Reply to reviewer 2

We sincerely thank the reviewer for their valuable and in-depth feedback, which has been instrumental in improving the clarity, rigor, and overall quality of the manuscript. Below, we provide a detailed point-by-point response to each comment, with our responses highlighted in blue.

The manuscript by Zinke et al. presents data of the concentrations (and fluxes) of the three major greenhouse gases CO2, CH4, and N2O in a variety of shallow bays in the wider Stockholm Archipelago, Baltic Sea, using continuous recording of a system with an air-sea equilibration system coupled to a Picarro G 2508 CEAS. Measurements were done during two seasons, around midday, and a variety of potentially controlling environmental parameters was recorded as well. The authors use these data to calculate fluxes using the ASE parameterization developed by Cole and Caracoa (1998) for lake environments.

While the data set is interesting and the question of GHG emissions from shallow coastal waters and its potential changes is timely, the paper unfortunately does not explain parts of the methods and approach, and despite the fact that the authors claim the need to find ways to address the large spatiotemporal variability of GHG fluxes from these environments, they do by far not fully exploit their data set to answer these questions by e.g. stringent correlation analysis. The only correlation they address in more detail is the one between N2O and CH4, while the physical drivers including the - nicely introduced – topographic openness index – are not addressed in a systematic way.

We thank the reviewer for their thorough review. In the revised version we included a Spearman correlation analysis. Since the environmental parameters were only measured in one location of each bay, correlating these parameters across all bays only yields four data points in April and six data points in September/October. Given the limited number of data points we have therefore decided to pool the data from both April and September/October. A more detailed discussion of this matter will be addressed in the respective comment below.

Therefore, I can recommend publication of the paper only after major revisions.

I will start with some general comments and then get more detailed .

Scope: in the last paragraph of the introduction, the authors indicate the scope of their study, in particular naming the spatio-temporal variability and the examination of potential control parameters. However, they do neither discuss the limits of their approach towards these goals nor fully exploit them.

We thank the reviewer for this important comment. In the revised manuscript, we have explicitly clarified the scope and limitations of the study in the final paragraph of the Introduction and expanded on these points in the discussion. Specifically, we now acknowledge that the temporal resolution is restricted to two seasonal snapshots and midday measurements, precluding a full assessment of diurnal variability. At the same time, we have strengthened the exploitation of the dataset by adding correlation analyses between GHG concentrations and key environmental variables, and by more explicitly linking observed spatial patterns to potential physical and biogeochemical drivers. These additions better align the analyses with the stated objectives while maintaining a clear distinction between supported findings and more speculative interpretations.

Variability: it is not really clear how and when the sampling was done exactly. The authors state that the measurements were done during midday (line 81), but also describe a 45min measurement cycle which was repeated several times. From the details of the measurements mapped in the Figs Appendix

A1-A6, it appears that for the inner bays, only the long-term measurement at the positions indicated by red triangles were used. For "outside Bay area", there is no stationary point indicated so apparently the data outside a line indicating the boundary between inner- and outside bay were pooled. The approach of these outside Bay area measurements is not explained at all in the text, nor the selection of the separation or the question how "out" these areas actually are (in terms of connection to the open Baltic, residence time, depth etc.). These approaches should be explained in more detail.

Continuous flow-through measurements of surface water GHG concentrations were conducted around noon (typically between 11 AM and 2 PM) over a time scale of 60-90 minutes (typically ~75 minutes) depending on the size of the bay. The long-term measurements marked with the red triangle in figures A1-A6 are only for seawater properties (including nutrient concentrations, dissolved oxygen, chlorophyll-a, LOI, turbidity, pH, temperature and salinity). Temperature and salinity were also measured in the WEGAS system using a thermosalinograph – this data is available for every GHG data point. No long-term monitoring was conducted outside the bay area. We have now expanded the description of how inner versus outer bay areas were defined by adding the following sentences to section 2.2.1: "To distinguish between "inner bay" and "outer bay" sampling points, we delineated the bay boundary at the narrowest part of the inlet connecting each bay to the open Baltic Sea. This location represents the transition in water exchange, residence time, and mixing characteristics."

Temporal variability: while it is correct that the instrumentation used by the authors can be used to tackle temporal variability, it is not used that way in this study. Basically, the authors claim to have done experiments over midday, and it is not clear what time frame is represented in their individual data sets (as shown in the Figures A1-A6 of the appendix; a few hours is my guess). There is a complete lack of discussion on diurnal cycling and potential bias, in particular on the fluxes, while very relevant studies on the topic exist, e.g. the study by Honkonen et al. 2021 on diurnal variability of pCO2 fluxes at Utö (really nearby) or the recent study of Pönisch et al. (2025) on summerly GHG fluxes from a rewetted peatland (shallow coastal water). The latter has a detailed analysis on diurnal variation and the effect and biases this might have on GHG flux calculations, which the authors should address.

In the context of our study "temporal variability" refers more to seasonal differences between April and September. However, since our study provides only snapshots from two months and does not cover a full annual cycle, we would like to refrain from using the wording "seasonal". As already addressed in the previous comment, measurements lasted between 60-90 minutes (typically ~75 min). As such, the reviewer is right to point out that our sampling approach of measuring around midday does not allow to resolve diurnal cycles, leading to a systematic bias. We have added a discussion of this limitation along with references to Honkanen et al. (2021) and Pönisch et al. (2025) to the discussion (see corresponding comment further below). Furthermore, to avoid any confusion we removed "temporal" from the title and from the research aims stated at the end of the introduction.

Flux calculations: The authors should consider whether the flux calculations should be part of the paper. The strength of the manuscript is the measurement of concentrations and potential relation to controlling parameters. The flux calculations are based on the concentrations measured, a chosen ASE model (without discussion of the choice), and with a lack of wind data. If the flux part should be incorporated, it would be essential to

- discuss the potential daily bias in these shallow systems (see above).

We added a discussion of the daily potential bias in the context of under-/overestimation of our fluxes. A discussion is added to the corresponding comment below.

- to at least get access of wind data from an adjacent wind station of a reliable wind product. Just assuming an average wind speed of 2m per second is not state of the art.

We agree that wind speed is an important driver of GHG exchange across air–water interfaces. However, as mentioned in the manuscript, we did not measure wind speed during our deployments. Furthermore, for several of the bays, no nearby meteorological stations exist that would represent local wind conditions. The closest stations were often >20–40 km away and located in very different exposure settings. Using these stations would likely introduce bias rather than reduce it.

For this reason, we decided to use wind estimates from the ICON EU model (DWD), obtained at 10 m height and ~7 km horizontal resolution. While the model cannot resolve small-scale variability, it provides a consistent and physically based estimate of the regional background wind forcing at the standard 10-m reference height.

