Referee #2

General Comments:

The manuscript presents an interesting comparison between measured methane concentrations and reported emission factors at a large metropolitan wastewater treatment lagoon. The use of Inverse-Dispersion Modelling (IDM) with open-path lasers (OPL) is a robust approach for non-intrusive monitoring; however, the current manuscript has significant gaps regarding spatial weighting, micrometeorological artifacts, and the representativeness of the upscaled annual data. Without addressing these potential biases, the conclusion that emissions are 2–3 times higher than NGERS reports may be premature.

We thank the Reviewer for their thorough and constructive review of our manuscript. Below, we address each of the reviewer's comments. We also outline additional corrections identified during the revision process, in addition to those required to address the comments from each referee. All the changes can be found in the revised manuscript.

Specific comments:

1. How does the "touchdown" coverage (mentioned as >20% in Section 2.5) vary across the pond surface? Does the weighting reflect the actual spatial concentration gradient?

It depends on the wind direction, but the source area can be seen when the wind's direction is within a range of >330° to < 40°. This is when the measurements can be considered for representing the source emission and not contaminated by nearby sources (e.g., eastern side and western side).

As the pond is particularly large and that the downwind open-path laser parallel to the pond, that is set with a typical pathlength of about ~200 m to maintain its stability and performance, we can only sample a portion of the pond emission, in addition to limited resources, we deployed two open-path lasers downwind the pond, by dividing the pond into two half, one laser can be focused on the pond area close to the anaerobic pod and the other one in contrast to the western side of the pond.

The touchdown coverage varies with wind direction and atmospheric conditions. We also analysed its spatial distribution. The footprints showing its spatial distribution indicate which parts of the lagoon contributed to the OPL-IDM estimate for each laser path.

However, the footprint does not provide spatially resolved concentrations within the sampled area. For each laser path and time step, we obtain one aggregated estimate for the area sampled by that footprint. Therefore, the footprint tells us where the measurement is coming from, but not how methane concentration varies within that area.

The weighting does not represent a measured spatial concentration gradient. Instead, it is used to scale the two aggregated estimates, one from the west footprint and one from the east footprint to the full lagoon area. We clarified this in the revised text and assessed the sensitivity of the full lagoon estimate to alternative weights.

2. Is the 8:00 AM peak a result of a biological process, or is it a micrometeorological artifact caused by the breakup of the nocturnal boundary layer? At 8:00 AM, the atmosphere often transitions from stable to unstable, which can "dump" accumulated methane toward the sensors. In the early morning, the "surface layer" might not be fully developed. Are

you seeing a real biological emission peak, or is it just the "fumigation" of methane trapped near the water surface overnight being released as the sun hits the pond?

This diurnal pattern on the eastern pond area, with peak emission at ~8:00, was associated with operation activities (aerator on/off and influent flow rate high/low) rather than micrometeorological artifact. A similar diurnal pattern has also been reported in the literature (Mannina et al., 2018; Bühler et al., 2022; Guisasola et al., 2008). Ebullition (bubbling) and diffusion (dissolved methane in the effluent) are the dominant pathways of methane emissions and can be influenced by a range of locations-specific, biotic and abiotic factors.

As sewage from the anaerobic digestion pod enter the Pond 1, the influent is saturated with dissolved methane. When the aerators are switched on, they not only rapidly strip dissolved methane from the water surface to the atmosphere but also generate turbulence in the sediment-water system in the pond, contributing to diffuse emission from the lower to the up layers of the pond and ultimately to the atmosphere. The influent flow rate plays an important role in regulating methane emissions: it exhibits a similar 24-hr diurnal variation as CH_4 emissions but with a time lag of 1-2 hrs. Furthermore, other studies suggest the peak emission in the morning could be due to the sewage system, where higher methane production with longer retention time in the effluent, and methane could also accumulate in sediments (Guisasola et al., 2008). If aerators are turned off in the afternoon, methane could build up overnight and then be released the following morning when aeration resumes (Bühler et al., 2022).

3. I am wondering why a simple linear weighting (0.67/0.33) is superior to a more robust spatial interpolation. If the wind direction shifts even slightly, the "footprint" of what those lasers "see" changes drastically. Did you perform a sensitivity analysis on those weights?

