The referees' comments were addressed point-by-point in the author comments published online October 31, 2014. Below we only summarise how we have modified the manuscript. Note that for clarity we have not retained deletions (corresponding to parts of text moved or deleted entirely) but only the additions can be followed in the revised marked-up manuscript version (bg-2014-277-manuscript-marked-up-28-11-2014.pdf) referred to below. We believe that it can be well followed how we have addressed and modified the MS in response to referees' comments.

With the help of numerous referee comments we have significantly revised the MS, paying special attention to section 4 (Discussion) and added a separate section for conclusions. We believe that the manuscript has much improved and is suitable for publication in BG.

Sincerely, Üllar Rannik

Anonymous Referee #1
Received and published: 13 September 2014

General comments:
Rannik et al write an interesting report on the performance of four different fast response N2O analysers in an eddy covariance field campaign over a several months'period. The instruments differed mainly in the sensor noise; the older makes had higher noise than the newer ones, which is good news, as it documents satisfying metrological progress. From their results The Authors defined the flux detection limit, which was in some cases higher than the flux level in usual conditions, i.e. when fertilisation didn't affect the N2O emissions anymore. Interestingly there were periods where the instruments measured very much the same, while in two periods systematic differences were found between two of the sensors and the others. For the latter the Authors speculated over possible reasons, but were not able to explain the differences.

The possible reasons for systematic differences in fluxes were analysed and discussed in revised section 4 Discussion.

The article is mostly easy to understand and the field data is very well suited given the objectives of the work. In its current form the manuscript stays a bit too technical, more like a report. Very good work, but, scientifically, not yet exceptionally inspiring. The reason is that the report stays rather descriptive. There are indeed some interesting, advanced, quantitative analyses that go beyond standard - I mean here the method, how the individual N2O sensors contributed to the flux noise - but this is mainly presented but not further discussed. There is little reference to other work, beyond methodology. Some conclusions are mentioned in the discussion, but not summarised. Below I will make some suggestions, on how this work can be improved. Finally, what good scientific yield essentially is, can be debated, and I leave it to the Editor to explain this further in this open discussion.

We have significantly amended the discussion section, adding relevant references to literature to compare with. Also, we have included a separate section 5 with conclusions.

Some details:

The site description:

Please mention the amount and type of fertiliser.

Added to Section 2.1 (p. 4, l. 11-13).

The measurement setup:

Reading the manuscript, I have got the impression that the experimental setup was generally chosen with great care. I was, however, surprised about the very low measurement height. If I'm right, the reed canarygrass canopy grew very quickly from 0.1 to 1.7 m and then more slowly up to 1.9 m. The measurement height was in the beginning 2.2 m and was raised by only 20 cm when the canopy reached 1.7 m, i.e. to only 50 cm above the canopy. If this is not a typographical error and rather, e.g., 3.4 meter was meant, this would be extremely narrow and should result in the disturbance of the spectra in the high-frequency range. This could, e.g., be the explanation for the observed high frequency noise in sonic temperature power spectra (Figs. 2 and 4). It would be interesting to think further about how this disturbance interferes with the analysis of flux and sensor noise.

We responded in detail to this question in our author comment. A paragraph was added into section 4 Discussion explaining our best understanding in this matter (p. 19 last paragraph – p. 20).

It was mentioned in the text that the vertical wind speed component (w) was expected not to be affected by noise. I suggest analysing the noise in was well to show this. And more generally, as total random noise in fluxes and that due to analyser noise is compared in the study, it would probably mean a difference, whether or not the turbulence is disturbed.

We demonstrated in our author comment that the anemometer's noise had much less impact on flux error than the analysers' noise. This is discussed now in section 4 of the revised MS (p. 22, l. 4-12).

Although it is said that the two involved universities were measuring with two separate EC systems (page 11751, line 18) the text never refers to the UEF system, i.e. its sonic anemometer. I guess this means that all 4 analysers were operated with one single sonic anemometer (the UH one). This makes perfect sense for avoiding another source for differences between the sensors. In this case the above mentioned sentence should be rephrased accordingly to not confuse the reader.

There were two sonic anemometers. Clarifying wording was added in the beginning of section 2.2 (p. 4, l. 26-27).

And, a bit more specific, what was the reason to run the LGR_CW_QCL with such a low flow rate? To my knowledge the supplier does recommend higher flow rates and a lower operating pressure.

The fact is stated in the section 2.2.

Finally, please mention how and when were the sensors were calibrated (see below).

The sensor calibration information was amended in section 2.2 (p. 5)

The data set:
The field data base is very well suited for such analysis. It is relatively long and covers periods with higher and lower flux levels. I think that the Authors have done a very good job when dealing with the limited availability of some of the sensors in some parts of the measurement period and also with the choice to avoid gap filling but confining the analysis only to data where all sensors were producing reasonable data. It became not entirely clear to me, whether the data quality criteria were also used, when the data were selected. In any case it would be desirable to only use the 'good' quality data here.

This information was moved from the beginning of section 3 into a separate (new in the revised MS) section 2.6 (p. 11).

What I was missing is any reference to the meteorological and soil conditions throughout the measurement campaigns. This could be interesting for others that are also interested in the control and interpretation of the N2O fluxes. Could you give the time series of soil temperatures and water contents under the measurements, especially during the period prior to and after the fertilisation.

Figure 2 presenting the time series of soil water content, precipitation and soil temperature was added.

Scientific analysis and discussion:

To begin with: I had the impression that the discussion of the results is the weakest part of the manuscript. A separate section on the scientific conclusions is also missing to take home a message from the work. The following comments will hopefully give The Authors some inspiration to help them increasing the scientific content and impact of this manuscript.

Section 4 Discussion was amended and re-written.

1. Random noise: To opt for separating the total random flux error into the part that is solely caused by the sensor and the rest, is very much reasonable in the context of the manuscript. It's a very strong aspect of the study and I did not read very often about it. The Authors used existing approaches to calculate and partition the random flux error. They mention 'Theoretically, there are several ways to approximate the same error estimate', but I missed a critical discussion on the usefulness of these approaches and a clear explanation for their choice. What is the uncertainty of the different methods / estimates? Wouldn't it be possible to try out more than one and compare the results? This debate can create some interesting scientific discussion. Such analyses are so far rare and exemplifying the methodical aspects and critically discuss them could add value to the article.

We believe that we have improved the section 2.5 describing the random errors with respect the clarity of presentation by considering the referee comments on this topic (also below). For interested readers to estimate the random errors by different approaches we refer to Rannik et al. (2009), but detailed analysis and comparison of different methods (also with different physical meaning) to calculate the random error estimates proposed in literature is left out of scope of the MS.

2. Why is the uncertainty of the instrumental noise estimate of LGR-CW-QCL so much higher than that of the other sensors (Fig. 5a), while both the total flux noise and the one related to sensor noise are about as certain as with the other sensors? This seems a bit inconsistent to me. For this comparison, a linear scale presentation of Fig. 5 would be better suited.