Importantly, field observations indicated that actual wind speeds at the water surface inside sheltered bays were low, often much lower than the 10-m model wind. This is expected due to sheltered nature of the bays, with limited fetch. Because gas-transfer parameterizations are conventionally defined at 10 m wind speed, model-derived 10-m wind is actually more representative for applying standard flux equations than the near-surface wind we experienced during sampling. As such, we consider the ICON-EU 10-m wind to be more appropriate than assuming a constant 2 m s^{-1} wind across all sites.

We have added the following text to the manuscript: “Wind speed at 10 m height was obtained from the ICON-EU numerical weather prediction model (Deutscher Wetterdienst, Germany). Model output at ~7 km horizontal resolution was accessed through the Ventusky online visualization platform (https://www.ventusky.com). We extracted 10 m wind values corresponding to the sampling dates and coordinates of each site. The derived wind speeds were 1.67 m s^{-1} (Sandviken), 7.0 m s^{-1} (Assöviken), 6.67 m s^{-1} (Högklykeviken), and 4.4 m s^{-1} (Bodviken) in April; 3.3 m s^{-1} (Sandviken), 3.9 m s^{-1} (Assöviken), 7.2 m s^{-1} (Högklykeviken), and 3.6 m s^{-1} (Bodviken) in September; and 2.5 m s^{-1} in both Östra Lermaren and Östra Myttingeviken in October.”

- To discuss the chosen wind model. This is of particular importance, as the wind model used in this study has a considerable wind-independent additive term, which dominates at windspeeds < 3.5 m/s. For systems like the ones investigated, there is no right or wrong choice, but some justification would need to be given. Way more important is that it would be important to look into the wind parameterizations used in the flux estimates worldwide the authors compare their results with (Table 5), which likely mostly used parameterizations without such a strong low-wind component.

We thank the reviewer for this important comment regarding the choice of gas-transfer velocity parameterization. Although our measurements were conducted in brackish waters (S ≈ 5 PSU), we selected the Cole & Caraco (1998) formulation because it was developed for shallow, sheltered, fetch-limited systems where gas exchange is not solely driven by wind speed. These physical conditions closely resemble those of our study bays, which are characterized by low currents, short fetch, and weak wind forcing.

Supersaturation of surface waters implies a positive air–water concentration gradient and thus outgassing even under low wind conditions; the Cole & Caraco parameterization accounts for this by allowing non-zero gas exchange at low wind speeds.

In contrast, parameterizations such as Wanninkhof (2014) are optimized for open-ocean conditions with large fetch and fully developed wave fields and are therefore likely to underestimate fluxes in

small, sheltered bays. Other parameterization developed for macrotidal estuaries (e.g. Borges et al., 2004) likely represent an upper bound for our system. A comparison of flux estimates using these three parameterizations shows that Cole & Caraco (1998) yields intermediate values, supporting its suitability for the physical setting of our study.

We have now added a discussion on the sensitivity of flux estimates to different gas-transfer parameterizations and highlight that methodological differences can substantially affect absolute flux values, thereby limiting direct comparability between studies:

"Estimates of air–water GHG fluxes are sensitive to the choice of gas-transfer velocity parameterization. In this study, we applied the formulation by Cole & Caraco (1998), which was developed for shallow, sheltered, fetch-limited systems and allows for non-zero gas exchange under low wind speeds. This is particularly relevant for the studied bays, which are characterized by weak currents and limited wind-driven turbulence. Alternative parameterizations such as the open-ocean parameterization of Wanninkhof (2014) or the estuarine parameterization of Borges et al. (2004) produce significantly lower or higher estimates, respectively. These differences highlight that absolute flux values are strongly dependent on the assumed turbulence regime and caution against direct inter-study comparisons without careful consideration of the underlying gas-transfer assumptions."

Assessment of drivers and potential controls:

The authors nicely introduced potential drivers of GHG production in the intro and beginning of the method, and took measures to investigate those (T,S, topographical openness, organic matter content, nutrient concentrations, …). However they report correlations only sporadically and without a sound estimate of significance etc. This is already documented in the separation of the GHG data and the "auxiliary" data in tables 2-4. The authors should try to give statistical underpinned information on correlations and potential driver analysis. A very good example for this is the use of Spearman correlation coefficients and effect size attribution as e.g. in Pönisch 2025.

We had initially considered to include a Spearman correlation analysis. However, since the environmental parameters were only measured in one location of each bay, correlating these parameters across all bays only yields four data points in April and six data points in September/October. Given the limited number of data points, most correlations were not significant if presented separately for April or September/October and we therefore initially decided not to include this analysis in the manuscript. In the revised version we have decided to combine the data from both periods and moved the discussion of environmental drivers from chapters 3.1.1-3.1.3 to a separate chapter. In cases were correlations were at or slightly above the significance level of 0.05, we refer to these correlations as "correlation trends".

"**Correlation of surface water GHG concentrations with environmental parameters across bays**

To identify environmental factors associated with variability in surface-water GHG concentrations, we conducted a Spearman's rank correlation analysis using bay-averaged GHG concentrations and environmental parameters measured in the center of each bay (Fig. 3). To increase statistical power and assess general trends, data from April and September/October were pooled.

CO_2 concentrations were positively correlated with LOI ($r = 0.67$, $p = 0.04$) and showed negative correlation trends with chlorophyll-a ($r = -0.64$, $p = 0.05$) and pH ($r = -0.59$, $p = 0.07$) as well as a positive correlation trend with rooted vegetation cover ($r = 0.60$, $p = 0.06$). The negative relationship with chlorophyll-a and pH suggests that periods or locations of enhanced primary production are associated with CO_2 drawdown and elevated pH, whereas positive correlations with LOI and

vegetation indicate that respiration and mineralization of organic matter—particularly from macrophyte-derived inputs—can offset photosynthetic uptake and elevate CO_2 concentrations in surface waters. This interpretation is supported by the observation that the bays with the highest CO_2 concentrations (Östra Lermaren, Östra Myttingeviken, and Bodviken) shared extensive rooted vegetation cover and elevated sediment organic carbon content. In Östra Lermaren and Östra Myttingeviken, which also exhibited the lowest eutrophication status as measured by TP and chlorophyll-a, high CO_2 concentrations may appear counter-intuitive but are likely driven by substantial autochthonous organic matter inputs from decaying vegetation, consistent with coastal studies documenting seasonal CO_2 hotspots linked to remineralization of organic-rich material (Amaral et al., 2021; Asmala and Scheinin, 2024). In contrast, Bodviken combined high CO_2 concentrations with comparatively higher eutrophication, suggesting that enhanced internal mineralization under nutrient-rich conditions may dominate CO_2 production in this system.
Although the correlations with pH and rooted vegetation were slightly above the conventional 5% significance threshold, they are consistent with the expected coupling between primary production, organic matter mineralization, and CO_2 dynamics in shallow coastal systems. Given the limited number of bays, these near-significant relationships should be interpreted as exploratory and warrant confirmation through studies with higher spatial and temporal resolution.