We agree that wind direction affects the area represented by each laser path. This is why we used the IDM footprints to interpret the two OPL measurements. However, a full spatial interpolation of methane concentrations across the lagoon is not possible with these data.

The method gives a single aggregated estimate for the area sampled by each laser path, that is, the laser footprint. We do not have spatially resolved concentrations within the footprints. At each time step, we therefore have two aggregated estimates for two different parts of the lagoon, not many point measurements that could be used to build a concentration map.

Applying a spatial interpolation method in this setting would require assumptions about the spatial autocorrelation structure, which cannot be estimated from only two aggregated observations per time step, and about emissions in unsampled portions of the lagoon. This would mostly be extrapolation rather than interpolation. For this reason, we used a simple weighting approach to scale the two estimates to the full lagoon. This method is transparent and does not imply more spatial detail than the data can support.

We also tested how sensitive the results were to the chosen weights. The main estimate used the 0.67/0.33 weighting, but we also recalculated emissions using 0.75/0.25 and 0.5/0.5. We report the results as a range, with 0.67/0.33 used as the most representative case based on the relative areas represented by the west and east laser footprints.

Please see the changes in the revised manuscript, Line 204-239.

"The influent in the eastern part of the pond is directly connected to the anaerobic pot, where dissolved methane is high. By contrast, the western part of the pond has lower dissolved methane. We expected this to create a gradient of emissions from the eastern to western part of the pond. Because of the size of the pond, a single laser was not sufficient to capture the full pond, and instead two lasers were used to measure emissions along the gradient. Formally, this is a type of stratified sampling that scales each laser's measurements (eastern pond area and western pond area) to the full-lagoon emissions. The classic way to do post-stratification when using stratified sampling is to weight the estimate by the area that each strata represents (i.e., take a weighted average). In the summer campaign, the pond was divided into two strata of similar sizes (~50 %-50 %). The total average emission was thus (0.5 times flux1 + 0.5 times flux2) × total pond area. The weights were varied by ± 5 % to test the sensitivity of the results to this assumption and to provide a range of emissions for the measurements. The total emission for the summer campaign thus ranged from a low estimate of (0.55 times flux1 + 0.45 times flux2) × total pond area to a high estimate of (0.45 times flux1 + 0.55 times flux2) × total pond area. In the winter campaign, the east laser was closer to the Pot end covered area. The pond was divided into two strata of varying sizes (~70 % west-30 % east). The total average emission was thus (0.70 times flux1 + 0.30 times flux2) × total pond area. After propagating uncertainty in the weights (± 5 %), the total emission for the winter campaign ranged from a low estimate of (0.75 times flux 1 + 0.25 times flux 2) × total pond area to a high estimate of (0.65 times flux 1 + 0.33 times flux 2) × total pond area.

2.7 CH_4 flux uncertainty

There are four sources of uncertainty in our flux estimate: instrument precision (± 2 %), inversion dispersion model (± 20 %) (Laubach and Kelliher, 2005), sampling uncertainty (i.e., the s.e. calculated using the number of observations and standard deviation among the 15-min measurements, which contributes around 1.2 to 4.0 %), and the pond area represented by each laser in our weighted average (2.0 % error in total flux for the summer campaign, and 5.2 % for the winter campaign, coming from varying the weight by ± 5 % around their centred value, see section 2.6, resulting in a range of total average emission). These four sources of uncertainty are added in quadrature (sqrt (e_1^2 + e_2^2 + e_3^2 + e_4^2), where the e_i's are the different error terms) to propagate the uncertainty and get our final estimates. Note that the total uncertainty is dominated by the inverse-dispersion model term as the terms are added in quadrature.

The 1-sigma figure was calculated using the definition for the coefficient of variation to represent the relative uncertainty (CV = sigma / mean). Solving for sigma gives: 1-sigma = CV × mean = 79.6 $\mu g\ m^{-2}\ s^{-1}$, for example, the mean flux from this study is the averaged Pond 1 flux, 386.5 $\mu g\ m^{-2}\ s^{-1}$".

4. Was the background laser (OPLC34) moved to account for different northerly wind angles (e.g., NNE vs. NNW)? If not, the $(C/Q)_{sim}$ could be biased by "dirty" upwind air that wasn't properly subtracted.

