Answer to comments of Anonymous Referee #2

General comments

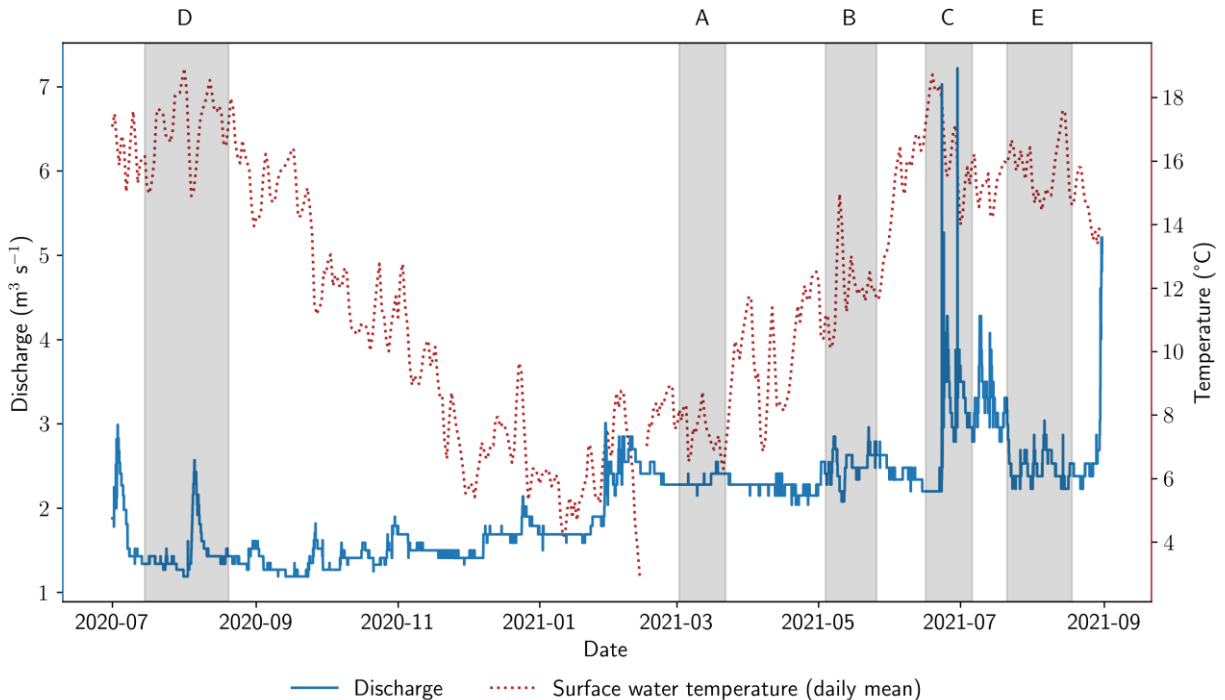
The authors want to thank the referee for the detailed review and the helpful comments. The referee clearly read the manuscript very carefully and could help to improve the paper. In the following, all comments and questions are answered in more detail. The updated manuscript with changes marked in blue color will be uploaded separately. Line numbers in this file refer to the updated manuscript. Significant changes are cited here.

The manuscript describes the results from 5 porewater peepers deployed in a small stream at different times and dates. From porewater profiles of different solutes the authors extract information about methane producing and consuming processes in the stream sediment. Results are supported by a molecular biological analysis of one of the five sites. The topic is interesting, innovative, and suitable for the journal. The paper is well written, methods seem to be carried out with great care (although I cannot judge the molecular methods). Major problems with the paper are related to the methods which make linking results and interpretation sometimes problematic or not possible:

1. Peepers integrate over long periods. I am not fully convinced that using them in a highly dynamic habitat is the best choice. Also the interpretation of the profiles relies on steady state assumptions. In order to judge this, it is necessary to have more information about temporal dynamics in that stream. I suggest to improve Figure 1 by showing discharge data with high temporal resolution (at least daily means) and to indicate peeper deployment periods in the figure. That enables judgement whether e.g. a flood occurred shortly before peeper retreaval. I think only stable conditions during peeper deployment allow the presented interpretation of the profiles.