CH_4 concentrations showed a significant negative correlation with dissolved oxygen ($r = -0.75$, $p = 0.03$) and a positive correlation with loss on ignition (LOI; $r = 0.67$, $p = 0.04$). These relationships are consistent with enhanced methanogenesis under low-oxygen conditions and increased availability of degradable organic substrates in the water column, which together promote CH_4 production and accumulation.

In contrast, N_2O concentrations exhibited significant negative correlations with temperature ($r = -0.82$, $p = 0.01$), total nitrogen (TN; $r = -0.72$, $p = 0.04$), total vegetation cover ($r = -0.71$, $p = 0.04$), and rooted vegetation ($r = -0.78$, $p = 0.02$). These patterns suggest that warmer, more vegetated bays with higher nitrogen availability may favor complete denitrification to N_2 or enhanced biological nitrogen uptake, thereby reducing N_2O accumulation in the water column. In addition, vegetation-associated oxygenation of surface sediments may suppress N_2O-producing pathways while promoting N_2O reduction, contributing to lower observed N_2O concentrations.

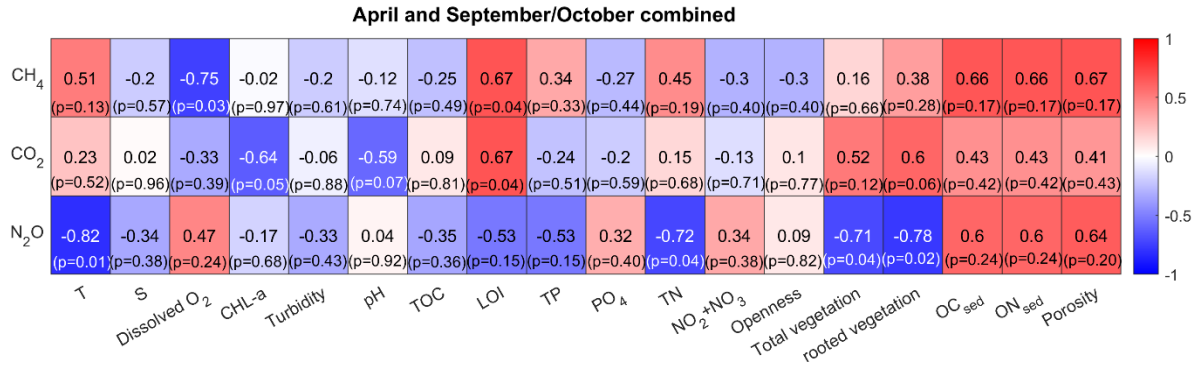

Figure 3 . Spearman correlation matrix between environmental parameters and CH_4, CO_2 and N_2O pooled for April and September/October. Blue indicates a negative correlation, red indicates a positive correlation. Significance (at the 95% confidence level) is indicated by p-values.

Additional to the heatmap plot, we added scatterplots to the appendix to visualize the relationships color-coded by season (see figures below).

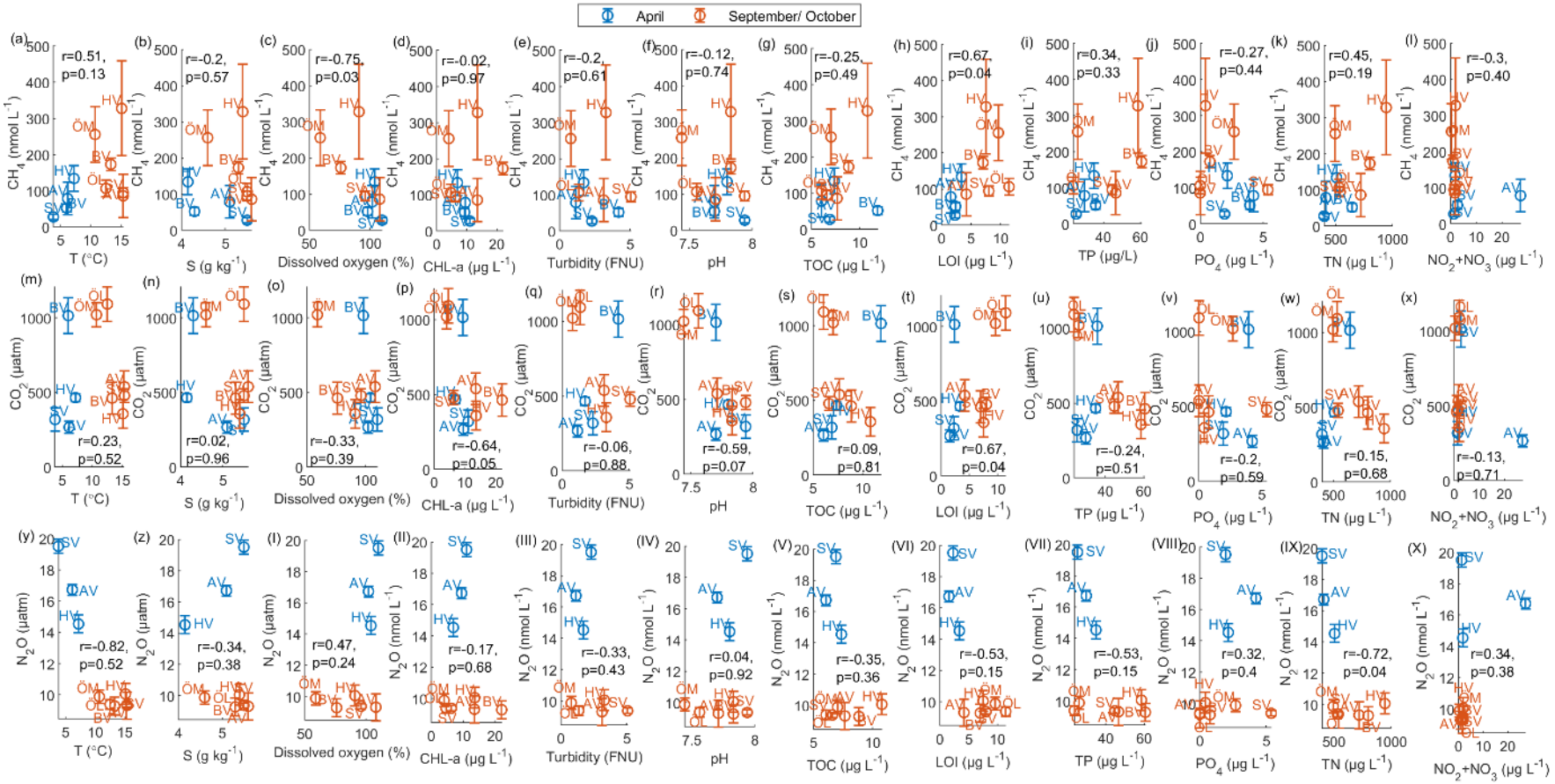

Figure A7: Relationships between seawater properties and GHG concentrations.