We kept the logarithmic presentation of Figure 5 to accommodate wide range of values for different instruments. We did not find inconsistency in the noise estimates for different instruments, see the author response.

3. The spectral analysis is straight forward but I wonder why the Authors perform it only at such a small data basis, i.e. a few hours. Please comment. This is probably the reason, why The Authors give one time constant value for H2O. This should rather vary with relative humidity as known from earlier work, by the way, even before the work by Mammarella et al. (2009).

We have explained the approach in the beginning of section 3.1 (p. 12, l. 12-16)

4. Flux intercomparison The analysis of the observed systematic differences between the analysers is not entirely satisfactory. For many biogeochemical analyses, such as annual GHG budgets, the systematic uncertainty is even more important than the random noise. There must be many raw and ancillary data that offer more analyses to finally find the reasons for these differences. As an example the systematic differences between the AR-P-QCL and AR-CW-QCL was discussed in terms of cross-sensitivity with water vapour. One should rather proof than indirectly conclude that the cross-sensitivity of the N2O spectra with H2O (and thus the estimated N2O flux with the H2O flux) is not /or unlikely to be the cause for observed differences. Which other reasons can explain the differences?

We have included more details of the analysis of systematic differences in the revised section 4 (p. 17-19), including the possible impact of calibration and water vapour interference with N2O fluxes.

The same applies for the discussion on sensor drift effects. The methods part does not refer to the calibration of the sensors. One could easily investigate the sensor drift with calibrations that were carried out during the course of the campaign. Or, if that wasn't done, one could compare the N2O concentrations from the 4 sensors with the mean from all measurements. I'm sure this will clarify the issues of sensor drift. Maybe you will even find a sensitivity drift, which could then very easily explain the observed absolute deviations of the flux estimates from the different sensors.

We performed such analysis of instrumental sensitivity drift based on available calibration information and comparison of concentrations, but were not able to explain the observed systematic differences. The sensor calibration is described more in section 2.2. and the discussion of possible reasons for systematic differences presented in section 4 (p. 17-19).

What is the reason for the large noise of fluxes from CS-TDL, when the AR-CW-QCL does not measure any flux (Fig. 6 left panels)? In my experience the absolute flux noise level is either constant, in the best case, or increasing with flux level. A higher absolute noise level at low or zero fluxes is on the contrary rather unusual. This must have a reason; any idea?

This is explained in terms of flux errors, see section 3.3 (p. 15 l. 31 to p.16 l. 7).

5. Comparison with published work. Comparison with literature, not only regarding the methodology used, but also the scientific results, e.g., about noise levels, fluxes etc., is missing. The discussion and conclusion section does only refer to one single publication, an own one that is cited for the third time in the same context (Mammarella et al., 2010) in this manuscript. I recommend using information from existing literature to define the state of art and describe the progress in this field and highlight this work's contributions.

We have discussed our results in context of literature, see revised section 4. We have added also the paragraph with precision analysis of the chamber system measurements, placing our results in the perspective for N2O exchange measurement in more general.

Finally I'd like to recommend thinking about which general scientific conclusions can be drawn from this study and summarising them in a short, concluding, final section.

A separate section 5 summarising the main conclusions was added.

Presentation of the material:

The manuscript is with a few exceptions very easy to read and understand. Even leaving out articles does not disturb very much, it is to my knowledge just not correct English. I would also suggest replacing 'multiplication' by 'convolution' (11755, 11 and 16)

Replaced as suggested (see p. 8).

Section 2.5 is the one most difficult to understand; in fact it became first clear to me what was actually done, after I read the results part. Probably using 'was' instead of 'can be' would already help a lot page 11756, lines 12 and 23, page 11757, line 11).

We replaced wording as suggested (p. 9, 10).

(11756, 16): 'The method evaluates the error in time domain through integration of the auto-covariance and cross-covariance functions of the vertical wind speed and the scalar concentration.' – integration over what? Integration of auto-and cross covariance (w,s) – the part is missing what follows to estimate a single noise value. An equation is needed to complete the text.

The equation (3) in the revised MS was added (p. 9).

(11756, 27) sigma_noise =’the standard deviation of instrumental noise as observed at frequency f’ and (11757, 5) What does ’instrumental random noise variance’ sigma_noise then mean? Probably not the variance of the noise but the noise expressed as a variance. Not sure of you even mean variance or rather standard deviation.

Wording corrected, see 2.5 (p. 10, l. 7).

(11757, 2-4): ’the method developed by Lenschow et al. (2000) and applied to EC fluxes by Mauder et al. (2013) to estimate the flux detection limit due to instrumental noise.’ Neither the study of Lenschow et al. (2000) nor the one from Mauder et al. (2013) contains the term ’detection limit’. Please define the term detection limit and show, how it is determined by sigma_noise.

We “use flux error due to instrumental noise” or similar wording throughout MS to avoid confusion. However, we have added the explanation that the error estimates gives the smallest error that the system is able to detect, see modifications in section 2.5 (p. 10, l. 17-24)

Structure / redundancy The main structure of the manuscript is good with a few exceptions:

1. the introduction to the section 3 (Results) that belongs into the methods description
And

Moved into separate section 2.6 under section 2 Materials and methods (p. 11).

2. the repetition of the methodology and the results in the discussion parts that should be avoided.

We have tried to avoid repetition and revised the section 4.

The presentation the data in tables and graphs is very good, with the exception of Figs 2, 4 and 6, where the single panels are not given letters for reference.

The letters were inserted for reference, see Figures numbered as 3, 5 and 7 in the revised MS.

Anonymous Referee #2
Received and published: 23 September 2014

P11753 Could the authors add some information (one sentence) about the calibration of all instruments, not only for the IRGA and pulsed QCL?

We have amended the section 2.2 with calibration information (p. 5). In addition, we have revised section 4 including discussion on the possible implication of calibration changes on the systematic differences in fluxes (p. 17-18).

P11759 L11 “the” is missing before method

Inserted.

P11759 L16-20 the sentence is not entirely clear. “higher values” refers to what?

The sentence was modified, see modifications in section 3.1 (p. 13, l. 6-8).

P11760 L8 “the” before expression

Inserted.

P11760 L18 I guess this is for H2O: could you repeat it in the text?

Yes for H2O, repeated for clarity in the revised MS (p. 14, l. 3).

P11760 L19 lines (plural)

Corrected.

P11760 L24 delete “frequency”

Deleted.

P11760 L26 add “the” before sampling.

Added.

