General comments

This is a nicely presented instrumentation paper describing an autonomous surface platform that combines chamber flux measurements with a broad ancillary and bathymetric payload. The study is the most valuable contribution, it gives a concrete, useful recommendation to the community.

However, my main concern is that, for AMT, the paper currently reads more as a capability demonstration than as a measurement-technique characterization. Here are my suggestions.

Major comments

1. Eq. (3) appears to compute blade-tip circumferential speed rather than axial jet velocity; please justify that this represents the water velocity, or recompute from thrust via momentum theory.
2. Wind bursts exceeding 30 m s^{-1}. This value (~108 km h^{-1}) is inconsistent with the campaign conditions described and, as far as I can tell, with the wind-speed distributions in Figs. 6 and C1. Please verify against the raw anemometer record. If these are anemometer spikes or a unit issue, plausible given the 1 m sensor height amid the platform superstructure, I'd recommend removing the claim rather than presenting it as evidence of robustness.
3. For a methods paper, each flux from Eq. (5) should carry a propagated uncertainty (regression slope error, sensor noise, T, p, and chamber geometry). Please also state the minimum detectable flux for CO_2 (both sensors) and CH_4 at your 5-min duration and chamber V/A. You discuss the EC analyzer's detection limit (Sect. 3.2) but never your own chamber's, which is precisely the number needed to support the claim that the lower-cost GMP343 is "sufficient."
4. The chamber is 25 cm tall with ~6 cm submerged, leaving a headspace of ~19 cm. Please state the numerical volume and area used, and confirm the effective headspace volume (not the full cylinder) was used, this is a ~30% effect on every reported flux.
5. CH_4 fluxes are discarded for $R^2 < 0.8$ (n = 22), but "all other CO_2 fluxes were kept." Please report the R^2 distribution for CO_2 and either apply a consistent linearity criterion or justify why one is unnecessary for CO_2. Given an opaque chamber and a 5-min enclosure, please also confirm whether linear vs. non-linear fits were tested for the higher CO_2 fluxes.
6. The consequences of using an opaque chamber are not discussed. This is most relevant to the CO_2 *uptake* reported at Inre Harrsjön (Fig. 7C), measured in daylight with a dark chamber. Please justify the opacity choice and discuss how it affects the representativeness of the mapped CO_2 fluxes.
7. Ebullition was detected once and excluded, yet at Inre Harrsjön ebullition is the dominant CH_4 pathway (your cited Wik et al., 2013: up to 567 nmol m^{-2} s^{-1}). A 5-min chamber rarely captures bubble events, so the CH_4 maps (Figs. 7-8) represent the diffusive component only. This limitation should be stated plainly in the abstract and conclusions, not only in passing.

8. Within the matched footprint the EC median is -0.27 (uptake) and the chamber median +0.29 (emission), opposite signs, while you conclude "no substantial systematic difference" from a non-significant Mann-Whitney ($p = 0.52$, $n = 17$ vs. 38). With EC scatter (SD 1.31) an order of magnitude larger than the chamber (0.11), the test is underpowered; non-significance does not demonstrate agreement. Please report the sign disagreement in the medians, temper "CO_2 fluxes matched well," and comment on why the chamber distribution is so much narrower than EC.
9. Please clarify whether EC half-hours and chamber points were compared as pooled distributions over May 13-16 using a maximum-extent footprint, or matched to concurrent half-hourly footprints. Since fluxes have diel structure and Dagow was sampled day and night, distributional comparison risks mismatched times of day. If feasible, restrict to concurrent half-hours.
10. The two sampling blocks a week apart, with bimodal water temperature, mean the mapped spatial patterns may partly reflect between-block differences rather than true spatial structure, most relevant to the headline CH_4 Moran's I = 0.33. Please test this directly, e.g. by colouring Fig. 7C/D by measurement day or reporting Moran's I within each block.
11. The eastern shore of Inre Harrsjön was deliberately unsampled yet Fig. 8C/D show filled fields there, and $\sigma = 30$ m at Dagow fills sizeable gaps. Please mask low-data-density cells or add a sampling-density/uncertainty layer so interpolated fill is not read as measured flux.

Minor comments

12. Abstract: "CH_4 fluxes diverged slightly" contradicts the >3× difference with large effect size. Please align.
13. The 5-min vs. 1.5-min comparison is confounded, the 1.5-min data are from one mission that also maximized travel distance and motor use. A cleaner test with existing data: truncate the 5-min records to their first 1.5 min and recompute, isolating the duration effect.
14. Bland-Altman: the near-zero mean difference at 1.5 min reflects cancellation of a magnitude-dependent bias; please emphasize the limits of agreement/RMSE instead.
15. Please state the calibration/zero-span protocol and any drift checks for both analyzers, and confirm both fluxes are on a consistent dry-mole-fraction basis.
16. Please state whether an initial mixing deadband is excluded from the dC/dt fit, and discuss chamber-induced disturbance of the near-surface layer (the ~6 cm submersion) as a possible contributor to the high chamber CH_4.
17. Given CEP up to 1.0 m and propeller station-keeping during loitering, please quantify how often propellers fired during active measurements and whether this correlates with flux R^2 or magnitude.
18. Bathymetry RMSE, clarify whether this is a fit residual or a cross-validation error, and justify the virtual-0.1 m near-shore depth assumption.
19. Justify the different k (5 vs. 8) in the Moran's I analysis and show the conclusions are not sensitive to it.

20. Eq. (5): write R as 8.314 J mol^{-1} K^{-1} (Pa m^3 mol^{-1} K^{-1}).
21. Kinematic viscosity is taken at 20 °C while water was 5-16 °C; use a representative value or note the approximation.
22. State the flux sign convention (positive = emission) in Sect. 2.6.
23. Label the "~170 per day / >1,000 per week" figures clearly as a theoretical ceiling and reconcile with the 6-8 h battery/swap constraint.
24. Make the novelty relative to Zug et al. (2025) and Pose et al. (2023) explicit early in the Introduction.