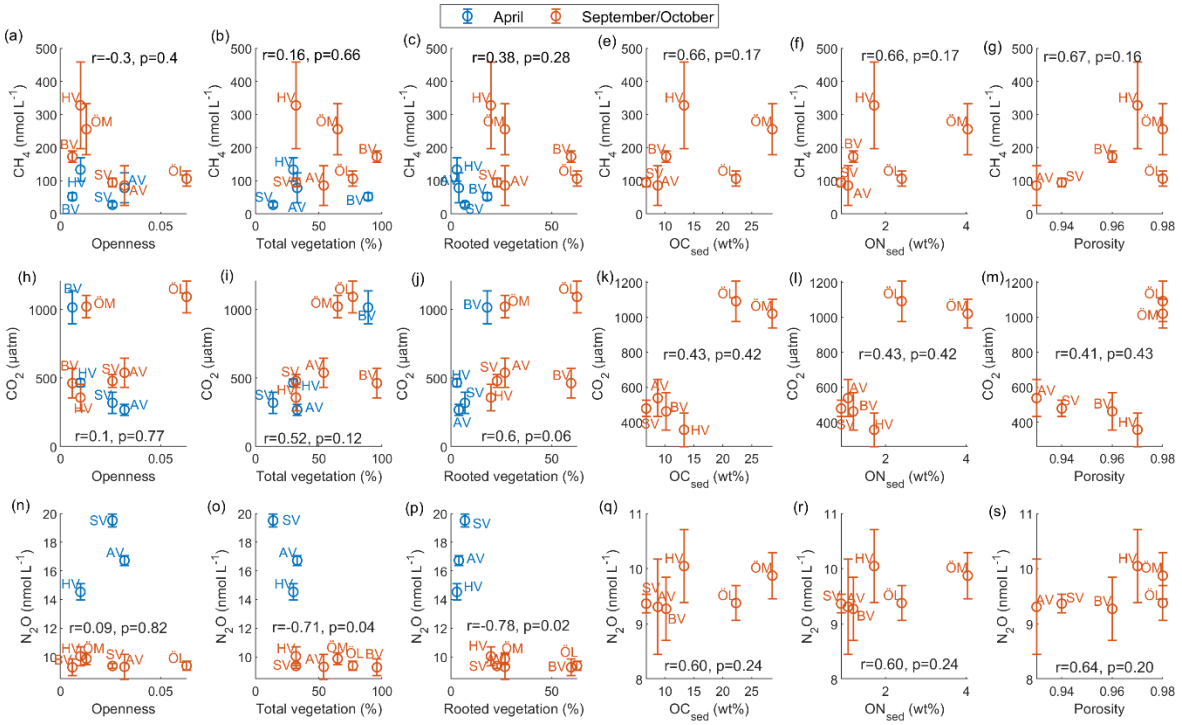

Figure A8: Relationships between bay characteristics (openness, vegetation cover and sediment properties) and GHG concentrations. Sediment data is only available for September/October.

The only quantitative analysis of correlations in this text is the one between N2O and CH4. While some of the discussion point in this section are really interesting, the interpretation of the correlation (and of the discussion on CH4 and N2O concentration suffers from the beforementioned lack of detailed assessment of the physical drivers. Here, for instant, it is clear that the main control of N2O is temperature, and this is likely also true (mentioned in the text) for methane, in the opposite way, as in the one case, control on solubility is dominant (N2O), while in the other, the control on production seems more important (CH4).
Stringent analysis could potentially help to reveal why and when the inverse correlation of CH4/N2O changes the direction at CH4 higher 250 nM.

We thank the reviewer for their valuable comment. We would like to clarify that the temperature for the measurements in Högklykeviken in September did not vary significantly (15.5℃ inside the bay versus ~14℃ outside the bay). As such we would like to argue, that the temperature effect on solubility and GHG production is likely minor. We added a discussion of alternative mechanisms that can change the CH_4–N_2O relationship in shallow bays. Unfortunately, sediment and seawater properties are only sampled in the center of each bay, limiting our discussion to speculations. In order to resolve which of these factors operates in our bays would require targeted process data. While porewater profiles and microbial gene assays were collected in the center of each bay on each sampling day, these yet remain to be analyzed and are beyond the scope of the present study.

We added the following text to the manuscript: "The different spatial distributions of CH_4 and N_2O may partly reflect their different optimal oxygen conditions: CH_4 production occurs mainly in anoxic regions, while N_2O production is maximal at suboxic levels near oxygen minimum zones where denitrification dominates (Naqvi et al., 2010; Ji et al., 2015, 2018; Foster and Fulweiler, 2016; Barnes and Upstill-Goddard, 2018; Tang et al., 2022). Although our measured dissolved oxygen levels

measurements in the central bay locations indicate generally oxic conditions in both Högklykeviken ($O_{2,dissolved}$ = 8.3 mg L^{-1} ≈ 91%) and Östra Myttingeviken ($O_{2,dissolved}$ = 5.6 mg L^{-1} ≈ 59%), we cannot resolve small-scale oxygen heterogeneity and therefore can only speculate that oxygen-reduced microenvironments may existed in areas of high CH_4 concentrations. Beyond oxygen availability, several additional mechanisms could explain the shift from a negative to a positive CH_4-N_2O correlation. As mentioned earlier, increased inputs of labile organic matter can stimulate methanogenesis further inside the bays, while changes in the availability of alternative electron acceptors (e.g., nitrate, sulfate, iron) alter competition among metabolic pathways, which can suppress or enhance methanogenesis and modulate N_2O production or consumption. Coupled processes such as nitrate-dependent anaerobic methane oxidation can also link CH_4 and N cycling in non-linear ways. Ebullition would provide a pathway for CH_4 accumulation by bypassing water-column oxidation and decoupling CH_4 from dissolved N_2O dynamics. However, as mentioned previously, our measurement set-up does not allow us to discern between bubble-mediated and diffusive CH_4. Changes in rooted vegetation and bioturbation may further modify sediment oxygen penetration and bubble release, influencing the relative dominance of CH_4 and N_2O-producing pathways. Finally, sediment disturbance from the research vessel in very shallow areas could explain these anomalous patterns. In order to resolve which of these factors operates in our bays would require targeted process data, limiting our discussion to speculations."

