# An expert survey on chamber measurement techniques and data handling procedures for methane fluxes

Katharina Jentzsch[1,2], Lona van Delden[1], Matthias Fuchs[3], Claire C. Treat[1,4]

[1]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Potsdam, Germany
[2]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany
[3]Renewable and Sustainable Energy Institute, University of Colorado Boulder, Boulder, CO, USA
[4]Department of Agroecology, Aarhus University, Denmark

*Correspondence to*: Katharina Jentzsch (katharina.jentzsch@awi.de)

**Abstract.**

Methane is an important greenhouse gas but the magnitude of global emissions from natural sources remains highly uncertain. To estimate methane emissions on large spatial scales, methane flux data sets from field measurements collected and processed by many different researchers must be combined. One common method for obtaining in-situ methane flux measurements are flux chambers. We hypothesize that considerable uncertainty might be introduced into data synthesis products derived from chamber measurements due to the variety of measurement setups, data processing and quality control approaches used within the chamber flux community. Existing guidelines on chamber measurements promote more standardized measurement and data processing techniques but to our knowledge, so far, no study has investigated which methods are actually used within the chamber flux community. Therefore, we aimed to identify the key discrepancies between the measurement and data handling procedures implemented for chamber methane fluxes by different researchers.

We conducted an expert survey to collect information on why, where, and how scientists conduct chamber-based methane flux measurements and how they handle the resulting data. We received 36 responses from researchers in North America, Europe, and Asia which revealed that 80% of respondents have adopted multi-gas analyzers to obtain high-frequency (< 1 Hz) methane concentration measurements over a total chamber closure time of typically between 2 and 5 minutes. Most but not all of the respondents use recommended chamber designs, including features such as airtight sealing, fans, and a pressure vent. We presented a standardized set of methane concentration timeseries recorded during chamber measurements and derived $CH_4$ flux estimates based on the processing and quality control approaches suggested by the survey participants. The responses showed broad disagreement among the experts concerning the processes that they consider responsible for nonlinear methane concentration increases. Furthermore, there was a tendency to discard low or negative $CH_4$ fluxes. Based on the expert responses, we estimated a variability of 28% introduced by different researchers deciding differently on discarding vs. accepting a measurement when processing a representative data set of chamber measurement. Different researchers choosing different time periods within the same measurement for flux calculation caused an additional variability of 17%. Our study highlights the importance to understand the processes causing the patterns in $CH_4$ concentrations visible from high-resolution

analyzers as well as the need for standardized data handling procedures in future chamber methane flux measurements. This is highly important to reliably quantify methane fluxes all over the world.

The survey results as well as the questionnaire are publicly available at https://doi.org/10.1594/PANGAEA.971695 (Jentzsch et al., 2024b).

## 1 Introduction

Methane ($CH_4$) is an important greenhouse gas with 45 times the global warming potential of carbon dioxide ($CO_2$) on a 100-year timescale (Neubauer, 2021). However, emission estimates differ largely between "top-down" atmospheric measurement inversions and "bottom-up" approaches using data-constrained or process-based models (Kirschke et al., 2013; Saunois et al., 2020). Natural emissions, especially bottom-up estimates of wetland emissions, are the largest source of uncertainty to the global $CH_4$ budget due to the poorly constrained areal extent of wetlands and other methane-producing ecosystems like lakes, streams, and reservoirs, highly uncertain $CH_4$ process parameterization, and a lack of validation data sets (Melton et al., 2013; Saunois et al., 2020).

One approach to obtain large-scale validation data sets for $CH_4$ fluxes has been to create synthesis datasets of measurements collected by multiple researchers using chamber-based methane flux measurements (Kuhn et al., 2021; Olefeldt et al., 2013; Treat et al., 2018). An advantage of using the closed-chamber technique over in-situ measurements operating on larger spatial scales is that the resulting data sets can capture the high spatial and temporal variability in natural $CH_4$ emissions with small-scale spatial changes in environmental and ecological conditions (Frenzel and Karofeld, 2000; Laine et al., 2007; Moore and Knowles, 1990; Waddington and Roulet, 1996). When applying the closed-chamber technique, a chamber is placed on top of the soil and the change in gas concentrations in the chamber headspace is monitored over time to estimate the exchange of $CH_4$ between soil, plants, and atmosphere on the microscale (e.g. Livingston and Hutchinson, 1995). The rate of change in gas concentrations, after correcting for temperature and pressure conditions using the ideal gas law, is then used to compute the flux of $CH_4$ through the surface area covered by the chamber (Holland et al., 1999). However, despite more than thirty years of chamber-based methane flux measurements from wetland ecosystems (Bartlett & Harriss, 1993; Harriss et al., 1985), developing large-scale methane validation data sets remains challenging.

Two approaches are typically used for measuring the $CH_4$ concentrations inside the chamber: manual sampling and in-line gas analyzers. Manual sampling for gas concentrations involves extracting gas samples from the chamber headspace in regular time intervals using syringes and subsequently analyzing them for $CH_4$ concentrations on a gas chromatograph. A linear fit is then usually applied to the $CH_4$ concentration measurements over time and its slope is used as the flux estimate after correction for the pressure and temperature inside the chamber (Holland et al., 1999). Manual sampling of the chamber headspace is typically characterized by a low sampling frequency which requires a relatively long chamber closure time. The considerations here are balancing the time needed to get a detectable change in $CH_4$ concentrations versus shorter measurement times to reduce chamber effects (Holland et al., 1999).

With the advances in laser spectroscopy, manual sampling is increasingly replaced by continuously circling chamber air through an in-line gas analyzer which performs high-frequency (>1 Hz), high-accuracy, real-time measurements of the $CH_4$ concentration. Through their portability and with reduced measurement times, such multigas analyzers have opened new possibilities, particularly for the analysis of key trace gases like $CH_4$ and $N_2O$. At the same time, the high frequency and high accuracy of the concentration measurements uncover chamber-induced artefacts and events of ebullitive $CH_4$ emission that are superimposed on the signal of natural diffusion of $CH_4$ between soil, plants, and atmosphere. Leakage of gas from the chamber (Hutchinson and Livingston, 2001), a saturation effect changing the concentration gradient between soil and chamber headspace over time (Livingston and Hutchinson, 1995), and natural $CH_4$ ebullition (Strack et al., 2005) as well as ebullition triggered by the chamber placement can all lead to a deviation of the concentration change from the linear increase expected for a constant diffusive flux. These observations call for a reassessment of measurement, processing, and quality control (QC) approaches to minimize the influence of chamber effects on the flux estimates.

Besides the general lack of validation data sets, existing data sets that combine flux data collected by different researchers likely include additional uncertainty due to the variety of measurement and data handling approaches used. Several studies have assessed the difference in flux estimates resulting from different chamber setups (Pihlatie et al., 2013; Pumpanen et al., 2004) and from different data processing approaches such as using nonlinear as compared to linear fits to the gas concentration measurements over time (Forbrich et al., 2010; Healy et al., 1996; Pirk et al., 2016). Such experimental and modelling studies have contributed to several guidelines for chamber measurements that were published in an attempt to establish a more standardized protocol for flux measurements. These best-practice guidelines for chamber measurements summarize recommendations on chamber designs (e.g., Clough et al., 2020) as well as on the entire workflow from measurement to data processing and quality control (e.g., de Klein and Harvey, 2012; Fiedler et al., 2022; Maier et al., 2022). While guidelines outlining best measurement practices for chamber measurements provide a well-founded summary of methods recommended to collect high-quality flux data, chamber-based flux data sets are often lacking detailed metadata reporting on chamber design, flux calculation and QC methods. This introduces substantial uncertainty to comprehensive comparisons of chamber-based data.

Given that the measures outlined in guidelines for chamber measurements have significant effects on the magnitude of $CH_4$ fluxes measured, we need to know how widely implemented these recommendations are and where key differences and knowledge gaps remain. Gathering scientific and technical information from experts is necessary to move beyond established theoretical knowledge and can offer further evidence to aid in decision-making (Morgan, 2014). Several studies have recently used expert assessments to gain valuable insights into topical climate-change-related issues (Macreadie et al., 2019; Rosentreter et al., 2024; Schuur et al., 2013). In this study, we use expert judgement derived from a questionnaire to identify the methods for chamber measurements, processing, and QC of $CH_4$ fluxes that are actually currently used within the flux community and to assess resulting variability and uncertainties.

This study aims to derive starting points for improving the usability of chamber $CH_4$ flux data sets for large scale synthesis studies through reducing the discrepancies between measurement and data handling approaches used within the chamber flux

community as identified from an expert survey. Our objectives were to (1) provide an overview of the chamber designs, measurement setups and routines, flux calculation and QC approaches that are currently used by scientists to quantify $CH_4$ fluxes; (2) estimate the variability that is introduced into $CH_4$ flux data sets by the variety of data handling approaches when a representative data set of chamber measurements is processed by different researchers. Our study raises awareness for differences in chamber methods used within the flux community – a potentially considerable but often neglected source of error in synthesis studies that combine flux data sets collected and processed by different researchers. Through identifying major sources of uncertainty resulting from the variety of measurement, calculation, and QC approaches used within the chamber flux community, we derive starting points for eliminating such error sources and rendering individual flux data sets more comparable and combinable and thus better suited for larger scale synthesis studies.

## 2 Methods

For this study, we evaluated an expert survey conducted in 2023 that consisted of two parts – the first part asking questions about the professional background of the participants and the field sites as well as the measurement, calculation and QC approaches that they use for their own chamber measurements of $CH_4$ fluxes and the second part being an exercise on visual QC of a given set of chamber measurements.

Experts were required to have a minimum expertise of one field season of chamber measurements of $CH_4$ fluxes. They were solicited using emails and conference poster presentations through professional networks, including the Permafrost Carbon Network, C-Peat network, ICOS, and through identification of experts not represented in these networks to increase the number and geographic background of the participants. Altogether, 46 experts were contacted via email. To capture the variety of chamber applications and methods used within the community, we selected the survey participants to be rather independent from each other in their choice of measurement and data handling approaches.

The survey was estimated to take 40 minutes to complete and the survey language was English. The survey was administered using LimeSurvey (Community Edition Version 5.6.68+240625). Survey participants were asked if they wished to be acknowledged or remain anonymous. Survey participation was voluntary and was not compensated. The survey has been legally checked by a data protection officer to comply with the EU data protection regulation and involved a privacy policy statement explaining the use and processing of the collected data that needed to be approved by every survey participant prior to participation. The complete, archived questionnaire and the survey responses are provided in Jentzsch et al. (2024b).

### 2.1 Methods of: Survey part 1 – The survey participants and their chamber measurements

In the first, informative part of the survey, we gathered information on the measurement, data processing and QC approaches that the participants use for their own chamber measurements. For this part of the survey we chose a combination of 20 choice questions (simple and multiple selection including seven yes/no questions), all of which offered to elaborate the selection(s) in a short free text comment, and 19 text entry questions. For a visual overview of the variety of measurement

setups used, we asked the survey participants to upload a photo of their chamber system. To assess the professional background of the group of participants we asked about their professional status, the country of their home institute as well as their educational and scientific background. For an overview of the area of application of chamber $CH_4$ flux measurements, we included questions on the participants' research questions and the regions and ecosystem types they usually work in. Questions on the chamber dimensions, the chamber equipment, measurement instruments, as well as photos thereof, together with questions on the measurement procedure and additional variables monitored showed us the variety of experimental designs used. Additionally, we asked the participants to describe their approaches for flux calculation, quality control, and uncertainty estimation of the flux estimates.

### 2.2. Methods of: Survey part 2 - Visual quality control of a standardized data set

To more directly assess the differences in interpretation of chamber data that lead to the discrepancies in measurement setups, data processing, and QC techniques as identified in the first survey part, we provided a standardized set of chamber measurements for visual QC by the survey participants and extrapolated the responses to a larger, representative data set. This second part of the survey included both qualitative and quantitative responses.

The standardized set of chamber $CH_4$ fluxes was composed of 12 selected chamber measurements from our field campaigns at Siikaneva bog (61°50'N, 24°12'E), Southern Finland, in summer 2021 and summer and fall 2022. The measurements were done using a manual chamber with a volume of 36 L and equipped with a cooling system to keep the chamber temperature close to constant, two fans to mix the air inside the chamber, and a small opening for pressure equilibration. For the measurements, the chamber was placed on collars that were permanently installed in the ground. In 2021, the connection between chamber and collar was sealed with a rubber skirt and in 2022 the rim between chamber and collar was filled with water to make the connection air tight. The gas concentrations inside the chamber were recorded with an in-line gas analyzer at a frequency of 1 Hz. Besides chamber measurements showing a linear increase in $CH_4$ concentration over time, we included examples showing a variety of deviations from the linear increase expected for constant diffusive wetland $CH_4$ emissions.

For visual QC of the measurements by the survey participants, we provided the concentrations of $CH_4$ over time as well as the simultaneously measured concentrations of $CO_2$ and $H_2O$ in the chamber, a photo of the chamber, and a description of the measurement setup as well as ,for each measurement example, information on dominant vegetation and water table depth at the measurement plot, date and time of the measurement, transparent vs. opaque chamber, gas analyzer model and a photo of the measurement plot (Figure A1a). We asked the participants if they would keep the respective measurement for flux calculation or if they would discard it and why they would do so (Figure A1b). If they decided to keep the measurement, we asked them to select the part of the measurement that they would use to calculate the $CH_4$ flux by submitting the start and end times of this period in seconds after chamber closure.

## 2.3. Statistical analyses

### 2.3.1 Cleaning of the data set

We anonymized the survey responses by separating the demographic information including the country of the home institute, the scientific background, the highest education level, the time since PhD completion, and the current professional role of the participants from each other and from the rest of the survey results. We furthermore removed the question for specific research sites before publishing the data and replaced two names of specific research sites given as a description of the main study regions by terms for a larger region. In one response, we removed the name of another researcher mentioned by one of the participants.

We harmonized and/or categorized certain free text responses including the responses on the chamber shape, the chamber area, chamber volume, the closure time of the chamber, and the frequency of the gas concentration measurements inside the chamber. From the chamber volume and chamber area we calculate the effective chamber height. We corrected obvious writing mistakes throughout the survey as part of the standardization. In questions on QC procedures, we standardized the information on the exclusion of the beginning of the measurement from flux calculation as well as the length of the excluded time period. We also adjusted the responses to questions whether to keep or to discard a measurement in the visual QC exercise when the free text responses clearly revealed that the wrong box had been ticked by mistake. We set the $CH_4$ flux to zero in two cases where survey participants clearly stated in their free text responses that this is how they would handle the presented measurement.

### 2.3.2 Evaluating the visual QC exercise

We quantitatively and qualitatively evaluated the responses to the visual QC portion of the survey. We summarized the reasons for keeping or discarding a measurement as elaborated in the free text responses to the visual QC part of the survey. Then, we numerically evaluated the visual QC performed on the 12 example measurements. This allowed us to quantify the variation in fluxes due to quality control and differences in fitting approaches among researchers. For this, we calculated the $CH_4$ fluxes for each researcher for each of the 12 example measurements using the time periods selected by the researcher.

To calculate the fluxes, we used a standard linear fitting approach and accounted for differences in temperature and pressure among the measurements (Holland et al., 1999). The ideal gas law was used to convert the rate of change in $CH_4$ concentrations ($\frac{dc_{CH_4}}{dt}$) in ppm $s^{-1}$ to the molecular $CH_4$ flux ($F_{CH_4}$) in mol $m^{-2}$ $s^{-1}$ for each measurement example *i* (i=1,…,n, where n=12) and each survey participant *j* (j=1,…,m, where m=36).

$$F_{CH_4\,i,j} = \frac{dc_{CH_4}}{dt}_{i,j} \times 10^{-6} \times \frac{p}{R \times T_i} \times \frac{V_i}{A},$$

where *p* represents the standard atmospheric pressure of 101325 Pa, *T* (degrees K) is the mean temperature inside the chamber during the closure, and *A* is the surface area of the chamber in $m^2$. $V_i$ is the volume of the chamber used in measurement *i*, calculated by $V_i = A \times h_i$, where $h_i$ is the effective height of the chamber headspace during measurement *i* (in m), calculated

as the mean of the height above the soil surface or vegetation cover that was measured at three points around the chamber for each measurement plot. $R$ is the Ideal Gas Constant of 8.314 kg $m^2$ $mol^{-1}$ K $s^{-2}$. We then converted the molecular $CH_4$ flux to the more commonly used mass flux of $CH_4$ using the molar mass of $CH_4$ of 16.04 g $mol^{-1}$. For each measurement example and each participant, $\frac{dc_{CH_4}}{dt}$ was estimated as the slope of a linear fit (lm function from stats package in R version 4.3.0) to the $CH_4$ concentrations within the time period selected by the researcher. For reasons of consistency, we used a linear fit even in the 12 cases that a participant suggested to use a nonlinear fit instead (7% of the total of 173 times that start and end times for flux calculation were given by a participant). When a measurement was accepted by an expert but no start and end time was given for flux calculation, we estimated the flux based on the entire chamber measurement.

We used the fluxes calculated from the quantitative responses to assess the variability in $CH_4$ flux estimates and QC procedures due to different researchers processing the measurement data, that is (1) the variability in flux estimates introduced by different researchers selecting time periods for flux calculation, and 2) the variability in the share of measurements kept for flux calculation during QC. In a representative data set of 788 chamber measurements, collected at Siikaneva bog in 2021 and 2022 (Jentzsch et al., 2024a), we visually identified and categorized the following eight classes of measurement scenarios based on the shape of the $CH_4$ concentrations measured in the chamber headspace over time: "Linear increase", "Linear decrease", "Nonlinear increase – decreasing slope", "Nonlinear increase – increasing slope", "Initial jump", "Jump(s)", "Inconsistent trend", and "Low variation". During the majority (60%) of measurements in the Siikaneva data set $CH_4$ concentrations increased linearly over time (Table A1). The second largest group, represented by 18% of the measurements, showed a nonlinear, weakening increase in $CH_4$ concentrations over the time of the chamber closure. During 8% of the measurements an abrupt jump in $CH_4$ concentrations in the beginning or one or several jumps at a later time during the measurement were detected, respectively. A nonlinear increase in $CH_4$ concentrations, that strengthened over time was found in 3% of the measurements and 2% of the measurements had an inconsistent and abruptly changing concentration trend. Low concentration changes, showing no clear trend, and a linear decrease in $CH_4$ concentrations, together, were represented by less than 1% of the measurements. From the Siikaneva data set, we selected 12 measurement examples so that each measurement scenario was represented at least once in the visual QC exercise (Table A1).

For each measurement scenario, we estimated the variability in flux estimates introduced by different researchers choosing different time periods within the same measurement for flux calculation using the coefficient of variance (CV) across the fluxes calculated for each survey participant. To extrapolate this variability to a representative data set (the presented fluxes were chosen to capture the range of observed behavior, rather than represent the observations as explained above), we calculated the weighted sum of the CVs based on the relative occurrence of each measurement scenario within the Siikaneva data set (Table A1). To assess the variability in QC procedures, we extrapolated the percentage of measurements kept for flux calculation to a representative data set for each participant, again using the relative occurrence of each measurement scenario within the Siikaneva data set. We then calculated the CV between the percentages of measurements kept across all survey participants.

## 3 Results

A total number of 36 expert researchers participated in the survey. All of them completed the survey parts on demographic information and their field sites for flux measurements. Most participants (35) answered the questions concerning their flux measurement setup, and 30 responded about their flux calculation and QC approach. Participation decreased to 28 experts for the visual QC part and an additional two participants dropped out after the second example measurement, resulting in a survey completion rate of 72%.

### 3.1 Results of: Survey part 1 - The survey participants and their chamber measurements

#### 3.1.1 Demography

The survey respondents work for universities (25 participants), research centers (11 participants), or companies (one participant) that are located in North America, central and northern Europe, and eastern Asia (Figure 1a). Most (89%) of the participants have a PhD title, 41% of whom completed their PhD within the last seven years, 25% between 7 and 15 year ago, and 34% more than 15 year ago (Figure 2a). Nearly all (94%) of the participants are researchers, two of whom are PhD students (Figure 2b). One participant each specified their current position as Bachelor student, professor, leader in industry, coordinator, and consultant, respectively. With 58%, the majority of the survey participants has a background in Geosciences, followed by biology (25%), ecology (11%), meteorology (8%), environmental sciences (6%), and physics (6%). One participant each has a background in forestry, biogeosciences, and agricultural sciences. Half of the participants (52%) are part of one or several of the flux networks and databases FluxNet, ICOS, AmeriFlux, OzFlux/TERN, European Fluxes Database Cluster, and LTER.

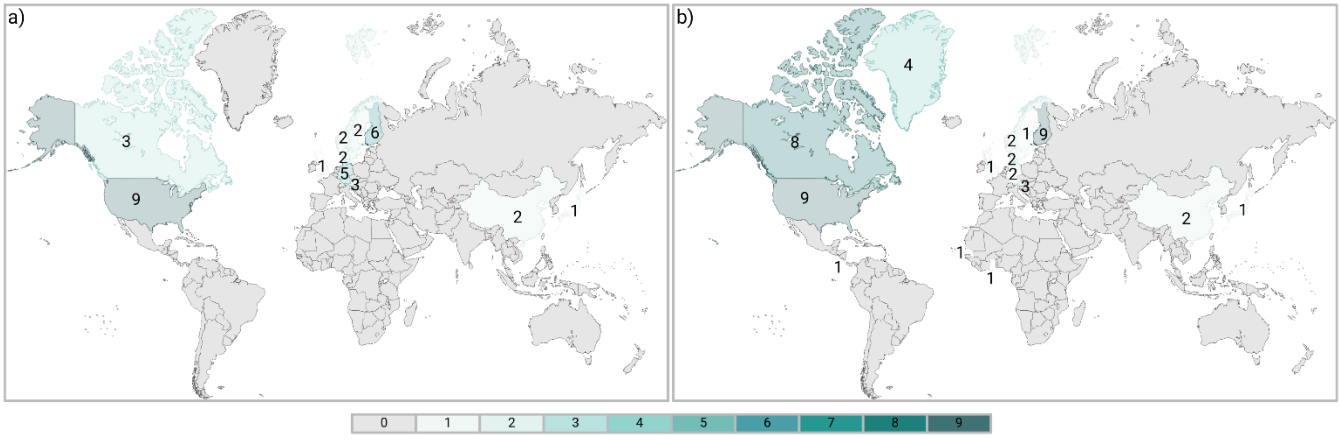



**Figure 1: Countries of the main institutes (a) and of the research sites (b) of the participants. Some participants gave multiple answers regarding the country of their research sites, causing the total number of responses to exceed the total number of 36 participants. This figure was created in BioRender.**

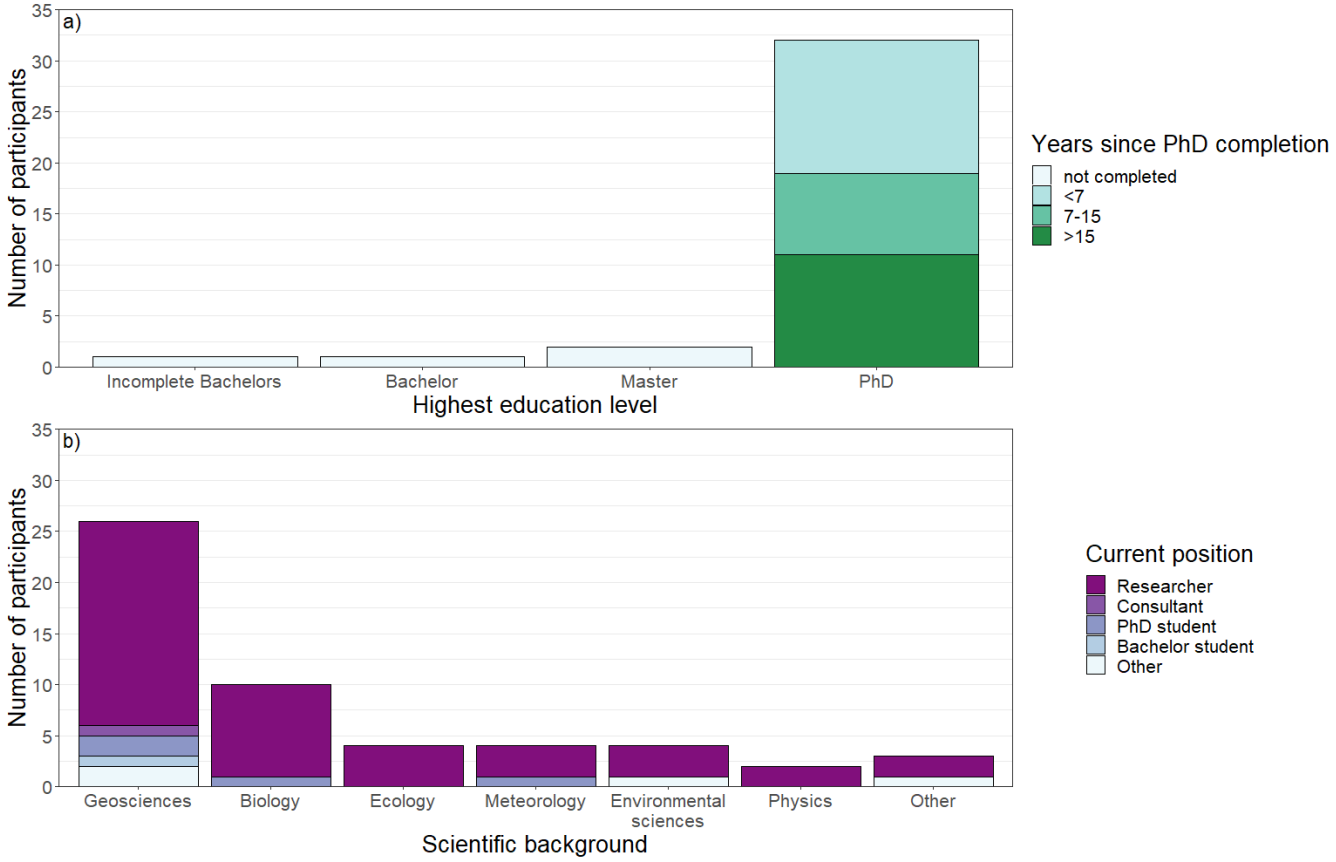



**Figure 2: Histograms of the highest education level of the participants, split by the years since their PhD completion (a) and of their scientific background by current position (b). Some participants gave multiple answers regarding their scientific background, causing the total number of responses to exceed the total number of 36 participants.**

### 3.1.2 Flux measurement sites

Most (83%) of the participants do field measurements in the same country as their home institute, among them all participants working for institutions in Asia, Canada, Finland, Norway, Denmark, Austria, and the United Kingdom (Figure 1b). Four participants additionally reported field measurements in Greenland and one participant in Ghana, Costa Rica, and Senegal, which were not among the countries of home institutes of the participants. Six participants from the US, Germany, and Sweden had their main research sites in Canada, Finland or Greenland, according to their research questions and ecosystems of interest. The majority (83%) of the participants focus their research on peatlands and wetlands, mainly fens or bogs (50%), and littoral wetlands (31%) (Figure 3). A few (14%) of the participants measure in (semi-)arid regions, upland areas, and at sites with mineral soil instead of or in addition to wetlands. Some (33%) of the participants explicitly mentioned field measurements in permafrost-affected landscapes; similarly, 33% of the participants explicitly mentioned that they measure in “northern”, “boreal”, “arctic”, or “subarctic” regions and 6% measure in “alpine” or “subalpine” terrain. Some

(25%) of the participants do aquatic measurements and 19% measure at anthropogenically managed sites such as on agricultural land, in drained and in rewetted peatlands. Specific ecosystems researched by two participants are rice paddies and reed ecosystems.

