

# Reply to reviewer 1

## Summary of the manuscript

The reviewed manuscript by Koschorrek et al. quantifies variability of methane and carbon dioxide fluxes between the atmosphere and a temperate low-land river at scales of hours and hundreds of meters. Based on a three-day sampling campaign, including flux chamber measurements in the river and in nearshore areas, the authors found considerable diurnal variability in carbon dioxide fluxes and variability from near-shore to off-shore areas in methane fluxes. The authors also discuss consequences of different sampling strategies for upscaled gas fluxes, concluding that accurate flux estimates require continuous measurements.

## Overall assessment

**Scientific significance:** As well introduced by the authors, rivers play an important component of the global carbon cycle and emit carbon gasses at globally significant rates. Yet, there are large uncertainties in emission estimates due to very large spatiotemporal variabilities. The research question on spatiotemporal greenhouse gas fluxes in rivers is not particularly novel, but the focus of this study on small-scale variations (diurnal, near-shore / off-shore) fills a poorly studied niche in the literature that is well worth investigating. I also appreciate the comparison of aquatic and terrestrial gas fluxes, which is rarely done, but highly relevant given that rivers can vary largely in their aerial extend, depending on discharge fluctuations.

**Scientific quality:** The scope of the study including 3 days of measurements in a 1 km river reach may not appear overly impressive and representative for other conditions. However, relative to many other studies, the authors managed to collect an impressive and interesting data set at very small spatial and temporal scales. Overall, the authors address the research question by using state-of the art techniques. The study design could be acceptable, overall, but some design-related questions should be addressed first (see major concerns below). I agree with most data interpretations and conclusions, but a mismatch in the results shown should be resolved (see major concerns below). I also have a few concerns about the statistical analysis of the data, as outlined below.

**Presentation quality:** The manuscript is well written, logically structured and clearly and concisely presented. Overall, the figures and tables, including the supplementary material, are adequately chosen and well designed, but I have some concerns and suggestions for improvements, as listed below.

Overall, I find that the manuscript is well within the scope of Biogeosciences.

## Major concerns

A main focus of the manuscript is to compare spatial and temporal variability in gas fluxes. I wonder to what extent this analysis may be biased by the fact that spatial and temporal assessments were not fully independent? I understand that for practical reasons (limited availability of gas analysers), it is impossible to perform simultaneous measurements at the different locations. However, I would expect a discussion on the consequences of the sampling design for the analysis of spatial and temporal variability in aquatic gas fluxes. For

example, I would like to see at what time the different floating chamber measurements were performed. Given that each measurement takes 2-5 min, I would expect that daytime may affect measurements, in addition to location. Did the authors account for time in their assessment of spatial variability?

The reviewer is right that it is challenging to measure simultaneously at several sites. We partly succeeded here by deploying 3 automatic chambers at different habitats on dry sediments. Thus, at the dry sites we think we adequately addressed spatiotemporal variability simultaneously. The reviewer is right that we did not do so at the aquatic sites. Chamber measurements were done only during a few hours during the day. This will probably not affect our results for  $k_{600}$  (because of rather constant wind and discharge conditions).  $\text{CH}_4$  fluxes will also not be affected, considering the very limited diurnal change of  $\text{CH}_4$  concentration. Regarding  $\text{CO}_2$  emissions one may argue that the diurnal amplitude of the  $\text{CO}_2$  concentration might differ between sites. For  $\text{CO}_2$  differences between the middle of the river and the groyne fields can be expected to be lower in the night because sediment driven  $\text{CO}_2$  production might increase  $\text{CO}_2$  concentrations in the groyne fields during the night. This would further decrease the already low spatial variability of aquatic  $\text{CO}_2$  emissions – supporting our conclusions. In a revised manuscript we will add these considerations to the discussion.

Related to the major comment above, it is unclear to me how potential temporal variability in the gas transfer velocity was accounted for in calculations of diel gas fluxes. I appreciate the high temporal resolution of dissolved gas concentrations, but for accurate calculations of gas fluxes, temporal variability in  $k$  should also be characterized.  $k$  may or may not vary on a diel basis (see e.g. Attermeyer et al. 2021 Comm. Earth&Env). Please clarify how time series fluxes were calculated and discuss any potential shortcomings, in case concentration and  $k$  estimates differ in temporal resolution.

We do not expect large differences in  $k$  because of rather low constant wind below 4 m/s and discharge. Furthermore existing literature suggests that in rivers wind speed (which is potentially variable during the day) has a small effect on  $k$  compared to hydrodynamic parameters (which are rather stable on the timescale of days) (Huotari et al., 2013; Molodtsov et al., 2022). A Figure with windspeed data will be added to the supplement.

