POLYTECH TOURS

université de TOURS

Rapport de stage individuel

4ème année

Investigating greenhouse gas emissions in pond ecological habitats

Utilizing a low-cost environmental data collection method

Aarhus University

Nordre Ringgade 1, 8000 Aarhus C,
DENMARK

AARHUS UNIVERSITY
DEPARTMENT OF ECOSCIENCE

Company supervisors:

Joachim Audet
Tuba Bucak
Thomas Davidson

Academic supervisor:

Sabine Greulich

Manon Dijoux

IMA

2023-2024

Acknowledgements

Before starting this report, I would like to thank all the people and organizations who made this internship possible.

First, I want to thank Aarhus University, and especially my professional supervisors, Mr. Joachim Audet, Ms. Tuba Bucak, and Mr. Thomas Davidson, whose trust and support gave me the opportunity to participate in this research project. I am also grateful to the entire Freshwater section, especially Mr. Marthin Pedersen, Ms. Eti Levi, and Mr. Dennis Hansen, with whom I had the pleasure of working directly for several weeks. Their kindness, support throughout these three months, and valuable advice were crucial for acquiring new skills, exploring the field of aquatic environmental research.

Finally, I wish to express my appreciation to Ms. Sabine Greulich, my academic supervisor, for her guidance throughout this project. I also thank Polytech Tours for giving me the opportunity to undertake this internship, which has been not only a unique experience but also an opportunity to meet many interesting people.

Use of AI

For transparency purposes it should be noted that ChatGPT was used in this report. It was used to help create R code and check the correct translation of the report into English.

Table of contents

I) Introduction 4
1) Presentation of the hosting structure 4
2) Context and objectives of the study project 4
a) Greenhouse gases and fresh water bodies 5
b) Main greenhouse gas production processes in fresh water 5
c) Focus on ponds and the challenges of their sampling 7
3) Presentation of my tasks and objectives within the project scope 7
II) Material and methods 8
1) Description and operation of the data collection chamber 8
a) Presentation of the automated flushing chamber 8
b) Calibration of the low-cost sensors for CH4 detection 12
c) Automation of CH_4 sensor calibration through the development of a R script for peak detection 14
2) Field measurements 15
a) Study sites 15
b) Field deployment of chambers 18
c) Data analysis for flux calculations using R 18
d) Statistical analysis using t-test to ensure significance of observed differences 19
III) Results 20
1) Calibration results 20
2) Greenhouse gas concentration over time 22
IV) Discussion and reflective feedback 27
1) Analysis of produced results 27
2) Suggestions for changes, improvements, continuation 29
3) Career projection (research) 31

I) Introduction

As part of my studies in Planning and Environmental Engineering, specializing in Aquatic Environment Engineering, I completed an internship to validate my fourth year of training. This internship took place over a period of 12 weeks, from 01/05/2024 to 23/07/2024, in Denmark, at Aarhus University, where I contributed to research on monitoring greenhouse gas concentrations emitted by water bodies.

The purpose of this report is to present and document my tasks and achievements during this internship, as well as the results obtained and the lessons learned from this experience. In the following sections of this introduction, I will outline the context in which my internship took place. Subsequently, in the main body of the report, I will detail the various tasks and missions I undertook, along with the results they generated. Finally, I will conclude by sharing the insights I gained from this experience.

1) Presentation of the hosting structure

I did my internship at Aarhus University (AU), Aarhus, Denmark. Aarhus is the second-largest city in Denmark, after Copenhagen, and AU is similarly the second-largest university in the country. In addition to offering a wide range of courses in numerous fields with an international influence, the university has a highly developed research activity covering various domains, with over 100 research centers distributed across its 5 faculties (Aarhus University website, n.d.).

I was assigned to the Department of Ecoscience, which is part of the Faculty of Technical Sciences. Specifically, I worked at the department's unit located on the AU campus ; there is also a second unit located in the city of Roskilde, with approximately 275 employees. The Department of Ecoscience specializes in teaching, research, and providing consultancy in the field of life in all its forms, covering topics such as microbiology or zoology, from the study of genes to ecosystems. Their research spans from fundamental research to applied biology in nature management and biotechnology (Aarhus University Ecoscience, 2024).

The Ecoscience unit on the AU campus is also divided into several sections. I worked in the research section dedicated to freshwater ecology, which covers streams, lakes, ponds, wetlands, and their surrounding areas, whether natural or artificial. This section conducts long-term studies, measurements, and monitoring of biological and chemical processes in freshwater ecosystems, as well as their interactions with nutrients and environmentally hazardous substances. They work on topics such as lake and stream restoration, nutrient dynamics, greenhouse gases, biodiversity, and the interaction between watersheds and freshwater areas (Aarhus University Ecoscience, 2024).

One of their main objectives is notably the development and use of new modeling, calculation, and monitoring tools. It is within the context of this objective, specifically, that the section focusing on lakes issued the internship project assigned to me, which I will now present.

2) Context and objectives of the study project

a) Greenhouse gases and fresh water bodies

The EcoScience Department at AU has been focusing on the subject of greenhouse gases (GHG) and conducting research in this field for nearly twenty years now. This particular attention to this subject is due to the fact that climate change represents a major challenge of the 21st century, requiring a deep understanding of its mechanisms. The greenhouse effect, resulting from the accumulation of certain gases in the atmosphere, notably carbon dioxide (CO_2) the most emitted one, and methane (CH_4) 34-times more potent than CO_2, is indeed one of the main drivers of this phenomenon, by absorbing and re-emitting energy in the low atmosphere (Li et al., 2021).

It has been shown that freshwater bodies have a significant impact on the greenhouse effect due to their production of GHG, the sum of which approaches 31% of the annual CO_2 emissions from fossil fuel combustion (Li et al., 2021). These observations partly stem from one of the longest long-term mesocosm experiments on a shallow lake, where GHG production was monitored over several years under various temperature and nutrient conditions. This experiment revealed the influence of interactions between various biotic factors (macrophytes, microbial metabolic processes, etc.) and abiotic factors (temperature, nutrients, etc.) on the production of GHG in these freshwater environments (Davidson et al., 2015). These findings highlight the complexity of the mechanisms at play and the importance of studying diverse ecological habitats and environments to understand how they function, thereby ensuring the implementation of appropriate and effective measures against these environmental issues.

b) Main greenhouse gas production processes in fresh water

In their publication published in 2021, Y. Li et al. provide a synthesis of the main phenomena and conditions explaining the significant production and emission of GHGs in freshwater bodies, summarized in figure 1. They particularly highlight eutrophication as a major driver of this production.

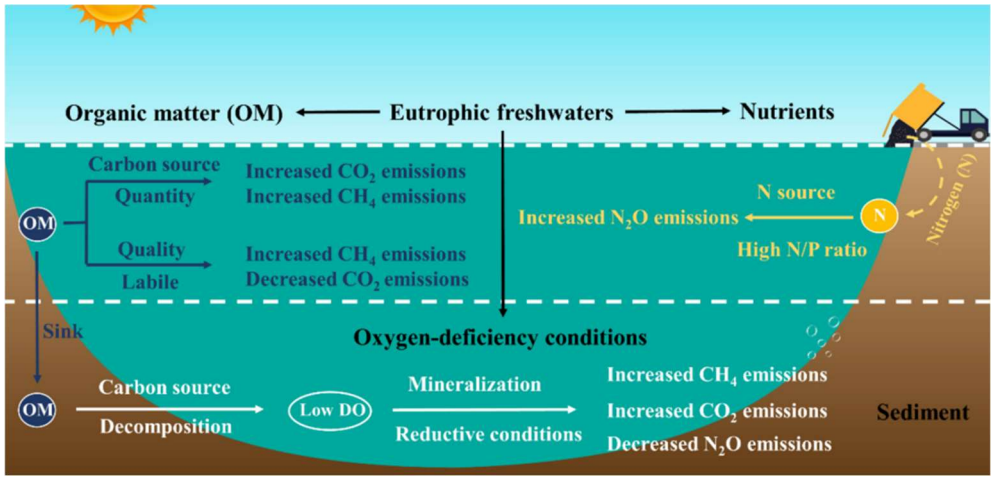


Fig. 1: Effects of freshwater eutrophication on Greenhouse gas emissions (Li et al., 2021)

Eutrophication, resulting from an excessive input of nutrients such as phosphorus (P) and nitrogen (N), leads to an overproduction of organic matter. As this organic matter is decomposed by bacteria, oxygen is consumed, causing low concentrations of dissolved oxygen (O_2) in bottom waters and the interstitial waters of sediments. These anaerobic conditions favor the

mineralization of organic matter (carbon) by bacteria, releasing other gases, primarily CH_4 and CO_2, which are GHGs that will subsequently be emitted into the atmosphere (Li et al., 2021).

For CH_4, there are several mechanisms by which it can migrate to the surface. After its production by archaea, CH_4 accumulates at the bottom, creating a concentration gradient that pushes it to rise to the surface and escape continuously into the atmosphere, a process known as diffusion. Another emission mechanism occurs when the gas accumulated in the sediments reaches saturation with a pressure higher than the hydrostatic pressure, leading to the formation of bubbles that periodically rise to the surface. Finally, methane can also reach the surface through plants that absorb it through their roots and release it through stomata, structures on their leaves that allow gas exchange (Bastviken et al., 2020).

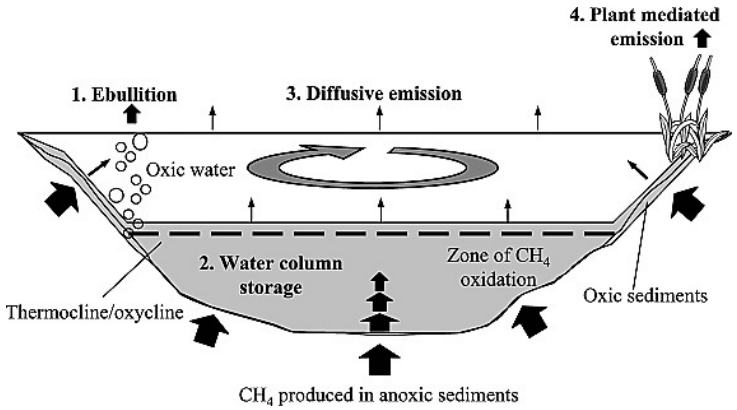


Fig. 2: Illustration of emission pathways and CH_4 dynamics in stratified lake (Bastviken et al., 2020)

Regarding CO_2, two divergent phenomena occur. On one hand, its production is observed through the previously mentioned mineralization; on the other hand, eutrophication leads to increased photosynthesis, which consumes CO_2 and thus reduces its concentration (Li et al., 2021). This example of CO_2 behavior illustrates the complexity of the phenomena related to GHG production in natural environments. Many other indirect mechanisms exist, and to this day, several unknowns persist, justifying the numerous ongoing studies on this subject.

For example, in a recent study researchers examined the impacts of thermal lake stratification on GHG emissions (Davidson et al., 2024). Stratification, characterized by the formation of thermally distinct water layers during warm periods, results in poor mixing between surface waters, which are in contact with air and oxygen, and bottom waters, where the chemical processes producing these gases occur in anoxic conditions. Additionally, a significant release of CH_4 and CO_2 is also observed at the end of the stratification (turnover), when mixing and water renewal resume, thereby releasing gases trapped in the lower layers and sediments (Davidson et al., 2024). These results highlight the indirect but significant impact of temperature on greenhouse gas emissions, underscoring the potential for climate warming to exacerbate these emissions. Indeed, the increase in greenhouse gas emissions tends to intensify the greenhouse effect and climate change, which in turn amplify these emissions.

c) Focus on ponds and the challenges of their sampling

It has been estimated that, despite their combined area representing only about 9% of all lentic freshwater bodies, ponds are responsible for 15% of all CO_2 produced and 40% of all diffusive CH_4, making them major sources of GHGs (Holgerson & Raymond, 2016). Small urban and agricultural ponds possess all the conditions favorable for the overproduction of GHGs: being shallow water bodies, they experience intense and frequent thermal stratifications due to their size, and they are located near activity zones where ambient temperatures are generally higher. The urban and agricultural environments in which they are found also make them more prone to eutrophication due to increased inputs of pollutants, nutrients, and various particles, which promotes plant and phytoplankton production, further stimulated by higher temperatures (Audet et al., 2020).

As we have observed, studies on the relationship between GHGs and natural water bodies are necessary, but they are also complex. For a reliable estimation of GHG emissions, spatially and temporally representative measurements are essential, which is achievable but expensive with current methods. For example, some phenomena, such as ebullition, are not continuous and can occur at any time, making their quantification difficult. Constant presence at the study site would be necessary to obtain representative measurements, but this is not feasible with current methods that include very expensive equipment (such as LGR-ICOS Ultraportable gaz analyzers and expensive precalibrated automated flushing chambers) (Davidson et al., 2024).

To address this issue, Aarhus University researchers have developed a new monitoring method, namely low-cost automated flushing GHG chambers that can be left continuously on water bodies. However, this approach involves significant technical and financial challenges. A device left unsupervised for an extended period, and on water, is really vulnerable and must be practical to deploy, durable, and inexpensive to minimize potential losses due to damage.

3) Presentation of my tasks and objectives within the project scope

My internship aimed to test this new low-cost automated flushing GHG chamber. During the internship, I contributed to the construction of the new chambers and calibration of the CH_4 sensors. To ensure their proper functioning and operational readiness, the second part of my internship involved deploying and testing them in real-world conditions on ponds. The objective was to analyze the results recorded by the chambers regarding the changes of GHG concentrations emitted in different environments and to compare them with the current knowledge on the dynamics of GHG production and emission in freshwater bodies, as presented earlier in the report.

Additionally, I had the opportunity to participate in a secondary mission lasting two and a half weeks as part of another project named Transponder (Transnational Biodiversity and Ecosystem Assessment Approaches for Ponds in Europe), an international project aiming to develop a standardized methodology for monitoring pond biodiversity (METU, 2024). I accompanied a research team during their sampling campaign on 30 ponds near the town of Laven and on the island of Avernakø in the south of the country. I assisted in carrying out various samplings, including eDNA, zooplankton, and physico-chemical parameters. This mission provided the opportunity to deploy automatic chambers on some of the ponds identified by the project and to gather environmental information such as water temperature, chlorophyll *a* and cyanobacteria

concentrations as well as dissolved oxygen, which are useful for interpreting the data obtained with the chambers.

