

Reply to editor comments

We added our point by point reply in red:

Dear Dr. Koschorreck, dear Matthias,

First, I have to apologize for the long response time. Too many things coming together at the same time.

Thank you for providing clear answers to the concerns of the two reviewers. I see clear merits in your study and support publication following a revision of the manuscript. I consider this a minor revision, even though I have to add a few points of concern besides those raised by reviewers.

Please realize your suggested actions to improve the manuscript. In particular I support the reviewer comments about a more modest title with regard to the spatial scale of the study. Representativeness of your study site for the "Elbe" may be supported by other studies on the river's morphology?

We added to the methods: "Such groyne fields are the dominant shore type along the German part of the river (Bussmann et al. 2022)."

I also agree with the reviewer wish for more statistical detail and inclusion of hysteresis results in the actual results chapter.

As explained in our reply to reviewer 2 we move a Figure from the supplement into the manuscript and extended the discussion of the hysteresis. We also added a table with results from the mixed effect models to the supplement.

The statistical methods chapter is poor and hard to understand. For instance, I also do not understand the question "whether samples originate from the sample distribution"?

We removed that sentence.

It is also unclear which groups were compared and why LME was used (i.e. which variables were fixed/random/covariates)?

The results of the LMEs including information about variables were added to the supplement.

In addition I ask you to address the following points:

*) It is not entirely clear to me how the influence of groundwater was hypothesised to actually affect CO₂ emissions. This may be because of a rather diverse vocabulary including groundwater evaporation, groundwater evasion, seepage and exfiltration. My understanding is seepage of CO₂-rich groundwater to the surface followed by diffusive emissions, but in various text sections actual evaporation is brought up as a mechanism as well?

As explained in the replies to the reviewers we tried to make this point more clear in the introduction (see below).

Overall, I am not quite sure why groundwater was speculated to potentially have such a strong influence on this particular site. None of the empirical data actually point to groundwater seepage. So, there isn't even hydraulic evidence that CO₂ could come from groundwater.

From other investigations we knew that at the Elbe both groundwater gaining and losing situations occur. In the floodplain near Magdeburg open waters are typically fed by groundwater. It was a very plausible assumption that at low river level a hydraulic gradient towards the river exists. But our measurements then showed that there was very little hydraulic exchange. Unfortunately we did not know that before we performed our study.

Also the small number of Rn measurements don't clearly identify groundwater influence in my opinion, as there is no measurement from river water in comparison.

Rn in the river water was actually measured. The Rn content of river water was more than one order of magnitude lower than that of the ground water. We added that data to Table 2 and added a

sentence to the text: “Groundwater contained more than one order of magnitude higher Rn concentrations than the river water (Table 2).”

I think the results are robust and clearly point to respiration as a source for CO₂ but the way groundwater influence is integrated into the manuscript may be reconsidered a bit.

*) I find the reference to a gauge 13 km downstream of the study site a bit difficult to follow. Is there a possibility to also provide information for how close the river water was to the chamber locations on the actual study site? Maybe I have misunderstood how this was done, then please see my comment as a hint for the need to clarify.

What we actually did is to derive a relation between the official river gauge 13 km downstream and the position of the water line at the study site. We expressed the chamber position as a height value and not as distance to the river, because distance was continuously changing.

*) Please provide clearer performance specifications of the FTIR gas analyzer and computations for respective data using the glimr package. This method section is hard to follow.

We changed the description to the FTIR analyser to: “. The change of concentrations in the chamber was monitored for every 30 s for ~ 5 minutes, with a multicomponent FTIR gas analyser (DX4000, Gaset Technologies GmbH, Helsinki, Finland). The FTIR gas analyser continuously measures CO₂, CH₄, and Nitrous oxide (N₂O) with an accuracy of ± 4 ppm CO₂, ± 0.1 ppm CH₄ and N₂O (Gaset Technologies GmbH 2018). Hence, the detection limit of the CO₂ flux was ~2 mmol m⁻² d⁻¹, while the CH₄ flux was detectable if above 0.12 mmol m⁻² d⁻¹, and N₂O if above 0.2 mmol m⁻² d⁻¹.” We added a reference where more details about the equation for flux calculation can be found.: “Fluxes were calculated from the linear increase of the respective gas mixing ratio (Gomez-Gener et al. 2015) with time using the R package glimr (Keller 2020).”

