Response to interactive comment (Anonymous Referee #1) on 'Gas chromatography vs. quantum cascade laser-based N_2O flux measurements using a novel chamber design'

[R#1.1] Brümmer et al. present a study analyzing linear and non-linear flux calculation methods under high and low flux rates of nitrous oxide and different scenarios of closure time. They use both traditional gas chromatography (GC) (low sampling number during closure time) and high-resolution quantum cascade laser (QCL) sampling. They find that non linear concentration changes are more clearly detectable during high emission scenarios and long chamber closure. Shortening of closure time results in a reasonable agreement between linear (3min) and non-linear (60min) flux estimates, but can only be applied when using the QCL set up. While under low flux conditions, GC measurements result in more scattered flux estimates, in both campaigns mean flux estimates of GC and QCL agreed well. Rare negative fluxes detected by GC measurements seem to be arbitrary and not caused by actual N2O uptake. The paper is well written and a good fit for the journal.

However, I could not help thinking that most of the results were as to be expected from literature and not 'radically' new.

[AC#1.1] We highly appreciate the comments and suggestions given by Anonymous Referee #1. We agree that some results like higher non-linearity in concentration changes under higher emission regimes and longer closure times have been hypothesized and reported earlier. The basic idea of this study is to give a concise overview by showing a side-by-side comparison of QCL vs. GC characteristics, low and high exchange regimes, linear and non-linear flux calculation methods alongside a presentation of our custom-built chamber design. Many other papers, however, usually deal with only a few of the above mentioned components, i.e. either low vs. high fluxes, or only with a GC vs. QCL comparison, or purely with different calculation methods. Therefore, we aimed at integrating the characterization of the measurement system, the exchange regime, and the flux calculation by means of two short campaigns without going into too extensive analyses.

Changes to the manuscript:
None specifically for this comment, but responses to comments R#2.4 and R#2.12 from Reviewer #2 deal with similar topics. See AC#2.4 and AC#2.12.

[R#1.2] Ultimately, the high temporal resolution of measurements possible with QCL (which do provide more sophisticated ways of flux data processing) and the fact that the concentration measurements are instantaneous, make these measurements desirable for exactly the long-term applications, the authors are suggesting.

[AC#1.2] The novel QCL application combines multiple advantages over traditional manual sampling systems. These are (amongst others)

- higher temporal resolution of concentration data leading to a higher number of flux rates per day,
- the possibility for robust application of flux calculation procedures,
- easy determination of system malfunction, e.g. caused by insufficiently closed chambers,
- low maintenance for laser operation,

- low uncertainty in flux estimates providing the opportunity for ecological process studies and calculating robust trace gas budgets

For those cases where QCL methodology cannot be applied, e.g. due to high initial investment costs, GC-based measurements may still be useful when investigating longer periods when the focus is not on short-term variability of gas exchange dynamics.

Changes to the manuscript:
None.

[R#1.3] It is not clear to me, what the accessibility of the described instrumentation is. Is there a plan to make it available for other users, i.e. to 'rent' it out or to make it available within the ICOS project? If that is the case, it should be pointed out more clearly.

[AC#1.3] We thank the reviewer for this comment. It is a good idea to promote the presented chamber design more clearly as it meets the anticipated standards listed in the ICOS protocol for chamber measurements. That protocol, which will be made publicly available soon by the Ecosystem Thematic Center of ICOS, does not explicitly state precise mandatory dimensions for chamber volume and design, but rather provides size ranges depending on ecosystem type. Information about our chamber system including the construction plan is open to the scientific community and can be requested from the authors. We add the respective information at the end of Chapter 2.1.

Changes to the manuscript:
Sentence added at the end of Chapter 2.1: 'Information about our chamber system including the construction plan is open to the scientific community and can be requested from the authors.'

[R#1.4] Overall, the most interesting aspect to me is the possibility to study ecological processes in a new way, as shown for the possible net N2O uptake and diurnal variability in emission rates. Interestingly, the study they compare their results to (Shurpali et al. 2016) is mostly an eddy covariance study. It would be interesting if the authors could comment on possible advantages of this automatic chamber against eddy covariance and whether other gases can be sampled in parallel to N2O (I am thinking mostly of CO2, considering the possible coupling of plant activity and N2O emission rates).

[AC#1.4] One advantage of chamber measurements in comparison with an eddy-covariance approach is the possibility to study small-scale spatial variability of greenhouse gas exchange. This can either be done in natural homogeneous environments or in specific trials at plot scale, e.g. when different types and amount of fertilizers are applied on relatively small plots of a few square meters where the eddy-covariance approach would fail as it requires a homogeneous fetch of up to a few hectares around the tower. Secondly, continuous automated measurements using QCL spectrometry for trace gas analysis like in our study do provide robust estimates of exchange fluxes in situations where assumptions of the eddy-covariance theory are violated. These situations are for example low atmospheric turbulence conditions that frequently occur during nighttime or when measurements are conducted in hilly terrain and advective flows cause significant bias in EC-based fluxes.

As many laser spectrometers that are currently available on the market allow for parallel detection of selected other trace gases – usually CH_4 and CO_2 – in one analyzer cell, our sampling setup can simultaneously provide concentrations and flux estimates of the chosen greenhouse gases to study coupled environmental processes such as effects of water table, soil moisture and temperature on the respective gases of interest.
We will add the information that parallel detection of different trace gases is possible with most common analyzers in combination with our chamber system.

Changes to the manuscript:
Sentence modified at the end of Chapter 3.4: 'Our study highlights that through its high time resolution QCL-based measurements will not only help enhance process understanding of N_2O exchange by disentangling the strength of different drivers of N_2O production like temperature, soil moisture, nitrogen availability, and microbial activity, but has also the potential to provide new insight into bidirectional exchange characteristics of other trace gases such as CH_4, which can be sampled simultaneously with our chamber system depending on analyzer type used.'

Response to interactive comment (Anonymous Referee #2) on 'Gas chromatography vs. quantum cascade laser-based N_2O flux measurements using a novel chamber design'

[R#2.1] The authors have tested the performance of a QCL analyzer connected to a new automated chamber, against a "conventional" GC + automated gas sampling unit system. Data from QCL system were used to observe the non-linearity in the concentration increase during the chamber closure. Based on two short campaigns, the paper gives recommendations how should the measurements, data screening and flux calculations be done. The new chamber design is interesting and the system coupled to QCL seems to be fluently producing nice data. Papers presenting new chamber designs, are always welcome, particularly if they can provide generalizations and recommendations which are useful for other chamber operators. The paper is fluently written, and the observation of different patterns in diurnal cycle is interesting and important. However, there are several deficiencies and pitfalls in the data treatment and the argumentation which need revision. The presentation quality would benefit from separating the results and discussion.

[AC#2.1] We sincerely thank Referee #2 for his/her thorough review. Through the consideration and inclusion of his/her meaningful comments and suggestions, we feel that the manuscript's quality has improved, particularly by shaping the main conclusions and take home messages. We also have streamlined the presentation of the main findings. Below we give our responses to all points raised by the reviewer plus short statements of the actual changes in the manuscript.

Changes to the manuscript:
Main changes are (see specific points below for details):

- Reformulation of the aims of the paper at the end of the Introduction
- Inclusion of a table summarizing main features of the chamber system by providing quantitative measures
- Streamlining the text through splitting up Results and Discussion sections
- Providing a clear story line of investigations (as can be seen by the reformulation of the aims and the newly structured table of contents)
- Shaping up some of the conclusions (see specific points below)
- Few changes to figures (see specific points below)

The restructured aims of the paper now read as follows:

(1) Presentation of a novel chamber design that is connected to both a vial air-sampling setup with subsequent GC analysis and a QCL spectrometer
    - Description of design and setup in Sections 2.1 and 2.2
    - New chamber system is used for the following investigations (aims 2 to 5)
(2) Characterization of the shape of the concentration increase
    - $\rightarrow$ Is the shape rather linear or non-linear?
    - Quantification of the curvature (κ) in concentration increase
    - Using κ to verify chamber density
        - $\rightarrow$ Is κ dependent on wind speed, wind direction, on the flux itself or on closure time?
(3) Comparison of N_2O fluxes and their associated standard errors from linear and non-linear regression models
(4) Testing the novel chamber system under high and low flux conditions and comparing GC vs. QCL-based flux estimates
(5) Investigation of ecosystem and climate-specific flux characteristics such as N_2O uptake and diurnal variation

The single paragraphs of the Results and Discussion section are now as follows:

3. Results
   - 3.1 Shape of concentration increase and curvature (κ) determination
   - 3.2 Comparison of N_2O fluxes and their associated errors from linear and non-linear regression models
   - 3.3 GC vs. QCL-based fluxes under high and low exchange regimes
   - 3.4 N_2O uptake and diurnal variation
4. Discussion
   - 4.1 The parameter κ as a chamber performance criteria
   - 4.2 Closure time and measurement frequency – How long and how often is enough?
   - 4.3 Differences between GC and QCL-based fluxes
   - 4.4 Enabling investigations of flux pattern characteristics

[R#2.2] First, the performance of the GC sampling system makes me wonder whether the comparison of two systems is meaningful. Before making any comparisons, the authors should find out the reason for the bad performance of the GC.

[AC#2.2] Please see detailed response to [R#2.6].

Changes to the manuscript:
Please see detailed response to [R#2.6].

[R#2.3] Secondly, there are several conclusions in the paper which are just qualitative, and as such they are vague and are not supported by the presented data.

[AC#2.3] We fully agree that some conclusions came a bit out of the blue. We have modified the respective sections as outlined in the specific comments to [R#2.7], [R#2.9], [R#2.11], [R#2.33], [R#2.34], and [R#2.39] below.

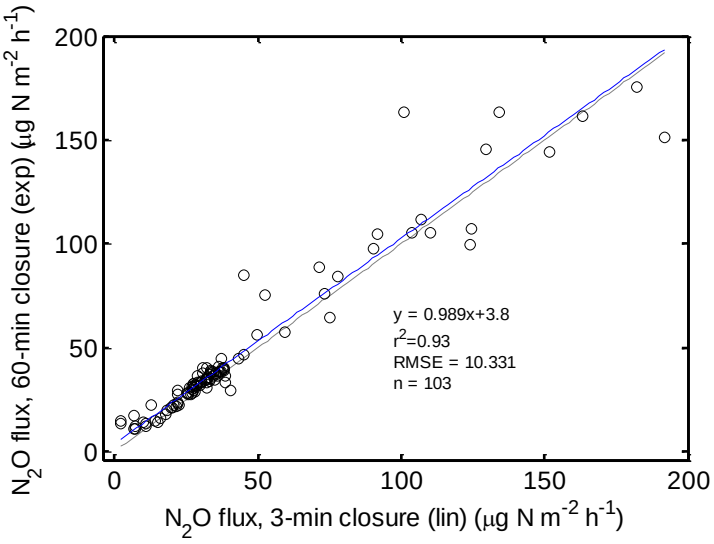
Changes to the manuscript:
See responses to comments [R#2.7], [R#2.9], [R#2.11], [R#2.33], [R#2.34], and [R#2.39] below.

[R#2.4] Third, I share the worry of the first reviewer that most of the results shown here are already well known. For example, it has been reported already in numerous papers that the curvature in concentration increase is higher with longer closure time, and that using the linear calculation instead of non-linear can result in great underestimate in flux rate. Instead of reporting curvature, it would be more useful to quantify what is the limit of curvature after which the authors recommend the use of non-linear fitting method.

[AC#2.4] We agree that higher curvature in concentration increase at longer closure time has been hypothesized and shown before. But to our knowledge, it hasn't been quantified and neither its dependency on N_2O fluxes (Fig. 3A) nor on chamber performance criteria like the insensitivity towards wind speed and direction (Fig. 3C and 3D) has been explicitly analyzed like in our study. This is clearly a new investigation alongside presenting a novel chamber design. Also, flux underestimation when using linear instead of non-linear regression may certainly be true for GC measurements when only a limited number of samples are available. But the point in our paper is that we on the hand highlight the advantages of QCL measurements (high time resolution, low standard errors of fluxes; *cf*. Figs. 2, 4, 5, 6) and then recommend to reduce chamber closure time to be able to apply linear regression (see modified Fig. 4D for better visualization of low differences between the application of linear

vs. non-linear regression for flux calculation). We also attach here graphs showing the flux difference, i.e. non-linear–linear, plotted against curvature (Fig. R1). Note that during shorter closure time (10 min; blue circles), relatively small (absolute) differences between the two calculation methods occur, although curvature was highly variable and single κ values up to –1000 μg N m^{-3} h^{-2} were found.

We fully agree that the manuscript would benefit from streamlining the aims, results, and messages towards a more concise overview of useful take home conclusions for the reader. See 'Changes to the manuscripts' at [AC#2.1] for the main modifications. However, regarding curvature in this study, it should not be used to define a threshold after which linear over non-linear flux calculation should be used as it is supposed to demonstrate chamber performance criteria as highlighted in Fig. 3C and 3D. Together with the new Fig. 4D and its slope close 1 in the most common flux range between 0 and 200 μg N m^{-2} h^{-1}, we prefer using the argument throughout the manuscript that reducing chamber closure time and applying linear regression for flux calculation is a valid approach.



New Figure 4D: Linear regression analysis of N_2O fluxes <200 μg N m^{-2} h^{-1} with adapted regression from the exponential vs. the linear model.

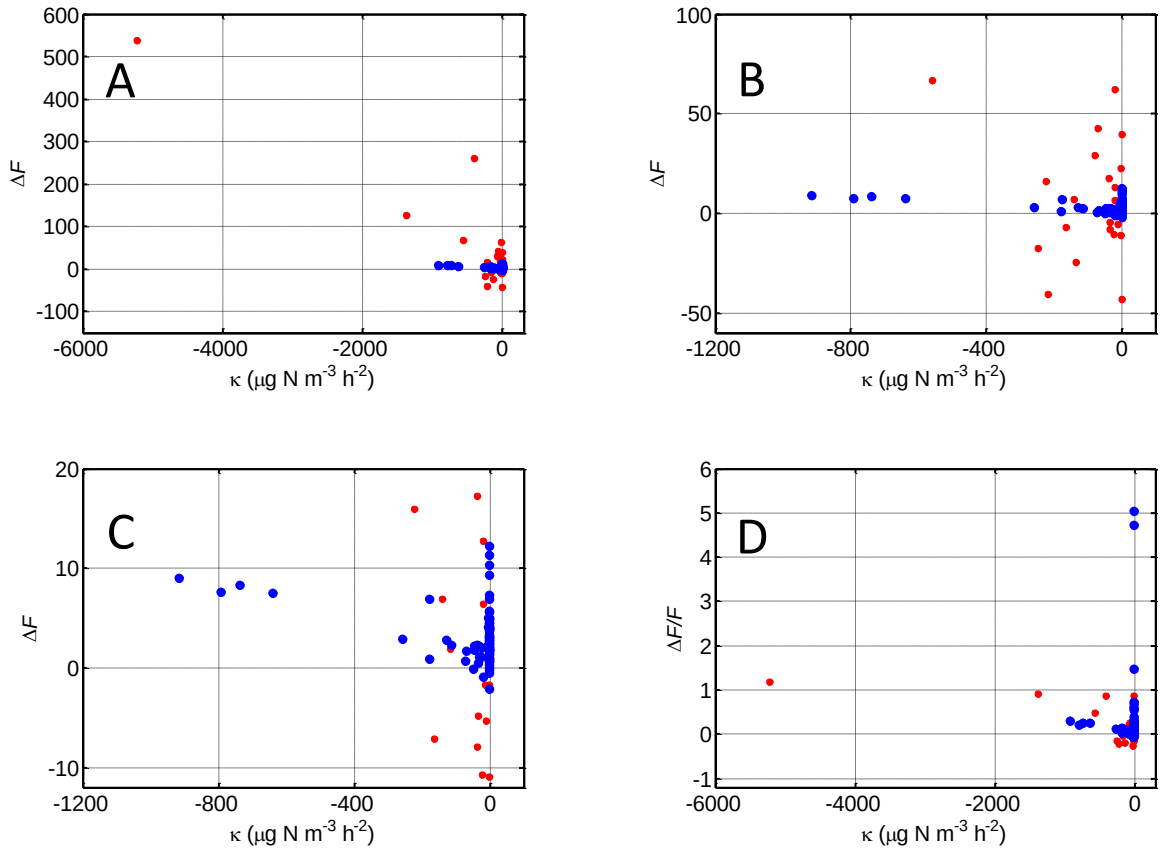


Figure R1 (will not be shown in the manuscript): Panels A, B, and C: Dependency of ΔF, i.e. N_2O fluxes from non-linear regression–linear regression, on κ values for different ranges. Panel D: Dependency of normalized flux difference on κ values.

Changes to the manuscript:
See [AC#2.1] plus revised Fig. 4D and 5.

[R#2.5] Also, it would have been interesting to learn more about the advantages and possible problems in the "novel" chamber design. In general, the paper could be more valuable would it provide more quantitative information and recommend some general tests which each chamber operator should run to ensure adequate data quality. It is also a bit questionable if the paper with such a short piece of data (25 + 6 days) is enough the draw firm conclusions. See more comments below.

[AC#2.5] We highly appreciate this comment and agree that a concise overview of system features will help the reader to get familiar with chamber and instrumentation characteristics. These information including flux detection limit, closure time, number of daily cycles, sampling frequency, etc., are summarized in Table 2. Other more qualitative features are given in Section 2.1. A general test every operator should perform is a (somewhat indirect) density test with a calculation of standard errors of fluxes under different flux magnitudes where the shape of the concentration increase/decrease appears to be valid such as in Fig. 2B at DOY 107.8. The standard errors of these 'good fluxes' should be taken as a reference. Operators should inspect all fluxes that deviate largely from those reference values. However, these absolute numbers predominantly depend on the precision of the analyzers that are used, thus making it difficult to provide specific thresholds of errors after which flux values should generally be discarded. In our study, QCL-based fluxes with a standard error >3 µg N m^{-2} h^{-1} have undergone further double-checking. Only a few of those remained

plausible as can be seen in Fig. 4C. Regarding the ‘short piece of data’, we think that the value of a methodological study is not necessarily depending on the length of the observation. In fact moving the same systems to different places instead of measuring longer at one site, increased the range of test conditions, which has added value to the study and made its conclusions more robust. The conditions varied from very low to high fluxes, high external wind speeds (Risø), moderate wind speeds (Braunschweig), lower and higher temperatures. The effects were clear and thus we feel that performing longer tests would not have considerably increased the information with respect to the objectives of the study.

Changes to the manuscript:
Inclusion of Table 2 with quantitative features; we also add some rather qualitative characteristics in Section 2.1 such as the fact that the size of the chamber allows for investigations including plants of considerable size (even up to rape seed; publication in preparation) and that lifting the chamber diagonally away from the soil frame reduces shading for radiation and precipitation, thereby keeping the measurement spots as natural as possible.

Table 2: Features of the chamber-analyzer system used in this study.

	GC* (model: Shimadzu GC-2014)	QCL* (model: Aerodyne Research Inc. mini-QCLAS)
No. of chambers	3	3
Chamber closure time	60 min	60 min 10 min (recommended)
Sampling frequency	every 20 min	0.1 sec (max) 5 sec (recommended)
No. of concentration records per chamber run	4	36000 in 60 min 6000 in 10 min
No. of chamber cycles per day	24 (max)	72 (recommended) 144 (max)
Maximum number of samples	168 (depending on autosampler size)	Limited only by data storage capacity of QCL's computer or external hard drive
Lag time	(~10 sec)	~10 sec
N_2O flux detection limit ($\mu g\ N\ m^{-2}\ h^{-1}$)	13.0	2.6
Mean campaign N_2O flux ($\mu g\ N\ m^{-2}\ h^{-1}$)	BS (pref.[1]): 6.42 Risø (pref.[1]): 77.40	BS (lin.): 7.77 Risø (lin.[2]): 122.95
Mean campaign SE of N_2O fluxes ($\mu g\ N\ m^{-2}\ h^{-1}$)	BS (pref.[1]): 5.98 Risø (pref.[1]): 8.17	BS (lin.): 0.13 Risø (lin.[2]): 0.21
Median campaign N_2O flux ($\mu g\ N\ m^{-2}\ h^{-1}$)	BS (pref.[1]): 5.15 Risø (pref.[1]): 64.80	BS (lin.): 7.38 Risø (lin.[2]): 105.43
Median campaign SE of N_2O fluxes ($\mu g\ N\ m^{-2}\ h^{-1}$)	BS (pref.[1]): 5.04 Risø (pref.[1]): 4.72	BS (lin.): 0.10 Risø (lin.[2]): 0.17
Percentage of flux estimates where HMR could be fitted	BS: 8.5 % Risø: 37.9 %	BS: 100 % Risø: 100 %

GC – Gas chromatograph, QCL – Quantum cascade laser spectrometer, [1]preferred means non-linear HMR model was used if applicable, otherwise robust linear regression was taken, [2]mean/median of DOY 105.5 to 108.5 to make it comparable to GC data set

MORE DETAILED COMMENTS:

[R#2.6] What is the reason for the very bad performance of the GC system? On p5 lines 7-10 it is said that the system was checked against ten samples of ambient air, and only if the CV falls <3%, the data is acceptable. Is this CV limit of 3% really acceptable for a GC system? From Figs. 2a and 7a it seems clear that the GC is not able to resolve concentration increases for fluxes < 20 µgN m-2 h-1. At least to my knowledge, much lower fluxes analyzed with the GC have been reliably reported. I think that a comparison between QCL and GC is not really meaningful if GC is not able to measure these "small" fluxes of N2O. However, Fig. 2a makes me doubt, whether the problem is in the autosampler, and not in the GC detection limit? In some cases the GC can quite perfectly detect a concentration increase of about 12 ppb's similarly to the QCL (second measurement of DOY 339), but during many other closures the data seems arbitrary. What is the reason for that?