In this report, we will focus on studying the impact of habitat on GHG production and utilize the first results obtained with the chambers to see if there are significant differences between various habitat types and environmental factors that particularly influence the recorded GHG concentrations in different water bodies.

The objectives of this internship are:

- The construction and calibration of operational data collection chambers.
- Using the initial results obtained with the chambers to determine if there are significant differences in GHG production between different habitat types within and between the ponds.
- Analyzing the environmental factors that particularly influence the GHG concentrations recorded in various water bodies.

We hypothesize that GHG emissions vary depending on the presence of vegetation, with vegetation potentially resulting in lower emissions.

II) Material and methods

1) Description and operation of the data collection chamber

a) Presentation of the automated flushing chamber

Components of the collection chamber

To estimate GHG fluxes, a floating chamber is used. The floating chamber is a data collection equipment that functions by being placed hermetically on the water's surface, capturing gases emitted by the water body and released into the air, which are then recorded by sensors. This device had already been used by the research team I worked with to monitor GHG emissions in the shallow lake mesocosm study. In this study, floating chambers, that can be seen in figure 3, are placed in experimental tanks, connected to a rope that allowed the chamber to be lifted at regular intervals to release the trapped gas, and connected to a fixed electrical cabinet.

Fig. 3: Example of a data collection chamber used in the mesocosm study

For natural environment studies, it was necessary to modify these chambers to make them more suitable. Specifically, it was essential to make them portable and autonomous. The new version of the collection chamber, which I helped construct, is presented in Figure 4.

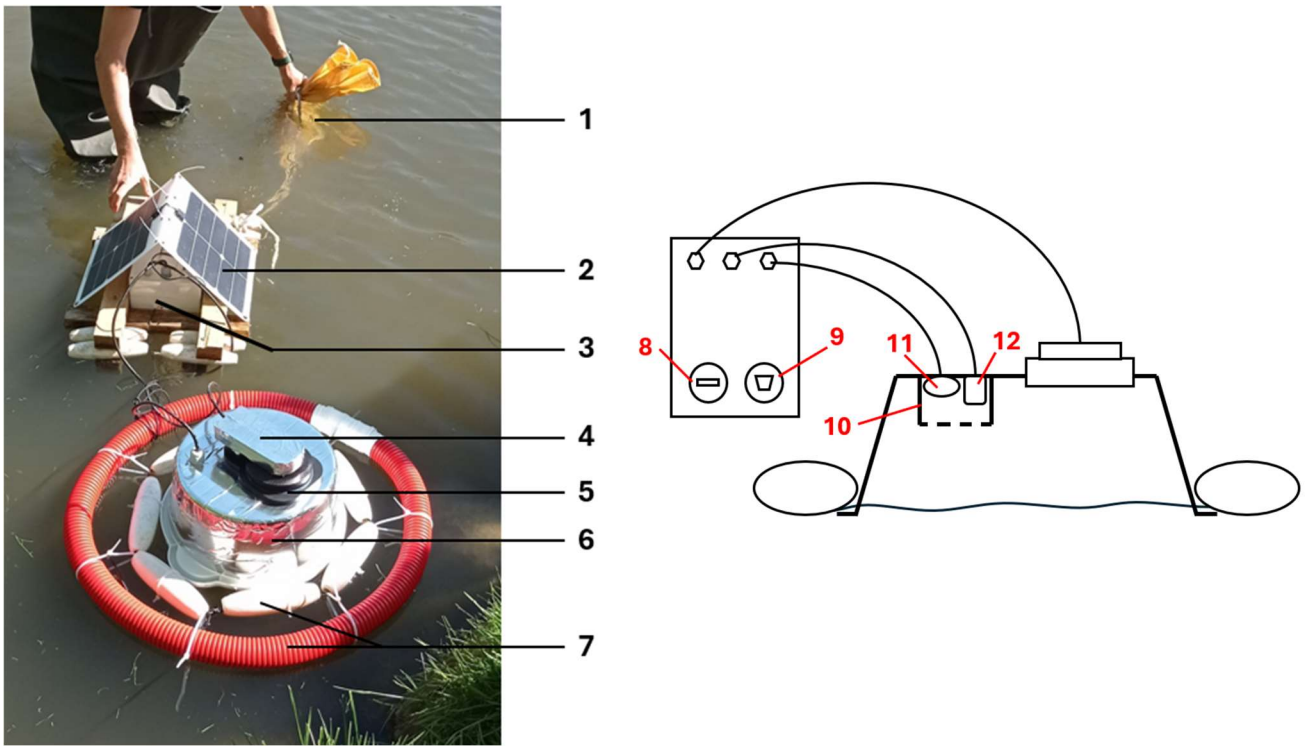


Fig. 4: Different compounds of the automated flushing chambers. 1: anchor ; 2: solar panels ; 3: control box ; 4: motor ; 5: fan ; 6: chamber ; 7: floaters ; 8: SD port ; 9: on and off button ; 10: perforated box ; 11: CH_4 sensor ; 12: CO2, temperature and relative humidity sensor.

This version of the collection chamber consists of several elements. The chamber (**6**) itself is an 8L plastic basin, covered with aluminum tape to reflect light and limit temperature variations inside the chamber. The basin is surrounded by two rings of floaters (**7**) to ensure its stability and prevent it from sinking. A device with a flap and a fan (**5**) is also installed at the top of the basin, controled by a motor (**4**) to allow the trapping or release of gas. Inside this chamber, there are two sensors: one for CH_4 (**11**) and one for CO_2, temperature and relative humidity (**12**), all protected from potential splashes by a perforated box (**10**).

The sensors and the motor are connected by cables to the second part of the chamber, which consists of a wooden raft supporting the control box (**3**). This box houses and protects all electronic components, including two 9V battery packs that power the system, an SD card module for storing sensor data and error logs, a real-time clock (RTC_DS3231) for timekeeping, an analog-to-digital converter (ADC - Adafruit_ADS1115) for reading and amplifying analog CH_4 sensor signals, a CO_2 sensor plug (SCD30), and an Arduino MEGA 2560 microcontroller board. It has a button (**9**) on its cover to turn the chamber on and off (start recording sensor data and automatic flushing cycles) as well as an SD port (**8**) where the memory card is inserted to save the data.

Solar panels (**2**) are placed above the box to power the device, making it autonomous, while also protecting the box from rain. Finally, anchors (**1**) are placed at two opposite ends of the wooden raft to secure the chamber and limit its drift caused by the wind.

Operation of the collection chamber

The operation of the chamber is controlled by an Arduino Mega 2560 program developed by the research team, which is initiated and reset each time the chamber is turned on. This program continuously cycles through four phases as shown in figure 5: a data acquisition phase of 3 hours and 54 minutes during which the chamber is hermetically sealed to allow gas accumulation inside, followed by a flap opening phase, then the activation of the fan for 175 seconds (approximately 3 minutes) to facilitate air renewal inside and get back to background atmospheric concentration level. Finally, the flap is closed again.

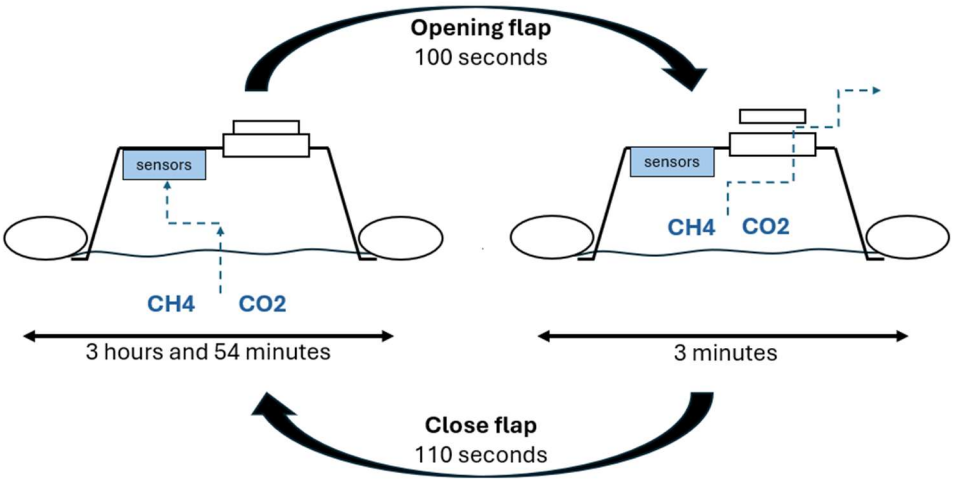


Fig. 5: Chamber operation cycle

Use of low-cost CH_4 sensor

One of the design requirements for the chamber was that it be low-cost, which presented certain technical challenges. Not only are CH_4 sensors expensive, but the entire measurement system as well. Its use would have been complicated in our situation, where multiple chambers need to be produced and left unattended on the water. Therefore, it was necessary to find an alternative for measuring CH_4 concentration. The chosen solution was to use a low-cost CH_4 sensor with a TGS sensor (Tin Oxide Gas Sensor), designed to detect the presence of gas in the air, following the technical note developed by Bastviken et al (Bastviken et al., 2020). The total material cost of the system is around 5000-6000 DKK (670-804 EUR), excluding labor costs, with the CO_2 sensor costing around 400-500 DKK (54-67 EUR) and the low-cost CH_4 sensor around 100-200 DKK (13-27 EUR).

The TGS sensor operates on the principle of gas absorption and desorption at its surface. As illustrated in figure 6, it consists of a reference resistor R_S with a voltage V_L and a detection zone R_0 with a voltage V_C, all connected in a circuit with a total voltage of 5V. When a gas comes into contact with the detection zone, its concentration determines how many electrons it captures from the system, thereby modifying the voltage V_C. A decrease in V_C will reduce the resistance R_0, resulting in an increase in voltage V_L and resistance R_S, according to the equation

$R_s = \left(\frac{V_C}{V_L} - 1\right) * R_0$. This will allow us to know the ratio $\frac{R_S}{R_0}$ associated with the gas concentration present in contact with the sensor.

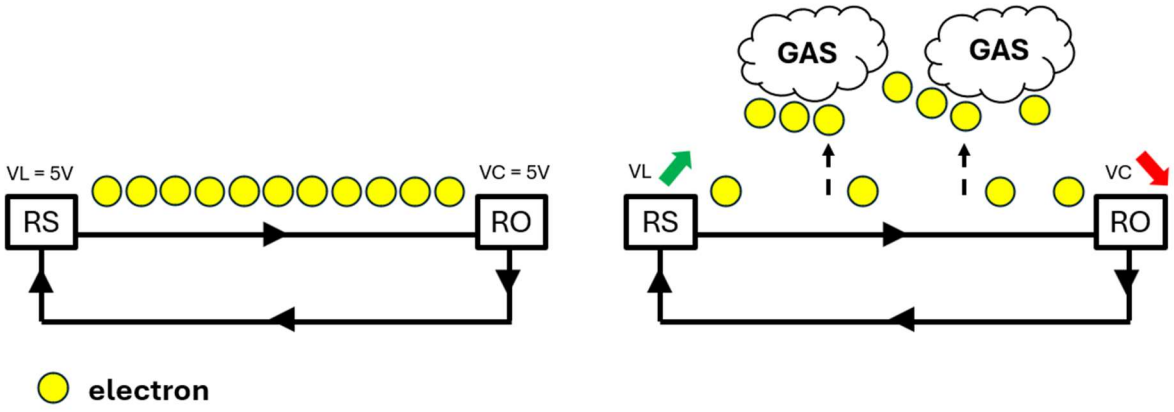


Fig. 6: Simplified diagram of the operation of the Tin Oxide Gas Sensor (TGS). RS: reference resistor ; RO: detection zone ; VL: reference voltage ; VC: detection voltage

A model developed by Bastviken et al. (2020) will be used to convert the ratio $\frac{R_S}{R_0}$ recorded by the sensor into CH_4 concentration in ppm based on a nonlinear relationship presented later. However, a technical challenge remains: the TGS sensor reacts not only to CH_4 but also to other gases, particularly water vapor and the high humidity that accumulates in the chamber placed on the water (Bastviken et al., 2020). Therefore, calibration is necessary to quantify the impact of these other gases and isolate CH_4.

The curves obtained during calibration, which show the increase in voltage related to humidity without CH_4, can be found in appendix 1.

b) Calibration of the low-cost sensors for CH4 detection

The calibration process of the CH_4 sensor is used to determine the relative response of the sensor without CH_4, considering the influence of water vapor. This will allow us to obtain accurate measurements of CH_4 concentration while accounting for variations in humidity. We will use a series of formulas and measurements according to Bastviken et al. (2020).

Theoretical principle

To calculate the relative response of the sensor to CH_4 $\left(\frac{RS}{RO}\right)$ without the influence of water vapor, we will use formula which incorporates all the parameters influencing the ratio, including V_0.

$$\frac{R_S}{R_0} = \frac{\left(\frac{V_C}{V_L} - 1\right)}{\left(\frac{V_C}{V_0} - 1\right)}$$

V_0 correspond to the baseline voltage for a given humidity level according to the next equation (Bastviken et al., 2020). We determine V_0 by measuring the voltage in the absence of CH_4 injections (baseline CH_4 concentrations). Then we will conduct multiple measurements of voltage and reference CH_4 concentration at different humidity levels to establish a relationship between the sensor voltage and humidity under different reference measurements.

$$V_0 Q = g * H_2O(ppm) + P$$

Using the last formula determined by Bastviken et al. (2020), we will then obtain the calibration coefficients a, b, c, and K through nonlinear least squares estimation. These parameters will allow us to obtain the adjusted CH_4 concentration for future recordings, taking humidity into account.

$$CH_4 = a * \left(\frac{R_S}{R_0}\right)^b + c * H_2Oppm * \left(\frac{R_S}{R_0}\right)^b + K$$

Application

To obtain the V_0 parameter necessary for calculating the ratio $(\frac{R_S}{R_0})$ and thereby the coefficients, it is essential to perform a series of measurements where the sensor is exposed to different concentrations of CH_4 under varying humidity and temperature conditions. These measurements establish a relationship between the sensor's voltage and humidity.