The background laser position was not moved during the summer and winter measurements. To the north of the pond (NNE-NNW) there was short grass adjacent to the pond, and a corn field a few hundred meters away; because the land cover was similar, the background condition remained stable during our measurements, and methane emissions from the corn field can be considered negligible compared with the large source of aerated pond.

Furthermore, downwind CH_4 concentrations on the western and eastern side of Pond 1 varied with time, while the background concentration on northern side of the pond remained constant at ~1.9 ppm under northerly winds. During the summer campaign, enhancement in CH_4 concentrations under northerly wind conditions were 3.6 ppm (western area) and 10.8 ppm (eastern area). During the winter campaign the averaged enhancements measured by laser 33 (western laser) and laser 1013 (eastern laser) was 2.3 ppm and 13.5 ppm, respectively. These enhanced downwind concentrations measured by both lasers were much higher than the change in background concentration.

Importantly, we applied filtering criteria to ensure that the periods with unreasonable (and high) values from the upwind, which do not represent the true background of the pond, are not included in the flux calculations: the difference in background concentration between the simulated background level and the measured one (upwind) is < 20 ppb, and wind directions >330° or < 40°.

We added a sentence in the revised manuscript, please see Line 136-138.

"Noted the upwind laser remained at the same location during the two measurement periods as the land cover was similar and the background condition remained stable during the measurements (Fig. 1)."

5. Did the cross-calibration (Section 2.2) account for the difference in lower-detection limits between the two brands? If the "East" laser is less precise, the uncertainty in the "high emission" zone is actually higher than in the "low emission" West zone.

Yes, this is an important point. We deliberately matched each instrument to the expected concentration range at its location. The higher-precision Unisearch laser (< 1 ppb at 100 m) was used in the western low-emission zone and the Boreal laser (~20 ppb at 100 m) in the eastern high-emission zone.

In practice, concentrations at both sites were well above detection limits throughout the campaign.

At the western site, the Unisearch laser measured enhancements of 0.12–10.8 ppm in winter and 0.96–14 ppm in summer, at least >100 time above its detection limit of <1 ppb. At the eastern site, enhanced concentrations ranged at 2.1–34 ppm in winter and 1.2–57 ppm in summer, or at least >50 times the Boreal laser's detection limit of 20 ppb.

Ultimately, instrument precision contributed negligibly to the total uncertainty in our flux estimates. Because flux is derived from a large number of measurements, random instrument noise averages out (SE = SD/sqrt(n), where n is the number of measurements), contributing only 2% of total flux uncertainty, which, as reported in the manuscript, is dominated by uncertainty in the wind dispersion model (~20%).

We revised the manuscript and changes can be found on Line129-135.

"In each campaign, two lasers were set up at the south of the pond to measure the downwind concentrations and one laser at the north of the pond to measure the upwind concentrations so that the enhance concentrations from the Pond 1 can be determined. Each instrument was deliberately matched to the expected CH_4 concentration range at its location. The higher-precision Unisearch laser was used in the western low-emission zone and the Boreal laser in the eastern high-emission zone: OPL_C33 was located the western side of the pond (west laser) and OPL_C1013 was located the eastern side (east laser), close to the anaerobic Pot cover area."

6. You calculated an annual emission based on a 5-week summer campaign and a 7-week winter campaign. But wastewater chemistry (BOD/COD) and microbial activity aren't just seasonal; they are operational. Was the "flow rate" or "aerator schedule" during these 12 weeks truly representative of the other 40 weeks of the year? If the facility had a "high load"

period during their measurement window, the "2–3 times higher than NGERS" claim might be an overestimation of the annual total.

The flow rate of wastewater and aerator schedule had significant variations on a diurnal timeframe as discussed in the paper.

While there are some minor seasonal changes between summer and winter in the flow of COD through Pond1, these variations are not very significant to the overall results of the paper. There were not changes to normal operations or unusually high loads during the 5 week and 7 week campaigns and results indicated that a longer campaign would be of little benefit. Furthermore, due to budget constraints a longer campaign period was not possible.

We carefully checked our calculation and corrected value to be 2.0-2.3 times higher than NGERS. Please see the changes in the revised manuscript, Line 342-344, and Line 366.