Figure 1 has been updated to show discharge data with a higher temporal resolution and sampling intervals (see below). Generally, the study site at river Moosach shows very low flow variability and low sediment dynamics compared to other rivers. The discharge is completely controlled by weirs and check dams leading to stable hydrologic conditions. Further, fine deposits with low hydraulic conductivities dominate this river segment resulting in reduced hyporheic exchange. After sampling at river Moosach for several years we know that patches of course and fine sediment are stable for long time periods. In summary, the system is not as dynamic as the hyporheic zone often is and as the reviewer has probably assumed for river Moosach. The updated Fig. 1 supports this. Only during the sampling period at site C, two peaks of more than $6 \, m^3 \, s^{-1}$, classified by the GKD Bayern as mean flood discharge, occurred. This actually is an interesting information, because the higher discharge might be responsible for the infiltration of increased amounts of O_2 and NO_3^- and thus, could be an explanation for the observed CH_4 consumption in this profile. We included this argument at two points in the discussion:

"For example, extreme events can alter the chemistry of infiltrating surface water, as well as hyporheic flow path lengths and residence times, thus impacting hyporheic geochemistry in multiple ways (Zimmer & Lautz, 2014). In this study in particular, location C might have been impacted by two high flow events during the sampling period." (L. 387-390)

"One reason for the observed methane oxidation in location C could be an increased supply of O_2 and NO_3^- during the two high-flow events in the sampling period." (L. 442-443)

2. The spatial resolution of the profiles is often not sufficient to allow the resolution of different biogeochemical zones

Indeed, geochemical gradients in the hyporheic zone tend to be very steep and this is one of the main challenges in the measurement and interpretation of geochemical profiles in this habitat. It is true that the 1 cm resolution of our data is in some cases not sufficient to clearly separate biogeochemical zones (for example aerobic oxidation of methane and anaerobic oxidation of methane coupled to denitrification). Thus, in general, a higher resolution makes sense. However, it cannot be implemented in practice with the required water volume. Already in this study, the low sample volumes constituted a challenge to analytical techniques. Increasing the spatial resolution would lead to even smaller sample volumes and thus, render the measurement of, for example, stable isotopes in methane impossible. Further, the time required for sampling would increase drastically with a higher number of samples making losses of CH_4 during sampling and O_2 contamination inevitable. Therefore, unfortunately, the measurement of geochemical gradients including dissolved O_2, relevant anions and cations, CH_4 concentrations and values for $\delta^{13}C$ in CH_4 with a higher spatial resolution is not realistic.

3. The study mixes spatial and temporal variability because peepers were deployed at different sides not simultaneously. As a result the study gives very limited information regarding both spatial and temporal differences between peepers. The most striking result from the study is probably that all peepers were unique. We cannot tell which part spatial and temporal factors play

but I think this has consequences for other studies: General conclusions about stream functioning and upscaling from such single spot data is simply not possible. One could guess that having 5 more peepers would have resulted in 5 more very unique datasets. I recommend to discuss the issue of variability more.

Thank you for raising this point. We added a section on the spatiotemporal variability and the implications for extrapolation from point measurements to our conclusion (L. 571-573)

"The uniqueness of the measured profiles underlines the high spatiotemporal variability in the hyporheic zone. Therefore, deriving general conclusions from point measurements is highly problematic and the representativeness of the available data should be critically questioned in future research on CH_4 emissions from rivers."

4. They applied a 2D model (PROFILE) to a 3D scenario. This means any deviation from what was expected could be attributed to horizontal heterogeneity, transport inhomogeneity etc.. That makes interpretation of the profiles with respect to vertical reaction profiles highly subjective. Who decides in which case a feature of the profile is due to vertical biogeochemical processes or rather an artefact caused e.g. by transport inhomogeneity?

Certainly, the 1D model (PROFILE) cannot represent the full complexity of the system. Especially, because only diffusion is considered as a transport process. We agree that the application of a too simple model demands a cautious interpretation of the modelling results and the conclusion that were drawn. The application of this model is a first approach to better understand the measured concentration gradients. Disadvantages of this method were clearly communicated to the reader. For applying a more complex model, more data would be needed that was not available to the authors and was not the main focus of this work. Given the stable conditions in river Moosach, the authors think that this simplified approach can help in identifying reactive zones and add some valuable information to the interpretation of the profiles.

The general comments of Anonymous Referee #2 did identify some of the weaknesses of our study. The authors are aware of these disadvantages of the used methods and tried to clearly communicate this in the paper. Results were discussed critically and no certain conclusions were drawn if there was no proof found in the data. Thus, we do agree to the referee that these problems need to be put up for discussion, but we still think our data is worth publishing with an adequate explanation and interpretation.