[AC#2.6] We appreciate the reviewer's concerns, but do not agree with the premise that a CV below 3 % (it was mainly close to 2 % in our study) at ambient concentrations is a 'very bad performance'. Inter-laboratory comparisons within Germany have shown that GC systems commonly exhibit CVs in this range during routine operations (publication in preparation). Also compare with Parkin et al. (2012), who show a CV of 4.4 % as an example in their Fig. 2. Based on this they calculated a detection limit of about 35 ppb h^{-1} for the linear flux model (see their Fig. 6, corresponds to about 40 µg N h^{-1}) and even higher detection limits for non-linear flux calculation schemes (which however reduce bias). Furthermore, detection limits should be determined based on statistics and not based on single flux measurements (e.g., on DOY 339). The median standard error of GC based flux measurements in our campaigns was SE = 6.5 µg N m^{-2} h^{-1}, thus the detection limit is approximately DL = 2 * SE = 13.0 µg N m^{-2} h^{-1}. See also response to Short Comment 3 for details.

Changes to the manuscript:
None.

[R#2.7] There are conclusions in the paper which are not supported by the presented data. For example: p.1 line 30: "new chamber design reduces the disturbance of the soil". There was nothing on that in the results. What are the possible disturbances? How can you detect that those can be omitted by your system?

[AC#2.7] This statement is simply related to the way the chamber is lifted and dragged away from the collar spot in a 45° angle. In comparison to many other chamber designs, soil and vegetation inside the soil collar are thereby kept under as natural conditions as possible, because the positions of the chambers when they are not operating largely prevent unintended shading and do not disturb throughfall, which is important when the chamber system is supposed to run for a longer time. This information was already given at the end of Section 2.1. We will slightly rephrase the sentence in the Abstract.

Changes to the manuscript:
Sentence modified to: 'Our new chamber design prevents the measurement spot from unintended shading and minimizes disturbance of throughfall, thereby complying with high quality requirements of long-term observation studies and research infrastructures.'

[R#2.8] Or: lines 25-26: GC was found to be a useful method to determine N2O fluxes at longer time scale". Where is the data to prove such a conclusion? There were no budgets

calculated. What happens with the low fluxes, how can you reliably determine budget if you cannot detect the flux?

[AC#2.8] See response to [AC#2.39].

Changes to the manuscript:
See response to [AC#2.39].

[R#2.9] Or p.8 line 31 forward: how do you justify the recommendation of removing the first 2 minutes of data?

[AC#2.9] We agree that the reader must have been puzzled by this sudden recommendation without showing any data. This statement arises from an observation we made in the increase pattern of the concentrations. In ~5 % of the cases, a somewhat irregular pattern as shown in the figure below was observed. It only happened right after setting the chamber onto the soil collar so maybe it was caused pressure fluctuations. We could not identify any correlations to either environmental or internal system conditions when this pattern was found. We therefore think it is a reasonable security procedure to remove the first two minutes (because it never exceeded this initial period) of data from a chamber cycle to ensure natural steady state soil efflux.

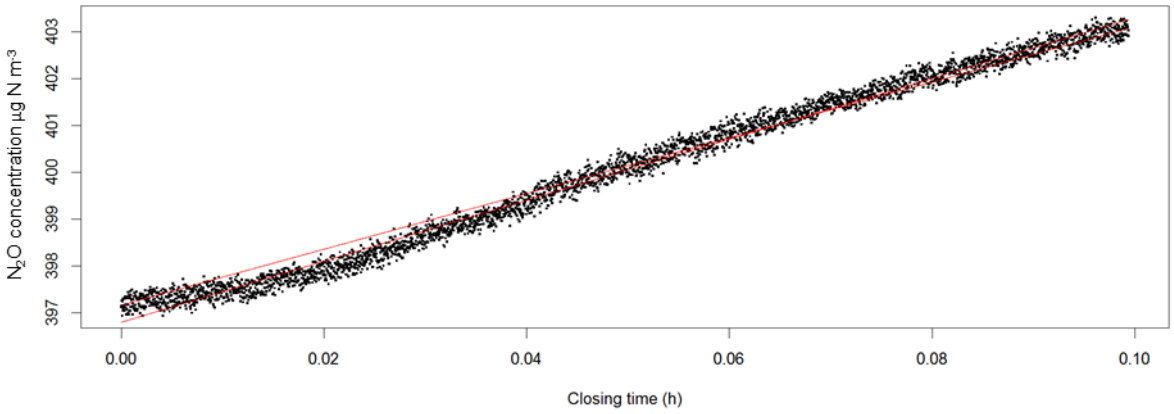


Fig. S1: Example of N_2O concentrations right after chamber closure up to 0.1 h (=6 minutes). Note the small dent at the beginning up to 0.03 h (=108 seconds).

Changes to the manuscript:
Inclusion of Fig. S1 into the supplementary material and reference to the figure on Page 9, Line 1.

[R#2.10] Or "1 to 5 s frequency was sufficient to keep SE on much lower level than in fluxes determined by the GC method" (p. 9). "sufficient" was not defined here.

[AC#2.10] See detailed responses to [AC#2.33] and [AC#2.34].

Changes to the manuscript:
See modified Fig. 5 detailed responses to [AC#2.33] and [AC#2.34].

[R#2.11] How do you justify the limit of using only the first 10 min of the data? Please give some argument based on the data, not just the feeling that this is good.

[AC#2.11] In [AC#2.9] we now point out that the first two minutes of data after chamber closure should be discarded and not used for the regressions. One of the main conclusions of the paper is that applying linear regression to only a short piece of QCL data is fully sufficient to reliably calculate the flux. We show this for periods of three minutes. Chamber operators can decide on their own whether they want to use 3 or 5 or 10 minutes for flux calculation or even extend the initial data that is discarded. Our point is that we clearly found that it does not take a long chamber deployment time to calculate robust fluxes. A period of 10 minutes gives the user enough tolerance for setting its own schedule. We will clarify this in the respective paragraphs of the manuscript.

Changes to the manuscript:
See [AC#2.9] and analyses of fluxes from 3-min linear regressions (Figs. 4, 5, 6, 7). Clarification will be provided in the newly arranged Section 4.2.

[R#2.12] Many of the conclusions of the paper follow those observed in previous studies and are already well known. The one exception is the data in Fig. 8 showing the diurnal variation in N2O flux, as this phenomenon is not much studied. In addition to such data providing information on GHG formation processes, the value of the paper could have been in showing how exactly this chamber system works and what are the special and quantified conditions needed to run the system and to screen the data in order to provide reliable flux data.

[AC#2.12] See responses to [RC#2.4] and [RC#2.5].

Changes to the manuscript:
See responses to [RC#2.4] and [RC#2.5].

[R#2.13] I strongly recommend to separate the results and discussion parts; presently it is difficult to follow the storyline.

[AC#2.13] We agree that splitting results and discussion into two parts may improve the readability of the paper.

Changes to the manuscript:
See [AC#2.1] for the newly structured results and discussion sections.

METHODS

[R#2.14] P.4 L.3 Why "semi-automatic"?

[AC#2.14] The term 'semi-automatic' refers to the operation mode when the system is connected to vial air-sampling, i.e. collecting air in the autosampler. It describes the fact that the sampling is automatic, but the actual gas analysis is done later in the lab. The term is explained in the Introduction, P.2, L.24.

Changes to the manuscript:
None.

[R#2.15] P.4 L.6-7 A volume of L x W x H does not result in 0.33 m3

[AC#2.15] It is true that the interior dimension we have given on Page 4, Line 6 would result in ~0.341 m^{-3}; however, the number 0.33 m^{-3} describes the real conditions as we needed to

subtract volume of some items inside the chamber such as the fan, different bigger screws and supporting racks and tubes. We clarify this in the manuscript.

Changes to the manuscript:
Sentence modified to: 'Subtracting inside items such as an axial fan, screws, supporting racks and tubes, the chambers have a headspace volume of 0.33 m^{-3} and covered a surface area of 0.56 m^{-2}.'

Chapter 2.3:

[R#2.16] p.5 L. 26: what are the conditions when HMR function cannot be fitted? ; L27, what is Akaike information criterion, please open this a bit, although there is the reference, the reader should get some kind of an idea just by reading the text here.

[AC#2.16] For clarification, we slightly rephrase this paragraph and provide some additional information as given below.

Changes to the manuscript:
The paragraph on Page 5, Lines 24 ff. has been modified to: 'Briefly, non-linear flux estimation with the HMR method (R Core Team, 2012; HMR package version 0.3.1) was performed when four data points were available and all of the following criteria were met, i.e. (1) the HMR function could be fitted, (2) Akaike information criterion (AIC; Burnham and Anderson, 2004), which is a measure of (relative) model quality, i.e., gives fit quality penalized by the model's degrees of freedom, and can be used to compare the quality of different model fits to the same dataset, was lower for HMR fit than for linear fit, (3) *p* value of flux calculated using HMR was lower than that from robust linear fit, and (4) the HMR flux was less than four times larger than the robust linear flux. Otherwise, robust linear regression or ordinary linear regression was used when four or three data points were available, respectively.'

[R#2.17] Equations 1-3 and the text related to them: add units.

[AC#2.17] To keep fluent readability, we will add units in a single sentence at the end of Section 2.3.

Changes to the manuscript:
Sentence added: 'Units for concentrations $c(t)$, c_{max}, and c_0 are g m^{-3}, units for k are g m^{-2} s^{-1}, and units for κ are g m^{-3} s^{-2}.'

RESULTS

[R#2.18] P.7 L.7: "low negative k values": care should be taken to express the relations between negative and more negative values. Perhaps more clear to speak about absolute values when comparing these.

[AC#2.18] We agree that this expression may lead to confusion and will refer to absolute values as suggested.

Changes to the manuscript:
Sentence changed to: 'Extremely low absolute κ values between -10^{-4} and -10^{0} – indicating quasi-linearity in $\partial c/\partial t$ – were almost exclusively found under low flux conditions, whereas…'

[R#2.19] P.7. L 10-11 "Near zero fluxes indicate no considerable changes in N2O concentration". Isn't this self-evident without any measurements? Also, what is "considerable change in N2O concentration"? Do you mean significant? If there's no significant increase in concentration, there is no flux, true? Remove or reword the sentence. The whole chapter (Lines 7-14) seem quite self-evident, as the authors hint in the last sentence of the chapter. From Line 15 onwards you say that application of linear model is acceptable in some cases. However, no quantification, i.e. limit below which this is acceptable, is given. I also do not understand how do you draw this conclusion from the results on Lines 7-14.

[AC#2.19] We agree that it would be self-evident when only taking the cited part as given above. In the manuscript, however, the sentence is clearly written in the context of the kappa discussion. Also, if there is no significant (with regard to being lower than the flux detection limit) increase in concentration, i.e. the flux is (close to) zero, which depicts a very important state of the ecosystem and should definitely be taken into account, the corresponding curvature is also marginal. In our opinion, this should at least be stated once. Again, a quantification of kappa hasn't been shown many times before and is used in our paper as a chamber performance criterion (*cf*. Fig. 3C and 3D; [AC#2.4]). We definitely like to stick with the description in Lines 7-14 as it is. Regarding the statement about the acceptance of linear regression at low fluxes, we will include the newly found relationship of Fig. 4D and rephrase the statement accordingly (see below).

Changes to the manuscript:
Sentence starting on Page 7, Line 15 modified to: 'Our results imply that at low to moderately high flux rates <200 µg N m^{-2} h^{-1} (*cf*. Fig. 4D) and/or short chamber closure, the slight non-linearity in concentration change when calculating fluxes is of minor importance and the application of linear models is acceptable, particularly with regard to other commonly observed errors such as those originating from soil disturbance, chamber placement (Christiansen et al., 2011), temperature, pressure and humidity perturbations, etc. (Parkin and Venterea, 2010).'

[R#2.20] P.7. L.23: what is meant with dispersion here?

[AC#2.20] The soil surface basically releases a dispersion plume to the chamber headspace, which eventually is being transported through tubing to the analyzer. If the dispersion of the elevated gas concentration is initially not uniformly mixed with the air inside the tubing, then a lagged concentration increase in the form of exponential analyzer readings (up to a certain point in time) may be observed.

Changes to the manuscript:
Additional information to the sentence starting on Page 7, Line 22: 'These are exponentially increasing N_2O concentrations after chamber closure due to possible dispersion effects leading to biased analyzer readings when the elevated gas concentration is initially not uniformly mixed with the air inside the tubing, placement of…'

[R#2.21] P.7 L. 29-30 "…outside the chamber and inside chamber conditions…" please reword

[AC#2.21] Sentence rephrased.

Changes to the manuscript:
Changed to: 'We also investigated the possible effect of ambient wind speed and direction on concentration build up characteristics (Figure 3C and 3D, respectively) as differences between the turbulence conditions outside the chamber may possibly vary from those conditions inside the chamber under changing wind speed.'

[R#2.22] P.7 L. 30-31 "…coupling of the flux under ambient conditions…" I do not understand this sentence, please reword

[AC#2.22] It means that placing a chamber on soil is a substantial interference with the local wind regime, particularly when wind speed is high and soil pores in the uppermost soil layer may have been ventilated under ambient conditions (i.e. conditions without a chamber) and it thus would take a while until a steady state flux is established. Sentence rephrased.

Changes to the manuscript:
Sentence(s) changed to: 'Theoretically, pores in the uppermost soil layer might be ventilated under high wind speed when no chamber is in place, thus a close coupling of the flux to the atmosphere exists. Consequently, the establishment of a steady state flux may be more postponed under these high wind speed conditions once the chamber is put onto the soil frame.'

[R#2.23] P.7 L.28 → What about the impact of the fan speed on curvature? Soil pores may be ventilated also by the fan (see for example Lai et al. 2012, BG).

[AC#2.23] Lai et al. (2012) found that 13 min of closure were needed before their fluxes (concentration increase) became constant and therefore they extended the deployment period to 30 min. In our study, we found in most cases clear linear or slightly saturating concentration increases right from the beginning. The few cases with 'irregular start patterns' are discussed under [AC#2.9] and in Section 4.1.
Not only fan speed, but also orientation may affect natural efflux from soil. Information on our fan operation is added to Section 2.1.

Changes to the manuscript:
Sentence on Page 4, Line 13 is modified to: 'Chambers were ventilated during measurements using an axial fan, which was mounted to produce a horizontally oriented airflow alongside chamber walls to minimize interference with the natural steady-state soil efflux, but to maximize proper mixing of the chamber headspace as was described in Drösler (2005).' See also [AC#2.9] with regard to the removal of the initial 2-min period.

Chapter 3.2

[R#2.24] Might be good to start with your own results, not with the literature review. For clarity, I strongly recommend separating results and discussion.

[AC#2.24] Results and discussion will be split up.

Changes to the manuscript:
See [AC#2.1] for newly arranged Results and Discussion section.

[R#2.25] P.8 L. 15 onwards: Figure 4B indicates that actually the 3-min/lin method produces higher fluxes than the 60 min/exp method in the lower flux regime (below 200 µg N m-2h-1). When taking into account also the higher fluxes (n=6), the relationship changes so that these 6 data points make a very strong impact, as the authors already discuss. Even though this is the case, the discussion here emphasizes continuously how the linear fluxes are smaller than exponential, although the results shown support this observation only for the few high flux points. This makes me to doubt how one can make generalizations about the validity of these two methods. I am missing discussion which tries to find explanation for the higher

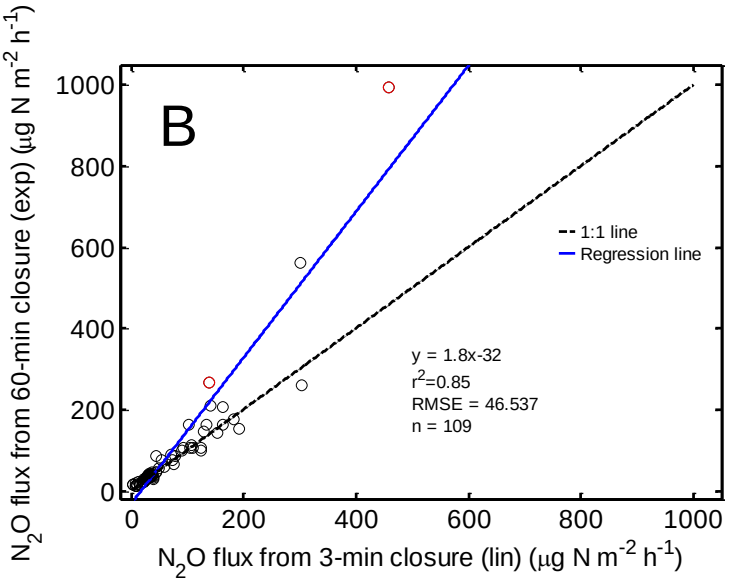
fluxes with 3-min/lin method. Is it so that the data set should be split, or is it far too small to make generalizations?

[AC#2.25] Looking at the newly arranged Fig. 4D (see under [AC#2.4] and [AC#2.44]), there is no indication that 3-min-lin fluxes result in systematically higher values than 60-min-exp fluxes. Only few data points are under the 1:1 line and the slope is 0.989. It is true that we discuss flux underestimation by using linear methods, but that is exactly what is mainly found in the literature (and caused the regression in Fig. 4B). We do not follow the quest of trying to explain why the 3-min-lin method produces higher numbers than the HMR method, because it is just simply not the case, neither in literature nor in the data shown in Fig. 4. We will, however, give a more detailed discussion why HMR-based fluxes are sometimes higher under high flux conditions.

Changes to the manuscript:
Discussion added in Section 4.2. Beside any form of unintended interferences to the 'natural steady-state flux' like for example disturbances through macrofauna, fluctuating pump performance or analyzer malfunctions due to internal re-calibration during chamber deployment, much higher 60-min-based HMR fluxes compared to 3-min-based linear fluxes may be observed when one of two following concentration increase patterns are observed.

1) Slow initial increase of concentrations followed by steeper rise after some minutes. Slope of the linear fit will then be much lower than the one from the HMR fit (lin fit at t_0).
2) Steady linear start of concentration increase followed by sudden relatively sharp bend with lower linear increase afterwards. HMR fit will also have a much steeper slope at t_0 than the linear fit, which will be on top of the data points for the first few minutes.

Red dots in Fig. 4B indicate situations similar to those described under (2) above.



Panel B of Figure 4: Linear regression analysis of N_2O fluxes from the exponential vs. the linear model. Red dots in Fig. 4B indicate situations where a steady linear start of concentration increase was followed by a sudden relatively sharp bend with lower linear increase afterwards.

[R#2.26] Would be also interesting to see, what happens to the SE/RMSE or similar, when apparently low fluxes are calculated with the exponential method. Is there perhaps a risk of

higher error /noisy flux data? Would it be possible to find a flux rate below which the linear method is working more reliably than the exponential?

[AC#2.26] See discussion above. For fluxes <200 µg N m^{-2} h^{-1} there is no significant deviation between the two methods. Even for the few high flux rates, standard errors are still on an acceptable level around 3 µg N m^{-2} h^{-1}.

Changes to the manuscript:
None in particular for this comment, but see also responses [AC#2.4] and [AC#2.25].

[R#2.27] L. 25 I do not understand this sentence

[AC#2.27] What sentence and what is unclear? In the study by Kroon et al. (2008), linear flux rates were underestimated by 60 % compared to those from an exponential function. This was the same order as the flux uncertainty due to temporal variation. Or was the following sentence meant? A simple description of mean and median values of standard errors of fluxes from both the linear and the non-linear model is given.

Changes to the manuscript:
As it is unclear to us what is meant, we stick with the given formulation as we feel the description is very clear.

[R#2.28] P.8 L.31 onwards, continuing on P.9: Here you give important recommendations, but show no data. Also the reference to Section 3.1 is strange, as I do not find anything about the delayed concentration increase in 3.1. This data should be definitely shown if such recommendations are given. You should justify the removal of the first 2 minutes of data: why exactly 2 minutes?

[AC#2.28] See [AC#2.9] for the reason of the removal of the initial 2-min period. The reference to Section 3.1 is indeed strange, because it is a relic from a former version when the kappa analysis looked slightly different than in the submitted version and will be removed.

Changes to the manuscript:
Inclusion of a graph showing delayed concentration increase for the justification to remove the first two minutes of data. Reference to Section 3.1 removed.

[R#2.29] P.8 L3. "…we also compared HMR-based fluxes from QCL? with robust linearly calculated…". How does this vary from that in Fig 7 upper right panel?

[AC#2.29] We assume the reviewer refers to Page 9, Line 3. The difference between the comparison described in Lines 3-10 and Fig. 7 (upper right panel) is that in Fig. 7 linear fluxes are based on 3 minutes of data (as everywhere else in the manuscript), whereas in the described comparison in Lines 3-10 (data not shown), the linear fluxes are – as mentioned in Line 4 – based on the full 60-min cycle of data. To avoid misunderstanding, we add information to the caption of Fig. 7. Please note that in the entire manuscript linear fluxes always refer to 3 minutes of data and HMR fluxes always refer to the full available data set (60 min in Braunschweig and in Risø from DOY 105.5 to DOY 108.5 and 10 min in Risø before DOY 105.5 and after DOY 108.5). This information is explicitly mentioned on Page 5, Lines 30-33.

Changes to the manuscript:

Modified Fig. 7 caption (other changes to Figure 7 are given in [AC#2.46]: 'Panels A and B: GC vs. QCL-based N_2O fluxes. Panels C and D: Relationships between standard errors (SE) of N_2O fluxes and the respective flux values. Blue markers indicate QCL data, which are all based on the 3-min linear calculation method. Red markers indicate GC data, which are based on the full 60-min data set. Crosses are plotted for GC data when all criteria for flux calculation using the exponential HMR model were met (see text for details), otherwise circles are plotted indicating the usage of a linear model for flux calculation.'

[R#2.30] A general comment/hint: there are many different comparisons with different analyzers, calculation methods and closure times, in which partly different data sets have been used (low and/or high flux) and it is not easy to follow how do all these small experiments differ from each other or support each other. A separate result section with subsections dedicated to each of these questions might help in that. Now there is lot of text (e.g. Chapter 3.2) and it is difficult to follow the argumentation on logics of the text. Also a clearer division into paragraphs would help the reader. And, as already pointed out, division into results and discussion is needed.

[AC#2.30] We fully agree and split Section 3 into two parts as outlined under [AC#2.1].

Changes to the manuscript:
Results and discussion will be split up.

[R#2.31] P.9 L. 3-10: In which figure are these shown? Are the slope of 0.97 (lin fluxes are independent) and the HMR fluxes being 22% higher in conflict with each other? How is it possible that now the linear and HMR based fluxes estimated from 60-min data are almost identical (slope=0.97), while earlier you have stated that linear method underestimates the fluxes?