*) You assume a "spatial dependence" between radon and co₂ fluxes. Isn't this just a correlation? Please also respect (discuss?) the limitations in this approach, the Rn-dataset is tiny.

We wrote “same spatial dependence of CO₂ and ²²²Rn fluxes would be expected”. Thus, we do not say that there is a dependence between CO₂ and Rn but that they show a similar dependence on location if they both (mainly) originate from groundwater (we expected sites with high CO₂ flux to also have high Rn flux).

*) Please provide computational details for data processing following headspace gas analysis after equilibration (river water, ground water). An inexperienced scientist would not be able to follow your approach.

We write “Dissolved gas concentrations were calculated using temperature dependent Henry coefficients (UNESCO/IHA 2010) and CO₂ concentrations were corrected for alkalinity (Koschorreck et al. 2021).”. Both references contain the necessary information and equations. Since concentration calculation from headspace samples is a standard method we think citing a good reference is enough here not to make the text too lengthy.

*) Please briefly summarize how CO₂ concentrations were "corrected for alkalinity" even if such information can be read in a different paper. It is not clear to me why and when such a correction was necessary.

CO₂ cannot be treated like other gases in headspace analysis because it chemically reacts with water. Thus, during headspace equilibration the total amount of CO₂ in the headspace vial may change due to changes in speciation and the CO₂ concentration in the original sample cannot be calculated from CO₂ in the headspace alone. A second parameter of the carbonate system is necessary to correctly quantify CO₂. This is nicely explained in Koschorreck et al. 2021. We improved the sentence to: “Because the carbonate system in the headspace vial may change during headspace equilibration CO₂ concentrations were corrected for alkalinity as described in Koschorreck et al. (2021).”

*) Temp dependence was computed using an obviously noisy dataset. Behind that noise is a hysteresis effect, however. What if temp dependence was computed based on data shifted by the assumed time lag between flux and temp data? Would be nice to know if there was any change. Yes – that would be interesting. However, it apparently does not make much sense to shift all data by an assumed fixed time shift. As now discussed in the manuscript, the time shift resulting in the best fit differed greatly between days. Also the R^2 of the best fit differed between 0.2 and 0.97. Thus, the hysteresis pattern obviously depended on the day of measurement. We address this now in the discussion: “However, the time shift which produced the best linear fit differed between days (min=0, max=10, mean \pm SD = 4.8 \pm 3.7 h) with a median of 4 hours and no apparent differences between sites. Also the R^2 of the best fit differed between 0.2 and 0.97. Thus, the hysteresis pattern obviously depended on the day of measurement and it is not possible to derive a general relation which then could be used to analyse temperature-flux relations of time-shift corrected data.

Last, I urge you to check language carefully. If possible, ask a native speaker for help. There are numerous grammatical errors that should be taken care of.

We carefully checked the text again and asked other colleagues to read through the text. We hope that we spotted all these errors and the text is fine now.

I am looking forward to reading a revised version of this manuscript! Sorry again for the delayed response.

Regards,

Gabriel

Reply to Referee #1

I have reviewed the manuscript entitled *Temporal patterns and potential drivers of CO₂ emission from dry sediments of a large river*. The authors perform several GHG measured with different techniques in the riparian area of a river, during different periods along a year. They complemented these measures with several other variables related to the sediment characteristics. Although I think the content of the study is novel and interesting and the topic is relevant for this journal, my major concern relates to the spatial scale of the study. It was performed in a specific reach of a large river, with very specific environmental characteristics. I think this limitation in the spatial scale should be acknowledged and taken into account while discussing the results.