Detailed comments

L.27: The abstract should end with a summarising/concluding sentence.

A concluding sentence has been added to the abstract.

L.36: Which % of natural sources are streams?

To answer this question, we used numbers provided by Saunois et al. (2020). Based on bottom-up estimations, rivers emit 27 Tg CH_4 yr^{-1}, all freshwaters together emit 159 Tg CH_4 yr^{-1}, and natural sources add up to 369 Tg CH_4 yr^{-1}. This has been integrated into the paper as follows (L. 37-39)

"Based on the evaluation of 385 globally distributed sites, rivers and streams are expected to emit 27 Tg CH_4 y^{-1} (Stanley et al., 2016) which is equal to 756 Tg CO_2 equivalents (IPCC, 2013) and constitutes approximately 17 % of freshwater emissions and 7 % of all natural sources (Saunois et al., 2020)."

L.47: One reference would be enough

We removed the second reference.

L.46-52: It is not really clear why this is relevant for the study

We think that this is information relevant, because later in the interpretation carbon stable isotopes in methane are used to determine the relevant production pathway. We moved this information to the results and discussion section (L. 423-424)).

"A kinetic isotope effect also occurs during CH_4 production and is larger for hydrogenotrophic than for acetoclastic methanogenesis (Krzycki et al., 1987). Here, $\delta^{13}C$-CH_4 values in the methanogenic zone were consistently lower than -60 ‰ which is characteristic for hydrogenotrophic methanogenesis (Whiticar, 1999)."

Introduction: There is lots of introduction about microbes but it is not really clear why. There are also lots of microbial references. I suggest to tailor the introduction more towards the aims and questions.

We tried to include all information that is needed to understand the processes mentioned in the results and discussion. In response to your comment we have removed and shortened some sections in order to get a better focus on the aims and questions and to reduce the number of references.

L.86: What were the findings of that study?

Ng et al. (2020) found that sulfate was the main terminal electron acceptor at the investigated wetland-stream system and that methane oxidation coupled to sulfate reduction (S-DAMO) could reduce methane concentrations close to the sediment-water interface. We included this result in the following sentence:

"Exceptions are for example the work of Villa et al. (2020) who measured vertical profiles of CH_4, CO_2 and N_2O at different beach positions and water stages to examine the relation of hyporheic exchange and GHG emissions, and Ng et al. (2020) who showed that S-DAMO could reduce CH_4 concentrations in a wetland-stream system by interpreting vertical geochemical with a multicomponent reactive transport model." (L. 76-79)

L.98: campaigns

This has been adopted.

L.112: Does that mean faster flow with macrophytes?

Thank you very much for this question. This is actually a mistake in our manuscript. Braun et al. (2012) found higher flow velocities above the substratum in cross sections without macrophytes. This has been corrected.

L.118: That information cannot be seen in Figure 1. Give discharge data with high temporal resolution.

Please check the updated Fig. 1.

L.120: How wide was the stream? Water depth?

This information has been added to the text with the following sentence:

"At this section, river Moosach is typically 10-12 m wide with a maximum water depth of approximately 1.3-1.4 m." (L. 113-114)

Figure 2: What are the two objects at the water surface? What are the 2 vertical lines separating the figure?

The objects at the water surface are the symbol for a free water table traditionally used in hydraulic engineering (see for example Novák et al. (2017) or the German norm DIN 19700). The vertical lines were intended to separate the profiles in a straight section from the profiles in a curved section of river Moosach. Since Fig. 2 should be clear without the symbols and separation lines, both were removed from the figure.

L.148: What was the orientation of the peepers relative to flow direction? Wasn´t there sediment erosion near to the peepers after deployment, because the peepers generate turbulence in flowing water.

The orientation was longitudinal to the flow direction to minimize disturbance of the flow. This was now also added to the text (see below). Erosion around the peepers was generally no problem. The reason might be that flow velocities in river Moosach are generally rather low. However, sediment erosion indeed occurred around the peeper at location C. As a consequence, profile C is shorter compared to the other profiles because sediment was removed from the upper part of the peeper. The erosion around the peeper occurred in the first days after insertion and the sediment surface stayed stable thereafter. Therefore, we assume that the measurement is still valid.

"To minimize flow disturbance, peepers were oriented longitudinal to the flow direction as indicated in Fig. 2a." (L. 141-142)

L.151-152: Be more specific. Why were 2 weeks not enough? How do you know?