"The average annual flux as a proportion of CH_4 production ranged from 22.0 to 25.3 %, reflecting that the measured emissions are approximately 25 % of the CH_4 captured by the anaerobic Pot. Our measurements are ~2.0–2.3 times higher than the NGERS reported emissions from Pond 1 (adopting the emissions factor of Unmanaged Aerobic Lagoon, Method 2, NGERS), which averaged 36,843 tCO_2-e yr^{-1} for financial years 2023 and 2024 (Table 2)."

"Our measurements are ~2 times higher than the NGERS estimate."

Referee #1

This manuscript presents direct methane flux measurements from an aerobic wastewater treatment lagoon and compares the observed fluxes with inventory-based estimates. The topic is important because wastewater lagoons remain poorly constrained sources of methane, and field measurements are still limited in the literature. The dataset has potential value, and the paper addresses a relevant gap for greenhouse gas accounting. However, in its current form, I think the manuscript requires major revision before publication. The main conclusions are potentially important, but several aspects of the analysis and interpretation need stronger justification. In particular, the representativeness of the measurements, uncertainty treatment, site characterization, and extrapolation to broader inventory implications require more detailed discussion.

We thank the Reviewer for the thorough and constructive review of our manuscript. Below, we address each of the reviewer's comments. We also outline additional corrections identified during the revision process, in addition to those required to address the comments from each referee. All the changes can be found in the revised manuscript.

My Specific comments are as follows:

1. What motivated the authors to study an aerobic pond instead of an anaerobic pond, given that the latter could be a larger concern for methane production? Is it possible to provide any comparison between these two types of ponds?

Pond 1 is largely an anaerobic pond and to some degree there are aerobic processes as well due to the *surface* aerators. However, the surface aerators only cover a small percentage of the entire pond area and do not deliver a satisfactory dissolved oxygen level that enables the categorising the Pond 1 as an aerobic pond. Other research using hoods and a mobile CH_4 sensor have shown that the Pond 1s have by far the highest CH_4 emissions at the WTP and therefore are the largest concern with respect to CH_4 emissions. This is because they are the first ponds after the covered anaerobic pots where anaerobic digestion of raw sewage takes place. Effluent from the anaerobic pots is saturated with dissolved methane, once agitated and/or under different atmospheric pressure (in Pond 1), releases CH_4 emissions to the environment. Wastewater in the downstream facultative ponds such as Pond 2 has a much lower COD than in Pond 1, and Pond 2 typically emits less than 1/10 of the CH_4 emissions compared to Pond 1.

We made changes in the revised manuscript, Line 99-106

"Other research using hoods and mobile survey techniques have shown that the Pond 1 have the highest CH_4 emissions at the WTP and therefore are the largest concern with respect to the CH_4 emissions. This is because it is the first pond after the covered anaerobic pot where anaerobic digestion of raw sewage takes place. Biogas is generated underneath the cover and is collected and sent to onsite power generation facilities for energy generation. Effluent from the anaerobic pot is saturated with dissolved methane, once agitated and/or under different atmospheric pressure (in Pond 1), releases CH_4 emissions to the environment. Wastewater in the downstream facultative ponds such as Pond 2 has a much lower COD than in Pond 1, and Pond 2 typically emits less than 1/10 of the CH_4 emissions compared to Pond 1."

2. The atmospheric transport model is 3-dimensional. However, the authors' measurements only happened at ground level. How could this fact be addressed in uncertainty analysis? Could drone-based vertical measurement of methane concentrations improve the credibility of the emission rate estimation?

The IDM is an inverse-dispersion, surface-layer model. In our study the pond-to-sensor distance was within 300 m, which is appropriate for a surface-layer model: emission plume emitted from the pond can be sampled by the ground-level sensor before it rises above the surface layer and escapes detection from the sensor.

In the previous section we described flux footprint. For each 15-minute period we used WindTrax to generate 150,000 Monte Carlo replicates to simulate the origin points of discrete particles (i.e., touchdowns) that might have contributed to the flux (in practice these particles are generated from the measurement to the source, hence WindTrax is said to be an inverse dispersion model). To determine the source area for each 15-minute measurement, we first filter the touchdowns that fall outside the pond (as they correspond to background concentrations). We aggregate the touchdowns into 10 by 10 m raster cells. We count the number of touchdowns per cell. As is customary in footprint analysis, we then estimate the footprint area as the area contributing to 80% of the touchdowns that fall within the pond (Kljun et al., 2015).