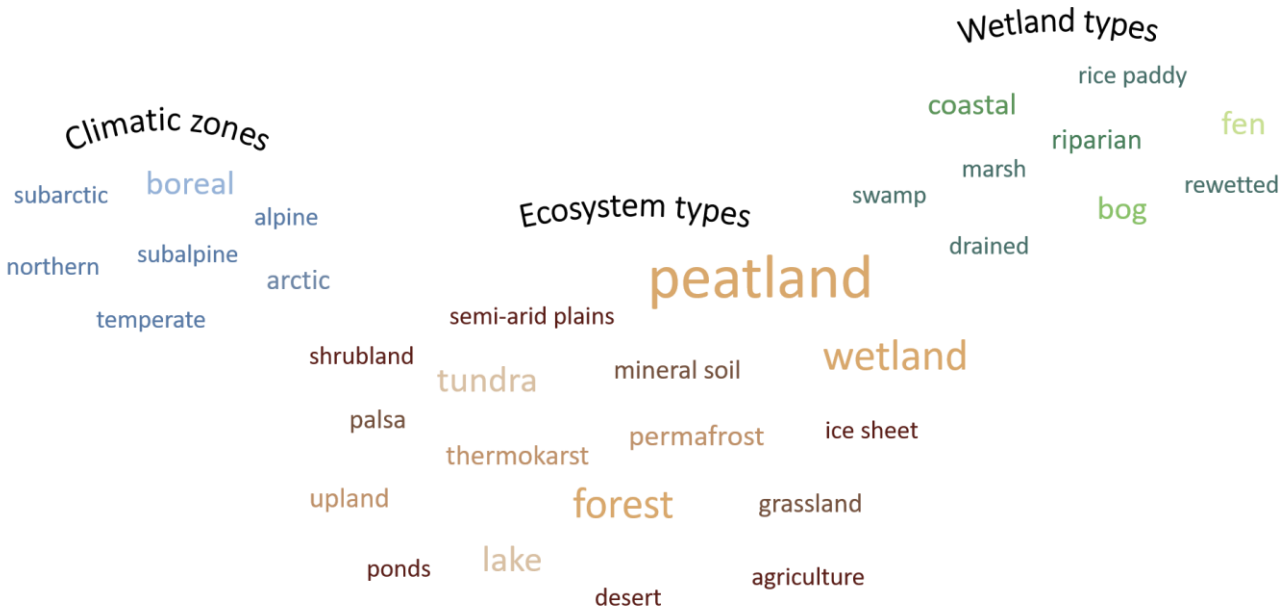



**Figure 3: Word clouds of the study areas representing the climatic zones of the study sites, the studied ecosystem types, and specifying the types of wetlands and peatlands that are researched by the participants.**

### 3.1.3 Research goals

The overarching research goals that the survey participants address with their flux measurements are to better understand the processes involved in greenhouse gas cycling, to better understand and quantify the effect of changes on greenhouse gas dynamics, to estimate greenhouse gas budgets, and to research the methodology for gas flux measurements. To investigate the environmental and ecological controls on the greenhouse gas exchange is the main goal of 28% of the participants, mainly in peatlands and wetlands and considering environmental conditions, vegetation properties, and the microbial community among others. The main aim of 53% of the participants is to understand and/or to quantify the effect of natural and anthropogenically induced change on greenhouse gas dynamics. The changes considered involve climate change, more specifically, warming, vegetation changes, elevated atmospheric $CO_2$ concentrations, permafrost thaw, and intensifying disturbances, such as wildfires, as well as peatland management, land-use change, and oil and gas exploration. Estimating greenhouse gas budgets is the goal of 22% of the participants but this goal varies in spatial and temporal scales from annual budgets of northern ecosystems to budgets of wetlands, microseepage, i.e. diffusive $CH_4$ fluxes over productive hydrocarbon basins, as an estimate of natural geologic $CH_4$ emissions, or permafrost and periglacial ecosystems, including thermokarst lakes, thawing permafrost peatlands, and degrading subaqueous permafrost. One participant uses the flux measurements to research methodologies for

gas flux measurements, investigating their accuracy, minimum detectable fluxes, curve fitting approaches, as well as engineering challenges around automation and minimizing measurement artefacts.

### 3.1.4 Flux measurement setup – Guidelines and implementation

There are several guidelines on best practices for chamber measurements that involve recommendations on the chamber setup (e.g. de Klein and Harvey, 2012; Fiedler et al., 2022). The aim of these guidelines is to keep the flux between soil, vegetation and chamber headspace as close as possible to the "real" flux that would be found in the absence of a chamber. This is achieved by minimizing chamber-induced artefacts. Such artefacts include an increasing deviation of environmental conditions inside the chamber from the ambient conditions over the time of the chamber closure and a disturbance of the system during the chamber placement. These chamber effects are reduced by equipping the chamber with additional features such as a vent and shading or active cooling to avoid a pressure and temperature change inside the chamber, respectively, a fan for mixing to avoid the buildup of a stable layering within the chamber (Clough et al., 2020). At the same time, the influence of remaining chamber artefacts can be reduced by a balanced combination of closure time and chamber dimensions and the remaining influence of chamber artefacts can be assessed depending on sampling frequency and additional variables measured. The efficiency of chamber setup recommendations at avoiding chamber artefacts have in part been demonstrated by experimental or modelling studies (e.g., Hutchinson and Livingston, 2001; Pumpanen et al., 2004). Our expert survey revealed that researchers use different instrumental setups, most of them implementing the recommended measures (Figure 4, Figure A2).

*Pressure vent*

A gas flux into or out of a closed chamber, would slowly alter the air pressure inside the chamber over time as well as more rapidly when the chamber is closed. As such as change in pressure can affect the gas flux between soil and chamber, it is recommended to install a vent, that is a small opening in the chamber, that allows for pressure equilibration but does not allow for significant mixing of ambient air into the chamber to keep the pressure inside the chamber close to ambient air pressure. Clough et al. (2020) recommend the simultaneous use of two types of vents as they tackle different pressure-related chamber artefacts – a larger one that is open only during chamber placement and a smaller one that remains open during the measurement. Vents for pressure equilibration are only used by half of the participants (Figure 4). Different methods for pressure equilibration employed by the respondents were a hole in the chamber that is sealed after chamber placement, explicitly mentioned by two participants, and a long line of tubing that is constantly open to the atmosphere allowing for pressure equilibration while preventing that too much ambient air enters the chamber, explicitly mentioned by one participant. The responses indicate that the two types of vents are considered rather as alternatives for vent designs than as two measures that tackle different pressure-related chamber artefacts and that should therefore both be applied simultaneously. One reason for the low implementation rate of pressure vents could be a fear of causing a so-called Venturi effect, where wind passing over the vent outlet can depressurize the chamber, leading to an increased gas flow from the soil into the chamber (Bain et al.,

2005; Conen and Smith, 1998). Clear guidelines exist however on how to avoid the Venturi effect by adjusting the vent design (Xu et al., 2006).

*Cooling*

Especially in summer, when air temperatures are high, a transparent chamber might act as a small greenhouse, causing the temperature inside the chamber to rise and increasingly deviate from the ambient air temperature over the time of the chamber closure, inducing a temperature gradient between the interior and the exterior of the chamber. A change in chamber temperature should be avoided as it can affect the gas flux through influencing processes like plant processes and evaporation or condensation. About one fifth of the survey participants addresses this issue in their chamber setup. As a way to avoid a temperature increase by insulation, 3% of the participants use opaque or reflecting chambers. Some applications, however, require the use of transparent chambers. This is the case for example when determining NEE. Furthermore, blocking out the incoming radiation can potentially reduce active $CH_4$ transport through plant aerenchyma thereby reducing the measured $CH_4$ emissions (Clough et al., 2020). 17% of the respondents therefore use active cooling of a non-insulated, transparent chamber. Types of cooling systems mentioned were Peltier elements, circulation of the chamber air through a tank filled with ice-water, and fans circulating the cold air from ice packs placed inside the chamber. However, an active cooling of the chamber air bears the risk of overcompensating for a temperature increase and causing condensation inside the chamber or sampling tubes (Fiedler et al., 2022). It is therefore recommended to use active cooling only if chamber cannot be insulated and/or if long chamber deployment periods are needed (Maier et al., 2022). The effectiveness of insulation or cooling should be evaluated by comparing surface soil temperatures inside and outside the chambers (Clough et al., 2020).

*Chamber pressure and temperature measurements*

Recording the temperature and the pressure inside the chamber over the time of the chamber closure is essential for correcting for temperature and pressure using the ideal gas law when calculating $CH_4$ fluxes as well as for detecting remaining changes in pressure and temperature over time that could not be eliminated with a pressure vent and insulation or cooling of the chamber headspace. Most participants record the temperature inside the chamber, while only a little less than one third of them measure the chamber pressure with measurement frequencies ranging from every second to once per chamber closure. While one temperature and pressure measurement during chamber closure might be sufficient for use in the ideal gas law, higher frequency measurements are needed in order to consider the stability in environmental conditions inside the chamber as an indicator of flux quality. Only two participants can therefore account for temperature and/or pressure changes over the time of the chamber closure by individually correcting each concentration measurement as they document the chamber temperature and/or pressure at the same frequency as the gas concentrations. Most notably, almost one fifth of the survey participants does not measure the chamber temperature at all (Figure 4), which can lead to large uncertainties considering the strong linear effect of temperature on the flux magnitude through the ideal gas law.

*Mixing*

In the absence of air movement in a closed chamber, a concentration gradient can develop inside the chamber, which might influence the further gas flux between soil and chamber headspace. A well-mixed headspace is furthermore needed to ensure that a representative gas sample can be taken. While most researchers use fans to mix the air inside their chamber, some researchers argued that the air flow from circulation through a closed loop with the gas analyzer was sufficient to mix the chamber air so that particularly small chambers did not need a fan. This statement highlights that further research is needed to investigate the strength of turbulence that is adequate for particular chamber dimensions to ensure proper mixing of the chamber air while preventing that additional gas is artificially released from the soil (Christiansen et al., 2011; Maier et al., 2022).

*Seal*

To reliably quantify the momentary gas exchange between a defined soil surface and the atmosphere, the mixing of chamber air with ambient air needs to be avoided. To achieve this, it is recommended to insert a chamber base into the ground to restrict lateral gas transport inside the soil and to additionally ensure an airtight connection between the chamber and its base (Clough et al., 2020). Two thirds of the participants follow this recommendation and place their chamber on top of a base that they inserted into the ground between one hour and one year before the measurement. The more time that passes between base insertion and measurement the less a potential disturbance of the ground and its concentration gradient will affect the measurement. The fact that the chamber setups employed by one third of the participants do not involve a collar or a seal might be less problematic than it appears since many participants measure in wetlands or on open water and the required insertion depth of the chamber into the soil as well as the necessity of a gastight seal are low under water saturated conditions and at low soil porosities (Clough et al., 2020). Two thirds of the participants aim to make the connection between the chamber and the collar or the soil gastight by using one or several types of sealing. Besides gaskets and water seals, a plastic sheet weighed down by a chain, a stocking filled with sand, and foam in the collar groove were mentioned as sealing methods. Every chamber setup should be tested for gas tightness before it is deployed in the field, as suggested by Clough et al. (2020).

*Chamber dimensions*

One challenge in developing chamber measurement protocols is to find a balance between a chamber closure time that is short enough to keep the influence of chamber artefacts low but long enough to reach gas concentrations within the chamber headspace that are above the detection limit of the gas analyzer or gas chromatograph used. One way to reduce the minimum time of chamber closure required to exceed the detection limit of the instrument is through reducing the chamber volume.

The volume of the chambers used by the participants ranged from 8 to 1800 L with a median of 64 L and an interquartile range (IQR) of 105 L. 93% of the chambers used are smaller than 260 L (Figure A3). However, more specific recommendations exist on the chamber dimensions besides requirements on its overall volume: To minimize the error caused by potential leakage

and maximize the area sampled, an area/perimeter ratio of ≥ 10 cm is recommended, which equates to a diameter of ≥ 40 cm for a cylindrical chamber. Two thirds of the chambers used by the survey participants respect this recommendation and the majority (75%) of chambers with a smaller-than-recommended area/perimeter ratio are cylindrical. Furthermore, a chamber height to deployment time of ≥ 40 cm $h^{-1}$ is recommended to maximize the flux detection while minimizing the perturbation of environmental variables. This recommendation is followed in 93% of the measurement setups used by the participants. The two remaining setups had too long closure times considering the relatively flat/low chambers. However, flexibility in chamber dimensions and closure time is often limited by the specific conditions of the research site: The minimum closure time needed depends on the flux magnitude of the gas of interest and on the sensitivity of the analyzer and the chamber height has to be chosen to accommodate the vegetation while its area might have to be adapted to the surface structure.

*Sampling techniques and chamber closure times*

Besides reducing the chamber volume, increasing the measurement frequency of the gas concentrations can reduce the required chamber closure time as in most researched environments $CH_4$ emissions are high enough so that the minimum detectable flux is reached rather quickly. Much higher sampling frequencies can be achieved through the use of in-line gas analyzers as opposed to manual sampling of the chamber headspace. The majority of the survey participants use an in-line gas analyzer for continuous and on-site measurements of the gas concentrations inside the chamber (Figure 4). All but one of these participants employ a closed sample loop which returns the air to the chamber after circulation through the gas analyzer. One participant uses open-path LI-COR gas analyzers installed inside a large chamber. The gas analyzers used by the respondents record the gas concentrations at frequencies between five times per second and once every 15 s. The chamber measurements therefore use shorter closure times of 0.5 to 12.5 min compared to the closure times of 16 to 50 min used by the fewer participants who manually sampling the chamber air every 4 to 10 min (Figure A4). Two participants using manual sampling keep their chamber closed for more than 40 min which is considered as too long by Rochette and Eriksen-Hamel (2008) while earlier guidelines allowed for up to 1 hour closure time (Holland et al., 1999). To avoid overly long closure times that promote chamber effects on the measured fluxes, the minimum required closure time should be determined considering the minimum detectible flux (MDF) based on the sensitivity of the analyzer and the chamber height (Christiansen et al., 2015; Nickerson, 2016).

An additional advantage of in-line gas analyzers over manual sampling, besides reducing the relevance of chamber artefacts through shortening closure times, is that the higher temporal resolution of the gas concentration recordings can reveal remaining chamber artefacts. This enhances the possibilities to evaluate the quality of a flux estimate or to exclude measurement periods affected by chamber artefacts at the stage of flux processing. In-line gas analyzers furthermore allow for the use of chambers that open and close automatically. Such automated chambers are used by one third of the survey respondents. While being more cost-intensive than manual chambers, automated chambers allow for continuous measurements at a higher temporal resolution.

The precision of the measured gas concentrations might differ between the survey participants as they calibrate their gas analyzers or gas chromatographs at different time intervals: Most respondents (58%) calibrate their instruments once per year and 24% do so once before each measurement campaign. A few (10%) of the participants calibrate the instrument less often e.g. when serviced every 1 to 3 years and 12% calibrate more frequently ranging from weekly to daily to calibration after each flux measurement.

*Reducing anthropogenic disturbance*

The survey participants take various precautions to minimize additional disturbances to their chamber measurements that can be caused by the presence of those who measure and their way to operate the chamber system. For wet, terrestrial sites, 28 participants stand on more stable ground while measuring, either by using permanently or temporarily installed boardwalks or wooden boards or by choosing a drier patch or a rock to stand on. Six participants furthermore mentioned that they make sure not to walk close to the measurement plots by using automated chambers or walking rules supported by warning tape. For aquatic measurements, participants avoid anthropogenic disturbance of the sediment and thus of the gas release by pulling the chamber into its measurement location with a rope or sitting in a boat while measuring. In addition, careful placement of the chamber, training of those who measure, maintenance of collars and sealing, and carefully keeping the vegetation away from the chamber sides were used to minimize disturbances to the chamber measurements.

*Ancillary data*

Recording additional variables alongside the chamber measurements can help explain the observed gas fluxes as well as identify potential disturbances to the measurements. The variety of variables measured by the survey participants (Figure 5) might indicate that, depending on their background and research questions, scientists consider different variables as important in controlling $CH_4$ fluxes. Almost all survey participants measure variables to characterize the soil, hydrological, and meteorological conditions, covering most of the ancillary data suggested by Maier et al. (2022). The potential effects of the vegetation cover were however considered by less than one sixth of the respondents.

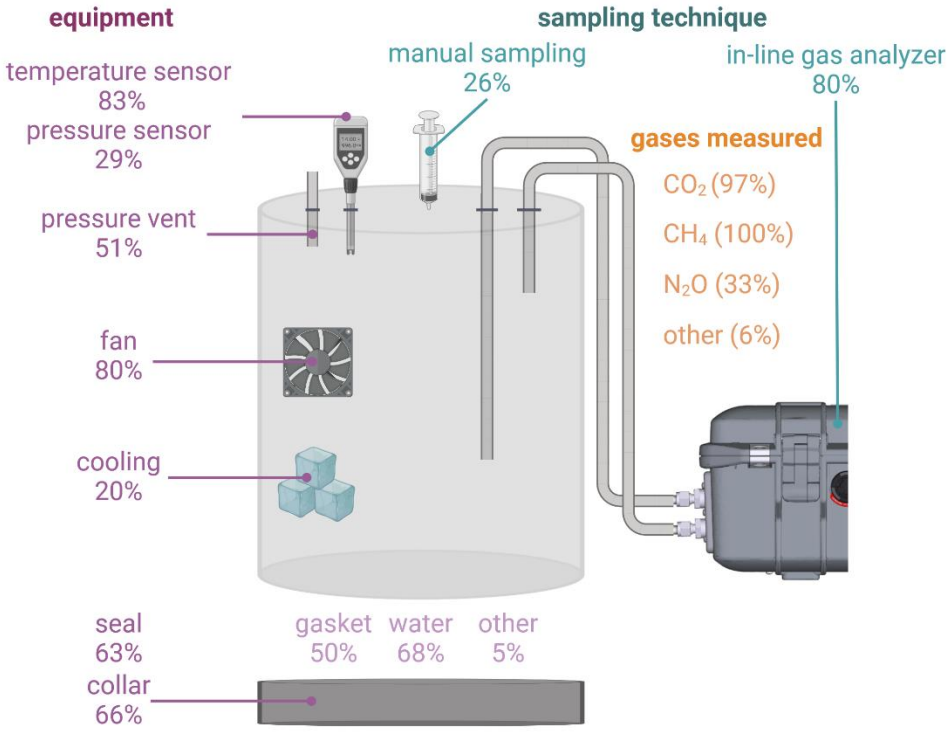



**Figure 4: Schematic chamber setup including the percentage of survey participants using certain types of chamber equipment and sampling approaches for the gas concentrations and measuring different greenhouse gases. Some participants use both manual sampling and in-line gas analyzer measurements at different applications, research sites or measurement campaigns, causing the total share of measurement methods used to exceed 100%. Other gases besides $CO_2$, $CH_4$, and $N_2O$ measured by one survey participant each are ethane and BVOC. This figure was created in BioRender.**

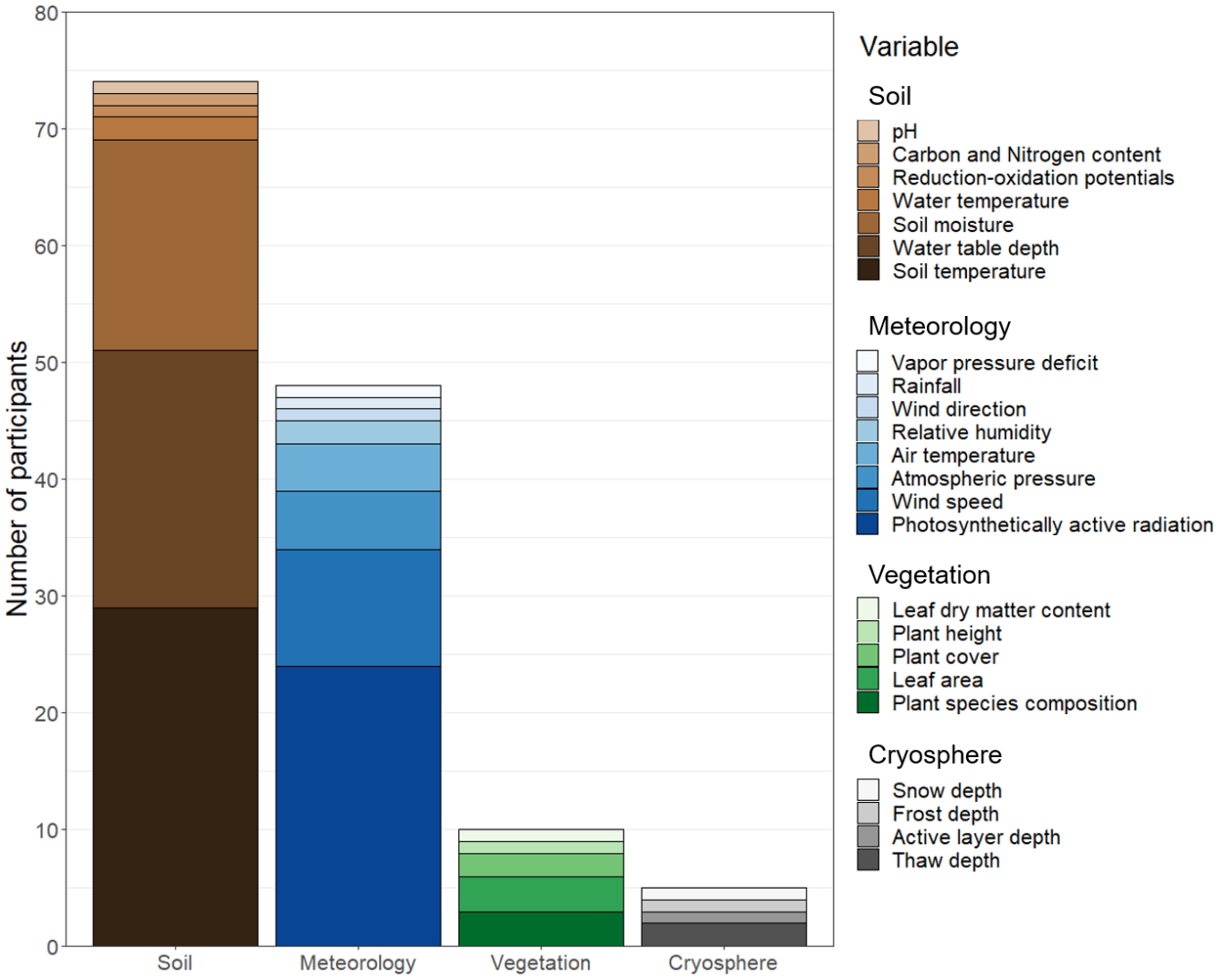



**Figure 5: Ancillary data recorded alongside the gas fluxes. Participants were permitted to select multiple variables, allowing the number of responses within one category of ancillary data to exceed the total number of 36 survey participants.**

### 3.1.5 Flux calculation and QC approaches

The qualitative responses on calculation approaches for $CH_4$ fluxes revealed differences in the flux processing and QC procedures that might result in considerable variation in the $CH_4$ fluxes among researchers. Gas fluxes are generally estimated from chamber measurements as the slope of the change in gas concentration over the time of the chamber closure and accounting for the water vapor concentrations, the temperature, and the pressure inside the chamber as well as for the chamber dimensions. This approach was modified by the survey participants mainly through selecting a time period of each chamber measurement for flux calculation, choosing a fit function to estimate the change in concentration over time, and determining

the accuracy of the temperature and pressure correction by selecting a measurement frequency for the two variables or deciding to use standard values instead (Figure 6). The majority of the participants (90%) use self-written scripts and functions for their flux calculation while 20% of the participants at least partly use existing and published R or Matlab scripts.

Selecting a time period within a chamber measurement for flux calculation, for many respondents, involves discarding the beginning of each measurement to exclude initial disturbances caused by the chamber placement. Most participants use a linear fit to estimate the change in gas concentration over the time of each chamber measurement. Most remaining respondents compute both a linear fit as well as the initial slope of an exponential fit, either deciding for one based on the goodness of the fit or using the difference between the two slopes as an uncertainty estimate for the final flux value. One participant each uses an exponential fit on all chamber measurements, considers the total change in gas concentration as the difference between the gas concentrations at the start and at the end of the chamber closure, or averages multiple linear fits on a one-minute window moving over the measurement at steps of 10 s.