Intercomparison of fast response commercial gas analysers for nitrous oxide flux measurements under field conditions

Ü. Rannik[1], S. Haapanala[1], N. J. Shurpali[2], I. Mammarella[1], S. Lind[2], N. Hyvönen[2], O. Peltola[1], M. Zahniser[3], P. J. Martikainen[2], and T. Vesala[1]

[1]Department of Physics, P.O. Box 48, FIN-00014 University of Helsinki, Finland

[2]Department of Environmental Science, University of Eastern Finland, Kuopio, Finland

[3]Center for Atmospheric and Environmental Chemistry, Aerodyne Research Inc., Billerica MA, USA

Correspondence to: Ü. Rannik (ullar.rannik@heuristica.ee)

Abstract

Four gas analysers capable of measuring nitrous oxide (N_2O) concentration at a response time necessary for eddy covariance flux measurements were operated from spring till winter 2011 over a field cultivated with reed canary grass (RCG, *Phalaris arundinaceae*, L.), a perennial bioenergy crop in Eastern Finland. The instruments were TGA100A (Campbell Scientific Inc.), CW-TILDAS-CS (Aerodyne Research Inc.), N2O/CO-23d (Los Gatos Research Inc.) and QC-TILDAS-76-CS (Aerodyne Research Inc.). The period with high emission, lasting for about two weeks after fertilization in late May, was characterised by an up to two orders of magnitude higher emission, whereas during the rest of the campaign the N_2O fluxes were small, from 0.01 to 1 nmol m^{-2} s^{-1}. Two instruments, CW-TILDAS-CS and N2O/CO-23d, determined the N_2O exchange with minor systematic difference throughout the campaign, when operated simultaneously. TGA100A produced cumulatively highest N_2O estimates (with 29% higher value during the period when all instruments were operational). QC-TILDAS-76-CS obtained 36% lower fluxes than CW-TILDAS-CS during the first period, including the emission episode, whereas the correspondence with other instruments during the rest of the campaign was good. The reasons for systematic differences were not identified, suggesting further need for detailed evaluation of instrument performance under field conditions with emphasis on stability, calibration and any other factors that can affect systematically the accuracy of flux measurements. The instrument CW-TILDAS-CS was characterised by the lowest noise level (with a standard deviation of around 0.12 ppb at 10 Hz

sampling rate), as compared to N2O/CO-23d and QC-TILDAS-76-CS (around 0.50 ppb) and TGA100A (around 2 ppb). We identified that for all instruments except CW-TILDAS-CS the random error due to instrumental noise was an important source of uncertainty at 30 min averaging level and the total stochastic error was frequently of the same magnitude as the fluxes when N_2O exchange was small at the measurement site. Both instruments based on Continuous-Wave Quantum Cascade Lasers, CW-TILDAS-CS and N2O/CO-23d, were able to determine the same sample of low N_2O fluxes with high mutual coefficient of determination at 30 min averaging level and with minor systematic difference over the observation period of several months. This enables us to conclude that the new generation instrumentation is capable of measuring small N_2O exchange with high precision and accuracy at sites with low fluxes.

Keywords: nitrous oxide, fast response instruments, eddy covariance, system performance.

1 Introduction

During the last years there has been a rapid development in the application of laser spectroscopy for greenhouse gas measurements. In particular, development of fast response N_2O analyzers based on spectroscopic techniques (e.g. tunable diode laser (TDL) and quantum cascade laser (QCL) spectrometers) has facilitated the eddy covariance (EC) measurements of N_2O exchange in different ecosystems. Such measurements have been reported in literature and they have been carried out in different ecosystems such as agricultural (Smith et al., 1994; Wienhold et al., 1994; Christensen et al., 1996; Laville et al., 1997; Scanlon and Kiely, 2003; Neftel et al., 2007; Kroon et al., 2007), forest (Pihlatie et al., 2005; Eugster et al., 2007) as well over urban canopies (Famulari *et al.* 2010; Järvi et al., 2014).

The observed N_2O emissions are episodic in nature, showing high spatial and temporal variability. Emission bursts of short duration, typically occurring after fertilizer application, or associated with thawing and rain events (Kroon et al., 2007, Pihlatie et al., 2010), are followed by long periods of small fluxes, when also uptake of N_2O has been observed (Flechard et al., 2005). Overall, N_2O fluxes reported by previous studies are characterised by large uncertainty and temporal variability, which are related to biogeochemical soil processes and several systematic and random error sources of the EC measurements. One of the sources

of uncertainty for the N_2O fluxes measured by the EC technique is the performance and stability of fast response gas analyzers. Some studies performed under field conditions (Eugster et al., 2007; Kroon et al., 2007; Neftel et al., 2009) have reported that the laser drift can cause occasional over- or under-estimation of EC flux. The instrumental drift typically characterizes TDL as well as QCL spectrometers (Werle et al., 1993; Nelson et al., 2002). Mammarella et al. (2010) thoroughly investigated the performance of TDL instruments in measurements of N_2O fluxes by the EC technique. They suggested that high pass filtering could be used to remove the low-frequency signal drifting, which could otherwise contaminate the detected concentration time series and significantly increase the flux uncertainty.

Apart from the episodic emissions, N_2O fluxes are typically small in magnitude (in the order of one to one hundred μg N m^{-2} h^{-1}, which corresponds to N_2O flux range from 10^{-2} to 1 nmol m^{-2} s^{-1} as presented in the units used in the current study), being on the detection limit of the EC systems (e.g. Pihlatie et al., 2005; Wang et al., 2013). Small fluxes imply small turbulent fluctuations of the concentration, requiring high precision of the instruments to be able to resolve those fluctuations. In other words, the signal (turbulent fluctuations) to instrumental noise ratio has to be high enough to achieve sufficiently low flux error arising due to the noise present in measured signals (Lenschow and Kristensen, 1985).

The goals of this study are to compare the available equipment for N_2O flux measurements employing the EC technique and to evaluate their performance, ability to detect small fluxes and long-term stability in determining the N_2O exchange. The instruments used were TGA100A (Campbell Scientific Inc.), CW-TILDAS-CS (Aerodyne Research Inc.), N2O/CO-23d (Los Gatos Research Inc.) and QC-TILDAS-76-CS (Aerodyne Research Inc.), which shall be further referred to as CS-TDL, AR-CW-QCL, LGR-CW-QCL and AR-P-QCL, respectively, throughout this study by using the combinations of acronyms for manufacturer and the laser type (see Table 1). In addition, the methods for flux calculation using the laser spectrometer data are evaluated and the magnitude and dynamics of N_2O fluxes during the RCG growing season are determined.