[AC#2.31] See [AC#2.29]. The analysis described in Lines 3-10 is not shown in a figure as mentioned in Line 4. The whole idea of this paragraph is to make these values, i.e. linear fluxes from 60-min closure comparable to other results presented in literature as for example Kroon et al. (2008) or Forbrich et al. (2010), which have been included in the discussion on Page 8, Line 19 ff.

Changes to the manuscript:
See [AC#2.29].

[R#2.32] P.9 L 11-18: How were the standard errors calculated?

[AC#2.32] In the whole manuscript, we deal with the standard error of the flux, not of individual concentrations. As the flux is a parameter in Equation (1), SE is the standard error of the parameter in the respective regression model (not of the residuals of the concentrations). The regression algorithm used is based on the Levenberg-Marquardt method and was taken from the R package 'minpack.lm' (https://cran.r-project.org/web/packages/minpack.lm/minpack.lm.pdf), function 'nlsLM'. The parameter errors are provided by the algorithm. Further details can be found in:

- https://www.rdocumentation.org/packages/minpack.lm/versions/1.2-0/topics/nlsLM
- Equation 22 in http://people.duke.edu/~hpgavin/ce281/lm.pdf
- Bates, D.M. and Watts, D.G.: Nonlinear Regression Analysis and Its Applications, Wiley, 1988.
- Moré, J.J.: The Levenberg-Marquardt algorithm: implementation and theory, in Lecture Notes in Mathematics 630: Numerical Analysis, G.A. Watson (Ed.), Springer: Berlin, 1978, pp.105-116.

Changes to the manuscript:
Additional information added at the end of Section 2.3: ‘Standard errors in this study were calculated as the parameter errors from the respective regression model with the algorithm being based on the Levenberg-Marquardt method (‘nlsLM function in R package ‘minpack.lm’, R Core Team, 2012).’

[R#2.33] This section and Figure 5: I think that SE is not an appropriate quantity when estimating the “sufficient” frequency of concentration data. By definition, SE is related to (the root square of) the number of observations. It is therefore evident that if you decrease the frequency and the number of data, you increase the SE. In a case where the random error of the concentration measurement during the chamber closure is constant, the SE will anyway increase in case the number of observations decreases, whereas the NRMSE, or the error in the flux will not increase. Therefore a better quantity to estimate the error related to the frequency of concentration data is RMSE (or NRMSE).

[AC#2.33] We sincerely thank the reviewer for this catch and fully agree that reducing sample size automatically leads to an increase in the error estimate when using the method given in [AC#2.32]. However, instead of taking RMSE – which wouldn’t work, because we deal with the SE of a parameter and not of residuals – we normalized the SE by multiplication with $\sqrt{n}$. This is now shown in the newly arranged Fig. 5 (see below). The result is fascinating: mean and median for both 3-min-lin and 60-min-exp fluxes are basically invariant with changing sampling time even up to a frequency of 0.03. Note that only 6 data points are left in that latter frequency class for the 3-min-lin fluxes. We rephrase our conclusions accordingly.

Changes to the manuscript:
Inclusion of newly arranged Fig. 5 and reformulation of the paragraph on Page 9, Lines 11-19: ‘A further intriguing analysis shows that standard errors were found to be invariant on QCL sampling frequency (Figure 5). We simulated different sampling times ranging from one tenth of a second to 25.6 sec, which corresponds to a frequency of 0.0390625 Hz, by excluding the respective intervals from the original 10-Hz dataset. Results show that the median of the standard error of the fluxes remains stable over a wide range of measurement frequencies. At a frequency class of 0.15 and lower (3 boxes on the right-hand side of Fig.5), which corresponds to a sampling time of ~5 sec and higher, lower and upper quartile values begin to deviate and the median changes slightly.’ (can now be found at the end of Section 3.2). Further, we add at the end of Section 4.2: ‘The conclusion we can draw from this finding is that chamber operators – in case an analyzer with a precision like the QCL presented in this study is available – can reduce their sampling time down to 5 seconds without risking an increase of the standard error of the flux, which would still be on a much lower level than those obtained from GC measurements.’

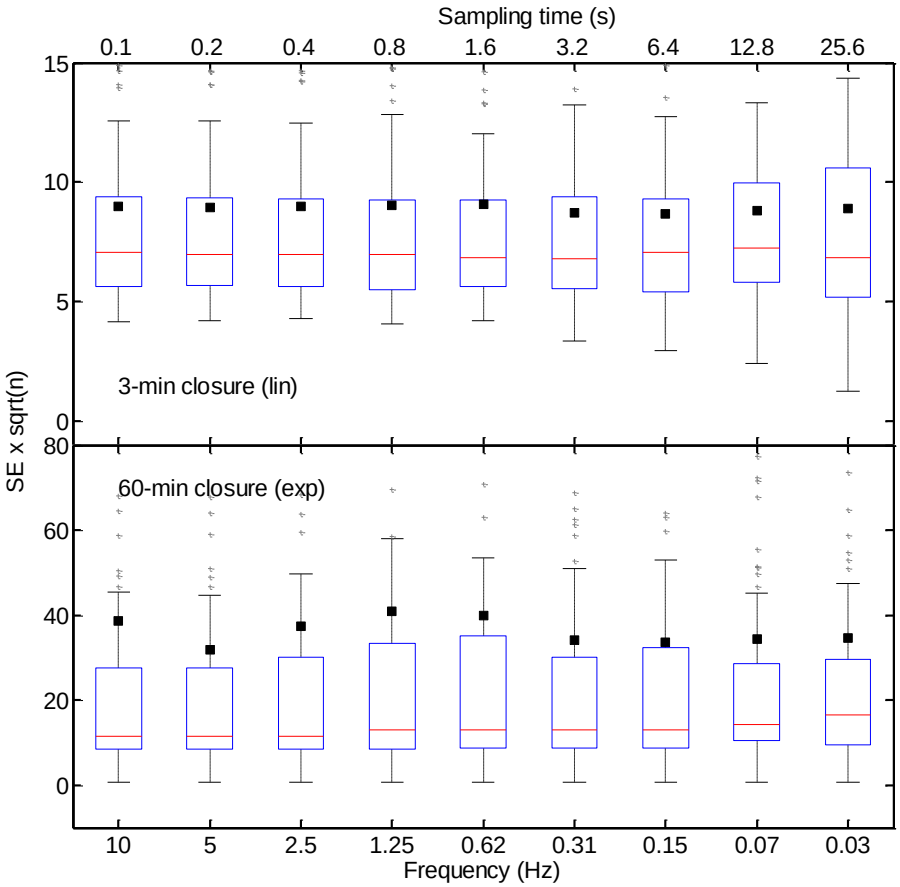


Figure 5. Boxplots of standard errors of N_2O fluxes for different frequency classes and regression models used, i.e. linear regression with 3 minutes of data (upper panel) and the exponential HMR model with 60 minutes of data (lower panel). To avoid a pseudo-dependency on sample size, the standard errors were normalized by multiplication with $\sqrt{n}$. Black squares represent the arithmetic mean, red horizontal lines indicate the median, blue horizontal lines indicate lower and upper quartile values, black whiskers represent the interquartile range and outliers from this range are plotted as grey crosses.

[R#2.34] Your argument "..sampling times between 1 and 5 sec are sufficient to keep SE of fluxes on a much lower level…" is vague. How do you justify that exactly the 1-5 sec limit is sufficient? What means sufficient? How much is "much lower level"? Please quantify and justify this with an appropriate and objective criteria.

[AC#2.34] See [AC#2.33]. The threshold will be set to 5 seconds.

Changes to the manuscript:
See [AC#2.33]. Rephrasing of occurrences where the threshold is given; now set to 5 seconds.

[R#2.35] P.9 L.12: "..to approx.. one minute,…" isn't it approx. half a minute (25.6 sec)?

[AC#2.35] Correct, but the paragraph has been rephrased anyway as mentioned under [AC#2.33].

Changes to the manuscript:
Rephrased as mentioned under [AC#2.33].

Chapter 3.3

[R#2.36] P.9 L. 23 To be exact, QCL fluxes are not explained by GC fluxes. They are correlated with GC fluxes.

[AC#2.36] We agree. The expression would fit better if we would look at a controlling factor (x-axis) of some dependent variable (y-axis). Sentence rephrased.

Changes to the manuscript:
Sentence rephrased to: 'A linear regression revealed no significant relationship between GC and QCL fluxes with a very low coefficient of determination of 0.036 (Figure 7A).' We also rephrased the sentence on Page 9, Lines 31-32 '48 % of the variance in QCL-based fluxes could be explained by fluxes from the GC method.'. It now reads 'A linear regression between GC and QCL fluxes revealed a coefficient of determination of 0.48 (Figure 7B).' as it deals with the same topic as the one above.

[R#2.37] P.10 L4 What does mean "…no dependency on flux value was observed…" Why should SE depend on flux value? Again, how was SE defined?

[AC#2.37] The standard error of a flux may depend on the flux itself for example when at very low fluxes (low concentration increases) the slope fit may be prone to much higher uncertainty than at larger fluxes when an analyzer with moderate or low precision is used. On the other hand high fluxes may show high standard errors for example when an analyzer is not well calibrated or not able to properly resolve certain concentration ranges. This is something we needed to investigate and a dependency that might explain faulty QCL calibration could not be found. See [AC#2.32] for SE calculation method.

Changes to the manuscript:
None.

[R#2.38] P.10 L 14 indicates → indicating

[AC#2.38] No, we don't change that. The word 'indicates' refers to 'The fact that…'. We think changing this to 'indicating' would lead to incorrect grammar (and/or different meaning).

Changes to the manuscript:
None.

[R#2.39] P.10 L 15 forward: "…GC is still useful method to determine soil-atmosphere exchange… at longer time scales.." What is your argument based on? There are no budget calculations in the paper. Averages were reported to be similar, particularly for the small flux regime, but at the same time the fluxes were hardly detected with the GC. Is it correct to say that GC fits for budget studies? How big errors are acceptable in budget studies?

[AC#2.39] We agree that 'at longer time scales' is a bit strong given the fact that we only show a few weeks of data. Nevertheless we need to point out that mean and median of the whole BS campaign where fluxes were on average quite low match pretty well. In Risø – although single flux values were closer to each other – deviation between GC and QCL mainly occurred at high fluxes (Fig. 7B) under the influence of fertilization. But this also

indicates that using a GC is still useful for a wide range of periods over an entire year. However, taking into account that the bulk of the annual efflux occurs after management events at a relatively short time scale, usage of a GC-based system will be prone to large uncertainties. Paragraph on Page 10, Lines 15-21 will be adjusted.

Changes to the manuscript:
Page 10, Lines 15-21 adjusted to: 'In summary, our comparison of GC vs. QCL fluxes revealed that despite much higher precision, robustness, and temporal resolution in QCL measurements, GC is still a useful method to determine the average campaign N_2O soil efflux. Although single flux values particularly under low exchange regimes did not match well, campaign means and medians were similar to those obtained by the QCL method. Under high exchange regimes, however, flux patterns matched considerably better, but resulted in larger absolute errors when comparing the campaign average, thereby leading to systematic errors (in our case an underestimation) when using the GC method at high N_2O fluxes for the assessment of N balances. However, given the fact that the bulk of the annual efflux occurs after management events at a relatively short time scale (Flechard et al., 2007; Skiba et al., 2013), usage of a GC-based system will be prone to large uncertainties (cf. Fig.7).'

[R#2.40] FIG 2 add A) and B) to panels and refer to them in the legend

[AC#2.40] Labels were added and were referred to in the text.

Changes to the manuscript:
Modified Figure 2 with modified caption:

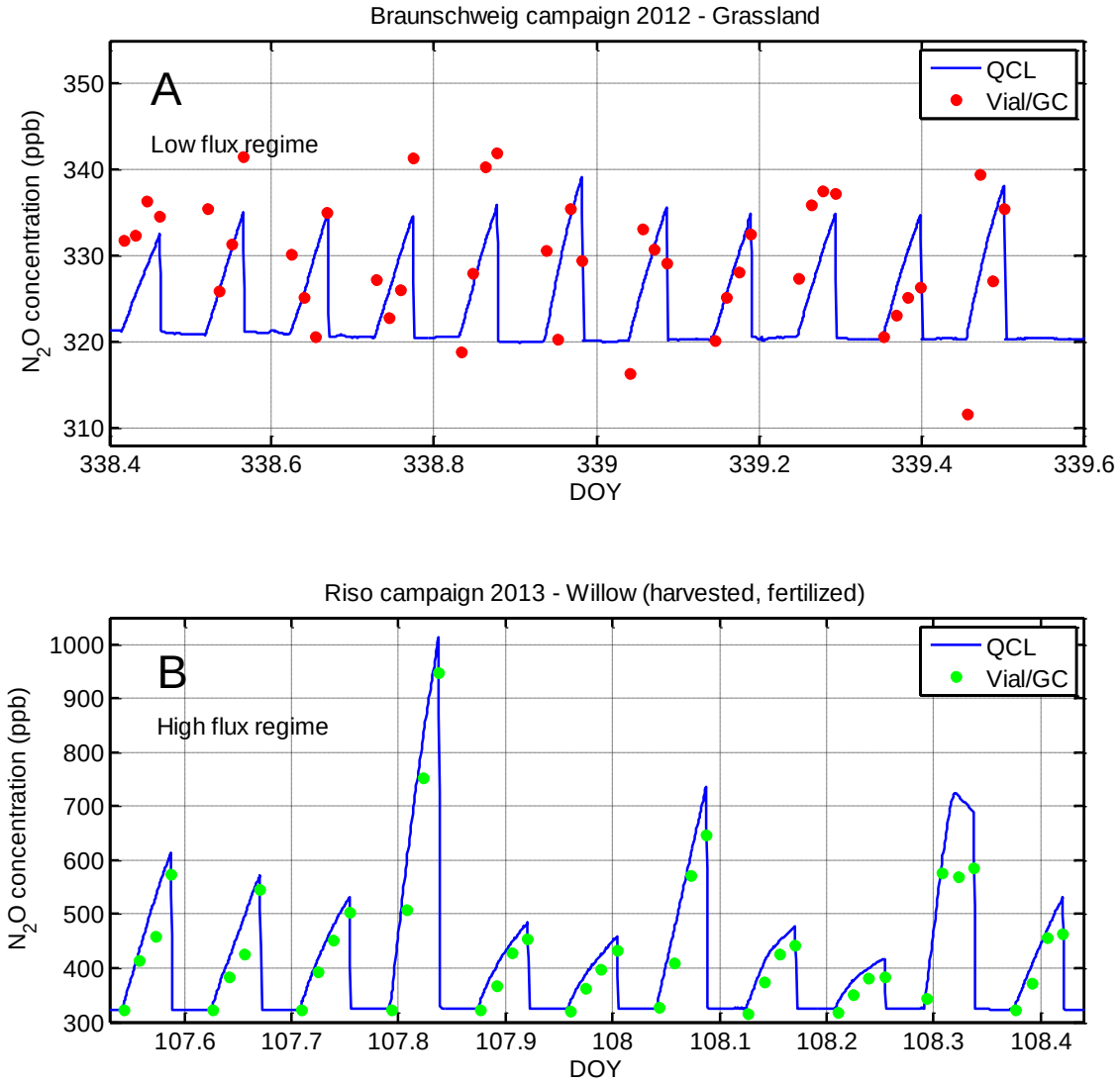


Figure 2. Examples of time series of N_2O chamber concentrations during the Braunschweig (Panel A) and Risø campaign (Panel B). Chambers were periodically closed for 60 minutes. Vials were filled with sample air at t_0, t_{20}, t_{40}, and t_{60}. The QCL system was operated at a sampling frequency of 10 Hz; plotted are 1-min means.

FIG 4

[R#2.41] - Refer to "A, B, C and D" before each legend text parts; "Figure 4. a) Comparison of N2O fluxes… b) Linear regression…"

[AC#2.41] Modified as suggested. Please also notice that Panel D has been changed for a better visualization of lower fluxes and its regression as a result to comment [R#2.26].

Changes to the manuscript:
Modified Figure 4 caption reads:

Figure 4. Panel A: Comparison of N_2O fluxes measured on a harvested willow field during the Risø campaign by the QCL system based on a linear model using only the first three minutes of data after chamber closure (filled blue circles) and an exponential model (open red circles) (see text) using either the full 60 minutes (DOY 105.5 to DOY 108.5) or the full 10 minutes of data (DOY <105.5 and DOY >108.5). Panel B: Linear regression analysis of N_2O fluxes from

the exponential vs. the linear model. Panel C: Standard errors of fluxes shown in Panel A. Panel D: Same as Panel B, but only for fluxes <200 μg N m^{-2} h^{-1} with adapted regression.

[R#2.42] - The legend text should be shortened. Remove phrases such as "Also shown is…" Figure 4b is showing 60-min fluxes plotted against 3 min fluxes.

[AC#2.42] Modified as suggested.

Changes to the manuscript:
See [AC#2.41].

[R#2.43] - Please remove the text "Riso campaign 2013, Willow…" from the top of each separate panel and add that part of information into the legend text which is not already there.

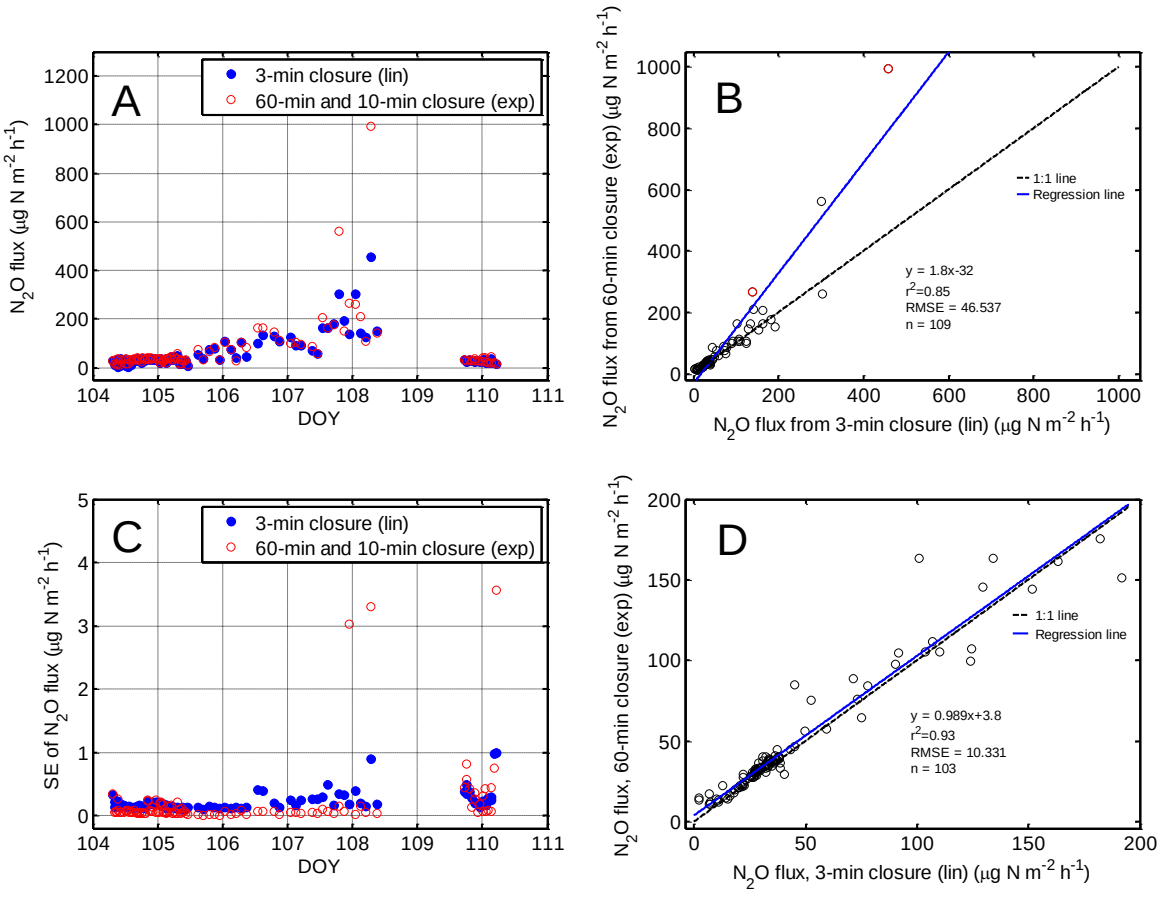
[AC#2.43] Modified as suggested.

Changes to the manuscript:
See [AC#2.41].

[R#2.44] - Panel B: indicate what are the two lines in the figure? Why are they not direct lines, but show some tiny variation?

[AC#2.44] The dashed black line is the 1:1 line and the blue solid line is the linear regression fit line. Line labels have been added to Panel B and D. They probably didn't appear as straight lines in the former version, because of the graphical resolution of the figure. This has now been improved.

Changes to the manuscript:
Modified Figure 4 (for caption see [AC#2.41]):



[R#2.45] - In Fig. 4 and Fig 5, what is the reason to compare 3-min linear and 60-min exponential fluxes? Why not to compare separately the lin vs exp AND 3-min vs 60 min closure times?

[AC#2.45] In Figures 4 and 5 we explicitly deal with high-resolution measurements of the QCL, which gives us the opportunity to use robust and precise data to compare the application of a linear model simulating short closure time (here 3 minutes) with a non-linear model representing long closure time (here 60 minutes). One of the main aims of the paper is to investigate whether it is suitable to reduce chamber closure time and to apply simple linear regression to calculate the N_2O flux. Through Figure 4 (particularly Panels A and D) and Figure R1 (see above), we demonstrate that the bulk of the flux differences between linear and non-linear models is in an acceptable range, keeping in mind that shorter closure times also have the advantage that plants and soil in the measurement plots are less affected in the long term. Hence, comparing fluxes from a linear model using 3 minutes of data with fluxes from a non-linear model using 60 minutes of data clearly supports our specific aim of the study, while for example applying a linear model to a 60-min data set that reveals obvious curvature (see κ values in Figure 3) or applying a non-linear model to only 3 minutes of data that are quasi linear would not give any further insights when investigating whether reducing chamber closure is acceptable or not.

Changes to the manuscript:
None.