In the step of correcting the measured gas concentrations for the temperature and pressure inside the chamber, most participants use one temperature value per chamber closure, that is either measured during one point of the chamber measurement or derived as the average of several temperature recordings over the time of the chamber closure. As less participants measure the pressure compared to the temperature inside the chamber, more have to rely on ambient pressure recordings or assume standard atmospheric pressure. As opposed to assuming constant conditions over the time of the chamber closure, two participants explicitly stated that they individually correct each gas concentration measurement for the chamber temperature and/or pressure measured at or interpolated to the same frequency as the concentration measurements.

Various approaches for QC of the flux estimates were mentioned by the participants. Most participants manually check each of their chamber measurements, while others use an automated procedure and some used a combination of both manual and automated diagnosis. Most participants use measures of the goodness of fit to evaluate the quality of their flux estimates, some of whom consider fixed cut-off values of these metrics that decide between keeping or discarding a flux measurement. Apart from two participants, the respondents typically discard up to 5% of their data.

The uncertainty of each individual flux estimate is assessed by 57% of the respondents, most of them using metrics for the goodness of fit or the variability between spatial or temporal measurement replicates. One participant each uses the difference between the slopes derived from a linear compared to an exponential fit and the variation in several one-minute linear fits in a moving time window as an uncertainty estimate, respectively.

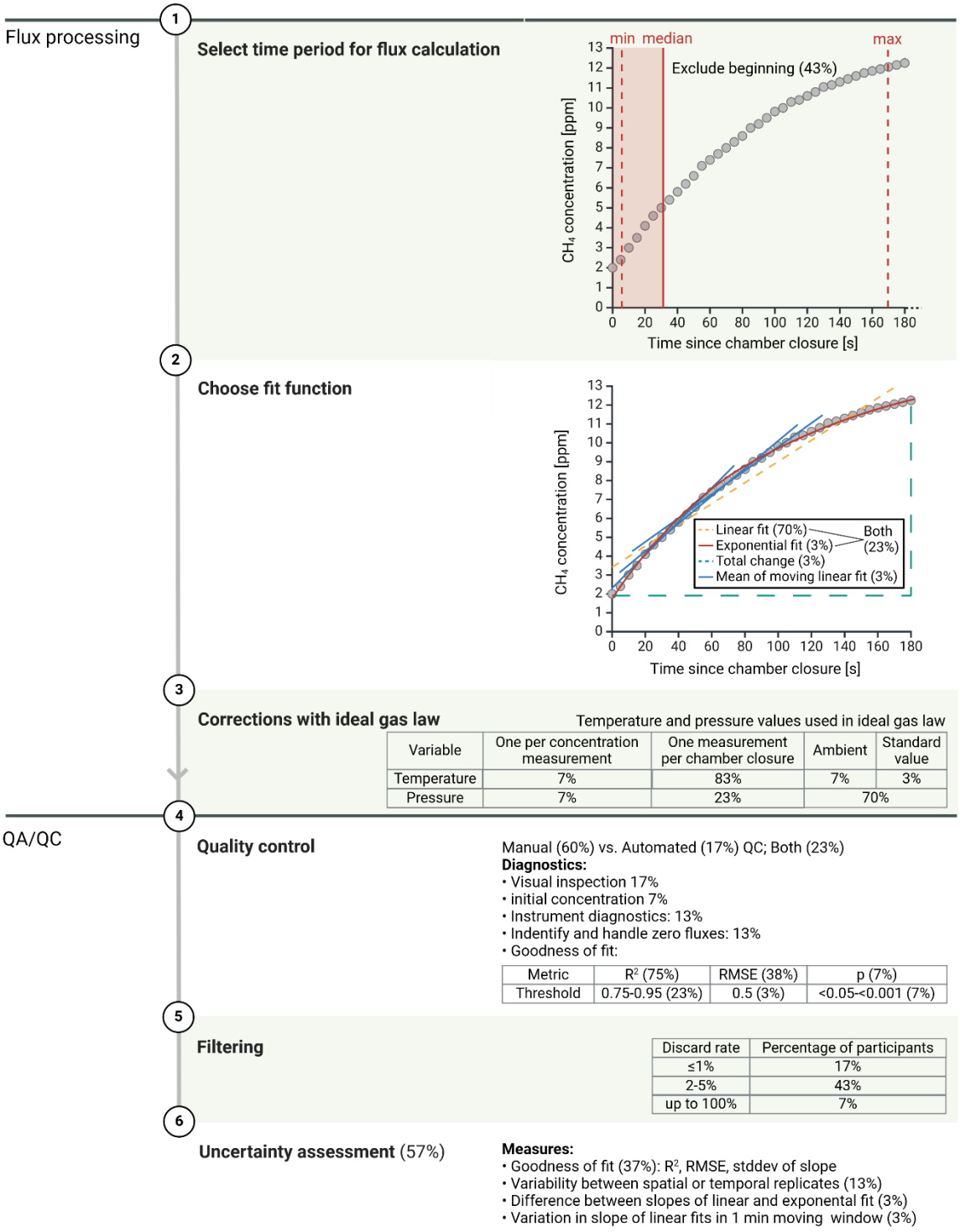



**Figure 6: Differences in the workflows used for flux processing, quality control (QC), and quality assurance (QA) by the survey participants. This figure was created in BioRender.**

### 3.2 Results of: Survey part 2 - Visual quality control of a standardized data set

The visual QC exercise revealed that the handling of the measurement examples (decision to keep or discard a measurement and choice of time period for flux calculation) differed between the survey participants depending on their interpretation of the $CH_4$ concentration change in the chamber headspace over time (Table 1). Depending on the shape of the concentration curve (linear or nonlinear), the choice of the time period used for flux calculation furthermore had a strong impact on the magnitude and in one case even on the direction of the estimated $CH_4$ flux (Figure 7, Table A2). Detailed descriptions of the individual measurement examples and their handling by the survey participants can be found in Text A1.

*Linear fluxes: Emission and uptake*

The majority of the participants (91%) decided to keep the measurements that showed a linear increase in $CH_4$ concentrations for flux calculation. Due to the linear behavior, these flux estimates were least affected by the time period that was chosen for the linear fit.

The latter also applied to measurement example showing a linear decrease in $CH_4$ concentrations over time. However, more participants decided to discard the entire measurement because they did not expect to find net uptake of $CH_4$ at a wetland site. The free text responses revealed that the conditions, and in particular the water table depths, under which net uptake of $CH_4$ can occur were debated among the participants.

*Nonlinear increase - decreasing slope*

Most participants (79%) also kept the measurement examples that showed a consistent but nonlinear and weakening increase in $CH_4$ concentrations over time. Here, the magnitude of fluxes estimated from the nonlinear concentration change strongly depended on the time period selected for the flux calculation. The selection of the time period in turn was influenced by how the participants explained the observed nonlinearity. There were two main reasonings among the participants for their choice of the time period with opposing effects on the flux magnitude: (1) About two thirds of explanations for the nonlinear behavior assumed that the increase in $CH_4$ concentrations was weakened by either $CH_4$ saturation of the chamber headspace or leakage of air from the chamber towards the end of the measurement. The participants concluded that this latter part of the measurement was disturbed and should therefore be excluded from the flux calculation, which resulted in higher flux estimates. (2) Conversely, the remaining third of explanations assumed that the stronger increase in $CH_4$ concentrations at the beginning of the measurement was caused by an initial disturbance such as ebullition, triggered by the chamber placement. A consequent exclusion of the strong initial increase in $CH_4$ concentrations from flux calculation resulted in lower flux estimates as the lower slope during the latter part of the measurement was preferentially selected.

*Nonlinear increase - increasing slope*

The survey participants were similarly divided on the appropriate handling of chamber measurements that show a nonlinear increase in $CH_4$ concentrations but with an increasing slope over time. Accordingly, half of the participants who discarded this measurement argued that they cannot justify choosing a time period for flux calculation as they cannot explain the observed shape in $CH_4$ concentrations and, considering the nonlinearity, an unsubstantiated selection of a time period could strongly bias the flux estimate. For those who kept the measurement and gave start and end times for flux calculation (65% of participants), the time period chosen significantly affected the flux estimate. This range between higher and lower flux estimates again resulted from contrasting explanations of the nonlinear concentration change: Higher flux estimates originated from explanations assuming an initial period of adjustment and disturbance caused by the chamber placement, through exclusion of the initial, lower slope in $CH_4$ concentrations. On the other hand, explanations involving chamber effects on $CH_4$ cycling processes through alteration of environmental conditions or interference of $CH_4$ measurements with high $H_2O$ concentrations led to lower flux estimates due to exclusion of the stronger increase in $CH_4$ concentrations towards the end of the measurement.

*Jumps*

The majority of the respondents (65%, 88%, 92%) interpreted the jumps showing in three of the measurement examples as episodic events of ebullitive $CH_4$ emission while one participant suggested a malfunctioning of the gas analyzer. The survey responses revealed uncertainty around the question under which water table conditions $CH_4$ ebullition is most likely to occur, indicating a fundamentally different understanding of the causes of ebullition evens among the participants. There were two major considerations concerning $CH_4$ ebullition during chamber measurements: First, the survey participants disagreed on whether ebullition events should be included in flux estimates from chamber measurements or if diffusive and ebullitive flux should be quantified separately, either by isolating periods of ebullitive and diffusive flux in one concentration time series or by separately measuring ebullition, for example using bubble traps. When accounting for both diffusive and ebullitive $CH_4$ emission by using a linear fit over an entire measurement containing ebullition events, as suggested by 4 to 8% of the participants, flux estimates were up to five times as high as the ones considering the diffusive flux only. Second, the respondents disagreed on whether the remaining part of a measurement after an ebullition event could still be used to quantify the diffusive flux. More than half of the survey participants (54%) kept the linear part of a measurement after an initial ebullition event for flux calculation while 38% of the participants discarded the entire measurement. The latter assumed that the high $CH_4$ concentrations in the chamber following the ebullition event would decrease the concentration gradient and thus reduce the $CH_4$ flux between soil and chamber headspace for the rest of the measurement. This decision also influenced the range of flux estimates derived from a measurement with repeated ebullition events. Flux estimates from the 15% of participants who used a shorter linear increase in $CH_4$ concentrations before the first ebullition event were three times as high as the flux estimates from the 19% of participants who fitted the longer linear increase after the first ebullition event.

*Low variation*

Another source of uncertainty in data handling among the survey participants lay in the identification and handling of so-called "zero fluxes". Two thirds of the survey participants discarded the measurement example showing only very low variation in $CH_4$ concentrations without a clear trend over the time of the chamber closure. The other third of the participants submitted a flux estimate, 20% of whom set the flux to zero and 80% calculated the small positive flux resulting from a nonlinear fit. One participant remarked that the magnitude in $CH_4$ variations would need to be compared to the instrument precision to decide whether a measurement can be classified as a "zero flux".

*Inconsistent trend*

Less than one quarter of the respondents kept the measurement example showing a reversing trend in $CH_4$ concentrations over the course of the measurement. Of all measurement examples, the resulting flux estimates varied strongest between the participants in this case, that is by more than the mean flux value and including both positive and one negative flux estimate. This indicates that in cases where the trend in $CH_4$ emissions changes between an increase and a decrease over the time of the measurement, interpretations of the concentration time series can make the difference between net $CH_4$ emission or uptake.

*Further considerations*

The survey participants repeatedly mentioned several other reasons to discard a measurement besides an overall nonlinear or otherwise unexpected behavior in $CH_4$ concentrations. These reasons included too high initial gas concentrations, assumed leakage of air from the chamber headspace, and a too short measurement time. Furthermore, some participants considered the simultaneously measured $H_2O$ and $CO_2$ concentrations as additional indicators of measurement quality: Nonlinear or otherwise unexpected behavior as well as high initial concentrations of $H_2O$ and $CO_2$ were mentioned as reasons to discard a measurement. Similarly, interference of $CH_4$ measurements with high $H_2O$ concentrations towards the end of a measurement was mentioned several times as an explanation for nonlinear behavior in $CH_4$ concentrations. Linear changes in $CO_2$ and $H_2O$ on the other hand were considered a proof the air tightness of the system. To avoid any initial disturbance caused by the chamber placement from influencing the flux estimate, almost half of the survey participants by default excluded the beginning of each measurement (30 ± 85 s (median ± IQR)) from their flux calculations.

*Effect of different flux calculations on an example flux dataset*

Using the prevalence of different measurement scenarios in the Siikaneva data set (Table A1), we estimated an overall variability in the calculated $CH_4$ fluxes due to difference in time periods used for fitting as well as an overall variability in the inclusion/exclusion of measurements. Different researchers chose different parts of the same measurement for flux calculation (Figure A1.1 – A1.12), which resulted in an overall flux difference of 17% across the Siikaneva data set (Table A2). The variation in the percentages of measurements in the Siikaneva data set passing the visual QC was 28% (Table 1, A2). These

estimated variability introduced by the selection of different time periods for flux calculation compares with the mean natural temporal variability of 19% but is lower than the mean natural spatial variability of 88% calculated from automated chamber measurements of $CH_4$ fluxes in five temperate and Arctic peatlands by Pirk et al. (2016). Pirk et al. (2016) similarly found that both natural spatial and temporal variability in $CH_4$ fluxes exceed the difference between fluxes estimated using different fit functions. However, it must be noted that the uncertainty estimates derived in our study consider only the effect of differences in visual QC and do not account for different measurement setups or different fit functions used for flux calculation, both of which can add variability to fluxes (e.g., Pihlatie et al., 2013).

**Table 1: Explanations of the overall shape in $CH_4$ concentrations during chamber closure, reasons to discard, and reasons and approaches to keep measurements as given by the participants as well as percentage of respondents that kept measurements and coefficient of variance (CV) of the flux estimates derived from the survey responses by measurement scenario (For details see Table A2).**

| Measurement scenario | ID | Explanations | Reasons to discard | Reasons to keep & approach | Kept [%] | CV |
|---|---|---|---|---|---|---|
| Linear increase | VQC1<br>VQC2 | • Net $CH_4$ production & diffusive emission | • Small nonlinearities due to ebullition or saturation<br>• High initial $CH_4$ and concentrations<br>• No shading | • Consistent linear increase in $CH_4$ concentrations<br>• No indications of significant disturbances or malfunctioning of the instruments<br>• Close-to-ambient initial $CH_4$ concentrations | 91 | 2.5 |
| Nonlinear increase – decreasing slope | VQC4<br>VQC5<br>VQC9 | • Saturation<br>• Initial disturbance<br>• Leakage<br>• Changing environmental conditions<br>• Unsure | • Saturation<br>• No steady state reached<br>• Initial disturbance<br>• Leakage<br>• Changing environmental conditions<br>• Unclear which part of the measurement represents real flux | • No clear disturbance<br>• Nonlinear fit<br>• Use more linear part in the beginning<br>• Use more linear part at the end | 79 | 44 |
| Initial jump | VQC7 | • Ebullition caused by chamber placement<br>• Malfunctioning of gas analyzer | • High $CH_4$ concentrations affect concentration gradient | • Use linear part after the jump | 62 | 5 |
| Jump(s) | VQC8<br>VQC12 | • Ebullition caused by (anthropogenic) disturbance<br>• Malfunctioning of gas analyzer | • High $CH_4$ concentrations affect concentration gradient<br>• Ebullition affects pressure inside the chamber | • Use measurement before first jump<br>• Use linear part after the jump(s)<br>• Use longest linear part in between jumps | 40 | 42.5 |
| Nonlinear increase – increasing slope | VQC10 | • Initial period of mixing or adjusting<br>• Increase in chamber temperature over time<br>• Disturbance of measurement plot/ concentration gradient during chamber placement<br>• Chamber affects plant-mediated $CH_4$ transport<br>• Leakage | • Shape of curve unexpected and strong curvature makes flux estimate depend strongly on selected time period | • Nonlinear fit<br>• Use more linear part in the beginning<br>• Use more linear part at the end | 76 | 19 |
| Inconsistent trend | VQC11 | • Net $CH_4$ consumption | • No consistent trend of sufficient length | • Keep increasing initial part of the measurement | 29 | 138 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | • Net $CH_4$ uptake unexpected → Measurement issue:<br>• Gas analyzer issue<br>• Condensation<br>• Leakage<br>• Malfunctioning fan<br>• Initial disturbance such as ebullition caused by chamber placement | • Unclear which part to use for flux calculation because reason for pattern / timing of disturbance unclear | • Keep later, decreasing part of the measurement | | |
| Linear decrease | VQC3 | • Net $CH_4$ uptake<br>• Leakage<br>• High initial $CH_4$ concentrations | • Net $CH_4$ uptake unexpected in wetland<br>• Initial ebullition<br>• Anthropogenic disturbance<br>• Leakage | • Net $CH_4$ uptake possible | 50 | 17 |
| Low variation | VQC6 | • Production and oxidation balance<br>• "zero flux" (uncertainty > flux)<br>• Leakage | • Leakage<br>• Changing trend in $CH_4$ concentrations | • Manually set flux to zero<br>• Small but real flux exceeding instrument precision | 38 | 57 |

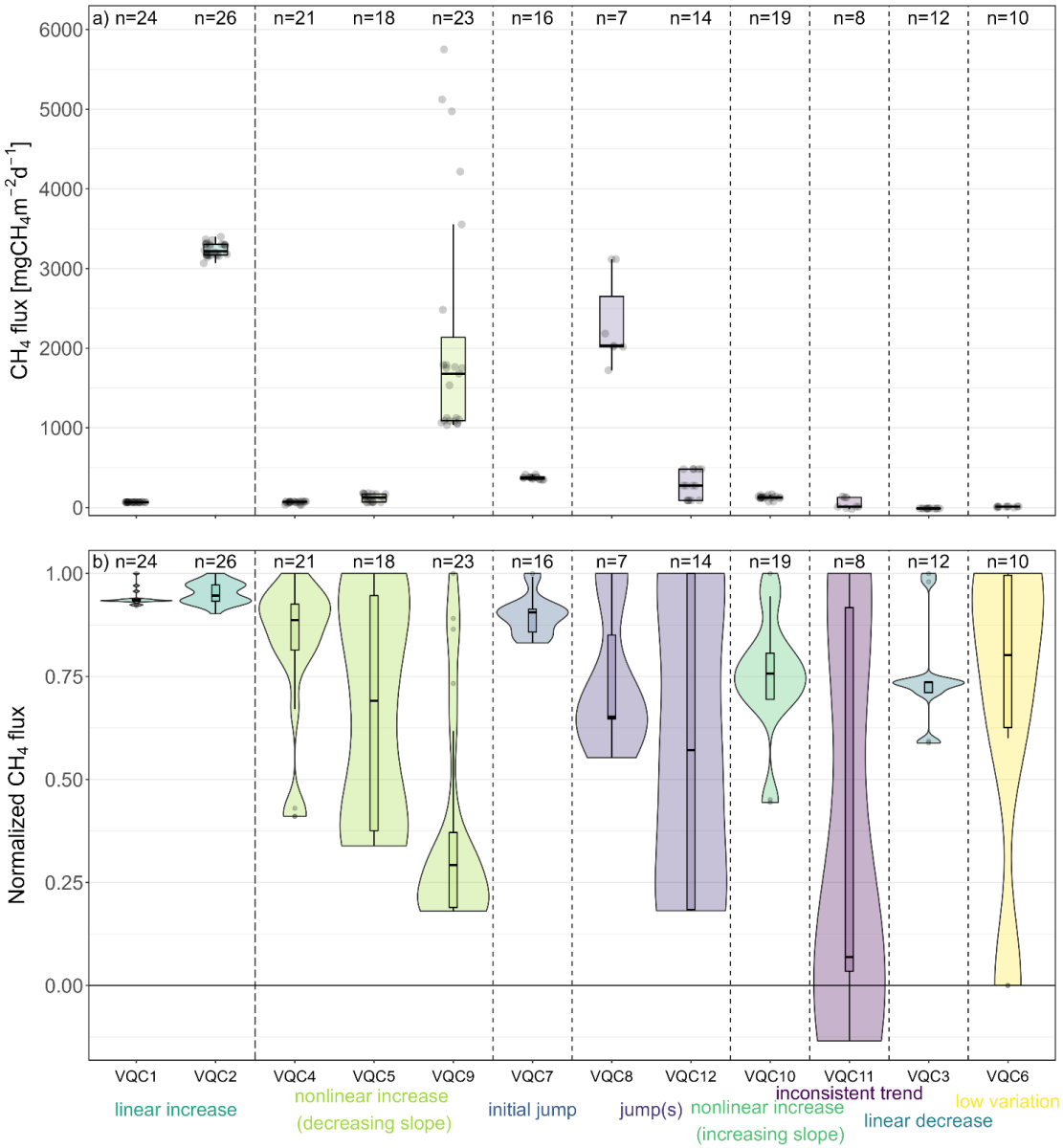



**Figure 7: Researcher variability in flux estimates for each measurement example in the visual QC exercise (VQC1 – VQC12) by measurement scenario. Range (a) and distribution (b) of flux estimates across the respondents. The number of survey participants (n) who contributed a flux estimate to the respective measurement example by selecting a time period for flux calculation is given on top of each boxplot (a) or violin (b), respectively. In (b), the flux estimates are normalized to the maximum flux estimate within each measurement example. The violins are scaled to all have the same maximum width. Violins crossing the zero line indicate that, for the respective measurement example, the selection of the time period for flux calculation made the difference between $CH_4$ emission and uptake.**

## 4 Evaluation of the survey methodology

### 4.1 Expert survey approach and insights

For our study, we used the method of an expert survey which allowed us to combine the accuracy of a literature review with the directness of an expert assessment. In general, a literature review might provide a more complete overview of the methods used and thus allow for more reliable statistical interpretation of the results. However, we found that published data sets and research articles involving chamber fluxes often lacked detailed information on measurement and data handling procedures – one of the current hurdles in interpreting, reusing, and combining existing chamber flux data sets. The expert survey, on the contrary, allowed us to obtain specific information directly from the scientists that might not be available in published literature but that might nonetheless significantly affect the $CH_4$ fluxes estimated from chamber measurements. In designing our survey, we had somewhat limited examples to follow as the approach of an expert survey rather than an expert assessment is not commonly employed. While the exact implementation of the survey could therefore surely be refined in future studies, we showed that surveying experts on their methods can be a useful approach and was strongly complementary to earlier reviews and recommendations of best measurement practices (e.g., Clough et al., 2020; de Klein and Harvey, 2012; Fiedler et al., 2022). The survey results clearly reveal that agreement on the measurement setup is high and generally in line with recommendations (Figure 4), but strong variability in the flux estimates is introduced at the data processing and analysis stage by the different researchers (Figure 7, Table A2). This provides an opportunity to re-focus the discussion from measurement setups and linear vs. exponential fitting approaches to a wider discussion about data workflows and uncertainty sources in chamber flux measurements that have emerged with new observational methods.

### 4.2 Representativeness of the survey respondents and questions

From the variety of survey responses, it becomes clear that evaluating the representativeness of the respondents of the chamber flux community as a whole is challenging. One reason is that the chamber flux measurement community remains less organized than the eddy covariance flux measurement community and is more fluid, potentially because the barriers for entry are lower, i.e. the cost of analysis. We recruited the survey participants from different places of employment assuming that this would make them rather independent in their choice of measurement and data handling approaches. The main strength of the collected data set therefore lies in representing a large range of measurement and data handling practices; indeed, there were substantial deviations in workflows within the part of the chamber flux community represented in this survey (Figures 4, 6). However, we did not reach all researchers using chamber fluxes with our survey; we likely underrepresented those working in agricultural ecosystems, disturbed sites, and tropical ecosystems. Overall, participants who had not encountered a certain shape in $CH_4$ concentrations in their own data sets before were more likely to discard the respective measurement example (Table 1, Table A2). For example, the measurement showing decreasing $CH_4$ concentrations over time was discarded by 50% of the current participants (Table 1), many of whom focus on wetland ecosystems (Figure 3), but is more likely to occur in

well-drained agricultural soils (Mosier et al., 1997). Thus, the background of the survey participants might have affected the outcome of the visual QC exercise with a bias towards expected (higher) fluxes.

Additionally, the question of number of survey participants is always a concern. While the number of researchers contacted (n=46) and the final maximum of 36 respondents might seem relatively low for a community survey, we estimate that this still represents a considerable extent of an estimated total number of several hundred chamber flux experts world-wide. Time is always a factor in voluntary survey participation; therefore, it was important to streamline questions to incentivize survey completion. In offering a diversity of question types, we attempted a balance between making the responses comparable and categorizable among the participants while still obtaining detailed information on their reasoning for the use of specific measurement and data handling techniques. The limited number of survey participants required a low number of possible responses in choice questions to allow for a meaningful statistical interpretation of the survey results; therefore, we used yes/no answers rather than scales of agreement. Yes/No questions further allowed us to draw conclusions on the prevalence of the implementation of recommended best measurement practices among the survey participants.

### 4.3 Assumptions in the flux calculations: site and researcher differences

Our estimates of researcher variability in flux data sets, derived from the visual QC exercise, strongly depended on the underlying reference data set collected at Siikaneva Bog (Table A1). Both natural processes and chamber-induced artefacts occur and their prevalence depends on both the environmental conditions of the research site as well as on the chamber design and measurement setup. Most measurements in the Siikaneva data set (~60%) showed the linear increase in $CH_4$ concentrations that is expected for an undisturbed measurement at a wetland site. However, a nonlinear, weakening increase in $CH_4$ concentrations was also represented by a rather high share of measurements (18%) and is also regularly observed at other sites (e.g. Pirk et al., 2016). The survey responses confirm that it is often unclear whether this shape is caused by an initial disturbance of the measurement or by $CH_4$ saturation of the chamber headspace over time (Table 1). Furthermore, this lack of process-understanding shows through in the high variance associated with the non-linear fluxes (Figure 7, Table A2). An initial disturbance, i.e. ebullition caused by the chamber placement, was a common explanation (Table 1) and might have occurred more frequently in the Siikaneva data set, than other sites, as roughly 60% of the measurements were obtained from vegetation removal plots. The removal of vascular plants and of the *Sphagnum* moss layer might have reduced both plant-mediated $CH_4$ transport and $CH_4$ oxidation, resulting in higher $CH_4$ concentrations in the pore water and thus increasing the probability of ebullition events (Jentzsch et al., 2024a). While $CH_4$ ebullition is a natural phenomenon often encountered in wetlands (Green and Baird, 2013), the increased probability of both natural and anthropogenically induced ebullition due to vegetation removal might have contributed to the high share of measurements (16%) showing abrupt jumps in $CH_4$ concentrations in the Siikaneva data set.