2 Materials and methods

2.1 Site

The measurement site was a 6.9 ha field cultivated with RCG, a perennial bioenergy crop. The site was located on the rural area of Maaninka, Eastern Finland (63° 9' 48.69" N, 27° 14' 3.29" E). Long-term (reference period 1981-2010; Pirinen et al., 2012) annual air temperature in the region is 3.2°C, the coldest month of the year is February and the warmest is July, with monthly mean air temperature being -9.4°C and 17.0°C, respectively. The annual precipitation in the region is 612 mm. Part of this precipitation amount falls as snow. Snow cover season starts in October and lasts until the end of April with a maximum snow cover of approximately 50 cm. The RCG crop at the Maaninka site was fertilized in the beginning of the growing season (late May), resulting in a large emission pulse of N_2O. The site was applied with an N-P-K-S fertilizer containing 76 kg N ha^{-1}, based on ammonium nitrate (NO_3-N : NH_4-N = 47:53). The canopy height developed throughout the growing season from about 10 cm in mid-May to 1.7 m by late June. The increase in plant height was almost linear in time between these periods and starting from July changed slowly up to 1.9 m.

The soil at the study site is classified as fine sand to coarse silt (particle size 0.03 – 0.06 mm). According to the World Reference Base for Soil Resources (WRB) system (FAO, 2006), the soil is classified as Regosol. The soil pH varies from 5.4 to 6.1 within the ploughing depth from the surface to about 30 cm, electrical conductivity between 960 to 3060 µS cm^{-1} and soil organic matter content between 3 and 11%. The average C/N ratio in the ploughing depth is 14.9 (ranging from 14.1 to 15.7). The soil particle density is about 2.65 g cm^{-3} within the soil depth from the surface to about 20 cm.

2.2 Measurements

Measurements were conducted by the University of Helsinki (UH) and by the University of Eastern Finland (UEF), operating separate EC systems based on two different sonic anemometers. The UH measurement setup included a 3-D ultrasonic anemometer (USA-1, METEK GmbH, Elmshorn, Germany) to acquire the wind components. The anemometer was installed on top of a pole, the measurement height being 2.2 m. The measurement height was raised to 2.4 m on 30.6.2011 due to the RCG growth. Gas analyzers were situated in an air conditioned cabin located about 15 m east from the anemometer pole. This wind direction (50-110° sector) was therefore discarded from further analysis due to possible disturbances to flux measurements. Sample inlets for gas analyzers were located 10 cm below the

anemometer. The N_2O instruments operated by the UH were the instrument based on tunable diode laser CS-TDL (model TGA100A, Campbell Scientific Inc.), and two instruments based on continuous wave quantum cascade lasers, AR-CW-QCL (models CW-TILDAS-CS, Aerodyne Research Inc., see e.g. Zahniser et al., 2009; Lee et al., 2011) and LGR-CW-QCL (model N2O/CO-23d, Los Gatos Research Inc., see e.g. Provencal et al., 2005). Sampling lines of AR-CW-QCL and LGR-CW-QCL were heated slightly above ambient temperature in order to avoid water from condensing to the lines. CS-TDL had a dryer just before the instrument and no sampling line heating was used. The flow rates and tube dimensions were chosen to correspond to turbulent flow regime except that the larger diameter of the sampling line of the LGR-CW-QCL analyser resulted in laminar tube flow for that instrument (see Section 3.1 below). Further details of the involved instruments are given in Table 1 and details of the different setups are given in Table 2.

The maintenance of CS-TDL was the most demanding of the compared instruments. It uses liquid nitrogen to keep the laser source at the operating temperature, and the Dewar was filled up twice a week. The instrument CS-TDL was calibrated in the beginning of the campaign. Further the operating parameters of the analyser, such as laser current and laser, housing and detector temperatures were checked once a week and after power failures. In addition, the shape and intensity of the absorption line were checked at the same time. These checks were assumed to guarantee calibration stability of the instrument to a reasonable degree. In addition, the inlet filter of CS-TDL was changed once a month.

The instrument AR-CW-QCL was calibrated and its operating parameters were fine-tuned at the site after instrument installation. The instrument manufacturer provided a software upgrade during the campaign to conduct the real-time water vapour correction to the trace gas concentration data analysed by the instrument. In addition, the operating parameters were fine-tuned a few times on-line by the instrument manufacturer during the campaign.

LGR-CW-QCL arrived in the campaign later (see Section 2.6 for details). The factory calibration of LGR-CW-QCL was checked but no deviation was observed within the uncertainty range of the calibration gases. After about two weeks of operation, the laser drifted out of the tuning range and the laser offset current was tuned manually to enable correct operation again. No calibration of the instruments based on CW-QCL-s was performed during the campaign as these analysers were expected to be very stable according to manufacturers' information.

The UEF set up included a pulsed quantum cascade laser spectrometer AR-P-QCL (Model QC-TILDAS-76-CS, Aerodyne Research Inc., Billerica, MS, USA, see McManus et al., 2005), an infrared gas analyser (IRGA, Model Li-6262) and a 3-D sonic anemometer (Model R3-50, Gill Instruments, Ltd., Hampshire, UK) for fast response gas concentration and wind component measurements (Tables 1 and 2). The heated intake tubes for the laser spectrometer and IRGA were installed on either sides of the sonic anemometer, all mounted on a boom on an adjustable instrument mast. The mast height was set at 2.0 m above the soil surface in the beginning of the campaign. To adjust to the increasing plant height, the mast was raised to 2.5 m during mid-June. AR-P-QCL was set up to measure simultaneously the N_2O, CO_2 and water vapour mixing ratios, while the IRGA was used to monitor the CO_2 and water vapour mixing ratios. Both trace gas analysers were calibrated against standard gases minimum once a month during the campaign, in particular AR-P-QCL was calibrated every 2-3 weeks with two standard gases 299 and 342 ppb. The calibration slope of AR-P-QCL did not change by more than 7.6% throughout the campaign and maximum 6.1% between consecutive calibrations. Thus 6.1% can be considered as the maximum flux systematic error arising from calibration accuracy of this instrument.

A weather station set up on another mast close to the EC mast monitored the supporting meteorological variables. The weather station mast height was also adjusted according to the changes in the EC mast height. Supporting measurements included air temperature and relative humidity (Model: HMP45C, Vaisala Inc.) using radiation shield, atmospheric pressure (Model CS106 Vaisala PTB110 Barometer), wind speed and direction (Model 03002-5, R.M. Young Company) and several other variables not used in current study. Data was collected using a datalogger (model CR 3000, Campbell Scientific Inc.). Except air pressure (stored as hourly averages), meteorological data was stored as 30 minute averages. Short gaps in the data were filled using linear interpolation, but when air temperature, relative humidity, pressure or rainfall data were missing for longer periods, data from Maaninka weather station operated by the Finnish Meteorological Institute located about 6 km to South-East from the site, was used.

2.3 Flux processing

Measurements were sampled at 10 Hz frequency. Filtering to eliminate spikes was performed according to standard approach (Vickers and Mahrt, 1997), where the high frequency EC data

were despiked by comparing two adjacent measurements. If the difference between two adjacent concentration measurements of N_2O was greater than 20 ppb, the following point was replaced with the same value as the previous point.