I think there is a mismatch in gas fluxes and concentrations shown in Table 1 and in Figures 3/4. According to Table 1,  $\text{CO}_2$  fluxes range up to 13.9  $\text{mmol m}^{-2} \text{h}^{-1}$ , with medians up to 2.8  $\text{mmol m}^{-2} \text{h}^{-1}$ . In contrast, the Figure 4 shows maximum fluxes of near 30  $\text{mmol m}^{-2} \text{h}^{-1}$  and medians of up to 10  $\text{mmol m}^{-2} \text{h}^{-1}$ . Also,  $\text{CH}_4$  concentrations in Figure 3 range up to 240  $\text{nmol/L}$ , compared to 320  $\text{nmol/L}$  in Table 1. Shouldn't the data shown in Table 1 and Figures 3/4 be the same? Table 1 suggests no considerable difference in  $\text{CO}_2$  fluxes between aquatic and terrestrial habitats, while Figure 4 does. The mismatch may have implications for results (L. 216) and conclusions (L. 432). This issue must be addressed, through corrections or clarifications, before the manuscript can be considered for publication.

In fact the data in Table 1 and Figure 5 are not the same. Table 1 shows the result of manual chamber measurements in the different habitat types including several sites per habitat type. In Figure 4 the temporal data at one site per habitat type (where the probe or automatic chamber was installed) is shown.

However, we discovered an error in table 1 where partly an earlier version of the table was included. The  $\text{CH}_4$  and  $\text{CO}_2$  emissions data of the terrestrial sites have the wrong unit

(nmol/m<sup>2</sup> s and Mmol/m<sup>2</sup> s) and an exponential instead of linear fit was used for flux calculations. Furthermore those data were not complete – a few sites which later were re-classified with respect to habitat type were missing. We also re-assessed the raw data of the chamber measurements and corrected the aquatic CH<sub>4</sub> emission data. The correct data (linear fit and unit as in the table column 1) are in the table below. We will correct Table 1 and Figure 4 (see below) as well as Figure 6 in the revision. We are very grateful that the reviewer discovered the wrong numbers.

	aquatic			terrestrial		
	middle	side	groyne fields	sandy	muddy	vegetated
CH <sub>4</sub> emissions [μmol m <sup>-2</sup> h <sup>-1</sup> ]	12.5 (4-20.8)	41.7 (21-75)	39.6 (12.5-54.2)	0.12 (-0.2-0.38)	0.03 (-0.83-1.13)	-0.006 (-2.9-0.5)
CO <sub>2</sub> emissions [mmol m <sup>-2</sup> h <sup>-1</sup> ]	2.2 (1.86-2.21)	1.5 (1.07-3.13)	0.8 (0.21-2.3)	1.9 (0.7-4.6)	4.9 (-4.1-16.8)	11.1 (-2.7– 45.1)
CH <sub>4</sub> concentration [μmol L <sup>-1</sup> ]	0.12 (0.11 - 0.16)	0.18 (0.17 - 0.32)	0.18 (0.17 - 0.21)	-	-	-
CO <sub>2</sub> concentration [μmol L <sup>-1</sup> ]	29.3 (28.4-29.6)	23.4 (15.4-24.2)	15.7 (15.4-28.8)	-	-	-
k600 [m d <sup>-1</sup> ]	2.6 (1.6-5.2)	5.2 (2.2-10.3)	4.3 (1.4-5.7)	-	-	-
velocity [m s <sup>-1</sup> ]	0.79 (0.72-0.81)	0.65 (0.22-0.84)	0.22 (0.09-0.45)	-	-	-
total area [m <sup>2</sup> ]	125,000	25,000	42,000	7,700	4,800	6,500
total CH <sub>4</sub> emissions [mol h <sup>-1</sup> ]	1.6 (0.5 - 2.6)	1.0 (0.2 - 1.8)	1.6 (0.5 - 2.3)	0.009 (-0.002- 0.003)	0.0002 (-0.004-0.005)	-0.00004 (-0.019 - 0.003)
total CO <sub>2</sub> emissions [mol h <sup>-1</sup> ]	275 (232 - 276)	37.5 (26.7 - 78.5)	33.6 (8.8 - 96.6)	14.3 (5.5-35.4)	23.6 (-19.3-80.7)	72.4 (-17 - 293)

As a result numbers in Table 3 will minimally change to:

	aquatic				terrestrial				all
	middle	side	groyne fields	total	sandy	muddy	vegetated	total	total

CH <sub>4</sub> [mol d <sup>-1</sup> ]	37.5	25.02	40	102	0.006	0.014	-0.001	0	91
CO <sub>2</sub> [mol d <sup>-1</sup> ]	9960	1992	3347	15299	425	137	1101	1663	16962
CO <sub>2</sub> -eq [mol d <sup>-1</sup> ]	10342	2247	3753	16342	425	137	1101	1664	18005

The authors mention the major effect of salty water inflow (river Saale) affecting water chemistry along the western shore (L. 203-206). The authors sampled the western shore and main part of the river, but not the eastern shore, which seems not to be affected by the salty water inflow. I understand that the focus of this study was on the Groynes located along the western shore. However, given the focus on spatial variability of this study, I think it would have been valuable to also study the eastern shore as a “reference” to better evaluate the effect of the Groynes and the salty inflow. Why did the authors did not do any attempt to also study the eastern shore? To what extent could the salty inflow have affected results? Would there be any way to disentangle the spatially overlapping effects of the salty water inflow and the groynes? I would appreciate a brief discussion on this issue.