Thank you for the constructive review

General comments:

#1. In the introduction, the objective of the study was presented as the determination of the origin of the CO₂ emitted by the dry sediment, saying that a possible source would be the seeping of ground water. When saying “seeping of ground water” the first thing that came to my mind was the presence of a ground water source (aquifer) in the catchment.

However, in this case, the term “seeping water” wanted to refer to the distance of the dry sediment to the flowing water in the river channel and to the saturated layer. I think this should be clear since from the very beginning, to clearly understand the purpose of the different measures that were performed and the experimental design in general.

We agree that we were not clear enough at this point. We rewrote that section in the introduction to: “Yet, recent findings revealed a spatial variability of CO₂ fluxes from dry river sediments with highest fluxes near to the river Mallast et al. (2020). As a possible explanation the authors hypothesize that at decreasing river water level a groundwater flow gradient towards the river would transport groundwater to the river (Peters et al. 2006). Groundwater is usually 10 to 100 fold over-saturated with CO₂ (Macpherson 2009). Near to the river the thickness of the unsaturated layer approaches zero and CO₂ rich groundwater reaches the surface sediment where CO₂ would eventually degas.”

#2. The title of the study says “...CO₂ emission from dry sediments of a large river” but the spatial scale of the study was small, measures were not taken all along the river but in a specific reach. Taking into account the spatial scale of the study, I would change the title to make it a bit less pretentious.

We agree that the title may raise expectations that we aim to budget the entire river. We changed the title to “Temporal patterns and drivers of CO₂ emission from dry sediments **in a groyne field** of a large river”.

However, we think that the study site was rather typical for the lowland part of the river Elbe. Thus, we are convinced that the results have general implications for the entire lowland part of the river and most probably also for dry sediment sites at other surface waters.

#3. The aim of the study was to elucidate the origin of the CO₂ emissions. The response to this question was discussed taking together all different measured variables (section 4.1). However, the relationship of all the measured variables with the aim of the study was not totally clear for me while reading the methods and results section. Due to the high number of variables, I suggest adding a small explanation of their purpose in the methods section.

OK. We will add some more explanation to the methods section. For example: “To investigate spatial variability on 11 occasions between May and September transects of sediment respiration were additionally measured with a portable soil respiration system...”, “To quantify CO₂ fluxes at different distance to the river and also check for CH₄ emissions, manual chamber measurements were done...”, “To assess groundwater degassing ²²²Rn measurements were performed.”

Moreover, finally not only the origin of the emitted CO₂ (ground water or respiration) was addressed, but also the drivers of the magnitude of the fluxes. I think that taking in account the extend of this other aspect, it should be also mentioned in the introduction and aim of the study.

We agree that the introduction focusses too much on the source of CO₂. Since a similar point was also raised by reviewer #2 we changed the last part of the introduction to “If groundwater was a significant source of CO₂ we hypothesize a only weak temperature dependence of CO₂ emissions. We applied a combination of automatic high frequency measurements and detailed studies using a variety of methods to identify the source of CO₂ emissions from dry sediments at a large German river and to understand their temporal dynamics and drivers.”.

Specific comments:

Introduction

L41: Large rivers with high-flow are also susceptible to seasonal dry (i.e., Albarine river catchment in sud-west France, where more than 80 km representing ~25 % of the catchment and including the most downstream part are intermittent)

We agree and re-formulated: “... which lead to low-water levels or desiccation in streams and rivers.”

L55: and what about long term dynamics?

There is very little known about temporal dynamics at all. The question is “what is long term?”. In rivers as the Elbe low water periods typically last for weeks-months. There are yet only few studies addressing seasonal differences by a manual chamber approach. We re-formulated: “Few studies did address temporal variability of CO₂ emissions but nothing is yet known about short term dynamics of GHG emissions from dry aquatic sediments.”

L65: see my general comment #1 about the use of the term “water seeping water”.

See our reply above.