The spectroscopic correction due to water vapour impact on the absorption line shape was applied along with the Webb-Pearman-Leuning (WPL) dilution correction due to water vapour on high-frequency raw concentration output X_C (mixing ratio with respect to moist air, uncorrected for spectroscopic effect) according to $\chi_C = \frac{X_C}{1-(1+b)\chi_V}$, where χ_C and χ_V are the instantaneous mixing ratios of N_2O and water vapour with respect to dry air and b is the spectroscopic correction coefficient determined experimentally for each instrument (Table 1) by measuring the response of instrument (output X_C) on sample air of standard gas (constant χ_C) with varying water content χ_V. The correction was not necessary for CS-TDL as a dryer installed after the air intake point on the sampling line dried the air sample before the optical cell. LGR-CW-QCL corrected for the water vapour effect by a built-in module in the LGR data acquisition software; the same applied to AR-CW-QCL after software update in July 2011.

Prior to calculating the turbulent fluxes, a 2-D rotation (mean lateral and vertical wind equal to zero) of sonic anemometer wind components was done according to Kaimal and Finnigan (1994) and all variables were linearly detrended. The EC fluxes were calculated as 30 min co-variances between the scalars and vertical wind velocity following commonly accepted procedures (e.g. Aubinet et al, 2000). Time lag between the concentration and wind measurements induced by the sampling lines was determined by maximizing the covariance. For CS-TDL the lag was determined by maximizing the covariance for high flux period only (day of year (DOY) 144-146) because in other periods the lag was not well defined by using this method. The final processing (instruments CS-TDL, AR-CW-QCL and LGR-CW-QCL) was done by fixing the time lag to avoid unphysical variation of lag occurring due to random flux errors. For AR-P-QCL system the lag was determined by maximising the covariance for CO_2 and the same lag was assigned to N_2O. This was to use the advantage that the instrument measured also CO_2 and therefore enabled to use much better signal-to-noise ratio in determination of the lag time. Spectral corrections were applied to account for the low and high frequency attenuation of the co-variances (Sect. 2.4). Then, the humidity effect on temperature flux was accounted for after Schotanus et al. (1983). All data processing was

performed with post-processing software EddyUH (http://www.atm.helsinki.fi/Eddy_Covariance/EddyUHsoftware.php).

2.4 Spectral corrections

Low and high frequency variations in the measured signal are attenuated due to data acquisition and processing, and by a non-ideal measurement system (e.g. Moore, 1986; Moncrieff et al., 1997; Rannik and Vesala, 1999; Massman, 2000). Block averaging and detrending of data acts as a high pass filter, thus damping low frequency fluctuations (Rannik and Vesala, 1999; Finnigan et al., 2003). Turbulent fluctuations occurring at high frequencies are attenuated due to the measurement system's limitations. Gas analyzer's finite frequency response, attenuation of fluctuations in the sampling line, spatial separation between the anemometer measurement head and sampling line inlet affect the attenuation of high frequency fluctuations in the signal.

The observed flux (F_m) can be formally presented as the integral over the convolution of the true co-spectrum (Co, unaffected by frequency attenuation) with the co-spectral transfer function as

$$F_m = \int_0^\infty T(f)Co(f)df \quad , \qquad (1)$$

where the co-spectral transfer function can be presented as the convolution of respective low-frequency $T_L(f)$ and high-frequency $T_H(f)$ transfer functions. For the low-frequency transfer function due to high-pass filtering and/or finite averaging period see Rannik and Vesala (1999).

For evaluation of the instrument frequency performance and subsequent high-frequency flux corrections during post-processing, the high-frequency transfer function of the EC-system was estimated (Aubinet et al., 2000) as the ratio of the observed and not-attenuated flux (Horst, 1997). The co-spectral transfer function $T_H(f)$ for a system behaving as a first order response sensor can be described by

$$T_H(f) = \frac{1}{1+(2\pi f\tau)^2} \quad , \qquad (2)$$

where f is the natural frequency and τ the (first order) response time of the attenuator (sensor or the system in total) (Horst, 1997). The effective transfer function of the EC system for

different instruments was estimated as the ratio of co-spectral density of scalar flux relative to co-spectrum of sensible heat flux (Aubinet *et al.* 2000). Such a procedure assumed that temperature measurements were not affected by attenuation (true for the sonic anemometer) and includes normalisation with integral over frequencies not affected by attenuation.

2.5 Estimation of random errors

Turbulent fluxes averaged over a limited time period have random errors because of the stochastic nature of turbulence (Lenschow et al. 1994; Rannik et al., 2006) as well as due to noise presented in measured signals (Lenschow and Kristensen, 1985).

The random error of the flux was evaluated as one standard deviation of the co-variance error, hereafter in the manuscript denoted by δ_F. It was defined through the variance of the distribution of the individual flux realization around the ensemble mean (e.g. Lenschow et al., 1994). Theoretically, there are several approaches to approximate the same error estimate, see e.g. Rannik et al. (2009). Currently, the flux random error was calculated according to the method implemented in EddyUH, the method proposed by Finkelstein and Sims (2001). The method evaluates the error in time domain through integration of the auto-covariance and cross-covariance functions of the vertical wind speed and the scalar concentration according to

$$\delta_F \approx \sqrt{\frac{1}{n}\left[\sum_{p=-m}^{m}\overline{w'w'}(p)\overline{c'c'}(p)+\sum_{p=-m}^{m}\overline{w'c'}(p)\overline{c'w'}(p)\right]}, \qquad (3)$$

where $\overline{w'w'}(p)=\frac{1}{n}\sum_{i=1}^{n-p}\left(w(t_i)-\overline{w}\right)\left(w(t_{i+p})-\overline{w}\right)$. In calculations we used m = 200 (corresponding to 20 sec) to ensure that integration of the covariance functions was performed over times exceeding the integral time scale of turbulence. This mathematically rigorous method provides estimates for the random uncertainty of the flux measurements for every flux averaging period.

Random uncertainty of the observed co-variance due to presence of noise in instruments signal, giving essentially the lowest limit of the flux that the system is able to measure, was expressed in its simplest form as

$$\delta_{F,noise}=\frac{\sigma_w\sigma_{noise}}{\sqrt{fT}}, \qquad (4)$$

where σ_w and σ_{noise} denote the standard deviation of the turbulent record of vertical wind speed and the standard deviation of instrumental noise as observed at frequency *f*, *T* denotes the flux averaging period. The expression above assumes that the noise component of the vertical wind speed measurement is negligible. In this study we use the method developed by Lenschow et al. (2000) and applied to EC fluxes by Mauder et al. (2013) to estimate the flux error due to instrumental noise. Lenschow et al. (2000) derived the method to estimate the instrumental random noise variance $(\sigma_{noise})^2$ from the auto-covariance function of the measured turbulent record close to zero-shift, enabling to determine the error for each half-hour flux averaging period.