Laboratory experiments conducted in our working group prior to this study indicated that 2 weeks are the minimal necessary equilibration time in water. Equilibration in the sediment will be slower than that and additionally, geochemical gradients need time to recover after insertion of

the peeper. Thus, we decided to leave the peeper for a time period of at least 3 weeks in the sediment.

L.161: Give type and size of vials. That means there was no water (except the 10µl NaOH) in the vials? There must be some small loss of sample gas using the described method. Did you check artefacts e.g. by preparing samples with known CH4 content?

We added this information to the text (L. 155). It is correct that there was no water in the vials. We added the samples slowly along the side of the glass vial to avoid turbulence and outgassing during injection. Both the injection needle and the smaller needle for pressure equalization were removed directly after sampling. We did not perform additional tests with samples of known CH_4 concentration.

L.211: I do not understand the boundary conditions chosen for CH4. Zero flux at top or bottom? Why can you assume that? Why not using concentration at the top and bottom as boundary conditions?

The model requires one concentration and one flux boundary. We assume no flux at the top of the profile, because the top of the measurements is above the sediment and in contact with the surface water. The stream water is well mixed and thus, we assume that there is no upward directed diffusive flux.

Results and discussion: I am not sure whether joining results and discussion are the best choice here. Jumping permanently between results and discussion is difficult for the reader. If a large revision is done I recommend to separate results and discussion. Use always past tense for results (e.g. L 252: depended).

Indeed, in our first draft of this paper we separated results and discussion. However, we found it more elegant to combine the two sections, because reading the discussion, one would always need to go back to the results section to check the measured profiles. This made reading the manuscript cumbersome. Thus, we would like to stick to the current structure. Thank you for the comment on the past tense, we will check the manuscript for this type of mistakes again.

L.256: Figure 3a and c.

Thank you.

L.275: What is the detection limit of the O_2 measurements and are <10 significantly different from zero?

According to the manufacturer, the detection limit of the O_2 micro-sensor (Unisense, Aarhus, Denmark) is 0.3 µM. However, O_2 very quickly diffuses through the membrane when removing the peeper from the sediment and thus, measured O_2 levels will almost always exceed the in-situ conditions. Unfortunately, due to the surface water flow and a water depth of up to 1.30 m we could not use the micro-sensor in-situ, but had to measure O_2 in the chambers of the peeper. The measurements do show a clear O_2 gradient and stagnating low concentrations (<10 µM) below.

Below this clear and sharp gradient, anaerobic processes like methanogenesis are taking place and thus, we assume that this zone beneath the measured O_2 gradient is completely anoxic.

L.314: Information on sediment composition would help a lot. Don´t you have e.g. LOI data for table A3?

Unfortunately, we have neither LOI or any other additional data on the sediment composition nor retained samples for conducting these analyses. All available information on the sediment composition is given in Appendix A.

L.324: "production" or rather "concentration"?

It is concentrations.

L.328: Why "seem". It should be possible to calculate CH4 partial pressure and compare with hydrostatic pressure.

The CH_4 saturation concentrations for the pressure and temperature conditions at each site were calculated using PhreeqC (Parkhurst & Appelo, 2013). Values for the sediment surface and 30 cm depth are given in the table below. At locations C, D and E the CH_4 partial pressures were above saturation pressure, i.e. the bubble point. It needs to be noted that bubble formation is also possible at lower CH_4 partial pressures if microstructures are present.

This was integrated in the paper as follows (L. 323-326):

"CH_4 concentrations in profiles C, D and E exceeded the CH_4 saturation concentrations of 170 µmol L^{-1}, 282 µmol L^{-1} and 202 µmol L^{-1}, respectively (calculated using PhreeqC and for the mean surface water temperature during the sampling period, a water depth of 0.6 m, 1.2 m and 0.7 m, respectively, and for a sediment depth of 30 cm). These concentrations are expected to cause the formation of gas bubbles."

Profil	T	Water depth (m)	Atmospheric pressure on sampling day (kPa)	Saturation conc. at the sediment-water interface (µmol/l)	Saturation conc. in 30 cm sediment depth (µmol/l)
A	7.5	0.7	102.2	151.64	216.62
B	12	1.1	101.7	222.47	283.15
C	16.6	0.6	101.1	113.64	170.46
D	17.1	1.2	95.4	225.72	282.15
E	15.8	0.7	101.8	141.76	202.52

L.333: "by"

Thanks, this has been changed.

