed to verify our findings. This should also include other cities and the determination of undisturbed local reference values which are required for a more reliable and also quantitative ecological assessment of urban aquifers. Moreover, a wider range of indicators should be considered in a classification scheme, such as temperature, porosity of the aquifer, groundwater flow, pollutants, nutrient supply, etc., especially when investigating urban areas. In addition, an important adaptation for an improved evaluation method is the determination of fauna at species level, which will provide more information (i.e. about Stygobionts, Stygophiles, Stygoxenes) and also consider the endemism of stygobiotic species. In this context, classification schemes should pay more attention to the different groundwater species and their potential use as indicator species."

References

Amrhein, V., Greenland, S. and Mchhane, B.: Retire statistical significance, *Nature*, 567, 305–307 [online] Available from: <https://www.nature.com/articles/d41586-019-00857-9>, 2019.

Brielmann, H., Griebler, C., Schmidt, S. I., Michel, R. and Lueders, T.: Effects of thermal energy discharge on shallow groundwater ecosystems, *FEMS Microbiol. Ecol.*, 68(3), 273–286, doi:10.1111/j.1574-6941.2009.00674.x, 2009.

Griebler, C., Stein, H., Hahn, H. J., Steube, C., Kellemmann, C., Fuchs, A., Berkhoff, S. and Brielmann, H.: Entwicklung biologischer Bewertungsmethoden und -kriterien für Grundwasserökosysteme, Umweltbundesamt., 2014.

Hahn, H. J. and Fuchs, A.: Distribution patterns of groundwater communities across aquifer types in south-western Germany, *Freshw. Biol.*, 54(4), 848–860, doi:10.1111/j.1365-2427.2008.02132.x, 2009.

C12

Hahn, H. J. and Gutjahr, S.: Bioindikation im Grundwasser funktioniert – Erwiderung zum Kommentar von T. Scheytt zum Beitrag "Grundwasserfauna als Indikator für komplexe hydrogeologische Verhältnisse am westlichen Kaiserstuhl" von Gutjahr, S., Bork, J. & Hahn, H.J. in *Grundwasser* 18, *Grundwasser*, 19(3), 215–218, doi:10.1007/s00767-014-0266-4, 2014.

Hahn, H. J. and Matzke, D.: A comparison of stygofauna communities inside and outside groundwater bores, *Limologica*, 35, 31–44, 2005.

Halsey, L. G., Curran-Everett, D., Vowler, S. L. and Drummond, G. B.: The fickle P value generates irreproducible results, *Nat. Methods*, 12(3), 179–185, doi:10.1038/nmeth.3288, 2015.

Korbel, K., Chariton, A., Stephenson, S., Greenfield, P. and Hose, G. C.: Wells provide a distorted view of life in the aquifer: Implications for sampling, monitoring and assessment of groundwater ecosystems, *Sci. Rep.*, 7(July 2016), 1–14, doi:10.1038/srep40702, 2017.

Kühlers, D., Maier, M. and Roth, K.: Sanierung im Verborgenen, *TerraTech Sanierungspraxis*, 3, 14–16, 2012.

di Lorenzo, T., Fiasca, B., di Camillo Tabilio, A., Murolo, A., di Cicco, M. and Galassi, D. M. P.: The weighted Groundwater Health Index (wGHI) by Korbel and Hose (2017) in European groundwater bodies in nitrate vulnerable zones, *Ecol. Indic.*, 116 [online] Available from: <https://doi.org/10.1016/j.ecolind.2020.106525>, 2020.

Spengler, C.: Die Auswirkungen von anthropogenen Temperaturerhöhungen auf die Crustaceagemeinschaften im Grundwasser, *Universität Koblenz-Landau.*, 2017.

Wickert, F., Müller, A., Schäfer, W. and Tiehm, A.: Vergleich hochauflösender Grundwasserprobennahmeverfahren zur Charakterisierung der vertikalen LCK- $\text{W\ddot{a}R}$ Verteilung im Grundwasserleiter, *Altlastenspektrum*, 01, 29–35, 2006.

C13

Please also note the supplement to this comment:

<https://hess.copernicus.org/preprints/hess-2020-151/hess-2020-151-AC1-supplement.pdf>

---

Interactive comment on *Hydrol. Earth Syst. Sci. Discuss.*, <https://doi.org/10.5194/hess-2020-151>, 2020.

C14