As pointed out correctly by the authors, the high variability in shallow coastal waters need more data, but also the best possible research to assign drivers and controlling parameters. Here, the paper could be largely improved.

We thank the reviewer for this comment. We agree that the high variability observed in shallow coastal systems highlights the need for both expanded datasets and robust analytical approaches to identify controlling drivers. In the revised manuscript, we have strengthened the analysis by adding correlation analyses between GHG concentrations and key environmental parameters, linking observed spatial patterns to potential physical and biogeochemical controls, and clearly distinguishing between statistically supported findings and more speculative interpretations in the Discussion. At the same time, we acknowledge the limitations of the present dataset with respect to fully resolving causal mechanisms.

Some areas of concern:

Askö Data: In Section 3.1.4, for the N2O -CH4 relationship, the data from observations in Asko are suddenly introduced, completely "out of the hat". Asko is a different setting, the method, sampling etc. has not been introduced, and there is no real link to the rest of the study. It honours the authors that they want to publish the data set, but it does not connect to the rest of the study. So I strongly recommend to remove this part, which is not introduced in the methods or site description, nor attached to the study. If these data are "left overs" they can be easily uploaded with relevant metadata to a publicly available data base. It appears that the authors were a little unsure here themselves, as if I am not mistaken, the data are not plotted correctly (Fig 3a). I am very sure that the data referenced as March data are actually September data and vice versa.

Based on the reviewer's suggestion, we removed the data from Askö in the revised version.

Method information:

There is a lack of information on some of the methodological aspects: how and when was the sensor calibrated with which kind of calibration gases? What is the expected accuracy / precision of the measurements? Which data were used for the calculation of the concentrations of the "inner bay" and the outside Bay (see further above). When calculating mean values (e.g. for fluxes), were they averaged simply over the No of regions, or were they area-weighted. Please add the information needed to assess the methodological part of the manuscript.

Thank you for this comment. We have now added detailed information on sensor calibration and instrument performance to the Methods section. The Picarro G2508 used in this study was factory-calibrated by the manufacturer in 2022, and the field campaign presented here represents its first deployment after calibration. According to the manufacturer's specifications, the 1-min precision is <300 ppb + 0.05 % of reading (typical 74 ppb) for CO_2, <7 ppb + 0.05 % (typical 0.1 ppb) for CH_4, and <10 ppb + 0.05 % (typical 1.1 ppb) for N_2O. These values are several orders of magnitude smaller than the concentration ranges measured in our study (CO_2: 300–1000 ppm; CH_4: up to 250 ppm; N_2O: 0.34–0.36 ppm), ensuring that instrumental precision does not constrain the interpretation of our results. The G2508 operates using cavity-ring-down spectroscopy (CRDS), which exhibits negligible long-term drift according to the manufacturer's documentation. Thus, the probability of significant drift between factory calibration and our field deployment is low. Furthermore, because our comparisons focus primarily on relative differences among bays measured with the same instrument under identical conditions, any potential systematic bias would not affect the observed spatial patterns. To further assess potential drift, we conducted a post-study calibration against certified standard gases (only for CO_2 and CH_4). We found that since the last calibration, the slope only changed by 0.2% for CO_2 and -0.5% for CH_4 suggesting that spatial and temporal differences are not affected by potential sensor drift. Assuming monotonic drift, the offset applicable to the April and September measurements was estimated by linear interpolation and ranged from -0.4 ppm for CO_2 and 0.03 ppm for CH_4 for the April measurements to -1.2 ppm CO_2 and 0.04 ppm CH_4 for the September measurements, which is much smaller than the measured values in this study. This information was added to the manuscript section 2.2.

We have added a description of how inner versus outer bay areas were defined (see response to comment above).

Furthermore, we have clarified how mean values were determined in the methods section and throughout the manuscript. The following sentence was added to the methods description: "For cross-bay comparisons, concentrations were first averaged within each bay, and summary statistics (e.g., median) were then calculated across bays using one value per bay, treating each bay as an independent unit rather than applying area-weighted averaging."

Missing N2O data for Bodviken: this should be explained. It is mentioned several times in the text, and difficult to understand, as the authors use an instrument measuring all 3 gases simultaneously, and usually the instrument fails completely or not; there is surely an explanation, but please explain in the text.

Due to technical issues, measurements in Bodviken in April were conducted with a G2201-i Picarro instead of a G2508. Since the former does not measure N_2O, this data is not available for this sampling day in April. We have added the following text to section 2.2: "April measurements in Bodviken were conducted using a Picarro G2201-i instead of the G2508, which measured the concentrations of CO_2 and CH_4 but not N_2O."

Minor issues:

Abstract:

Line 7: “seasonal variation concentrations” – Wording

We have corrected the wording to “seasonal variations **of** concentrations”

Line 10: It is unclear why the finding of a shift in N2O-CH4 relation slope indicates a shift in biogeochemical processes; see also comments further above

We agree that this point required clarification. We have added a more in-depth discussion of different factors that could potentially impact the CH_4-N_2O relationship in chapter 3.1.4 (see earlier comment). In order to keep the abstract concise, we refrained from discussing the biogeochemical processes in detail and instead rephrased the text as follows: “CH_4 concentrations below 250 nmol L^{-1} negatively correlated with N_2O, while higher CH_4 levels showed a positive correlation, suggesting differences in the dominant sedimentary microbial pathways**.**”

Line 11: what is meant by “anthropogenically degraded”?

This particular bay is highly impacted by anthropogenic influences such as dredging, input of nutrients and high boat traffic, as reflected in the elevated TN concentrations measured in this bay. We have rephrased the text as follows: “One bay that is subject to substantial human impacts (e.g. dredging, high nutrient loading, reduced vegetation cover) […]”.

Line 12:”methane emissions that surpasses CO2 update”: first of all, most regions had no CO2 update; second: this is not clear without introducing the fact that the authors calculated CO2 equivalent fluxes – WORDING

We added the following text (in bold) to be more specific: “**CO_2-equivalent** CH_4 emissions that surpassed CO_2 uptake **in this particular bay**”

Intro:

Line 33: strictly speaking, N2O is a by-product of nitrification, but an intermediate of denitrification

We now specify this is in the text: “N_2O is generated as a by-product of nitrification or as an intermediate of denitrification”

Line 37-38: statement on aerobic oxidation needs a reference

We have added the following references:

Hanson, R. S., & Hanson, T. E. (1996). Methanotrophic bacteria. *Microbiological reviews*, *60*(2), 439-471.