The random flux error δ_F is the results of limited sampling in time and/or in space of a stochastic turbulence realization. Its expression includes the covariance and cross-covariance functions of turbulent records, therefore in addition to variances and co-variances accounting for respective integral time scales of turbulent records. The error δ_F incorporates also the contribution due to instrumental noise and is therefore larger from the latter.

The error $\delta_{F,noise}$ instead does not depend on the intergral time scale of turbulence, being therefore mainly determined by the instrumental noise characteristics and less on the observation conditions (only via σ_w). Assuming no true turbulent variation of concentration and thus zero flux, the calculated flux will be generally non-zero due to noise in instrumental signal. Evidently the system will not be able to detect the fluxes smaller than the ones obtained from the expression for $\delta_{F,noise}$. Therefore this is the minimum flux that the EC system can detect and $\delta_{F,noise}$ serves useful in characterising the instrumental limitation to detect small fluxes.

If an average over fluxes F_i (i = *1..N*) is calculated, each of these representing a flux value observed over averaging period *T* and being characterised by an error $\delta_{F,i}$, then the error of the average flux $\langle F \rangle = \frac{1}{N}\sum_{i=1}^{N} F_i$ was expressed as

$$\Delta_{<F>} = \sqrt{\frac{\sum_{i=1}^{N}(\delta_{F,i})^2}{N^2}} \quad . \qquad (5)$$

This expression will be used to estimate the random errors of the average fluxes in Sect. 3.4.

2.6 Periods of analysis and quality screening

The intercomparison measurements were performed from the beginning of the growing season in April till November 2011. According to instrumental data coverage, the period was divided into three sub-periods for the instrument evaluation and flux analysis purposes. During the period I, DOY 110-181 (20.04-30.06.2011), the measurements of CS-TDL, AR-CW-QCL and AR-P-QCL were available, during the period II, DOY 206-271 (25.07-28.09.2011), all instruments were measuring and during period III, DOY 272 – 324 (29.09-20.11.2011), all other except CS-TDL were operational. Prior to analysis data quality screening was performed. The measurements corresponding to wind direction interval 50-110° were excluded as possibly affected by instrumental cabin. In addition, quality screening was performed according to Vickers and Mahrt (1997) by applying the following statistics and selection thresholds: data with N_2O concentration skewness outside (-2, 2), or kurtosis outside (1, 8), or Haar mean and Haar variance exceeding 3 were rejected. Applying the same statistics and thresholds as for N_2O, additional quality screening of N_2O fluxes was performed according to H_2O concentration statistics for AR-CW-QCL and AR-P-QCL due to the impact of the spectroscopic and dilution corrections on fluxes and according to CO_2 concentration statistics for AR-P-QCL because the lag obtained for CO_2 was assigned to N_2O in case of this instrument.

The applied quality criteria were used to ensure exclusion of the system malfunctioning as well as unphysical and/or unusual occasions. No quality screening for stationarity was performed as the focus of the study was the instrumental intercomparison, which was not affected by occasional non-stationary conditions included in the analysed data set.

3 Results

The fluxes obtained for three periods are presented in Fig. 1, being averaged over daily period for the clarity of presentation. No gap-filling was used and for each day only the existing measurements, after applying data quality screening described above, were averaged. In May the fluxes increased significantly after the fertilization and then decreased back to low, although clearly positive level after a few weeks. This was the only occasion of high N_2O emission followed by continuous decrease of fluxes towards the autumn. The soil temperature had increasing trend until about DOY 205 (24 July, 2011) and since August declining

seasonal trend (Fig. 2). SWC increased with occasional rain events. During the high emission, starting from DOY 144 (24 July, 2011) and lasting until approximately DOY 155 (4 June, 2011), the SWC was approximately 0.3 m^3 m^{-3}, being relatively high.

The high fluxes observed during that period enabled to evaluate the frequency performance of three systems including CS-TDL, AR-CW-QCL and AR-P-QCL. The LGR-CW-QCL instrument was not operational then and the frequency response analysis for this instrument was performed based on the concurrently measured H_2O and CO signal analysis.

3.1 Spectral characteristics of instruments

Spectral analysis was performed to study the frequency performance of the instruments. In general, averaging over long periods should lead to better spectral statistics. However, aggregating over different periods might lead to biased results as the spectra do not necessary follow the idealised normalizations in frequency scale, considering also that spectral scaling depends on stability. Therefore we aimed to use optimal averaging period over several hours for similar conditions in terms of wind speed and stability. For the period May 26th, from 7:00 to 13:00 EET (Eastern European Time) when the conditions corresponded to moderately unstable (average wind speed of the period 3.2 m s^{-1} and sensible heat flux 50 W m^{-2}), the calculated spectra exhibited very clear and systematic patterns for temperature as well as N_2O concentration records measured by three instruments (Fig. 3). In spite of high fluxes registered by the instruments during this period, CS-TDL N_2O signal was dominated by noise almost over the whole frequency range presented. For AR-CW-QCL, almost no evidence of noise could be observed in the power spectral plot (multiplied with frequency). The older version by Aerodyne, the AR-P-QCL instrument, revealed increase of the spectral density only at the high-frequency end of the power spectrum, being characteristic to some noise contribution. The co-spectra of all three instruments showed smooth patterns, the shape being consistent with the co-spectral model by Kaimal et al. (1972) but slightly shifted in frequency scale. At the high frequency ends of the presented co-spectra the N_2O signal curves deviate from the theoretical as well as from temperature co-spectra, indicating attenuation of signals at high frequencies by the measurement systems.

The same time period was used to estimate the frequency response of the N_2O eddy covariance systems according to the method described in Sect. 2.4 (Fig. 4). The time constants estimated by making use of the co-spectra presented in Fig. 3 and eq. (2) for CS-

TDL, AR-CW-QCL and AR-P-QCL were 0.12, 0.07 and 0.08 seconds, respectively. Note that these time constants characterise the frequency response of the systems in total.

Although the response time obtained for AR-P-QCL system from high flux period was 0.08 sec, the analysis of the response time from measured CO_2 signal for several other periods yielded the average response time 0.15 sec. The N_2O signal was synchronised with CO_2 by using the lag determined for CO_2 and theoretically the N_2O response time does not differ from that of CO_2 under turbulent tube flow regime, hence we choose the constant value 0.15 sec for co-spectral corrections throughout the campaign for this instrument.

Spectral analysis was performed also for the period when LGR-CW-QCL measurements were available. For the comparison purpose, the results for a time period in August 4th form 00:30 to 4:00 EET are presented for AR-CW-QCL and LGR-CW-QCL instruments (Fig. 5). The period was chosen with relatively high fluxes (with LGR-CW-QCL measurements available) and similar stability and wind conditions (average wind speed of the period 0.94 m s^{-1} and sensible heat flux -37.5 W m^{-2}). The power spectra of both instruments revealed contribution of noise at high frequency ends of the spectra, being more pronounced for LGR-CW-QCL. The co-spectra were more scattered when compared to high flux period (Fig. 3). Estimation of the frequency response of the systems based on this period was uncertain due to scatter and could not be used as the basis for co-spectral corrections for LGR-CW-QCL.