FIG 7
[R#2.46] - Please use A-D notations, not left/right/upper/lower explanations

[AC#2.46] We added Labels A to D and included them in the figure caption to keep it consistent with Figure 4.

Changes to the manuscript:
Modified Figure 7 with modified caption:

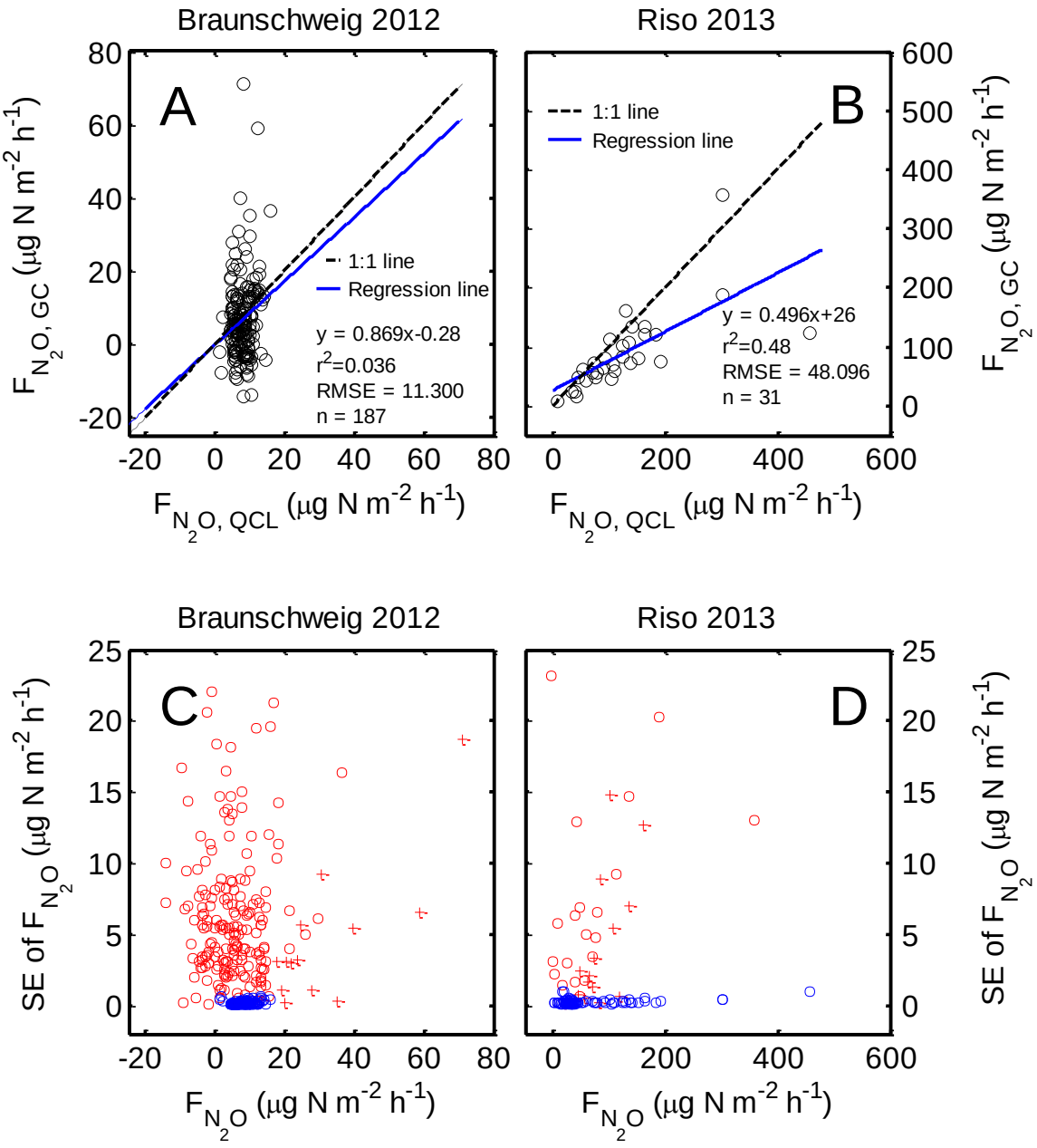


Figure 7. Panels A and B: GC vs. QCL-based N_2O fluxes. Panels C and D: Relationships between standard errors (SE) of N_2O fluxes and the respective flux values. Blue markers indicate QCL data, which are all based on the 3-min linear calculation method. Red markers indicate GC data, which are based on the full 60-min data set. Crosses are plotted for GC data when all criteria for flux calculation using the exponential HMR model were met (see text for details), otherwise circles are plotted indicating the usage of a linear model for flux calculation.

[R#2.47] - Upper panel: define the lines

[AC#2.47] Line labels added.

Changes to the manuscript:
See [AC#2.46].

FIG 8
[R#2.48] - An interesting Figure. What does the error bar denote? Is the diurnal variation significant? Why is the hourly data not shown? Are the points averages from many hours? What was actually the frequency of measurements in both campaigns, I did not find it, but I assumed you measured hourly?

[AC#2.48] The error bar indicates the standard error of the mean from all flux values in each bin. Each bin contains fluxes from 3-hour periods, i.e. from 00:00 to 03:00, 03:00 to 06:00, 06:00 to 09:00 and so on. The mean values in Figure 8 are plotted in the center of each bin. Fluxes were binned due to irregular starting times of new chamber cycles. In general, a new chamber cycle could be started each full hour, but to get a more robust diurnal pattern, we decided to bin data in the above-mentioned 3-hour containers. While the diurnal variation of N_2O fluxes from the Risø campaign is significant (*p*-value = 0.0059), the diurnal variation found during the Braunschweig campaign is not as the difference between mean minimum and maximum values is lower than the upper flux detection limit of ~2.6 μg N m^{-2} s^{-1} (*cf*. response to SC3).

Changes to the manuscript:
Information on data handling and significance of the diurnal variation is added to Section 3.4.

References cited in this document:

Bates, D.M. and Watts, D.G.: Nonlinear Regression Analysis and Its Applications, Wiley, 1988.

Drösler, M., 2005. Trace gas exchange and climatic relevance of bog ecosystems, Southern Germany. Doctoral thesis, TU München, 1-182 pp.

Flechard, C.R., Ambus, P., Skiba, U., Rees, R.M., Hensen, A., van Amstel, A., van den Pol-van Dasselaar, A., Soussana, J.-F., Jones, M., Clifton-Brown, J., Raschi, A., Horvath, L., Neftel, A., Jocher, M., Ammann, C., Leifeld, J., Fuhrer, J., Calanca, P., Thalman, E., Pilegaard, K., Di Marco, C., Campbell, C., Nemitz, E., Hargreaves, K.J., Levy, P.E., Ball, B.C., Jones, S.K., van de Bulk, W.C.M., Groot, T., Blom, M., Domingues, R., Kasper, G., Allard, V., Ceschia, E., Cellier, P., Laville, P., Henault, C., Bizouard, F., Abdalla, M., Williams, M., Baronti, S., Berretti, F., Grosz, B.: Effects of climate and management intensity on nitrous oxide emissions in grassland systems across Europe, Agric. Ecosyst. Environ., 121, 135–152, 2007.

Lai, D.Y.F., Roulet, N.T., Humphreys, E.R., Moore, T.R., and Dalva, M.: The effect of atmospheric turbulence and chamber deployment period on autochamber CO_2 and CH_4 flux measurements in an ombrotrophic peatland, Biogeosciences, 9, 3305–3322, doi:10.5194/bg-9-3305-2012, 2012

Moré, J.J.: The Levenberg-Marquardt algorithm: implementation and theory, in Lecture Notes in Mathematics 630: Numerical Analysis, G.A. Watson (Ed.), Springer: Berlin, 1978, pp.105-116.

Parkin, T.B., Venterea, R. T., and S. K. Hargreaves: Calculating the detection limits of chamber-based soil greenhouse gas flux measurements, J. Environ. Qual., 41, 705–715, doi:10.2134/jeq2011.0394, 2012.

Skiba, U., Jones, S.K., Drewer, J., Helfter, C., Anderson, M., Dinsmore, K., McKenzie, R., Nemitz, E., Sutton, M.A.: Comparison of soil greenhouse gas fluxes from extensive and intensive grazing in a temperate maritime climate, Biogeosciences, 10, 1231–1241, doi:10.5194/bg-10-1231-2013, 2013.

Gas chromatography vs. quantum cascade laser-based N_2O flux measurements using a novel chamber design

Christian Brümmer[1,*], Bjarne Lyshede[1], Dirk Lempio[1], Jean-Pierre Delorme[1], Jeremy J. ~~Smith[1]~~Rüffer[1,*], Roland Fuß[1], Antje M. Moffat[1], Miriam Hurkuck[1], Andreas Ibrom[2], Per Ambus[3], Heinz Flessa[1], Werner L. Kutsch[1,4]

[1]Thünen Institute of Climate-Smart Agriculture, Braunschweig, Germany
[2]Department of Environmental Engineering, Technical University of Denmark, Lyngby, Denmark
[3]Department of Geosciences and Natural Resource Management, University of Copenhagen, Denmark
[4]Integrated Carbon Observation System ~~(ICOS)~~, ICOS ERIC Head office, Helsinki, Finland
[*]previously published under the name: Jeremy Smith

Correspondence to: Christian Brümmer (christian.bruemmer@thuenen.de)

Abstract. Recent advances in laser spectrometry offer new opportunities to investigate soil-atmosphere exchange of nitrous oxide. During two field campaigns conducted at a grassland site and a willow field, we tested the performance of a quantum cascade laser (QCL) connected to a newly developed automated chamber system against a conventional gas chromatography (GC) approach using the same chambers plus an automated gas sampling unit with septum capped vials and subsequent laboratory GC analysis. Through its high precision and time resolution, data of the QCL system were used for quantifying the commonly observed non-linearity in concentration changes during chamber deployment, making the calculation of exchange fluxes more accurate by the application of exponential models. As expected, the curvature in the concentration increase was higher during long (60 min) chamber closure times and under high flux conditions (F_{N2O}>150 µg N m^{-2} h^{-1}) than those that were found when chambers were closed for only 10 min and/or when fluxes were in a typical range of 2 to 50 µg N m^{-2} h^{-1}. Extremely low standard errors of fluxes, i.e. from ~0.~~1~~ 2 to 1.7 % of the flux value, were observed regardless of linear or exponential flux calculation when using QCL data. Thus, we recommend reducing chamber closure times to a maximum of 10 min when a fast-response analyzer~~high-frequency measurements are~~ is available to keep soil disturbance low and conditions around the chamber plot as natural as possible. ~~If instrumentation with high-frequency resolution is available~~Further, a sampling time~~s from 1 to~~of every 5 s proved to be sufficient for robust flux determination assuring standard errors of N_2O fluxes still being on a relatively low level. Despite low signal to noise ratios, GC was still found to be a useful method to determine mean soil-atmosphere exchange of N_2O at longer time scales during specific campaigns~~, i.e. seasons to years, e.g., when the main focus of the study is to investigate site budgets~~. Intriguingly, the consistency between GC and QCL-based campaign averages was better under low than under high N_2O efflux conditions, although single flux values were highly scattered during the low efflux campaign. Furthermore, the QCL technology provides a useful tool to accurately investigate the highly debated topic of diurnal courses of N_2O fluxes and its controlling factors. Our new chamber design prevents the measurement spot from unintended shading and minimizes ~~reduces the~~ disturbance of throughfall,

Formatiert: Hochgestellt

thereby ~~of the soil and complies~~ complying with high quality requirements of long-term observation studies and research infrastructures.

1 Introduction

Accurate determination of ambient nitrous oxide (N_2O) concentrations and the associated exchange between soil and atmosphere has been in the focus of environmental research for several years. Nitrous oxide is of high relevance for the Earth's greenhouse gas budget due to its long residence time in the troposphere and its relatively large energy absorption capacity per molecule, resulting in a cumulative radiative forcing almost 300 times higher than the same mass unit of carbon dioxide over a 100-year period when climate-carbon feedbacks are included (IPCC, 2013). It is predominantly emitted as a byproduct of nitrification and an intermediate product of denitrification and nitrifier denitrification, which are key microbiological processes in the soil nitrogen (N) cycle (Firestone and Davidson, 1989; Wrage et al., 2001; Thomson et al., 2012; Butterbach-Bahl et al., 2013). Main N_2O sources are agricultural activities in the form of N fertilization. In smaller quantities, N_2O is also produced through biomass burning, degassing of irrigation water, and industrial processes (Seinfeld and Pandis, 2006). On the other hand, some field studies report that soils can also consume N_2O, albeit the strength of this sink has not yet been thoroughly evaluated (Donoso et al., 1993; IPCC 2007; Chapuis-Lardy et al., 2007).

Precise measurements of N_2O – particularly at the field scale – are therefore essential for specific applications in ecosystem research such as the study of N cycling, fertilization effects, and for the compilation of full greenhouse gas budgets. The most common method to measure soil-atmosphere exchange of N_2O is the operation of static chambers (Hutchinson and Mosier, 1981; Schiller and Hastie, 1996). The N_2O flux is calculated from the concentration increase (or decrease) over time in a gas-tight chamber, which is usually attached to a collar that is permanently inserted into the soil. A number of approaches have emerged over the last years where the air sample is either manually collected using a syringe through a septum and/or directly inserted into sample vials (e.g., Castaldi et al., 2010; Jassal et al., 2008, 2011; Livesley et al., 2011; Lohila et al., 2010; Parkin and Venterea, 2010 and references therein) with subsequent analysis on gas chromatography (GC) systems using ^{63}Ni electron capture detectors for N_2O detection. Different chamber designs and air sampling procedures exist, either with manual, semi-automated, i.e. automatic sampling but manual transport of air samples in syringes or vials to the GC (this study), or fully automated gas collection, where the air samples are directly pumped (or sucked) via carrier gas to a temperature-stable housing equipped with a GC in the field (e.g., Brümmer et al., 2008; Butterbach-Bahl et al., 1997; Dannenmann et al., 2006; Flessa et al., 2002; Papen and Butterbach-Bahl, 1999; Rosenkranz et al., 2006).

In the last decade, substantial progress has been made in the development of fast-response technologies for analyzing a variety of N and carbon (C) trace gases. These are tunable diode laser absorption spectrometers (TDLAS), quantum cascade lasers (QCL) and devices originating from individual applications such as Fourier transform infrared (FTIR) spectrometers or custom-made converters coupled to chemiluminescence detectors (CLD). These robust, fast and precise analyzers are essential for long-term monitoring of biosphere-atmosphere exchange and have even allowed first eddy covariance (EC)

measurements of field-scale N_2O, methane (CH_4) (e.g., Rinne et al., 2005; Denmead et al., 2010; Kroon et al., 2010; Neftel et al., 2010; Tuzson et al., 2010; Jones et al., 2011; Merbold et al., 2014), and reactive N fluxes (Horii et al., 2004; Ammann et al., 2012; Brümmer et al., 2013). Continuous observations of trace gas exchange over time scales from hours to decades enable researchers to evaluate diurnal, seasonal and interannual variability and trends as well as the elucidation of climatic and management controls on gas exchange patterns (e.g., Baldocchi et al., 2001; Brümmer et al., 2012; Kutsch et al., 2010). With regard to chamber measurements, it is expected that the precision and time resolution of the above-mentioned technologies may considerably reduce the chamber closure duration for single flux measurement events, thereby minimizing plot disturbance and allowing for a significant increase in repeated measurements leading to more robust databases, which are required for reliable greenhouse gas budgets. Although the EC methodology provides near-continuous time series of greenhouse gas concentrations and exchange, chamber measurements will certainly still be required in the future as prerequisites for EC measurements are sometimes not fulfilled (for example through insufficient turbulent mixing, complex terrain, inhomogeneous fetch) and small-scale spatial variability or emissions from replicated field plot experiments can only be determined by chamber measurements. Some first examples of high-resolution chamber measurements using fast-response analyzers can be found in Cowan et al. (2014a; 2014b), Hensen et al. (2006), Laville et al. (2011), Sakabe et al. (2015), and Savage et al. (2014).

The comparability, applicability and uncertainty associated with the respective approach are currently debated in the ecosystem research community, e.g. when comparing fluxes from GC/vial systems with those from more recent continuous setups such as QCL systems. In this context, the flux determination method was found to be an important factor (e.g., Kroon et al., 2008; Forbrich et al., 2010). Fluxes are often calculated using a linear regression of the change in headspace concentration over time and are scaled to the collar area, including a temperature and pressure correction (e.g. Savage et al., 2014). However, several other studies demonstrate the need for non-linear models for soil-atmosphere trace gas flux estimation (Hutchinson and Mosier, 1981; Livingston et al., 2006; Kutzbach et al., 2007; Kroon et al., 2008; Pedersen et al., 2010; Pihlatie et al., 2013). It has been argued that molecular diffusion theory states that chamber effects lead to declining gradients in the relationship between concentration and time and that slight chamber leakages create the same effect (Hutchinson and Mosier, 1981; Livingston et al., 2006; Pedersen et al., 2010). Nevertheless, linear concentration data often predominates (e.g. Forbrich et al., 2010), which may not necessarily be in conflict with the theory as non-linearity is sometimes not visible in data series with only a limited number of samples (mostly due to noisy concentration measurements or effects of small chambers; Pedersen et al., 2010).

To further investigate effects of flux estimation methods on the one hand and the use of different gas analyzer types on the other hand, our study comprises N_2O chamber flux measurements from two campaigns conducted with a newly developed chamber system under different environmental conditions. The aims of this study were as follows: ~~with the aim to (1) present a novel chamber design used in both an automated vial air-sampling setup followed by gas chromatography analysis in the laboratory and in combination with a directly connected QCL system, (2) test different flux calculation methods, i.e. conventional linear regression models vs. a slightly modified HMR procedure (Hutchinson-Mosier Regression, *cf.* Pedersen~~

et al., 2010), including different chamber closure durations, and sampling frequencies to find an optimal setup particularly for the QCL system, and to (3) compare single and mean fluxes of the field campaigns to help recommending an appropriate setup for common field applications.

- Presentation of a novel chamber design that is connected to both a vial air-sampling setup with subsequent GC analysis and a QCL spectrometer. Description of design and setup can be found in Sections 2.1 and 2.2. The new chamber system was used for the following investigations:
- Characterization of the shape of the concentration increase/decrease to identify whether $\partial c/\partial t$ is rather linear or non-linear including a quantification of the curvature (κ) in concentration increase/decrease (Section 3.1). The parameter κ was further used to verify chamber density by checking its dependency on wind speed, wind direction, on the flux itself and on closure time.
- Comparison of N_2O fluxes and their associated standard errors from linear and non-linear regression models (Section 3.2).
- Testing the novel chamber system under high and low flux conditions and comparing GC vs. QCL-based flux estimates (Section 3.3).
- Investigation of ecosystem and climate-specific flux characteristics such as N_2O uptake and diurnal variation (Section 3.4).

2 Methods

2.1 Chamber design

Nitrous oxide measurements were carried out using a newly developed semi-automatic chamber system (Figure 1). It consisted of aluminum guiding racks (length 2121 mm, width 936 mm, height 3033 mm) with aluminum soil collars (length 750 mm, width 750 mm, height 160 mm, inserted 0.10 m into the soil), and opaque PVC chambers (color white, interior dimensions: length 777 mm, width 777 mm, height 565 mm) (Ps-plastic, Eching, Germany). Subtracting inside items such as an axial fan, screws, supporting racks and tubes, the ~~The~~ chambers have a headspace volume of 0.33 m^3 and covered a surface area of 0.56 m^2. Depending on vegetation height, extension modules (interior dimensions: length 730 mm, width 730 mm, height 360 mm) can be connected to the chambers (total headspace volume was then 0.54 m^3) if needed over taller vegetation, but were not used in this study. EPDM-gaskets (20 mm x 15 mm) were attached to the bottom of each chamber in an aluminum u-channel to ensure gas tight closure when chambers were operating. Up to three chambers can be combined to one system (Figure 1B) with a joint control unit and autosampler or analyzer. Two custom made temperature probes (Pt100) were installed inside and outside of each chamber to measure ambient air temperatures. Chambers were ventilated

during measurements using an axial fan~~.~~, which was mounted to produce a horizontally oriented airflow alongside chamber walls to minimize interference with the natural steady-state soil efflux, but to maximize proper mixing of the chamber headspace as was described in Drösler (2005). The air was sampled from the top centre of the lids. Chamber operation was controlled by a logic module (Millenium 3, Crouzet, Hilden, Germany). An autosampler consisting of a membrane pump (operated at 0.8 L min^{-1}, NMP 830 KNDC, KNF Neuberger, Freiburg, Germany), an absorber to avoid water condensation within tubes (3.2 mm ID, 6.4 mm OD) (BEV-A-Line, ProLiquid GmbH, Überlingen, Germany) and valves as well as an exchangeable rack for 162 headspace vials (20 mL, WICOM WIC 43200, Maienfeld, Germany) were connected to the chamber system. Chambers were lifted and moved down by a 24 V (DC) motor winch and were directed to the soil collar by the aluminum rack. After measurement events, the chamber was lifted to 1.18 m above ground and dragged backwards in a 45° angle to keep the soil and vegetation inside the soil collar under as natural conditions as possible (e.g. prevention of shading and undisturbed throughfall). To avoid pressure changes when setting the chamber on the collar, the chamber had a 1.5 m pressure compensation tube leading from the inside through the side wall of the chamber to the outside. Information about our chamber system including the construction plan is open to the scientific community and can be requested from the authors.

2.2 Campaigns and measurement setup

Two field campaigns were conducted in fall 2012 in Braunschweig, Germany, and in spring 2013, at Risø Campus, Technical University of Denmark, using both GC and QCL chamber setups (see Table 1 for additional information). The chamber architecture was identical during the two campaigns. Sites and time periods were selected with the aim to compare chamber system performance under high and low flux conditions. Due to low temperatures and no fertilizer applied, we expected a low exchange regime during the Braunschweig campaign, whereas higher fluxes were expected at Risø (higher temperatures and a substantial amount of fertilizer applied).