Venetz, J., Żygadłowska, O.M., Dotsios, N., Wallenius, A.J., van Helmond, N.A., Lenstra, W.K., Klomp, R., Slomp, C.P., Jetten, M.S. and Veraart, A.J., 2024. Seasonal dynamics of the microbial methane filter in the water column of a eutrophic coastal basin. *FEMS microbiology ecology*, *100*(3), p.fiae007

Line 39-44: the manuscript completely ignores anaerobic methane oxidation, both here, and also in discussing the effect of T on methanogenesis; in fact, both methanogenesis and methanotrophy are T-dependent.

Here, we added: “dissolved CH_4 may be aerobically oxidized by methanotrophic bacteria (Hanson, 1996) **or consumed by anaerobic methanotrophic archaea (Knittel & Boetius, 2009).**

Knittel, K., & Boetius, A. (2009). Anaerobic oxidation of methane: progress with an unknown process. *Annual review of microbiology*, *63*(1), 311-334.

Line 55-64: see comment above; the authors did not really fully exploit the possibility to scaling based on driver analysis, nor did they really make use of the high temporal resolution in this study.

In the revised version we have added a Spearman correlation analysis. However, contrary to our expectations we could not show distinct correlations between GHG concentrations and seawater/sediment properties indicative of an eutrophication gradient. This is likely due to the limited number of data points - seawater/sediment properties were only measured in one central location of each bay. With respect to temporal resolution, we have now addressed the limitation of our study that our measurements were always conducted during daytime and thus cannot resolve a diurnal cycle, leading to a systematic bias when upscaling our measurements to fluxes. We have addressed this issue in more detail in the corresponding comment above.

Methods:

Chapters 2.1 and 2.2.: as stated above, give more details on method, and introduce the outer Bay sites; also please explain which data were used for the inner and outer Bay mean concentration value (just on red triangle spots, or all data in the inner Bay; outer Bay I cannot tell, it is so far not indicated in the text at all. Also, be more specific about the timing and duration of the data acquisition (midday vs. several hours of measurements).

We added the following text to section 2.2.1: “Sampling durations lasted between 60 and 90 minutes (typically ~75 minutes). Measurements were conducted both inside and outside bay areas. To distinguish between “inner bay” and “outer bay” sampling points, we delineated the bay boundary at the narrowest part of the inlet connecting each bay to the open Baltic Sea. This location represents the transition in water exchange, residence time, and mixing characteristics.”

For continuous GHG measurements, all data points with in inner and outer bay areas were used. Seawater and sediment properties were only sampled in one location in the center of each bay (as indicated by the red triangle and grey dot in Figures A1-A6). To conduct the Spearman correlation analysis, GHG measurements were averaged per bay and correlated with the seawater/sediment properties.

Line 95: Partial pressures do NEVER have the unit ppm

We adapted the text as follows: “Mole fractions (in ppm)”

Line 98: “ atm=106 ppm” ; same, please be scientifically correct here

We adapted the text as follows: “1 ppmv corresponds to 1 µatm at an ambient pressure of 1 atm”

Section 2.2.3. See remarks above on wind speed data availability, discussion of use of ASE exchange parameterization and its implication for comparison of fluxes with other studies

As addressed above, we now use wind data from the ICON-EU model for the respective sampling locations and days. In the response to the earlier comment on the k-parameterization, we also added

a discussion why we chose this particular parameterization. Furthermore, we now address the use of different parameterizations in comparison to other studies (table 5).

Section 2.3.4: Line 158: "..most relevant for …" - not fully true for methane, where deeper sediment information might me needed to understand flux or ebullition behavior

We agree that for methane, processes occurring deeper in the sediment column, such as methanogenesis and ebullition, can play a critical role in determining fluxes to the water column and atmosphere. We have therefore revised the wording. In the revised manuscript, we clarify that the uppermost sediment layer (0–1 cm) is most relevant for characterizing sediment–water exchange processes and redox-sensitive transformations directly affecting dissolved GHG concentrations in surface waters, while acknowledging that deeper sediment layers may be important for understanding total CH_4 production and ebullition dynamics, which were not explicitly resolved in this study.

Results and Discussion

3.1 Outside Bay areas not mentioned or described before or anywhere after in the text (se general comment further up)

Thank you for pointing out the need for clarification. We have now expanded the description of how inner versus outer bay areas were defined by adding the following sentences to section 2.2.1: "To distinguish between "inner bay" and "outer bay" sampling points, we delineated the bay boundary at the narrowest part of the inlet connecting each bay to the open Baltic Sea. This location represents the transition in water exchange, residence time, and mixing characteristics."

Lines 171-173: The lack of consistent pattern between bay openness and CO2 concentrations (see more general comment above on quantitative correlation analysis) indicates that there is no major control of this parameter on CO2 concentration, but not that there are high spatial and temporal variability in CO2dynamics". This is no substantial statement.

We now removed the second part of the sentence "~~indicating high spatial and temporal variability in CO_2 dynamics~~"

Lines 184 and 185 „... share ...GHG emissions" ; Bold and unsupported statement, based just on the similarity of a mean concentration value (for CO2). So why is this used to speculate on GHGs in general. For CH4, for instance, it is well established that freshwater and brackish water are quite distinct due to the role of processes involving sulphate.

We thank the reviewer for their comment and agree that the respective sentence is an overstatement. We have now replaced it with the following sentence: "The sheltered nature of the bays may resemble lake-like conditions with respect to air–water CO_2 exchange, but not necessarily other gases"

Lines 188-190. This might be a good place to extend on related work who actually did that, also showing how important it is, like e.g. Honkanenen et al. 2021 or Pönisch et al. 2025. However, this is again a statement true for all three GHGs, so maybe the best place to discuss this, and with it the limits of the study here, would be after the reports on the individual GHGs.

We have added the following text to discuss diurnal variability and potential bias of sampling at the same time: "Recent studies by Honkanen et al. (2021) and Pönisch et al. (2025) reported diurnal variability in surface-water pCO_2 and CH_4 in the Baltic Sea that could be linked to biological and physical drivers such as solar radiation, temperature or biological activity. We acknowledge that our

measurements, which were always conducted around noon, do not capture these diurnal fluctuations and thus likely introduce a small but systematic bias relative to true daily mean conditions. While such measurements remain valuable, more extensive, long-term monitoring is required to identify the environmental parameters that drive these systems to function as CO_2 sources or sinks across different temporal scales."

Line 204-206: „while ...aerobic methanotrophs: Missing reference. Also, there is also a wealth of literature on rooted vegetation actively transporting methane, which would escape the flux measurements. This should be briefly mentioned as well.

This sentence was removed in the revised version as it no longer aligns with the results of the Spearman correlation analysis of the pooled data.

Lines 216-217: while it is true that the system would measure dissolved methane from the seafloor and dissolved by bubble dissolution, it would be then important to state that bubble-mediated transport itself escapes the device, and in fact that trapped bubbles could be an issue for the measurements. Maybe the authors did some work to protect the inlet from ascending bubbles? If so, it would be good to mention this in the method section.