Although some measurement scenarios included in the visual QC exercise are relatively uncommon, it is still important to evaluate how these scenarios would be handled by different researchers as they showed large sources of disagreement (Table A2). Many survey participants stated that the nonlinear increase in $CH_4$ concentrations with an increasing slope over time was

unexpected. However, this shape was reported surprisingly often in other studies and occurred during several of our measurements (Table A1). Overall, this behavior of $CH_4$ concentrations in the chamber headspace is not consistent with diffusion theory (Kutzbach et al., 2007), indicating the influence of other processes. Similarly, both low changes in $CH_4$ concentrations without a clear trend and a decrease in $CH_4$ concentrations over time occurred infrequently in the Siikaneva data set (<1% of measurements) but were scenarios with high variability in the calculated fluxes (Table A2). Still, small fluxes might be expected at higher and drier wetland microtopographical features (e.g., Laine et al., 2007), while low, close-to-zero fluxes or $CH_4$ uptake are more commonly observed at upland sites (Virkkala et al., 2024; Voigt et al., 2023).

Overall, the Siikaneva data set might have contained more non-linear measurements than data collected by the survey participants due to the selected experimental setup as well as site-specific environmental conditions. This theory is difficult to test as this information is not often available for other sites but might be a reason for the high discard rate in the visual QC exercise. While the median percentage of measurements that the researchers said they discarded from their own data sets was 5%, they discarded 19% in the visual QC exercise when weighted by the prevalence of measurement scenarios within the Siikaneva data set (Tables A1, A2). Another reason for the high discard rate might be that the survey participants did not do the measurements themselves. They did not have the option to redo a measurement that they diagnosed as disturbed and they lacked an overall view of the dataset. Several participants mentioned that they would like to see the entire data set before deciding on keeping or discarding an individual measurement as they did not know the prevalence of the different measurement scenarios; the decisions for processing an entire data set might differ from the limited number of example measurements presented here. Processing the full dataset as a common dataset rather than a small subset would also eliminate the assumptions with the visual classification of measurement scenarios (Table A1); however, this might also have decreased the number of respondents as this is a relatively intensive exercise. If respondents did their own flux calculations, this would allow for non-linear fitting methods, which we did not use in our exercise despite being occasionally suggested by a participant (7% of responses). While our fitting and calculation approach may have been overly simplistic, post-hoc assumptions of how many participants would have used a non-linear fit and the different fitting options (such as exponential, quadratic, or logarithmic functions) would introduce substantial additional uncertainty into our estimates of researcher variability. Reproducing the calculation approaches of every respondent would have required additional, very detailed information from the survey participants, likely reducing the number of completed surveys and making our uncertainty estimates less representative of the entire chamber flux community. However, this type of exercise might be worth undertaking in the future.

## 5 Visions for improving chamber $CH_4$ flux measurements and data sets

### 5.1 Recommendations for high-frequency measurements of $CH_4$ fluxes from chambers

Earlier studies have highlighted variability in $CH_4$ fluxes due to chamber design and fitting approaches (e.g., Fiedler et al., 2022; Maier et al., 2022; Pihlatie et al., 2013; Pirk et al., 2016). Here, we show that many researchers have adopted the recommended measurement techniques and setup (Figure 4). The relatively widespread adoption of high frequency $CH_4$

analyzers provides new challenges and illustrates a need to move focus from measurement setup and curve fitting considerations to data handling as the disagreement in QC approaches varies widely among the survey participants and explanations for some observed behaviors remain inconsistent (Tables 1, A2). While broader discussions about QC approaches are warranted, some simple steps may help to improve data quality:

1) **Calculation and implementation of a minimum detectable flux given the analyzer precision and chamber height.** Use this to determine the measurement length and to determine when fluxes are below detection limit. Use short measurement times to avoid chamber effects.
2) **Do not discard fluxes, including ebullition fluxes, low fluxes or zero fluxes.** Instead, we should move towards a standardized QC flagging system. Ebullition fluxes, low and zero fluxes should be preserved and can be flagged in archived data. $CO_2$ concentrations can be used in addition to $CH_4$ concentrations to determine measurement quality (Pirk et al., 2016). This will work best in dark chambers as a net emission is expected. $H_2O$ vapor is less reliable as an indicator of flux quality.
3) **Report all data for archival purposes and implement data quality flagging.** A flagging system will indicate to others interested to re-use the data where uncertainties lie and has been implemented in eddy-covariance networks. Ideally, raw concentration data will be archived as well as processed data. This will allow reprocessing of data in the future as needed.

On the longer term, we need to develop new tools and networks to figure out how we can best leverage the new possibilities of high-frequency gas concentration measurements. Key steps are underway to allow easier operation, analysis, and standardization of flux calculation, for example the GoFlux Package for R (Rheault et al., 2024). In earlier times, ebullition was difficult to identify using GC analysis but can be seen in the high frequency concentration time series (Fig. A1.6-A1.8), allowing the separation of ebullition from diffusive fluxes (e.g., Hoffmann et al., 2017). The survey showed strong disagreement with how to handle these measurements, sometimes resulting in quite large variations in flux magnitude (Figure 7), suggesting that this new insight into $CH_4$ transport pathways is not fully utilized. Overall, more discussion and exploration about this crucial measurement approach is needed to fully leverage the technological developments of the past decade.

### 5.2 Establish a formal trace gas chamber flux network

One reason for the large variability in chamber methods revealed in this survey could be a lack of exchange between the researchers working with chamber measurements of $CH_4$ fluxes. Only half of the survey participants are part of a flux monitoring network, such as FluxNet or AmeriFlux, both of which strongly focus on eddy covariance measurements. This indicates that the exchange within the chamber flux community might be impeded by a lack of suitable networking platforms. Chamber technique-focused conference sessions and workshops to further develop approaches and revise methodologies would be beneficial. Further discussion and recommendations toward a more rigorous standardization of flux calculations by identifying the best fit based on objective criteria (e.g., Pedersen et al., 2010; Rheault et al., 2024) would be domain of such a network. While much work has already gone into developing chamber-based approaches and recommendations for

measurements, the substantial (and potentially novel) uncertainty in fluxes calculated among researchers here indicates that this matter is not yet settled (Figure 7). Furthermore, there was never complete agreement on whether to keep or exclude the fluxes included in the survey (Table 1).

Introducing a chamber flux network and data platform might speed adoption of a more standardized measurement protocol (although many recommended chamber components are widely adopted, Figure 4), improve metadata and ancillary measurements quality, spur development of a data quality flagging system that could foster a transparent exchange between researchers on measurement and data handling procedures and ultimately enhance the compatibility of individual flux data sets. Such a chamber flux network could build on existing research infrastructures such as the LTER sites, the ICOS sites in Europe and NEON sites in North America. Some examples for chamber databases have been developed like the Soil Respiration Data Base (SRDB), which includes chamber measurements of ecosystem respiration and has been widely cited, particularly for the response of ecosystem respiration to warming (Bond-Lamberty and Thomson, 2010; Bond-Lamberty and Thomson, 2010; Jian et al., 2020). This requires the open sharing of data, both raw chamber measurements and the calculated flux estimates using quality flagging rather than pre-filtered flux data sets. This way, reanalysis of existing data sets can be facilitated and all chamber measurements contributing to larger scale synthesis studies can be reprocessed using a uniform calculation and QC approach to remove differences among researchers, which we have shown to significantly affect $CH_4$ flux estimates (Figure 7).

### 5.3 Develop and adopt tools for uncertainty estimation in data processing

We demonstrated the potential of using a common data set to assess the variance in flux estimates caused by different data processing and QC approaches (Figure 7). This approach could also be implemented more broadly to build consensus on calculation methods, quality control, and data quality indices. The survey participants performed a detailed quantitative and qualitative evaluation of example measurements and explained their decisions made in data handling (Figure 6). This helps to identify differences and rationales for the processing and QC approaches they used; these differences affect fluxes even in the small number of example fluxes in this survey (Figure 7). The resultant variability in fluxes could be assessed in a more rigorous way by distributing an entire data set of raw chamber measurements with flux experts and the community as a whole to process using their own calculation and QC approaches. Using a common dataset could provide a more accurate estimate of the uncertainty due to data processing by different researchers and add insights into the relevance of this additional source of variability in $CH_4$ fluxes through comparison with the natural spatial and temporal variability in the data set. Sharing an entire data set would also eliminate some assumptions made in data processing in this survey (Section 4.3). The fluxes estimated from the raw data set could then be uploaded to a chamber flux network website or an existing platform, such as GitHub, to add to a growing pool of flux data sets computed from the same chamber measurements by different researchers. A flagging metric could be added to indicate data quality, which would aid modelers and others in interpreting and understanding noisy observations, which are common in ecosystems with high spatial and temporal heterogeneity. This way,

every interested researcher could assess how the flux estimates based on their processing techniques relate to those calculated by other flux experts and could help to build agreement on how to handle unclear cases of non-linear concentration changes.

Additional synthetic data could be an important addition to chamber measurements in the reference dataset and would clarify the processes resulting in strongly divergent flux estimates (Figure 7). A forward model could be developed to simulate the change in $CH_4$ concentrations in a chamber headspace as the real flux overlaid by a combination of chamber-induced artefacts in response to environmental conditions (Hutchinson and Livingston, 2001). This way, researchers could compare their flux estimates to the known "real" flux underlying a simulated measurement to see if they succeed in detecting it against background noise and artefacts.

Measurement simulations can also help to understand how $CH_4$ (and covarying $CO_2$) concentrations change in response to commonly cited measurement issues that might result in non-linear fluxes (Table 1). The simulations can therefore answer the key questions that appeared in the visual QC exercise – Which part of a nonlinear measurement should be used for flux calculation? Can the linear part of a measurement following an ebullition event still be used for flux calculation? How should close-to-zero fluxes be identified and handled? Does a decrease in headspace $CH_4$ concentrations indicate actual $CH_4$ uptake? Pirk et al. (2016) demonstrated that applying a nonlinear model for flux calculations can lead to an overestimation of $CH_4$ emissions if the nonlinear change in $CH_4$ concentrations was not, as assumed, caused by a change in the gas concentration gradient over time. Improved process understanding will help to avoid introducing bias into flux data sets through unsubstantiated handling of non-linear measurements.

Models and experimental studies to assess the effect of instrumentation on the flux estimates have previously been used to derive guidelines for certain aspects of measurement setups, such as chamber dimensions, pressure vents, and airtight seals (Christiansen et al., 2015, 2011; Hutchinson and Livingston, 2001; Pumpanen et al., 2004). Additionally, model-derived metrics can be used in post-hoc quality control, as was demonstrated for the minimum detectible flux (MDF) metric by Nickerson (2016). Such metrics will help identify a standardized set of required metadata on chamber setup and experimental design and ancillary measurements that should be takes alongside $CH_4$ fluxes in addition to the various variables currently recorded for the specific applications of the survey participants (Figure 5). For example, using a model with inputs of the air-soil $CH_4$ concentration gradient in the soil together with soil porosity and other soil properties can help to assess the potential effect of headspace saturation on the $CH_4$ flux (Pirk et al., 2016), which was commonly cited as problematic in this analysis (Table 1). Introducing model-derived metrics will therefore help scientists to make more informed decisions in selecting a time period within a chamber measurement for flux calculation, choosing a fit function, and filtering the data set. From the metrics, diagnostics for quality flagging can be derived that foster standardization of quality control procedures while factoring in site-specific conditions.

## 6. Conclusions

Chamber flux measurements are crucial for capturing spatial variability in ecosystems and quantifying treatment effects on greenhouse gas emissions. At broader spatial scales, synthesis datasets of chamber $CH_4$ fluxes show differences among but also high variability within wetland classes (Kuhn et al., 2021; Treat et al., 2018).The high variability within classes has been attributed to high spatial and temporal variability and can be partly compensated for by using longer integration times (Treat et al., 2018) or by capturing spatial variability due to microtopography effects (Virkkala et al., 2024). The results from this expert survey show that differences in methodology may be an additional factor contributing to high variability in $CH_4$ fluxes across sites and datasets. For example, the handling of low positive and negative fluxes can significantly affect estimates of $CH_4$ budgets, particularly in high latitude and upland regions where low $CH_4$ emissions and/or uptake of $CH_4$ can be expected during large parts of the year. Discarding low or zero fluxes can lead to a bias towards higher $CH_4$ emissions and potentially make the difference between a net annual uptake or a net emission of $CH_4$ in low-flux regions. Ebullition events may also comprise a substantial fraction of emissions; discarding these may lead to an underestimation of ecosystem $CH_4$ fluxes.

Our assessment of flux variability points towards the questions of where and when we introduce the largest error into our flux estimates – is it when we choose our measurement setup and processing approaches or do the location and the timing together with the spatial and temporal resolution of the measurements matter more? Answering this question will help identify the most relevant starting points for improving the accuracy of flux estimates and help lower uncertainties for flux syntheses. In any case, the survey shows that our human decision making introduces uncertainties that can obscure natural spatial and temporal variability in $CH_4$ fluxes. This might make it harder to identify relevant spatial and temporal environmental drivers of $CH_4$ fluxes, which is crucial for model development and $CH_4$ budget estimations.

*Data availability.* The results of the expert survey described in this paper are available from PANGAEA: https://doi.org/10.1594/PANGAEA.971695 (Jentzsch et al., 2024b).

## Appendix

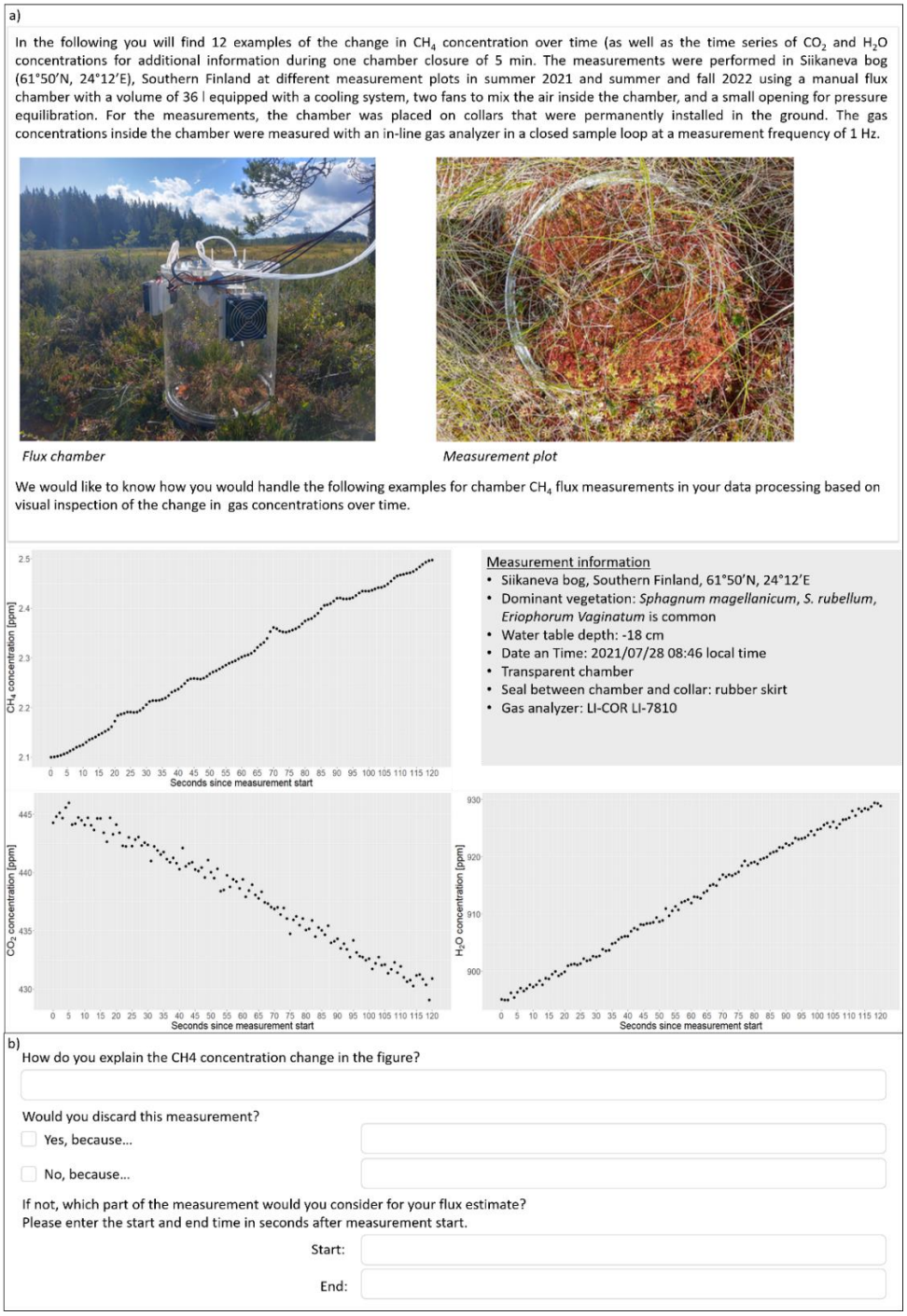



**Figure A1: Example of the information provided (a) and the questions asked (b) for the visual QC of one of the 12 chamber measurements under discussion.**

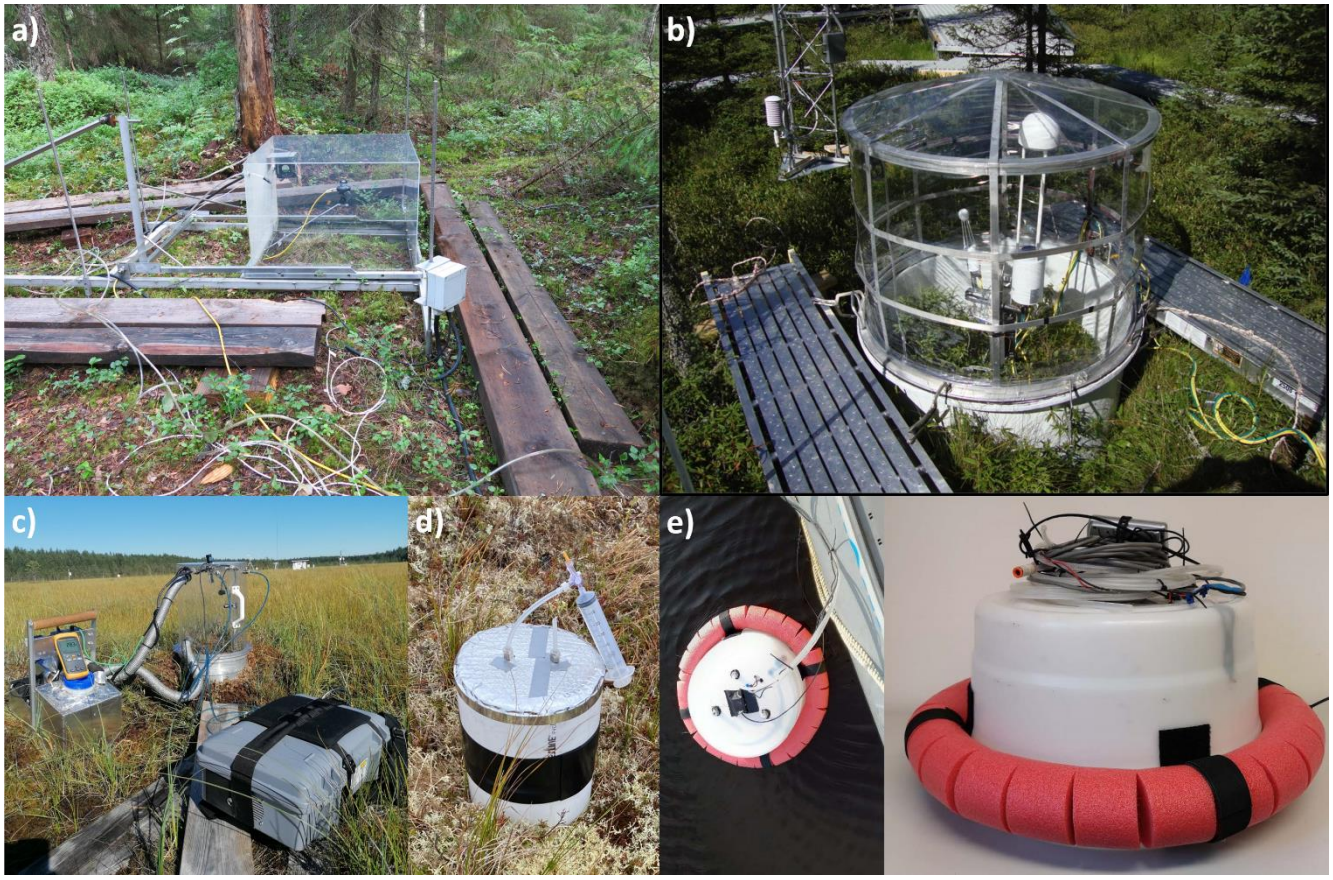



**Figure A2: Examples for different chamber setups in different environments. Automated chamber in a boreal forest (a), large manual chamber with gas analyser insider the chamber (b), transparent manual chamber with in-line gas analyser and cooling unit in a boreal fen (c), opaque manual chamber with a syringe for manual gas sampling and a tube for pressure equilibration (d), floating chamber connected to in-line gas analyser and deployed from a boat for aquatic measurements (e).**

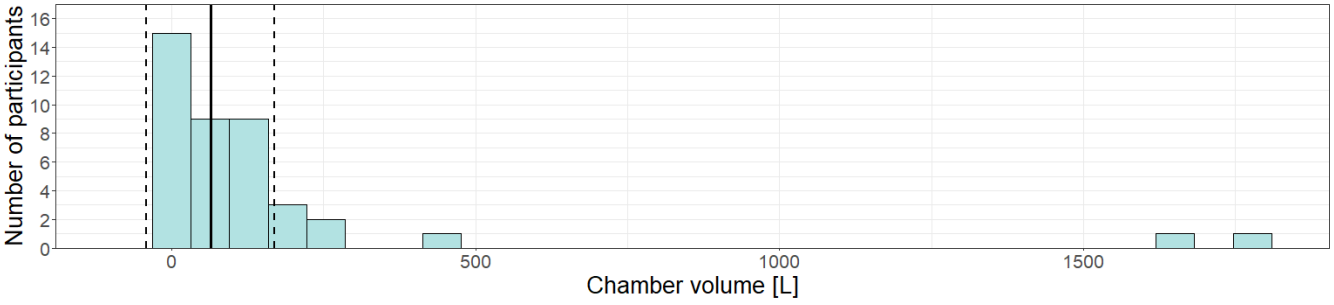



**Figure A3: Histogram of the chamber volumes. Black solid line: median, black dashed lines: IQR.**

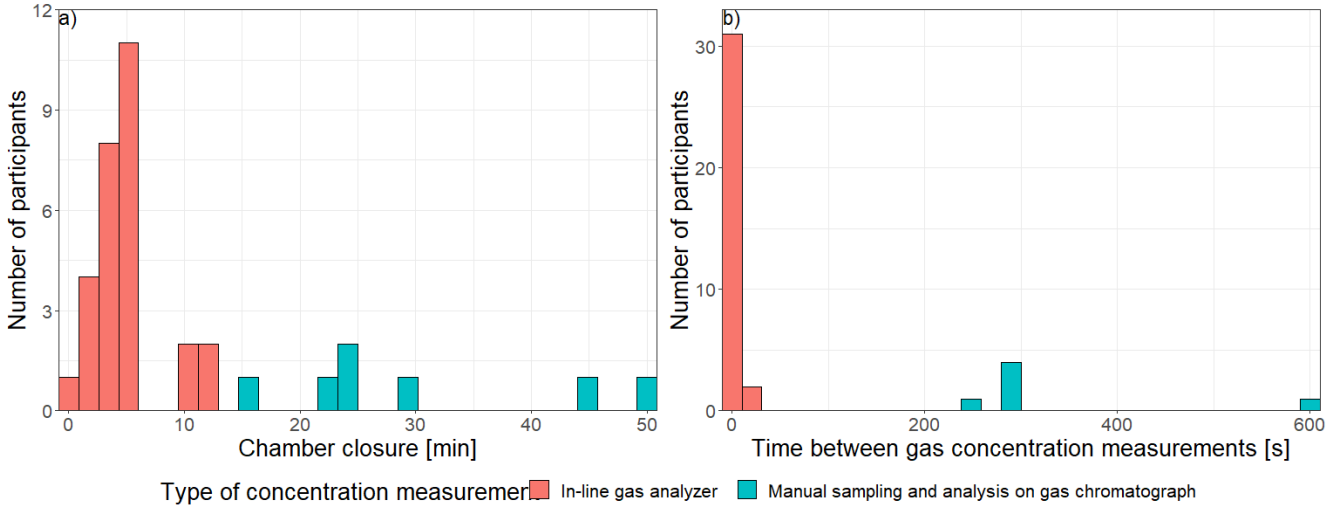



**Figure A4: Duration of chamber closure (a) and frequency at which the gas concentration inside the chamber is recorded, as time between two measurements (b) by type of concentrations measurement.**