We did in fact perform chamber measurements on both sides of the river but lumped the data together in the analysis because otherwise our n would be quite low. When looking at the data from both sides separately we do not see large differences. Conductivity differed between sides but the difference was rather small (1100 versus 1300  $\mu\text{S}/\text{cm}$ ). Thus, we would not expect significant differences in microbial processes at both sides. We think the small conductivity difference as well as the low number of chamber measurements prevents any further analysis of the potential effect of slightly different salt concentrations on GHG emissions. As we write in the manuscript the conductivity difference is a good indicator for limited lateral mixing of the river water.

What is the role of ebullition for gas fluxes in the studied system? Given the potentially large role for total fluxes as well as spatial and temporal variability of methane fluxes, I think this should be discussed more in the manuscript (extending the statement in L. 422). In particular, did you observe sudden jumps in the within-chamber gas measurements that would indicate ebullition? If so, how did you treat such data and how would the exclusion of ebullition affect gas flux estimates?

The reviewer is right that ebullition would be a game-changer for CH<sub>4</sub> emissions. We actually did not observe ebullition in our chamber measurements. We cannot fully exclude that ebullition might occur at other sites or times. However, Matousu et al (2019) did also not observe ebullition in the Elbe (with the exception of one harbor). Thus, ebullition does not seem to be very relevant in the Elbe. Although Matousu et al did not observe ebullition in dammed sections of the river we would not exclude ebullition at river sections upstream of weirs (as frequently shown in other studies) and our results might not be valid directly upstream of the only weir in the German part of River Elbe in Geesthacht (where Bussmann et al. 2022 showed elevated CH<sub>4</sub> concentrations). We will add a sentence on this to the discussion.

I would like to see more details on the statistical analyses used. For example, the choice of methods described in L. 187-192 should be justified and the used R functions / packages should be explained/cited. What explanatory variables (fixed and random effects) were investigated in the Linear mixed models (L. 191)? Were fluxes always positive so that log-transformation is justified (L. 190)? How was temporal / spatial autocorrelation tested/accounted for in the analyses? How does the correlation analysis and linear mixed

effects modelling help to address the stated research question? Can you please add details of statistical analysis (Wilcox test statistics, mixed effects model parameters / AIC, degrees of freedom). This could be added in the main text or as tables, e.g. in the supplementary material.

We used base-R functions – thus we do not see the need to cite packages. The explanatory variables used in the mixed linear models are mentioned in the results section (1.260-271). We think that having that information in the results part makes the results easier to read.

There were indeed some negative fluxes in our timeseries at the muddy site and a few at the sandy site (see line 246). For those we added the most negative flux as a constant to the data before log calculation. Autocorrelation was inspected visually and choice of variables done based on expert knowledge. We will add that information to the methods section. In the revision we will also add a table with data from the statistical analysis to the supplement.

### Specific comments

L. 14: Can the authors motivate their statement that most existing studies were carried out in small streams? Perhaps by referring to published work (review, metaanalysis). Personally, I don't have a complete / up-to date overview of the existing literature, but I don't necessarily have the impression that smaller streams are represented more than larger rivers. For air-water gas exchange work in larger rivers, see e.g. Yao et al. (2007, *Sci Total Environ*), Alin et al. (2011, *JGR*), Hall et al. (2012, *L&O*), Beaulieu et al. (2012, *JGR*), Striegl et al. (2012, *GBC*), Huotari et al. (2013, *GRL*), Borges et al. (2016, *Nat. Geosci.*), Qu et al. (2017, *Sci. Reports*), Rosentreter et al. (2017, *L&O*), Paranaíba et al. (2018, *ES&T*).

We did not perform a robust literature analysis to clarify this point. In the review by Stanley et al. (2016) CH<sub>4</sub> concentrations in 652 small to medium size streams compared to 265 in large streams are reported and there are probably much more data on tracer addition derived k values from small streams compared to k values measured with alternative methods in larger systems. However, we agree with the reviewer that the point is probably not so clear and the answer depends a lot on the subject studied. Process studies and studies on the gas transfer velocity are often carried out (for practical reasons) in small streams. On the other hand in global upscaling, larger systems are better represented because the surface area is easier to quantify (Marx et al., 2017). We are aware that there are several studies on river GHG emissions but there are also many studies on smaller streams. In our eyes it does not make much sense to cite a small selection of the existing literature at this point. Since the point is not crucial for our study we would remove the sentence from the abstract.

L. 28 This may be a matter of taste, but could the title “Necessity of upscaling/quantification of GHG emissions from rivers” be shortened? Starting the manuscript with a less bulky title may approach a wider readership.

