



A mid-cost automated floating chamber system for high-frequency CO₂ and CH₄ flux measurements: design, calibration, and comparison with infrequent measurement strategies

Ralf C.H. Aben¹, Judith van der Knaap¹, Floris Hahn², Gerben Wulterkens², Dániel Szalas-Motesiczky²,
5 Emiel Vissers², Arno Engels², Arjan de Kleine², Christian Fritz¹ & Sarian Kosten¹

¹Department of Ecology, Radboud Institute for Biological and Environmental Sciences, Radboud University, Nijmegen, 6525AJ, the Netherlands

²FNWI TechnoCentre, Radboud University, Nijmegen, 6525ED, the Netherlands

Correspondence to: Ralf C.H. Aben (ralf.aben@ru.nl)

10 **Abstract.** Small inland waters are important greenhouse gas (GHG) sources; however, most field campaigns rely on daytime-only measurements collected at multi-week intervals, potentially biasing flux estimates. We present the Buoyant Emission Recorder for Temporal High-Frequency Analyses (BERTHA), a mid-cost automated floating chamber system for high-frequency measurements of CO₂ and CH₄ fluxes on multiple spatial scales. We describe the system's design and test how empirically derived pressure and water vapour correction factors for the mid-cost CO₂ and CH₄ sensors increase the
15 accuracy of reported concentrations and derived flux estimates. Here, we found a 17–37 % (CO₂) and 4–9 % (CH₄) higher mean flux estimate with pressure compensation during the 97-day measuring period. Compensation for water vapour resulted in much smaller differences (~1–2 % increase). Additionally, we calculated the mean CO₂ and CH₄ emission using high-frequency measurements during 97 days and simulated alternative measurement scenarios based on the full dataset and evaluated how well these 'reference means' could be approached by 1) single, daytime-only chamber flux measurements at
20 daily to monthly intervals, and 2) short-term (2–3 days) automated chamber deployments every two–three weeks combined with machine learning-based gap-filling. Results of the simulation analysis showed that daytime-only measurements consistently and strongly underestimated the reference CO₂ mean. CH₄ emissions matched the reference mean well for daily, but not for monthly measurements (>90 % and ~23 % of simulations deviated <20 % of the reference mean, respectively). Short-term automated chamber campaigns, gap-filled using machine learning closely reproduced the CO₂ reference mean,
25 with 89 % and 48 % of means being within ±20 % of the reference mean, for campaign intervals of two and three weeks, respectively. We conclude that high-frequency chamber measurements using mid-cost GHG analysers combined with empirical pressure corrections provide accurate emission estimates, a robust benchmark for evaluating reduced



measurements strategies and gap-filling approaches, and potential for improved process-based understanding of GHG fluxes when combined with ancillary high-frequency measurements.

30 1 Introduction

Despite their limited surface area, drainage ditches, ponds, and channels are disproportionately strong sources of carbon dioxide (CO₂), methane (CH₄) (Van Der Knaap et al., 2025; Peacock et al., 2021a), and nitrous oxide (N₂O) (Silverthorn et al., 2025). This is driven by high nutrient and carbon inputs (Peacock et al., 2021b), shallow depths, and high sediment and edge-to-water volume ratios (Holgerson and Raymond, 2016). Emissions occur through diffusive fluxes across the water-air
 35 interface, ebullition from the sediment, and plant-mediated transport. The relative importance of each pathway is influenced by the morphology, hydrology, and biogeochemical conditions.

Small inland waters emit greenhouse gases (GHGs) at rates that fluctuate significantly over time, making it difficult to accurately quantify them. Emissions vary on both diel (De Eyto et al., 2025; Attermeyer et al., 2021) and seasonal timescales (Reed et al., 2018; Raymond et al., 2013) and are influenced by multiple factors, including agricultural activities such as
 40 fertilisation (Mwanake et al., 2023), weather conditions such as precipitation (Piatka et al., 2024; White et al., 2021) and solar radiation (Golub et al., 2023), hydrology (Peacock et al., 2019), and biological processes related to plants (Bodmer et al., 2024; Shen et al., 2024) and algae (Theus et al., 2023). As a result, campaign-based measurements conducted at biweekly or monthly intervals—typically limited to daytime—introduce large uncertainty in flux estimates and increase the risk of bias associated with the chosen measurement strategy (such as daytime-only measurements). While such measurements provide
 45 valuable data, their low temporal resolution restricts solid process-based understanding, underscoring the need for more (prolonged) high-frequency measurements.

Studies on temporal variability in GHG fluxes are relatively scarce, and even fewer rely on automated high-frequency measurements. Diel cycles often show elevated CO₂ emissions during the night (Liu et al., 2016; Paranaíba et al., 2025; De Eyto et al., 2025), whereas reported findings for CH₄ are more variable across aquatic systems, with both higher nighttime
 50 (Paranaíba et al., 2025) and higher daytime (Bastviken et al., 2010; Sieczko et al., 2020) fluxes reported. While previous work has demonstrated short-term and diel variability, its implications for the estimation of mean flux intensities under reduced measurement frequency remains poorly quantified.

Currently, GHG flux measurements are typically conducted over short intervals (~minutes) and at infrequent time steps (e.g. weekly or monthly) using floating chambers with laboratory-grade gas analysers (see e.g. Deemer et al. (2016) & Wik et al.
 55 (2016)). These instruments provide highly accurate concentration measurements but are costly. In recent years, low-cost sensor technology and (custom-built) automated chambers have enabled high-frequency monitoring at substantially lower cost. This advantage, however, often comes at the expense of measurement quality and reliability. Sensor accuracy and



precision is lower, and long chamber closure times can substantially alter the chamber microclimate (Matthias et al., 1980; Pumpanen et al., 2010). Together, these approaches represent two extremes between measurement accuracy and temporal
60 coverage. A gap remains for mid-cost approaches that combine sufficient sensor precision, accuracy and reliability with sustained, high temporal resolution.

Here, we present BERTHA - Buoyant Emission Recorder for Temporal High-frequency Analyses. BERTHA is a mid-cost automated floating chamber system (~€3k materials, ~€3k CO₂ analyser, ~€5k CH₄ analyser; excluding optional costs for autonomous power supply) designed to autonomously measure CO₂ and CH₄ fluxes in small inland waters at high frequency
65 and increased spatial resolution. The measurement frequency is set by the user to allow high or lower frequency measurements, depending on the research questions and potential power limitations. We describe the system's design, calibration, and (in-situ) validation, evaluate its performance, and demonstrate its capacity to capture temporal patterns that would be missed by conventional snapshot-like measurements. Additionally, using the collected data, we quantitatively assess how different measurement strategies affect the accuracy of mean flux estimates. Specifically, we compare (i) typical
70 short-term daytime measurements and (ii) periodic multi-day automated chamber deployment combined with machine-learning-based gap-filling. This dual-scenario framework provides a quantitative basis for understanding how reduced measurement frequency and incomplete temporal coverage influence emission estimates. By bridging the gap between high-precision infrequent, manual measurements (with lab-grade analysers) and lower-precision, automated measurements (with low-cost sensor systems), our approach offers a practical solution for expanding GHG monitoring networks, linking high
75 frequency measurements to field and catchment scale events (e.g. fertilisation, erosion, rain, drought) and improving process-based understanding of emissions from small inland waters.

2 Methods

2.1 Chamber measurements

The automated transparent chamber system (BERTHA) (Fig. 1) was deployed in a ditch in a drained peatland in Zegveld
80 (52.138°, 4.841°) from 30 May 2024 until 10 September 2024. The area is used as grassland for dairy farming. The system consists of three units of floating chambers with PVC floaters. The chambers are opened and closed using a linear actuator. The central unit contains electronics for chamber control, multiplexing of the gas flow, a mid-cost non-dispersive infrared (NDIR) CO₂ sensor (GMP343, Vaisala; calibrated range 0–2,000 ppm), and a tunable diode laser spectrometer (TDLS) CH₄ sensor (LGD Compact-A, Axetris; calibrated range 0–100 ppm). The cylindrical chambers have an inner diameter of 29.2
85 cm and a height of 9.8 cm, with the lower 1.5 cm below the waterline during chamber closure. Each chamber contained a sensor for air temperature and relative humidity (RH; EZO-HUM™, Atlas Scientific). Before gas reached the analysers, any condensate was trapped using a rigid water trap (AF series 5µm, SMC). During prototype testing in an experimental ditch at



the university's experimental garden, erroneous CH₄ measurements (i.e. noise with a magnitude of ten ppm, indicative of interference) were consistently observed during nights and mornings when the temperature dropped below the dewpoint. This interference disappeared after installing the water trap. Gas pressure, temperature, and RH were measured just before and directly after gas entered and exited the analysers using a BME680 sensor (Bosch Sensortec) built into the plumbing system (Fig. 2 and Fig. S1). Gas recirculation was achieved using polyurethane tubing and a membrane pump (NMP830KPDC-B4 HP 12V, KNF) set at a flow rate of $\sim 0.75 \text{ L min}^{-1}$. To achieve a balance between power consumption and desired measurement frequency, the three chambers were set to measure sequentially every 36 minutes using a closure time of 250 seconds and a 12-second flushing time before and after each chamber closure. Power consumption with these settings was $\sim 18 \text{ W}$. All chamber systems were equipped with a low-flow fan to achieve a well-mixed headspace (Christiansen et al., 2011; Rochette and Hutchinson, 2005). Measurements of above-mentioned sensors were logged on an SD card every 2.5 seconds as well as messaged wirelessly via the MQTT protocol using an embedded modem with SIM card.



Figure 1. Impression of the automated floating chamber system. An extra floater—mounted on top of the centre floater—carries the box containing the multiplexer and its components (Fig. 2) as well as a light sensor (measuring photosynthetically active radiation) mounted on top of the box.

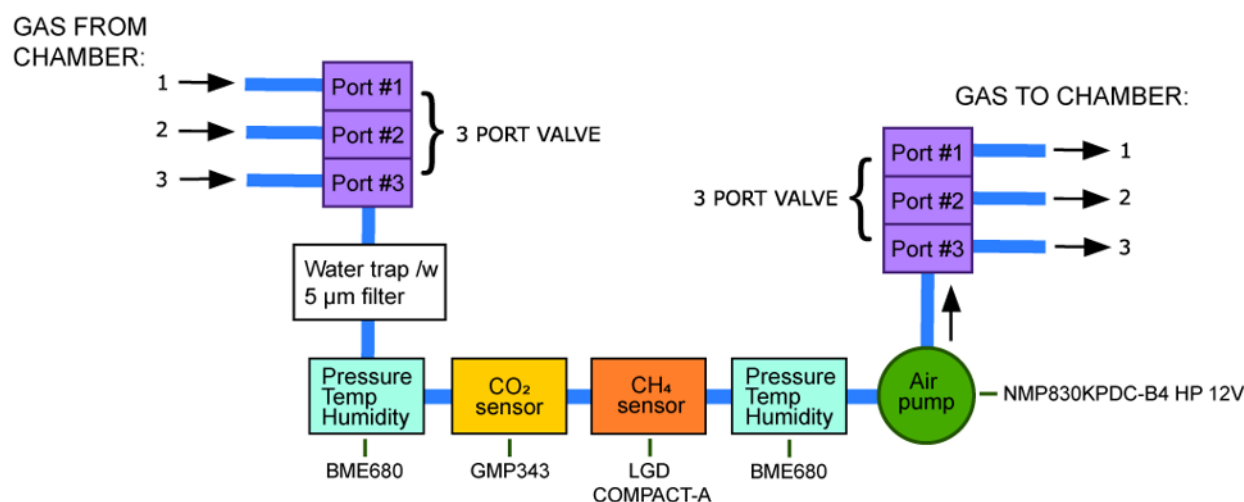


Figure 2. Plumbing diagram of the multiplexer and its components.