During parallel operation of GC and QCL, chambers were closed for 60 minutes at both sites to measure the concentration increase. When only QCL measurements were conducted, i.e. at Risø at DOY <105.5 and DOY >108.5, chambers were closed for only 10 minutes. For the GC setup, four air samples were taken at 0, 20, 40, and 60 minutes after chamber closure to calculate one flux rate. Air samples (20 mL) were pumped through the tubing system using a membrane pump (3.2 L min^{-1}, NMP 830 KNDC, KNF Neuberger, Freiburg, Germany) and were injected into septum-capped vials. Two cannulas were automatically inserted through the septum, one cannula acting as sample air inlet until overpressure was established, and the other cannula acting as outlet for cycling the air back to the chamber. Air samples were stored in the exchangeable rack of the autosampler unit and were analyzed in the GC-lab of the Thünen Institute using a GC-2014 (Shimadzu, Duisburg, Germany; modified according to Loftfield et al., 1997) with an electron capture detector for N_2O analysis. Performance of the GC-system was checked weekly by conducting ten consecutive measurements of a standard gas with ambient N_2O concentration (320 ppb). Samples were only analyzed if the coefficient of variation of peak areas during this test was below 3 %.

Parallel to the autosampler setup for GC analysis, we operated our chamber system directly connected to a QCL (continuous-wave quantum cascade laser absorption spectrometer, model mini-QCLAS, Aerodyne Research Inc., Billerica, Massachusetts, USA; see Nelson et al. (2004) for principle of operation) in a thermo-controlled housing. Briefly, the laser is thermoelectrically cooled (Thermocube) to 25 °C, uses a 76-m path length, 0.5 L volume and multiple pass absorption cell for sampling, and operates at 40 Torr. It provides a measurement precision of 0.04 ppb (1σ) within an averaging interval of one second. Calibration is performed by continuously aligning the N_2O absorption peak of the sampled air to the standard of the HITRAN database (Rothman et al., 2009). A dry vacuum scroll pump (BOC Edwards XDS10, Sussex, UK) maintained a steady flow rate of 1.0 L min^{-1}. After passing the QCL cell, the sample air was cycled back to the respective chamber to avoid underpressure conditions and unintentional sucking of soil air into chambers. Data was stored on the QCL's internal hard drive at a frequency of 10 Hz.

The detection limit (LoD) of our QCL and GC setups could be estimated using our campaign data assuming stationary conditions during the low flux campaign in Braunschweig. Taking the whole campaign into account, the calculated standard deviations were 2.5 µg m^{-2} h^{-1} and 7.5 µg m^{-2} h^{-1} for QCL and GC measurements, respectively. Thus, the resulting 2-σ uncertainty range for QCL was 5.0 µg m^{-2} h^{-1} and for GC 15.0 µg m^{-2} h^{-1}. If only the first quarter of the Braunschweig campaign data are taken, i.e., a period where environmental conditions were less variable than over the whole campaign, the calculated standard deviations were 1.3 µg m^{-2} h^{-1} and 6.5 µg m^{-2} h^{-1} for QCL and GC measurements, respectively. Thus, the resulting 2-σ uncertainty range for QCL was 2.6 µg m^{-2} h^{-1} and for GC 13.0 µg m^{-2} h^{-1}. These estimates can be regarded as an upper flux detection limit. A supposable lower flux detection limit solely depends on the sensitivity of the analyzers. Precision of the QCL is 0.03 and 0.01 ppb when averaging over 1 and 60 s, respectively. Table 2 summarizes features of the chamber-analyzer system used in this study.

2.3 Flux calculation

GC-based N_2O fluxes using linear, robust linear (Huber, 1981), and modified Hutchinson-Mosier regression (HMR; *cf.* Pedersen et al., 2010) were calculated as described in Leiber-Sauheitl et al. (2014) after converting molar concentrations into mass concentrations using temperature but no pressure correction. Briefly, non-linear flux estimation with the HMR method (R Core Team, 2012; HMR package version 0.3.1) was performed when four data points were available and all of the following criteria were met, i.e. (1) the HMR function could be fitted, (2) Akaike information criterion (AIC; Burnham and Anderson, 2004), which is a measure of (relative) model quality, i.e., gives fit quality penalized by the model's degrees of freedom, and can be used to compare the quality of different model fits to the same dataset, was lower for HMR fit than for linear fit, (3) p value of flux calculated using HMR was lower than that from robust linear fit, and (4) the HMR flux was less than four times larger than the robust linear flux. Otherwise, robust linear regression or ordinary linear regression was used when four or three data points were available, respectively. ~~Briefly, robust linear regression was used as default when four data points were available. Non-linear flux estimation with the HMR method (R Core Team, 2012; HMR package version~~

~~0.3.1) was performed when all of the following criteria were met, i.e. (1) the HMR function could be fitted, (2) Akaike information criterion (AIC; Burnham and Anderson, 2004) was lower for HMR fit than for linear fit, (3) *p* value of flux calculated using HMR was lower than that from robust linear fit, and (4) the HMR flux was less than four times larger than the robust linear flux.~~

QCL-based fluxes were estimated using two different methods. We applied the non-linear HMR model with a slightly modified parameterization (Equation 1 this study; *cf.* Moffat, 2012) to the 60-min dataset of a full chamber cycle (10-min cycle in Risø at DOY <105.5 and DOY >108.5) and compared these fluxes with those resulting from an application of linear regression when only the first three minutes of data after chamber closure were used (*cf.* Section 3.2).

To investigate the frequently observed non-linearity in chamber field data, we computed a quantitative parameter κ describing the curvature in N_2O concentration increase (or decrease) over time (60-min and 10-min QCL data only). Based on the assumption of exponential gas concentration changes in the chamber (*cf.* Nakano et al., 2004) using

$$c(t) = c_{max}\left(1 - \exp\left(\frac{-k}{c_{max}}t\right)\right) + c_0 \quad (1)$$

with $c(t)$ being the N_2O concentration in the chamber at a certain point in time, c_{max} the maximum possible concentration, c_0 the measured concentration at t=0, and k the initial flux F_0 divided by the effective chamber height h, we estimated the N_2O soil-atmosphere flux as the first derivative of Equation (1) evaluated at t=0, i.e.

$$c'(t)|_{t=0} = k \quad (2)$$

and the curvature parameter κ as the second derivative of Equation (1) evaluated at t=0, i.e.

$$c''(t)|_{t=0} = -\frac{k^2}{c_{max}} = \kappa. \quad (3)$$

Units for concentrations $c(t)$, c_{max}, and c_0 are g m^{-3}, units for k are g m^{-2} s^{-1}, and units for κ are g m^{-3} s^{-2}. Negative values of κ correspond to concave curvature indicating a plateauing, i.e. saturating concentration increase over time. Standard errors in this study were calculated as the parameter errors from the respective regression model with the algorithm being based on the Levenberg-Marquardt method ('nlsLM function in R package 'minpack.lm', R Core Team, 2012). Standard errors are solely associated with the flux calculation method and not with any kind of observational errors or issues related to measurement performance such as changes in flow rate, temperature sensitivity of the QCL, pump performance, or changes in chamber volume due to rough soil surfaces or plants in the chamber.

3 Results ~~and discussion~~

3.1 Shape of concentration increase and curvature (κ) determination ~~Non-linear concentration increase and curvature determination~~

Significantly different patterns in chamber N_2O concentration changes during the Braunschweig and Risø campaigns were observed (Figure 2). While increases in the order of 10 to 20 ppb per hour (one chamber cycle) were found for the grassland site in Braunschweig, steep concentration increases measured on the harvested willow field at Risø were almost exclusively higher than 100 ppb per hour and reached maximum rates of over 650 ppb per hour in the period from DOY 105.5 to DOY 108.5. For the low exchange regime in Braunschweig, GC-based data points were highly scattered and rarely showed a clear increasing (or decreasing) tendency making flux calculations difficult. For the high exchange regime at Risø, GC-based concentration data were mostly showing well-defined increases and were similar to those obtained by the QCL system (*cf.* Section 3.3). The latter showed a precise and robust performance with clear base line levels and obvious chamber cycles during both campaigns. None of the QCL-based measurements revealed concentration decreases, i.e. negative fluxes (N_2O uptake) while chambers were closed.

~~The high time resolution of QCL data allowed for a closer look at the increase characteristics and showed that only in a very few cases $c'(t)$ was strictly constant when taking the full 60-min cycle into account. This was expected as the rate of transport of a diffusing trace gas into the chamber headspace necessarily declines throughout deployment because any increase in the headspace concentration results in a corresponding decline in the vertical concentration gradient driving that transport (Rolston, 1986; Hutchinson et al., 2000; Livingston et al., 2006).~~ Results of the investigation on quantifying the curvature in $c(t)$, expressed as κ, are given in Figure 3.

Extremely low ~~negative~~ absolute κ values between -10^{-4} and -10^{0} – indicating quasi linearity in $\partial c/\partial t$ – were almost exclusively found under low flux conditions, whereas fluxes >100 µg N m^{-2} h^{-1} were only observed when κ was < -10^{1} (Figure 3A). ~~This means that at higher fluxes the curvature in $c(t)$ is concave suggesting concentrations that tend to plateau over time with the saturation effect becoming larger at higher flux rates. Near-zero fluxes, however, corresponding to κ values around zero, indicate no considerable changes in N_2O concentrations, thus, hardly any alteration of the vertical concentration gradient over time. Furthermore, closure time was found to have an impact on the magnitude of κ (Figure 3B). Longer chamber deployment led to higher curvature in $c(t)$, which was expected as concentration gradients decline over time when a considerable flux is measured (*cf.* Hutchinson and Mosier, 1981; Livingston et al., 2006; Pedersen et al., 2010).~~

While closure time was found to have an impact on the magnitude of κ (Figure 3B), i.e. higher κ coincided as expected with longer closure times, κ values and its corresponding concentration build up characteristics were invariant among a broad range of wind speed and wind direction classes (Figure 3C and 3D). ~~Our results imply that at low flux rates and/or short chamber closure, the slight non-linearity in concentration change when calculating fluxes is of minor importance and the application of linear models is acceptable, particularly with regard to other commonly observed errors such as those originating from soil disturbance, chamber placement (Christiansen et al., 2011), temperature, pressure and humidity~~

perturbations, etc. (Parkin and Venterea, 2010). At higher fluxes, however, significant curvature in *c(t)* expressed by large negative κ values will most likely lead to a substantial underestimation of fluxes when using linear regression instead of applying an exponential model for flux calculation (*cf.* Matthias et al., 1978; Jury et al., 1982; Anthony et al., 1995; Kroon et al., 2008; Section 3.2). In principle, several other reasons making flux determination with linear or exponential models problematic, may technically be found. These are exponentially increasing N_2O concentrations after chamber closure due to possible dispersion effects leading to biased analyzer readings, placement of the sample tube inlet at the top of the chamber lid leading to an establishment of a temporary concentration gradient in a weakly mixed chamber atmosphere or an insufficient dimension of the pressure compensation tube leading to a push back of air into the uppermost soil layer in the moment when the chamber is set onto the lid. However, none of these were observed during our campaigns, thereby indicating a robust setup and chamber design for reliable N_2O flux calculations.

We also investigated the possible effect of ambient wind speed and direction on concentration build up characteristics (Figure 3C and 3D, respectively) as differences between the turbulence regime outside the chamber and inside chamber conditions vary under changing wind speed. Theoretically, soil pores might be more ventilated and the coupling of the flux under ambient conditions to the atmosphere is much closer. Consequently, the establishment of a steady state flux may be more postponed under higher wind speed. Such time delay caused by slow filling up of the previously ventilated pore space in parallel to the diffusion into the chamber might in principle explain exponentially increasing concentrations. However, κ values (Figure 3C and D) and fluxes (not shown) were independent of both wind speed and direction, which is a further indicator that the chosen chamber design and setup can be used over a wide range of environmental conditions and neither seem to affect concentration build up characteristics nor resulting flux magnitudes.

3.2 Comparison of N_2O fluxes and their associated errors from linear and non-linear regression models~~3 min vs. 60 min closure time in QCL measurements~~

~~Reviewing past decades of field chamber measurements for studying soil-atmosphere exchange of N_2O, several challenges and shortcomings emerged such as limited number of replicates or disturbance of the soil micro-environment due to chamber coverage and soil collar insertion (e.g., Hutchinson and Mosier, 1981; Parkin and Venterea, 2010). One way of getting a higher temporal resolution, thereby a higher number of replicates and keeping soil disturbance as low as possible is to reduce the chamber closure period, which also is expected to decrease deviation from linearity in concentration increase.~~ With the QCL's high time resolution – in this study operated at the analyzer's maximum frequency of 10 Hz – we compared N_2O flux estimates based on 60-min (DOY 105.5 to DOY 108.5) and 10-min (DOY <105.5 and DOY >108.5) closure periods calculated by the modified HMR approach with those flux estimates that are based on the first three minutes of concentration data only and were calculated by linear regression. The Risø dataset was used for this comparison, because both high and low fluxes were observed.

Flux estimates of the two approaches matched reasonably well; significant differences were only observed at very high rates (Figure 4A and 4B). 85 % of the variance in N_2O fluxes from 3-min closure could be explained by fluxes from 60-min and

10-min closure (Figure 4B). The relatively high slope of 1.80 was mainly caused by three exceptionally high fluxes where the 60-min method considerably overestimated values of the 3-min method~~, which could have been caused by disturbances during the 60-min chamber deployment due to soil macrofauna. Recent work e.g. by Kroon et al. (2008) and Forbrich et al. (2010) demonstrated that emission estimates from closed chamber measurements were significantly underestimated when using linear regression methods instead of the slope of an exponential function at the beginning of chamber closure. However, their linear regression models were applied to longer periods, i.e. to 10-min periods by Kroon et al. (2008) also using an Aerodyne QCL spectrometer and to 25-min periods by Forbrich et al. (2010) using a GC setup. Kroon et al. (2008) also showed that linear estimates differed by up to 60 % compared to those from exponential methods with a systematic error due to linear regression being in the same order as the estimated uncertainty due to temporal variation.~~. ~~In our study,~~ ~~s~~Standard errors of N_2O fluxes from both 3-min and 60-min closure were extremely low, i.e. in the order of 0.2~~1~~ % of the fluxes (Figure 4C) with median values of 0.17 and 0.06 µg N m^{-2} h^{-1} and arithmetic means of 0.21 and 0.20 µg N m^{-2} h^{-1} for the 3 and 60-min closure flux estimates, respectively. ~~This implies that limiting the chamber closure period to 3 minutes benefits in two ways. On the one hand, the soil column of interest is less disturbed by shorter coverage and/or the number of replicates can be significantly increased. As these measurements are automated, no further manual work is required. On the other hand, standard errors of fluxes remain extremely low. However, it is recommended to extend the chamber closure period to a minimum of five minutes as slightly delayed concentration increases under low flux regimes (*cf.* Section 3.1) may occur and would lead to an underestimation of 3-min linear fluxes. We therefore recommend skipping the first 2 minutes of data to guarantee undisturbed conditions that might have been caused in the moment when the chamber is set on the soil collar.~~

For better comparison with other studies, we also compared HMR-based fluxes with robust linearly calculated fluxes from our GC measurements when the full 60-min cycle was taken into account. A linear regression analysis (data not shown) resulted in a slope of 0.97 and an R^2 value of 0.86 under the high flux regime in Risø with the data set of robust linearly calculated fluxes being the independent variable. The mean campaign flux value from HMR-based calculations was 22 % higher than the average campaign value of the robust linear method. The difference between the two methods was even higher under the low flux regime in Braunschweig. Slope and R^2 value of a linear regression analysis were 1.82 and 0.42, respectively. Despite the high slope value, the mean campaign value of the robust linear method only reached 51 % of the value obtained from the HMR method.

A further intriguing analysis shows that standard errors were found to be invariant on QCL sampling frequency (Figure 5). We simulated different sampling times ranging from one tenth of a second to 25.6 sec, which corresponds to a frequency of 0.0390625 Hz, by excluding the respective intervals from the original 10-Hz dataset. Results show that the median of the standard error of the fluxes remains stable over a wide range of measurement frequencies. At a frequency class of 0.15 and lower (3 boxes on the right-hand side of Figure 5), which corresponds to a sampling time of ~5 sec and higher, lower and upper quartile values begin to deviate and the median changes slightly. ~~Furthermore, standard errors were found to depend on QCL sampling frequency (Figure 5). We simulated different sampling times ranging from one tenth of a second to~~

approx. one minute, by excluding the respective intervals from the original 10-Hz dataset. Results showed a clear tendency towards higher standard errors, when fewer data points were included. Arithmetic means of standard errors increased steadily from 0.21 µg N m^{-2} h^{-1} at 10 Hz to 3.36 µg N m^{-2} h^{-1} at 0.03 Hz and from 0.20 µg N m^{-2} h^{-1} at 10 Hz to 2.93 µg N m^{-2} h^{-1} at 0.03 Hz for the 3-min and 60-min method, respectively. These results suggest that – in case no high-frequency measurement devices are available or data storage capacity is limited – sampling times between 1 and 5 seconds are sufficient to keep standard errors of N_2O fluxes on a much lower level than those obtained from GC measurements (*cf.* Section 3.3).

3.3 GC vs. QCL-based fluxes under low and high exchange regimes

Time series of N_2O fluxes and their associated standard errors using both the GC and the 3-min QCL linear regression method during the Braunschweig and Risø campaigns are given in Figure 6. QCL fluxes in Braunschweig were on a constantly low level ranging between 2 and 16 µg N m^{-2} h^{-1}, whereas GC-based fluxes at the same site were scattered between –13 and 39 µg N m^{-2} h^{-1}. A linear regression revealed no significant relationship between GC and QCL fluxes with a very low coefficient of determination of 0.036 (Figure 7A). Less than 1 % of the variance in QCL fluxes could be explained by GC fluxes (Figure 7). While standard errors of the QCL method were always below 0.6 µg N m^{-2} h^{-1}, values of the GC method were distributed between 0.5 and 22.0 µg N m^{-2} h^{-1}. Although higher variability and higher standard errors in GC-based fluxes were evident, mean N_2O flux rates of the entire observation period were almost identical when comparing the two analyzer types. 6.42 ± 5.98 µg N m^{-2} h^{-1} and 7.77 ± 0.13 µg N m^{-2} h^{-1} were found for the GC and the QCL method, respectively.

Under the high exchange regime at Risø, N_2O fluxes of the two analyzer types matched considerably better (Figure 6D). Although the willow field was already fertilized on DOY 99, N_2O fluxes did not start to increase until DOY 105 when a sharp rise in air temperature was observed. GC-based fluxes were lower than QCL-based fluxes (slope=0.50) as in most cases a non-linear model could not be fitted with only four data points. A linear regression between GC and QCL fluxes revealed a coefficient of determination of 0.48 (Figure 7B). 48 % of the variance in QCL-based fluxes could be explained by fluxes from the GC method. Standard errors of the QCL method were again extremely low, i.e. <1 % of the flux value and were always below 1.0 µg N m^{-2} h^{-1}, while those from the GC method were on average in the range of 5 to 10 % of the flux value. Parallel operation of both methods was conducted from DOY 105 to 108. During this period, the campaign means were 117.8 ± 0.2 and 77.4 ± 8.2 µg N m^{-2} h^{-1} for the QCL and GC method, respectively.

As standard errors of QCL-based N_2O fluxes were on a constantly low level, no dependency on flux value was observed in any of the campaigns (Figure 7). The same was evident for GC-based fluxes in Braunschweig. At Risø, however, a slight but non-significant tendency of higher standard errors at higher flux rates was found. Only 8 % of GC data from Braunschweig met the criteria for flux calculation using the HMR model. At Risø, 38 % of GC data allowed for HMR flux calculation indicating that higher exchange regimes favor the usage of an exponential model when using the GC method. Similar

findings (37 % allowance for non-linear model application) were reported by Petersen et al. (2011). Forbrich et al. (2010) found percentages of 13.6 %, 19.2 %, and 9.8 % of GC measurements on hummocks, lawns and flarks, respectively, which were best fitted with an exponential model. Their measurements, however, were made for methane fluxes and under an even larger $\partial c/\partial t$ range than was prevalent in our two campaigns. The fact that higher fluxes in our study were associated with lower standard errors and accepted HMR application corresponds well with κ findings in Section 3.1 indicates that higher curvature in $c(t)$ coincided with higher fluxes.

In summary, our comparison of GC vs. QCL fluxes revealed that despite much higher precision, robustness, and temporal resolution in QCL measurements, GC is still a useful method to determine soil-atmosphere exchange of N_2O at longer time scales (here several weeks) when aiming at investigating N_2O budgets. Although single flux values particularly under low exchange regimes did not match well, campaign means were similar to those obtained by the QCL method. Under high exchange regimes, however, flux patterns matched considerably better, but resulted in larger absolute errors when comparing the campaign average, thereby leading to systematic errors (in our case an underestimation) when using the GC method at high N_2O fluxes for the assessment of N balances.

In general, chamber architecture is essential for headspace concentration buildup patterns given certain enclosure times, activity levels and headspace mixing. Our new chamber system performed well during the two campaigns for both analyzer setups. Through its specific design with not only vertically but also horizontally moving chambers, it will keep the soil column under relatively natural conditions. The only problem emerged at Risø when the guiding racks were slightly distorted under high wind speed conditions, i.e. when half-hourly means of wind speed were higher than 6 m s^{-1}. However, this problem could easily be fixed by tightening the guy wires that are attached to the aluminum rack. Commonly observed winter problems such as unnatural accumulation of snow inside the chamber and rime ice formation on the guiding racks and soil frame were not tested within this study, but will likely affect the ease of operation during harsh winter conditions.

3.4 N_2O uptake and diurnal variation

From an ecological point of view, QCL measurements offer a new opportunity for robust quantification of soil N_2O consumption. As N_2O uptake via denitrification exists in theory and could be shown under controlled lab conditions (e.g., Firestone and Davidson, 1989), it has been a major challenge to measure reliable fluxes in the field due to the fact that the magnitude of N_2O uptake by soils is usually very low (Schlesinger, 2013), thereby problematic to be determined by GC measurements that are subjected to low signal to noise ratios (e.g., Brümmer et al., 2008).

In our study, nNeither at Risø nor during the Braunschweig campaign soil N_2O uptake was observed when using QCL based measurements. In only very few cases (n=5) $c'(t)$ were initially found to be negative, however, these data, which exhibited abnormally high standard errors, were discarded due to mechanical malfunctioning of the chamber system as a result of non-closure caused by distorted guiding racks through very high wind speeds at Risø (*cf.* Section 3.3).