Good point. We will shorten to “Greenhouse gas emission from rivers”

L. 37 Raymond et al. (2013) relied mainly on calculated CO<sub>2</sub> based on pH, alkalinity and temperature, not “measured concentrations” as written here.

It was our intention to include both directly and indirectly measured GHG concentrations without going into detail on how concentrations are derived. We will clarify this to

“...measured in a restricted number of water samples or calculated from other parameters of the carbonate system (pH, alkalinity, and/or DIC).”

L. 38 Perhaps “gas transfer velocities” could be defined/introduced to make the manuscript more accessible for a wider readership?

We agree and will add a sentence explaining “gas transfer velocities”.

L. 38 The term “multiplied” confuses me, because the other terms of the equation that is referred to here (concentrations, gas transfer velocity) are simply mentioned without any mathematical characterization of their relationship. I suggest to rephrase the statement to be more consistent in the language.

We will add the actual equation to make this point clear.

L. 38 I agree that most datasets seem to contain weekly or monthly data, but could the authors provide (a) reference(s) for their statement? Perhaps a metaanalysis/review? For example, Marx et al. (2017, Reviews of Geophysics) mentions “knowledge gaps with respect to high-resolution temporal (i.e., diurnal) and spatial variations of carbon fluxes”.

Since we talk about typical datasets we think it is maybe more appropriate to cite a few such studies. In the revision we will add a few references.

L. 40/ L. 124 I agree with Lorke et al. (2015, Biogeosciences) that floating chamber measurements can be problematic in flowing water. This has also been evaluated by Vingiani et al. (2021, Biogeosciences) under a range of hydraulic conditions. I would appreciate if the authors could give more details in the methods section on their floating chamber design. How did the authors minimize potential experimental artifacts (e.g. by using “flying” chambers such as described by Lorke et al. and Vingiani et al.)?

We used a “drifting chamber” similar to the ones used in Lorke et al 2015. We will explain our chamber method in more detail in the method section.

L. 53-54 Please provide (a) reference(s) to support the statement “While a single water sample might be representative of a certain specific reach in a small stream this is undoubtedly not the case in larger rivers.” Why would spatial variability be higher in larger systems? Greenhouse gas fluxes can be highly heterogeneous in headwater systems (see e.g. Marx et al. 2017, Reviews of Geophysics; Lupon et al. 2019 L&O; Horgby et al. 2019, JGR). I am not aware of any systematic analysis of variability relative to system size, but I would be happy if the authors can substantiate their statement.

This point was also raised by another reviewer. We agree that spatial heterogeneity in rivers is not necessary larger than in small streams. Actually, we consider the comparison of spatial variability on different scales as a very interesting point. In a revised manuscript we will reformulate our statement in the sense that “less is known on spatial heterogeneity in rivers compared to what is known in streams”.

L. 80 Elsewhere in the manuscript it says the campaign was 3 days long, but here it says 4 days. Can you clarify this difference, please?

The duration of our campaign was in fact 4 days but some time was needed for installation and removal of instruments. Thus, time series data comprise up to 3 days of continuous data. We will change it to 3 days here.

L. 91 Why was the outer boundary of the groyne fields set to 15 m into the river? Is this based on previous research?

This is a misunderstanding. As written in our manuscript the outer boundary of the groyne fields were the line between two neighboring groyne heads. We decided to define the side area of the river extending 15m from that line into the river. We indeed had long discussion how to define the side area. We checked flow velocity and water depth but both were not systematically different between the middle of the river and the side. We finally decided to use a fixed distance because this enables easy quantification of areas in the GIS. We choose 15 m because that are approximately 10% of the river width. Visual inspection confirmed that 15m fully included the turbulent areas below the groyne heads.

L. 116 Measurements of turbidity and chlorophyll are mentioned here but there is no data shown. This should be consistent. I would be happy to see data on chlorophyll as it could indicate the level of primary production and hence provide important context to diel CO<sub>2</sub> concentrations.

A supplementary figure (S5-b) showing chlorophyll and turbidity will be added.

L. 118 I had to look up the term “moon pool of the albis”. I can imagine that there are more potential readers that are unfamiliar with this term. Consider clarification.

We will replace “moon pool” by "ship's duct with direct water supply".

L. 123-174 The authors used three different portable gas analyzers and a gas chromatograph. Have you performed cross-characterizations of the analyzers to make sure that concentration measurements and flux estimates are comparable between the study systems?

Yes. As written in the method section the measurements by the automatic equilibrator were corrected using GC samples. Before deployment the probes were checked in the laboratory by comparing probe readings with GC samples and/or separate measurements using a membrane equilibrator connected to an NDIR analyzer as explained in Koschorreck et al. 2021. The automatic chambers on land were not directly compared to the other analyzers but differences of atmospheric background concentrations measured by the various instruments were quite small (see figure below).