L75-77: related again to the general comment #1, make clear what you mean with “ground water” to clearly understand the aim of the study

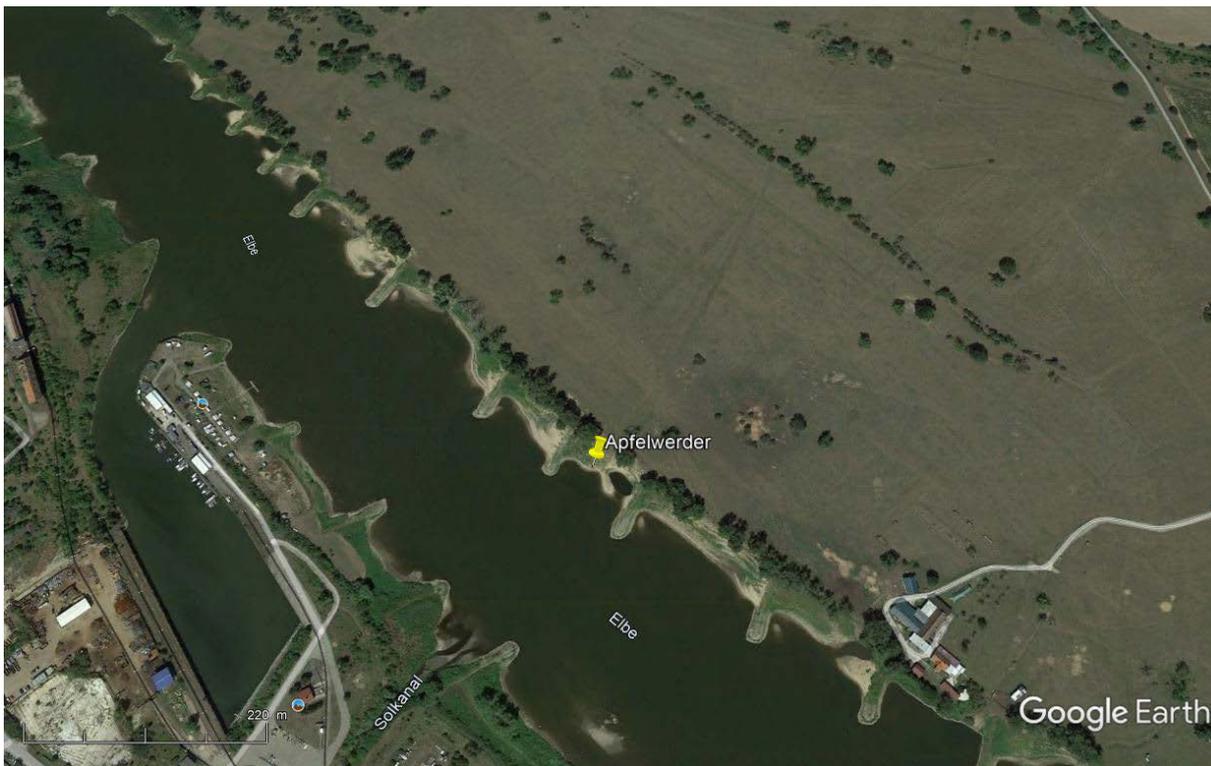
dito

Methods

L87: Could you specify the length of the reach? Maybe it's included in figure S1 but I was unable to download the supplementary material.

There is indeed a photograph of the site in the supplement. To describe the study site better we will add a short description of those groyne fields to the method section including e.g. “Groynes

extended about 50 m in to the river and distance between groynes was 130 ± 37 m". We also added a Google-Earth image to the SI showing the groynes around our study site:



L96: I don't understand what do you mean with "the chambers measured hourly CO₂ fluxes". Did you measure 5 minutes intervals during a period of 1 hour?

We measured for 5 minutes. The chamber was then opened and the system waited for 55 minutes before starting the next measurement. We will reformulate to: "The chambers measured CO₂ fluxes once every hour. Each flux measurement lasted 5 minutes and between flux measurements the chambers were open for 55 min."