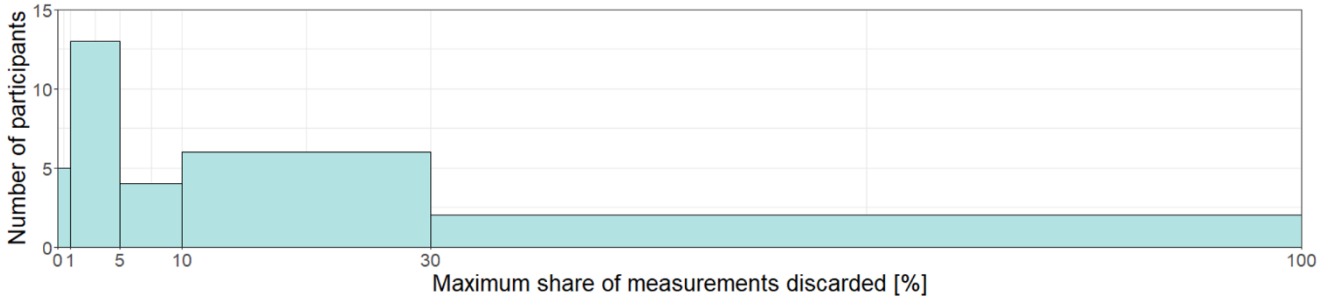



**Figure A5: Maximum percentage of measurements that the survey participants usually discard from their own data sets.**

**Table A1: Number and percentage of occurrence of identified measurement scenarios in a data set of chamber measurements collected at Siikaneva bog, Southern Finland, in 2021 and 2022. Categorization of the 12 measurement examples used in the visual QC exercise into these measurement scenarios and weighting factors calculated for each measurement example for the extrapolation of CVs derived from the 12 measurement examples to a representative data set (Table A2). Weighting factors were derived as the relative occurrence of the respective measurement scenario in the Siikaneva data set (788 measurement in total) divided by the number of occurrences in the visual QC exercise.**

| Measurement scenario | Linear increase | | Nonlinear increase – decreasing slope | | | Initial jump | Jump(s) | | Nonlinear increase – increasing slope | Inconsistent trend | Linear decrease | Low variation |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Occurrence in Siikaneva data set (%) | 468 (59.4) | | 144 (18.3) | | | 66 (8.4) | 62 (7.9) | | 25 (3.2) | 16 (2.0) | 4 (0.5) | 3 (0.4) |
| Measurement ID | VQC1 | VQC2 | VQC4 | VQC5 | VQC9 | VQC7 | VQC8 | VQC12 | VQC10 | VQC11 | VQC3 | VQC6 |
| Weight [%] | 29.7 | 29.7 | 6.1 | 6.1 | 6.1 | 8.4 | 4.0 | 4.0 | 3.2 | 2.0 | 0.5 | 0.4 |

**Table A2: Visual QC exercise (blue shading): Fluxes in mg $CH_4$ $m^{-2}$ $d^{-1}$ calculated for each example chamber measurement (columns) by each survey participant (rows) based on the time periods chosen for flux calculation in the visual QC part of the survey. Coefficients of variance (CV) are given for each example measurement ($CV_{ex}$) and each measurement scenario ($CV_{scenario}$) across the participants. "NR" indicates that a participant did not respond to the respective measurement example, so neither kept nor discarded the measurement and "D" indicates that a participant discarded a measurement example. Percentage of measurements kept for flux calculation (purple shading): Percentage of measurement examples within the visual QC exercise kept for flux calculation by participant ($Kept_{VQC}$, relative to total number of measurement examples that a participant responded to), measurement example ($Kept_{ex}$), and measurement scenario ($Kept_{scenario}$). Extrapolation to Siikaneva data set and uncertainty estimates (red shading): Extrapolation of the visual QC results to the entire Siikaneva data set through weighting by the frequency of occurrence of each measurement scenario in the data set (Table A1). Answers from participants who gave no response to more than one measurement example were excluded from the calculations of $Kept_{VQC}$ and $Kept_{Sii}$. Participants who did not participate at all in the visual QC exercise have been excluded from the table.**

| Measurement scenario | Linear increase | | Nonlinear increase – decreasing slope | | | Initial jump | Jump(s) | | Nonlinear increase – increasing slope | Inconsis-tent trend | Linear decrease | Low variation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Measurement ID / Participant ID | VQC1 | VQC2 | VQC4 | VQC5 | VQC9 | VQC7 | VQC8 | VQC12 | VQC10 | VQC11 | VQC3 | VQC6 | $Kept_{VQC}$ [%] | $Kept_{Sii}$ [%] |
| 1 | 67.31 | 3204.96 | 68.56 | 171.47 | 1102.58 | 351.70 | 2183.91 | 87.58 | 137.75 | 127.25 | -8.99 | 14.17 | 100 | 100 |
| 2 | 66.80 | 3169.76 | D | D | 1049.99 | 377.13 | D | D | 131.51 | D | D | D | 42 | 77 |
| 3 | 68.68 | 3367.11 | 73.95 | 169.61 | 1036.42 | 371.20 | D | 88.95 | 158.75 | -18.64 | -10.78 | 13.96 | 92 | 96 |
| 4 | 67.02 | 3294.94 | 66.92 | 171.47 | 1790.25 | 350.85 | 3117.04 | 482.53 | 127.40 | D | -11.16 | D | 83 | 98 |
| 5 | 67.02[a] | 3166.90[a] | 32.80[a] | 67.90[a] | 1751.88[a] | D | D | D | 116.73[a] | D | -11.16[a] | 0 | 67 | 81 |
| 6 | 66.34 | 3156.77 | 65.24[n.l.] | 61.26 | 1535.05 | D | D | 275.62 | D | D | D | D | 50 | 52 |
| 7 | 67.17 | 3067.51 | 61.86 | D | 1057.30 | D | 2031.93 | D | 133.38 | D | D | D | 50 | 79 |
| 8 | 67.02 | 3169.76 | D | D | 1091.37 | 378.93 | D | D | 116.73 | D | -11.16 | D | 50 | 78 |
| 9 | 67.02 | 3169.76[n.l.] | 71.13 | 67.90[a, n.l] | 1751.88[a] | D | D | D | D | D | -11.08 | 0 | 58 | 79 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 67.02 | 3304.26 | 75.34 | D | D | 380.34 | D | D | 158.75 | D | D | D | 42 | 77 |
| 14 | D | 3166.90[a] | 77.53 | 180.71 | 4216.12 | D | D | D | 168.07 | D | D | D | 42 | 51 |
| 16 | 67.02[a] | 3166.90[a] | 72.72 | 110.49 | 3553.67 | 380.34 | 2016.94 | 482.53[a] | 125.25 | D | -11.16[a] | D | 83 | 98 |
| 17 | D | NR | NR | NR | NR | NR | NR | NR | NR | NR | NR | NR | NA | NA |
| 18 | 66.14 | 3293.94 | 32.80[a] | 67.90[a] | 1122.55 | 357.19 | D | 482.53[a] | 131.32 | D / 9.58[a] | D | D | 71 | 94 |
| 19 | 67.02 | 3169.76 | 34.38 | 69.94 | 1790.25 | 416.49 | 3117.04 | 482.53 | 116.73 | 9.58 | -11.16 | 9.95 | 100 | 100 |
| 21 | 67.31 | 3229.48 | 66.81 | 171.47 | 1084.14 | 357.19 | D | 88.36 | 125.19 | 127.25 | -8.92 | 14.17 | 92 | 96 |
| 22 | 71.78 | 3301.49 | 79.90[n.l.] | 67.90[a] | 4973.26 | D | D | D | 75.82[n.l.] | D | D | D | 50 | 81 |
| 23 | 66.34 | 3141.12 | 70.84 | 130.21[n.l.] | 1680.02 | 371.20 | 2016.94 | 88.95 | 142.01 | -6.17 | -14.85 | 14.16 | 100 | 100 |
| 24 | 67.31 | 3399.49 | 73.94 | 169.61 | 5122.50 | 380.27 | D | 273.84 | 74.68 | D | -10.76 | D | 75 | 94 |
| 25 | D | D | D | D | D | D | D | D | NR | D | D | D | 0 | 0 |
| 26 | 67.02 | 3204.96 | 76.38 | D | D | D | D | D | D | D | -11.16 | D | 33 | 66 |
| 28 | 67.02 | 3286.36 | 53.58[n.l.] | 119.60 | 2482.08 | D | D | 275.62 | D | D | D | D | 50 | 82 |
| 29 | 67.17 | 3178.85[n.l.] | D | 66.39[n.l.] | 1764.55 | 377.13 | 1723.41 | D | 127.24 | 8.45 | D | 12.77 | 75 | 90 |
| 30 | D | 3358.45 | D | D | 1088.47 | 346.32 | D | D | D | D | D | 8.51 | 33 | 45 |
| 31 | 67.02[a] | 3312.29 | 72.72 | 180.71 | 5750.52 | 180.71[a] | D | 482.53[a] | 116.73[a] | 138.74 | D | D | 75 | 95 |
| 32 | 67.02[a] | 3304.26 | NR | NR | NR | NR | NR | NR | NR | NR | NR | NR | 100 | NA |
| 34 | 69.67 | 3304.26 | 65.05 | 138.58 | 1065.04 | 380.34 | D | 275.62 | D | D | -15.16 | 9.95 | 75 | 91 |
| 36 | 67.31 | 3321.94 | 73.94 | D | 1124.44 | D | D | 88.95 | 120.74[n.l.] | D | D | D | 50 | 79 |
| Mean±SD | 67.36 ±1.18 | 3238.93 ±84.93 | 65.07 ±14.52 | 121.29 ±48.77 | 2129.75 ±1490.01 | 374.37 ±19.88 | 2315.31 ±564.43 | 282.58 ±172.84 | 126.57 ±23.50 | 49.50 ±68.31 | -11.36 ±1.89 | 9.76 ±5.55 | Weighted sum: | 80 ±23 |
| $CV_{ex}$ [%] | 2 | 3 | 22 | 40 | 70 | 5 | 24 | 61 | 19 | 138 | 17 | 57 | 17 | 28 |
| $CV_{scenario}$ [%] | 2.5 | | 44 | | | 5 | 42.5 | | 19 | 138 | 17 | 57 | | |
| $Kept_{ex}$ [%] | 86 | 96 | 81 | 69 | 88 | 62 | 27 | 54 | 76 | 29 | 50 | 38 | | |
| $Kept_{scenario}$ [%] | 91 | | 79 | | | 62 | 40 | | 76 | 29 | 50 | 38 | | |

[a] accepted but no time period was given for flux calculation. The flux was therefore estimated based on the entire measurement.
[n.l.] Flux was estimated based on a linear fit although the participant suggested to use a nonlinear model instead.

## Text A1

### A1.1 Linear increase

Two example measurements included in the visual QC exercise showed a linear increase in $CH_4$ concentrations over the entire time of the chamber closure. In the first example (measurement ID VQC1, Figure A1.1), the $CH_4$ concentrations increased by only 0.4 ppm from a starting concentration of 2.1 ppm to a concentration of 2.5 ppm at the end of the measurement while in the second example (measurement ID VQC2, Figure A1.2) the starting concentration of 3.8 ppm was higher and increased by 38.2 ppm to reach 42.0 ppm over the course of the measurement. The $CO_2$ concentration in the chamber decreased during both measurements by 13.4 ppm and 9.1 ppm, respectively. The starting concentration of $H_2O$ in VQC2 was more than 10 times higher than for VQC1 and decreased strongly over the course of the measurement with an abrupt decrease at around 50 s after chamber closure, while $H_2O$ concentrations increased slightly in VQC1.

The participants described the trend in $CH_4$ concentrations in both VQC1 and VQC2 as a linear increase which they explained by net $CH_4$ production and diffusive emission. The $CH_4$ flux in VQC2 was additionally classified as large with one participant concluding that the measurement plot was a “hotspot” for $CH_4$ emission. For both VQC1 and VQC2, some participants additionally noticed slight deviations from the linear behavior of the $CH_4$ concentrations. Minor jumps in the $CH_4$

concentration in VQC1 were mentioned by 17 participants (61%), which they related to $CH_4$ ebullition (9 participants), insufficient mixing due a defective fan (3 participants), wind (1 participant), wind-induced pressure changes (1 participant), changes in atmospheric pressure influencing the ground diffusion rates and/or atmospheric pressure gradient (1 participant), boundary layer disturbance (1 participant), leakage (2 participants), disturbance (1 participant) caused by chamber placement or footsteps (1 participant). For VQC2, half of the participants pointed out a decrease in the slope of $CH_4$ concentrations starting between 250 and 260 s after the chamber closure, 21% of whom also noticed a simultaneous decrease in the slope of $CO_2$ concentrations. As explanations the participants mentioned saturation of the chamber headspace decreasing the concentration gradient over time (5 participants), a build-up of pressure (2 participants) potentially due to a defective pressure valve towards the end of the measurement (1 participant), a change in temperature over the course of the measurement (1 participant), or a small leak (1 participant) probably combined with windy conditions (1 participant). Many participants furthermore discussed the change in $CO_2$ and $H_2O$ concentrations over the time of the chamber closure. For VQC1, three participants mentioned that the $CO_2$ and $H_2O$ concentrations show a linear change, two of whom concluded that there was no air leaking from the chamber. Three participants on the other hand were concerned about the $H_2O$ measurements due to the high and increasing concentrations, and due to an assumed saturation and therefore decreasing slope towards the end of the measurement. Leakage from the chamber was suspected by three participants, two of whom explained this presumption with vegetation overgrowing the collar and one with the use of a less airtight rubber seal as opposed to a water seal. For VQC2, 18% of the participants picked up on the drop in $H_2O$ concentrations occurring around 40 s after the chamber closure, 40% of whom additionally mentioned a simultaneous change in the slope of $CO_2$ concentrations. Their reasoning included water condensing on the chamber walls and changing light conditions. Few participants decided to discard the two measurement examples. Measurement VQC1 was discarded by 4 participants (14%) suspecting $CH_4$ ebullition or stating that the starting concentrations of $CO_2$ were too high above ambient concentrations or that all chamber measurements generally need to be shaded. One participant excluded VQC2 due to an assumed saturation effect and one additional participant mentioned ebullition and a high initial concentration of $CH_4$ as potential reasons to exclude the measurement from flux calculations. 86% and 89% of the participants decided to keep VQC1 and VQC2 for flux calculation, respectively, due to the consistent linear increase in $CH_4$ concentrations without clear indications of significant disturbances or any malfunctioning of the instruments. For VQC1, the participants further supported their decision with the linear change in $CO_2$ and $H_2O$ concentrations making leakage from the chamber unlikely as well as with near-ambient $CH_4$ concentrations at the measurement start. For both VQC1 and VQC2 most participants who gave start and end times for flux calculation chose the middle part of the measurement, discarding the beginning and the end without mentioning a specific reason. The remaining participants considered the $CO_2$ and/or $H_2O$ concentrations in their choice of the time period for curve fitting. For VQC1, three participants chose the beginning of the measurement only, resulting in slightly higher flux estimates, two of whom assumed that $H_2O$ saturation diminished the increase in $CH_4$ concentrations towards the end of the measurement. For VQC2, some participants acknowledged the strong drop in $H_2O$ concentrations. Having no further information on potential reasons three of them decided not to let this unexpected behavior in $H_2O$ concentrations make them discard the $CH_4$ measurements while other participants reacted by excluding the

time of the drop in $H_2O$ concentrations from their calculation of the $CH_4$ flux through either using the part of the measurement after the drop (7 participants) or before the drop (1 participant). 61% of the 23 participants who entered start and end times for flux calculation discarded the end of the measurement where $CH_4$ and $CO_2$ concentrations increased at a lower rate, resulting in slightly higher flux estimates above 3200 mg $CH_4$ $m^{-2}$ $d^{-1}$. Two participants suggested to use a nonlinear fit which one of them specified as exponential.

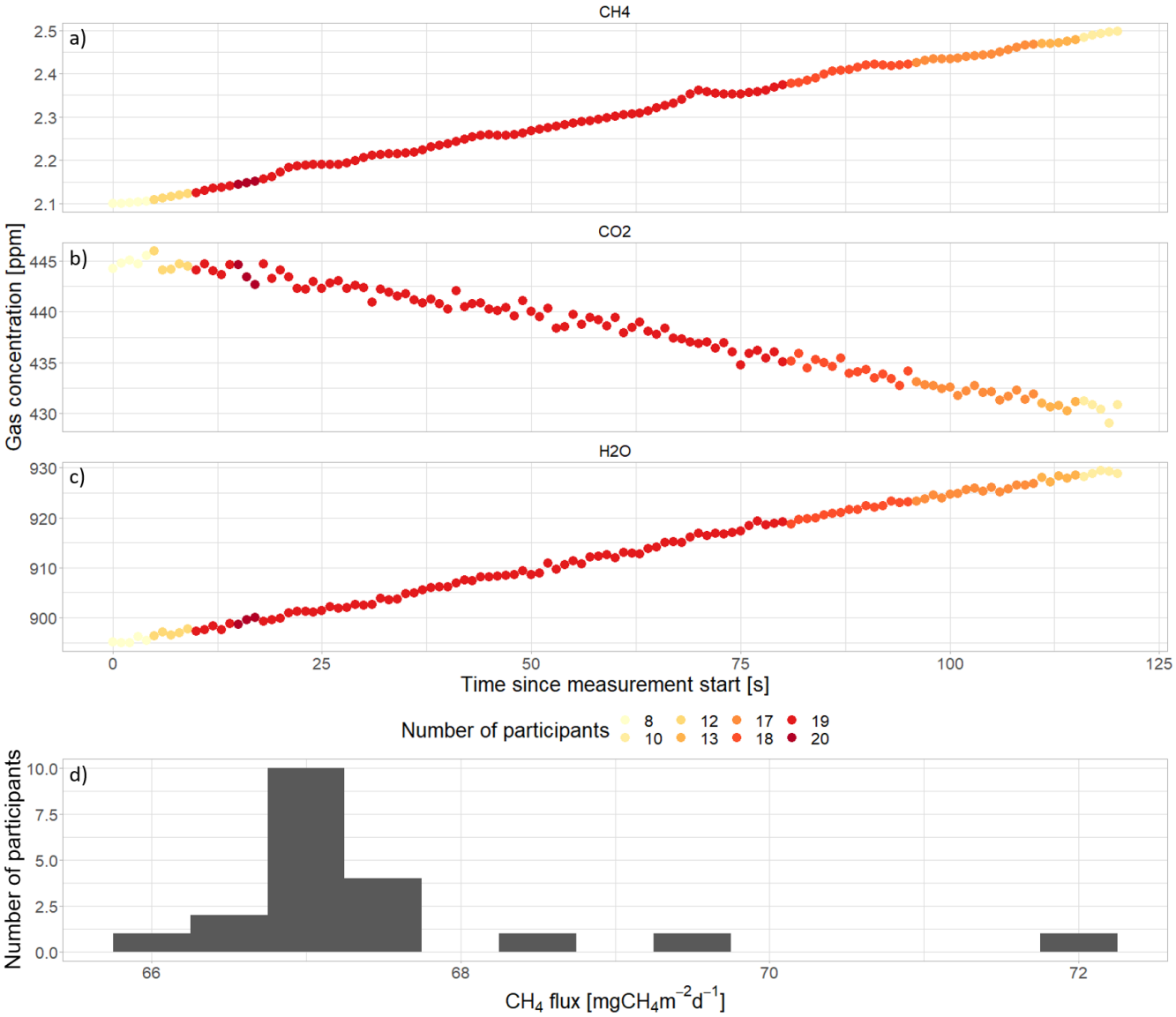



**Figure A1.1: Measurement example VQC1 of a small linear increase in $CH_4$ concentrations over the time of the chamber closure. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

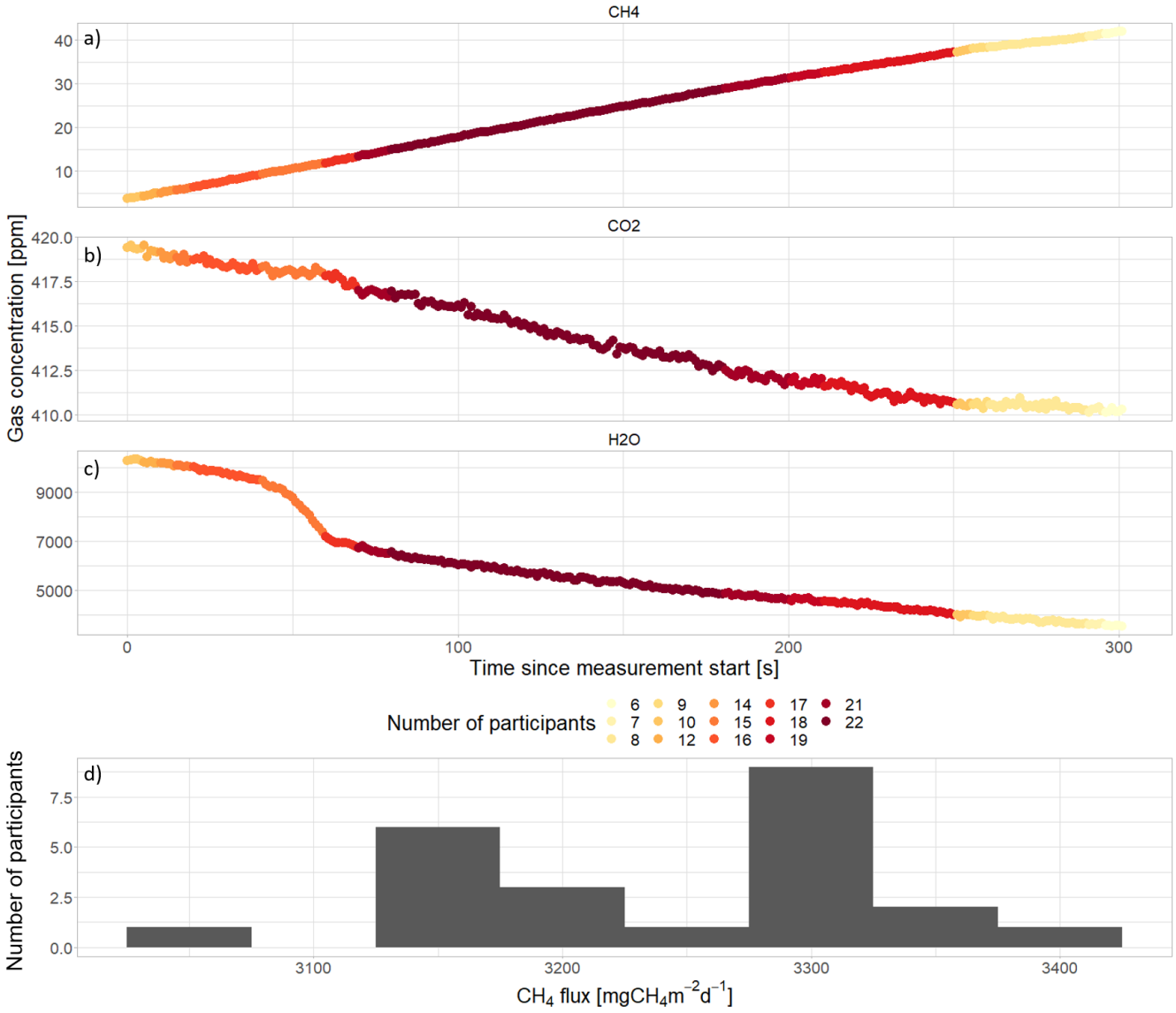



**Figure A1.2: Measurement example VQC2 of a strong linear increase in $CH_4$ concentrations over the time of the chamber closure. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

### A1.2 Nonlinear increase – decreasing slope

In the visual QC exercise, we included three examples of measurements that feature a nonlinear increase in $CH_4$ concentrations during the chamber closure with the rate of increase flattening out over time. Two examples show a small

915 nonlinear increase in $CH_4$ concentrations (measurement IDs VQC4 and VQC5, Figures A1.3 and A1.4) simultaneous with linearly decreasing $CO_2$ concentrations. $H_2O$ concentrations increased over the time of the chamber closure in VQC4 but decreased in VQC5. The third example (measurement ID VQC9, Figure A1.5) shows a stronger increase in $CH_4$ concentrations with intermittent jumps, linearly increasing $CO_2$ concentrations and $H_2O$ concentrations that fluctuate without a clear trend.

The participants classified the $CH_4$ measurements in VQC5 as a small flux that resulted from a balance between $CH_4$ 920 production and oxidation while VQC9 was identified as large emission indicating a $CH_4$ hotspot. The majority of the participants (85%, 85%, and 81%) discussed the nonlinear behavior of the $CH_4$ concentrations in VQC4, VQC5, and VQC9, respectively, offering various explanations for the decreasing rate of increase over time that were mainly related to chamber saturation, chamber leakage or an initial disturbance (Table A1.1).

Most participants (9, 10, and 8) suspected a saturation of the chamber headspace, while two participants stated that 925 saturation was unlikely to be reached during a measurement as short as VQC4 (330 seconds) and one participant explicitly mentioned that the changing slope in VQC9 did not look like a saturation effect. Just as many participants (9) suggested leakage through a weakening seal as the reason for the decreasing slope in VQC4, as supported by the simultaneously decreasing slope in $CO_2$ and $H_2O$ concentrations while other participants explicitly stated that $CO_2$ and $H_2O$ concentrations did not indicate a leak in this measurement. Due to the consistently linear $CO_2$ concentrations in VQC5 and VQC9, only one participant each 930 suspected leakage during these measurement examples. For VQC4, three participants further suspected that the high $H_2O$ concentrations at the end of the measurement influenced the $CH_4$ measurements, for example through condensation inside the chamber or in the gas flow line, one participant suggested a varying performance of the chamber fan, and two participants assumed that the nonlinearity was a phenomenon specific to *Sphagnum* moss. One participant more generally suggested that vegetation effects changed over the course of the measurement due to changing light conditions, affecting the $CO_2$ and $H_2O$ 935 concentrations in VQC4 and the $CH_4$ concentrations in VQC5.