L. 139 Which “instruments” are referred to here?

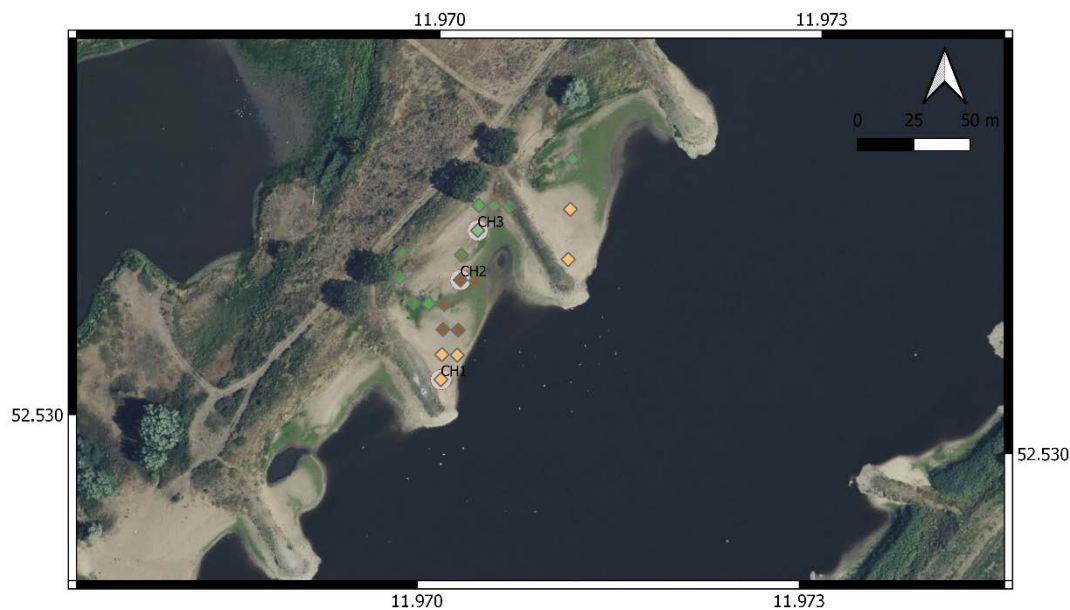
The degassing unit and greenhouse gas analyzer. Will be specified in the revision.

L. 153 I appreciate that CO<sub>2</sub> was measured in the air continuously. However, I cannot find any data on this in the manuscript or any statement on how the data was used. Did you use this data in flux calculations?

Atmospheric CO<sub>2</sub> data were indeed used to calculate fluxes from aquatic concentrations. This was necessary because there was a diurnal change of atmospheric CO<sub>2</sub> (see figure above). We will add this information to the method section.

L. 158 “Sampling points” are mentioned here, but I would appreciate a clarification of the exact sampling setup, perhaps already in the section with the study site description. How many chambers were deployed in total / per vegetation zone? This is implicit in Figure 1, but it is not clear to me until this point, whether chambers were deployed in all Groyne fields. Also, based on what criteria was the location of the soil flux chamber chosen? Fig. 1c) suggests that the vegetated site C3 was located very close to muddy area, which makes me wonder how representative this site was for the vegetated area?

The sampling points for the chamber were chosen to represent the three different habitats and to assess the temporal and spatial variation. We defined 19 sampling points representing the different habitats (yellow: sandy; green: vegetated, brown: muddy) including the points of the continuous soil flux chambers C1, C2 and C3. Measurements were taken at different times during the campaign in order to cover the diurnal variability. We will include the Figure below in the SI (Figure S9). The typical features of the different habitats are shown in figure S1. Muddy and sandy areas were free from vegetation and could be clearly distinguished from vegetated zones, which were widely covered by typical herbaceous plants such as *Persicaria*, *Inula britannica*, and *Xanthium strumarium*.





L. 179-182 Gas transfer coefficients were calculated from CH<sub>4</sub> fluxes and then converted to CO<sub>2</sub>. This conversion could potentially be erroneous in the presence of bubbles (Klaus et al. 2022, JGR), so I would appreciate a brief note on the role of bubbles in gas exchange in the study system.

As explained above we did not detect any ebullition. But surface bubbles might result from breaking waves. In our case the water surface was rather smooth without breaking waves. We will add this information to the manuscript. We also tried to quantify k<sub>600</sub> from our CO<sub>2</sub> data. But CO<sub>2</sub> concentrations were sometimes close to equilibrium resulting in large uncertainty in kCO<sub>2</sub> calculations. That's why we decided to use CH<sub>4</sub> derived k<sub>600</sub> also for CO<sub>2</sub>.

L. 185-186 It is unclear to me why "Probe measurements of CO<sub>2</sub> and CH<sub>4</sub> concentrations were converted to fluxes using the measured gas transfer velocity of k<sub>600</sub> = 5.5 m d<sup>-1</sup> (Table 1)." Why was this constant value chosen here, given that it varies substantially, as the data in Table 1 suggests.