2.2 Compensation for effects of pressure

CO₂ and CH₄ mole fractions reported by the sensors are influenced by the gas pressure, as pressure affects both the molecular number density along the optical path and the shape of the absorption spectrum of the measured gas through collisional broadening (Castner, 2005; Gaynullin et al., 2023). Fluctuations in ambient air pressure are relatively small at sea level, but pressure inside the system's multiplexer and its sensors (Fig. 2 and Fig. S1) can vary more substantially, mainly due to recirculation of gas combined with restrictions in the gas flow (e.g. through flow resistance introduced by the plumbing, potentially worsened by partial clogging of particle filters or valves), which can lead to the internal pressure being significantly lower than the ambient pressure. This, in turn, can introduce significant errors in reported concentrations (i.e. mole fractions). To empirically derive pressure correction factors, we followed the procedure reported by Gaynullin et al. (2023). Here, pressure compensation factors were determined in a two-dimensional way, taking into account that pressure compensation factors are not always universal and can depend on the concentration of the measured gas. To this end, we introduced a stream of calibration gas to the plumbing system (Fig. 2) and altered the gas pressure inside the plumbing system to observe its effect on the reported CO₂ and CH₄ concentrations. To lower the gas pressure, we used a Bronkhorst D-6311 Mass Flow Controller on the inlet tubing to introduce a stepwise reduction in flow rate. This, in turn, resulted in a stepwise pressure reduction within the system, since the built-in membrane pump continuously drew air from the inlet tubing. This procedure was repeated three times, each time with a different calibration gas (synthetic air with concentrations of 181, 588, and 980 ppm for CO₂ and, respectively, 1.99, 5.00, and 9.91 ppm for CH₄). The internal pressure was measured using BME680 sensors incorporated into the plumbing system, as described earlier (Fig. 2 and Fig. S1).



To determine the concentration (c_{P0} ; ppm) reported by the sensor at the standard pressure (P_0) of 1 atm, we fitted a cubic polynomial between the measured concentration and pressure values. Pressure values were lowered by 1 atm so that the intercept of the model fit represented c_{P0} . Subsequently, the pressure compensation factor (K) was calculated for each measured concentration (c_{meas}) as:

$$K = \frac{c_{meas}}{c_{P0}}. \quad (1)$$

The data were then fitted via nonlinear regression analysis in R using the `nlsLM` function of the `minpack.lm` package (Elzhov et al., 2016) to calculate K for a given pressure:

$$K = A(P - P_0)^2 + B(P - P_0) + 1, \quad (2)$$

where A and B are fitted parameter values. Following this procedure for each of the three calibration gases concludes the first dimension of the pressure compensation algorithm. The second dimension of the pressure compensation algorithm involves taking the concentration dependence of the pressure compensation into account. To do so, K was calculated for a given pressure for each of the three calibration gases using the model fit of Eq. (2) of each calibration gas. For each calibration gas n , the calculated K was multiplied by c_{P0} to obtain the pressure-normalised value ($c_{n_norm_P}$) of c_{P0} for the given pressure condition:

$$c_{n_norm_P} = c_{P0}K_n. \quad (3)$$

$c_{n_norm_P}$ thus is the concentration expected to be reported by the sensor for c_{P0} at a given pressure for reference gas n . Next, to calculate K for a given reported concentration (K_n) at the given pressure, we used nonlinear regression to relate $c_{n_norm_P}$ and K_n , which forms the second dimension of the pressure compensation algorithm:

$$K_n = C(c_{n_norm_P})^2 - D(c_{n_norm_P}) + E, \quad (4)$$

where C , D , and E are fitted parameter values. These fitted parameter values could then be used to calculate K for any reported concentration at the given pressure via:

$$K_n = C(c_{meas})^2 - D(c_{meas}) + E. \quad (5)$$

Lastly, the obtained K values combined with c_{meas} were converted to pressure-compensated concentrations (c_{comp}) using a modification of Eq. (1), where c_{P0} becomes c_{comp} . The nonlinear regression in Eq. (2) was run using the pressure data from the BME680 sensor located near the CO₂ sensor's inlet, the one located near the CH₄ sensor's outlet, and using the average values of the two pressure sensors. Since the root mean squared error (RMSE) of the regression was lowest when using the pressure data of the BME680 located near the CO₂ sensor's gas inlet, we used the data of that sensor for the empirical fit and as input for the pressure correction function for our field data. When applying the pressure correction function to our field data, extra noise was introduced which likely resulted from pressure fluctuations induced by the membrane pump. As these pressure fluctuations occur at high frequencies, an individual CO₂ or CH₄ concentration measurement may not be



representative of the pressure logged alongside it. We found that the noise was reduced by calculating the mean pressure for each chamber closure and using this mean pressure in the equations for pressure correction of the individual reported CO₂ and CH₄ concentrations. Like pressure, temperature also affects reported CO₂ and CH₄ concentrations. However, both the GMP343 and LGD Compact-A have an internal temperature sensor with a built-in temperature compensation feature.

160 2.3 Compensation for effects of water vapour

Water vapour affects the reported CO₂ and CH₄ concentrations via dilution and broadening of the spectral lines. To correct for this and obtain dry-mole fractions of CO₂ and CH₄, we followed the procedure described by Rella et al. (2013). We humidified compressed air from a SCUBA tank with a gas washing bottle containing ultrapure water (Milli-Q, Millipore). The humidified air, after passing through a 1 µm Gelman filter assembly, was fed into the gas inlet of our system. After
 165 stabilising at maximum water vapour concentrations, the washing bottle was excluded from the gas stream, leading to a gradual decrease in water vapour concentration as the residual moisture was slowly stripped from the filter assembly. After standardising the measured gas concentrations to 1 atm using the above-described pressure compensation method, water vapour correction coefficients were obtained using Eq. (4) of Rella et al. (2013) and nonlinear regression in R (function nlsLM):

$$170 \quad \frac{c_{\text{wet}}}{c_{\text{dry}}} = 1 + aH_{\text{gas_calc}} + bH_{\text{gas_calc}}^2, \quad (6)$$

where c_{wet} is the uncorrected concentration measured by the CO₂ or CH₄ sensor, and $H_{\text{gas_calc}}$ is the water vapour concentration (as a mole fraction in %) of the gas sample calculated from the pressure, %RH, and temperature measurements of the BME680 sensor located after the CO₂ and CH₄ sensors in the plumbing system (Fig. 2), using the equations in Chapter 4 (Volume I) of the 2023 edition of the Guide to instruments and Methods of Observation (WMO-No. 8); a , b , and c_{dry} are
 175 constants obtained from the model fit, where c_{dry} represents the gas concentration reported by the sensor when $H_{\text{gas_calc}}$ is 0. We then used this function with its fitted coefficients to obtain the water vapour-corrected mole fractions of CO₂ and CH₄ (i.e. c_{dry}) for our field measurements. In contrast to the pressure correction, we used data from the BME680 sensor that measured gas exiting the CO₂ and CH₄ sensor, as using these data resulted in a slightly better model fit (i.e. RMSE of predicted vs. observed values) than using the data of the BME680 sensor that was located before the inlet of the CO₂ and
 180 CH₄ sensor.

2.4 Flux estimation

The water-atmosphere diffusive CO₂ flux (in µmol CO₂ m⁻² s⁻¹) was calculated as

$$CO_{2\text{-diff}} = \frac{VP_{\text{air}} \left(1 - \frac{H_{\text{chamb_calc}}}{100} \right) f}{RS(T+273.15)}, \quad (7)$$



where V (m^3) is the sum of the chamber volume above the water and internal volume of the tubing and analysers, P_{air} (Pa) is
 185 the air pressure measured by a nearby weather station (170 m distance), $H_{\text{chamb_calc}}$ is the water vapour mole fraction (%) in
 the chamber headspace calculated from P_{air} and the chamber's headspace RH and temperature using the equations in Chapter
 4 (Volume I) of the 2023 edition of the Guide to instruments and Methods of Observation (WMO-No. 8), f is the rate of
 change in water-corrected CO_2 mole fraction ($\mu\text{mol mol}^{-1} \text{s}^{-1}$) inside the closed chamber, R is the ideal gas constant (8.314 Pa
 $\text{m}^3 \text{K}^{-1} \text{mol}^{-1}$), S (m^2) is the surface area covered by the chamber, and T ($^{\circ}\text{C}$) is the chamber headspace temperature. To
 190 determine f , we used robust linear regression in R using the gasfluxes package (Fuss, 2024) with a regression length of 240 s.
 For V , P_{air} , $H_{\text{chamb_calc}}$, and T , the values represent the mean of the measurements during the regression period.

For CH_4 , we report total emissions rather than separating diffusive and ebullitive fluxes. Total CH_4 flux was calculated using
 Eq. (7), with f determined as the mean of the last six concentration measurements minus the mean of the first six
 195 concentration measurements, divided by the corresponding time interval. We chose this approach because diffusive and
 ebullitive fluxes could not be reliably partitioned since concentrations—even if (pseudo)linear—were large and erratic in
 magnitude, suggesting interference from gas-atmosphere diffusion of CH_4 bubbles trapped at the surface by primary
 producers (Aben et al., 2022). Rather than reporting water-atmosphere diffusive CH_4 fluxes of which a significant part is
 most likely biased (i.e. not originating from dissolved CH_4), we decided to only report estimates of total CH_4 emissions. Part
 200 of the CO_2 measurements were also influenced by ebullition. Hence, we also calculated total CO_2 emissions in the same way
 as described for CH_4 .

2.5 Data exclusion & data loss

Having automated measurements enabled us to collect a large number of flux estimates. However, due to various reasons we
 experienced periods of data loss for one or more chambers. Flux measurements of the first five days after deploying the
 205 chamber system were discarded because they may have been biased by disturbance effects, such as us entering the ditch to
 install bamboo rods as support for various other long-term instrumentation. Furthermore, we excluded data from chamber 2
 between 20 June and 4 July, because a foam kickboard with other instruments had drifted underneath the chamber,
 preventing the chamber from fully closing. Data from all chambers were lost between 3 July and 10 July as we accidentally
 damaged a PCB while installing a new SIM card. From 13 July until the end of the timeseries, a faulty actuator caused
 210 chamber 3 to be frozen in open or closed position for prolonged periods of time. When this occurred, the system detected the
 problem and followed up with a built-in system reset. This resulted in multiple normal chamber closures of chamber 3 during
 the day and night. However, we took a conservative approach by excluding these normal chamber closures from our
 analysis, as the prolonged chamber closures that occurred between good measurements of chamber 3 may have caused
 artefacts on aquatic vegetation or the processes driving the GHG fluxes below that chamber. Our time series measurements



215 unintentionally stopped on 10 September 2024 because a cow had fallen into the ditch on top of the chamber system (Fig. S2), forcing us to retrieve the system for repairs.