L111: Why were these periods chosen? Looking at Figure 1, there were also other time-frames in which the water level was lower than in those selected periods

The river level changed quite dynamically. It was our aim to cover as much low water periods as possible without losing the instruments. Since prediction of water level changes was related to some uncertainty and there were also logistic and time constraints due to the travel distance we did not place the chambers immediately but to be on the safe site a few days later. We mostly removed the chambers very shortly before they would have been flooded. In one instance there was already river water entering the chambers when we removed them in the last second. To deal with such dynamics was a rather high effort and to capture these periods was a challenge. We see that it would always be desirable to have more data or longer time periods, but this was not possible in this first study.

What the reviewer probably asks is, why we did not place chambers at all water level positions during low water levels consistently. E.g. at beginning of September we did not put chambers at the 95 cm position. The reason is that we were most interested in the lower lying sediment areas and we speculated that the river level would further drop and we would then move chambers further down. Unfortunately, that summer did not show such a clear low water period as the preceding years.

There were a few periods during high water level when the chambers were placed high up just to put them in a safe place. However, these rarely flooded sites were not the focus of our research and we did not intend to monitor those sites.

All in all we hope that Figure 1 and the explanation in L.99-103 sufficiently explain why those measurement periods were chosen.

L152: The significance of the acronym “Bq” is not specified

Bq is the SI derived unit for radioactive decay (Becquerel). Since the text explains that we are measuring radioactive decay we did not think it was necessary to specify this more. To clarify, we specify that Bq means radioactive decay we will add to the method section: “The counts were measured over one hour and averaged, with a standard deviation of one sigma **and expressed as decays per second [Bq].**”

L154: How/Where (flowing channel? ground water?) were these water samples taken?

Samples were both taken from the groundwater and the river. River data are now included in the manuscript in Table 2. Changed to “The ²²²Rn concentration in 300 mL samples from ground water (2.3.3) and the river was measured with the Wat250 mode.”

L189: The significance of the acronym “g-dw” has not been specified

Changed to “... rates of respiration per gram dry weight [$\mu\text{mol g-dw d}^{-1}$] were converted to fluxes by multiplying with sediment...”

L207: For how long were the drying and the ignition?

Changed to: “...after drying for at least 2 days to constant weight at 105 °C and loss on ignition (LOI) at 550°C, respectively.”

L252-253: As the variability of the data shows a temporal pattern, it would be interesting to analyse the potential drivers of this changes.

Indeed, the pattern is very interesting and motivated us to follow up on this study. The discussion can be found in section 3.1.1 and most illustrative when discussing Figure 4.

L268: And what about precipitation? Why didn't you include it in the mixed model?

Precipitation and moisture showed considerable co-linearity. We decided to keep moisture because it was measured directly at the site while precipitation data were from a weather station.

Figure 3. The way the hour is indicated in legend (0-20) seems a bit strange

We changed the color scale to cover 0-24 h.

Results

Related to the general comment #3, at a first reading only the section 3.2.1 of the results seems directly related to the current aim of the study (source of the CO₂).

As explained above we rewrote the introduction and expand the aim of the study by including the drivers underlying the fluxes and patterns.

Discussion

The first point of the discussion (4.1) perfectly addresses the aim of the study and summarises all the results obtained. I really like the way it's written, very clear and direct.

Thanks.

In relationship with the general comment #2, along the discussion, the influence of different factors (dependence of temperature, sediment characteristics, thickness of the unsaturated layer...) in the CO₂ emissions were discussed. However, although performed in a big river, the spatial scale of the study was small. Those factors can substantially change along the river, from the headwaters (typically at higher altitude, with more forested and closer riparian areas) to downstream areas (wider reaches, less forested riparian areas, more exposed to solar radiation). I think this spatial scale (together with the already addressed temporal scale) should be at least mentioned in the discussion.