We used the k<sub>600</sub> value measured at the side of the river since the probes were also installed at the sides. Since we did not measure k<sub>600</sub> exactly where the probe was installed and we wanted to be representative for the "site habitat" we used the mean k<sub>600</sub> measured in the "side habitat".

L. 186 Please clarify "converted to kCO<sub>2</sub> and kCH<sub>4</sub> as explained in Striegl et al. (2012)". Do you mean the Schmidt number conversion as explained in L. 182? What is the difference between the conversions you mention in L. 182 and 186?

There was actually no difference. We will replace "as explained in Striegl et al" by "as explained above."

L. 187-188 Please clarify the definition of "day" and "night". Some details are given in Table 2, but they should also appear or be moved to the methods section. Figure 5 suggests some offset between the timing of day-night shifts and changes in PAR. What is the reasoning behind this offset?

Based on sunrise and sunset we defined day as the period between 6:00 and 20:30. We checked our data and found that the PAR data indeed had a different time zone (CET) than the other data (UTC). We will correct Figure 5 accordingly and re-run the statistical analysis. We are sorry and thank you for discovering this.

Figure 3 I appreciate this map, but I wonder why spatial patterns are only shown for CH<sub>4</sub> and not for CO<sub>2</sub>? I would like to see a map of CO<sub>2</sub> measurements.

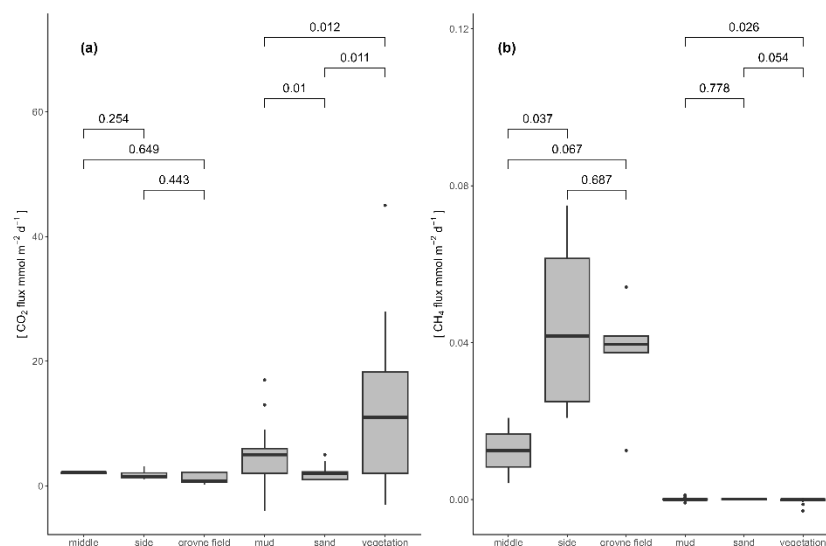
The green house gas analyzer could also analyze CO<sub>2</sub>, however the instrument has never been tested and calibrated for CO<sub>2</sub>. Thus, no such figure can be provided. As written in the caption of Table 1 the CO<sub>2</sub> concentrations were obtained from GC samples. We will check our raw data and figure out whether the CO<sub>2</sub> data of the GHG analyser can be trusted and eventually be used in the paper.

L. 226 What criteria did you use to extract the areas manually from a google earth image? Was there a clear division between the different areas or is the manual extraction prone to uncertainties/errors?

The criteria for the selection of the aquatic habitats was based on the groyne characteristics (groyne heads delimit the groynes, 10% transition area for side habitat, river area for middle habitat). For the terrestrial areas, the corresponding habitats were inferred on the basis of the structures on the land surface and the shade. We agree, there is a slight uncertainty in the manual determination of the areas. In order to minimize the error we based our estimation not only a single image (as written in the text). We derived the estimation from 5 scenes taken during different water level situations in the summertime from 2016 to 2022. In a new version of the manuscript, corresponding uncertainties regarding the area estimation will be included in Table 1.

Figure 4 It is not totally clear to me what the p-values of pair-wise comparisons refer to. Three values are given, but they are all aligned with the same arrow. Please modify the arrows so it becomes clear which p-value belongs to which comparison.

We will improve figure 4 as shown:



L. 235-236 Data on pH and O<sub>2</sub> is mentioned here, but I cannot find this data in Figure S5. Please add the data to the figure.

This figure will be added to the supplement as Figure S5c.

Figure 5 Please add units of light and temperature to panel c).

Will be added.

L. 259 / Figure S6 Why did you perform the correlation analysis only for CO<sub>2</sub> fluxes, but not for CH<sub>4</sub> fluxes?

Because we did not measure CH<sub>4</sub> fluxes with the automatic chambers. We will clarify this in the figure legend.

L. 260 Please clarify how you treated the “high” autocorrelation of light and temperature in the statistical analysis.

We did not treat it because both parameters were not part of the final model.