2.6 Quality control

We filtered the data for CO₂ concentrations below 2,000 ppm to avoid including measurements outside the analogue output range (any value above 2,000 ppm is logged as 2,000 ppm). We also excluded values below 100 ppm, as this is unlikely to occur within the set chamber closure time. Both concentration filters helped identify erroneous measurements of the chamber that experienced a problematic actuator during part of the study period (as described in the previous section). Time vs CO₂ concentration plots of each chamber measurement were visually inspected to explore the data, identify periods with erroneous data (chamber not [fully] closing or not opening as discussed in the previous paragraph), and as a check for our data filters. CO₂ concentrations outside the above-mentioned filtering range only occurred during the period when chamber 3 had actuator problems, causing it to freeze in a closed position. Measurements that were affected by ebullition (characterised by a sudden jump in concentration) were only used for calculation of the total CO₂ flux and removed from calculation of the diffusive flux. Resulting gaps in diffusive fluxes from bubble-affected measurements were gap-filled via linear interpolation of the high-frequency measurements only for the purpose of estimating the relative magnitude of CO₂ ebullition (i.e., all visualisations and statistics for diffusive CO₂ fluxes are based on quality-controlled, non-gap-filled measurements).

230 We filtered the CH₄ concentrations to the calibrated operating range of the sensor (0–100 ppm), thereby excluding implausible zeroes and values ≥ 100 ppm. Because the analogue output range of the CH₄ sensor was limited to 100 ppm, concentrations above this threshold (mainly due to ebullition events) were logged as 100 ppm. This plateau without doubt caused an underestimation of the 24 h mean flux and dampened the measured variability in CH₄ fluxes. Therefore, the reported CH₄ fluxes should be regarded as conservative in terms of both magnitude and temporal dynamics. Duplicate timestamps were removed. To limit biased flux estimates, we excluded measurements of which the start or end coincided with an ebullition event (e.g. rapid increase in CH₄ concentration due to bubble entry or a strong decrease in concentration due to headspace mixing following bubble entry). To do so, closures were retained only when short-term variability was low (SD < 2 ppm) among the first six and last six concentration measurements. To further constrain potentially biased flux estimates, we excluded measurements where the mean of the first six concentration measurements > 10 ppm. Periods linked to known system issues, as described above, were also excluded. Additionally, fluxes were discarded when CH₄ fluxes were negative while initial and/or final concentrations were high (> 10 ppm). Specifically, we checked whether a statistically significant ($p < 0.05$) decrease occurred within the first 50 s at elevated starting concentrations (> 5 ppm); indicative for mixing after bubble entry. For the calculation of 24 h mean CO₂ fluxes, only days with a minimum of 32 (out of a maximum of 39) measurements were included to avoid biased averages due to having disproportionate daytime/nighttime measurements. For 24 h mean CH₄ fluxes, only days with at least 20 valid fluxes were included. A lower threshold was



deemed acceptable given the low diel variability in total CH₄ fluxes, with no detectable differences between day and night, and because fewer valid observations were obtained due to the 100 ppm analogue output limit of the sensor.

2.7 Calibration and validation of concentrations and fluxes

250 The calibration of the gas sensors was validated using the three calibration gases mentioned in Sect. 2.2. Since the CO₂ and CH₄ concentrations on the labels of the calibration tanks came with a 5 % analysis uncertainty from the manufacturer, we used a cavity ringdown spectroscopy analyser (G2508, Picarro) with recent factory calibration to determine the CO₂ and CH₄ concentrations of the calibration gases with high accuracy (these are the concentration values mentioned in Sect. 2.2). We found that the slope of the linear relationship between the known calibration values and the GMP343's reported CO₂ concentrations—after two-dimensional pressure compensation—was 1.004. For the LGD Compact-A the slope of the relationship between the pressure-corrected CH₄ concentrations and the calibration values was 1.033. All pressure- and water vapour-compensated CO₂ and CH₄ concentrations were corrected using the slopes and intercepts obtained from these calibration procedures before calculation of the fluxes.

260 An extra validation procedure was performed in late September 2025, when the floating automated chamber system was deployed in an experimental ditch at the university campus. A Microportable Greenhouse Gas Analyser (LGR-ICOS™ M-GGA-918 [CO₂/CH₄/H₂O]; ABB Inc.) was temporarily added as a reference analyser in a parallel plumbing configuration on the tubing of one of the three automated chambers, allowing it to sample gas that was returning to the chamber after passing through the floating system's multiplexer and internal sensors.

265 2.8 Auxiliary measurements

Photosynthetically active radiation (PAR) was measured using an SQ-110 sensor (Apogee Instruments) mounted on the chamber system (Fig. 1). Water and sediment temperatures were measured using a custom-made temperature probe containing 10 temperature sensors (DS18B20, Dallas Semiconductor; ±0.5 °C accuracy between -10 °C and +85 °C), spaced at 10 cm distance between each other (Fig. S3). The central unit of the system also has integrated electronics to connect a pH and dissolved oxygen sensor (ENV-40-pH and ENV-40-DOX, Atlas Scientific). We had these sensors deployed just below the surface of the water. However, because both sensors were not regularly calibrated, we did not include their data in this study.

2.9 Comparison with alternative flux measurement strategies

We assessed how different measurement strategies affect the accuracy of diffusive CO₂ and total CH₄ flux estimates using the high-frequency dataset from chamber 1, which provided the most complete record. The mean flux intensity of the full 97-day time series served as a reference representing high-frequency observations.



These references served as the basis for evaluating to what extent the mean emission intensities produced by the simulated measurement strategies approached the corresponding reference means. We evaluated two general approaches that represent common trade-offs between measurement effort, temporal resolution, and spatial coverage:

1. short-term daytime measurements with manually operated closed chambers, and
2. periodic multi-day measurements with automated chambers.

Both approaches were evaluated using Monte Carlo simulations to test how reduced measurement frequency or restricted measurement periods affect the accuracy of mean flux estimates relative to the high-frequency references.

2.9.1 Periodic single daytime measurements

Infrequent daytime measurements are typically conducted during daytime hours. To quantify the potential bias introduced by extrapolating infrequent and/or daytime-only measurements (as in, for example, Van der Knaap et al. (2025)), we evaluated a range of measurement scenarios differing in measurement frequency and timing. We derived two reference means from the 97-day time series, which capture distinct observational perspectives: a daytime mean (08:00–18:00) reflecting typical manual chamber measurements, and a 24 h mean representing the true daily flux.

Simulations were conducted in R (version 4.3.1) using a Monte Carlo framework with 1,000 iterations (random seed = 42 for reproducibility). Measurement intervals ranged from $f = 1$ –30 days, including commonly used daily ($f = 1$), weekly ($f = 7$), biweekly ($f = 14$), and monthly ($f = 30$) measurement schemes. For each interval, approximately N/f measurement days ($N = 97$ days total) were selected and distributed as evenly as possible across the period to avoid clusters of consecutive days. On each selected day, one flux measurement was randomly drawn from the daytime period (08:00–18:00) of our high-frequency dataset. The mean flux over the 97-day period was then estimated for each simulation by averaging these single-day values.

We performed 1,000 Monte Carlo simulations for each measurement scenario. The resulting distributions of mean flux estimates were summarised by their mean and central 90 % interval (5th–95th percentiles) and compared with the daytime and 24 h reference means based on the full high-frequency dataset. Accuracy was evaluated as the proportion of Monte Carlo estimates falling within ± 20 % of the respective reference means (daytime or 24 h). This framework therefore quantifies how a measurement strategy affects the accuracy of mean flux estimates, indicates how much effort is required to approach the reference means (here we used ± 20 % of the 24 h mean), and directly assesses the bias introduced by restricting measurements to daytime only.



2.9.2 Periodic multi-day measurements using automated chambers

An alternative strategy involves performing multi-day automated chamber campaigns at multiple sites rather than maintaining continuous operation at a single location. This approach trades temporal detail for broader spatial coverage: more sites can be measured, but at each site fewer direct flux measurements are conducted. Such strategies are useful when 1) insight in spatial coverage outweighs the accuracy of the mean emission estimate for a single site, and 2) gaps in flux estimates in between measurement campaigns can be gap-filled with a high likelihood of obtaining a fairly accurate mean flux estimate (e.g. within 20 % of the 24 h reference mean).

We evaluated this approach by simulating periodic campaigns with automated chambers in R (version 4.0.4) following the design of Nijman et al. (2024), who rotated automated chamber systems among grassland sites and filled the data gaps using Random Forest (RF) machine learning. Two measurement scenarios were simulated: measurement campaigns taking place every two and every three weeks. For the former, random intervals between subsequent campaigns were allowed to vary between 10 and 18 days to approach reality, rather than a fixed interval of 14 days. For the latter, the random interval varied between 17 and 25 days. The first measurement campaign was forced to take place within the first week after the start of the full dataset. The duration of a campaign was randomly chosen as ~2–3 days, following the campaign-based measurements in Nijman et al. (2024). The simulated deployment start and end of the chamber system was allowed to happen (randomly) between the hours of 08:00 and 18:00 on weekdays only, where Thursday starts were disallowed, Wednesday starts were forced to end on Fridays, and Friday starts were forced to end the following Monday, given the aforementioned operational constraints. Based on the resulting chamber system deployment periods of each simulation, fluxes were subsampled from the full dataset and gap-filled as explained below.

For each of the two scenarios (i.e., measurement campaigns happening every two or every three weeks) we ran 5,000 simulations, each with a unique and fixed seed number for reproducibility. For each simulation, CO₂ fluxes of the full dataset that were outside of the randomly generated deployment periods were set to missing, resulting in a dataset with artificial gaps. A RF regression model was then trained on the subset with remaining observations using PAR, chamber headspace temperature, sediment temperature, solar azimuth, solar elevation, cumulative PAR since the start of measurements, and mean rainfall of the last hour as predictor variables. For each simulation, hyperparameters of the RF model ('mtry', 'splitrule', and 'min.node.size') were automatically tuned via the 'caret' package (Kuhn, 2008), using 5-fold cross-validation with 500 trees, tuneLength = 8, method = "ranger", and metric = "RMSE" as arguments. Subsequently, a final RF model with 1,000 trees was fit to the training subset using the auto-tuned hyperparameters via the 'ranger' package (Wright and Ziegler, 2017). Performance of the final RF model was evaluated using RMSE, MAE and R², which were obtained using ranger's aggregated out-of-bag (OOB) predictions for each observation via the 'MLmetrics' package (Yan, 2024). The RF model was then used to only fill the artificial gaps resulting from the subsampling procedure (i.e., randomly generated



340 chamber system deployment periods). Finally, the arithmetic mean of the gap-filled CO₂ fluxes was calculated and saved, together with model information (e.g., seed and simulation number, model RMSE and R², and number of gap-filled observations). We also saved the feature importance of the final RF model of each simulation. The distribution of mean gap-filled CO₂ fluxes, based on 5,000 Monte Carlo simulations, was then compared against the true arithmetic mean (based on the full dataset) to assess how intermittent multi-day measurement scenarios affect overall flux accuracy. The analyses
 345 described in this paragraph were only performed for the diffusive CO₂ fluxes as an illustration. An identical or similar approach could be used for CH₄, which was not done here considering the aforementioned limitations of our CH₄ measurements.