L.334: Can you show the correlation between CH4 and NH3, e.g. in the supplement?

Scatter plots of CH_4 and NH_4^+ concentrations have been added as Appendix D. This has been integrated into the paper as follows:

"Appendix D: Correlation between CH_4 and NH_4^+ data

Figure D1 shows the correlation between CH_4 and NH_4^+ concentrations. In profile A, no NH_4^+ could be detected, therefore only data for profiles B-E is displayed. A clear positive linear correlation can be observed for profiles C and D. In profile B, NH_4^+ was mostly below the detection limit of 0.005 mmol L^{-1}. Higher concentrations were only found between 6-14 cm depth, above the zone where CH_4 concentrations peaked (increased CH_4 concentrations between 5-23 cm depth with a peak at 15 cm depth). A negative correlation between CH_4 and NH_4^+ concentrations seems to exist between 6-14 cm depth. In profile E, no correlation can be observed, NH_4^+ concentrations are generally very low compared to the other profiles.

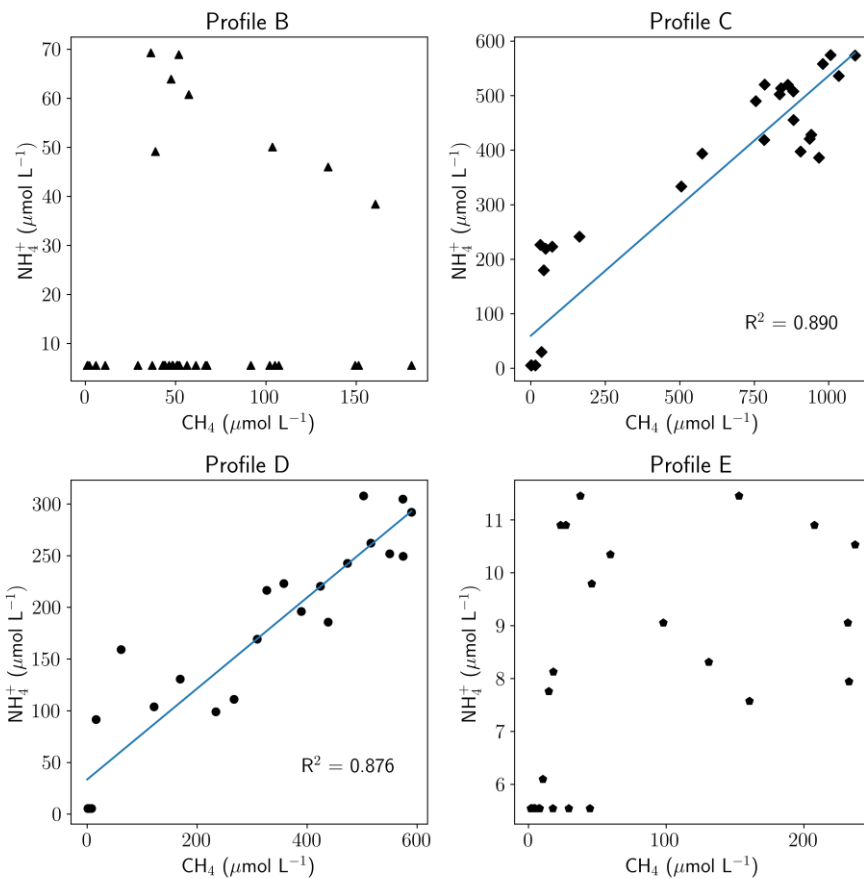



Figure D1: Correlation between CH_4 and NH_4^+. Each panel corresponds to one measured profile. Axes are scaled in the range of the data. For profiles C and D, a linear regression was performed and R^2 values are given in the respective plots."

L.380: So what? How is this sentence related to your study?

This sentence has been removed.

L391: delete "measured"

Ok.

L.396: ad a reference for this statement.

A reference to Fig. 3 (3b) has been added.

L.408-409: Difficult to understand

This has been rephrased to (L. 413-414):

"Apparently, microbial consumption only impacts the upper part of the gradient, while diffusive transport shapes the lower part of the gradient."

L.412: Why can you conclude that CH4 oxidation was not relevant at site D?

If CH_4 would be consumed at significant rates, this would lead to an isotopic enrichment of the remaining CH_4 pool, because organisms preferably take up lighter isotopes. However, our isotopic data does not show the expected shift in $\delta^{13}C$ in CH_4.