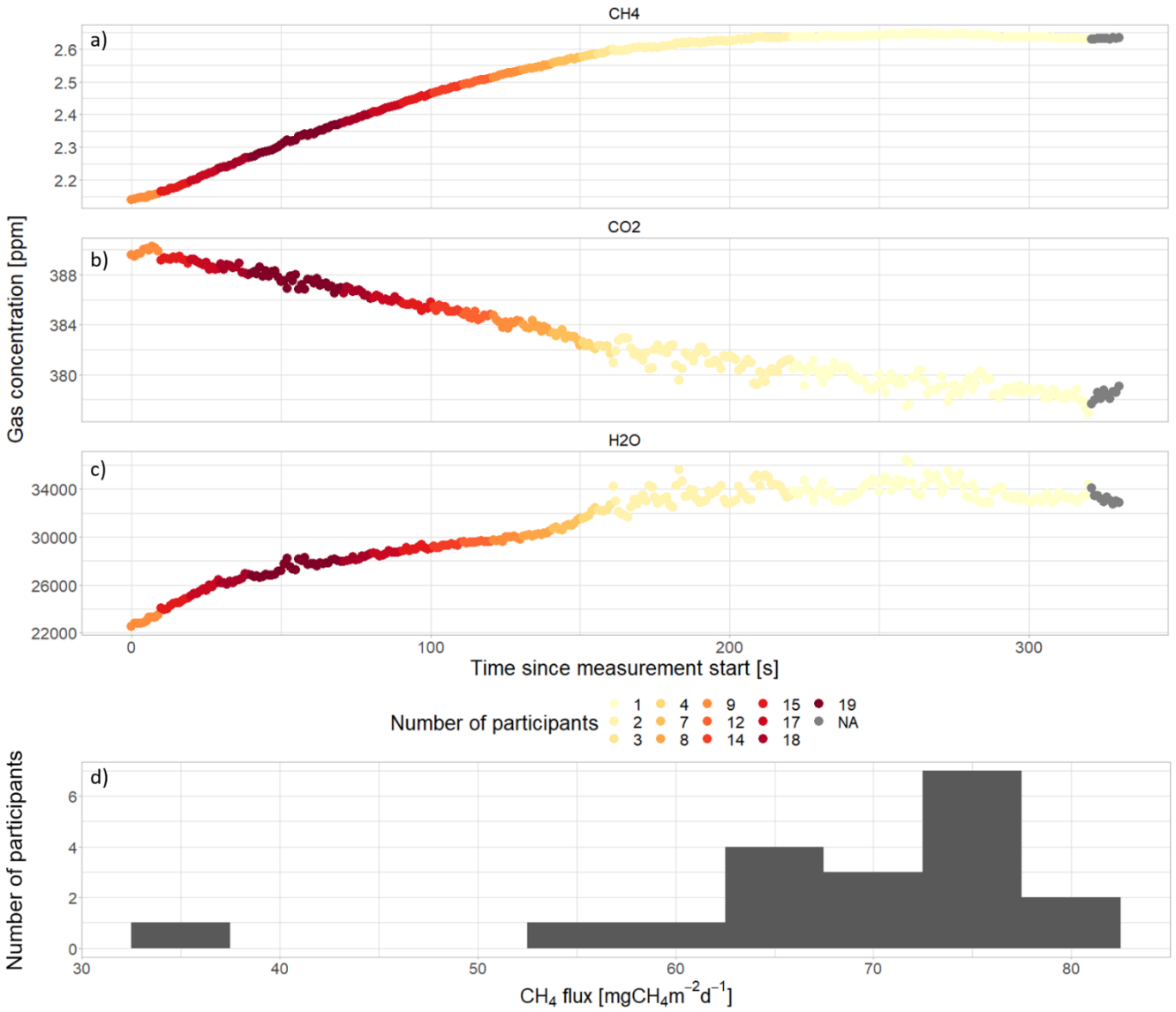
Besides a saturation effect or a weakening seal that would cause a decreasing slope in $CH_4$ concentrations towards the end of the measurement, many participants (3, 6 and 8) suggested that an initial disturbance such as ebullition triggered by the chamber placement had caused the stronger increase in the beginning of measurement examples VQC4, VQC5, and VQC9. For VQC9, 31% of the participants additionally pointed out minor fluctuations superimposed on the overall nonlinear increase 940 in $CH_4$ concentrations. Two thirds of them referred to the fluctuations as minor ebullition events while the others suggested episodic leakage from the chamber potentially caused by gusts of wind lifting the chamber sides or a malfunctioning pressure gauge. One participant pointed out that the $CH_4$ fluctuations cooccurred with fluctuations in the $H_2O$ concentrations and therefore suspected an instrument issue that could be related to spikes in the instrument cavity pressure.

The nonlinearity in the $CH_4$ concentrations resulted in 15%, 31%, and 12% of the participants deciding to discard the entire 945 measurement example VQC4, VQC5, and VQC9, respectively (Table A1.1). The reasons mentioned for the exclusion of the measurements again reflected the different interpretations of the participants on which part of the measurement represented the real flux. This disagreement shows less strongly in the range of flux estimates since participants who suspected an initial disturbance of the measurement disproportionately often discarded the entire measurement as they assumed that an initial

disturbance would also affect the remaining part of the measurement. For VQC4 and VQC5, all 54% of the participants who provided start and end times for flux calculations agreed that the beginning of the measurement should be used for or at least be included in the flux calculation, with three participants suggesting a nonlinear fit for both measurement examples. This resulted in smaller ranges of flux estimates compared to VQC9 (Table A1.1) which instead reflects the fundamentally different interpretations among the participants on which part of the measurement should be used for flux calculation. Here, half of the 21% of participants who gave start and end times for flux calculation chose a later part of the measurement where $CH_4$ concentrations appeared linear over a longer time period. This resulted in lower flux estimates (between 1000 and 1200 mg $CH_4$ $m^{-2}$ $d^{-1}$) compared to the flux estimates larger than 3500 mg $CH_4$ $m^{-2}$ $d^{-1}$ derived for the one quarter of participants who instead chose the beginning of the measurement (Figure A1.5).

**Table A1.1: Explanations for the nonlinear increase in $CH_4$ concentrations, reasons to discard, and reasons and ways to keep measurements showing an increase in $CH_4$ concentrations as given by the participants. The responses were categorized based on the free text entries for measurement examples VQC4, VQC5, and VQC9. The number of responses given in the respective category are provided in brackets.**

| Explanations for nonlinearity | Reasons to discard | Reasons to keep |
|---|---|---|
| Saturation (23) | Saturation (2)<br>Nonlinearity – no steady state reached (3) | A nonlinear fit can be used (9) |
| Initial disturbance (16) | Initial disturbance biases flux later on (2) | A (linear) part of the curve can still be used (41) |
| Bad seal / Leakage from the chamber (8) | Bad seal / Leakage from the chamber (4) | |
| Unsure (12) | Unclear which part of the measurement represents the real flux (3) | No clear disturbance of the measurement (9) |
| Changing environmental conditions (1) | Changing environmental conditions (1) | Linear trend in $CO_2$ concentrations (5) |




965 **Figure A1.3: Measurement example VQC4 of a small nonlinear increase in $CH_4$ concentrations with decreasing slope over the time of the chamber closure. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

970

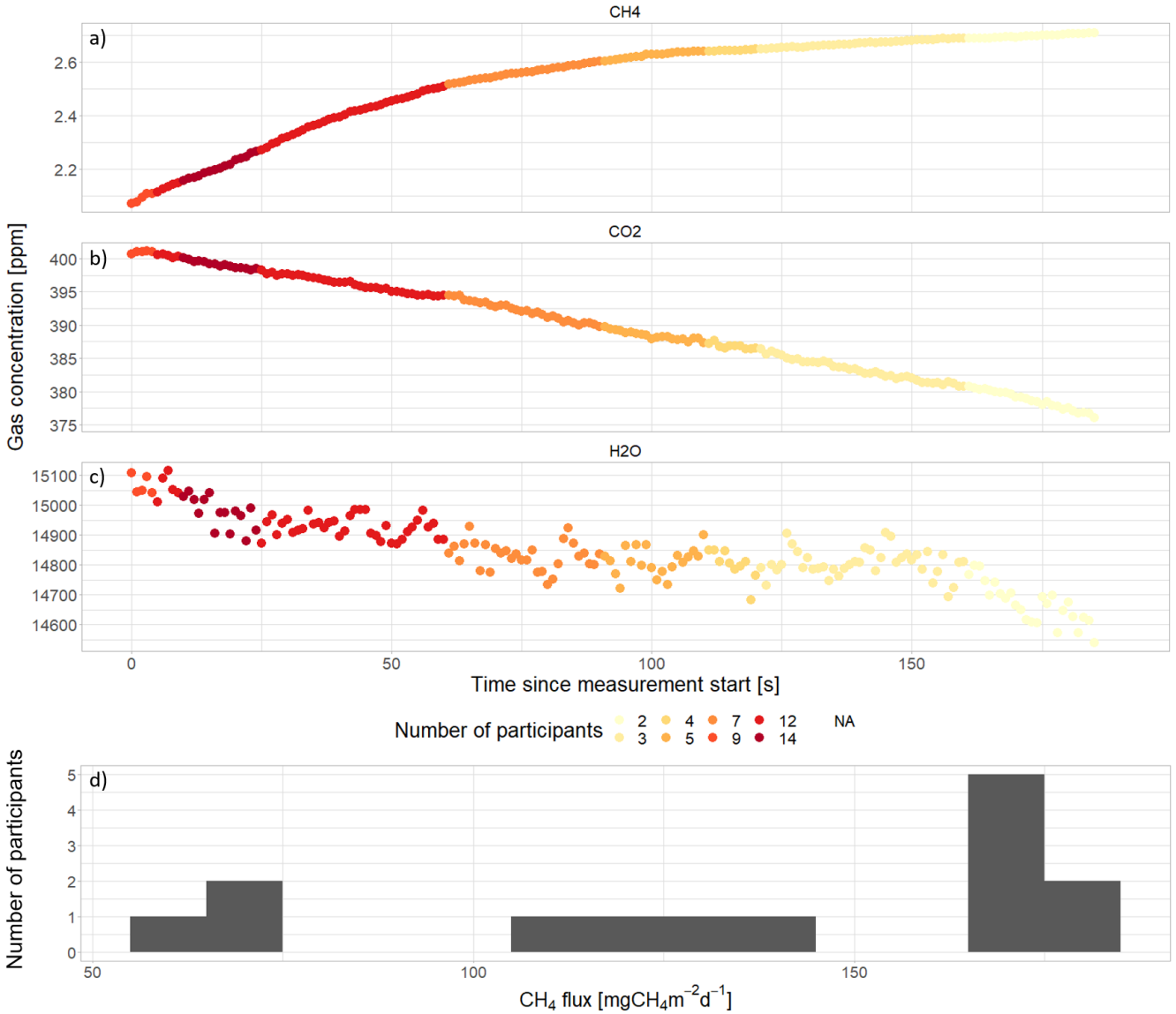



**Figure A1.4: Measurement example VQC5 of a small nonlinear increase in $CH_4$ concentrations with decreasing slope over the time of the chamber closure. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

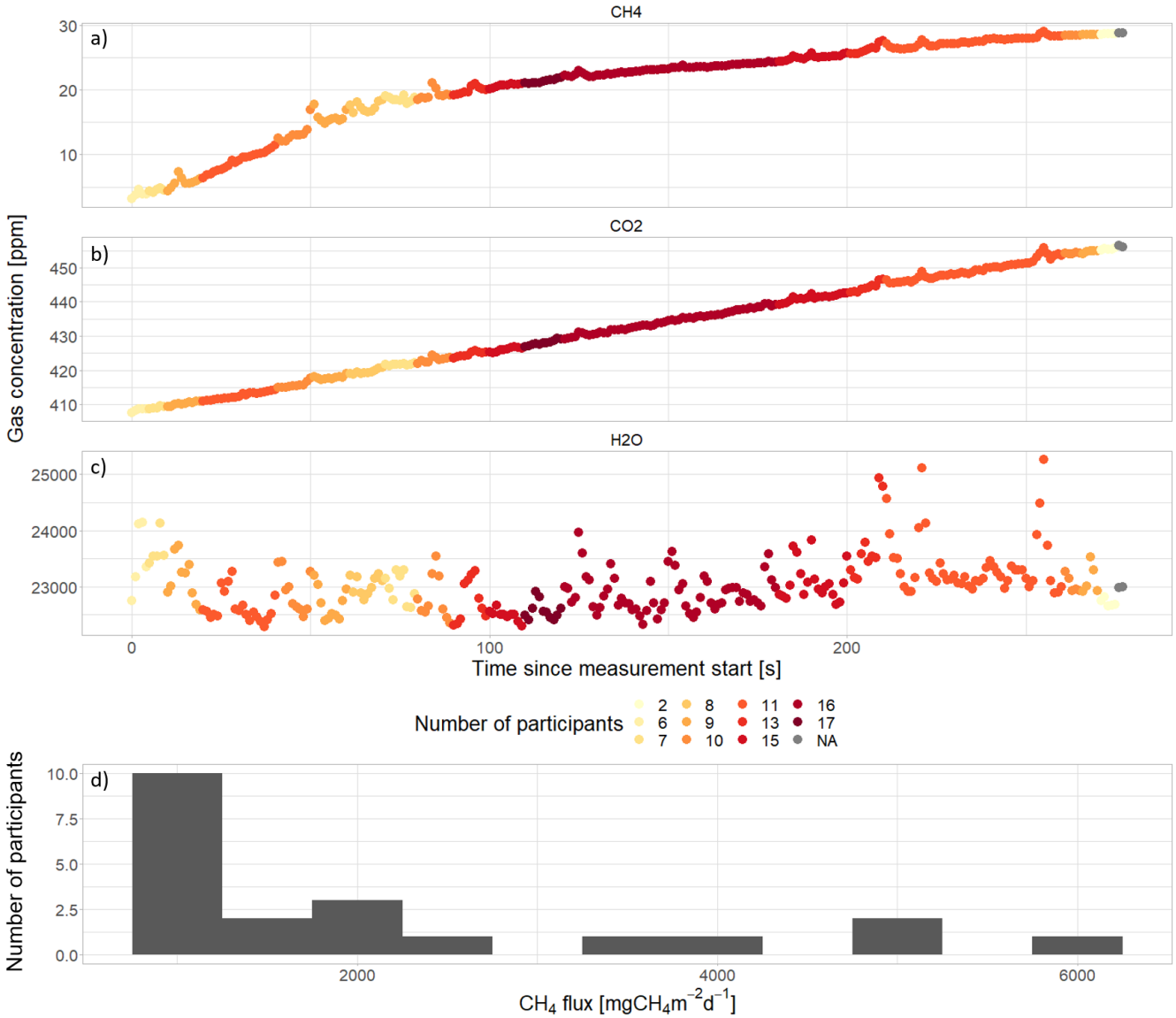



**Figure A1.5: Measurement example VQC9 of a strong nonlinear and jumpy increase in $CH_4$ concentrations with decreasing slope over the time of the chamber closure. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

### A1.3 Jump(s) at the beginning or in the middle of measurements

In our visual QC exercise, we included three example measurements that showed a relatively linear increase in $CH_4$ concentrations that was interrupted by one (VQC8, Figure A1.6 and VQC7, Figure A1.7) or more (VQC12, Figure A1.8) sudden increases in the concentration. In one, this occurred at the beginning (VQC7) and in the others, the middle of the measurement (VQC8, VQC12). In examples VQC8 and VQC7, these sudden jumps appeared in all three gases with $CH_4$ and $H_2O$ concentrations showing a sudden increase while $CO_2$ concentrations dropped simultaneously. In VQC12, on the contrary, $CO_2$ and $H_2O$ showed no equivalent to the jumps in the $CH_4$ concentration. In VQC8, a strong decrease in $CH_4$ concentrations directly followed the sudden increase, while in VQC7 and VQC12 the concentrations continued to increase at a lower rate starting close to the high concentration level after the jump.

Nearly all (100%, 65%, and 92%) of the participants mentioned the jump(s) in $CH_4$ concentration when discussing the measurement examples VQC8, VQC7, and VQC12, respectively. For all measurement examples, the majority of these participants explained their observation with episodic events of ebullitive $CH_4$ emission (VQC8: 65%, VQC7: 88%, VQC12: 92%) with only one participant each suggesting a malfunctioning of the gas analyzer as a reason for the sudden increase in $CH_4$ concentrations in VQC7 and VQC12. Some (35%) of the participants assuming ebullition stated that the ebullition event in VQC8 was caused by a disturbance and all agreed that the chamber placement caused the ebullition for VQC7. Only one participant (5%) mentioned anthropogenic disturbance as the reason for the ebullition events in VQC12. For VQC8, 12% of the participants pointed out the sudden changes in $CO_2$ and $H_2O$ concentrations along with the jumps in $CH_4$. Reasons mentioned by one participant each were a malfunctioning of the gas analyzer and an overpressure caused by the bubble release while another participant suggested the release of gas bubbles with high $CH_4$ but low $CO_2$ concentrations as a natural cause for this observation. Similarly, one of the two participants who mentioned the absence of a simultaneous change in the other gases in VQC12, assumed a release of bubbles with high $CH_4$ concentration but $CO_2$ concentrations close to ambient conditions due to the different production depths of the two gases. For VQC8, 41% of the participants discussed the decrease in $CH_4$ concentrations following the assumed ebullition event and suggested leakage of air from the chamber, potentially combined with wind as a potential cause. In the discussion of VQC7, two participants disagreed on the effect of the water table on $CH_4$ ebullition, one mentioning that in the measurement $CH_4$ ebullition was more likely to happen because of the high water table while the other stated that ebullition happened despite the high water table, indicating a fundamentally different understanding of the causes of $CH_4$ ebullition among the participants. Two participants of VQC7 furthermore classified the measurement as an example of strong $CH_4$ emission which they explained by strong anaerobic $CH_4$ production related to the high water table and by the vegetation providing substrate for acetoclastic $CH_4$ production, respectively.

Of the three measurements with jumps in $CH_4$ concentrations that we included in the visual QC exercise, VQC8 raised the most concern with the highest number of participants excluding the example (Table A2) and with the largest variety of reasons mentioned for the discard, including the inconsistent trend in $CH_4$ and $CO_2$ concentrations making them wonder which part of the measurement to use for flux calculation, ebullition affecting the pressure inside the chamber, too much variation in $CH_4$

and $CO_2$ concentrations even after the jump, chamber leakage and too high initial $CH_4$ and $CO_2$ concentrations. Leakage was also suggested by one participant for VQC7, who suspected that Sphagnum moss might have obstructed the chamber seal with the collar. VQC12 was classified as too short of a measurement by one participant and discarded by another for too high initial $CO_2$ concentrations.

There was disagreement among the participants on whether the remaining part of a measurement after a jump in the $CH_4$ concentration could still be used for flux calculation. For 10 of 11 participants discarding VQC7, the main concern was high concentrations having a lasting effect on the concentration gradient and thus on the diffusive $CH_4$ flux during the rest of the measurement while only one of 21 and 15 participants discarded measurements VQC8 and VQC12 for that reason. For VQC12, four of the 11 participants who kept the measurement, all of whom also gave start and end times, avoided this problem by using the beginning of the measurement before the first jump for flux calculation. On the contrary, for VQC8 and VQC7 five and 14 of the seven and 15 participants who kept the measurement and/or gave start and end times for flux calculation decided that the measurement after the jump in $CH_4$ concentrations could still be used for flux calculation, respectively, and five participants in VQC12 preferred to use the part between the first two jumps because it showed a longer linear increase.

The choice of different time periods for flux calculation resulted in two and three different classes of flux magnitudes for VQC8 and for VQC7 and VQC12, respectively. The highest flux estimates of more than 3000 mg $CH_4$ $m^{-2}$ $d^{-1}$,483 mg $CH_4$ $m^{-2}$ $d^{-1}$, and 416 mg $CH_4$ $m^{-2}$ $d^{-1}$ stemmed from the two, one, and two participants who used the whole measurement example VQC7, VQC8, and VQC12, respectively, for flux calculation because these estimates also included ebullitive in addition to diffusive $CH_4$ emissions, reflecting the general disagreement on whether $CH_4$ ebullition should directly be included in the flux estimates derived from chamber measurements. For VQC7 and VQC12, the flux estimates from the participants who excluded the jumps in $CH_4$ concentration from the time period for flux calculation can further be split into two classes. For VQC7, nine participants excluded only the very beginning of the measurement, while five participants only used a later part starting at about 50 s into the measurement when $CO_2$ concentrations decreased at a higher rate, resulting in slightly lower $CH_4$ fluxes. For VQC12, when excluding the jump in $CH_4$ concentrations the flux estimates were higher for four participants who chose the measurement period before the first jump, reaching up to 275 mg $CH_4$ $m^{-2}$ $d^{-1}$ compared to the five participants who chose the longest linear part of the measurement leading to flux as low as 88 mg $CH_4$ $m^{-2}$ $d^{-1}$. Due to the very linear behavior of the $CH_4$ concentrations following the initial jump and the higher agreement on the time period used for flux calculation, the CV of 5% for VQC7 was much lower than for the CVs of 24 and 61% for VQC8 and VQC12, respectively.

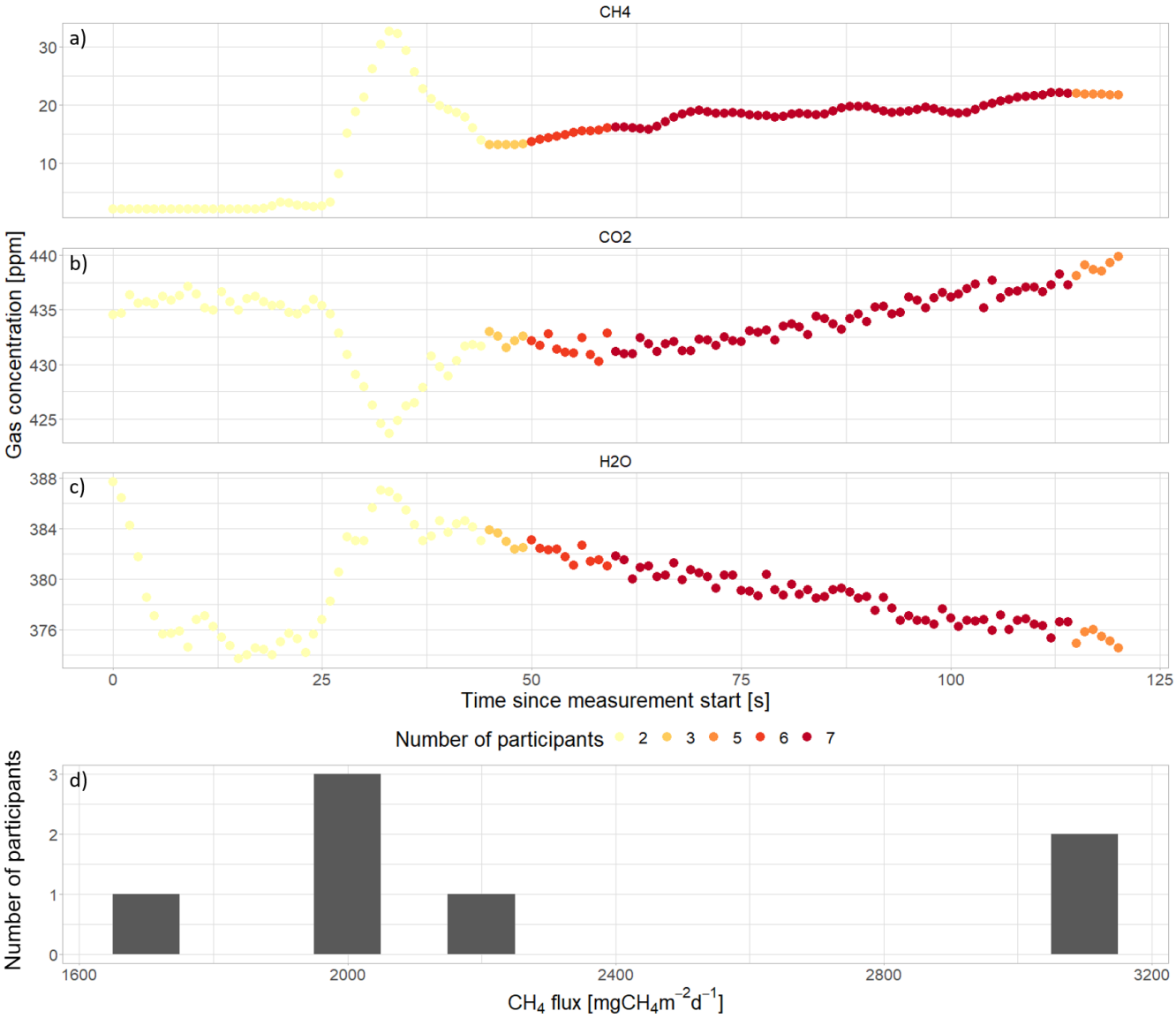



**Figure A1.6: Measurement example VQC8 of an overall increase in $CH_4$ concentrations over the time of the chamber closure after a strong jump in $CH_4$ concentrations. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