In this study we did not take measures to protect the inlet from ascending bubbles. We have added the following sentence to the text "Another factor that can contribute substantially to CH_4 emissions in shallow, organic rich sediments is ebullition (McGinnis et al., 2006; Hermans et al., 2024; Bisander et al., 2025). Recently, Bisander et al. (2025) showed that ebullition from sandy sediments can be substantial. The WEGAS system measures CH4 from both benthic diffusion and bubble dissolution. Consequently, the observed CH4 concentrations represent the combined effect of these pathways, and without isotopic information we cannot distinguish between diffusive transport and ebullition. Although no visible bubbling was observed during sampling, we cannot exclude the possibility that episodic ebullition events might have impacted our measurements. This measurement limitation should be considered when interpreting the relationships between CH4 and the environmental parameters described above."

Line 210: enhanced temperatures also enhance aerobic oxidation rates, though apparently the effect on methanogenesis "wins" here.

We added "Elevated temperatures accelerate […] methanogenic rates [...] **as well as aerobic oxidation rates**."

Line 211-212: Stratification would enhance concentrations below the stratification gradient, but lead to lower concentrations in the top layer. Please be clear in your argument here.

The reviewer is right to point this out. We have now rephrased the text as follows: "Warmer seawater temperatures also decrease CH_4 solubility and enhance stratification, resulting in elevated CH_4 concentrations below the thermocline and lower concentrations above it."

Lines 216-217: while it is true that the system would measure dissolved methane from the seafloor and dissolved by bubble dissolution, it would be then important to state that bubble-mediated transport itself escapes the device, and in fact that trapped bubbles could be an issue for the measurements. Maybe the authors did some work to protect the inlet from ascending bubbles? If so, it would be good to mention this in the method section.

See answer above.

Chapter 3.1.3 as mentioned in general comments; a quantitative discussion of N2O concentrations in relation to temperature is needed, which is clearly the dominant driver. Or the discussion could be based on the calculated N2O saturations which eliminate this effect. Here, it would help to distinguish whether a part of the difference is temperature rather than production related.

We have added the following discussion of the temperature effect on N_2O to the manuscript (in context of the added Spearman correlation analysis: "The negative relationship of N_2O and temperature is likely driven by two key factors: (1) increased N_2O solubility at lower temperatures, and (2) the temperature sensitivity of denitrification enzymes. Under low-temperature conditions, enzymatic activity of N_2O reductase may be reduced, potentially slowing conversion of N_2O to N_2 and thereby increasing net N_2O emissions (Wang et al., 2014)."

Tables 2-4: again consider the general comment on integrated research on the drivers and the concentrations.

Following the reviewer's general comment on integrating drivers with concentrations, we have now included a Spearman correlation analysis to contextualize the values presented as environmental drivers of GHG concentrations across the bays.

Chapter 3.1.4. while some aspects of the discussion are quite interesting, the analysis suffers from the investigation of T as mean driver before; here again having a look on whether deviations could be found on the saturation level could overcome this. Most of the N2O data will be on a straight line then, and it would be interesting to see how the trend at CH4 > 250 nmol displays on the saturation level.

Changes in temperature could be a reason for the difference in N_2O concentrations in April versus September/October. However, the temperature range encountered in Högklykeviken in September was narrow (14–15.5 °C), and therefore solubility effects were limited. Presenting the relationships in saturations instead of concentrations exhibits the same patterns (see figure below). As such we have decided to keep presenting the data in concentrations.

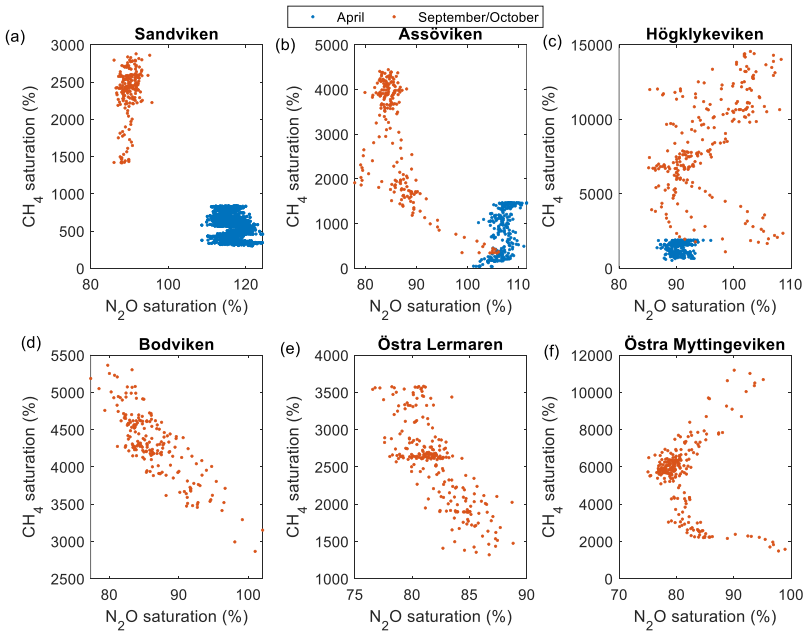

Figure: Correlations between N2O and CH4 saturations across different bays and seasons

Still, see my general comment on semi-quantitative statistical evaluation.

We understand this comment as referring to the limited quantitative attribution of drivers underlying the observed CH_4–CO_2 relationships. In the revised manuscript, we have addressed this by adding a correlation analysis between GHG concentrations and key environmental parameters. We have also revised the discussion to avoid causal interpretations and instead discuss the observed relationships in the context of established biogeochemical processes reported in the literature.

Line 268-272: "Since …. . … anomalous patterns". This section is purely speculative.

We thank the reviewer for pointing this out. We agree that the original text implied mechanistic explanations that cannot be supported by the available data. Oxygen was measured only at the central bay stations, and therefore local conditions at the CH_4 and N_2O peaks are unknown. To avoid speculation, we have revised the paragraph to remove unsupported mechanisms and now emphasize the limits of inference and the need for additional measurements to resolve these small-scale anomalies: "Although our measured dissolved oxygen levels measurements in the central bay locations indicate generally oxic conditions in both Högklykeviken ($O_{2,dissolved}$ = 8.3 mg L^{-1} ≈ 91%) and Östra Myttingeviken ($O_{2,dissolved}$ = 5.6 mg L^{-1} ≈ 59%) we cannot resolve small-scale oxygen heterogeneity and therefore can only speculate that oxygen-reduced microenvironments may existed in areas of high CH_4 concentrations. Beyond oxygen availability, several additional mechanisms could explain the shift from a negative to a positive CH_4-N_2O correlation. As mentioned earlier, increased inputs of labile organic matter can stimulate methanogenesis further inside the bays, while changes in the availability of alternative electron acceptors (e.g., nitrate, sulfate, iron) alter competition among metabolic pathways, which can suppress or enhance methanogenesis and modulate N_2O production or consumption. Coupled processes such as nitrate-dependent anaerobic methane oxidation can also link CH_4 and N cycling in non-linear ways. Ebullition would provide a pathway for CH_4 accumulation by bypassing water-column oxidation and decoupling CH_4 from dissolved N_2O dynamics. However, as mentioned previously, our measurement set-up does not allow us to discern between bubble-mediated and diffusive CH_4. Changes in rooted vegetation and bioturbation may further modify sediment oxygen penetration and bubble release, influencing the relative dominance of CH_4 and N_2O-producing pathways. Finally, sediment disturbance from the research vessel in very shallow areas could explain these anomalous patterns. In order to resolve which of these factors operates in our bays would require targeted process data, limiting our discussion to speculations. "