In practical terms, the calibration involves placing a floating chamber containing the sensors to be calibrated, along with a LGR Ultraportable sensor Greenhouse Gas Analyze (Los Gatos Research (ABB, n.d.)) capable of detecting CH_4 as well as CO_2 and H_2O. Inside this chamber, which will first be at room temperature (20-24°C) and then in a cold room temperature (4°C), known quantities of CH_4 will be regularly injected. Ideally, 20 to 30 measurements ranging from 2 to 200 ppm at different temperatures should be conducted. The recorded results will be analyzed using an R script developed by Ms. Bucak Onay, which automatically provides the calibration coefficients a, b, c, and K, as well as control graphs that will be presented in the Result section. The more detailed steps of this calibration process are available in appendix 2.

It is important to note that each sensor has unique characteristics and resistances, requiring individual calibration. Thus, each sensor will have unique calibration coefficients. It is also important to regularly check baseline readings because bad connections are common and can distort the calibration requiring the operation to be restarted from the beginning. Figure 7 illustrates the arrangement of the setup calibration.

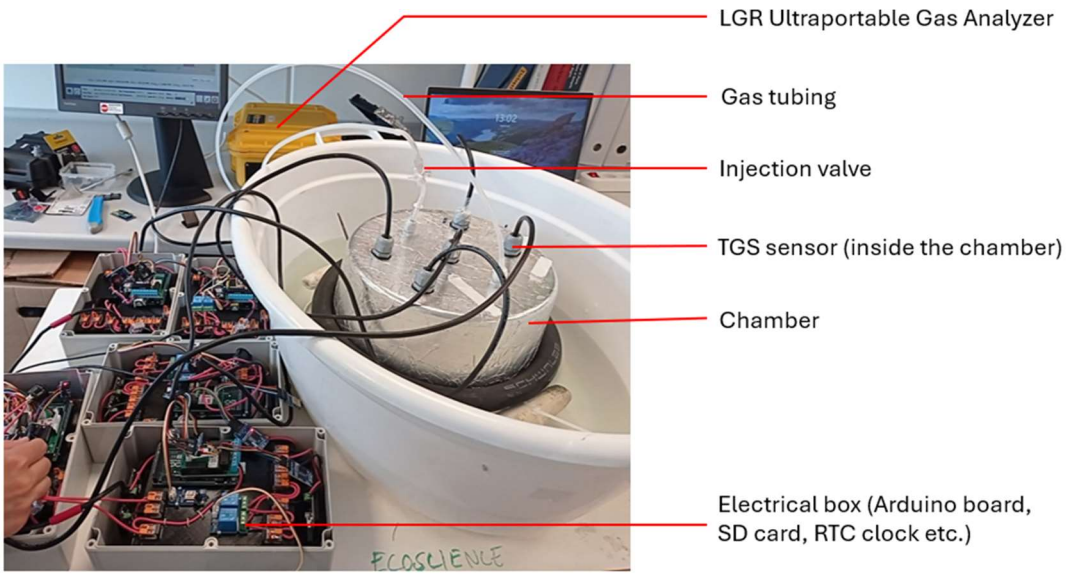


Fig. 7: Calibration setup

During this internship, I calibrated 12 sensors, excluding those that encountered errors during the process. Among the 5 sensors that could be calibrated simultaneously, there was always at least one sensor that was not functional at the end. This calibration task took me about a month.

Verification of calibration quality

To evaluate the quality of the calibration, two key statistical indicators were used: RMSE (Root Mean Square Error) and the coefficient of determination (R^2):

- The R^2 measures the proportion of the variance in the observed data explained by the model. An R^2 close to 1 indicates that the model explains the variability of the data well.

- The RMSE measures the average error between the predicted and observed values. A lower RMSE indicates greater precision in the model's predictions.

Good calibration is indicated by a low RMSE and a high R^2, meaning the model's predictions are accurate and the model captures the variance in the observed data effectively.

c) Automation of CH_4 sensor calibration through the development of a R script for peak detection

It was previously mentioned that after each calibration, it was necessary to sort the measurements recorded by the LGR eqipment to isolate those corresponding to the times of the CH_4 injections. This sorting is done manually and involves sifting through numerous lines of data for each sensor, as the device records a measurement approximately every two seconds, while the system often runs all day.

So during this internship, I attempted to automate this task by developing an R script capable of detecting peaks and providing the associated value list. A secondary objective of this task was also to learn and familiarize myself with this software for the first time.

During the development, I created three versions of the script using three different methods. I will present here the final version of the script that was selected.

The objective of this script is to find a way for the program to identify the presence of a peak in CH_4 addition, meaning a significant increase in the values recorded by the sensor. The main challenge is to define what constitutes a significant increase with the appropriate sensitivity threshold. Indeed, if the sensitivity is too low, smaller peaks corresponding to low concentrations of injected CH_4 may go undetected ; on the contrary, if the sensitivity is too high, even minor noise in the recording may be identified as a peak. The initial versions of the script did not succeed due to these difficulties, which made them neither precise nor practical to use.

For the final version of the script, I concluded that it was simpler, faster and more reliable to continue to rely on human detection of peaks. Instead of trying to completely automate peak detection, I automated and facilitated the necessary manipulations so that a user could do it more efficiently.

I automated the steps using the Shiny package, which allows you to build an interactive interface in R. On this interface, it is possible to load your raw dataset, which displays the graph of values and makes it possible to better visually detect peaks. Clicking on each point on the graph displays the associated date, time and concentration. “Previous Point” and “Next Point” buttons allow you to navigate in the graph. Once the desired point is selected, just press the "Confirm Point" button to save it. Then the point appear in the console to confirm that it was saved and allow you to move on to the next one.

Finally, by clicking on "Download Data", you obtain a CSV file with the list of chosen points, directly formatted as required for further data processing. The points are sorted to appear in chronological order in the graph and assigned a rank based on this order, according to the requested format. An illustration of this interactive interface can be seen in figure 8.

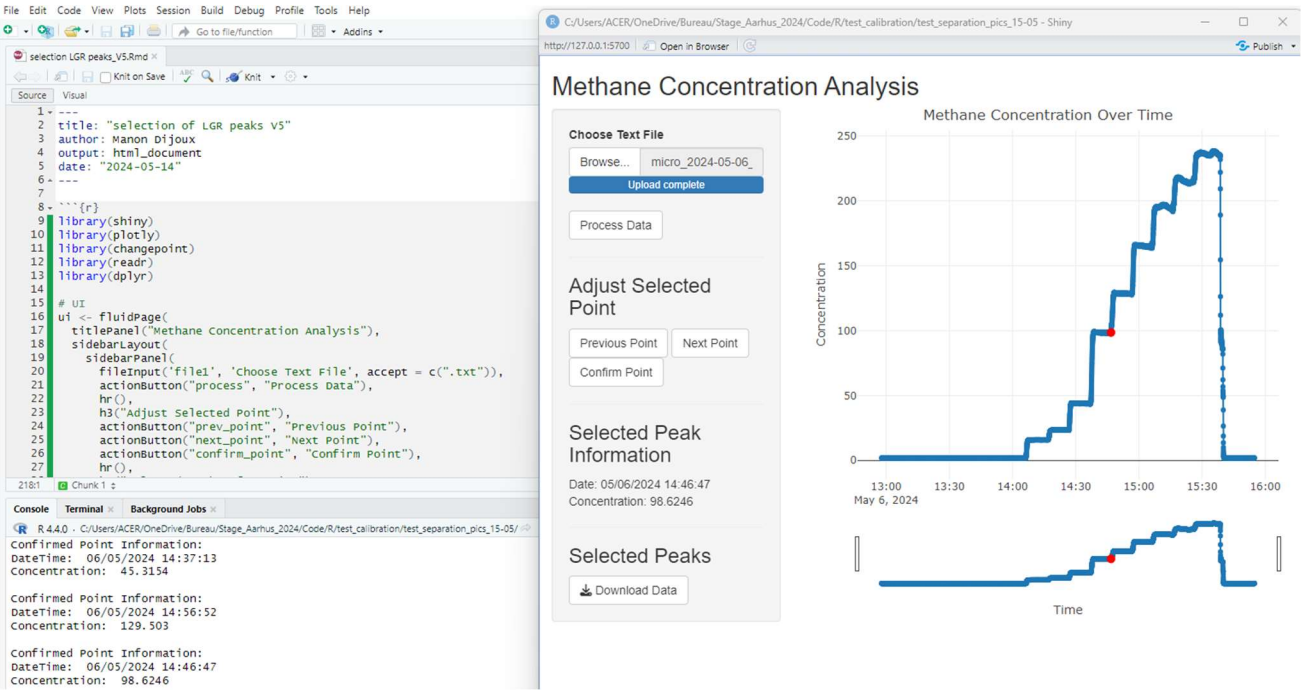


Fig. 8: Shiny interface developed for peak detection

2) Field measurements

a) Study sites

The three ponds selected for the deployment of the chambers were chosen from those identified and studied as part of the Transponder research project. They were selected because a measurement campaign was already scheduled for them during our data acquisition period. This allowed us to access sites that were already identified with previously collected environmental data (physico-chemical), while also taking advantage of a planed day in the field to install our chambers in exchange for our help with the planned sampling. These three ponds were specifically chosen because all the samples for the Transponder project had already been collected there, thereby limiting the interference of our deployment with their data, particularly for eDNA sampling, which requires no contact with the water.

The ponds studied are located about half an hour from Aarhus by car (Fig. 9).

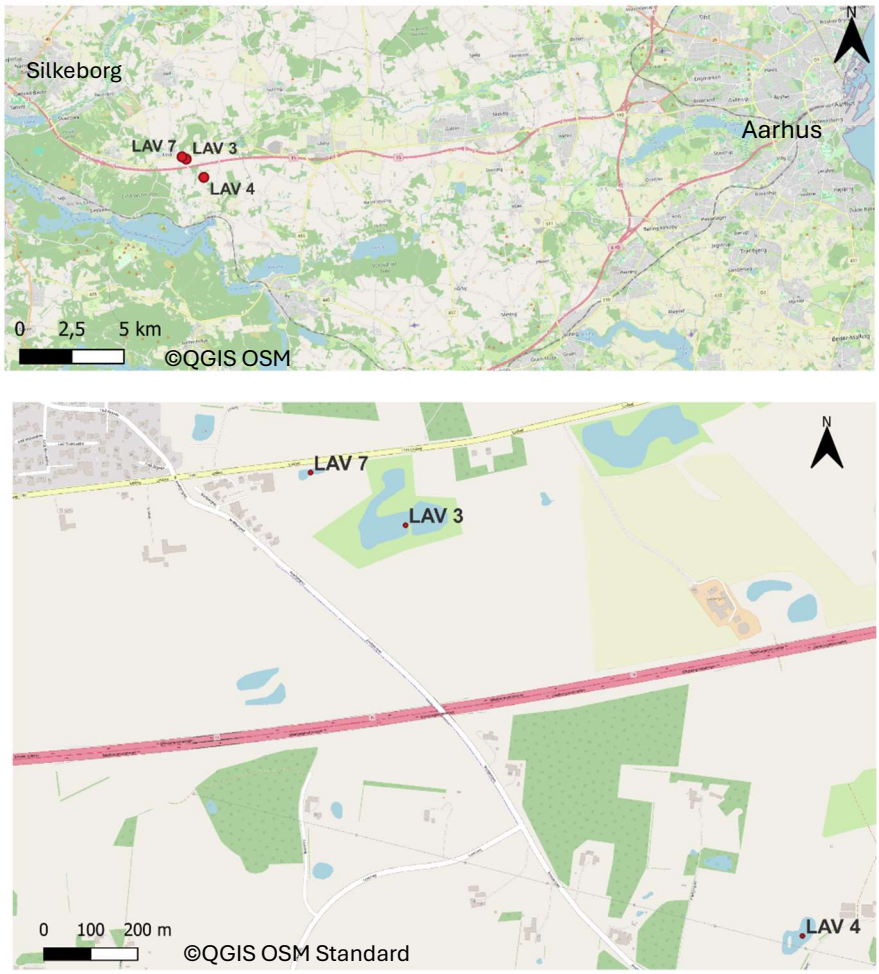


Fig. 9: Location of the studied ponds

Figure 10 shows photos of the three ponds. It can be observed that LAV3 and LAV4 have two types of habitats: a central area without vegetation and shores covered with reedbeds and submerged plants. The main difference between these two ponds lies in their size: LAV3 measures approximately 3245 m^2 with a maximum depth of 70 cm, while LAV4 covers about 14028 m^2 with a maximum depth of 3 m. In contrast, the last pond, LAV7, is completely covered with vegetation, with an area of 855 m^2 and a maximum depth of 60 cm.

Fig. 10: Photos of the ponds studied

Environmental data were collected from measurements previously conducted by the Transponder team, which I will now present in Table 1.

Tab. 1: Environmental data from the ponds studied. RFU: relative fluorescence units (median) ; Chla: clorophyll *a* ; Phyco: cyanobacteria ; DO: dissolved oxygen

	Chla (RFU)	Phyco (RFU)	Temperature (°C)		DO (mg/L)	
			Surface	Bottom	Surface	Bottom
LAV3	36,2	3,3	13,4	13,2	7,35	7,28
LAV4	128,1	20,6	14,8	12,8	11,39	0,4
LAV7	62,6	8,5	13,9	13	9,32	3,3

This table shows the median RFU for Chla and Phyco, calculated from 10 field measurements, as well as the temperatures and DO concentrations at the surface and bottom of each pond at the deepest point. For Chla and Phyco, RFU is used to provide a standardized unit for comparing quantities across the different ponds. This table shows that the three ponds have varied and distinct environmental parameters, despite the initially similar habitats of two of them.

b) Field deployment of chambers

At the time of data collection, six chambers were constructed, calibrated, and ready for deployment in the field. The deployment took place from June 12 to June 19, 2024, across the three ponds.

The objective was to compare the potential impact of different environments on recorded GHG concentrations within the same pond. Therefore, two chambers per pond were installed based on the present habitats.

In the field, one of the six chambers encountered a technical issue and could not be deployed, reducing the total number of chambers to five. It was decided to install only one chamber in pond LAV7, because there was only one habitat present, thus requiring only one chamber. The decision was made to place two chambers in each pond with reed beds to ensure at least some possible comparison between the two.

To deploy the chambers, a paddle board was used to navigate easily on the water without disturbing the environment (figure 11).