The anaerobic pond was treated as an area source, and the inverse-dispersion model was used to measure gas emissions from the pond. IDM has been well reported in the literature, it requires the horizontal concentration difference between the upwind and downwind of the pond, coupled with three-dimensional wind variables to characterize the atmospheric turbulence. Each measurement technique has advantages and limitations, and emission estimates could be substantial differences using different measurement approaches, due to their spatial footprint and capability of long-term measurement. Using drone technique is a top-down measurement approach. It may be well suitable for targeting emission hotspots within a complex source configuration, but footprint is usually limited by the flying height due to nationwide regulations of flying zone. It is usually used for mobile survey and data quality assessment but spatial interpretability of concentration changes relies on wind regime and atmospheric stability (Dash et al., 2026). Therefore, interfering emission sources, shorter period on site, and requirement of consistent weather conditions limit the use of this method (Gong et al., 2023; De Jong et al., 2026). Recently Flesch et al. (2023) used IDM technique coupled with concentrations measured by uncrewed aerial vehicle (UAV) to quantify CH_4 flux from waste pond.

We modified the uncertainty calculations. The changes can be found in the manuscript Line 224-239.

“

2.7 CH_4 flux uncertainty

There are four sources of uncertainty in our flux estimate: instrument precision (± 2 %), inversion dispersion model (± 20 %) (Laubach and Kelliher, 2005), sampling uncertainty (i.e., the s.e. calculated using the number of observations and standard deviation among the 15-min measurements, which contributes around 1.2 to 4.0 %), and the pond area represented by each laser in our weighted average (2.0 % error in total flux for the summer campaign, and 5.2 % for the winter campaign, coming from varying the weight by ± 5 % around their centred value, see section 2.6, resulting in a range of total average emission). These four sources of uncertainty are added in quadrature (sqrt (e_1^2 + e_2^2 + e_3^2 + e_4^2), where the e_i's are the different error terms) to propagate the uncertainty and get our final estimates. Note that the total uncertainty is dominated by the inverse-dispersion model term as the terms are added in quadrature.

The 1-sigma figure was calculated using the definition for the coefficient of variation to represent the relative uncertainty (CV = sigma / mean). Solving for sigma gives: 1-sigma = CV × mean = 79.6 μg m^{-2} s^{-1}, for example, the mean flux from this study is the averaged Pond 1 flux, 386.5 μg m^{-2} s^{-1}.”

3. The measurement only happened in Summer and Winter, which was not representative enough for the annual case. Also, the manuscript would benefit from a clearer justification of whether the monitoring period is representative of longer-term methane emissions.

I agree with the reviewer’s comment. A longer-term measurement would be ideal; however, the funding was short and resources are limited, importantly the snapshot measurement can indicate seasonal emissions variability. Our measurements in each season obtained over 25% valid data over the measurement period, which is sufficient to represent the emissions status.

4. In L151-L154, could you elaborate more on how the WindTrax model works, instead of treating it like a black box? What assumptions does the model use? What uncertainties does it bring? And does it work equally well on point source and areal source?

Sure, WindTrax is a software (www.thunderbeachscientific.com) based on the backward Lagrangian stochastic dispersion (bLs) model for calculating gas emission rate from a source area (Flesch et al., 1995). It is described in a few scientific publications (e.g., (Usepa); Wilson et al. (2012); Bai et al. (2025)), and has a substantial record of use with IDM (Bai et al., 2023). It includes a graphical interface that creates a map showing the target source and sensor locations, including sensor heights. By default, WindTrax releases 50,000 trajectories from 30 points- particles distributed along the measurement path. Trajectories travel in the dispersion plume upwind, starting at the senor position backwards towards the source surface; some trajectories intersect with the source surface, while others avoid the source surface. When trajectories touch the ground, surface information including the particle's touchdown position and vertical velocity (x,y,w) is recorded.