We thought that talking about a “river” would already make clear that the study is not about “headwaters”. We will clarify that this study is not about headwaters but about the lowland section and that we focus on a (typical) groyne field. Although it is a rather small scale of observation, the chosen groyne field is representative for large sections of the river. The entire river upstream of km 130 has been modified by groynes (Bussmann et al. 2022) – thus, the site can be considered typical for the lowland part of River Elbe from km 130 to km 580. We add to the methods section: “Such groyne fields are the dominant shore type along the German part of the river (Bussmann et al. 2022).”

Typing errors:

L207: repeated words: loss after

deleted

L243: May

Corrected.

Reply to referee #2 (Kenneth Thorø Martinsen)

General comments:

The authors investigate CO₂ emissions from dry sediments at one site in a large German river. High frequency automatic flux measurements provide an excellent view into the temporal dynamics of CO₂ emissions. Additionally, measurements across transects provide information on spatial variability and the contribution of groundwater is assessed using Rn as a tracer. The CO₂ emissions are primarily driven by microbial respiration. Furthermore, there interesting descriptions of hysteresis and dark CO₂ uptake. The study appear thorough, methods appropriate, and results are well presented and discussed. Unfortunately, I was not able to access the supplementary material.

Thanks Kenneth for the helpful review.

Specific comments:

- I miss some explicit hypothesis. The aims (1.4) are presented in a broad sense, and test of the groundwater hypothesis is mentioned but so much more data is presented in the manuscript which is why a think specific hypothesis should be included.
We agree. We added a hypothesis addressing the drivers of CO₂ emissions “If groundwater was a significant source of CO₂ we hypothesize a only weak temperature dependence of CO₂ emissions.”.
- How are the flux chamber data quality checked (L 103)? I think this should be described.
When waves reached the chambers sand was washed away and in some occasions the chambers were not tight any more. This immediately led to concentration data fluctuating around atmospheric concentration which results in zero flux with low R² of the linear fit. Such data was discarded. We will change to: “Automatic flux chamber data were discarded when the collar was flooded or the sand was washed away by waves removed, which resulted in CO₂ concentrations fluctuating around ambient concentration”.
- L 216, following the ANOVA test I would have expected something like a Tukey post hoc test adjusted for multiple comparisons and not repeated pairwise t-tests.
Right – the Tukey test would have been the right choice here. However, the ANOVA was only used to detect differences between days. Since this test does not really provide important information and the results are not presented we remove the sentence from the methods section.
- Regarding LME, how was model selection performed? In general, I miss some more details on the modeling procedure.
We first tried a complete model including temperature, moisture, and thickness of the unsaturated zone as predictors. We then removed single predictors and compared R² values. We specify in the method section: “Linear mixed-effects models (lme) were applied to predict the influence of the environmental variables on the CO₂ flux at the study site for variables for which a linear relationship with the CO₂ flux was presumed. Model selection was done by removing predictors and comparing conditional R² values of different models.”
- Also regarding LME, I miss a more detailed description of LME results. Currently, only the R² values are presented but a table (supplementary perhaps) with model coefficients etc. would be welcome.
We added LME model results in table S1 in the supplement.
- Figure 5, I had a difficult time understanding this figure. Could this alternatively be shown using lines in a CO₂ flux (y) vs distance (x) type plot. Something is also wrong in the legend, i.e. “NA” values.
We agree that the figure is difficult to read. We improve it by replacing the symbols with a “plant line”. We also improved the color scale by reducing the number of colors.