Fig. 11: Deployment of a chamber near reedbeds ©Tuba Bucak Onay

The chambers were then activated and left in the water for seven days. At the end of the seventh day, the chambers were retrieved and turned off. The SD cards from each chamber were removed for data analysis.

c) Data analysis for flux calculations using R

Following data acquisition, it is possible to convert the voltage to ppm using the formulas presented earlier, automatically calculated in an R script. The same way, it is also possible to calculate the flux using an R script developed by Ms. Bucak Onay.

The diffusive flux is calculated for CO_2 and CH_4 using the same equation:

$$F_{dif} = \frac{\Delta C_i}{\Delta t} * \frac{P*M}{R*T} * \frac{V_i}{A_i} *1000$$

Where $Fdif$ represents the diffusive flux (mg C m^{-2}, h^{-1}), $\frac{\Delta Ci}{\Delta t}$ is the change in concentrations (ppm* 10-6), P is the atmospheric pressure (atm), R is the gas constant (L*atm/mol*K), T is the temperature (K), M is the molar mass of Carbon (g mol^{-1}), V_i is the volume of the chamber (L), and A_i is the area of the chamber. The value 1000 is for unit conversion from grams to miligrams.

CO_2 Flux: For CO_2, for each 4-hour measurement cycle, the flux is calculated by determining the slope between the first and last measurement after flushing.

CH_4 Flux: For CH_4, three different calculations are necessary.

Total flux: The same steps as for CO_2 are performed. The total flux for CH_4 is calculated in the same manner as the diffusive flux for CO_2 because the diffusive flux for CO_2 represents the total flux.

Diffusive flux: The 4-hour data is divided into 150 sub-datasets, each corresponding to about 10 minutes of data. Then, the flux is calculated for each sub-dataset, and the average of these fluxes is taken. This average is used to calculate the 4-hour diffusive flux. Ms. Bucak Onay determined that this process estimates values representing only the diffusive flux, excluding ebullition.

Ebullitive flux: It is calculated by subtracting the diffusive flux from the total flux.

Following the flux calculation, a data filtering is necessary because some 4-hour cycles are incomplete (beginning and end) or yield unusable results for other reasons, and therefore should not be included in the average calculations. These data result in irregular curves, showing fluctuations rather than a simple increase. Indeed, a decrease should not occur for CH_4 in a hermetically sealed chamber accumulating gases, except during gas dissolution in the water, which is not the case for the phenomena observed in these graphs. The difference between a normal graph and a graph to be excluded is visible in appendix 3.

d) Statistical analysis using t-test to ensure significance of observed differences

One of the objectives is to determine if there is a significant difference in GHG production between different types of habitats. To achieve this, the emissions recorded in two habitats within the same pond was compared. To ascertain whether the observed differences are significant, a statistical test is realised.

For this, a t-test in R was performed to compare two means. Specifically, wel used a paired sample t-test, which is used to see if there is a difference in a group between two points in time (DataNovia, n.d.).

To apply this test, certain conditions must be met:

The two groups are paired, which is our case.

No significant extreme outliers in the two groups.

Normality: The data for each group should be approximately normally distributed.

To ensure that we can use this test on our samples, we perform preliminary tests:

- First, we use the « identify_outliers » function in R to identify the presence of any significant extreme outliers.
- Next, we perform a Shapiro-Wilk test to check for normality.

If these preliminary tests confirm that our data is suitable, we proceed with the final t-test.

Null hypothesis (H0): The difference between the means of the two groups is zero (there is no significant difference).

Alternative hypothesis (H1): The difference between the means of the two groups is not zero (there is a significant difference).

If the p-value is less than the significance level, we reject the null hypothesis and conclude that the difference between the means is statistically significant. If the p-value is greater than the significance level, we do not reject the null hypothesis and conclude that there is no significant difference between the means of the two groups (DataNovia, n.d.). The chosen p-value here is 0.01, which is commonly associated with a highly significant result (DATAtab, n.d.).

III) Results

1) Calibration results

Statistical indicators

Following the calibration, several key results were obtained. The first results examinated indicate whether the calibration was successful, which helps avoid using poorly calibrated or faulty sensors that could skew the analyses.

For the 12 sensors we calibrated, we obtained an average R^2 of 0.987, with a maximum value of 0.99 and a minimum value of 0.97. Regarding RMSE, we have an average of 4.13, which means that, on average, we are 4.13 ppm above or below what was predicted. The maximum RMSE value obtained is 10.16 and the minimum value is 1.07. It should be noted that the lowest R^2 value (0.97) and the highest RMSE value (10.16) both correspond to sensor O29, which is therefore less accurate than the others, although these values are still acceptable. The R^2 and RMSE values obtained for each calibrated sensor can be found in appendix 4.

It is also possible to watch the accuracy of the calibrations more quickly and visually using the graphs generated by the script. One graph shows the nonlinear relationship between the recorded CH_4 concentration and the R_S/R_0 ratio, while a second graph compares the predicted and observed CH_4 concentrations. Figure 12 presents an example for sensor 21, compared with the results from Bastviken showing similar tendencies.

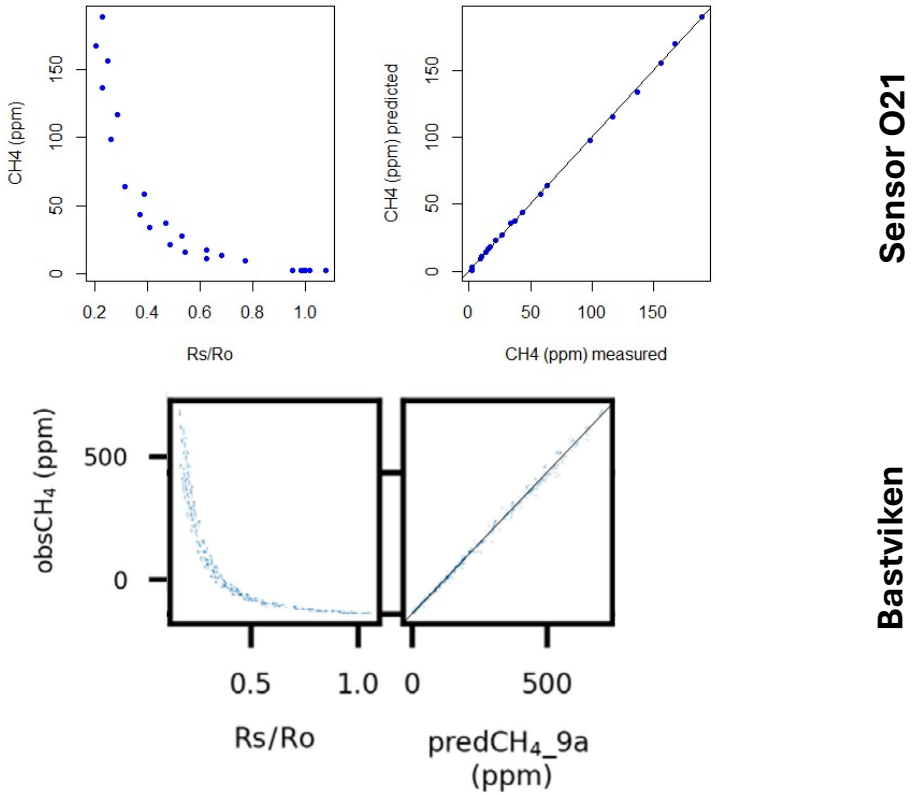


Fig. 12: Relations to visually ensure the accuracy of the calibrations carried out. Comparison of the results presented in the Bastviken et al. publication with the results obtained for the calibration of one of the chambers

Calibration coefficients

To conclude on the calibration, the most important result obtained is the calibration coefficients, which will be used in the R script for converting volts to ppm and calculating flux, correcting the field results based on humidity. Table 2 displays the coefficients obtained for each sensor, along with the gravity (g) and pressure (P) associated with each one. For each sensor, the values of the different coefficients are of a similar order of magnitude but are all different, providing each with a unique combination that reflects its specific manufacturing characteristics.

Table 2: Calibration coefficients for Low-Cost CH_4 Sensors

Sensor	a	b	c	K	g	P
O21	1.21525606	-2.8736003	0.00017849	-3.0245501	0.00824975	95.0374085
O22	1.94415612	-2.9836313	0.00012173	-3.2767883	0.00858744	117.048657
O23	1.73410716	-2.7053608	0.00012153	-3.4007789	0.00770421	86.2150926
O24	4.08115830	-1.9998065	0.00003473	-6.9088588	0.00831238	101.082406
O25	3.73878166	-2.1268358	0.00002975	-6.4885643	0.00801503	114.319302
O26	1.44717909	-2.7566617	0.00011318	-3.0813862	0.00706687	89.4125536
O27	1.67066394	-2.8842481	0.00007621	-2.2348539	0.00674761	82.7016541
O28	1.77533344	-2.7789579	0.00006669	-2.6481566	0.00695650	89.2810698
O29	7.00717356	-3.5409959	0.00007.88	-12.078765	0.00886064	55.7707483
O30	2.21207004	-2.7611711	0.00006164	-4.0878773	0.00682616	83.4332735
O31	5.24788666	-2.0787808	0.00002843	-6.5206799	0.00624540	91.0654527
O32	0.02721349	-3.0629516	0.01070351	-2.3075849	0.00970126	74.5455792

2) Greenhouse gas concentration over time

Based on the recordings obtained in the field using the chambers and the coefficients calculated during the calibration, it is possible to calculate the fluxes of CH_4 and CO_2 emitted from the three studied lakes over a period of 7 days. The main results obtained are presented now.

Average GHG emission

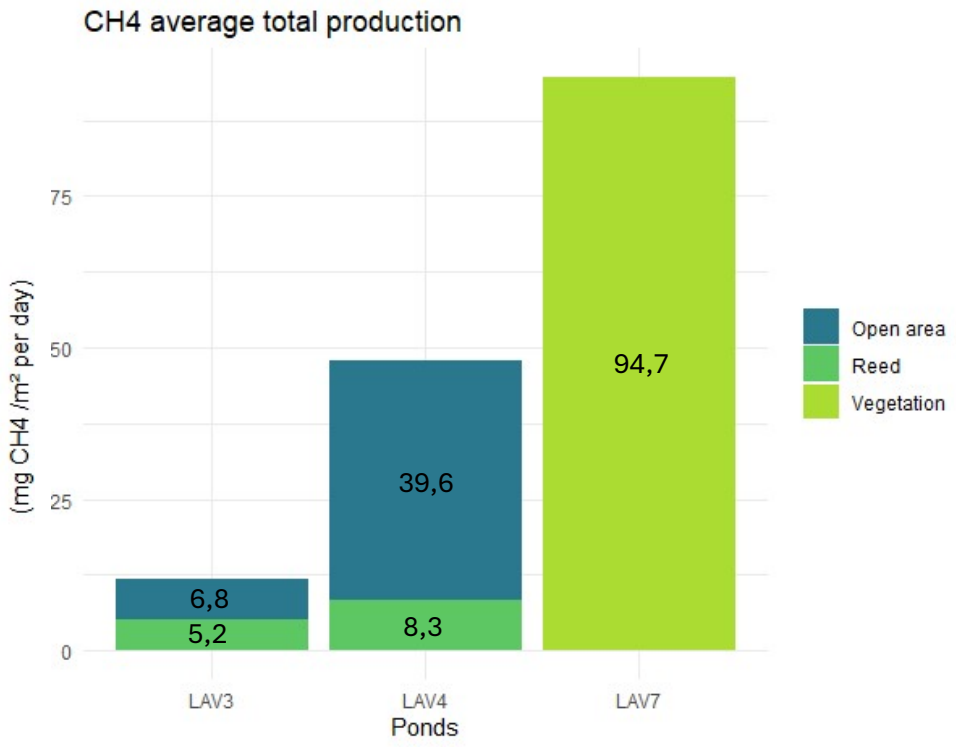


Fig. 13: Average total CH_4 flux per pound and per habitat per day

The figure 13 shows different trends for the three studied ponds. For pond LAV3, there is a higher CH_4 emission of 1.6 mg m^{-2} day^{-1} in the open area compared to reeds. In contrast, a marked difference is observed between the two habitats in pond LAV4, with a higher average CH_4 emission in the open area, showing a difference of 31.3 mg m^{-2} day^{-1}. Pond LAV7, with a vegetation habitat, records the highest average CH_4 emission, much greater than all the other ponds and habitats studied.

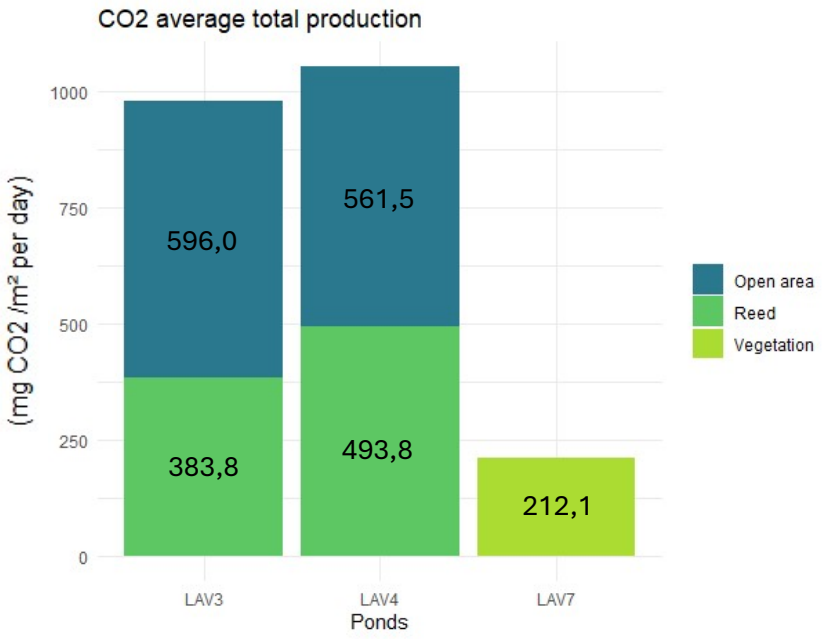


Fig. 14: Average total CO_2 flux per pound and per habitat per day

Regarding the average CO_2 concentrations, we observe relatively similar results between LAV3 and LAV4 ponds in terms of emissions for both Open Area and Reed habitats. For these two ponds, there is a notable difference in emissions between the Reed and Open Area habitats, with a higher emission at the open area level compared to the reeds of 212 mg m^{-2} day^{-1} for LAV3 and 68 mg m^{-2} day^{-1} for LAV4. Proportionally, the LAV7 pond, which has the Vegetation habitat, shows the lowest CO_2 emissions among the ponds studied.