L.423: There is a problem of logic: Diffusion is a transport process and cannot reduce a concentration in the profile. If CH4 disappears you need a CH4 consumption process.

Actually, it is an important finding of this study, that CH_4 is not always consumed where concentrations decline. As stated above, microbial consumption leads to an isotopic enrichment which was not found in most of the profiles. Thus, consumption cannot be responsible for the diminishing concentrations. If there is a high concentration at the bottom (methanogenic zone) and a low concentration at the top (stream water), a concentration gradient will build up due to diffusion. This is the best explanation we have for the observed phenomenon.

L.429: Unknown? Is there really no literature about CH4 ebullition in streams?

Of course there is literature on methane ebullition from streams. However, there is no data from the Moosach river and the sampling sites of the peepers. As it is very hard to extrapolate from any point measurements in streams (which is also demonstrated by the variability found in this study) we cannot quantify the influence of ebullition in comparison to diffusive fluxes. We changed the sentence to make more clear that we are talking about our specific site and added some references on ebullition from streams (L. 434-436):

"The contribution of these bubbles to total CH_4 fluxes across the sediment-water interface at river Moosach remains unknown, but ebullition might be a significant contributor to CH_4 effluxes as suggested in the literature (DelSontro et al., 2010; McGinnis et al., 2016)."

L.438-439: Of course because that is what the PROFILE software is doing: Interpreting changes in slope as production/consumption processes.

Yes, exactly. We included the sentence because not every reader might be as familiar with the modeling tool PROFILE. But you are right that this is not actually a result and we removed the sentence.

L.452: This is a dangerous argument. The model is a quantitative one and give concrete numbers. How can you judge which numbers you trust and which not? This argumentation may question the entire quantitative interpretation of your profiles.

This is true. With this argument we wanted to highlight the fact, that calculated reaction rates should not be over interpreted since hyporheic exchange fluxes are not considered and their influence could not be quantified. The model can still be used to find the reactive depths (slope changes) which is its main purpose in this study. We re-phrased the section to:

"The model is applied to find the depths of reactive production and consumption zones. Calculated reaction rates are used to compare profiles, but due to the limitations described above, absolute values should not be over-interpreted." (L. 459-460)

L.462: These O_2 fluxes look extremely low. I would guess that the spatial resolution of the profiles was either not sufficient to model proper O_2 fluxes or that assumption about transport coefficients were not met.

We agree that the spatial resolution was not high enough to fully capture the O_2 gradients. This also affects the modeled O_2 fluxes. A second factor is surely that O_2 supply by hyporheic fluxes cannot be incorporated into the model. It was clearly mentioned in the text that uptake rates are most probably underestimated.

Fig.6: Is it possible to compare different groups quantitatively? It is striking that there were more methanotrophs than methanogens and that there were much more SRB. This brings also up the idea whether it makes sense to compare sulfate reduction and methane production rates from the PROFILE analysis to get information about the contribution of the different processes to total organic matter mineralisation.

The data cannot be compared in absolute quantitative abundance. 16S data is semi quantitative, and the percentages on the x axis correspond to relative abundances as a percent of total sequences obtained per sample. Therefore, if a percentage is higher it could be due to the decreasing of a different group (as opposed to an increase in the quantitative abundance of the group of interest). Further, relative abundance might not be directly linked to microbial activity. High relative abundance of a certain group of microorganisms, or an increase in relative abundance in particular geochemical zones defined to specific sediment depths (e.g., the methanogenic zone), can be seen as a hint that they may play a role in catalyzing certain processes in the specific environment they were found in. For example, an increased relative abundance of 16S rRNA gene sequences from methanogenic archaea in the methane zone allows for the logical conclusion that there is more methane in this zone due to the higher relative abundance of the methanogens (because they produce methane as they grow, and the increased growth causes a higher relative abundance of their 16S rRNA genes to be detected). Yet, their abundance alone cannot be used to quantify the amount of substrates that are converted. Therefore, we prefer to stick to our more careful interpretation.

L.542: The molecular analysis also integrates over a larger timescale. Without having information about short term dynamics of e.g. redox conditions it is difficult to interpret the findings.