L. 267-268 The statement on fixed and random factors should be moved to the methods section.

We will move that sentence to the methods section.

L. 269 Do you mean the most parsimonious model?

Yes. Will be changed.

L. 279 Why did you not measure temporal changes in CH<sub>4</sub> flux at the dry sites?

We did not have an automatic CH<sub>4</sub> flux measuring equipment for dry sites. Because CH<sub>4</sub> fluxes at the dry sites were very low and previous research suggested very little temporal variability we decided not to do manual measurements in the night. A significant temporal variability of the methane concentration during the survey chamber measurements can be excluded. During the measurements, the continuous chamber in the muddy habitat was used as a base station and was approached and measured several times during a single measurement interval (approx. 2 hours). No temporal trend was detected.

L. 302 and L. 383 Can you provide a reference for the statement that CH<sub>4</sub> is primarily produced in the sediment? Methane can potentially also be produced in the water column, even under oxic conditions (e.g. Guentzel et al. 2019, Nat. Comm.). Is anything known about the sources of CH<sub>4</sub> in the Elbe river or other lowland rivers?

Yes – CH<sub>4</sub> can also be produced in oxic waters but rates are usually much lower than in anoxic sediments. We actually did water and sediment incubations to check for methane production. And yes, the data for methane production in water were much lower ( $1.73 \pm 0.46$  pmol/mL/h) than from the sediment slurries ( $1.07 \pm 1.69$  nmol/g dry weight/h). We did not add these data to the manuscript because this would blow up the methods section and we felt this information was not essential for the story of the manuscript. We will check whether these data can be included in a revision without making the paper too complicated.

L. 307 I cannot find any statistical support for the statement that there was “a significant difference between aquatic and terrestrial CH<sub>4</sub> emissions”. Please provide this support (e.g. in Fig. 4).

We will add “(Wilcox Test,  $p < 0.05$ )” to the text.

L. 319 I appreciate the pioneering effort on small-scale spatial variability in gas transfer velocities. However, the spatial variability is not explicitly shown in the manuscript. Data on k<sub>600</sub> is given in Table 1, but only median values and ranges are shown and it remains unclear whether they represent spatial or temporal variability. While showing variability in k<sub>600</sub> is not critical to the focus on fluxes in this paper, it might still be interesting to provide more detailed information on spatial vs temporal variability (e.g. CV) in k<sub>600</sub> and gas concentrations in the river. I leave it to the authors to decide whether they want to add this information to substantiate the statement in L. 319, or whether they want to leave this part of the story out.

We actually did not measure temporal variability of k<sub>600</sub> in our study because hydrodynamic conditions were rather constant during our study period. We will make this clearer in the text. It is actually a weak point of our study that we did not deploy GHG probes at different

locations to cover temporal variability at different sites. We think this is not a serious problem given the rather small differences in concentration between sites. As written in the manuscript to our knowledge our study is the first to look for small scale spatial variability of  $k$  in a river. We think it is a good idea to look for spatio-temporal variability of  $k_{600}$  on small scales in future studies.

L. 319-321 The floating chamber is not the only method applicable to rivers. See Huotari et al. 2013 (GRL) for deployment of the eddy-covariance technique in a river. Also, what is a “large stream” relative to a “river”? I suggest to use consistent terms throughout the manuscript.

We agree that eddy covariance can be used in (really) large rivers. However, river Elbe is only 150 m wide which means that the footprint of EC measurements would in most cases be “contaminated” by the river banks. In the study of Huotari et al. for example the river was more than 2 times wider than Elbe. Reliable data could probably only be obtained if the wind would blow exactly parallel to the river. In a river like Elbe this situation is rarely found given the fact that the river is meandering heavily. We will add a sentence discussing this.

According to Vannote et al. (1980) we define lotic waters larger than 6<sup>th</sup> order (Strahler) as rivers. Since the Elbe has Strahler order 8 we call it a “river”. Unfortunately, there is no common and sharp definition of a “large” stream - some scientists (probably used to very large rivers) do not even consider River Elbe to be large. However, River Elbe is among the list of large rivers (length >1.000 km) in Wikipedia, and fulfills the catchment size criterion (148.000 km<sup>2</sup> is >50.000 km<sup>2</sup>) (where discharge is dependent on precipitation). We will check the manuscript for consistency avoiding the term “large stream”.

L. 326-329 The statement “While higher  $k_{600}$  values at the side of the river were expected,” leaves me to wonder why you expected this. The explanation comes indirectly in the following sentence, but perhaps you can rephrase / reorder the sentence to improve logic?

We will rephrase this.

L. 331-333 Why do the authors refer to stream metabolism here? I agree that  $k$  is critical to metabolism calculations based on the free-water oxygen method, but there could be many other examples on exchange of other gases (e.g. Hg, Rn) where  $k$  values are relevant. Perhaps rephrase the sentence to reflect that metabolism calculations is just one example where  $k$  is relevant?

We will rephrase.