- I think the hysteresis results (L 430-432) should be presented the Results section. The hysteresis is interesting and could potentially be further elaborated in the discussion, where there any differences between sites?
The hysteresis part was indeed developing during writing of the manuscript. As the reviewer states it is now an important result. We had scanned the hysteresis figures for several dates. Four additional dates were already shown in the supplement (which for some reason the reviewer could not access) indicating that the general appearance of these curves was quite similar between days. To better illustrate the time shift we moved Figure S5 from the supplement into the main manuscript as Figure 8. We also extended our analysis by calculating the time delay which produces the best linear fit for each single day: “However, the time shift which produced the best linear fit differed between days (min=0, max=10, mean±SD = 4.8±3.7 h) with a median of 4 hours and no apparent differences between sites.”
- An admittedly minor thing perhaps, but please be consistent with capitalization of axis and legend labels in all figures. Also for figure references, e.g. Figure 5 (L 304), figure S1 (L 91) and Fig. S1 b in (L 135). Please correct throughout the manuscript.
We unified to “Figure”.
- Date formatting in tables and figures differ, e.g. month-day in figure 5 and day.month.year in table 1, at least month-day or day-month order should be consistent. Please correct throughout the manuscript.
We unified this to mm-dd.

Technical comments:

L28 Replace “largely” with “greatly” or other.
Replaced by “greatly”.

L55-57 Awkward sentence, please rephrase.
Changed to: "Investigating temporal variability of CO₂ fluxes should provide information about the potential sources of emitted CO₂. Knowing sources of emitted CO₂ from dry sediments is crucial to be able to model or scale up GHG emissions from these systems."

L64 Is something missing e.g. “In contrast to respiration”? Please rephrase.
We rephrased to: “In contrast to respiration, abiotic processes are rarely taken into account as sources of CO₂ (Rey 2015)”.

L71 Replace over-saturated with super-saturated
replaced.

L131 Replace “manual” with “Manual”
replaced.

L186-188 and 232-234 Same paragraph occurring twice
Sorry. L232-234 are deleted.

L230 regard log-transformation, there are also negative fluxes how were they treated.
We added to the methods section: “Because of occasional small negative fluxes we shifted all fluxes to positive values by adding 121 mmol m⁻² d⁻¹ prior to transformation (120 was is the value of the largest negative flux).”

L242 +/- what – standard error? Please write.
Standard Deviation. We added “±SD”

L243 Replace “Mai” with “May”
replaced.

L260 Just write LOESS smoother with span 0.1. The gray confidence region around the smoothers are confidence intervals or standard errors? And not SD?
The gray area is the confidence interval. We changed the legend.

L262 What is “HF” in title?
Its “High Frequency”. Removed.

L266 Details regarding modelling, e.g. chamber as a random effect should be in methods.
We think if we do not list the predictors here it will be difficult for the reader to figure out about what model we are talking. We would prefer keeping it as it is.

L267 What are the R^2 values for the mixed models? Often they are conditional/marginal depending on whether they include random effects or not.
As written in the method section we used the conditional R^2 which includes random effects.

L291 Replace “spatial” with “Spatial”
Replaced.

L306 Awkward sentence, please rephrase.
We rephrased to: “In sum, our field based measurements provide strong evidence that respiration in the sediment was the major driver of the observed CO₂ flux.”.

L318 Description of texture method should be in Methods.
Moved to Methods.

L550 “Short term temporal dynamics” Maybe replace “dynamics” with “variation”?
We agree. Changed.

references

- Bussmann, I., U. Koedel, C. Schütze, N. Kamjunke, and M. Koschorreck. 2022. Spatial Variability and Hotspots of Methane Concentrations in a Large Temperate River. *Front Env Sci-Switz* **10**.
- Gomez-Gener, L. and others 2015. Hot spots for carbon emissions from Mediterranean fluvial networks during summer drought. *Biogeochemistry* **125**: 409-426.
- Keller, P. S. 2020. glimmr: Compute gasfluxes with R. *Gas Fluxes and Dynamic Chamber Measurements*.
- Koschorreck, M., Y. T. Prairie, J. Kim, and R. Marce. 2021. Technical note: CO₂ is not like CH₄ - limits of and corrections to the headspace method to analyse pCO₂ in fresh water. *Biogeosciences* **18**: 1619-1627.
- UNESCO/IHA. 2010. GHG Measurement Guidelines for Freshwater Reservoirs, p. 138. *In* J. A. Goldenfum [ed.]. UNESCO.