The main difference in the flow setups of the systems concerned LGR-CW-QCL. With larger tube diameter and slightly lower flow rate the flow regime was likely laminar (Re $\approx$ 2 000), whereas for other instruments it was clearly turbulent (Re $\geq$ 4 600). It is well established that under laminar flow regime tube flow attenuates turbulent fluctuations of concentration much more than under turbulent flow. According to the expression for tube attenuation in laminar flow regime (Foken et al., 2012) the first order response time for LGR-CW-QCL flow setup would be 0.37 sec (estimated for N_2O). For turbulent flow (ARI-CW-QCL setup) the theoretical response time for tube damping is much smaller (0.01 sec) than the response time obtained from the co-spectra (0.07 sec), suggesting that the system's response was dominated by the instrumental response.

The frequency response of the LGR-CW-QCL system was further determined from the co-spectral analysis of the CO signal and we obtained the value 0.26 sec. We determined also the

experimental response time for water vapour from several periods corresponding to low humidity conditions (RH<40%) and we consistently found the value around 0.35 sec (for LGR-CW-QCL system). For comparison, the response time for H_2O measured by ARI-CW-QCL system was determined to be 0.10 sec. Damping of water fluctuations in sampling lines is stronger than for other scalars as evidenced by experimental studies (e.g. Mammarella et al., 2009). This is due to adsorption/desorption of water molecules on tube walls. This explains the difference between the response times obtained from CO and H_2O. Thus we believe that a value of 0.26 sec characterises well the first order response time of LGR-CW-QCL setup for N_2O and we use this value in co-spectral corrections. Note, however, that a higher response time of the LGR-CW-QCL system does not mean a slower instrument performance because the system has more damping primarily in the sampling line due to lower flow rate and larger tube diameter (Table 2).

The frequency response times determined in this section were used in performing the co-spectral corrections (Table 2) as described in Sect. 2.4, typical magnitudes of these corrections are presented in Table 3.

3.2 Random uncertainty of fluxes and instrumental noise

The method by Lenschow et al. (2000) described in Sect. 2.5 enabled to calculate the instrumental noise for each 30 min period and the resulting flux uncertainty due to instrumental noise. Fig. 6a shows the estimated signal noise statistics with upper and lower percentiles and quantiles (boxes) with a median value in the middle. For all instruments except LGR-CW-QCL the distributions are very narrow and different percentiles cannot be separated from the plot (for values see Table 1). This tells us that the noise levels of the three instruments are very stable, but the noise level of LGR-CW-QCL somewhat varies. In comparison of the instruments, AR-CW-QCL has by far the lowest noise level of around 0.12 ppb (standard deviation of the signal noise at 10 Hz frequency). The two instruments, LGR-CW-QCL and AR-P-QCL, are characterised by a similar noise level (around 0.5 ppb), while CS-TDL signals show the highest noise level (2 ppb). Consequently, these instrumental noise levels are reflected in the random errors of fluxes, determining essentially the minimum flux level that each instrument is able to measure at a given flux averaging interval (30 min period). For AR-CW-QCL the respective lowest flux is around 10^{-2} nmol m^{-2} s^{-1} (as given by median in Fig. 6b), for LGR-CW-QCL and AR-P-QCL around $4x10^{-2}$ nmol m^{-2} s^{-1} and for CS-TDL 0.15 nmol m^{-2} s^{-1}.

The frequency distributions of the total flux random errors, calculated according to Eq. (3), are naturally higher than the flux error due to instrumental noise only. It can be observed that in case of full flux random error the difference between different instruments is reduced (Fig. 6b) because in addition to instrumental noise impact this error statistic also incorporates the flux uncertainty due to stochastic nature of turbulence. The relative random errors (Fig. 6c) are the largest for CS-TDL (being in the order of 100% and in most cases less than ±300%) and the smallest for AR-CW-QCL (median around 30% and mostly the error being less than 100%) instruments. It is the signal noise of the instrument that contributes to the random error of the flux, determining which instrument is able to detect lowest fluxes. In case of CS-TDL the low-frequency signal drifting can also enlarge the total random error of the calculated flux.

3.3 Intercomparison of fluxes averaged over turbulent spectrum

It was observed that the fluxes calculated from CS-TDL measurements during the low flux period were dominated by stochastic uncertainty, being frequently in the order of the random uncertainties of fluxes (Sect. 3.2). Therefore, the fluxes averaged over 30 min period were compared for this instrument with AR-CW-QCL results over the period DOY 110-182, which included the high emission episode starting from DOY 144 and exhibiting elevated fluxes until approximately DOY 155. In general the fluxes with high magnitude obtained by CS-TDL compared well with those of obtained by AR-CW-QCL (Fig. 7a). The AR-P-QCL system, as compared with AR-CW-QCL, showed systematically lower fluxes during the given period of high fluxes (slope 0.70). In spite of lower noise level of this instrument, the coefficient of determination for this instrument (0.63) was lower than that for CS-TDL (0.77) in comparison to the fluxes as measured by AR-CW-QCL.

During the second observation period, when fluxes were much lower, CS-TDL was not able to determine fluxes with sufficiently small error and the correlation with AR-CW-QCL at 30 min averaging level was very low (Fig. 7c). At around zero fluxes as measured by AR-CW-QCL, the results by CS-TDL showed scattered values visually between ±2 nmol m^{-2} s^{-1}. The noise level of CS-TDL around 2 ppb translates into flux uncertainty due to instrumental noise of about 0.05 to 0.3 nmol $m^{-2}s^{-1}$. The total flux error δ_F was within the range from 0.1 to 0.45 nmol $m^{-2}s^{-1}$ (upper and lower quantiles of the distribution in Fig. 6b). We analysed the range of variation of CS-TDL fluxes during the given period DOY 206-272, conditionally selecting

the observations when the observed fluxes by AR-CW-QCL were absolutely smaller than 0.15 nmol m^{-2} s^{-1} (90% of N_2O flux random errors for AR-CW-QCL less than this value during the given period). The respective N_2O fluxes as determined by CS-TDL were characterised by the upper and lower quantiles of -0.27 and 0.52 nmol m^{-2} s^{-1}. This is consistent with the upper quantile of the flux error distribution for CS-TDL. Therefore the fluxes of CS-TDL, corresponding to close-to-zero fluxes as determined by AR-CW-QCL, were consistent with the flux error estimates.