By comparing these two graphs, we notice several points. Firstly, the quantities of CH_4 recorded are lower compared to CO_2; however, CH_4 has a global warming potential (GWP) at least 28 times higher than that of CO_2. Therefore, if we want to compare their GWP, we multiply the average CH_4 emissions obtained from the three ponds (51,53 mg m^{-2} day^{-1}) by 28, which gives 1442,93 CO_2eq mg m^{-2} day^{-1}. Compared to the average CO_2 quantities recorded from the three ponds, which is 749,1 mg m^{-2} day^{-1}, CH_4 is the gas with an impact that is 1,93 times greater.

We can also observe that, proportionally, while the CH_4 production in pond LAV7 is significantly higher than in the other ponds, it has on the contrary the lowest CO_2 emission.

Comparison between two habitats on the same pond

To observe potential differences in emissions based on different habitat types, recorded in the vegetation-proximate area (Reed) and the non-vegetated area (Open Area) for ponds LAV3 and LAV4.

- Comparison of CO_2 emissions

LAV4 CO_2 emissions

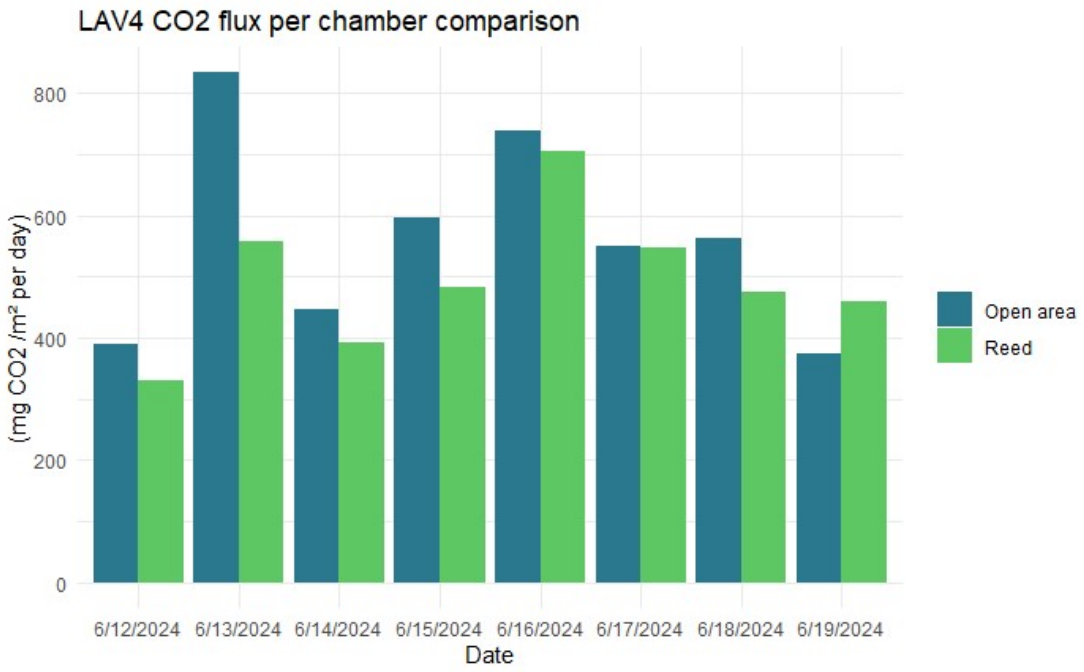


Fig. 15: Total CO_2 flux per chamber per day in two habitats in pond LAV4

For figure 15 regarding CO_2 in pond LAV4, we generally observe higher concentrations emitted in the Open Area compared to the area near the Reeds. However, there are days when the emissions from the two habitats are very similar (on 06/17/2024), or even when the Reeds exhibit higher emissions than the Open Area, as seen on the last observed day (06/19/2024). Overall, there was no statistical difference between habitats (p-value: 0,107). The results of the preliminary tests conducted to ensure the t-test is applicable are available in appendix 5.

LAV3 CO_2 emissions

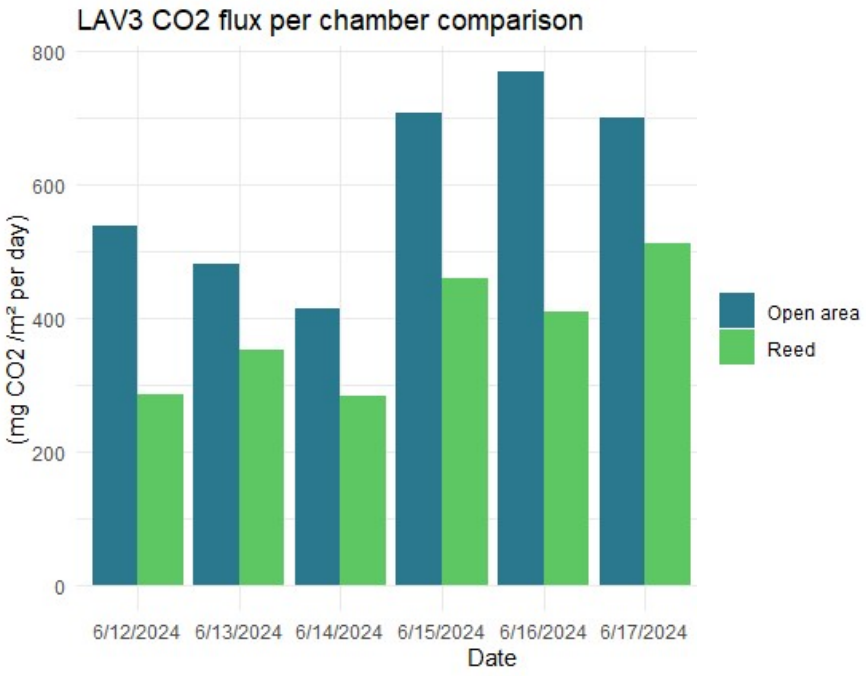


Fig. 16: Total CO_2 flux per chamber per day in two habitats in pond LAV3

In figure 16, which presents CO_2 emissions in pond LAV3, we observe generally higher concentrations in the Open Area compared to the area near the Reeds on all days. On the last two days (June 18 and 19, 2024), emissions were recorded only in the Open Area. This is due to the fact that the chamber located near the reeds ceased functioning on those two days.

There is a significant difference in CO_2 emissions between the two habitats in the LAV3 pond (p-value: 0,00174). Detailed results of the preliminary tests are presented in appendix 6

- Comparison of CH_4 emissions

LAV4 CH_4 emissions

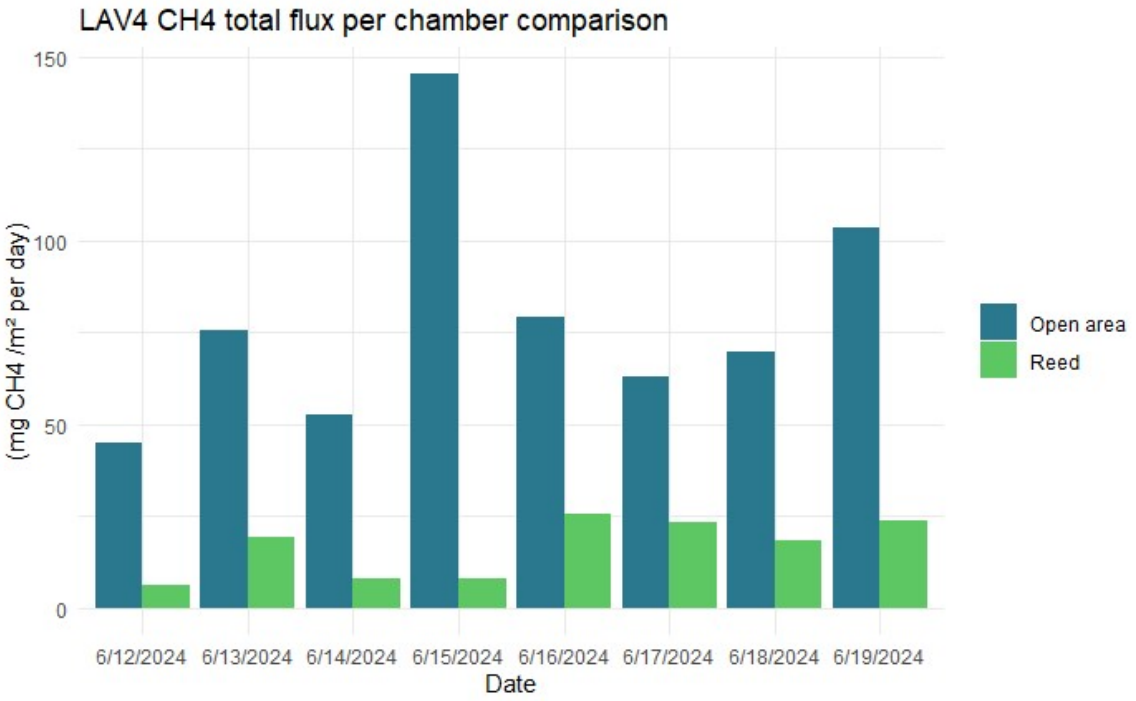


Fig. 17: Total CH_4 flux per chamber per day in two habitats in pond LAV4

In figure 17, we observe a significant difference in CH_4 emissions between the two habitats in the LAV4 pond, with notably higher emissions in the Open Area without vegetation. The t-test was performed following the preliminary tests presented in appendix 7. The results of the t-test confirm a significative difference between the two habitats (p-value: 0,00101).

LAV3 CH_4 emissions

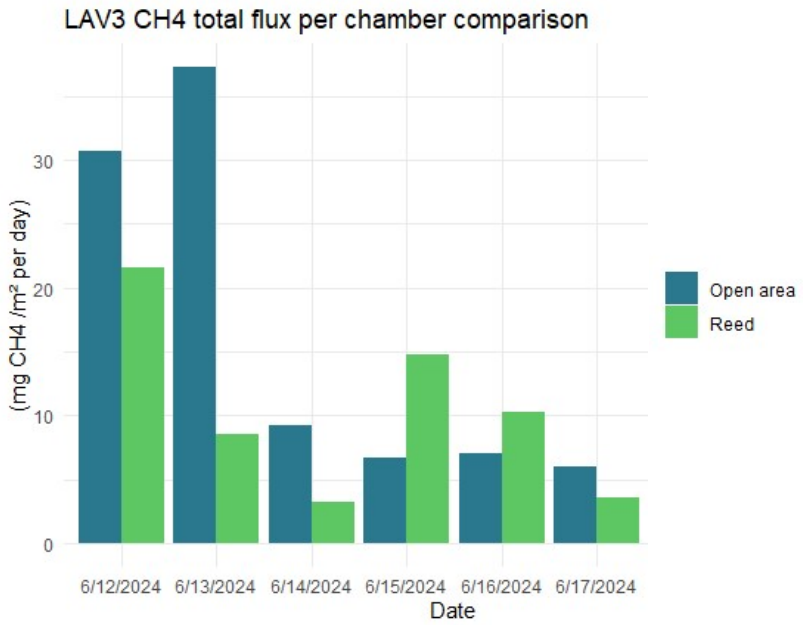


Fig. 18: Total CH_4 flux per chamber per day in two habitats in pond LAV3

The results of CH_4 emissions in the LAV3 pond are different from those of the LAV4 pond. There is actually not one habitat producing more than the other, with notably higher emissions in the reeds compared to the open area on June 15 and 16. This dataset also stands out because the Shapiro test, detailed along with other preliminary tests in appendix 8, indicates that it is not normally distributed. A log10 transformation was performed, and while the data was still not normally distributed, it was closer to normality. Given the robustness of the t-test, we proceeded with it. The results confirm that there is no significant difference in emissions between the two habitats (p-value: 0,315).

Other results

Following the data collection in the field, the proportions of emissions by diffusion and by ebullition at each collection point was calculated. However, given the complexity and vast scope of the subject, it was necessary to focus the analysis on only a portion of the results for the purposes of this internship. Since the primary objective of the internship was to study whether there are significant differences in GHG production between different types of habitats, we concentrated solely on the results presented above. Other graphs generated during data processing are available in appendix 9 for reference.

IV) Discussion and reflective feedback

1) Analysis of produced results

Average GHG emission

Based on the results summarized again in table 19, we observe different trends in CH_4 flux across the three ponds, which could be explained by their environmental differences. In the LAV3 pond, CH_4 emissions are the lowest. This could be due to the fact that it has the lowest quantities of chlorophyll *a* and cyanobacteria, indicating lower availability of nutrients, as well as little variation in temperature and dissolved oxygen between the surface and the bottom: conditions less favorable for CH_4 production. In contrast, LAV7 shows the highest CH_4 fluxes, likely due to higher levels of chlorophyll *a* and cyanobacteria (more significant eutrophication) than LAV3, and lower dissolved oxygen at the bottom. Additionally, as the smallest pond, LAV7 experiences greater temperature variations and warming, which favor GHG production. The vegetation in LAV7 also plays a role: oxygen levels are low at the bottom (3 mg), and the bottom layer may be anoxic. High vegetation can increase labile organic matter and stop light from entering the water, favoring CH_4 production, and dense vegetation beds can create anoxic conditions. However, vegetation may also promote oxidation, reducing CH_4 levels. Finally, the case of LAV4 is more complex. It shows the highest concentrations of chlorophyll *a* and cyanobacteria, as well as very low dissolved oxygen levels at the bottom (0,4 mg/L). These conditions should theoretically favor the highest GHG production, and we should observe the highest concentration among the three ponds, but LAV4 has lower CH_4 emissions than LAV7. This could be explained by the stratification observed in this much deeper pond. Unlike LAV7, where the water is homogeneous, the CH4 emitted in LAV4 remains at the bottom due to the absence of mixing. To verify this, the chambers would have

needed to be left longer until the appearance of a turnover to observe whether or not there was a significant increase in emitted CH4.