WindTrax uses the bLs method, which is based on Monin-Obukhov Similarity Theory (MOST), and simulates the relationship between concentration and emission rate, $(C/Q)_{sim}$, based on the vertical velocity of the trajectories (w_0) that infers the pond surface ("touchdown").

$(C/Q)_{sim} = 1/N\ \Sigma|2/w_0|$

Where $(C/Q)_{sim}$ is the simulated relationship ratio, N is the total number of particles released, w_0 is the vertical 'touchdown' velocity, which is the function of u^* (friction velocity, m s^{-1}), L (Obukhov stability length, m), z_0 (surface roughness length, m) and β (average wind direction). These parameters can be derived from a 3-D sonic anemometer. The summation includes all 'touchdown' trajectories within the source area from each release point.

Assuming a horizontally homogenous surface layer (Flesch et al., 2004), and given the downwind concentration increase above the background level ($C_{downwind}$-C_{upwind}), the unknown source emission rate (e.g., the effluent pond) Q_{IDM} is determined using:

$Q_{IDM} = \Delta C/\ (C/Q)_{sim}$

Where Q_{IDM} is the pond emission rate, ΔC is the concentration enhancement between upwind (background) C_{upwind} and downwind $C_{downwind}$, $\Delta C = C_{downwind}$-C_{upwind}, $(C/Q)_{sim}$ is the simulated concentration and emission rate relationship ratio, it is associated with source geometry, wind variables, surface layer meteorology, and the sensor coordinates.

Wind variables, including ambient temperature and pressure, and the standard deviations of the velocity fluctuations in the three directional components ($\sigma_{u,v,w}$), can be derived from a 3-D sonic anemometer.

As the IDM is based on the Monin–Obukhov similarity theory (MOST), which assumes a horizontally homogenous terrain with no tree lines, tall buildings, or other nearby sources, it also requires a well-defined, spatially uniform source area. As a result, the accuracy of the IDM can be sensitive to the conditions that are not ideal (for example, wind complexity or a non-uniform source area) (Gao et al., 2010). Bai (2010) discussed several factors that contribute to the Windtrax uncertainty, such as the number of trajectories releases, the number of particles distributed along the measurement path, the averaging interval for flux estimates, the distance of the sensors to the source, and boundary layer stability. Flesch et al. (2007) noted that in many situations MOST may not be well adapted, accurate IDM calculation can be obtained with assumptions of winds under ideal conditions. Harper et al. (2010) and Laubach and Kelliher (2005) reported uncertainties of approximately 10% and 20% in IDM flux estimates, respectively. Wilson et al. (2001) tested IDM accuracy by introducing disturbed winds around a lagoon to create wind complexity and found good agreement between IDM and a mass-balance method, and the IDM measurements were within 15% of the true emission rate. Bühler et al. (2021) reported an uncertainty range of 14-21% when using IDM coupled with line-average concentrations in a dairy cow study but rising to as much as 36% when external sources were included. However, in this study we applied filtering criteria so that external sources are not included in the flux calculation. An average of 20% uncertainty from IDM simulation is applied in our uncertainty estimation.

The IDM technique has been shown to offer the efficiency (over survey approach), simplicity (single gas concentration sensor and basic wind information) and flexibility (any location at any arbitrary shape). IDM techniques are suitable for well-defined surface source; however, for some point sources, such as flares, leaky pipes, biodigesters, or a barn, small animal pen, it can be challenging because they create wind complexity. Flesch et al. (2011) approved to overcome these complications by adjusting the distance between sensor and point source until concentration is measured far enough so that the IMD is insensitive to the complications. As a rule of thumb for a single source, the distance from the point source should be ten times the height of the largest wind obstacle, and roughly two times the maximum distance between potential sources. The IDM method has also been widely used in intensive feedlots and grazing dairy cows where individual cattle is treated as a point source, but a group of cattle can be treated as a source area.

We added more information about WindTrax in the manuscript, Line 161-166.

"WindTrax is a software (www.thunderbeachscientific.com) based on the backward Lagrangian stochastic dispersion model (bLs) for calculating gas emission rate from a source area (Flesch et al., 1995). WindTrax uses the bLs method, which is based on Monin-Obukhov Similarity Theory, and simulates the relationship between concentration ($C-C_b$) and emission rate Q, $(C/Q)_{sim}$. Because the Q v.s. $(C-C_b)$ relationship depends on wind conditions, one must also make wind measurements."