Figure 3: Plot a should be removed, as not part of the presented study (and likely wrong seasonal attribution)

We now removed subplot a) based on the reviewer's suggestion.

3.2.

Line 283-284 „... similar to the complex ..." This similarity is an artifact of using basically a mostly constant transfer coefficient (i.e. fixing the wind speed); see general comment.

In the revised version, we use wind speeds from the ICON-EU model at the respective sampling locations and days and have now removed the sentence highlighted by the reviewer.

295-296: Again: if flux discussion remains, nearby wind speed should be addressed, as well as the model of use and its impact for the flux comparison of table 5.

We added a discussion of the k-model used in this study in comparison to the parameterizations used in other studies: “Furthermore, estimates of air–water GHG fluxes are highly sensitive to the choice of gas-transfer velocity parameterization. In this study, we applied the formulation by Cole and Caraco (1998), which was developed for shallow, sheltered, fetch-limited systems and allows for non-zero gas exchange under low wind speeds. This is particularly relevant for the studied bays, which are characterized by weak currents and limited wind-driven turbulence. Alternative parameterizations such as the open-ocean parameterization of Wanninkhof (2014) or the estuarine parameterization of Borges et al. (2004) produce significantly lower or higher estimates, respectively. These differences highlight that absolute flux values are strongly dependent on the assumed turbulence regime and caution against direct inter-study comparisons without careful consideration of the underlying gas transfer assumptions.”

3.3.

Line 301 “ … with a median …” Here and elsewhere: please describe how data were averaged (just one No pe area, or “area weighted …”.

For each bay, we first calculated the mean using all measurements collected within that bay. To summarize conditions across bays, we then calculated the median of these six bay-level values, treating each bay as one independent unit. This approach is intentionally not area-weighted, as our aim was to compare bays as ecological entities. We now explicitly state this in the revised manuscript.

Line 309: To assess the regional significance, one would need to know typical flux estimates from similar areas on land or over open water.

We thank the reviewer for this important comment. However, a direct comparison of area-integrated CO_2-equivalent fluxes across ecosystems is challenging, as studies often report different subsets of greenhouse gases and cover vastly different spatial extents. In section 3.2 and table 5 we present areal flux intensities in comparison with other aquatic systems while the area-integrated fluxes presented in section 3.3. are intended only to provide a first-order estimate of potential regional relevance and are not directly comparable across studies or ecosystems due to differences in spatial extent and the inclusion of different greenhouse gases. We have added the following sentence to the discussion of the area-integrated fluxes: “Scaling these fluxes to the estimated total area of shallow enclosed bays provides a first-order indication of their potential regional relevance, but should not be interpreted as a closed regional budget due to spatial heterogeneity and limited spatial coverage.”

Figure 4: I found it really interesting that Högklyeviken is characterized by highest CH4 fluxes, but also as the only region with a net carbon dioxide uptake in both seasons; do the authors see any reason for that (I cannot see a link to the fact that the phosphate-binding experiment took place there …). This is just reviewer’s curiosity, so ignore if not worthwhile pursuing …

We thank the reviewer for their comment and agree that the high CH_4 fluxes combined with net CO_2 uptake in Högklykeviken warrant explanation. The CO_2 uptake likely reflects short-term dominance of pelagic photosynthesis during sampling, while the elevated CH_4 emissions point to active anaerobic processes in the sediments—processes that can decouple in time and space. Högklykeviken was selected for the phosphate-binding experiment because it was expected to be the most disturbed bay, but our environmental measurements show no significant differences to the other bays in terms of nutrient status or other indicators of anthropogenic degradation.

Conclusion

Line315: „.. shallow Baltic Sea“ – I think the authors should make clear that the data are from a very small part of the Baltic Sea and surely not fully representative for the (entire) Baltic Sea.

We have now replaced “Baltic Sea” with “wider Stockholm archipelago”

Line320 : …dominates CO2-equivalent fluxes.

We added “CO2-equivalent”

Line 324-329 surely need revision after addressing the review’s comments.

In response to the reviewer’s comments, we have revised this paragraph to better align our conclusions with the semi-quantitative nature of the analysis, to moderate claims of novelty, and to more clearly distinguish observational patterns from mechanistic attribution: “In the two bays with the highest concentrations of CH_4, we observed a change in the relationship between CH_4 and N_2O, with negative correlations at CH_4 concentrations below 250 nmol L^{-1} and positive correlations at higher concentrations. To our knowledge, such a pattern has rarely been reported for shallow coastal bay environments and highlights the complexity of coupled nitrogen and carbon cycling under variable redox and hydrodynamic conditions. This shift likely reflects a transition from conditions where nitrification and coupled nitrification–denitrification dominate to more reduced, microbially active regimes in which methanogenesis become more prevalent.

By placing GHG concentrations and fluxes in the context of measured environmental parameters, this study identifies observational relationships between bay characteristics and seawater properties with variability in coastal GHG dynamics. CO_2 was negatively correlated with chlorophyll a and pH and positively correlated with LOI and rooted vegetation, while CH_4 was negatively correlated with dissolved oxygen and positively correlated with LOI. N_2O was negatively correlated with seawater temperature, TN, total vegetation and rooted vegetation. At the same time, the pronounced spatial and temporal heterogeneity across bays and seasons, together with the limited number of study sites, constrained our ability to quantitatively attribute individual drivers, underscoring the need for targeted process-based studies to resolve the mechanisms underlying these patterns.”

Line 332: “However …. budgets” - shallow bays are not really overlooked, and so far, the data presented were quite in line with earlier findings; so I do not understand the statement

We removed “overlooked”

Line 347: Data Availability:It is of utmost importance that all data , also the auxiliary data, are publicly available upon publication of the paper, once accepted.

We will submit our data to the SOCAT and MEMENTO database once the manuscript revision has been finalized.