Tab. 19: Summary of results

	Chla (RFU)	Phyco (RFU)	Temperature (°C)		DO (mg/L)		Average total CH_4 (mg m^{-2} day^{-1})	Average total CO_2 (mg/m^{-2} day^{-1})
			Surface	Bottom	Surface	Bottom		
LAV3	36,2	3,3	13,4	13,2	7,35	7,28	12	979,8
LAV7	62,6	8,5	13,9	13	9,32	3,3	94,7	212,1
LAV4	128,1	20,6	14,8	12,2	11,39	0,4	47,9	1055,3

Regarding CO_2, as shown in the results of table 19, we observed different trends, with higher emission levels in ponds LAV3 and LAV4 compared with LAV7, which had the highest concentration of CH_4. These observations can be explained by two phenomena. Firstly, CO_2 is consumed by plants during respiration. The pond LAV7, which is entirely covered with vegetation, might explain the observed lower CO_2 fluxes, even though floating plants primarily absorb CO_2 from the air. Additionally, different types of bacteria (and archaea for CH_4) are involved in the production of CO_2 and CH_4, and the presence of oxygen favors CO_2-producing bacteria compared with anoxic conditions. This might explain why, in ponds LAV3 and LAV7, where there is no stratification, mixing and greater oxygen availability at the bottom lead to higher CO_2 production. Furthermore, stratification in pond LAV4 could limit the emission of produced CO_2.

These preliminary results already suggest the influence of environmental factors on the amount of greenhouse gases emitted into the atmosphere, particularly regarding the impact of temperature, pond depth, and eutrophication, as highlighted in the literature. However, beyond the overall emissions by pond, initial analyses also show that, for some ponds like LAV3, the concentrations of CH_4 emitted are similar across the two habitats. In contrast, LAV4 exhibits a significant difference between the emissions from the two habitats which might be due to depth and DO levels at the sediment, stratification being much rarer in the shallow edges. We will explore these observations in more detail in the following section of the report.

Comparison between two habitats on the same pond

The comparative analysis of CO_2 and CH_4 emissions between the two habitats (open area and near the reeds) for ponds LAV4 and LAV3 produced mixed results. Significant differences were observed in CO_2 emissions for LAV3 and CH_4 emissions for LAV4, with markedly higher emissions in the open areas lacking vegetation. However, it is important to note that LAV3 was not stratified, meaning both the middle and the edge were more similar as both were oxic. In contrast, LAV4 was stratified in the deep part (middle) but probably not at the edge, so bubble emissions were much lower at the edge and higher in the middle. Therefore, the comparison is not merely between the edge and middle but rather between anoxic and oxic conditions. This suggests a notable impact of habitat, particularly stratification, on GHG emissions in ponds.

However, contradictory results were observed for other measurements, such as CH_4 emission in LAV3 and CO_2 production in LAV4, where the differences were not statistically significant and emissions varied by habitat depending on the day. These variations may be attributed to sporadic events affecting the chambers during data collection (such as the presence of animals, for example) or other unidentified phenomena influencing emissions. Additionally, LAV3 is a lot smaller than LAV4, so the habitats within LAV3 were not so different from each other, with reedbeds close to the chamber in the open area and similar depths. So, the results may be normal for LAV3. In contrast, LAV4 has more distinct habitats, with a significant difference in depth, almost 2 meters, between the littoral zone and the open area. To gain further insights, it would be necessary to continue measurements and also examine additional data collected, such as the differences between CH_4 ebullition and diffusion, to better understand the underlying phenomena.

2) Suggestions for changes, improvements, continuation

Continuation of the analyzes started during the internship

During this internship, five chambers were deployed across three different ponds for 7 days. The results obtained allowed us to identify some initial trends, such as potential variations in GHG emissions based on habitat types and environmental conditions. However, it is important to note that this sampling for 1 week is limited and may not be representative in the long term or on a larger geographical scale.

Sampling in only three ponds over a period of seven days limits our ability to draw definitive conclusions and establish a definitive trend. To obtain more representative results, it would be necessary to extend the sampling duration and diversify the sampling sites and habitats.

Since the end of the internship, this continuation is already underway with the construction of a total of 25 chambers, some of which have already been deployed.

Limitations of this collection method regarding vegetation

It is important to note that although this collection method becomes more accessible thanks to its price and therefore facilitates the acquisition of data, it still does not account for all GHG emission phenomena. As mentioned in the introduction, plants are a significant vector for GHG emissions, especially hollow-stemmed plants like reeds that transport gas from their roots directly to the surface without passing through water, thus avoiding oxidation (Li et al., 2021). Since the chambers cannot be placed on such tall vegetation, some gas emissions are not recorded. These observations call into question the previous findings that suggested reduced emissions near vegetation, highlighting the importance of conducting additional measurements.

This vegetation also poses another problem during data collection as it creates shade and blocks sunlight from reaching the solar panels, causing the chamber to shut down when it is not charged. These points indicate that this methode, in its current state, is not yet well-suited for all types of environments.

Consequences of using low cost equipment

Calibration and data acquisition are time-consuming processes that take several days each time. Errors or problems during these stages are often only detected at the end of the process, requiring a restart from the beginning. Low-cost equipment, while attractive in terms of cost, remains highly sensitive, fragile, and sometimes unreliable, and such issues are recurring. During the internship, it was common to have to repeat manipulations or repair chambers several time.

To give concrete examples, during every calibration, out of the 5 sensors tested simultaneously, at least one provided incorrect results. Additionally, among the 6 chambers that were supposed to be deployed on the field, one did not function from the outset, even before its deployment in the water. And out of the remaining 5, two stopped working before the end of the experiment. This resulted in the loss of two days of comparative data collection, thereby reducing the amount of data available for analysis.

To improve the reliability of the chambers, it might be considered to use slightly more expensive equipment. However, this contradicts the initial objective of cost reduction. It is important to note that the process of constant repair and improvement is an integral part of research and development. At the beginning, many problems may arise, but eventually, an optimal solution will be reached. Moreover, current alternative solutions are very expensive and are not free from problems. They also do not allow for simultaneous measurements on the same lake as our current solutions do. Therefore, it is important not to focus solely on the negative aspects. To estimate GHG emissions from various habitats globally and develop restoration measures, we truly need affordable solutions. It would thus be preferable to invest this time in repeating manipulations and frequently repairing the devices in order to find a balanced solution in the end.

Integration of the observations into planning decisions

Given the environmental challenges and climate change, the recreation of natural areas, particularly wetlands and small water bodies, has been widely encouraged. These initiatives are justified for reasons such as preserving habitat and species diversity, improving water and soil quality, creating cool spots, regulating runoff to meet environmental expectations, as well as for aesthetic and recreational purposes (Audet et al., 2020).

However, research on GHG emissions shows that these water bodies can sometimes have counterproductive effects under some conditions. Human influence, through the addition of nutrients, urbanization, and other degrading factors, promotes eutrophication and GHG production.

While planning our territories, it is crucial to consider this aspect when making decisions and developing planning documents or projects. It is essential to avoid inappropriate solutions, such as creating ponds in heavily urbanized areas, where conditions may not allow the ecological benefits without generating adverse effects.

It is worth noting that these considerations are already integrated into some planning discussions, such as in the Loir-et-Cher decree for example, regarding the obligations and framework to be followed when creating a pond: "It should not be ignored that the creation of a pond impacts aquatic environments (modification of water quality by increasing temperature and turbidity...). The creation of a pond requires careful consideration of its technical characteristics to ensure it

integrates perfectly into its environment both biologically and aesthetically." (Préfet de Loire-et-Cher, 2023). It is important to continue evaluating and adjusting our practices to ensure that environmental initiatives effectively contribute to combating climate change and preserving our ecosystem without creating new sources of problems.

3) Career projection (research)

This internship provided me with a valuable opportunity to gain a deeper understanding of applied research, particularly in the fields of environmental science and aquatic ecosystems. It not only allowed me to explore aspects of research that were not covered in my academic courses but also enabled me to practically apply the theoretical knowledge I had acquired. I developed a range of skills, from participating in data collection campaigns and creating electronic measurement devices to using the R software.

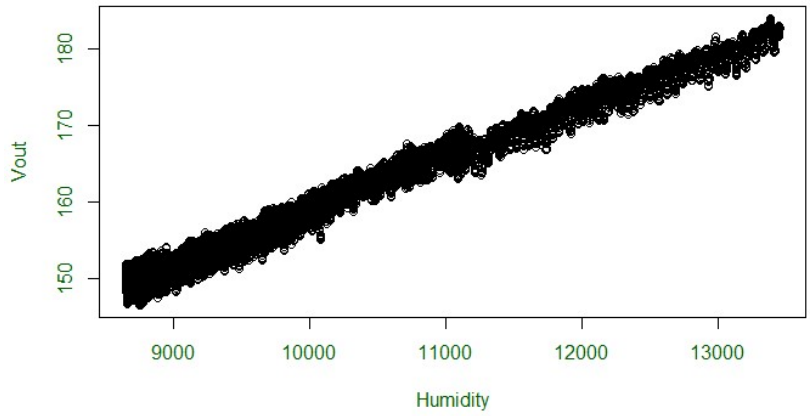
Beyond the knowledge I gained about greenhouse gases, I also had the opportunity to interact with numerous researchers and participate in the weekly "water talks," which offered me a comprehensive view of the diversity of research topics and the people involved in the field of water.

Through this internship, I gained a better understanding of research methodologies, data analysis techniques, and experimental approaches. I also improved my scientific reasoning skills. These skills are crucial not only for a career in research but also for any role requiring an analytical and critical approach, which is essential for engineering professions.

Finally, this internship gave me a concrete sense of the numerous opportunities available in research and engineering within the fields of water and environmental science. For these reasons, I am deeply grateful to all the people I had the pleasure of meeting during this experience and whom I wish to thank once again.

Appendix

Appendix 1



Appendix 1: relationship between voltage (V) and humidity (ppm) in the absence of CH_4

Appendix 2

Detailed steps of the calibration procedure:

Preparation: Place the sensor to be calibrated in a chamber positioned on water at ambient temperature, accompanied by a CH_4 sensor in ppm serving as a reference. It is advisable to leave the sensors on in the chamber for at least 4 hours before calibration to allow them to warm up.

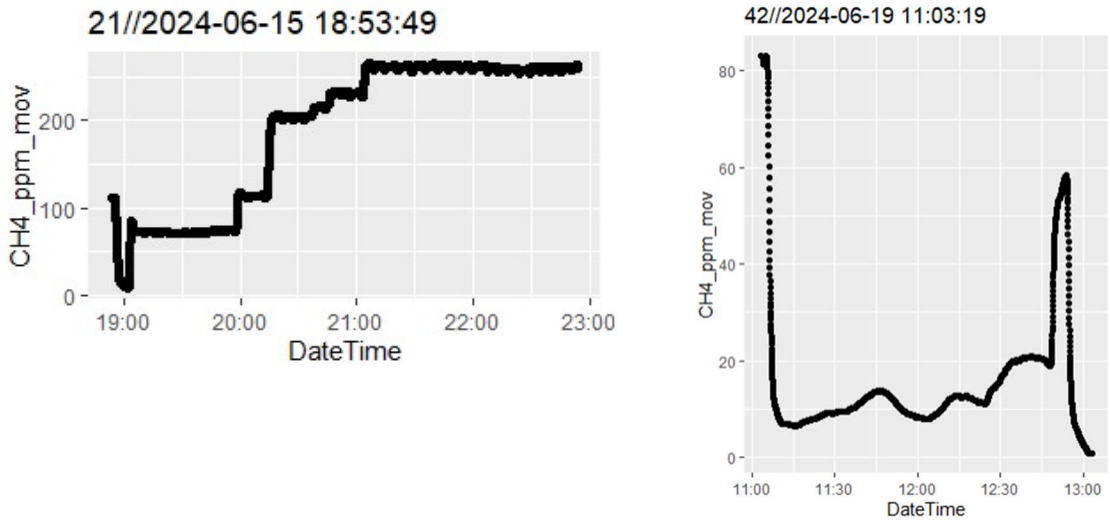
Methane injection: Inject a known quantity of methane using a syringe into the chamber via a valve. Wait approximately ten minutes for stabilization before the next injection. Repeat this process several times until reaching a concentration of about 200 ppm.

Temperature variation: After taking measurements at ambient temperature, empty the chamber and repeat the operation in a cold chamber to obtain temperature variations.

Data sorting: The CH_4 sensor data should be sorted to isolate and retain only the data points at the moment of CH_4 injection. It is also necessary at this stage to match the DateTime corresponding to the injections recorded by the CH_4 sensor and the low-cost sensors, as they can differ by a few seconds to several minutes. Identify the measurement corresponding to the same injection in the data from both sensors and calculate the time difference in seconds to determine the time lag between the two sensors and adjust accordingly.

Data analysis: Implement the collected data into an R script developped by ms Bucak Onay, which automatically calculates the calibration coefficients a, b, c, and k, and generates control charts that will be presented in the results section.