3. Results

3.1 Flux estimates

350 After applying the filtering procedure described in the Sect. 2.6, we collected a total of 7614 CO₂ and 6448 CH₄ flux estimates using three chambers over 97 days (covering the height of the growing season and presumed peak emission). Individual CO₂ flux estimates ranged from -1.63 to 3.26 μmol CO₂ m⁻² h⁻¹ (Fig. S4). The individual CH₄ fluxes ranged from -0.02 to 1.75 μmol CH₄ m⁻² s⁻¹. The 24 h average CO₂ flux estimates for individual chambers ranged from 0.06 to 6.54 g CO₂ m⁻² d⁻¹ (Fig. 3), with a median value of 1.83. The 24 h average CH₄ fluxes ranged from 0.08 to 3.38 g CH₄ m⁻² d⁻¹, with a
 355 median value of 1.25 (Fig. 4). CO₂ ebullition was observed on multiple occasions and contributed 2.0 %, 4.2 %, and 2.3 % to the mean total CO₂ flux for chamber 1–3, respectively (Fig. S5). The coefficient of variation (CV) of the 24 h average CO₂ flux estimates among the three chambers (only for dates when all three chambers had data, i.e. Fig. 3b) varied from 19 to 75 % (median and mean: 42 % and 44 %). When considering the average CO₂ flux of the full period (i.e. Fig. 3b) for each chamber, the CV among the three chambers was 27 %. For the 24 h average CH₄ flux estimates the CV ranged from 20 to 84
 360 % (median and mean: 62 % and 59 %). For the means of the full period, the CV was 53 %.

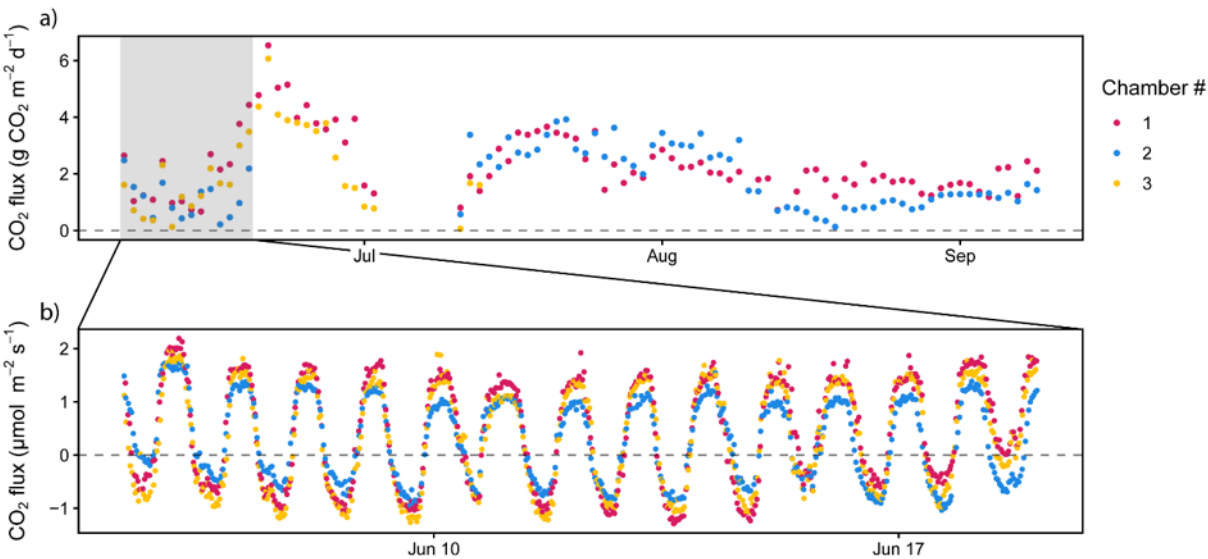


Figure 3. Temporal and spatial variation in CO₂ fluxes: **(a)** 24 h mean CO₂ fluxes for each chamber, and **(b)** individual CO₂ flux estimates for each chamber within the shaded area in panel (a). Colours represent different chambers. The dashed horizontal line represents zero emissions.

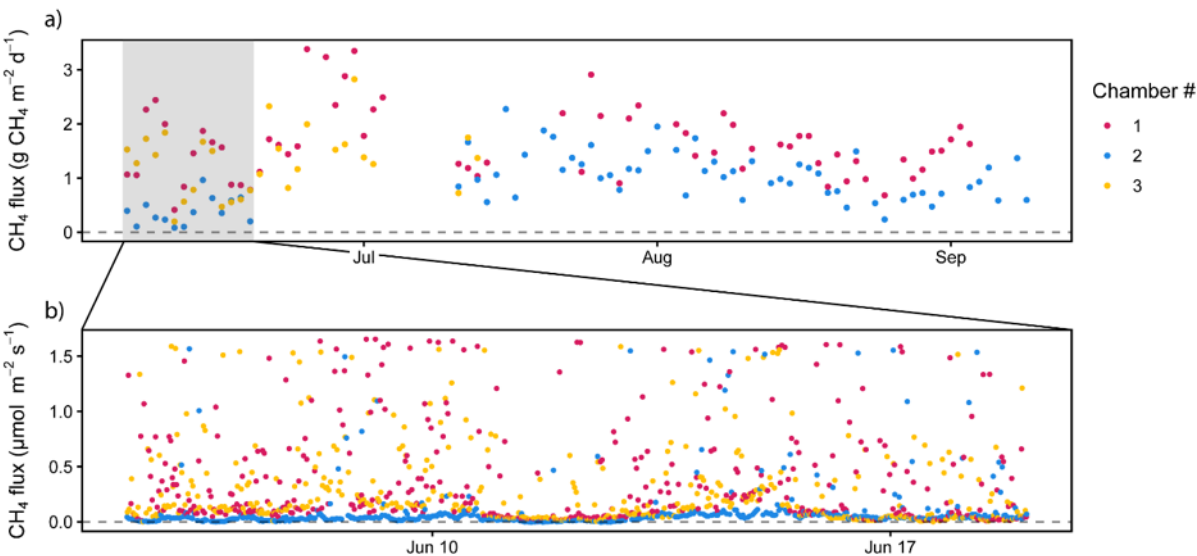


Figure 4. Temporal and spatial variation in CH₄ fluxes: **(a)** 24 h mean CH₄ fluxes for each chamber and **(b)** individual CH₄ flux estimates for each chamber within the shaded area in panel (a). Colours represent different chambers. The dashed horizontal line represents zero emissions.



3.2 Compensation for effects of pressure

The impact of pressure on reported CO₂ concentrations depends on the CO₂ concentration of the measured gas (Fig. 5). After applying the two-dimensional pressure compensation algorithm to the reported CO₂ concentrations used to calculate the flux estimates, we found that the mean CO₂ flux was 37 % (chamber 1) and 17 % (chambers 2 and 3) higher with the two-dimensional pressure compensation than without it. In addition, we found that the mean CO₂ flux was 14 % (chamber 1), 6 % (chamber 2), and 7 % (chamber 3) higher with our empirically derived pressure compensation than with pressure compensation based on applying the ideal gas law. Because the membrane pump draws in air, the pressure inside the CO₂ and CH₄ sensors is always slightly lower than the atmospheric pressure. However, for chamber 1, this pressure was substantially lower (Fig. S6), likely due to gradually increased clogging or malfunctioning of a valve, thereby explaining the larger effect of applying pressure compensation on the mean CO₂ flux for this chamber as compared to the other two chambers. This highlights the importance of measuring the internal pressure when using CO₂ and CH₄ sensors without internal pressure control in flow-through applications. The effect of pressure on the CO₂ flux estimates is shown in Fig. S7.

For CH₄, we also found a clear pressure dependence, although the effect was less strong than for CO₂. We did not observe evidence for an interactive effect of pressure and the CH₄ concentration of the gas on the reported CH₄ concentrations (Fig. S8). Although the fitted line of the 2 ppm CH₄ calibration gas shows a slight deviation from the ones based on the 5 and 10 ppm CH₄ calibration gases, the noise-to-signal ratio (and thereby the uncertainty in the fit) for the same procedure using 2 ppm CH₄ was rather high. Our fitted lines for 5 and 10 ppm agree well with each other and with a fit based on the pressure test data for 100 ppm CH₄ gas (Fig. S8) that was shared with us by Axetris (Kaegiswil, Switzerland). Hence, we pooled the pressure experiment data of the 5 and 10 ppm CH₄ gas and modelled the relationship between pressure and the reported CH₄ concentration using Eq. (2), to calculate the CH₄ concentration-independent *K*-values for the pressure compensation of the field data. Here, we found that the mean total CH₄ flux was 9 % (chamber 1) and 4 % chamber (2 and 3) higher with our empirically derived pressure compensation than without it. Additionally, the mean total CH₄ flux was 11 % (chamber 1) and 6 % (chamber 2 and 3) lower with our empirically derived pressure compensation as compared to using the ideal gas law for pressure compensation.

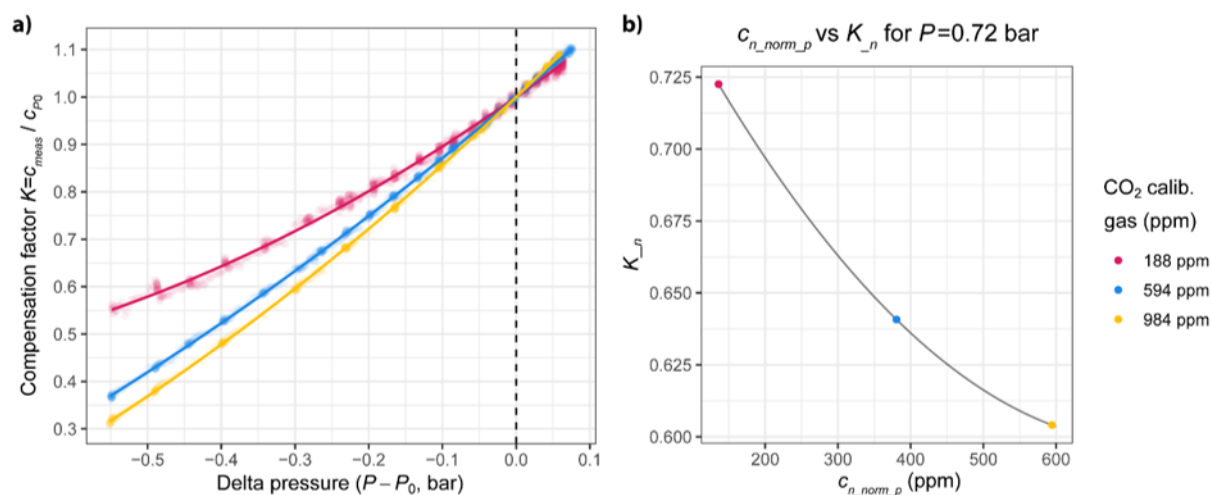


Figure 5. a) Compensation factor K for CO₂ as a function of delta pressure for each of the three calibration gases. The dots represent the values of K for individual measurements (Eq. (1)). Lines represent fitted relationships based on Eq. (2). **b)** An example of the second dimension of the pressure correction algorithm, where $c_{n_norm_P}$ is the expected reported concentration for reference gas n at a given pressure (in this example, 0.72 bar), and K_n is the compensation factor for the given reference gas and pressure. After relating the two variables via nonlinear regression (Eq. (4)), K_n can be accurately calculated for any reported concentration within the calibrated range, such that the reported concentration can be converted to a concentration at standard pressure using Eq. (3).