5. In L151-L154, did the authors treat the entire pond as a homogeneous areal source? If so, what uncertainty can that bring?

No. We did not treat the entire pond as one homogeneous source. The two OPL paths sampled two different footprint areas, one associated with the west laser and one with the east laser. At each time step, each footprint gives one aggregated estimate for the area it samples. This does not mean that emissions within the footprint are homogeneous, only that the method measures an average response over that footprint.

To scale these estimates to the full lagoon, we used a weighted average of the west and east estimates, where the weights represent the relative lagoon areas associated with each laser footprint. The uncertainty from this scaling was assessed using a sensitivity analysis with different weights. We report the resulting range in the manuscript.

We made changes in the revised manicurist, please see Line 204-239.

"

2.6 Average Pond 1 CH_4 flux calculation

The influent in the eastern part of the pond is directly connected to the anaerobic pot, where dissolved methane is high. By contrast, the western part of the pond has lower dissolved methane. We expected this to create a gradient of emissions from the eastern to western part of the pond. Because of the size of the pond, a single laser was not sufficient to capture the full pond, and instead two lasers were used to measure emissions along the gradient. Formally, this is a type of stratified sampling that scales each laser's measurements (eastern pond area and western pond area) to the full-lagoon emissions. The classic way to do post-stratification when using stratified sampling is to weight the estimate by the area that each strata represents (i.e., take a weighted average). In the summer campaign, the pond was divided into two strata of similar sizes (~50 %-50 %). The total average emission was thus (0.5 times flux1 + 0.5 times flux2) × total pond area. The weights were varied by ± 5 % to test the sensitivity of the results to this assumption and to provide a range of emissions for the measurements. The total emission for the summer campaign thus ranged from a low estimate of (0.55 times flux1 + 0.45 times flux2) × total pond area to a high estimate of (0.45 times flux1 + 0.55 times flux2) × total pond area. In the winter campaign, the east laser was closer to the Pot end covered area. The pond was divided into two strata of varying sizes (~70 % west-30 % east). The total average emission was thus (0.70 times flux1 + 0.30 times flux2) × total pond area. After propagating uncertainty in the weights (± 5 %), the total emission for the winter campaign ranged from a low estimate of (0.75 times flux 1 + 0.25 times flux 2) × total pond area to a high estimate of (0.65 times flux 1 + 0.33 times flux 2) × total pond area.

2.7 CH_4 flux uncertainty

There are four sources of uncertainty in our flux estimate: instrument precision (± 2 %), inversion dispersion model (± 20 %) (Laubach and Kelliher, 2005), sampling uncertainty (i.e., the s.e. calculated using the number of observations and standard deviation among the 15-min measurements, which contributes around 1.2 to 4.0 %), and the pond area represented by each laser in our weighted average (2.0 % error in total flux for the summer campaign, and 5.2 % for the winter campaign, coming from varying the weight by ± 5 % around their centred value, see section 2.6, resulting in a range of total average emission). These four sources of uncertainty are added in quadrature (sqrt (e_1^2 + e_2^2 + e_3^2 + e_4^2), where the e_i's are the different error terms) to propagate the uncertainty and get our final estimates. Note that the total uncertainty is dominated by the inverse-dispersion model term as the terms are added in quadrature.

The 1-sigma figure was calculated using the definition for the coefficient of variation to represent the relative uncertainty (CV = sigma / mean). Solving for sigma gives: 1-sigma = CV × mean = 79.6 μg m^{-2} s^{-1}, for example, the mean flux from this study is the averaged Pond 1 flux, 386.5 μg m^{-2} s^{-1}."

6. In L182, what does "stratified sampling" mean?

Stratified sampling means that the lagoon was divided into different sampling groups, or strata, before calculating the overall estimate. In our case, the strata correspond to the west and east footprint areas sampled by the two laser paths. Each stratum was represented by its own OPL-IDM estimate, and the full-lagoon estimate was then calculated by weighting these estimates according to the relative area represented by each footprint. This is a common approach in survey sampling when different parts of a study area are sampled separately and then combined using weights.