Appendix 3



Appendix 3: Comparison between a normal CH_4 flux graph and a non-usable one

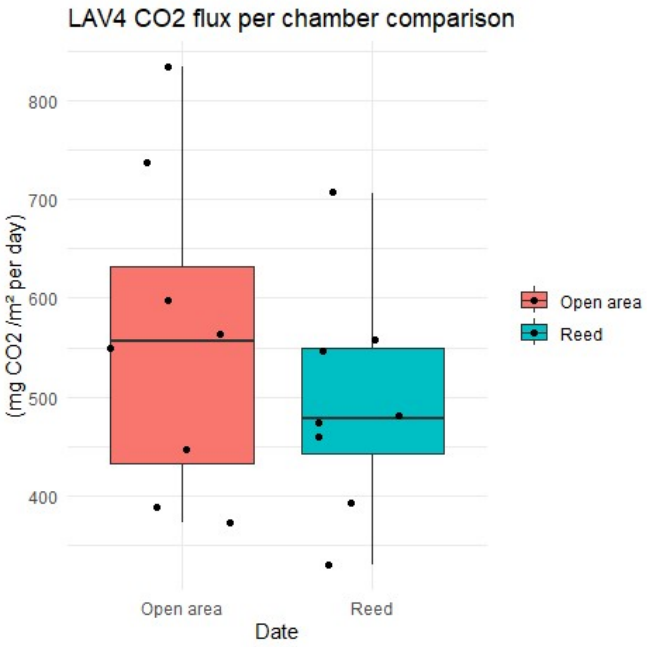
Appendix 4

Appendix 4: Calibration statistics for low-cost methane sensors

Sensor	R^2	RMSE
O21	0,99	1,5
O22	0,99	4,06
O23	0,99	2,89
O24	0,99	3,65
O25	0,99	4,25
O26	0,99	4,63
O27	0,99	2,75
O28	0,99	4,03
O29	0,97	10,16
O30	0,98	7,69
O31	0,99	1,07
O32	0,99	2,89
Mean	0,9875	4,1308

Appendix 5

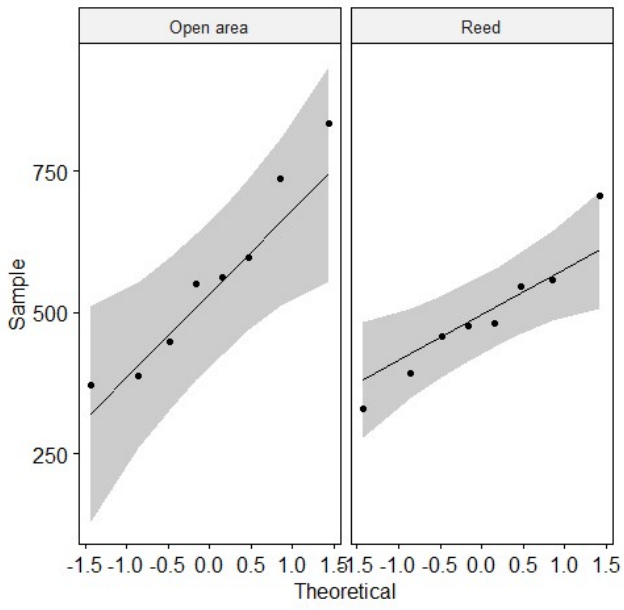
Preliminary test for CO_2 emissions in pond LAV4



Appendix 5 a: Box plot of CO_2 emissions for the two habitats in LAV4

```
> #Outliers
> LAV4_CO2 %>%
+   group_by(Habitat) %>%
+   identify_outliers(flux_mgC_m2_day)
[1] Habitat          Date             flux_mgC_m2_day Chamber          is.outlier       is.extreme
<0 lignes> (ou 'row.names' de longueur nulle)
```

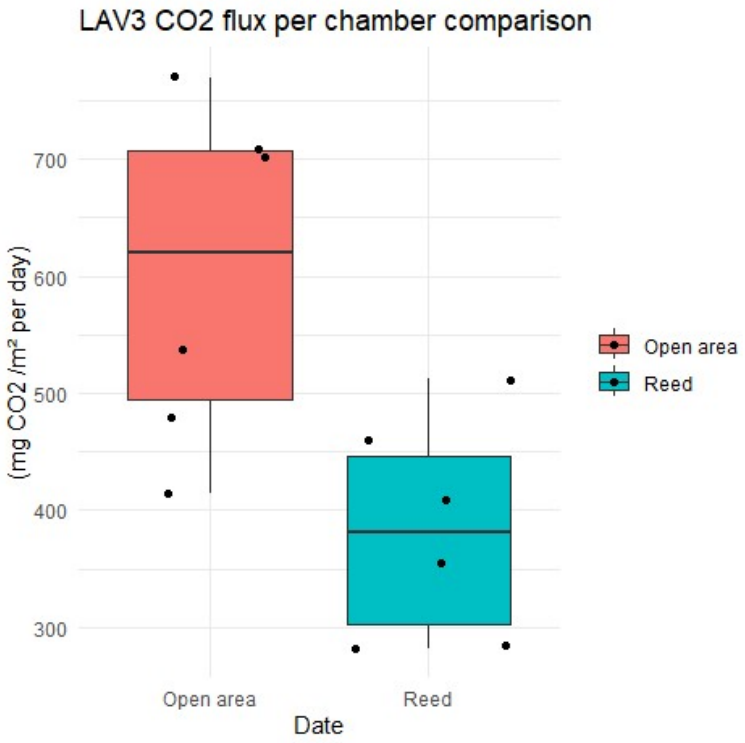
Appendix 5 b: Results of the outlier test for CO_2 emissions in the LAV4 pond



Appendix 5 c: Results of the Shapiro test assessing the normality of the dataset for CO_2 emissions in the LAV4 pond

Appendix 6

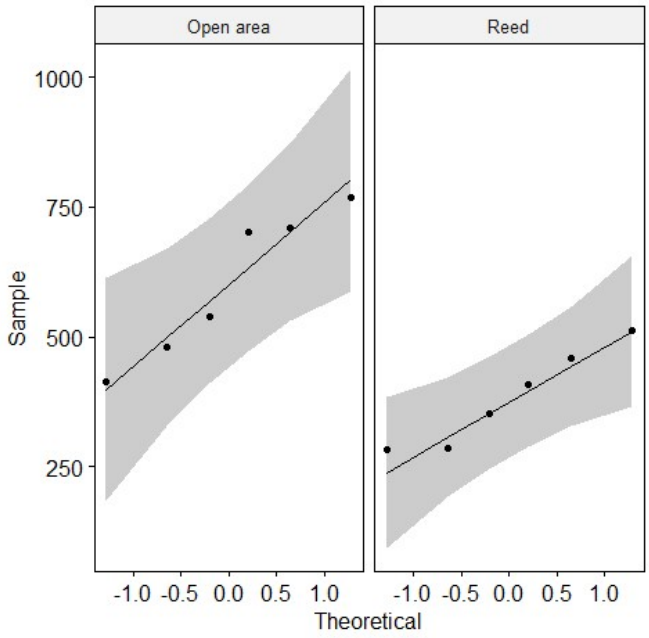
Preliminary test for CO_2 emissions in pond LAV3



Appendix 6 a: Box plot of CO_2 emissions for the two habitats in LAV3

```
> #Outliers
> LAV3_CO2 %>%
+   group_by(Habitat) %>%
+   identify_outliers(flux_mgC_m2_day)
[1] Habitat          Date             flux_mgC_m2_day Chamber          is.outlier       is.extreme
<0 lignes> (ou 'row.names' de longueur nulle)
```

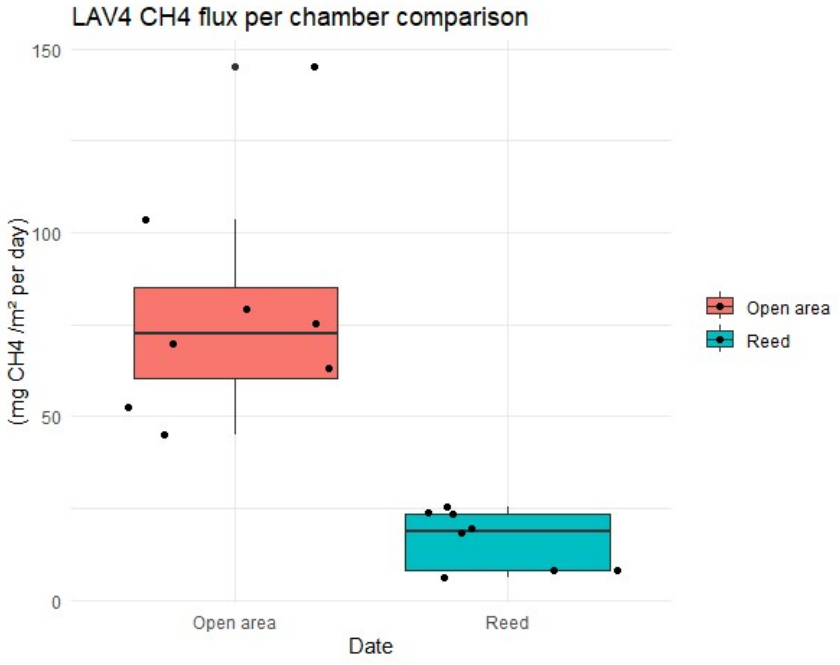
Appendix 6 b: Results of the outlier test for CO_2 emissions in the LAV3 pond



Appendix 6 c: Results of the Shapiro test assessing the normality of the dataset for CO_2 emissions in the LAV3 pond

Appendix 7

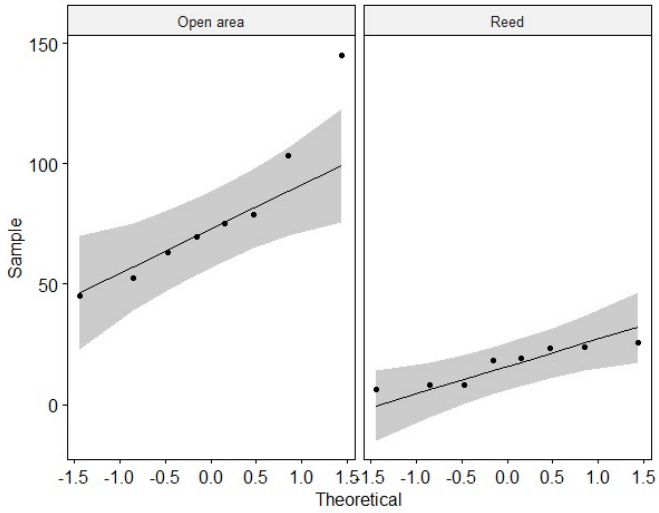
Preliminary test for CH_4 emissions in pond LAV4



Appendix 7 a: Box plot of CH_4 emissions for the two habitats in LAV4

```
> #Outliers
> LAV4_CH4_total %>%
+   group_by(Habitat) %>%
+   identify_outliers(flux_mgC_m2_day)
# A tibble: 1 x 6
  Habitat   Date       Chamber flux_mgC_m2_day is.outlier is.extreme
  <chr>     <chr>      <chr>             <dbl> <lgl>      <lgl>
1 Open area 6/15/2024 LAV_O6             145. TRUE       FALSE
```

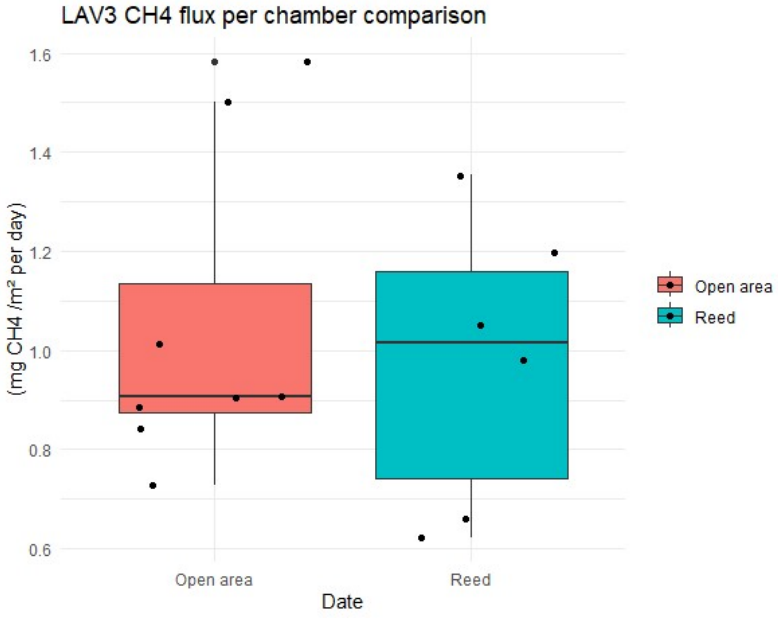
Appendix 7 b: Results of the outlier test for CH_4 emissions in the LAV4 pond



Appendix 7 c: Results of the Shapiro test assessing the normality of the dataset for CH_4 emissions in the LAV4 pond

Appendix 8

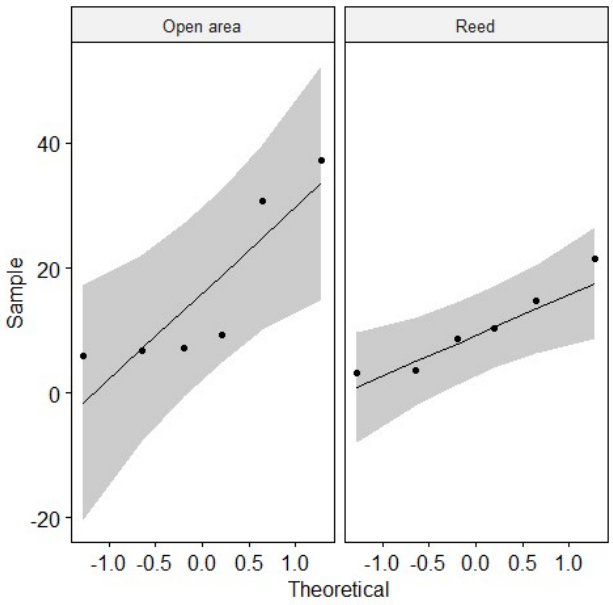
Preliminary test for CH_4 emissions in pond LAV3



Appendix 8 a: Box plot of CH_4 emissions for the two habitats in LAV3

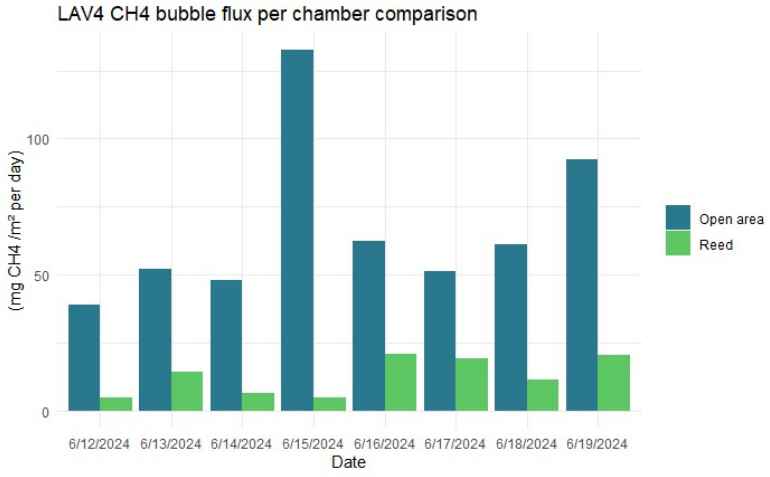
```
> #Outliers
> LAV3_CH4_V2 %>%
+   group_by(Habitat) %>%
+   identify_outliers(flux_mgC_m2_day)
[1] Habitat         Date            Chamber         flux_mgC_m2_day is.outlier      is.extreme
<0 lignes> (ou 'row.names' de longueur nulle)
```