3.3 Compensation for effects of water vapour

The effect of water vapour concentration on measured CO₂ mole fractions is smaller than the dilution-only effect of water vapour (Fig. 6). Hence, using a dilution-only correction would overcompensate for the true effect of water vapour on the measured CO₂ mole fractions. When comparing the (pressure-compensated) mean CO₂ flux with and without our empirically derived water vapour correction, we found that the mean CO₂ flux with water vapour correction was 0.7 %, 1.5 %, and 1.0 % higher than the flux without correction for chambers 1, 2, and 3, respectively. Compared to a dilution-only water vapour correction, the mean CO₂ flux—based on the empirically derived water vapour correction—was 0.4 %, 0.7 %, and 0.5 % lower for chambers 1, 2, and 3, respectively. More details about the effects of water vapour on the CO₂ flux estimates are shown in Fig. S9–S11. For CH₄, the effect of water vapour on the measured concentrations was also relatively small (Fig. S12). The mean CH₄ flux with the empirically derived water vapour correction was 1.9 %, 1.9 % and 2.4 % higher for chambers 1, 2, and 3, respectively, than the flux uncorrected for effects of water vapour. When compared to a dilution-only water vapour correction, the mean CH₄ flux—based on the empirically derived water vapour correction—was 0.6 %, 0.5 %, and 0.4 % lower for chambers 1, 2, and 3, respectively.



The range in water vapour concentration (0.76–5.41 %) used for our calibration encompassed 99 % of our measurements. Only during some very humid events the water vapour concentration fell outside the calibrated range. The accuracy of our water vapour concentrations, calculated from measurements of relative humidity, temperature, and pressure using a BME680 sensor (see Sect. 2.3), was validated in a separate experiment using a reference instrument (Fig. S13).

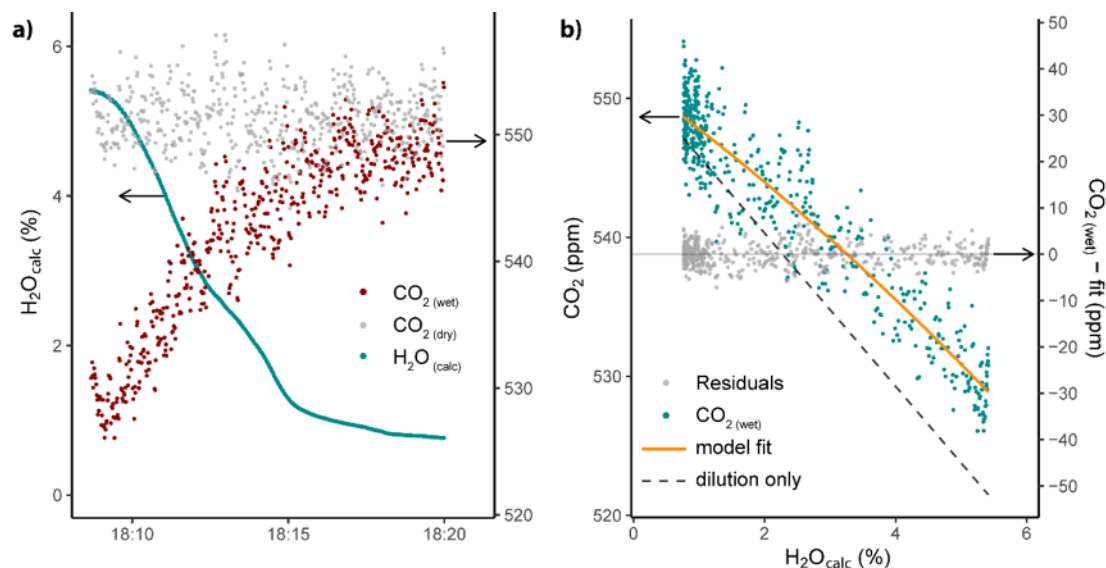


Figure 6. a) Changes in water vapour uncorrected reported CO_2 mole fractions ($CO_{2(wet)}$; red dots) and H_2O fractions (H_2O_{calc} , cyan line; calculated from measurements of relative humidity, temperature, and pressure [see Sect. 2.3 for details]) over time during the water vapour interference experiment. The grey dots denote the water vapour-corrected CO_2 concentrations ($CO_{2(dry)}$) based on the parameters obtained from the model fit (Eq. (6)). **b)** Water vapour-uncorrected reported CO_2 mole fractions ($CO_{2(wet)}$, cyan dots) against H_2O fractions. The orange line represents the regression line (i.e., the fitted correction function). The grey dots denote the residuals (right y-axis). The black dashed line demonstrates the bias caused when only accounting for the dilution effect of H_2O .

3.4 Validation of flux estimates using a reference analyser

The flux estimates calculated using the concentration data of the floating system's internal CH_4 and CO_2 analysers (compensated for effects of pressure and water vapour) compared well against those calculated from the data of the Microportable Greenhouse Gas Analyser (MGGA) that was added temporarily to the system as a higher quality reference instrument (Fig. S14–S16). The mean CO_2 flux during the multiday validation test was -0.055 and $-0.063 \mu mol CO_2 m^{-2} s^{-1}$ for the GMP343 and MGGA, respectively (a discrepancy of 12 %). The mean CH_4 flux was 0.0266 and $0.0259 \mu mol CH_4 m^{-2} s^{-1}$ for the LGD Compact-A and MGGA, respectively (a discrepancy of 2 %).



3.5 Measurement strategy simulations

For the first measurement strategy (single daytime measurements with manually operated closed chambers), measurement frequency strongly affected the agreement between mean fluxes of the simulated measurement campaigns and the high-frequency reference mean (Fig. 7). In each scenario, one random daytime flux measurement (08:00–18:00) was selected for each sampled day, and these values were averaged to obtain a 97-day mean flux estimate. For each measurement frequency we repeated this 1,000 times. In the scenario in which we sampled daily (i.e. 97 times), 62 % of simulated CO₂ means fell within ± 20 % of the daytime reference mean. In contrast, none of the simulated mean daily flux estimates reproduced the 24 h reference mean within ± 20 %, indicating a systematic bias when nighttime emissions were excluded. As measurement frequency decreased, agreement with the daytime reference declined sharply. Only 21 % of weekly, 14 % of biweekly, and 8 % of the estimated means of the monthly measurement scenarios produced the daytime reference mean within ± 20 %. Agreement with the 24 h reference remained minimal across all reduced scenarios (< 2 %), demonstrating that infrequent and daytime-only measurements consistently underestimated the true 24 h reference CO₂ mean.

Total CH₄ emissions were less sensitive to reduced measurement frequency. With daily measurements, 94 % of simulated mean flux estimates fell within ± 20 % of the daytime reference mean and 90 % were within ± 20 % of the 24 h reference mean. With weekly measurements, still 48 % of the estimated flux means were within ± 20 % of the daytime reference mean and 46 % of the 24 h reference. Biweekly measurements yielded 35 % and 36 %, and monthly measurements 23 % and 23 %, respectively. Agreement with daytime and 24 h reference means was therefore similar across measurement frequencies, suggesting smaller diel variability in total CH₄ emissions compared to CO₂ emissions. At the lowest measurement frequencies—particularly monthly—the simulated means tended to underestimate rather than overestimate the 24 h reference mean, indicating a slight directional bias.

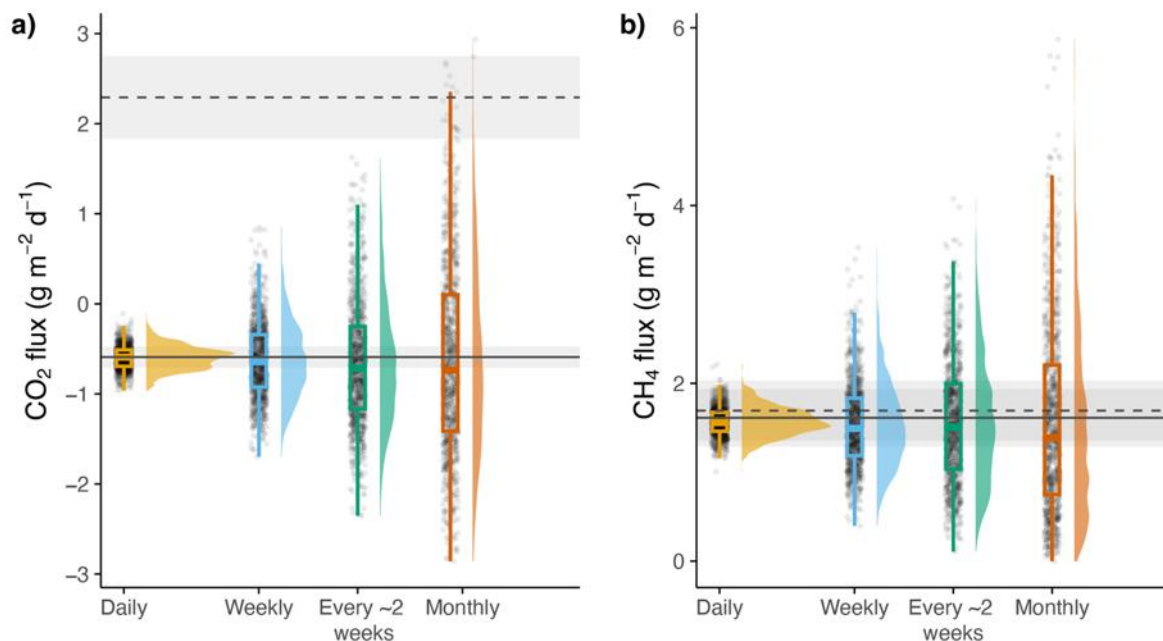


Figure 7. Monte Carlo simulations of 97-day mean **a)** diffusive CO₂ and **b)** total CH₄ fluxes under different measurement scenarios (one random daytime flux measurement between hours of 08:00 and 18:00 on a daily, weekly, every ~two weeks, or monthly basis). Boxplots show the median and interquartile range of simulated means ($n = 1,000$), whiskers extend to $1.5 \times \text{IQR}$, and raincloud plots display the full distributions. Shaded grey bands denote the $\pm 20\%$ range around the daytime reference mean (08:00–18:00; solid line) and the 24 h reference mean (dashed line).