As discussed above, conditions in the sediment of the Moosach river are relatively stable. We do think that the molecular analysis is applicable here. Knowing the microbial community distribution helps in understanding and interpreting measured geochemical gradients. However, the results cannot break down which processes happen exactly at which time and in which exact depth. Results are used just as an aid for the interpretation of the geochemical profiles and this is clearly communicated to the reader.

L554: Delete "can"

Ok.

equation C3: Explain symbols

Thank you. The parameters are now explained below the equation.

refs: 112 references are a lot. I suggest to critically check the necessity of all reference. There is potential for shortening esp. in the introduction. On the other hand I wonder if at least some discussion of temporal dynamics (e.g. the work of https://www.ufz.de/index.php?en=38353) might be helpful for interpretation of the data.

We have shortened the introduction and removed some of the references where appropriate. We have discussed temporal dynamics as follows (L. 386-393):

"When discussing the influence of hyporheic fluxes on redox zonation, it needs to be noted that not only spatial heterogeneities, but also temporal dynamics may play a key role. For example, extreme events can alter the chemistry of infiltrating surface water, as well as hyporheic flow path lengths and residence times, thus impacting hyporheic geochemistry in multiple ways (Zimmer & Lautz, 2014). In this study in particular, location C might have been impacted by two high flow events during the sampling period. Further, seasonal changes in river-groundwater mixing can potentially impact redox conditions and microbial populations (Danczak et al., 2016). However, fine sediments have been shown to reduce hyporheic exchange (Sunjidmaa et al., 2022). The combination of very fine deposits and stable, controlled hydrologic conditions is expected to limit hyporheic exchange and temper temporal dynamics in the HZ of river Moosach."

We cite the following literature in these answers:

Braun, A., Auerswald, K., and Geist, J.: Drivers and spatio-temporal extent of hyporheic patch variation: implications for sampling, 2012.

Danczak, R. E., Sawyer, A. H., Williams, K. H., Stegen, J. C., Hobson, C., and Wilkins, M. J.: Seasonal hyporheic dynamics control coupled microbiology and geochemistry in Colorado River sediments, Journal of Geophysical Research: Biogeosciences, 121, 2976-2987, 2016.

DelSontro, T., McGinnis, D. F., Sobek, S., Ostrovsky, I., and Wehrli, B.: Extreme Methane Emissions from a Swiss Hydropower Reservoir: Contribution from Bubbling Sediments, Environ Sci Technol, 44, 2419-2425, 10.1021/es9031369, 2010.

McGinnis, D. F., Bilsley, N., Schmidt, M., Fietzek, P., Bodmer, P., Premke, K., Lorke, A., and Flury, S.: Deconstructing methane emissions from a small Northern European river: Hydrodynamics and temperature as key drivers, Environ Sci Technol, 50, 11680-11687, 2016.

Ng, G. H. C., Rosenfeld, C. E., Santelli, C. M., Yourd, A. R., Lange, J., Duhn, K., and Johnson, N. W.: Microbial and reactive transport modeling evidence for hyporheic flux-driven cryptic sulfur cycling and anaerobic methane oxidation in a sulfate-impacted wetland-stream system, Journal of Geophysical Research: Biogeosciences, 125, e2019JG005185, 2020.

Novák, P., Moffat, A. I. B., Nalluri, C., & Narayanan, R. A. I. B. Hydraulic structures, CRC Press, 2017.

Parkhurst, D. L. and Appelo, C.: Description of input and examples for PHREEQC version 3 - a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations, US geological survey techniques and methods, 6, 497, 2013.

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., and Patra, P. K.: The global methane budget 2000–2017, Earth System Science Data, 12, 1561-1623, 2020.

Sunjidmaa, N., Mendoza-Lera, C., Hille, S., Schmidt, C., Borchardt, D., and Graeber, D.: Carbon limitation may override fine-sediment induced alterations of hyporheic nitrogen and phosphorus dynamics, Sci Total Environ, 837, 155689, https://doi.org/10.1016/j.scitotenv.2022.155689, 2022.

Villa, J. A., Smith, G. J., Ju, Y., Renteria, L., Angle, J. C., Arntzen, E., Harding, S. F., Ren, H., Chen, X., and Sawyer, A. H.: Methane and nitrous oxide porewater concentrations and surface fluxes of a regulated river, Sci Total Environ, 715, 136920, 2020.

Zimmer, M. A. and Lautz, L. K.: Temporal and spatial response of hyporheic zone geochemistry to a storm event, Hydrological Processes, 28, 2324-2337, 2014.