We revised the manuscript and made changes, please see Line 204-233.

"The influent in the eastern part of the pond is directly connected to the anaerobic pot, where dissolved methane is high. By contrast, the western part of the pond has lower dissolved methane. We expected this to create a gradient of emissions from the eastern to western part of the pond. Because of the size of the pond, a single laser was not sufficient to capture the full pond, and instead two lasers were used to measure emissions along the gradient. Formally, this is a type of stratified sampling that scales each laser's measurements (eastern pond area and western pond area) to the full-lagoon emissions. The classic way to do post-stratification when using stratified sampling is to weight the estimate by the area that each strata represents (i.e., take a weighted average). In the summer campaign, the pond was divided into two strata of similar sizes (~50 %-50 %). The total average emission was thus (0.5 times flux1 + 0.5 times flux2) × total pond area. The weights were varied by ± 5 % to test the sensitivity of the results to this assumption and to provide a range of emissions for the measurements. The total emission for the summer campaign thus ranged from a low estimate of (0.55 times flux1 + 0.45 times flux2) ×

total pond area to a high estimate of (0.45 times flux1 + 0.55 times flux2) × total pond area. In the winter campaign, the east laser was closer to the Pot end covered area. The pond was divided into two strata of varying sizes (~70 % west-30 % east). The total average emission was thus (0.70 times flux1 + 0.30 times flux2) × total pond area. After propagating uncertainty in the weights (± 5 %), the total emission for the winter campaign ranged from a low estimate of (0.75 times flux 1 + 0.25 times flux 2) × total pond area to a high estimate of (0.65 times flux 1 + 0.33 times flux 2) × total pond area."

References

Bühler, M., Häni, C., Ammann, C., Brönnimann, S., and Kupper, T.: Using the inverse dispersion method to determine methane emissions from biogas plants and wastewater treatment plants with complex source configurations, Atmospheric Environment: X, 13, 100161, https://doi.org/10.1016/j.aeaoa.2022.100161, 2022.

Guisasola, A., de Haas, D., Keller, J., and Yuan, Z.: Methane formation in sewer systems, Water Research, 42, 1421-1430, https://doi.org/10.1016/j.watres.2007.10.014, 2008.

Mannina, G., Butler, D., Benedetti, L., Deletic, A., Fowdar, H., Fu, G., Kleidorfer, M., McCarthy, D., Steen Mikkelsen, P., Rauch, W., Sweetapple, C., Vezzaro, L., Yuan, Z., and Willems, P.: Greenhouse gas emissions from integrated urban drainage systems: Where do we stand?, Journal of Hydrology, 559, 307-314, https://doi.org/10.1016/j.jhydrol.2018.02.058, 2018.

Bai, M.: Methane emissions from livestock measured by novel spectroscopic techniques, PhD School of Chemistry, University of Wollongong, University of Wollongong, NSW, Australia, 303 pp., 2010.

Bai, M., Wang, Z., Lloyd, J., Seneviratne, D., Flesch, T., Yuan, Z., and Chen, D.: Long-term onsite monitoring of a sewage sludge drying pan finds methane emissions consistent with IPCC default emission factor, Water research X, 19, 100184, https://doi.org/10.1016/j.wroa.2023.100184, 2023.

Bai, M., Wang, Z., Seneviratne, D., Lloyd, J., De Jong, P., Ye, L., and Chen, D.: Substantial ammonia emissions from sludge drying pans in wastewater treatment plants, Nature Water, 10.1038/s44221-025-00479-8, 2025.

Bühler, M., Häni, C., Ammann, C., Mohn, J., Neftel, A., Schrade, S., Zähner, M., Zeyer, K., Brönnimann, S., and Kupper, T.: Assessment of the inverse dispersion method for the determination of methane emissions from a dairy housing, Agricultural and Forest Meteorology, 307, 108501, https://doi.org/10.1016/j.agrformet.2021.108501, 2021.

Dash, S. S., Coates, T. W., and Madramootoo, C. A.: UAV-Based Measurements of Methane Enhancements Reveal Hotspot Structure and Wind Effects, Environmental Science & Technology, 60, 13980-13996, 10.1021/acs.est.5c18173, 2026.