For the second measurement strategy (periodic multi-day measurements with automated chambers combined with machine learning-based gap-filling), we found a much higher likelihood of obtaining a mean flux estimate that is close to the high-frequency 24 h reference (Fig. 8). When chamber campaigns were conducted about every two weeks, 89 % of simulated 97-day mean fluxes fell within $\pm 20\%$ of the 24 h reference mean. When campaigns were conducted about every three weeks, this proportion decreased to 48 %. Across the 5,000 Monte Carlo simulations of the two-week interval measurement campaign scenario, variables related to light availability consistently had the highest predictive importance. Median feature importance values were 0.40 for solar azimuth, 0.24 for PAR, 0.21 for solar elevation and 0.18 for cumulative PAR, followed by chamber temperature (0.15), sediment temperature (0.11), and rainfall in the previous hour (0.06). Similar rankings were obtained when considering the 10 % of models with the lowest RMSE. An example of the relationship between PAR and diffusive CO₂ fluxes is shown in Fig. 9 based on high-frequency measurements from chamber 1 during a four-day period in August 2024.

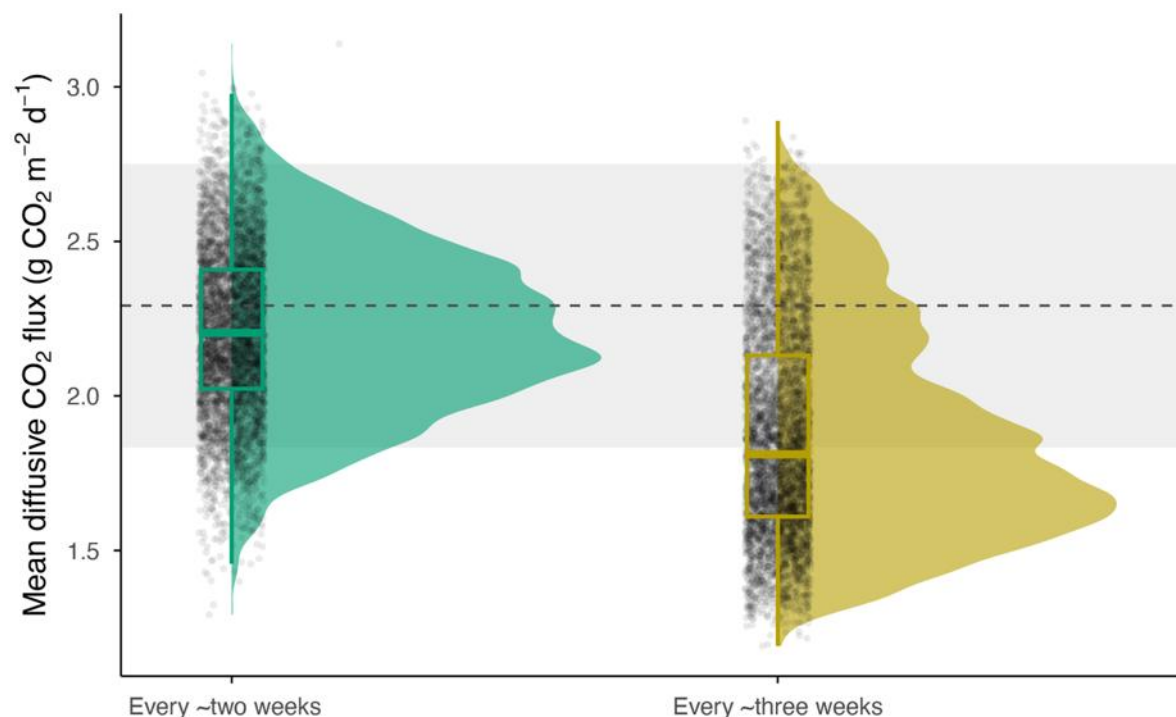


Figure 8. Monte Carlo simulations of 97-day mean diffusive CO₂ fluxes under periodic multi-day measurement scenarios using automated chambers combined with machine-learning-based gap-filling. Raincloud plots show the distributions of simulated mean fluxes for every ~two weeks (green) and every ~two weeks (olive) measurement scenarios ($n = 5,000$). Boxplots indicate the median and interquartile range, with whiskers extending to $1.5 \times \text{IQR}$. The dashed line represents the 24 h reference mean derived from the 97-day, high-frequency dataset, and the shaded grey band denotes the $\pm 20\%$ range around this reference.

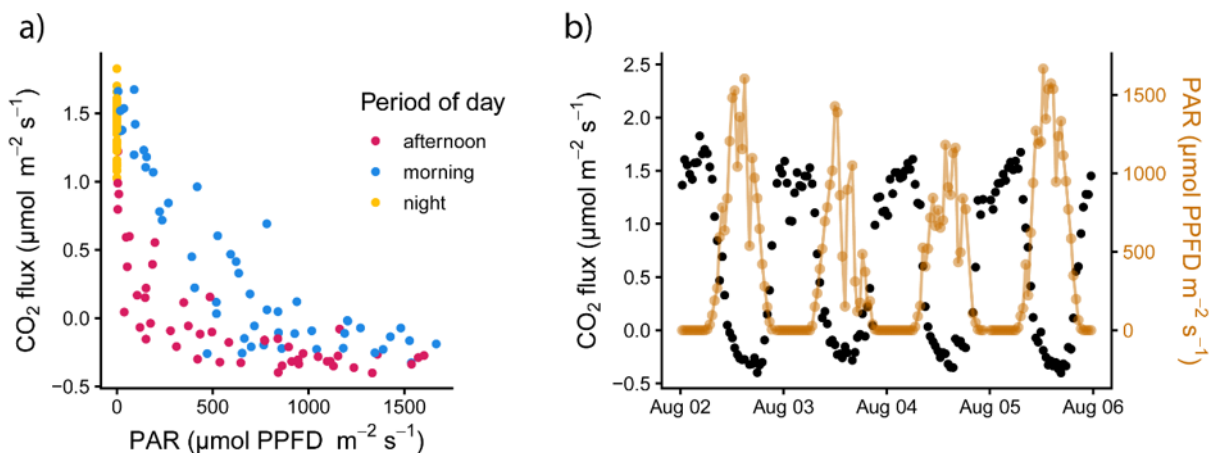




Figure 9. a) An example of the relationship between diffusive CO₂ fluxes and photosynthetically active radiation (PAR) measured during a four-day period in August 2024. Each point represents a flux measurement from chamber 1. **b)** Time series of CO₂ fluxes (black) and PAR (orange) for the same period.

4. Discussion

Our automated chamber system BERTHA has demonstrated its ability to accurately measure CO₂ and CH₄ fluxes at high resolution using relatively affordable sensors. The integration of automated control, high measurement frequency, and comprehensive correction procedures facilitates stable long-term, high-frequent monitoring without the need for frequent manual intervention. This capability is particularly valuable for studying short-term (sub-daily) flux dynamics and process controls in wetland systems, where GHG exchanges can vary substantially over the course of a day. It also overcomes the high likelihood of obtaining strongly biased flux estimates (due to temporal variation) using manual chamber measurements. Several technical and analytical challenges were identified and addressed during the development and validation of the system. These include pressure compensation (Sect 4.1), water vapour compensation (Sect 4.2), validation of flux estimates against a reference analyser (Sect 4.3), and quantifying the likelihood of biased results when using alternative measurement strategies (Sect 4.4). Each of these aspects is discussed in detail in the following sections.

4.1 Pressure compensation

We found that our CO₂ flux estimates would have had a significant bias if no compensation for pressure effects had been applied, whereas for CH₄ effects were smaller. Pressure compensation of measurements by the GMP343 has been applied by other researchers using a theoretical approach based on the ideal gas law (e.g. Sharma et al. (2014) and Shusterman et al. (2016)) and by empirical determination of the pressure dependence based on a single calibration standard, thereby assuming pressure compensation to be independent of the measured concentration (e.g. Kulmala et al. (2009)). The latter method likely provides fairly accurate pressure compensation when the measured CO₂ concentrations deviate only slightly from the concentration at which the pressure dependence relationship was determined. In situations like ours, where a large range of CO₂ concentrations occurs, both methods will produce an inaccurate pressure compensation given our finding that the relationship between pressure and reported CO₂ concentrations of the GMP343 is clearly concentration-dependent, similar to what Gaynullin et al. (2023) found for other NDIR CO₂ sensors. Hence, future studies using the GMP343 should account for the concentration dependence of the pressure response. Assuming negligible sensor-to-sensor variation in the pressure response, the empirically derived compensation factors reported here for the two-dimensional pressure compensation algorithm (Table S1) may be used by other researchers employing the GMP343 sensor. The same applies to the LGD Compact-A CH₄ sensor, albeit here only a one-dimensional pressure compensation is needed.



4.2 Water vapour compensation

Our finding that the effect of water vapour concentration on measured CO₂ mole fractions is smaller than the dilution-only effect of water vapour for the GMP343 contrasts with results for other analysers, such as for a Picarro G2401 CRDS analyser (Rella et al., 2013), where the true water vapour effect was found to be larger than that of dilution only. This also means that attempts to correct GMP343's reported CO₂ concentration for the effects of water vapour via a dilution-only correction, such as in Sharma et al., 2014, overcompensate for the true effects of water vapour. Nevertheless, our results show that accounting only for the dilution effect of water vapour is better than not accounting for effects of water vapour at all. Similar to pressure compensation, researchers employing the GMP343 may use the empirically determined water vapour correction reported here (Table S2), assuming negligible sensor-to-sensor variation in the response to water vapour.

For the LGD Compact-A CH₄ sensor, our results differ from those of Lin et al. (2024), who quantified the effect of water vapour on reported CH₄ concentrations on a combined CH₄/CO₂ version of the LGD Compact-A. Their results show a decrease in reported CH₄ concentrations with increasingly humidified calibration gas, whereas in our case, a nonlinear increase in reported concentrations is seen. A likely explanation for the difference is that the CH₄ concentration used in our water vapour compensation procedure was too low and the effect of water vapour too small to reliably separate signal from noise—similar to what we described for the results of the pressure compensation procedure when using the ~2 ppm CH₄ calibration gas. For both CH₄ and especially for CO₂, the effect of the water vapour correction on the flux estimates was relatively small and of much less importance than correcting for pressure.

4.3 Validation of gas flux estimates using a reference analyser

The agreement between fluxes calculated using data from the internal CO₂ and CH₄ analysers (GMP343 and LGD Compact-A, respectively) and the reference analyser (MGGA) demonstrates that the mid-cost analysers integrated in our chamber system—when combined with the correction procedures that we described—are a fair alternative to the lab-grade analysers that are commonly used for flux measurements that are at least €25k more expensive. The mid-cost analysers allow for accurate flux measurements without the need for undesirable compromises that often come with the use of low-cost sensors, such as the need to use long chamber closure times to obtain sufficient signal-to-noise ratios. Being able to keep chamber closure times short by using mid-cost sensors minimises the disturbance of the studied ecosystem and will limit deviation of the chamber's microclimate (e.g. temperature and pressure) relative to the surrounding climate (Maier et al., 2022), which promotes making representative measurements. The 12 % difference found in the mean CO₂ flux based on the GMP343's and the MGGA's data (Fig. S14 and Fig. S16) can most likely be attributed to 1) the noise of the GMP343 that was relatively high compared to the signal (flux magnitude) in the experimental ditch where the comparison data was collected (the range in flux magnitude was about 10 times smaller in the experimental ditch than in our field site [Fig. S4 and Fig. S14]); and 2) the very small magnitude of the mean CO₂ flux (~0.2 g CO₂ m⁻² d⁻¹) that causes any small differences in flux estimates



545 between the GMP343 and MGGA's data to result in a relatively large percent difference in means. Additionally, both the GMP343 and MGGA were calibrated using the same calibration standards and are therefore not expected to give significantly different flux estimates during the validation procedure in the experimental ditch.