L. 342 Why did you exclude plants from the chambers?

This is an interesting point. Existing studies on GHG emissions from dry sediments largely exclude plants – often for practical reasons. This is why in terrestrial studies “ecosystem emissions” are separated from “soil respiration”. In case of river sediments we prefer to talk about “sediment GHG emissions” rather than “sediment respiration” because what we directly measure are emissions and these emissions can be affected by other processes than respiration (see Marcé et al 2019). Integrating temporary vegetation on dry sediments into the whole system carbon budget is tricky and would need a completely different approach (e.g. tracing the fate of vegetation after flooding). We will add 1-2 sentences about the complex role of vegetation to the discussion.

We are currently trying to investigate the effect of plants on dry sediments in the dryflux network (<https://www.ufz.de/dryflux/>).

L. 357 I would replace “obviously” by “most likely”, to reflect uncertainties.

We will replace it.

L. 362 Abbreviation “DIC” should be explained.

It’s dissolved inorganic carbon. Will be added.

L. 385 I don’t quite understand why the CH<sub>4</sub> pool in the water would buffer fluctuations in CH<sub>4</sub> emissions caused by CH<sub>4</sub> oxidation at the sediment interface. CH<sub>4</sub> could also potentially be oxidized in the water column. Some of the co-authors show this for the Elbe river estuary (Matousu et al. 2017, Aqua Sci). Can you please substantiate your argumentation, e.g. by referring to references?

The statement about buffering means that the pool of CH<sub>4</sub> dissolved in the water column is not much affected by hypothetical short term changes of sediment-water exchange of CH<sub>4</sub>. Thus, even if sediment CH<sub>4</sub> emissions would change diurnally this would not change much the aquatic CH<sub>4</sub> concentration and thus, the CH<sub>4</sub> emissions from the water surface.

Regarding CH<sub>4</sub> oxidation we will change the text to: “Methane consumption (oxidation) can occur either at the sediment surface or in the water column (Matousu et al 2019). A recent study however, suggests that this process is not influenced by light and thus daily variations (Broman et al., 2023).”

L. 405 / 414 Please use consistent terms to describe the “optimal” or “perfect” approach.

We will unify terms.

Supporting Information: Please clarify the symbols and units of the data included in the excel sheet that contains Time series data. Why are is the spatially resolved data not provided?

We will provide the spatially resolved data in a revision.

Figure S2 I think it would be useful to show a length metric as x-axis.

We will revise the figure accordingly.

Figure S4 Please add letters to the panels. Also, I think that either the panel headings or the the panel descriptions for b and c in the figure caption are mixed up.

Thank you for your recommendation. In the revision we will add the letters to the panels and revise the panel description.

Figure S6 Please give a clarification of the symbols / abbreviations used in the figure.

We will improve the figure legend.

**Technical notes**

We will address all technical points in the revision.

L. 21 improve -> improve

L. 54 or those or -> delete "or"

L. 98 The lines in Fig. 1 b are orange, not red as indicated here

L. 187 Here, both "emissions" and "fluxes" are used. I would suggest to use consistent terms throughout the manuscript. Personally, I prefer the more neutral term "fluxes".

L. 259 Remove "tried". It is apparent that you did the analysis.

L. 395 Remove "very"

L. 399 logyrhythmic -> logarithmic

L. 406 under estimation -> underestimation

## References

- Broman, E., Barua, R., Donald, D., Roth, F., Humborg, C., Norkko, A., Jilbert, T., Bonaglia, S. and Nascimento, F.J.A. 2023. No evidence of light inhibition on aerobic methanotrophs in coastal sediments using eDNA and eRNA. *Environmental DNA* 5(4), 766-781.
- Huotari, J., Haapanala, S., Pumpanen, J., Vesala, T. and Ojala, A. 2013. Efficient gas exchange between a boreal river and the atmosphere. *Geophys Res Lett* 40(21), 5683-5686.
- Marx, A., Dusek, J., Jankovec, J., Sanda, M., Vogel, T., van Geldern, R., Hartmann, J. and Barth, J.A.C. 2017. A review of CO<sub>2</sub> and associated carbon dynamics in headwater streams: A global perspective. *Reviews of Geophysics* 55(2), 560-585.
- Molodtsov, S., Anis, A., Li, D., Korets, M., Panov, A., Prokushkin, A., Yvon-Lewis, S. and Amon, R.M.W. 2022. Estimation of gas exchange coefficients from observations on the Yenisei River, Russia. *Limnol. Oceanogr. Methods* 20(12), 781-788.
- Stanley, E.H., Casson, N.J., Christel, S.T., Crawford, J.T., Loken, L.C. and Oliver, S.K. 2016. The ecology of methane in streams and rivers: patterns, controls, and global significance. *Ecological Monographs* 86(2), 146-171.
- Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. and Cushing, C.E. 1980. River Continuum Concept. *Can. J. Fish. Aquat. Sci.* 37(1), 130-137.