Regarding GC-based data in our study, 2 out of 37 fluxes in Risø were negative. Note that GC-based fluxes in Risø were only determined between DOY 105.5 and 108.5 when fluxes were elevated due to fertilizer application. In Braunschweig,

however, nearly 25 %, i.e. 50 out of 201 flux rates from the GC setup were showing N_2O uptake with only 3 of the 50 negative flux rates being significant (p<0.05, p-values not corrected for multiple testing). ~~, thus stressing the challenge of a robust determination of soil consumption of this important greenhouse gas when using the common vial-GC approach. Due to the fact that in this study no N_2O soil uptake was found when using the much more reliable QCL setup, a further investigation of this topic on a variety of soil types under different land uses, plant communities, and climatic conditions is highly desired.~~

An investigation of the diurnal variability of N_2O fluxes showed that during the Braunschweig campaign – although only small differences were observed – highest fluxes were found during midday and early afternoon (~8.7 µg N m^{-2} h^{-1}), while lowest N_2O efflux was measured shortly before midnight and before sunrise (~7.2 and 7.3 µg N m^{-2} h^{-1}, respectively; Figure 8), thereby following a commonly observed temperature-driven pattern (*cf.* Section 4.4). In Risø, however, we found lowest fluxes of ~18.2 µg N m^{-2} h^{-1} at midday and highest fluxes when it was dark peaking before midnight at ~32.0 µg N m^{-2} h^{-1} (only data of DOY <105.5 and DOY >108.5 were taken to exclude fertilizer effects). Error bars in Figure 8 indicate the standard error of the mean from all flux values in each bin. Each bin contains fluxes from 3-hour periods, i.e. from 00:00 to 03:00, 03:00 to 06:00, 06:00 to 09:00, etc. The mean values in Figure 8 are plotted in the center of each bin. Fluxes were binned due to irregular starting times of new chamber cycles. In general, a new chamber cycle could be started each full hour, but to get a more robust diurnal pattern, we decided to bin data in the above-mentioned 3-hour containers. While the diurnal variation of N_2O fluxes from the Risø campaign is significant (p-value = 0.0059), the diurnal variation found during the Braunschweig campaign is not as the difference between mean minimum and maximum values is lower than the upper flux detection limit of ~2.6 µg N m^{-2} s^{-1}.

~~Besides investigating possible N_2O soil uptake, the QCL methodology offers the opportunity to study diurnal variability of N_2O fluxes. In a recent study by Shurpali et al. (2016) it has been pointed out that neglecting diurnal variations leads to uncertainties in terrestrial N_2O emission estimates and should therefore be carefully taken into account when calculating nitrogen budgets. Similar to our study (Figure 8), they found reversed diurnal patterns under differing flux magnitudes. Intriguingly, when mean N_2O fluxes were in a range between 12 and 35 µg N m^{-2} h^{-1}, in both this study (Risø low flux regime) and Shurpali et al. (2016) highest fluxes were found during nighttime and lowest fluxes around midday. A contrasting diurnal pattern was observed when fluxes were lower than during the Risø campaign, i.e. in Braunschweig (7.2 to 8.7 µg N m^{-2} h^{-1}) or much higher due to fertilizer application as in Shurpali et al. (2016) (230 to 475 µg N m^{-2} h^{-1}). In the latter campaigns, mean N_2O fluxes were highest at midday and lowest during nighttime, which corresponds to earlier findings (*e.g.*, Christensen, 1983; Du et al., 2006; Parkin and Kaspar, 2006; Brümmer et al., 2008; Alves et al., 2012) where temperature was proved to be the main controlling factor for N_2O soil-atmosphere exchange. Our study highlights that through its high time resolution QCL-based measurements will not only help enhance process understanding of N_2O exchange by disentangling the strength of different drivers of N_2O production like temperature, soil moisture, nitrogen~~

~~availability, and microbial activity, but has also the potential to provide new insight into bidirectional exchange characteristics of other trace gases such as CH_4.~~

4 Discussion

4.1 The curvature parameter κ as a chamber performance criteria

The high time resolution of QCL data allowed for a closer look at the increase characteristics and showed that only in a very few cases *c'(t)* was strictly constant when taking the full 60-min cycle into account. This was expected as the rate of transport of a diffusing trace gas into the chamber headspace necessarily declines throughout deployment because any increase in the headspace concentration results in a corresponding decline in the vertical concentration gradient driving that transport (Rolston, 1986; Hutchinson et al., 2000; Livingston et al., 2006).

The fact that extremely low negative κ values between -10^{-4} and -10^{0} – indicating quasi linearity in $\partial c/\partial t$ – were almost exclusively found under low flux conditions, whereas fluxes >100 µg N m^{-2} h^{-1} were only observed when κ was < -10^{1} (Figure 3A)~~This~~ means that at higher fluxes the curvature in *c(t)* is concave suggesting concentrations that tend to plateau over time with the saturation effect becoming larger at higher flux rates. Near zero fluxes, however, corresponding to κ values around zero, indicate no considerable changes in N_2O concentrations, thus, hardly any alteration of the vertical concentration gradient over time. Furthermore, closure time was found to have an impact on the magnitude of κ (Figure 3B). Longer chamber deployment led to higher curvature in *c(t)*, which was expected as concentration gradients decline over time when a considerable flux is measured (*cf.* Hutchinson and Mosier, 1981; Livingston et al., 2006; Pedersen et al., 2010).

Our results imply that at low to moderately high flux rates <200 µg N m^{-2} h^{-1} (*cf.* Figure 4D) and/or short chamber closure, the slight non-linearity in concentration change when calculating fluxes is of minor importance and the application of linear models is acceptable, particularly with regard to other commonly observed errors such as those originating from soil disturbance, chamber placement (Christiansen et al., 2011), temperature, pressure and humidity perturbations, etc. (Parkin and Venterea, 2010).~~Our results imply that at low flux rates and/or short chamber closure, the slight non-linearity in concentration change when calculating fluxes is of minor importance and the application of linear models is acceptable, particularly with regard to other commonly observed errors such as those originating from soil disturbance, chamber placement (Christiansen et al., 2011), temperature, pressure and humidity perturbations, etc. (Parkin and Venterea, 2010).~~ At higher fluxes, however, significant curvature in *c(t)* expressed by large negative κ values will most likely lead to a substantial underestimation of fluxes when using linear regression instead of applying an exponential model for flux calculation (*cf.* Matthias et al., 1978; Jury et al., 1982; Anthony et al., 1995; Kroon et al., 2008; Section 3.2). In principle, several other reasons making flux determination with linear or exponential models problematic, may technically be found. These are exponentially increasing N_2O concentrations after chamber closure due to possible dispersion effects leading to biased analyzer readings when the elevated gas concentration is initially not uniformly mixed with the air inside the tubing,

placement of the sample tube inlet at the top of the chamber lid leading to an establishment of a temporary concentration gradient in a weakly mixed chamber atmosphere or an insufficient dimension of the pressure compensation tube leading to a push back of air into the uppermost soil layer in the moment when the chamber is set onto the lid. However, none of these were observed during our campaigns, thereby indicating a robust setup and chamber design for reliable N_2O flux calculations.

We also investigated the possible effect of ambient wind speed and direction on concentration build up characteristics (Figure 3C and 3D, respectively) as differences between ~~the~~ turbulence conditions~~regime~~ outside the chamber may possibly vary from those conditions~~and~~ inside the chamber ~~conditions vary~~ under changing wind speed. Theoretically, pores in the uppermost soil layer might be ventilated under high wind speed when no chamber is in place, thus a close coupling of the flux to the atmosphere exists. Consequently, the establishment of a steady state flux may be more postponed under these high wind speed conditions once the chamber is put onto the soil frame.~~Theoretically, soil pores might be more ventilated and the coupling of the flux under ambient conditions to the atmosphere is much closer. Consequently, the establishment of a steady state flux may be more postponed under higher wind speed.~~ Such time delay caused by slow filling up of the previously ventilated pore space in parallel to the diffusion into the chamber might in principle explain exponentially increasing concentrations. However, κ values (Figure 3C and D) and fluxes (not shown) were independent of both wind speed and direction, which is a further indicator that the chosen chamber design and setup can be used over a wide range of environmental conditions and neither seem to affect concentration build up characteristics nor resulting flux magnitudes.

4.2 Closure time and measurement frequency – How long and how often is enough?

Reviewing past decades of field chamber measurements for studying soil-atmosphere exchange of N_2O, several challenges and shortcomings emerged such as limited number of replicates or disturbance of the soil micro-environment due to chamber coverage and soil collar insertion (e.g., Hutchinson and Mosier, 1981; Parkin and Venterea, 2010). One way of getting a higher temporal resolution, thereby a higher number of replicates and keeping soil disturbance as low as possible is to reduce the chamber closure period, which also is expected to decrease deviation from linearity in concentration increase.

The overestimation of the 60-min method compared to the 3-min method as shown in Figure 4B causing a relatively high slope of 1.80 was mainly caused by three exceptionally high fluxes. Beside any form of unintended interferences to the 'natural steady-state flux' like for example disturbances through macrofauna, fluctuating pump performance or analyzer malfunctions due to internal re-calibration during chamber deployment, much higher 60-min-based HMR fluxes compared to 3-min-based linear fluxes may be observed when one of two following concentration increase patterns are observed.

(1) Slow initial increase of concentrations followed by steeper rise after some minutes. Slope of the linear fit will then be much lower than the one from the HMR fit (linear fit at t_0).

(2) Steady linear start of concentration increase followed by sudden relatively sharp bend with lower linear increase afterwards. HMR fit will also have a much steeper slope at t_0 than the linear fit, which will be on top of the data points for the first few minutes.

Red dots in Fig. 4B indicate situations similar to those described under (2) above. ~~, which could have been caused by disturbances during the 60-min chamber deployment due to soil macrofauna.~~ Recent work e.g. by Kroon et al. (2008) and Forbrich et al. (2010) demonstrated that emission estimates from closed chamber measurements were significantly underestimated when using linear regression methods instead of the slope of an exponential function at the beginning of chamber closure. However, their linear regression models were applied to longer periods, i.e. to 10-min periods by Kroon et al. (2008) also using an Aerodyne QCL spectrometer and to 25-min periods by Forbrich et al. (2010) using a GC setup. Kroon et al. (2008) also showed that linear estimates differed by up to 60 % compared to those from exponential methods with a systematic error due to linear regression being in the same order as the estimated uncertainty due to temporal variation.

As shown in Figure 4C, standard errors of N_2O fluxes from both 3-min and 60-min closure were extremely low, i.e. in the order of 0.2 % of the fluxes with median values of 0.17 and 0.06 µg N m^{-2} h^{-1} and arithmetic means of 0.21 and 0.20 µg N m^{-2} h^{-1} for the 3 and 60-min closure flux estimates, respectively. In comparison, Cowan et al. (2014a) also find low flux uncertainty of less than 1 to 2 µg N m^{-2} h^{-1}. This implies that limiting the chamber closure period to 3 minutes benefits in two ways. On the one hand, the soil column of interest is less disturbed by shorter coverage and/or the number of replicates can be significantly increased. As these measurements are automated, no further manual work is required. On the other hand, standard errors of fluxes remain extremely low. However, it is recommended to extend the chamber closure period to a minimum of five minutes as slightly delayed concentration increases under low flux regimes may occur (in ~5 % of the cases in our study) and would lead to an underestimation of 3-min linear fluxes (see Figure S1 in the Supplement). We therefore recommend skipping the first 2 minutes of data to guarantee undisturbed conditions that might have been caused in the moment when the chamber is set on the soil collar. The 'dead time' of the system, i.e. the time that passes between the moment when an air sample leaves the chamber and the moment when it reaches the analyzer, was ~10 s – given a tube length of 10 m, a flow rate of 1 L min^{-1}, and an ID of the tube of 4.6 mm – and was already considered in the recommendation.

Standard errors of N_2O fluxes were found to be invariant on QCL sampling frequency (Figure 5). The conclusion we can draw from this finding is that chamber operators – in case an analyzer with a precision like the QCL presented in this study is available – can reduce their sampling time down to 5 seconds without risking an increase of the standard error of the flux, which would still be on a much lower level than those obtained from GC measurements (*cf.* results in Section 3.2).~~These results suggest that – in case no high-frequency measurement devices are available or data storage capacity is limited – sampling times between 1 and 5 seconds are sufficient to keep standard errors of N_2O fluxes on a much lower level than those obtained from GC measurements (*cf.* Section 3.3).~~

4.3 Differences between GC and QCL-based fluxes

~~Similar findings (37 % allowance for non-linear model application) were reported by Petersen et al. (2011). Forbrich et al. (2010) found percentages of 13.6 %, 19.2 %, and 9.8 % of GC measurements on hummocks, lawns and flarks, respectively, which were best fitted with an exponential model. Their measurements, however, were made for methane fluxes and under an even larger $\partial c/\partial t$ range than was prevalent in our two campaigns. The fact that higher fluxes in our study were associated with lower standard errors and accepted HMR application corresponds well with κ findings in Section 3.1 indicates that higher curvature in $c(t)$ coincided with higher fluxes.~~

Our comparison of GC vs. QCL fluxes revealed that despite much higher precision, robustness, and temporal resolution in QCL measurements, GC is still a useful method to determine the average campaign N_2O soil efflux. Although single flux values particularly under low exchange regimes did not match well, campaign means and medians were similar to those obtained by the QCL method. Under high exchange regimes, however, flux patterns matched considerably better, but resulted in larger absolute errors when comparing the campaign average, thereby leading to systematic errors (in our case an underestimation) when using the GC method at high N_2O fluxes for the assessment of N balances. However, given the fact that the bulk of the annual efflux occurs after management events at a relatively short time scale (Flechard et al., 2007; Skiba et al., 2013), usage of a GC-based system will be prone to large uncertainties (*cf.* Figure 7).~~In summary, our comparison of GC vs. QCL fluxes revealed that despite much higher precision, robustness, and temporal resolution in QCL measurements, GC is still a useful method to determine soil-atmosphere exchange of N_2O at longer time scales (here several weeks) when aiming at investigating N_2O budgets. Although single flux values particularly under low exchange regimes did not match well, campaign means were similar to those obtained by the QCL method. Under high exchange regimes, however, flux patterns matched considerably better, but resulted in larger absolute errors when comparing the campaign average, thereby leading to systematic errors (in our case an underestimation) when using the GC method at high N_2O fluxes for the assessment of N balances.~~

While only 8 % of GC data from the Braunschweig campaign met the criteria for flux calculation using the HMR model, 38 % of GC from the Risø data allowed for HMR flux calculation indicating that higher exchange regimes favor the usage of an exponential model when using the GC method.

Similar findings (37 % allowance for non-linear model application) were reported by Petersen et al. (2011). Forbrich et al. (2010) found percentages of 13.6 %, 19.2 %, and 9.8 % of GC measurements on hummocks, lawns and flarks, respectively, which were best fitted with an exponential model. Their measurements, however, were made for methane fluxes and under an even larger $\partial c/\partial t$ range than was prevalent in our two campaigns. The fact that higher fluxes in our study were associated with lower standard errors and accepted HMR application corresponds well with κ findings in Section 3.1 indicates that higher curvature in $c(t)$ coincided with higher fluxes.

In general, chamber architecture is essential for headspace concentration buildup patterns given certain enclosure times, activity levels and headspace mixing. Our new chamber system performed well during the two campaigns for both analyzer

setups. Through its specific design with not only vertically but also horizontally moving chambers, it will keep the soil column under relatively natural conditions. The only problem emerged at Risø when the guiding racks were slightly distorted under high wind speed conditions, i.e. when half-hourly means of wind speed were higher than 6 m s^{-1}. However, this problem could easily be fixed by tightening the guy wires that are attached to the aluminum rack. Commonly observed winter problems such as unnatural accumulation of snow inside the chamber and rime ice formation on the guiding racks and soil frame were not tested within this study, but will likely affect the ease of operation during harsh winter conditions.

4.4 Enabling investigations of flux pattern characteristics

From an ecological point of view, QCL measurements offer a new opportunity for robust quantification of soil N_2O consumption. As N_2O uptake via denitrification exists in theory and could be shown under controlled lab conditions (e.g., Firestone and Davidson, 1989), it has been a major challenge to measure reliable fluxes in the field due to the fact that the magnitude of N_2O uptake by soils is usually very low (Schlesinger, 2013), thereby problematic to be determined by GC measurements that are subjected to low signal to noise ratios (e.g., Brümmer et al., 2008).

Our QCL-based measurements under the given soil, temperature and moisture conditions at Risø and Braunschweig did not result in any soil N_2O uptake fluxes. In the study by Cowan et al. (2014b), approx. 10% of their fluxes on grazed grassland and barley sites were negative. However, only 4 out of 115 negative fluxes were above the LoD of the method, which was estimated to be 4 µg N m^{-2} h^{-1}, thus being similar to ours (cf. Table 2).

GC-based data in our study showed 2 out of 37 and 50 out of 201 negative fluxes in Risø and Braunschweig, respectively. In Risø, only 3 of the 50 negative flux rates were found to be significant (p<0.05, p-values not corrected for multiple testing), thus stressing the challenge of a robust determination of soil consumption of this important greenhouse gas when using the common vial-GC approach. Due to the fact that in this study no N_2O soil uptake was found when using the much more reliable QCL setup, a further investigation of this topic on a variety of soil types under different land uses, plant communities, and climatic conditions is highly desired.

Besides investigating possible N_2O soil uptake, the QCL methodology offers the opportunity to study diurnal variability of N_2O fluxes. In a recent study by Shurpali et al. (2016) it has been pointed out that neglecting diurnal variations leads to uncertainties in terrestrial N_2O emission estimates and should therefore be carefully taken into account when calculating nitrogen budgets. Similar to our study (Figure 8), they found reversed diurnal patterns under differing flux magnitudes. Intriguingly, when mean N_2O fluxes were in a range between 12 and 35 µg N m^{-2} h^{-1}, in both this study (Risø low flux regime) and Shurpali et al. (2016) highest fluxes were found during nighttime and lowest fluxes around midday. A contrasting diurnal pattern was observed when fluxes were lower than during the Risø campaign, i.e. in Braunschweig (7.2 to 8.7 µg N m^{-2} h^{-1}) or much higher due to fertilizer application as in Shurpali et al. (2016) (230 to 475 µg N m^{-2} h^{-1}). In the latter campaigns, mean N_2O fluxes were highest at midday and lowest during nighttime, which corresponds to earlier findings (*e.g.*, Christensen, 1983; Du et al., 2006; Parkin and Kaspar, 2006; Brümmer et al., 2008; Alves et al., 2012) where

temperature was proved to be the main controlling factor for N_2O soil-atmosphere exchange. Our study highlights that through its high time resolution QCL-based measurements will not only help enhance process understanding of N_2O exchange by disentangling the strength of different drivers of N_2O production like temperature, soil moisture, nitrogen availability, and microbial activity, but has also the potential to provide new insight into bidirectional exchange characteristics of other trace gases such as CH_4, which can be sampled simultaneously with our chamber system depending on analyzer type used.

5 Conclusions

A new chamber system for automated measurements of soil-atmosphere trace gas exchange was developed. The system was tested for N_2O flux detection in a conventional vial air sampling setup and with a directly connected QCL spectrometer under low and high exchange regimes. Through its specific design, the system prevents measurement spots from unintended shading and minimizes disturbance of throughfall, thereby complying with high quality requirements of long-term observation studies and research infrastructures. Curvature in $\partial c/\partial t$ proved to be invariant with wind speed and direction. High correlation (slope=0.99; R^2=0.93) was found when comparing linear vs. modified HMR flux calculation methods for F_{N2O}<200 µg N m^{-2} h^{-1}. Intriguingly, mean campaign N_2O fluxes measured by GC and QCL of 6.42 and 7.77 µg N m^{-2} h^{-1}, respectively, matched fairly well under low flux conditions, whereas under high flux conditions a significant deviation was observed (77.40 vs. 122.95 µg N m^{-2} h^{-1} from GC and QCL, respectively). While mean standard errors were in a range of 10 to 93 % of the N_2O flux in low to high exchange regimes when using the GC approach, extremely low values for standard errors of 0.2 to 1.7 % of the flux under different exchange conditions were found for QCL measurements. When a fast-response analyzer is available, we recommend reducing chamber closure time to 10 minutes and a measurement frequency of 0.2 Hz, i.e. a sampling output of every 5 seconds. With its high precision and temporal resolution, QCL technology provides furthermore a powerful tool to investigate highly debated topics such as diurnal flux variability and soil N_2O uptake.

Acknowledgements

This study was supported by the Thünen Institute of Climate-Smart Agriculture through the German Federal Ministry of Food and Agriculture (BMEL) as well as the Integrated Carbon Observation System (ICOS) infrastructure through the Federal Ministry of Research and Education (BMBF). Funding for CB and JJSR by the BMBF Junior Research Group NITROSPHERE under support code FKZ 01LN1308A is greatly acknowledged. We highly appreciate logistical support during the Risø measurements, which were conducted in the framework of the InGOS project. The excellent introduction and valuable help on laser spectrometer operation and maintenance by David D. Nelson and Mark Zahniser is gratefully acknowledged. Many thanks are owed to Florian Hackelsperger from the experimental research station of the Institute of

Animal Nutrition for technical support during the Braunschweig campaign as well as Kerstin Gilke and Andrea Oehns-Rittgerodt for N_2O analyses in the GC lab. Michel Bechtold is thanked for discussing and calculating dispersion estimates.

References

Alves, B.J.R., Smith, K.A., Flores, R.A., Cardoso, A.S., Oliveira, W.R.D., Jantalia, C.P., Urquiaga, S., and Boddey, R.M.: Selection of the most suitable sampling time for static chambers for the estimation of daily mean N_2O flux from soils. Soil Biol. Biochem. 46, 129–135, 2012.

Ammann, C., Wolff, V., Marx, O., Brümmer, C., and Neftel, A.: Measuring the biosphere-atmosphere exchange of total reactive nitrogen by eddy covariance. Biogeosciences 9, 4247–4261, 2012.

Anthony, W.H., Hutchinson, G.L., and Livingston, G.P.: Chamber measurement of soil–atmosphere gas exchange: Linear vs. diffusion-based flux models. Soil Sci. Soc. Am. J. 59, 1308–1310, 1995.