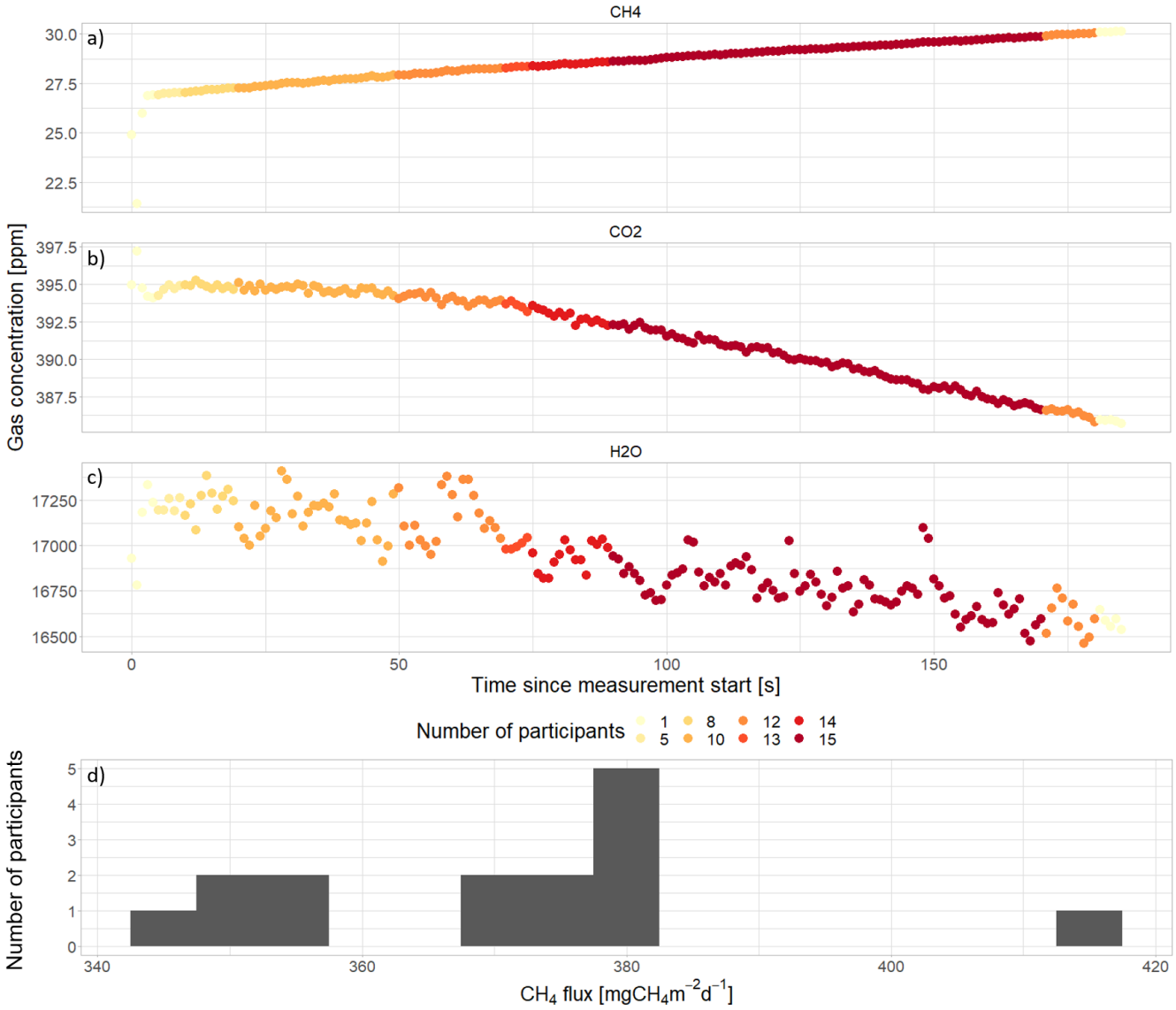



**Figure A1.7: Measurement example VQC7 of a linear increase in $CH_4$ concentrations of the chamber closure after an initial jump in $CH_4$ concentrations. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

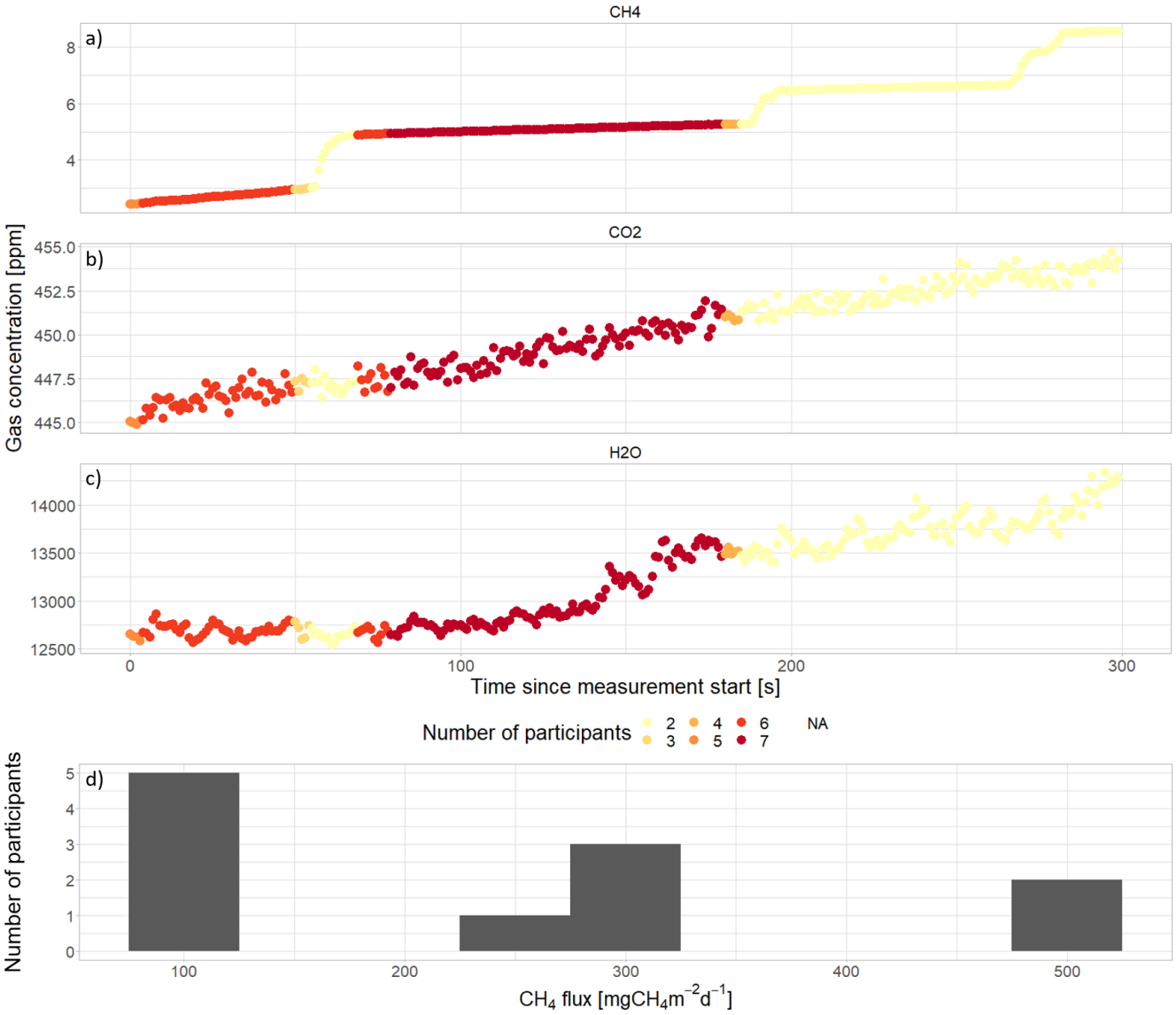



**Figure A1.8: Measurement example VQC12 of a linear increase in $CH_4$ concentrations between repeated jumps in the $CH_4$ concentration over the time of the chamber closure. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

**A1.4 Nonlinear increase - increasing slope**

One example included in the visual QC exercise showed a nonlinear increase in $CH_4$ concentrations over the chamber closure with the rate of increase becoming stronger over time (measurement ID VQC10, Figure A1.9). 15% of the participants classified the measurement as a diffusive emission of $CH_4$ without mentioning further details while 65% discussed the increasing slope in $CH_4$ concentrations over time, suggesting various reasons that could have caused the observed shape of the curve. The reasons suggested included an initial period of mixing or adjusting, an increase in chamber temperature over time, a disturbance of the measurement plot, a disturbance of the concentration gradient in the soil during chamber placement, an influence of the chamber on plant-mediated $CH_4$ transport, an incomplete seal of the chamber, incomplete mixing, and an interference with the simultaneously increasing $H_2O$ concentrations. Two participants mentioned that they had not seen such a shape in $CH_4$ concentrations from chamber measurements before. Regarding the magnitude of $CH_4$ emissions, three participants pointed out the strong increase in $CH_4$ concentrations despite the relatively low water table, which they related to plant-mediated $CH_4$ transport. One participant further mentioned that also the emission of $CO_2$ was high, indicating warm peat conditions. Two participants mentioned the higher and decreasing $CO_2$ concentrations in the beginning of the measurement which one of them related to the chamber placement, pushing more gases out of the ground. One participant furthermore mentioned that the chamber seal seemed to be intact.

Six participants decided to discard the measurement, three of whom did so because they could not explain the shape of the curve and stated that the curvature was so strong that the flux estimate would strongly depend on the time period chosen for flux calculation. The three remaining participants mentioned similarly unexpected shapes of $CO_2$ and $H_2O$ concentrations, higher $H_2O$ concentrations towards the end of the chamber closure which might have interfered with the $CH_4$ measurements, and high initial $CH_4$ concentrations as reasons to discard the measurement. 19 participants kept the measurement for flux calculation. The flux estimates for the 17 participants who gave start and end times for flux calculation strongly depended on the time period they chose which in turn depended on their interpretation of the measurement resulting in three distinct classes of flux magnitudes. Two participants decided to use the entire measurement, resulting in intermediate flux estimates of 117 mg $CH_4$ $m^{-2}$ $d^{-1}$. The majority (13) decided to remove the first 20 to 120 s of the measurement to keep only the more linear part of the $CH_4$ concentrations in the end, resulting in the highest flux estimates between 125 and 170 mg $CH_4$ $m^{-2}$ $d^{-1}$. The two remaining participants chose only the linear first 60 or 70 s of the measurement for flux calculation resulting in lower flux estimates of 75 and 76 mg $CH_4$ $m^{-2}$ $d^{-1}$, respectively, due to the lower rate of increase. Two participants suggested to use a nonlinear fit which one of them specified as exponential.

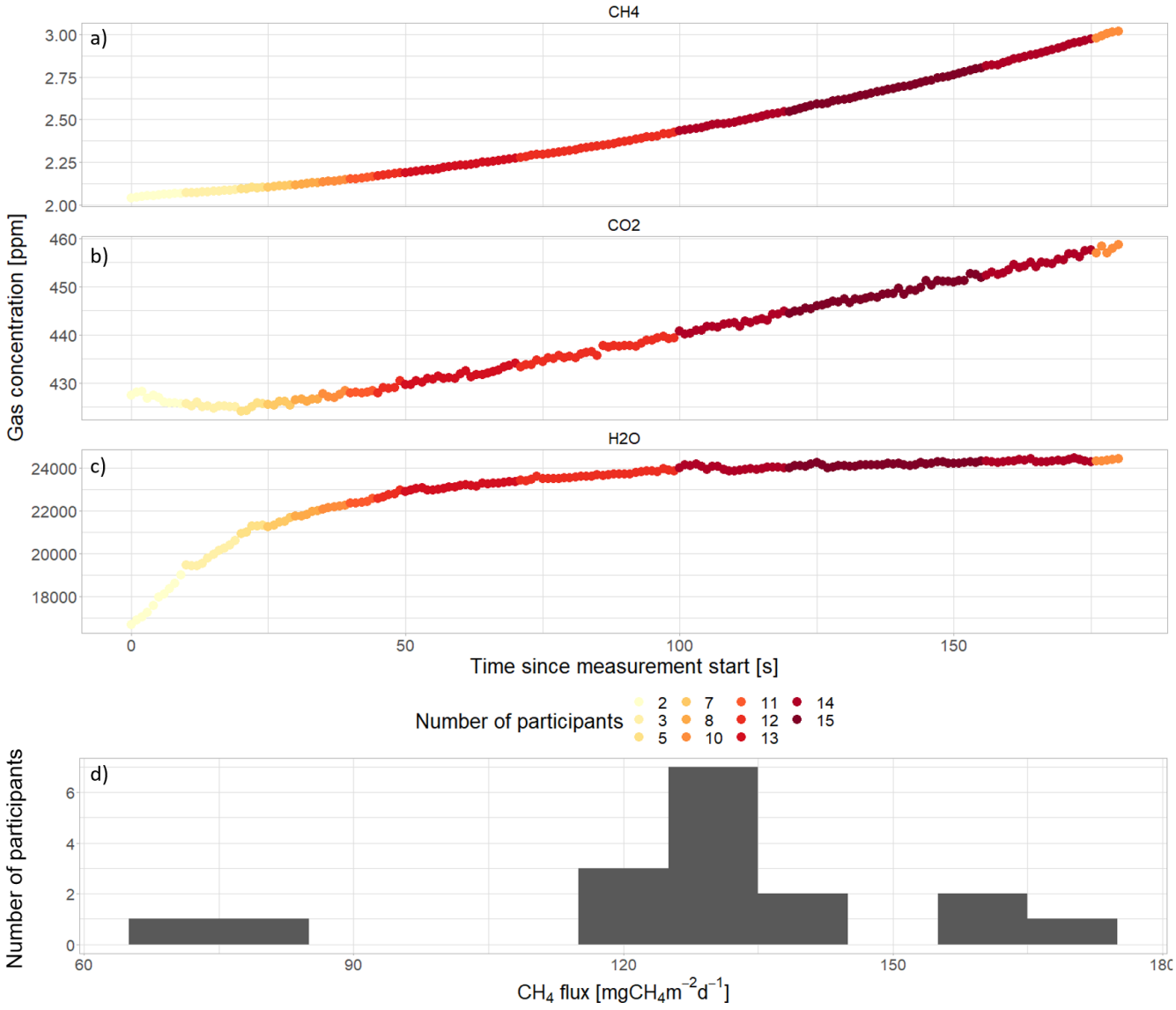



**Figure A1.9: Measurement example VQC10 of a small nonlinear increase in $CH_4$ concentrations with increasing slope over the time of the chamber closure. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

### A1.5 Inconsistent trend

One example included in our visual QC exercise showed an inconsistent trend in $CH_4$ with a change from increasing to decreasing concentrations over the time of the chamber closure (measurement ID VQC11, Figure A1.10). The survey participants disagreed on the reason for this behavior of the $CH_4$ concentrations. One part of the participants stated that $CH_4$ oxidation as indicated by the decrease in $CH_4$ concentrations towards the end of the measurement was unexpected and suggested that measurement issues were responsible for the inconsistent trend in $CH_4$ concentrations. They had different opinions however on the timing of the disturbance and therefore on which part of the measurement represented the actual $CH_4$ flux. Some participants suggested an initial disturbance such as $CH_4$ ebullition caused by the chamber placement while others assumed that the measurement was disturbed at a later point by a problem with the $CH_4$ analyzer like saturation of the detector or $H_2O$ interference due to the high concentrations towards the end of the measurement and potentially condensation of water vapor, or leakage or a malfunctioning fan after about 50 s into the measurement.

Most participants (66%) discarded the measurement because they missed a consistent trend of sufficient length in the $CH_4$ concentrations. Since the changing trend was either related to a disturbance or the reason was described as unclear, the participants did not know which part of the measurement to use for the flux calculation. Two participants additionally discarded the measurement because they considered the changes in the $CH_4$ concentration as too close to zero and another participant mentioned that the $CO_2$ and $H_2O$ concentrations did not show a steady trend over time either. Some (23%) of the participants decided to keep the measurement for flux calculation, all of whom provided start and end times for flux calculation as well as one additional participant who was uncertain whether to keep or to discard the measurement. The choice of the time periods used for flux calculation depended on the interpretation of the observed pattern in $CH_4$ concentrations and thus strongly influenced the resulting flux estimate ranging between a $CH_4$ uptake of -19 mg $CH_4$ $m^{-2}$ $d^{-1}$ to $CH_4$ emissions of up to 139 mg $CH_4$ $m^{-2}$ $d^{-1}$ and splitting the flux histogram into three distinct modes. Two participants chose to keep the entire measurement, resulting in a small positive flux indicating small net $CH_4$ emission of 8 to 10 mg $CH_4$ $m^{-2}$ $d^{-1}$. Three participants decided to use the stronger increase in $CH_4$ concentrations in the beginning of the measurement, resulting in the highest $CH_4$ emissions between 127 and 139 mg $CH_4$ $m^{-2}$ $d^{-1}$ while two participants assumed that $CH_4$ was consumed at the plot, using the later decreasing part of the $CH_4$ concentrations, resulting in negative flux estimates between -6 and -19 mg $CH_4$ $m^{-2}$ $d^{-1}$. This resulted in the highest CV among the measurement scenarios, estimated at 138%.

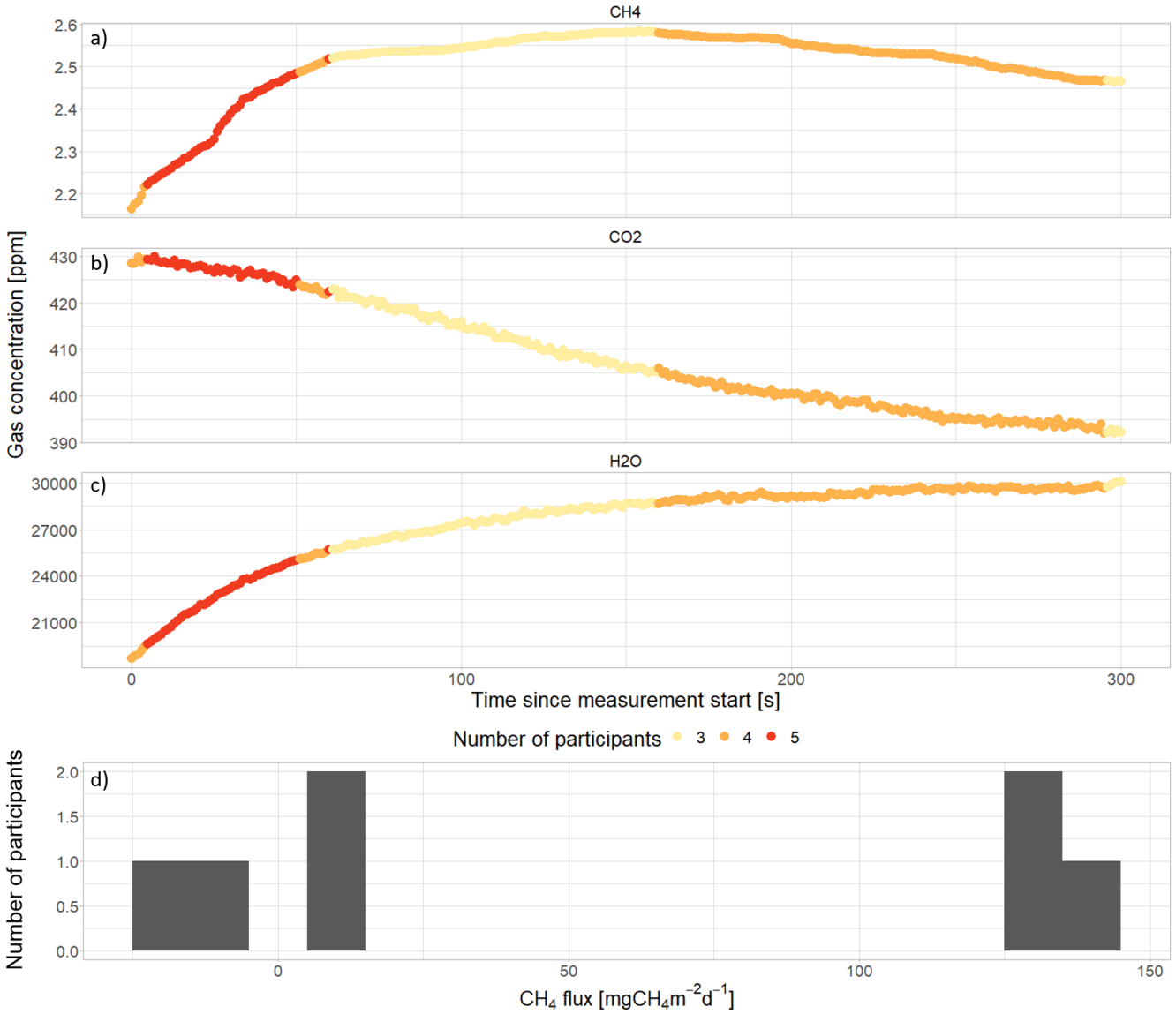



**Figure A1.10: Measurement example VQC11 of $CH_4$ concentrations showing an inconsistent trend over the time of the chamber closure. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

### A1.6 Linear decrease

One of the measurements in the visual QC exercise showed a small linear decrease in $CH_4$ concentrations over time (measurement ID VQC3, Figure A1.11). The survey participants largely disagreed on whether this measurement represented a real $CH_4$ flux. The majority (65%) of the participants assumed real net $CH_4$ uptake due to $CH_4$ oxidation dominating over $CH_4$ production while some (19%) of the participants referred to leakage and too high initial $CH_4$ concentrations in the chamber as technical problems causing a false apparent uptake of $CH_4$. The remaining 15% of the participants explicitly stated that they were unsure if the measurement represented a real flux. 23% of the participants more specifically mentioned an inconsistent trend in the $CH_4$ concentrations referring to three different stages of $CH_4$ flux or nonlinearities at the beginning and at the end of the measurement. As explanations, they offered initial $CH_4$ ebullition caused by the chamber placement, changes in the chamber temperature, changes in wind speed combined with chamber leakage, or changes in PAR potentially due to a changing cloud cover or due to condensation inside the chamber indicated by the trend in $CO_2$ concentrations changing along with the $CH_4$ trend as well as by high $H_2O$ concentrations.

A slim majority (54%) of the participants discarded the measurement because they did not expect $CH_4$ uptake in the given environmental (despite the relatively low water table), or because of the inconsistent trend in $CH_4$ concentrations which makes them unsure which part of the measurement to use for flux calculation, or because of too high initial concentrations of $CH_4$ and/or $CO_2$, or because they suspected anthropogenic disturbance from footprints and compacted vegetation or leakage. The flux estimates derived from the start and end times given by 11 of the 12 participants who decided to keep the measurement (46%) differed between the time periods chosen for flux calculation. While five participants chose the entire measurement, resulting in intermediate values of $CH_4$ uptake, the remaining six participants chose the time period for curve fitting based on the $CO_2$ concentrations. The middle part of the measurement with linearly decreasing $CO_2$ concentrations, the beginning of the measurement with stable $CO_2$ concentrations, and the end of the measurement with linearly increasing $CO_2$ concentrations were chosen by one, two, and one participant, respectively, while two participants excluded the end of the measurement resulting in strongly negative, lower negative, stronger negative and intermediate $CH_4$ fluxes, respectively. Overall, the mean of the flux calculated by the 12 experts keeping this flux was 11.36 mg $CH_4$ $m^{-2}$ $d^{-1}$ with a CV of 17%.

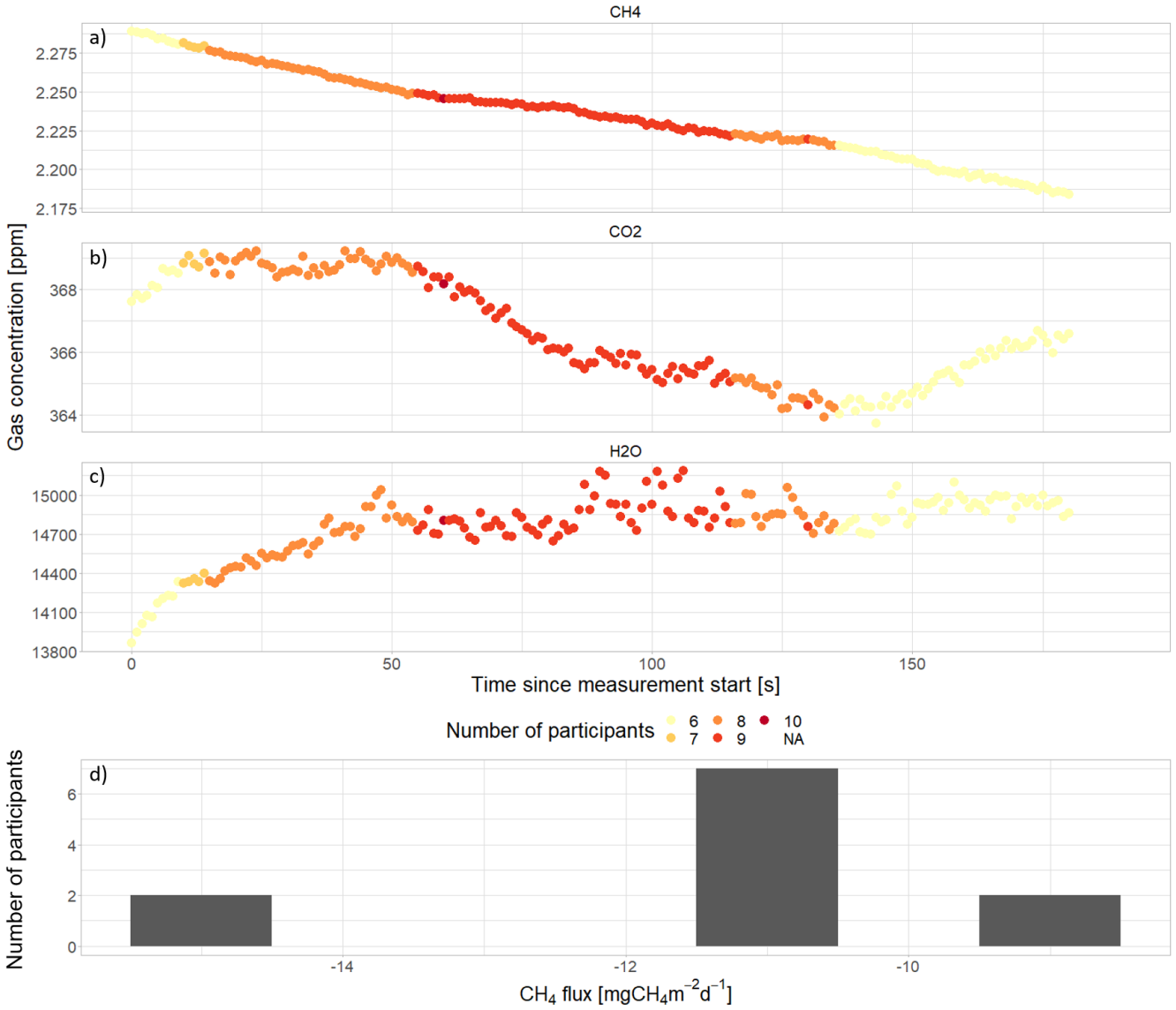



**Figure A1.11: Measurement example VQC3 of a small linear decrease in $CH_4$ concentrations over the time of the chamber closure. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants (d).**

### A1.7 No trend

In the visual QC exercise, we included one measurement example for which the $CH_4$ concentrations did not show a clear trend and varied only little over the time of the chamber closure (measurement ID VQC6, Figure A1.12). Most participants (69%) noticed the very small change in $CH_4$ concentrations over the whole measurement but they disagreed on whether the concentration measurements represented a real flux. Half of them suspected a real emission that remained very small because of $CH_4$ production and oxidation cancelling each other out at a low water table and two more participants called it a "zero flux" where the uncertainty would likely exceed the flux magnitude. Some (39%) of the participants, however, explained the low change in $CH_4$ concentrations by air leaking from the chamber, two of whom related the leak to vegetation obstructing the chamber seal and one to lateral diffusion into the chamber from the surrounding area.