Appendix 8 b: Results of the outlier test for CH_4 emissions in the LAV3 pond

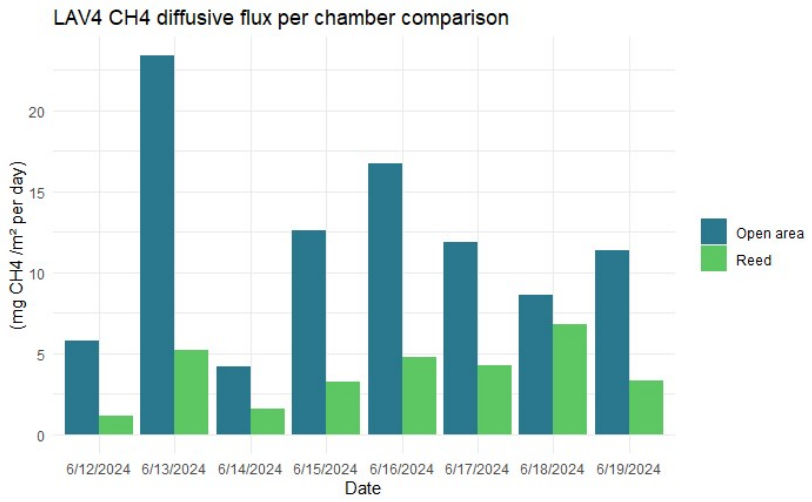


Appendix 8 c: Results of the Shapiro test assessing the normality of the dataset for CH_4 emissions in the LAV3 pond

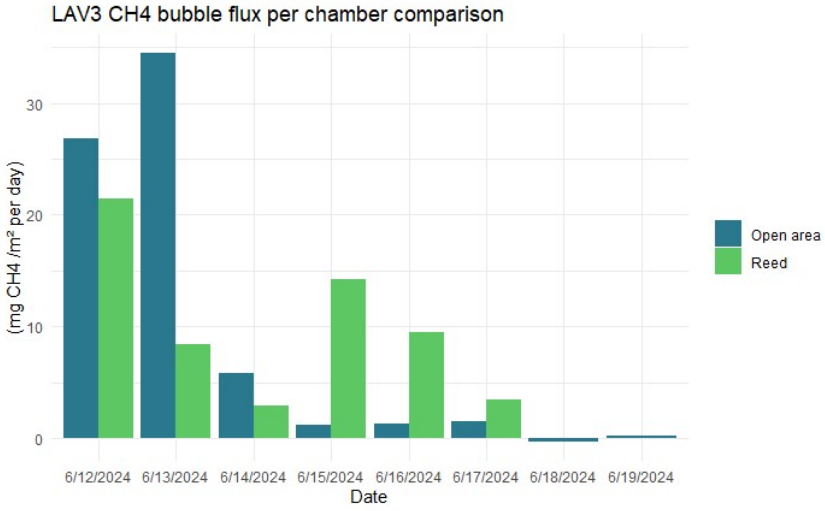
Appendix 9



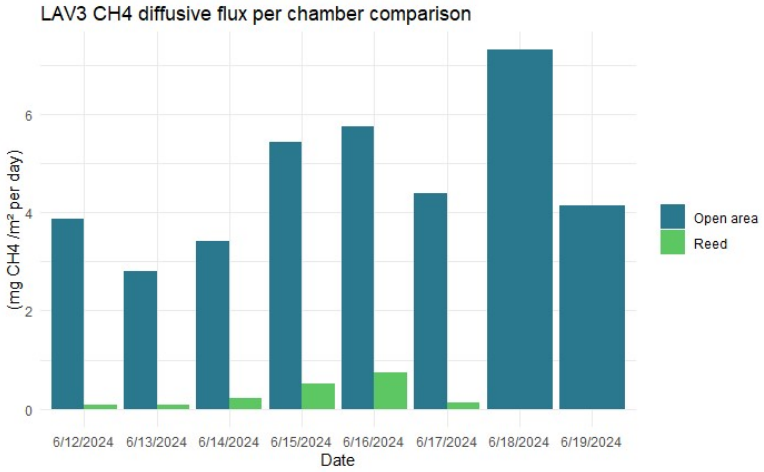
Appendix 9 a: CH_4 bubble flux per chamber per day in two habitats in pond LAV4



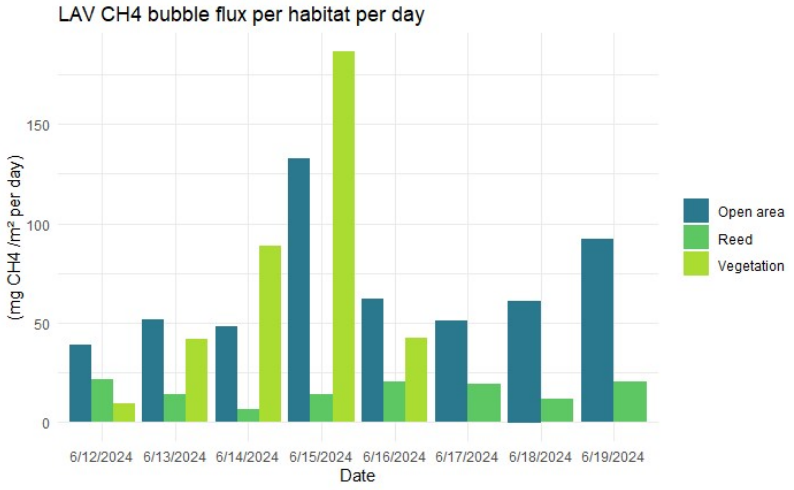
Appendix 9 b: CH_4 diffisive flux per chamber per day in two habitats in pond LAV4



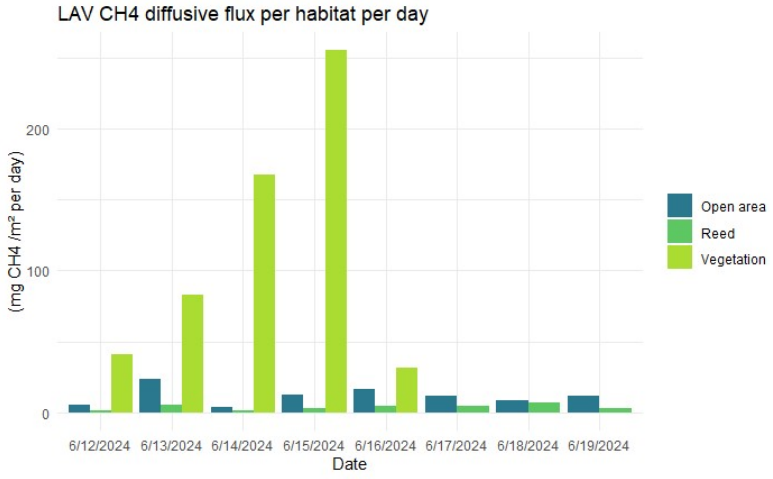
Appendix 9 c: CH_4 bubble flux per chamber per day in two habitats in pond LAV3



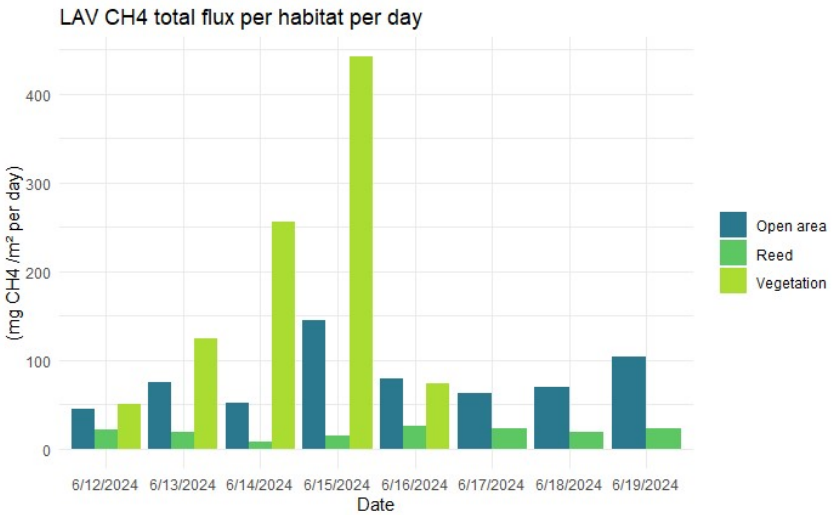
Appendix 9 d: CH_4 diffusive flux per chamber per day in two habitats in pond LAV3



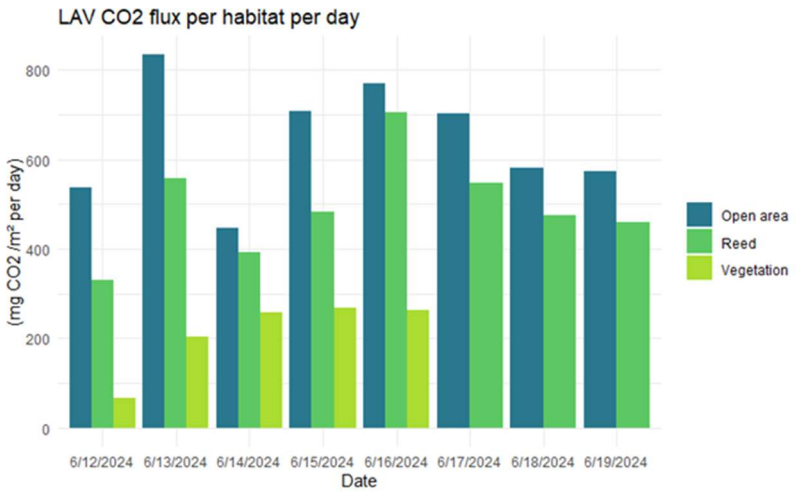
Appendix 9 e: CH_4 bubble flux per habitat per day



Appendix 9 f: CH_4 diffusive flux per habitat per day



Appendix 9 g: total CH_4 flux per habitat per day



Appendix 9 h: CO_2 flux per habitat per day

References

Aarhus University Ecoscience. (2024, July 1). *Department of Ecoscience*. https://ecos.au.dk/en/

Aarhus University website. (n.d.). *Aarhus University – website*. Retrieved July 10, 2024, from https://international.au.dk/

ABB. (n.d.). *LGR-ICOS Portable analyzers*. Measurement Products. Retrieved July 30, 2024, from https://new.abb.com/products/measurement-products/analytical/laser-gas-analyzers/laser-analyzers/lgr-icos-portable-analyzers

Audet, J., Carstensen, M. V., Hoffmann, C. C., Lavaux, L., Thiemer, K., & Davidson, T. A. (2020). Greenhouse gas emissions from urban ponds in Denmark. *Inland Waters*, *10*(3), 373–385. https://doi.org/10.1080/20442041.2020.1730680

Bastviken, D., Nygren, J., Schenk, J., Massana, R., & Duc, N. (2020). Technical note: Facilitating the use of low-cost methane (CH4) sensors in flux chambers – calibration, data processing, and an open-source make-it-yourself logger. *Biogeosciences*, *17*, 3659–3667. https://doi.org/10.5194/bg-17-3659-2020

DataNovia. (n.d.). T-test in R: The Ultimate Guide. *Datanovia*. Retrieved July 21, 2024, from https://www.datanovia.com/en/lessons/t-test-in-r/

DATAtab. (n.d.). *T-Test, khi-deux, ANOVA, Régression, Corrélation…* Retrieved July 21, 2024, from https://datatab.fr/tutorial/p-value

Davidson, Søndergaard, M., Audet, J., Levi, E., Esposito, C., Bucak, T., & Nielsen, A. (2024). Temporary stratification promotes large greenhouse gas emissions in a shallow eutrophic lake. *Biogeosciences*, *21*(1), 93–107. https://doi.org/10.5194/bg-21-93-2024

Davidson, T., Audet, J., Svenning, J.-C., Lauridsen, T., Søndergaard, M., Landkildehus, F., Larsen, S., & Jeppesen, E. (2015). Eutrophication effects on greenhouse gas fluxes from shallow lake mesocosms override those of climate warming. *Global Change Biology*, *21*. https://doi.org/10.1111/gcb.13062

Li, Y., Shang, J., Zhang, C., Zhang, W., Niu, L., Wang, L., & Zhang, H. (2021). The role of freshwater eutrophication in greenhouse gas emissions: A review. *Science of The Total Environment*, *768*, 144582. https://doi.org/10.1016/j.scitotenv.2020.144582

METU. (2024). *TRANSPONDER | ECOSYSTEM IMPLEMENTATION AND RESEARCH CENTER*. https://ekosam.metu.edu.tr/en/transponder

Préfet de Loire-et-Cher. (2023, December 10). *Création de plan d'eau—Le cadre à respecter*. Les services de l'État en Loir-et-Cher. https://www.loir-et-cher.gouv.fr/Actions-de-l-Etat/Environnement/Eaux-et-milieux-aquatiques/Plan-d-eau/Demande-prealable-a-la-creation-ou-la-modification-d-un-Plan-d-Eau/Creation-de-plan-d-eau-Le-cadre-a-respecter

POLYTECH
TOURS

35 ALLÉE FERDINAND DE LESSEPS
37200 TOURS

Investigating greenhouse gas emissions in pond ecological habitats: Utilizing a low-cost environmental data collection method

Manon Dijoux

2023-2024

Abstract: This report details the development and calibration of a low-cost CH_4 sensor used in floating chambers for monitoring greenhouse gas emissions. The primary objective was to establish an effective and economical calibration method to accurately measure CH_4 concentrations in varying environmental conditions. This involved building the chambers from scratch and calibrating the sensors. The initial results seems to reveal an impact of the environment and habitats on greenhouse gas emissions; however, many questions remain, necessitating further research.

Keywords: calibration, collection chambers, greenhouse gases, habitats, ponds

Aarhus University: Nordre Ringgade 1, 8000 Aarhus C, DENMARK

Company supervisor:
Joachim Audet
Tuba Bucak
Thomas Davidson

Academic supervisor:
Sabine Greulich