Baldocchi, D.D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B.E., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.C.: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor and energy flux densities. B. Am. Meteorol. Soc. 82, 2415–2434, 2001.

Brümmer, C., Brüggemann, N., Butterbach-Bahl, K., Falk, U., Szarzynski, J., Vielhauer, K., Wassmann, R., and Papen, H.: Soil-atmosphere exchange of N_2O and NO in near-natural savanna and agricultural land in Burkina Faso (W. Africa). Ecosystems 11, 582–600, 2008.

Brümmer, C., Black, T.A., Jassal, R.S., Grant, N.J., Spittlehouse, D.L., Chen, B., Nesic, Z., Amiro, B.D., Arain, M.A., Barr, A.G., Bourque, C.P.A., Coursolle, C., Dunn, A.L., Flanagan, L.B., Humphreys, E.R., Lafleur, P.M., Margolis, H.A., McCaughey, J.H., and Wofsy, S.C.: How climate and vegetation type influence evapotranspiration and water use efficiency in Canadian forest, peatland and grassland ecosystems. Agric. For. Meteorol. 153, 14–30, 2012.

Brümmer, C., Marx, O., Kutsch, W.L., Ammann, C., Wolff, V., Fléchard, C.R., and Freibauer, A.: Fluxes of total reactive atmospheric nitrogen ($\sum N_r$) using eddy covariance above arable land. Tellus B 65, 19770, http://dx.doi.org/10.3402/tellusb.v65i0.19770, 2013.

Burnham, K.P. and Anderson, D.R.: Multimodel Inference: Understanding AIC and BIC in Model Selection, Sociol. Method. Res. 33, 261–304, 2004.

Butterbach-Bahl, K., Baggs, E.M., Dannenmann, M., Kiese, R., and Zechmeister-Boltenstern, S.: Nitrous oxide emissions from soils: how well do we understand the processes and their controls? Philos. T. Roy. Soc. B 368, doi: 10.1098/rstb.2013.0122, 2013.

Castaldi, S., de Grandcourt, A., Rasile, A., Skiba, U., and Valentini, R.: CO_2, CH_4 and N_2O fluxes from soil of a burned grassland in Central Africa. Biogeosciences 7, 3459–3471, 2010.

Chapuis-Lardy, L., Wrage, N., Metay, A., Chotte, J.L., and Bernoux, M.: Soils, a sink for N_2O? A review. Glob. Change Biol. 13, 1–17, 2007.

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 CH_4 fluxes in a laboratory experiment. Plant Soil 343, 171–185, 2011.

Christensen, S.: Nitrous oxide emission from a soil under permanent grass: Seasonal and diurnal fluctuations as influenced by manuring and fertilization. Soil Biol. Biochem. 15(5), 531–536, 1983.

Cowan, N. J., Famulari, D., Levy, P. E., Anderson, M., Bell, M. J., Rees, R. M., Reay, D. S., and Skiba, U. M.: An improved method for measuring soil N_2O fluxes using a quantum cascade laser with a dynamic chamber. Eur. J. Soil Sci., 65(5), 643–652, doi:10.1111/ejss.12168, 2014a.

Cowan, N. J., Famulari, D., Levy, P. E., Anderson, M., Reay, D. S. and Skiba, U. M.: Investigating uptake of N_2O in agricultural soils using a high-precision dynamic chamber method. Atmospheric Meas. Tech., 7(12), 4455–4462, doi:10.5194/amt-7-4455-2014, 2014b.

Dannenmann, M., Gasche, R., Ledebuhr, A., Holst, T., Mayer, H., and Papen, H.: The effect of forest management on trace gas exchange at the pedosphere-atmosphere interface in beech (Fagus sylvatica L.) forests stocking on calcareous soils. Eur. J. Forest Res. 126, 331–346, 2007.

Denmead, O.T., MacDonald, B.C.T., Bryant, G., Naylor, T., Wilson, S., Griffith, D.W.T., Wang, W.J., Salter, B., White, I., and Moody, P.W.: Emissions of methane and nitrous oxide from Australian sugarcane soils. Agric. For. Meteorol. 150, 748–756, 2010.

Donoso, L., Santana, R., and Sanhueza, E.: Seasonal variation in N_2O fluxes at a tropical savanna site: soil consumption of N_2O during the dry season. Geophys. Res. Lett. 20, 1379–1382, 1993.

Drösler, M., 2005. Trace gas exchange and climatic relevance of bog ecosystems, Southern Germany. Doctoral thesis, TU München, 1–182 pp.

Du, R., Lu, D., and Wang, G.: Diurnal, seasonal, and inter-annual variations of N_2O fluxes from native semi-arid grassland soils of Inner Mongolia. Soil Biol. Biochem. 38, 3474–3482, 2006.

Firestone, M.K. and Davidson, E.A.: Microbiological basis of NO and N_2O production and consumption in soil. In: Andreae, M.O., Schimel, D.S., Eds. Exchange of trace gases between terrestrial ecosystems and the atmosphere. Chichester: John Wiley and Sons Ltd. pp 7–21, 1989.

Flechard, C.R., Ambus, P., Skiba, U., Rees, R.M., Hensen, A., van Amstel, A., van den Pol-van Dasselaar, A., Soussana, J.-F., Jones, M., Clifton-Brown, J., Raschi, A., Horvath, L., Neftel, A., Jocher, M., Ammann, C., Leifeld, J., Fuhrer, J., Calanca, P., Thalman, E., Pilegaard, K., Di Marco, C., Campbell, C., Nemitz, E., Hargreaves, K.J., Levy, P.E., Ball, B.C., Jones, S.K., van de Bulk, W.C.M., Groot, T., Blom, M., Domingues, R., Kasper, G., Allard, V., Ceschia, E.,

Formatiert: Tiefgestellt

Formatiert: Tiefgestellt

Cellier, P., Laville, P., Henault, C., Bizouard, F., Abdalla, M., Williams, M., Baronti, S., Berretti, F., Grosz, B.: Effects of climate and management intensity on nitrous oxide emissions in grassland systems across Europe, Agric. Ecosyst. Environ., 121, 135–152, 2007.

Flessa, H. Ruser, R., Schilling, R., Loftfield, N., Munch, J.C., Kaiser, E.A., and Beese, F.: N_2O and CH_4 fluxes in potato fields: automated measurement, management effects and temporal variation. Geoderma 105, 307–325, 2002.

Forbrich, I., Kutzbach, L., Hormann, A., and Wilmking, M.: A comparison of linear and exponential regression for estimating diffusive CH_4 fluxes by closed-chambers in peatlands. Soil Biol. Biochem. 42, 507–515, 2010.

Hensen, A., Groot, T.T., van den Bulk, W.C.M., Vermeulen, A.T., Olesen, J.E., Schelde, K.: Dairy farm CH_4 and N_2O emissions, from one square metre to the full farm scale. Agr. Ecosyst. Environ. 112, 146–152, 2006.

Horii, C.V., Munger, J.W., and Wofsy, S.C.: Fluxes of nitrogen oxides over a temperate deciduous forest. J. Geophys. Res. 109, D08305, doi:10.1029/2003JD004326, 2004.

Huber, P.J.: Robust Statistics. J. Wiley, New York, 1981.

Hutchinson, G.L. and Mosier, A.R.: Improved soil cover method for field measurement of nitrous oxide fluxes. Soil Sci. Soc. Am. J. 45, 311–316, 1981.

Hutchinson, G.L., Livingston, G.P., Healy, R.W., and Striegl., R.G.: Chamber measurement of surface–atmosphere trace gas exchange: Dependence on soil, interfacial layer, and source/sink properties. J. Geophys. Res. 105, 8865–8875, 2000.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In: Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Jassal, R.S., Black, T.A., Chen, B., Roy, R., Nesic, Z., Spittlehouse, D.L., and Trofymow, J.A.: N_2O emissions and carbon sequestration in a nitrogen-fertilized Douglas fir stand. J. Geophys. Res. 113, G04013, DOI: 10.1029/2008JG000764, 2008.

Jassal, R.S., Black, T.A., Roy, R., and Ethier, G.: Effect of nitrogen fertilization on CH_4 and N_2O fluxes, and bole and soil respiration. Geoderma 162, 182–186, 2011.

Jones, S.K., Famulari, D., Di Marco, C.F., Nemitz, E., Skiba, U.M., Rees, R.M., and Sutton, M.A.: Nitrous oxide emissions from managed grassland: a comparison of eddy covariance and static chamber measurements. Atmos. Meas. Tech. 4, 2179–2194, 2011.

Jury, W.A., Letey, J., and Collins, T.: Analysis of chamber methods used for measuring nitrous oxide production in the field. Soil Sci. Soc. Am. J. 46, 250–256, 1982.

Formatiert: Englisch (Großbritannien)

Formatiert: Tiefgestellt

Formatiert: Tiefgestellt

Kroon, P.S., Hensen, A., van den Bulk, W.C.M., Jongejan, P.A.C., and Vermeulen, A.T.: The importance of reducing the systematic error due to non-linearity in N_2O flux measurements by static chambers. Nutr. Cycl. Agroecosys. 82, 175–186, 2008.

Kroon, P.S., Schuitmaker, A., Jonker, H.J.J., Tummers, M.J., Hensen, A., and Bosveld, F.C.: An evaluation by laser Doppler anemometry of the correction algorithm based on Kaimal cospectra for high frequency losses of EC flux measurements of CH_4 and N_2O. Agric. For. Meteorol. 150, 794–805, 2010.

Kutsch, W.L., Aubinet, M., Buchmann, N., Smith, P., Osborne, B., Eugster, W., Wattenbach, M., Schrumpf, M., Schulze, E.D., Tomelleri, E., Ceschia, E., Bernhofer, C., Beziat, P., Carrara, A., Di Tommasi, P., Grünwald, T., Jones, M., Magliulo, V., Marloie, O., Moureaux, C., Olioso, A., Sanz, M.J., Saunders, M., Sogaard, H., and Ziegler, W.: The net biome production of full crop rotations in Europe. Agr. Ecosyst. Environ. 139, 336–345, 2010.

Kutzbach, L., Schneider, J., Sachs, T., Giebels, M., Nykänen, H., Shurpali, N. J., Martikainen, P. J., Alm, J., and Wilmking, M.: CO_2 flux determination by closed-chamber methods can be seriously biased by inappropriate application of linear regression. Biogeosciences 4, 1005–1025, 2007.

Laville, P., Lehuger, S., Loubet, B., Chaumartin, F., and Cellier, P.: Effect of management, climate and soil conditions on N_2O and NO emissions from an arable crop rotation using high temporal resolution measurements. Agric. For. Meteorol. 151, 228–240, 2011.

Leiber-Sauheitl, K., Fuß, R., Voigt, C., and Freibauer, A.: High CO_2 fluxes from grassland on histic Gleysol along soil carbon and drainage gradients. Biogeosciences 11, 749–761, 2014.

Livesley, S.J., Grover, S., Hutley, L.B., Jamali, H., Butterbach-Bahl, K., Fest, B., Beringer, J., and Arndt, S.: Seasonal variation and fire effects on CH_4, N_2O and CO_2 exchange in savanna soils of northern Australia. Agric. For. Meteorol. 151, 1440–1452, 2011.

Livingston, G.P., Hutchinson, G.L., and Spartalian, K.: Trace gas emission in chambers: a non-steady-state diffusion model. Soil Sci. Soc. Am. J. 70, 1459–1469, 2006.

Loftfield, N., Flessa, H., Augustin, J., and Beese, F.: Automated gas chromatographic system for rapid analysis of the trace gases methane, carbon dioxide, and nitrous oxide. J. Environ. Qual. 26, 560–564, 1997.

Lohila, A., Aurela, M., Hatakka, J., Pihlatie, M., Minkkinen, K., Penttil, T., and Laurila, T.: Responses of N_2O fluxes to temperature, water table and N deposition in a northern boreal fen. Eur. J. Soil Sci. 61, 651–661, 2010.

Matthias, A.D., Yarger, D.N., and Weinbeck, R.S.: A numerical evaluation of chamber methods for determining gas fluxes. Geophys. Res. Lett. 5, 765–768, 1978.

Merbold, L., Eugster, W., Stieger, J., Zahniser, M., Nelson, D. D., and Buchmann, N.: Greenhouse gas budget (CO_2, CH_4 and N_2O) of intensively managed grassland following restoration. Glob. Change Biol. 20, 1913–1928, 2014.

Formatiert: Tiefgestellt

Moffat, A.M.: A new methodology to interpret high resolution measurements of net carbon fluxes between terrestrial ecosystems and the atmosphere. Doctoral Thesis, Department of Mathematics and Computer Science, Friedrich Schiller University, Jena, 2012.

Neftel, A., Ammann, C., Fischer, C., Spirig, C., Conen, F., Emmenegger, L., Tuzson, B., and Wahlen, S.: N_2O exchange over managed grassland: Application of a quantum cascade laser spectrometer for micrometeorological flux measurements. Agric. For. Meteorol. 150, 775–785, 2010.

Nakano, T., Sawamoto, T., Morishita, T., Inoue, G., and Hatano, R.: A comparison of regression methods for estimating soil–atmosphere diffusion gas fluxes by a closed-chamber technique. Soil Biol. Biochem. 36, 107–113, 2004.

Nelson, D.D., McManus, B., Urbanski, S., Herndon, S., and Zahniser, M.S.: High precision measurements of atmospheric nitrous oxide and methane using thermoelectrically cooled mid-infrared quantum cascade lasers and detectors. Spectrochim. Acta A 60, 3325–3335, 2004.

Papen, H. and Butterbach-Bahl, K.: A 3-year continuous record of nitrogen trace gas fluxes from untreated and limed soil of a N-saturated spruce and beech forest ecosystem in Germany, 1. N_2O emissions. J. Geophys. Res. 104, 18487–18503, 1999.

Parkin, T.B. and Kaspar, T.C.: Nitrous Oxide Emissions from Corn–Soybean Systems in the Midwest. J. Environ. Qual. 35, 1496–1506, 2006.

Parkin, T.B. and Venterea, R.T.: Chamber-based trace gas flux measurements. USDA-ARS GRACEnet Project Protocols Chapter 3. Web access: http://www.ars.usda.gov/SP2UserFiles/Program/212/Chapter%203.%20GRACEnet%20Trace%20Gas%20Sampling%20Protocols.pdf, 2010.

Pedersen, A.R., Petersen, S.O., and Schelde, K.: A comprehensive approach to soil-atmosphere trace-gas flux estimation with static chambers. Eur. J. Soil Sci. 61, 888–902, 2010.

Petersen, S.O., Mutegi, J.K., Hansen, E.M., and Munkholm, L.J.: Tillage effects on N_2O emissions as influenced by a winter cover crop. Soil Biol. Biochem. 43, 1509–1517, 2011.

Pihlatie, M.K., Christiansen, J.R., Aaltonen, H., Korhonen, J.F., 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., Serca, D., Vicca, S., Wolf, B., and Pumpanen, J.: Comparison of static chambers to measure CH_4 emissions from soils. Agric. For. Meteorol. 171, 124–136, 2013.

R Core Team: R: A language and environment for statistical computing, available at: http://www.R-project.org/, 2012.

Rinne, J., Pihlatie, M., Lohila, A., Thum, T., Aurela, M., Tuovinen, J., Laurila, T., and Vesala, T.: Nitrous oxide emissions from a municipal landfill. Environ. Sci. Technol. 39, 7790–7793, 2005.

Rolston, D.E.: Gas diffusivity. In A. Klute (ed.) Methods of soil analysis. Part 1. Physical and mineralogical methods. Agron. Monogr. 9., p. 1089–1102. 2nd ed. ASA and SSSA, Madison, WI, 1986.

Rosenkranz, P., Brüggemann, N., Papen, H., Xu, Z., Horváth, L., and Butterbach-Bahl, K.: Soil N and C trace gas fluxes and microbial soil N turnover in a sessile oak (Quercus petraea (Matt.) Liebl.) forest in Hungary. Plant Soil 286, 301–22, 2006.

Rothman, L.S., Gordon, I.E., Barbe, A., Benner, D.C., Bernath, P.F., Birk, M., Boudon, V., Brown, L.R., Campargue, A., Champion, J.-P., Chance, K., Coudert, L.H., Dana, V., Devi, V.M., Fally, S., Flaud, J.-M., Gamache, R.R., Goldman, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W., Mandin, J.-Y., Massie, S.T., Mikhailenko, S.N., Miller, C.E., Moazzen-Ahmadi, N., Naumenko, O.V., Nikitin, A.V., Orphal, J., Perevalov, V.I., Perrin, A., Predoi-Cross, A., Rinsland, C.P., Rotger, M., Simeckova, M., Smith, M.A.H., Sung, K., Tashkun, S.A., Tennyson, J., Toth, R.A., Vandaele, A.C., and Vander Auwera J.: The HITRAN 2008 molecular spectroscopic database. J. Quant. Spectrosc. Ra. 110, 533–572, 2009.

Sakabe, A., Kosugi, Y., Takahashi, K., Itoh, M., Kanazawa, A., Makita, N., Ataka, M.: One year of continuous measurements of soil CH_4 and CO_2 fluxes in a Japanese cypress forest: Temporal and spatial variations associated with Asian monsoon rainfall, J. Geophys. Res. – Biogeosciences, 120(4), 585–599, 2015.

Savage, K., Phillips, R., and Davidson, E.: High temporal frequency measurements of greenhouse gas emissions from soils. Biogeosciences 11, 2709–2720, 2014.

Schiller, C.L. and Hastie, D.R.: Nitrous oxide and methane fluxes from perturbed and unperturbed boreal forest sites in northern Ontario. J. Geophys. Res. 101, 22,767– 22,774, 1996.

Schlesinger, W. H.: An estimate of the global sink for nitrous oxide in soils. Glob. Change Biol. 19(10), 2929–2931, 2013.

Seinfeld, J.H. and Pandis, S.N.: Atmospheric Chemistry and Physics – From Air Pollution to Climate Change. 2nd ed. John Wiley & Sons, Inc., Hoboken, New Jersey, 1232 pp, 2006.

Shurpali, N.J., Rannik, Ü., Jokinen, S., Lind, S., Biasi, C., Mammarella, I., Peltola, O., Pihlatie, M., Hyvönen, N., Räty, M., Haapanala, S., Zahniser, M., Virkajärvi, P., Vesala, T., and Martikainen, P.J.: Neglecting diurnal variations leads to uncertainties in terrestrial nitrous oxide emissions. Scientific Reports 6:25739, DOI: 10.1038/srep25739, 2016.

Skiba, U., Jones, S.K., Drewer, J., Helfter, C., Anderson, M., Dinsmore, K., McKenzie, R., Nemitz, E., Sutton, M.A.: Comparison of soil greenhouse gas fluxes from extensive and intensive grazing in a temperate maritime climate, Biogeosciences, 10, 1231–1241, doi:10.5194/bg-10-1231-2013, 2013.

Thomson, A.J., Giannopoulos, G., Pretty, J., Baggs, E.M., and Richardson, D.J.: Biological sources and sinks of nitrous oxide and strategies to mitigate emissions. Philos. T. Roy. Soc. B 367, 1157–1168, 2012.

Tuzson, B., Hiller, R.V., Zeyer, K., Eugster, W., Neftel, A., Ammann, C., and Emmenegger, L.: Field intercomparison of two optical analyzers for CH_4 eddy covariance flux measurements. Atmos. Meas. Tech. 3, 1519–1531, 2010.

Wrage, N., Velthof, G.L., van Beusichem, M.L., and Oenema, O.: Role of nitrifier denitrification in the production of nitrous oxide. Soil Biol. Biochem. 33(12), 1723–1732, 2001.

Formatiert: Tiefgestellt

Formatiert: Tiefgestellt

Table 1. Supplementary information on field campaigns.

	Braunschweig	**Risø**
Coordinates	52°17‘52‘‘N, 10°26‘36‘‘~~O~~E	55°40‘50‘‘N, 12°06‘05‘‘~~O~~E
Start observation period	Nov 13, 2012	Apr 10, 2013
End observation period	Dec 12, 2012	Apr 24, 2013
Total GC flux rates (n)	201	37
Total QCL flux rates (n)	187	158
Land use	Grassland	Willow field (harvested)
Fertilization, date	No fertilization	Apr 9, 2013
Fertilization, amount	No fertilization	120 kg N ha^{-1}
Fertilization, type	No fertilization	Mineral (ammonium nitrate), N-P-K 21-3-10
Soil texture	Silty sand	Sandy loam
Soil type	Cambisol	Luvisol

Table 2: Features of the chamber-analyzer system used in this study.