Some (19%) of the participants furthermore pointed out an inconsistent trend in the $CH_4$ concentrations which they related to a changing balance between $CH_4$ production and oxidation over time, noisy measurements due to a low precision of the gas analyzer, or a bad chamber seal combined with wind disturbance. According to one participant the latter was supported by the fluctuations appearing in the concentrations of all three gases, while two other participants mentioned that the $CO_2$ concentrations looked linear, at least after 30 to 40 s into the measurement, indicating an intact chamber seal.

The majority of the participants (62%) decided to discard the measurement due to leakage from the chamber (38%), a changing trend in the $CH_4$ concentrations (44%), a too short measurement time (13%), or too high initial concentrations of $CH_4$ and $CO_2$ (13%). While two of these participants manually set the $CH_4$ flux to zero, one participant pointed out that the concentration changes were too large to be below the precision of the instrument so that the measurement should not be accepted as a zero flux. Some (31%) of the participants kept the measurement assuming a small but nonetheless real $CH_4$ flux and gave start and end times for flux calculation. Half of them discarded the beginning of the measurement as a period of initial equilibration, while the other half kept the entire measurement. The choice of different time periods for flux calculation by the participants resulted in a CV of 57% for this measurement example.

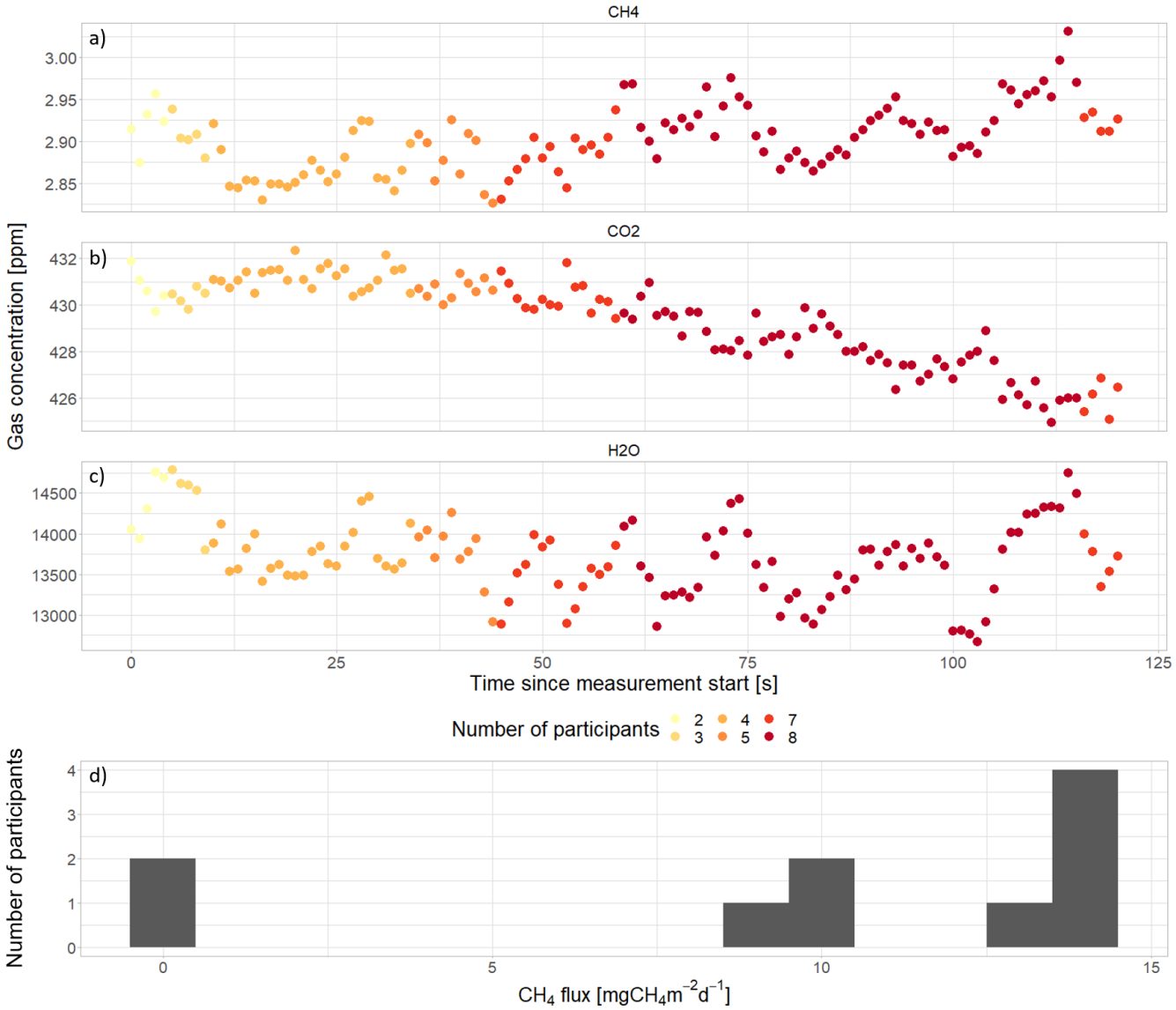



**Figure A1.12: Measurement example VQC6 with the $CH_4$ concentrations showing little variation without a clear trend over the time of the chamber closure. Simultaneous measurements of $CH_4$ (a), $CO_2$ (b), and $H_2O$ (c) concentrations over time during chamber closure with the colors of the data points indicating how many participants included the respective data point in the time period that they chose for flux calculation. Only the responses by participants who gave start and end times for flux calculation were considered for this figure. Histogram of $CH_4$ fluxes calculated based on the time period chosen for flux calculation by the participants including the two participants who set the flux to zero (d).**

*Author contributions.* CCT, MF, and KJ conceived the project idea. KJ drafted the questions for the expert survey, set up the final online version of the survey, and collected the Siikaneva data set from which a subset of measurements was used as examples for the visual QC exercise. CCT, LvD, and MF reviewed the survey questions and tested the survey. KJ analysed the survey responses and created the figures. The manuscript was written by KJ and commented on by all authors. CCT supervised the project.

*Competing interests*. The authors declare that they have no conflict of interest.

*Acknowledgements.* We thank Kathleen Hall, Katey Walter Anthony, Clayton Elder, Vasilii Petrenko for helpful discussions on methods that resulted in the conception of this project. We would like to thank Pamela Baur, Julia Boike, Jill Bubier, Jesper Christiansen, Scott J. Davidson, Lona van Delden, Bo Elberling, Kathleen Hall, Paul Hanson, Nicholas Hasson, Liam Heffernan, Jacqueline Hung, Vytas Huth, Gerald Jurasinski, Sari Juutinen, Masako Kajiura, Evan Kane, Aino Korrensalo, Elisa Männistö, Nicholas Nickerson, Genevieve Noyce, Frans-Jan Parmentier, Matthias Peichl, Norbert Pirk, Maria Strack, Eeva-Stiina Tuittila, Anna-Maria Virkkala, Carolina Voigt, Lei Wang as well as seven anonymous participants for completing the survey.

*Financial support.* The contribution of Katharina Jentzsch, Lona van Delden, and Claire C. Treat is part of the FluxWIN project, funded with a Starting Grant by the European Research Council (ERC) (ID 851181). The work by Katharina Jentzsch was supported by a fellowship of the German Academic Exchange Service (DAAD). CT also acknowledges support from the Pioneer Center Land-CRAFT at AU.

**References**

Bain, W.G., Hutyra, L., Patterson, D.C., Bright, A.V., Daube, B.C., Munger, J.W., and Wofsy, S.C.: Wind-induced error in the measurement of soil respiration using closed dynamic chambers, Agricultural and Forest Meteorology 131, 225–232, https://doi.org/10.1016/j.agrformet.2005.06.004, 2005.

Bond-Lamberty, B. and Thomson, A.: A global database of soil respiration data, Biogeosciences 7, 1915–1926, https://doi.org/10.5194/bg-7-1915-2010, 2010.

Bond-Lamberty, B. and Thomson, A.: Temperature-associated increases in the global soil respiration record, Nature 464, 579–582, https://doi.org/10.1038/nature08930, 2010.

Christiansen, J.R., Korhonen, J.F.J., 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 Soil 343, 171–185, https://doi.org/10.1007/s11104-010-0701-y, 2011.

Christiansen, J.R., Outhwaite, J., and Smukler, S.M.: Comparison of CO2, CH4 and N2O soil-atmosphere exchange measured in static chambers with cavity ring-down spectroscopy and gas chromatography, Agricultural and Forest Meteorology 211–212, 48–57, https://doi.org/10.1016/j.agrformet.2015.06.004, 2015.

Clough, T.J., Rochette, P., Thomas, S.M., Pihlatie, M., Christiansen, J.R., and Thorman, R.E.: Global Research Alliance $N_2O$ chamber methodology guidelines: Design considerations, J of Env Quality 49, 1081–1091, https://doi.org/10.1002/jeq2.20117, 2020.

Conen, F. and Smith, K.A.: A re-examination of closed flux chamber methods for the measurement of trace gas emissions from soils to the atmosphere, European J Soil Science 49, 701–707, https://doi.org/10.1046/j.1365-2389.1998.4940701.x, 1998.

de Klein, C.A. and Harvey, M.J., 2012. Nitrous oxide chamber methodology guidelines de Klein, C. A., & Harvey, M. J.

Fiedler, J., Fuß, R., Glatzel, S., Hagemann, U., Huth, V., Jordan, S., Jurasinski, G., Kutzbach, L., Maier, M., Schäfer, K., Weber, T., and Weymann, D.: BEST PRACTICE GUIDELINE Measurement of carbon dioxide, methane and nitrous oxide fluxes between soil-vegetation-systems and the atmosphere using non-steady state chambers, Arbeitsgruppe Bodengase, Deutsche Bodenkundliche Gesellschaft, https://doi.org/10.23689/FIDGEO-5422, 2012.

Forbrich, I., Kutzbach, L., Hormann, A., and Wilmking, M.: A comparison of linear and exponential regression for estimating diffusive CH4 fluxes by closed-chambers in peatlands, Soil Biology and Biochemistry 42, 507–515, https://doi.org/10.1016/j.soilbio.2009.12.004, 2010.

Frenzel, P. and Karofeld, E.: CH4 emission from a hollow-ridge complex in a raised bog: The role of CH4 production and oxidation, Biogeochemistry 51, 91–112, https://doi.org/10.1023/A:1006351118347, 2000.

Healy, R.W., Striegl, R.G., Russell, T.F., Hutchinson, G.L., and Livingston, G.P.: Numerical Evaluation of Static-Chamber Measurements of Soil-Atmosphere Gas Exchange: Identification of Physical Processes, Soil Science Society of America Journal 60, 740–747, https://doi.org/10.2136/sssaj1996.03615995006000030009x, 1996.

Hoffmann, M., Schulz-Hanke, M., Garcia Alba, J., Jurisch, N., Hagemann, U., Sachs, T., Sommer, M., and Augustin, J.: A simple calculation algorithm to separate high-resolution CH4 flux measurements into ebullition- and diffusion-derived components, Atmos. Meas. Tech. 10, 109–118, https://doi.org/10.5194/amt-10-109-2017, 2017.

Holland, E.A., Robertson, G.P., Greenberg, J.P., Groffman, P.M., Boone, R.D., and Gosz, J.R.: Soil CO2, N2O, and CH4 exchange, in: Standard Soil Methods for Long-Term Ecological Research, edited by: Robertson, G.P., Coleman, D.C., Bledsoe, C.S., and Sollins, P., Oxford University Press, Oxford, UK, pp. 185–201, 1999.

Hutchinson, G.L. and Livingston, G.P.: Vents and seals in non-steady-state chambers used for measuring gas exchange between soil and the atmosphere, European J Soil Science 52, 675–682, https://doi.org/10.1046/j.1365-2389.2001.00415.x, 2001.

Jentzsch, K., Männistö, E., Marushchak, M.E., Korrensalo, A., Van Delden, L., Tuittila, E.-S., Knoblauch, C., and Treat, C.C.: Shoulder season controls on methane emissions from a boreal peatland, Biogeosciences 21, 3761–3788, https://doi.org/10.5194/bg-21-3761-2024, 2024a.

Jentzsch, K., Van Delden, L., Fuchs, M., and Treat, C.C.: An expert survey on chamber measurement techniques and approaches for calculation and quality control of methane fluxes, https://doi.org/10.1594/PANGAEA.971695, 2024b.

Jian, J., Vargas, R., Anderson-Teixeira, K., Stell, E., Herrmann, V., Horn, M., Kholod, N., Manzon, J., Marchesi, R., Paredes, D., and Bond-Lamberty, B.: A restructured and updated global soil respiration database (SRDB-V5), https://doi.org/10.5194/essd-2020-136, 2020.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J.G., Dlugokencky, E.J., Bergamaschi, P., Bergmann, D., Blake, D.R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E.L., Houweling, S., Josse, B., Fraser, P.J., Krummel, P.B., Lamarque, J.-F., Langenfelds, R.L., Le Quéré, C., Naik, V., O'Doherty, S., Palmer, P.I., Pison, I., Plummer, D., Poulter, B., Prinn, R.G., Rigby, M., Ringeval, B., Santini, M., Schmidt, M., Shindell, D.T., Simpson, I.J., Spahni, R., Steele, L.P., Strode, S.A., Sudo, K., Szopa, S., Van Der Werf, G.R., Voulgarakis, A., Van Weele, M., Weiss, R.F., Williams, J.E., and Zeng, G.: Three decades of global methane sources and sinks, Nature Geosci 6, 813–823, https://doi.org/10.1038/ngeo1955, 2013.

Kuhn, M.A., Varner, R.K., Bastviken, D., Crill, P., MacIntyre, S., Turetsky, M., Walter Anthony, K., McGuire, A.D., and Olefeldt, D.: BAWLD-CH4: a comprehensive dataset of methane fluxes from boreal and arctic ecosystems, Earth Syst. Sci. Data 13, 5151–5189, https://doi.org/10.5194/essd-13-5151-2021, 2021.

Kutzbach, L., Schneider, J., Sachs, T., Giebels, M., Nykänen, H., Shurpali, N.J., Martikainen, P.J., Alm, J., and Wilmking, M.: CO2 flux determination by closed-chamber methods can be seriously biased by inappropriate application of linear regression, Biogeosciences 4, 1005–1025, https://doi.org/10.5194/bg-4-1005-2007, 2007.

Laine, A., Wilson, D., Kiely, G., and Byrne, K.A.: Methane flux dynamics in an Irish lowland blanket bog, Plant Soil 299, 181–193, https://doi.org/10.1007/s11104-007-9374-6, 2007.

Livingston, G.P. and Hutchinson, G.L.: Enclosure-based measurement of trace gas exchange: applications and sources of error, in: Biogenic Trace Gases: Measuring Emissions from Soil and Water, edited by: Matson, P.A., Harriss, R.C., Blackwell Science Ltd., Oxford, UK, pp. 14–51, 1995.

Macreadie, P.I., Anton, A., Raven, J.A., Beaumont, N., Connolly, R.M., Friess, D.A., Kelleway, J.J., Kennedy, H., Kuwae, T., Lavery, P.S., Lovelock, C.E., Smale, D.A., Apostolaki, E.T., Atwood, T.B., Baldock, J., Bianchi, T.S., Chmura, G.L., Eyre, B.D., Fourqurean, J.W., Hall-Spencer, J.M., Huxham, M., Hendriks, I.E., Krause-Jensen, D., Laffoley, D., Luisetti, T., Marbà, N., Masque, P., McGlathery, K.J., Megonigal, J.P., Murdiyarso, D., Russell, B.D., Santos, R., Serrano, O., Silliman, B.R., Watanabe, K., and Duarte, C.M.: The future of Blue Carbon science, Nat Commun 10, 3998, https://doi.org/10.1038/s41467-019-11693-w, 2019.

Maier, M., Weber, T.K.D., Fiedler, J., Fuß, R., Glatzel, S., Huth, V., Jordan, S., Jurasinski, G., Kutzbach, L., Schäfer, K., Weymann, D., and Hagemann, U.: Introduction of a guideline for measurements of greenhouse gas fluxes from soils using non-steady-state chambers, J. Plant Nutr. Soil Sci. 185, 447–461, https://doi.org/10.1002/jpln.202200199, 2022.

Melton, J.R., Wania, R., Hodson, E.L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C.A., Beerling, D.J., Chen, G., Eliseev, A.V., Denisov, S.N., Hopcroft, P.O., Lettenmaier, D.P., Riley, W.J., Singarayer, J.S., Subin, Z.M., Tian, H., Zürcher, S., Brovkin, V., Van Bodegom, P.M., Kleinen, T., Yu, Z.C., and Kaplan, J.O.: Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP), Biogeosciences 10, 753–788, https://doi.org/10.5194/bg-10-753-2013, 2013.

Moore, T.R. and Knowles, R.: Methane emissions from fen, bog and swamp peatlands in Quebec, Biogeochemistry 11, https://doi.org/10.1007/BF00000851, 1990.

Morgan, M.G.: Use (and abuse) of expert elicitation in support of decision making for public policy, Proc. Natl. Acad. Sci. U.S.A. 111, 7176–7184, https://doi.org/10.1073/pnas.1319946111, 2014.

Mosier, A.R., Delgado, J.A., Cochran, V.L., Valentine, D.W., and Parton, W.J.: Impact of agriculture on soil consumption of atmospheric CH4 and a comparison of CH4 and N2O flux in subarctic, temperate and tropical grasslands, Nutrient Cycling in Agroecosystems 49, 71–83, https://doi.org/10.1023/A:1009754207548, 1997.

Neubauer, S.C.: Global Warming Potential Is Not an Ecosystem Property, Ecosystems 24, 2079–2089, https://doi.org/10.1007/s10021-021-00631-x, 2021.

Nickerson, N.R.: Evaluating gas emission measurements using Minimum Detectable Flux (MDF), https://doi.org/10.13140/RG.2.1.4149.2089, 2016.

Pedersen, A.R., Petersen, S.O., and Schelde, K.: A comprehensive approach to soil-atmosphere trace-gas flux estimation with static chambers, European Journal of Soil Science 61, 888–902, https://doi.org/10.1111/j.1365-2389.2010.01291.x, 2010.

Pihlatie, M.K., Christiansen, J.R., Aaltonen, H., Korhonen, J.F.J., Nordbo, A., Rasilo, T., Benanti, G., Giebels, M., Helmy, M., Sheehy, J., Jones, S., Juszczak, R., Klefoth, R., Lobo-do-Vale, R., Rosa, A.P., Schreiber, P., Serça, D., Vicca, S., Wolf, B., and Pumpanen, J.: Comparison of static chambers to measure CH4 emissions from soils, Agricultural and Forest Meteorology 171–172, 124–136, https://doi.org/10.1016/j.agrformet.2012.11.008, 2013.

Pirk, N., Mastepanov, M., Parmentier, F.-J.W., Lund, M., Crill, P., and Christensen, T.R.: Calculations of automatic chamber flux measurements of methane and carbon dioxide using short time series of concentrations, Biogeosciences 13, 903–912, https://doi.org/10.5194/bg-13-903-2016, 2016.

Pumpanen, J., Kolari, P., Ilvesniemi, H., Minkkinen, K., Vesala, T., Niinistö, S., Lohila, A., Larmola, T., Morero, M., Pihlatie, M., Janssens, I., Yuste, J.C., Grünzweig, J.M., Reth, S., Subke, J.-A., Savage, K., Kutsch, W., Østreng, G., Ziegler, W., Anthoni, P., Lindroth, A., and Hari, P.: Comparison of different chamber techniques for measuring soil CO2 efflux, Agricultural and Forest Meteorology 123, 159–176, https://doi.org/10.1016/j.agrformet.2003.12.001, 2004.

Rheault, K., Christiansen, J.R., and Larsen, K.S.: goFlux: A user-friendly way to calculate GHG fluxesyourself, regardless of user experience, JOSS 9, 6393, https://doi.org/10.21105/joss.06393, 2024.

Rochette, P. and Eriksen-Hamel, N.S.: Chamber Measurements of Soil Nitrous Oxide Flux: Are Absolute Values Reliable? Soil Science Soc of Amer J 72, 331–342, https://doi.org/10.2136/sssaj2007.0215, 2008.

Rosentreter, J.A., Alcott, L., Maavara, T., Sun, X., Zhou, Y., Planavsky, N.J., and Raymond, P.A.: Revisiting the Global Methane Cycle Through Expert Opinion, Earth's Future 12, e2023EF004234, https://doi.org/10.1029/2023EF004234, 2024.

Saunois, M., Stavert, A.R., Poulter, B., Bousquet, P., Canadell, J.G., Jackson, R.B., Raymond, P.A., Dlugokencky, E.J., Houweling, S., Patra, P.K., Ciais, P., Arora, V.K., Bastviken, D., Bergamaschi, P., Blake, D.R., Brailsford, G., Bruhwiler, L., Carlson, K.M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P.M., Covey, K., Curry, C.L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M.I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K.M., Joos, F., Kleinen, T., Krummel, P.B., Langenfelds, R.L., Laruelle, G.G., Liu, L., Machida, T., Maksyutov, S., McDonald, K.C., McNorton, J., Miller, P.A., Melton, J.R., Morino, I., Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R.J., Peng, C., Peng, S., Peters, G.P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W.J., Rosentreter, J.A., Segers, A., Simpson, I.J., Shi, H., Smith, S.J., Steele, L.P., Thornton, B.F., Tian, H., Tohjima, Y., Tubiello, F.N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T.S., Van Weele, M., Van Der Werf, G.R., Weiss, R.F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Qing, Zhu, Qiuan, and Zhuang, Q.: The Global Methane Budget 2000–2017, Earth Syst. Sci. Data 12, 1561–1623, https://doi.org/10.5194/essd-12-1561-2020, 2020.

Schuur, E.A.G., Abbott, B.W., Bowden, W.B., Brovkin, V., Camill, P., Canadell, J.G., Chanton, J.P., Chapin, F.S., Christensen, T.R., Ciais, P., Crosby, B.T., Czimczik, C.I., Grosse, G., Harden, J., Hayes, D.J., Hugelius, G., Jastrow, J.D., Jones, J.B., Kleinen, T., Koven, C.D., Krinner, G., Kuhry, P., Lawrence, D.M., McGuire, A.D., Natali, S.M., O'Donnell, J.A., Ping, C.L., Riley, W.J., Rinke, A., Romanovsky, V.E., Sannel, A.B.K., Schädel, C., Schaefer, K., Sky, J., Subin, Z.M., Tarnocai, C., Turetsky, M.R., Waldrop, M.P., Walter Anthony, K.M., Wickland, K.P., Wilson, C.J., and Zimov, S.A.: Expert assessment of vulnerability of permafrost carbon to climate change, Climatic Change 119, 359–374, https://doi.org/10.1007/s10584-013-0730-7z, 2013.

Strack, M., Kellner, E., Waddington, J.M.: Dynamics of biogenic gas bubbles in peat and their effects on peatland biogeochemistry, Global Biogeochemical Cycles 19, 2004GB002330, https://doi.org/10.1029/2004GB002330, 2005.

Treat, C.C., Bloom, A.A., and Marushchak, M.E.: Nongrowing season methane emissions–a significant component of annual emissions across northern ecosystems, Global Change Biology 24, 3331–3343, https://doi.org/10.1111/gcb.14137, 2018.

Virkkala, A.-M., Niittynen, P., Kemppinen, J., Marushchak, M.E., Voigt, C., Hensgens, G., Kerttula, J., Happonen, K., Tyystjärvi, V., Biasi, C., Hultman, J., Rinne, J., and Luoto, M.: High-resolution spatial patterns and drivers of terrestrial ecosystem carbon dioxide, methane, and nitrous oxide fluxes in the tundra, Biogeosciences 21, 335–355, https://doi.org/10.5194/bg-21-335-2024, 2024.

Voigt, C., Virkkala, A.-M., Hould Gosselin, G., Bennett, K.A., Black, T.A., Detto, M., Chevrier-Dion, C., Guggenberger, G., Hashmi, W., Kohl, L., Kou, D., Marquis, C., Marsh, P., Marushchak, M.E., Nesic, Z., Nykänen, H., Saarela, T., Sauheitl, L., Walker, B., Weiss, N., Wilcox, E.J., and Sonnentag, O.: Arctic soil methane sink increases with drier conditions and higher ecosystem respiration, Nat. Clim. Chang. 13, 1095–1104, https://doi.org/10.1038/s41558-023-01785-3, 2023.

Waddington, J.M. and Roulet, N.T.: Atmosphere-wetland carbon exchanges: Scale dependency of $CO_2$ and $CH_4$ exchange on the developmental topography of a peatland, Global Biogeochemical Cycles 10, 233–245, https://doi.org/10.1029/95GB03871, 1996.

Xu, L., Furtaw, M.D., Madsen, R.A., Garcia, R.L., Anderson, D.J., and McDermitt, D.K.: On maintaining pressure equilibrium between a soil $CO_2$ flux chamber and the ambient air, J. Geophys. Res. 111, 2005JD006435, https://doi.org/10.1029/2005JD006435, 2006.