4.4 Measurement strategy evaluation

Our simulations indicate that infrequent and daytime-only measurements introduce substantial uncertainty and systematic
 550 bias in CO₂ flux estimates from agricultural ditches. Reduced measurement frequency and daytime only simulations consistently lead to underestimation of the 24 h reference mean derived from the full 97-day high-frequency dataset. In contrast, periodic multi-day automated-chamber campaigns conducted approximately every two weeks, combined with machine-learning-based gap filling, had a high likelihood of producing mean flux estimates within ± 20 % of the 24 h reference mean. Extending the interval to three weeks distinctly reduced this likelihood, although the performance remained
 555 considerably better than the infrequent single daytime-only manual chamber measurements.

This systematic underestimation in daytime-only scenarios results from pronounced diel variability driven by photosynthetic uptake of CO₂ (Podgrajsek et al., 2015; Miller et al., 2024; De Eyto et al., 2025). These findings align with high-frequency eddy covariance observations across lakes and reservoirs (Golub et al., 2023), which report pronounced diel and sub-
 560 monthly variability, with daytime CO₂ fluxes 30–40 % lower than nocturnal values on average. The Monte Carlo simulations further showed that the share of multi-week mean estimates that resembled the reference mean decreased rapidly as measurement frequency decreased. Even when using only daytime observations, longer temporal gaps widened the range of plausible 97-day mean values, demonstrating how conducting sparse measurements amplifies uncertainty in upscaled CO₂ fluxes. Using infrequent, daytime only measurements for CO₂ flux estimates introduces systematic underestimation not only
 565 due to not accounting for diel dynamics, but also due to not accounting for other causes of temporal variability (e.g. changes in vegetation and rainfall-driven contributions). In practice, field measurements schedules are often constrained by weather, logistics, and personnel availability, which reinforces the dominance of daytime observations and irregular temporal coverage.

570 For CH₄, only the simulations using single daytime measurements with various intervals were evaluated. In this scenario, individual daytime measurements were compared with both the daytime and the 24 h reference means for the full 97-day period. The CH₄ sensor's upper analogue output limit of 100 ppm truncated ebullition-related peaks and reduced observable short-term variability in the high-frequency record, making CH₄ unsuitable for the periodic multi-day Monte Carlo simulations, which rely on gap-filled time series. Diel diffusive CH₄ patterns are known to be highly variable, with studies
 575 reporting higher nighttime emissions (Podgrajsek et al., 2014; Paranaíba et al., 2025), higher daytime emissions (Bastviken et al., 2010; Sieczko et al., 2020), or no consistent diel signal (Dalbem Barbosa, 2026, Chapter IV, pp. 64–77). Ebullition



further complicates interpretation because it is governed by sediment gas content and episodic processes—such as hydrostatic pressure changes and mechanical disturbances—that may or may not follow a diel cycle. When these processes do not follow a diel cycle, daytime measurements could still approximate the 24 h total CH₄ reference mean reasonably well, whereas accuracy declined rapidly at weekly to monthly intervals, indicating that stochastic short-term variability cannot be reliably captured by sparse measurements alone.

In addition to temporal variability within individual chambers (as discussed above), spatial variation must also be considered, since flux estimates among our three chambers varied substantially despite their close proximity. Temporal variability was observed not only in the magnitude of differences among chamber flux estimates, but also in which chamber exhibited the highest and lowest emissions. Variation in flux estimates among chambers is likely driven by spatial heterogeneity in features as floating algae beds, submerged and floating vegetation or small-scale differences in sediment characteristics. This shows that representative flux estimates from a given ditch may require spatially replicated chamber deployments in addition to sufficient temporal cover. Wik et al. (2016) showed for their post-glacial lakes that 3 and 11 sampling locations were needed for high confidence of accurate diffusive and ebullitive fluxes, respectively. Lastly, the system and its analysers being mid-cost offer potential for increased spatial cover by having multiple systems deployed at multiple sites in addition to potential rotation of a single system among multiple sites (i.e., the simulation results shown in Fig. 8).

4.5 Recommendations

Iterative deployment of BERTHA in small inland waters (e.g. ditches) combined with targeted testing and refinement will help improve system performance and robustness. Further development of BERTHA will focus on enhancing its ability to resolve distinct emission pathways and link them to environmental drivers. In particular, differentiating between diffusive and ebullitive CH₄ fluxes will enable a more mechanistic interpretation of the observed dynamics and strengthen the connections to underlying physical and biogeochemical processes. This can be achieved through hardware improvements, such as implementing an automated bubble trap and changing the Axetris CH₄ sensor's output method from analogue to digital to overcome the 100 ppm analogue upper output limit. The latter we already implemented, which now enables accurate quantification of high-concentration ebullition events. An automated bubble trap with CH₄ sensor was included in each floater of our system but failed to work reliably. Hence it needs further development and testing. Complementary measurements, such as vertical temperature profiling in the water column (Fig. S17), oxygen and pH measurements, will further support process-based interpretation by linking emission variability to stratification, mixing events, external drivers such as solar radiation and wind, and aquatic metabolism. Integrating such high-frequency driver data with high-frequency flux observations will facilitate machine learning-based gap filling, improve model calibration, and enhance predictive assessments of GHG dynamics under changing environmental conditions.



5. Conclusions

610 Automated and high-frequency monitoring systems significantly enhance the precision and (spatial) accuracy of aquatic CO₂ and CH₄ flux estimates while uncovering diel and sub-seasonal dynamics that conventional short-term methods fail to capture. Our system, BERTHA, exemplifies how high frequency greenhouse gas (GHG) flux measurements can be used to evaluate and optimise measurement strategies. In addition, the high-frequency measurements can support the development of correction factors for extrapolating infrequently measured fluxes to full 24-hour periods or even longer time scales. Analyses
615 of high frequency GHG measurements in conjunction with environmental variables can advance a more process-based understanding of aquatic GHG dynamics. Integrating automated chamber systems into aquatic GHG research, management, and national inventories will enhance the reliability of emission assessments across different temporal and spatial scales. Ultimately, this approach enables a more accurate representation of inland water carbon fluxes in regional and global GHG budgets and supports a more effective evaluation of mitigation strategies and ecosystem management practices.

620 Code and data availability

All underlying code and data used in this study are publicly accessible via <https://doi.org/10.17026/LS/HMOC5M>.

Author contributions

Resources and Funding acquisition: CF, RA and SK. Conceptualization: RA, JvdK, CF and SK. Methodology: RA, JvdK, FH, GW, DS-M, EV, AE, AdK. Investigation: RA and JvdK. Data curation: RA and JvdK. Formal analysis: RA and JvdK.
625 Software: RA, JvdK and DS-M. Validation: RA. Visualisation: RA and JvdK. Writing – Original Draft: RA and JvdK. Writing – Review & Editing: CF and SK. Supervision: SK, CF and RA. Project administration: CF, RA and SK.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgements

630 We thank B-WARE research centre for borrowing their Picarro G2508. We thank Natalia Rudik for assistance during the procedures for water vapour and pressure compensation of the sensors. We are grateful for the technical support of Coline Boonman during early testing of gas analysers and sensors. We thank Peter Cruijssen and Stefan Weideveld for technical assistance at the field site.



Financial support

635 This research was funded by the Netherlands Research Programme on Greenhouse gas Dynamics in Peatlands and organic Soils (NOBV). Sarian Kosten was supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek NWO-TTW project under [grant number 18661] and NWO-VIDI under [grant number 203.098]. Christian Fritz was supported by the European Union's Horizon Europe programme (PaluWise, GAP-101181479).

References

- 640 Aben, R. C., Velthuis, M., Kazanjian, G., Frenken, T., Peeters, E. T., Van de Waal, D. B., Hilt, S., de Senerpont Domis, L. N., Lamers, L. P., and Kosten, S.: Temperature response of aquatic greenhouse gas emissions differs between dominant plant types, *Water research*, 226, 119251, <https://doi.org/10.1016/j.watres.2022.119251>, 2022.
- Attermeyer, K., Casas-Ruiz, J. P., Fuss, T., Pastor, A., Cauvy-Fraunie, S., Sheath, D., Nydahl, A. C., Doretto, A., Portela, A. P., and Doyle, B. C.: Carbon dioxide fluxes increase from day to night across European streams, *Communications Earth & Environment*, 2, 118, <https://doi.org/10.1038/s43247-021-00192-w>, 2021.
- 645 Bastviken, D., Santoro, A. L., Marotta, H., Pinho, L. Q., Calheiros, D. F., Crill, P., and Enrich-Prast, A.: Methane emissions from Pantanal, South America, during the low water season: toward more comprehensive sampling, *Environmental science & technology*, 44, 5450-5455, <https://doi.org/10.1021/es1005048>, 2010.
- Bodmer, P., Vroom, R. J., Stepina, T., Del Giorgio, P. A., and Kosten, S.: Methane dynamics in vegetated habitats in inland waters: quantification, regulation, and global significance, *Frontiers in Water*, 5, 1332968, <https://doi.org/10.3389/frwa.2023.1332968>, 2024.
- 650 Castner, E. W., Jr.: Modern Spectroscopy, 4th Edition (J. Michael Hollas), *Journal of Chemical Education*, 82, 43-43, <https://doi.org/10.1021/ed082p43.1>, 2005.
- Christiansen, J. R., Korhonen, J. F., Juszczak, R., Giebels, M., and Pihlatie, M.: Assessing the effects of chamber placement, manual sampling and headspace mixing on CH₄ fluxes in a laboratory experiment, *Plant and soil*, 343, 171-185, <https://doi.org/10.1007/s11104-010-0701-y>, 2011.
- Dalbem Barbosa, A. P.: Nutrient fluxes in aquaculture systems: From feed to fish and the environment, Doctoral dissertation, Department of Ecology, Radboud University, <https://hdl.handle.net/2066/327467>, 2026.
- De Eyto, E., Smyth, R. L., Pilla, R. M., Laas, A., Shahabinia, A. R., Baldocchi, A., Desai, A. R., Lupon, A., Lohila, A., and 660 Obrador, B.: Diel variation in CO₂ flux is substantial in many lakes, *Limnology and Oceanography Letters*, <https://doi.org/10.1002/lol2.70066>, 2025.
- Deemer, B. R., Harrison, J. A., Li, S., Beaulieu, J. J., DelSontro, T., Barros, N., Bezerra-Neto, J. F., Powers, S. M., Dos Santos, M. A., and Vonk, J. A.: Greenhouse gas emissions from reservoir water surfaces: a new global synthesis, *BioScience*, 66, 949-964, <https://doi.org/10.1093/biosci/biw117>, 2016.