	GC* (model: Shimadzu GC-2014)	QCL* (model: Aerodyne Research Inc. mini-QCLAS)
No. of chambers	3	3
Chamber closure time	60 min	60 min 10 min (recommended)
Sampling frequency	every 20 min	0.1 sec (max) 5 sec (recommended)
No. of concentration records per chamber run	4	36000 in 60 min 6000 in 10 min
No. of chamber cycles per day	24 (max)	72 (recommended) 144 (max)
Maximum number of samples	168 (depending on autosampler size)	Limited only by data storage capacity of QCL's computer or external hard drive
Lag time	(~10 sec)	~10 sec
N_2O flux detection limit ($\mu g\ N\ m^{-2}\ h^{-1}$)	13.0	2.6
Mean campaign N_2O flux ($\mu g\ N\ m^{-2}\ h^{-1}$)	BS (pref.[1]): 6.42 Risø (pref.[1]): 77.40	BS (lin.): 7.77 Risø (lin.[2]): 122.95
Mean campaign SE of N_2O fluxes ($\mu g\ N\ m^{-2}\ h^{-1}$)	BS (pref.[1]): 5.98 Risø (pref.[1]): 8.17	BS (lin.): 0.13 Risø (lin.[2]): 0.21
Median campaign N_2O flux ($\mu g\ N\ m^{-2}\ h^{-1}$)	BS (pref.[1]): 5.15 Risø (pref.[1]): 64.80	BS (lin.): 7.38 Risø (lin.[2]): 105.43
Median campaign SE of N_2O fluxes ($\mu g\ N\ m^{-2}\ h^{-1}$)	BS (pref.[1]): 5.04 Risø (pref.[1]): 4.72	BS (lin.): 0.10 Risø (lin.[2]): 0.17
Percentage of flux estimates where HMR could be fitted	BS: 8.5 % Risø: 37.9 %	BS: 100 % Risø: 100 %

GC – Gas chromatograph, QCL – Quantum cascade laser spectrometer, [1]preferred means non-linear HMR model was used if applicable, otherwise robust linear regression was taken, [2]mean/median of DOY 105.5 to 108.5 to make it comparable to GC data set

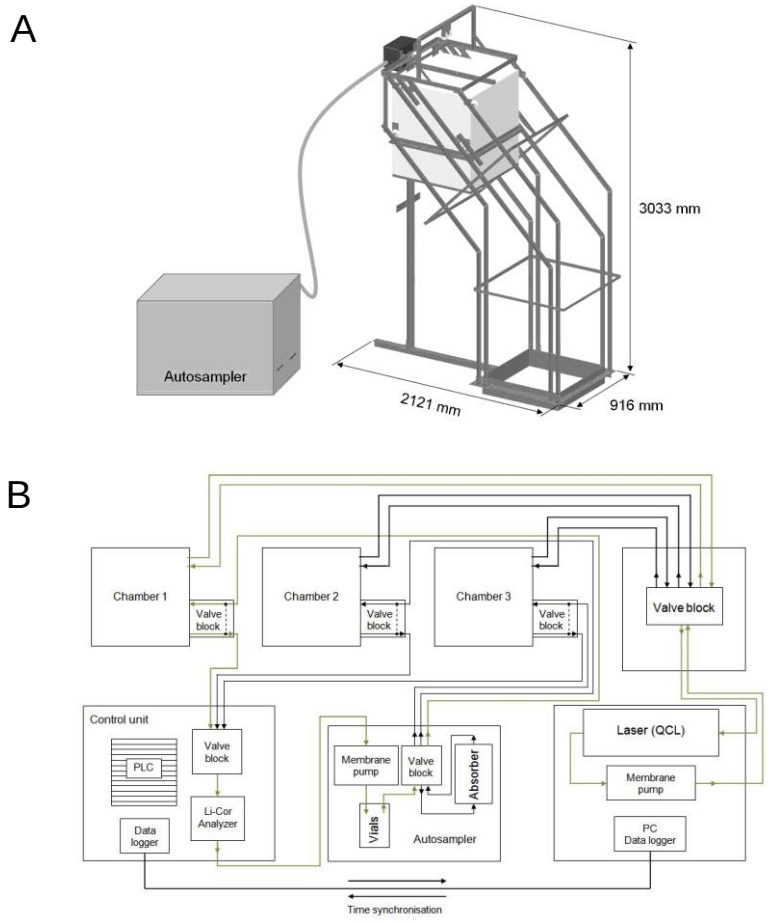
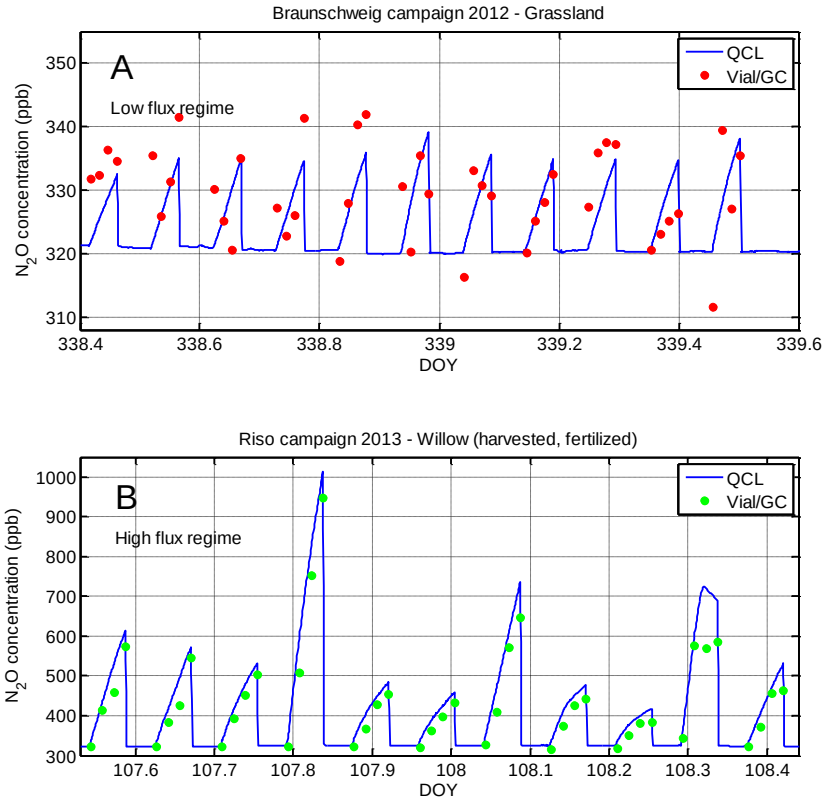


Figure 1. Schematic diagrams of an automated chamber connected to an autosampler unit (A) and of the entire chamber system (B). Green lines indicate that Chamber 1 is currently in measurement mode. See text for detailed description.



5 **Figure 2. Examples of time series of N_2O chamber concentrations during the Braunschweig (~~upper panel~~Panel A) and Risø campaign (~~lower panel~~Panel B). Chambers were periodically closed for 60 minutes. Vials were filled with sample air at t_0, t_{20}, t_{40}, and t_{60}. The QCL system was operated at a sampling frequency of 10 Hz; plotted are 1-min means.**

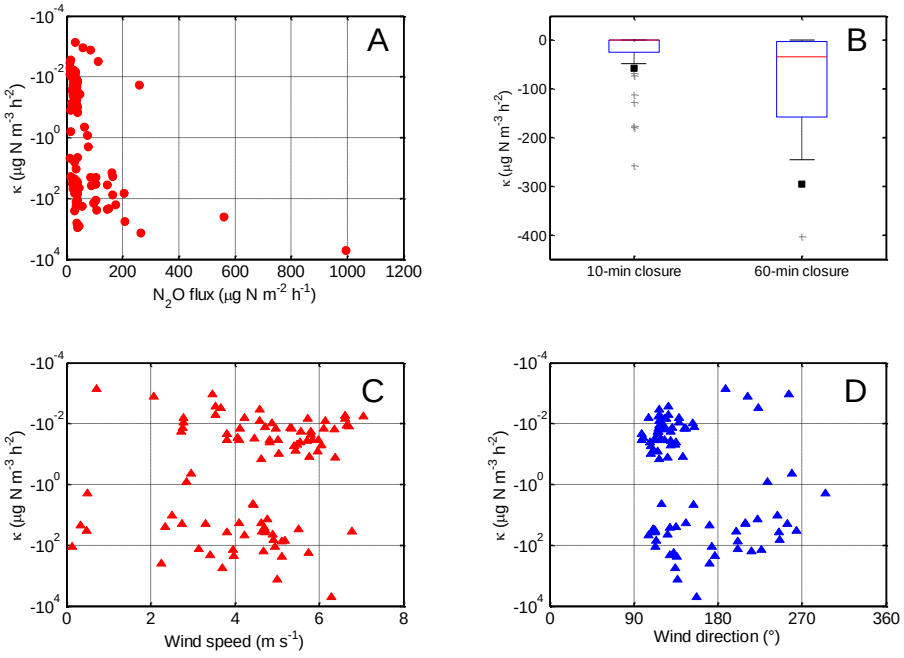


Figure 3. Panels A: Relationship between κ and N_2O fluxes calculated with an exponential model (see text for details). The parameter κ indicates the curvature, i.e. the second derivative of the exponential model used for flux calculation. Negative κ values correspond to concave functions, i.e. plateauing (saturating) N_2O concentration increases (*cf.* Figure 2). Panel B: Box plot of κ values showing the difference between 10 and 60-min closure where black squares represent the arithmetic mean, red horizontal lines indicate the median, blue horizontal lines indicate lower and upper quartile values, black whiskers represent the interquartile range and outliers from this range are plotted as grey crosses. To ensure better readability, the y-axis is truncated at –450 µg N m^{-3} h^{-2}. Thus, some outliers between –450 and -10^4 µg N m^{-3} h^{-2} are not shown. Panel C and D: Relationship between κ and wind speed (C) as well as κ and wind direction (D). All data are taken from the quantum cascade laser system operated during the Risø campaign. Chambers were closed for 10 minutes at DOY <105.5 and for 60 minutes at DOY >105.5.

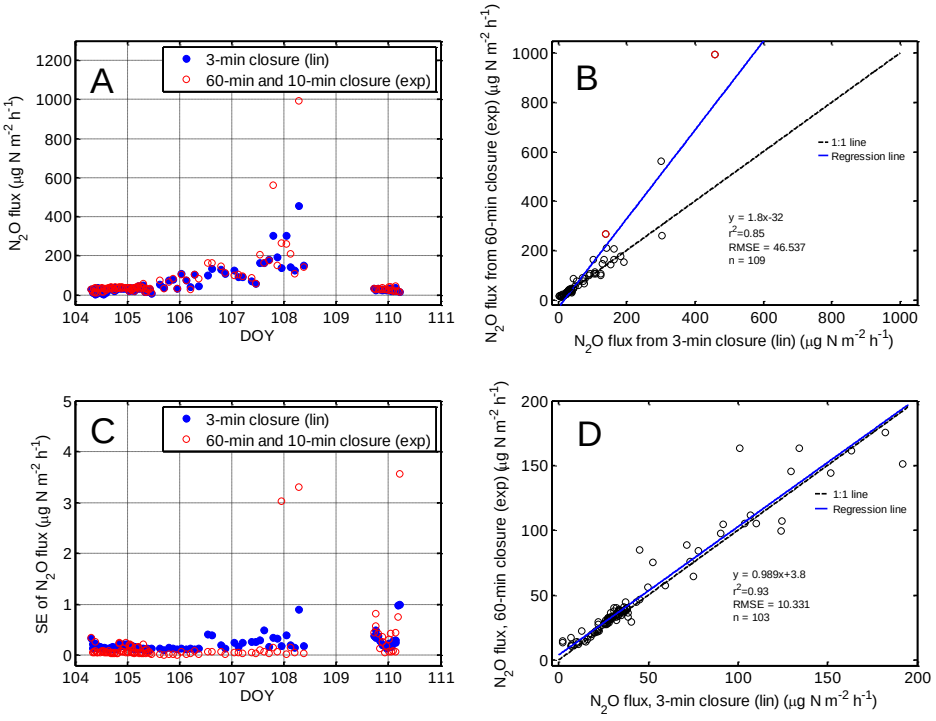


Figure 4. ~~Comparison of N_2O fluxes measured during the Risø campaign by the QCL system based on different calculation methods, i.e. a linear model using only the first three minutes of data after chamber closure (filled blue circles) and an exponential model (open red circles) (see text) using either the full 60 minutes (DOY 105.5 to DOY 108.5) or the full 10 minutes of data (DOY <105.5 and DOY >108.5) (Panel A) with associated standard errors (SE) of N_2O fluxes (Panel C). Also shown is a linear regression analysis (orthogonal) of N_2O fluxes calculated by the exponential model (60 and 10 minutes of data) vs. N_2O fluxes calculated by the linear model (3 minutes of data) (Panel B) and a box plot of associated standard errors of N_2O fluxes (Panel D) showing the difference between 3 and 60 min closure. Black squares represent the arithmetic mean, red horizontal lines indicate the median, blue horizontal lines indicate lower and upper quartile values, black whiskers represent the interquartile range and outliers from this range are plotted as grey crosses. To ensure better readability, the y-axis is truncated at +0.7 µg N m^{-2} h^{-1}. Thus, few outliers~~

~~between 0.7 and 4 µg N m^{-2} h^{-1} are not shown.~~ Panel A: Comparison of N_2O fluxes measured on a harvested willow field during the Risø campaign by the QCL system based on a linear model using only the first three minutes of data after chamber closure (filled blue circles) and an exponential model (open red circles) (see text) using either the full 60 minutes (DOY 105.5 to DOY 108.5) or the full 10 minutes of data (DOY <105.5 and DOY >108.5). Panel B: Linear regression analysis of N_2O fluxes from the exponential vs. the linear model. Panel C: Standard errors of fluxes shown in Panel A. Panel D: Same as Panel B, but only for fluxes <200 µg N m^{-2} h^{-1} with adapted regression.

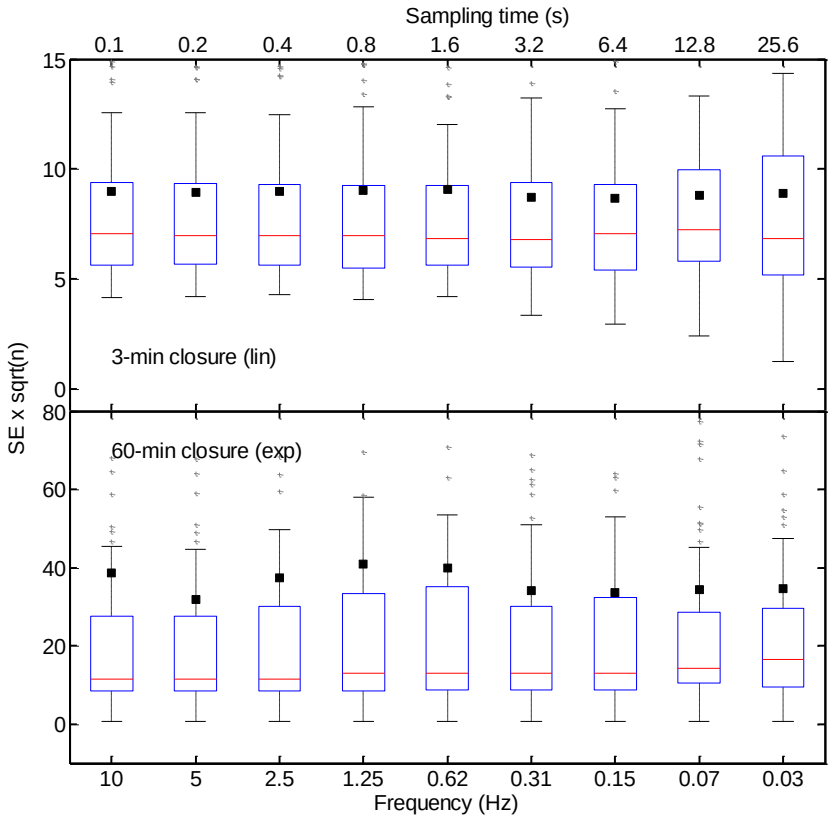


Figure 5. ~~Dependency of standard errors of N_2O fluxes on measurement frequency when using only the first 3 minutes of data after chamber closure (linear model; upper panel) and the full 60-minute dataset (exponential model; lower panel). Black squares represent the arithmetic mean, red horizontal lines indicate the median, blue horizontal lines indicate lower and upper quartile values, black whiskers represent the interquartile range and outliers from this range are plotted as grey crosses. For better readability, the y-axis is truncated at +6 µg N m^{-2} h^{-1}. Thus, few outliers are not shown.~~ Boxplots of standard errors of N_2O fluxes for different frequency classes and regression models used, i.e. linear regression with 3 minutes of data (upper panel) and the

exponential HMR model with 60 minutes of data (lower panel). To avoid a pseudo-dependency on sample size, the standard errors were normalized by multiplication with $\sqrt{n}$. Black squares represent the arithmetic mean, red horizontal lines indicate the median, blue horizontal lines indicate lower and upper quartile values, black whiskers represent the interquartile range and outliers from this range are plotted as grey crosses.

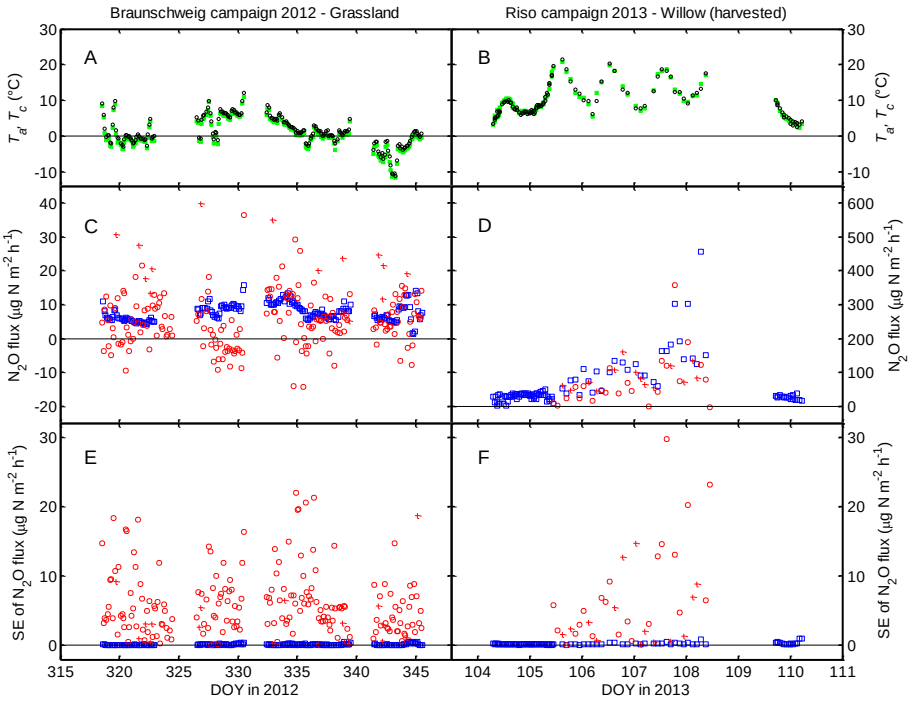
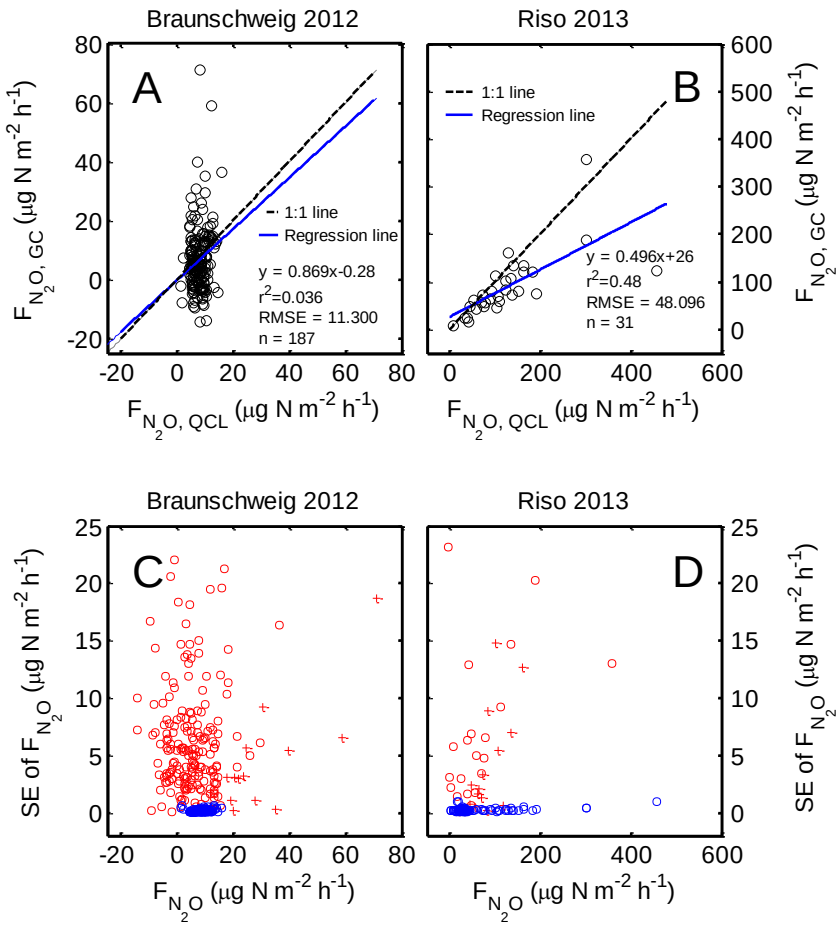


Figure 6. Time series of air (T_a, green markers) and chamber temperatures (T_c, black markers)(panels A and B), N_2O fluxes and the respective standard errors of N_2O fluxes during the Braunschweig (panels C and E) and the Risø campaign (panels D and F). Blue markers indicate QCL data, red markers indicate GC data. Crosses are plotted for GC data when all criteria for flux calculation using the exponential HMR model were met (see text for details), otherwise circles are plotted indicating the usage of a linear model for flux calculation.



5 **Figure 7.** Panels A and B: GC vs. QCL-based N_2O fluxes. Panels C and D: Relationships between standard errors (SE) of N_2O fluxes and the respective flux values. Blue markers indicate QCL data, which are all based on the 3-min linear calculation method.

Red markers indicate GC data, which are based on the full 60-min data set. Crosses are plotted for GC data when all criteria for flux calculation using the exponential HMR model were met (see text for details), otherwise circles are plotted indicating the usage of a linear model for flux calculation. ~~**GC vs. QCL-based N_2O fluxes (upper panels) and relationships between standard errors (SE) of N_2O fluxes and the respective flux values (lower panels) during the Braunschweig (panels on left-hand side) and the Risø campaign (panels on the right-hand side). Blue markers indicate QCL data, red markers indicate GC data. Crosses are plotted for GC data when all criteria for flux calculation using the exponential HMR model were met (see text for details), otherwise circles are plotted indicating the usage of a linear model for flux calculation.**~~

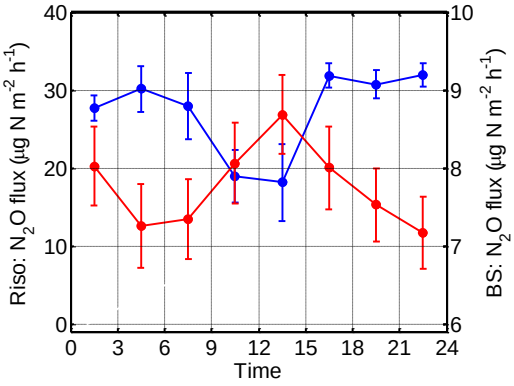


Figure 8. Mean diurnal courses of N_2O fluxes derived from QCL flux measurements during the Risø (blue line) and Braunschweig (red line) campaign. To exclude fertilization effects in Risø, only data from the low flux period (DOY<105.5 and >108.5) were taken.