- 665 Fuss, R.: gasfluxes: Greenhouse Gas Flux Calculation from Chamber Measurements (0.7), CRAN [code],
 10.32614/CRAN.package.gasfluxes, 2024.
- Gaynullin, B., Hummelgård, C., Mattsson, C., Thungström, G., and Rödjegård, H.: Advanced pressure compensation in high
 accuracy NDIR sensors for environmental studies, *Sensors*, 23, 2872, <https://doi.org/10.3390/s23052872>, 2023.
- Golub, M., Koupaei-Abyazani, N., Vesala, T., Mammarella, I., Ojala, A., Bohrer, G., Weyhenmeyer, G. A., Blanken, P. D.,
 670 Eugster, W., and Koebisch, F.: Diel, seasonal, and inter-annual variation in carbon dioxide effluxes from lakes and reservoirs,
Environmental Research Letters, 18, 034046, <https://doi.org/10.1088/1748-9326/acb834>, 2023.
- Holgerson, M. A. and Raymond, P. A.: Large contribution to inland water CO₂ and CH₄ emissions from very small ponds,
Nature Geoscience, 9, 222–226, <https://doi.org/10.1038/ngeo2654>, 2016.
- Kuhn, M.: Building predictive models in R using the caret package, *Journal of statistical software*, 28, 1–26, <https://doi.org/10.18637/jss.v028.i05>, 2008.
 675 <https://doi.org/10.18637/jss.v028.i05>, 2008.
- Kulmala, L., Grönholm, T., Laakso, L., Keronen, P., Pumpanen, J., Vesala, T., and Hari, P.: Pressure responses of portable
 CO₂ concentration sensors, *Boreal Environment Research*, 14, 754, <https://doi.org/10.60910/k6dq-2dvz>, 2009.
- Lin, Y., Manalili, D., Khodabakhsh, A., and Cristescu, S. M.: Real-Time Measurement of CH₄ in Human Breath Using a
 Compact CH₄/CO₂ Sensor, *Sensors*, 24, 1077, <https://doi.org/10.3390/s24041077>, 2024.
- 680 Liu, H., Zhang, Q., Katul, G. G., Cole, J. J., Chapin, F. S., and MacIntyre, S.: Large CO₂ effluxes at night and during
 synoptic weather events significantly contribute to CO₂ emissions from a reservoir, *Environmental Research Letters*, 11,
 064001, <https://doi.org/10.1088/1748-9326/11/6/064001>, 2016.
- Maier, M., Weber, T. K., Fiedler, J., Fuss, R., Glatzel, S., Huth, V., Jordan, S., Jurasinski, G., Kutzbach, L., and Schäfer, K.:
 Introduction of a guideline for measurements of greenhouse gas fluxes from soils using non-steady-state chambers, *Journal*
 685 *of Plant Nutrition and Soil Science*, 185, 447–461, <https://doi.org/10.1002/jpln.202200199>, 2022.
- Matthias, A., Blackmer, A., and Bremner, J.: A simple chamber technique for field measurement of emissions of nitrous
 oxide from soils, *Wiley Online Library*0047–2425, <https://doi.org/10.2134/jeq1980.00472425000900020017x>, 1980.
- Miller, A. W., Muirhead, J. R., Reynolds, A. C., Minton, M. S., and Klug, K. J.: High-frequency continuous measurements
 reveal strong diel and seasonal cycling of pCO₂ and CO₂ flux in a mesohaline reach of the Chesapeake Bay,
 690 *Biogeosciences*, 21, 3717–3734, <https://doi.org/10.5194/bg-21-3717-2024>, 2024.
- Mwanake, R. M., Gettel, G. M., Wangari, E. G., Glaser, C., Houska, T., Breuer, L., Butterbach-Bahl, K., and Kiese, R.:
 Anthropogenic activities significantly increase annual greenhouse gas (GHG) fluxes from temperate headwater streams in
 Germany, *Biogeosciences*, 20, 3395–3422, <https://doi.org/10.5194/bg-20-3395-2023>, 2023.
- Paranaíba, J. R., Struik, Q., Shendurnikar, S., Ma, Y., Quadra, G. R., and Kosten, S.: Summer CH₄ ebullition strongly
 695 determines year-round greenhouse gas emissions from agricultural ditches despite frequent dredging, *Journal of*
Environmental Management, 373, 123813, <https://doi.org/10.1016/j.jenvman.2024.123813>, 2025.
- Peacock, M., Audet, J., Jordan, S., Smeds, J., and Wallin, M. B.: Greenhouse gas emissions from urban ponds are driven by
 nutrient status and hydrology, *Ecosphere*, 10, e02643, <https://doi.org/10.1002/ecs2.2643>, 2019.



- Peacock, M., Audet, J., Bastviken, D., Cook, S., Evans, C., Grinham, A., Holgerson, M., Högbom, L., Pickard, A., and
 700 Zieliński, P.: Small artificial waterbodies are widespread and persistent emitters of methane and carbon dioxide, *Global
 Change Biology*, 27, 5109-5123, <https://doi.org/10.1111/gcb.15762>, 2021a.
- Peacock, M., Audet, J., Bastviken, D., Futter, M. N., Gauci, V., Grinham, A., Harrison, J. A., Kent, M. S., Kosten, S.,
 Lovelock, C. E., Veraart, A., and Evans, C.: Global importance of methane emissions from drainage ditches and canals,
Environmental Research Letters, 16, 044010, <https://doi.org/10.1088/1748-9326/abeb36>, 2021b.
- 705 Piatka, D. R., Nánási, R. L., Mwanake, R. M., Engelsberger, F., Willibald, G., Neidl, F., and Kiese, R.: Precipitation fuels
 dissolved greenhouse gas (CO₂, CH₄, N₂O) dynamics in a peatland-dominated headwater stream: results from a continuous
 monitoring setup, *Frontiers in Water*, 5, 1321137, <https://doi.org/10.3389/frwa.2023.1321137>, 2024.
- Podgrajsek, E., Sahlée, E., and Rutgersson, A.: Diurnal cycle of lake methane flux, *Journal of Geophysical Research:*
Biogeosciences, 119, 236-248, <https://doi.org/10.1002/2013JG002327>, 2014.
- 710 Podgrajsek, E., Sahlée, E., and Rutgersson, A.: Diel cycle of lake-air CO₂ flux from a shallow lake and the impact of
 waterside convection on the transfer velocity, *Journal of Geophysical Research: Biogeosciences*, 120, 29-38,
<https://doi.org/10.1002/2014JG002781>, 2015.
- Pumpanen, J., Longdoz, B., and Kutsch, W. L.: Field measurements of soil respiration: principles and constraints, potentials
 and limitations of different methods, 978-0-521-86561-6, <https://doi.org/10.1017/CBO9780511711794.003>, 2010.
- 715 Raymond, P. A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E.,
 and Humborg, C.: Global carbon dioxide emissions from inland waters, *Nature*, 503, 355-359,
<https://doi.org/10.1038/nature13142>, 2013.
- Reed, D. E., Dugan, H. A., Flannery, A. L., and Desai, A. R.: Carbon sink and source dynamics of a eutrophic deep lake
 using multiple flux observations over multiple years, *Limnology and Oceanography Letters*, 3, 285-292,
 720 <https://doi.org/10.1002/lol2.10075>, 2018.
- Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., Karion, A., Miles, N. L., Richardson, S. J., and
 Steinbacher, M.: High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air,
Atmospheric Measurement Techniques, 6, 837-860, <https://doi.org/10.5194/amt-6-837-2013>, 2013.
- Rochette, P. and Hutchinson, G. L.: Measurement of soil respiration in situ: chamber techniques, *Micrometeorology in*
 725 *agricultural systems*, 47, 247-286, <https://doi.org/10.2134/agronmonogr47.c12>, 2005.
- Sharma, N., Dadhwal, V., Kant, Y., Mahesh, P., Mallikarjun, K., Gadavi, H., Sharma, A., and Ali, M.: Atmospheric CO₂
 variations in two contrasting environmental sites over India, *Air, Soil and Water Research*, 7,
<https://doi.org/10.4137/ASWR.S13987>, 2014.
- Shen, C., Qian, M., Song, Y., Chen, B., and Yang, J.: Surface pCO₂ and air-water CO₂ fluxes dominated by submerged
 730 aquatic vegetation: Implications for carbon flux in shallow lakes, *Journal of Environmental Management*, 370, 122839,
<https://doi.org/10.1016/j.jenvman.2024.122839>, 2024.



- Shusterman, A. A., Teige, V. E., Turner, A. J., Newman, C., Kim, J., and Cohen, R. C.: The Berkeley atmospheric co₂ observation network: Initial evaluation, *Atmospheric Chemistry and Physics*, 16, 13449-13463, <https://doi.org/10.5194/acp-16-13449-2016>, 2016.
- 735 Sieczko, A. K., Duc, N. T., Schenk, J., Pajala, G., Rudberg, D., Sawakuchi, H. O., and Bastviken, D.: Diel variability of methane emissions from lakes, *Proceedings of the National Academy of Sciences*, 117, 21488-21494, <https://doi.org/10.1073/pnas.2006024117>, 2020.
- Silverthorn, T., Audet, J., Evans, C. D., van Der Knaap, J., Kosten, S., Paranaíba, J., Struik, Q., Webb, J., Wu, W., and Yan, Z.: The importance of ditches and canals in global inland water CO₂ and N₂O budgets, *Global change biology*, 31, e70079, <https://doi.org/10.1111/gcb.70079>, 2025.
- 740 Theus, M. E., Ray, N. E., Bansal, S., and Holgerson, M. A.: Submersed macrophyte density regulates aquatic greenhouse gas emissions, *Journal of Geophysical Research: Biogeosciences*, 128, e2023JG007758, <https://doi.org/10.1029/2023JG007758>, 2023.
- Van der Knaap, J., Harpenslager, S. F., Aben, R. C., Weideveld, S. T., van Giersbergen, Q., van Dijk, G., Wintjen, P., Buzacott, A. J., Fritz, C., and Kruijt, B.: Disproportionately High Contribution of Ditches to Landscape Greenhouse Gas Emissions in Drained Peatlands, *Ecosystems*, 28, 1-16, <https://doi.org/10.1007/s10021-025-01005-3>, 2025.
- 745 White, S. A., Santos, I. R., Conrad, S. R., Sanders, C. J., and Hessey, S.: Large aquatic nitrous oxide emissions downstream of intensive horticulture driven by rain events, *Journal of Hydrology*, 596, 126066, <https://doi.org/10.1016/j.jhydrol.2021.126066>, 2021.
- 750 Wik, M., Thornton, B. F., Bastviken, D., Uhlbäck, J., and Crill, P. M.: Biased sampling of methane release from northern lakes: A problem for extrapolation, *Geophysical Research Letters*, 43, 1256-1262, <https://doi.org/10.1002/2015GL066501>, 2016.
- Wright, M. N. and Ziegler, A.: ranger: A fast implementation of random forests for high dimensional data in C++ and R, *Journal of statistical software*, 77, 1-17, <https://doi.org/10.18637/jss.v077.i01>, 2017.
- 755 Yan, Y.: MLmetrics: Machine Learning Evaluation Metrics, CRAN [code], 10.32614/CRAN.package.MLmetrics, 2024.