The comparison of the 30 min average fluxes calculated from two instruments, AR-CW-QCL and LGR-CW-QCL, revealed very good correspondence and high correlation (R^2 = 0.90) even though those measurements corresponded to very low N_2O fluxes. The slope close to unity and negligible intercept indicates no systematic bias between the measurements of these systems (Fig. 7d).

3.4 Long-term averages and systematic differences

In order to evaluate the possible systematic differences, cumulative curves of the flux observations were calculated. No gap-filling of missing data was done but instead only the half-hour periods were used when the results for all instruments were available. Thus the cumulative sums do not assume representing the total emissions over the given periods, although rough estimates could be calculated by accounting in total sums with the data coverage percentage presented in Table 4. The summation of fluxes over the first and second periods reveals that CS-TDL gives the highest flux sums and AR-P-QCL the lowest, in particular during the first period (Fig. 8). The cumulative sums for fluxes obtained from AR-CW-QCL and LGR-CW-QCL measurements converge over 2nd and 3rd periods and show only small differences. Also the cumulative fluxes measured by AR-P-QCL during these periods are very close to fluxes measured by the two other instruments. In order to assess the magnitude of the random errors in these differences, the random errors of the fluxes averaged over three periods were calculated according to Eq. (5). The analysis revealed that the average fluxes for period II, obtained from the measurements of AR-CW-QCL and LGR-CW-QCL instruments did not differ within calculated error limits, and were very close during the period III with the result for AR-P-QCL (Table 4).

However, CS-TDL produced a 7% higher total sum for the period of high fluxes (DOY 110-181 with an average flux of 0.87 nmol m^{-2} s^{-1} as determined by AR-CW-QCL) and a 29%

higher sum for the second period (DOY 206-271) compared to an average flux 0.142 nmol m^{-2} s^{-1} (average of AR-CW-QCL and LGR-CW-QCL results). The AR-P-QCL instrument determined for these two periods 36% and 13% lower average fluxes, respectively. The possible reasons for this will be discussed in the next section. For the third period, the results for AR-P-QCL did not differ much from the results of the other two instruments.

4 Discussion

Performance of four instruments (see Tables 1 and 2) capable of fast response measurement of N_2O was studied throughout the 2011 growing season over a field cultivated with RCG in Eastern Finland. The N_2O fluxes were small in the beginning of the season, increased significantly after the fertilization (late May) and then decreased back to low, positive values after a few weeks. Three instruments, CS-TDL, AR-CW-QCL and AR-P-QCL were operational during this high emission period. During this period, all instruments detected the same flux dynamics, whereas the fluxes obtained by AR-P-QCL, the previous instrument version by Aerodyne, were lower compared to the other two instruments.

For many applications the systematic errors of micrometeorological flux measurements of atmospheric trace gases are more important than the random errors. For example, for determination of annual balances (e.g. Kroon et al., 2010b) or for the comparison of exchange of different ecosystems (e.g. Nicolini et al., 2013) the systematic errors become very important. The two CW-QCL instruments compared very well on half-hour basis as well as produced statistically close cumulative fluxes over the period when the two instruments were simultaneously operational (25.07.2011-20.11.2011). The cumulative emission estimate obtained by CS-TDL for the same period was 29% higher than the average result for instruments based on the continuous wave quantum cascade lasers, AR-CW-QCL and LGR-CW-QCL. AR-P-QCL obtained 36% lower fluxes than AR-CW-QCL during the first period including the emission episode, whereas the correspondence with other instruments during the rest of the campaign was relatively good.

The systematic differences in fluxes could be the result of calibration and/or limited stability of the system over time. The impact of the instruments calibration (sensitivity shift) impact on flux systematic differences can be assessed by using calibration information (Section 2.2) as well as comparison of average concentrations measured by different instruments. The two analysers based on CW-QCL-s are expected to be very stable, which was confirmed by the

measurements: The concentrations measured by these two instruments were very consistent and the slope (characterising sensitivity) of the 30 min average concentration comparison did not deviate from unity by more than 5% (with the coefficient of determination of linear regression $R^2 = 0.86$).

The sensitivity of AR-P-QCL did not change more than 6.1% between consecutive calibrations and this can be considered as the maximum flux error arising from calibration accuracy of this instrument (Section 2.2). Nevertheless, the correlation of the 30 min average concentration measured by this instrument as compared to AR-CW-QCL was not as good (for the period DOY 206-272 slope 1.05 was determined with $R^2 = 0.63$). The concentration comparison presented here does not reveal that the calibration bias was the reason for the observed flux systematic difference for the instrument AR-P-QCL.

The analyser CS-TDL is known for its signal drifting as illustrated and discussed by Mammarella et al. (2010) and the absolute concentrations were not well determined during our campaign. Therefore accurate measurement of absolute concentration by this instrument over a long period of time cannot be expected and the concentration comparison was not used as the method for evaluation of the instrument's calibration impact on flux systematic bias. Note that signal drifting makes the time series produced by the instrument essentially non-stationary and therefore enhances the random variability of the flux estimate around the true value. However, such enhanced random uncertainty does not affect systematically the cumulative sums over longer periods.

In case of low fluxes the water vapour dilution and spectral line broadening effects are the primary suspects for the reasons in systematic differences in fluxes (e.g. Peltola et al., 2014). Close correspondence of the concentrations and fluxes as measured by AR-CW-QCL and LGR-CW-QCL let us conclude that the spectroscopic and water vapour dilution corrections for these instruments were adequate. Note that those corrections were done by built in functionality in case of LGR-CW-QCL. For AR-CW-QCL the respective corrections were done in post processing phase for the period I and by built-in software for the rest of the campaign.

The only evident systematic flux error source that could affect performance of CS-TDL would be incomplete drying of sample air. If that was the case, then the calculated fluxes had

suffered from missing partial density and spectroscopic corrections. Since the water fluxes are dominantly upward, a respective correction would tend to increase the flux values, therefore increasing even more the systematic difference relative to other instruments.

The instrument ARI-P-QCL is based on the pulsed quantum cascade laser. For this instrument the experimentally determined spectroscopic correction coefficient was much lower than the coefficient for AR-CW-QCL (Table 1). The reason for systematically lower values of fluxes determined by AR-P-QCL from the beginning of the experiment in April till June 2011, but subsequent relatively good comparison with other instruments till the end of the experiment in November 2011, is not known. Two types of corrections were applied to N_2O fluxes: the spectroscopic correction to account for the impact of water vapour on the absorption line shape, and the co-spectral correction. The latter correction was comparable to all instruments (Table 3) and does not introduce significant difference between instruments. The spectroscopic correction was applied together with the water vapour dilution correction (Sect. 2.3) and can constitute a major correction depending on the value of the coefficient b. The correction is related to the water vapour flux, which was during the day time on the average around 100 Wm^{-2} (periods I and II, Table 5), with mid-day averages around 150 to 200 Wm^{-2}. Considering the average concentration of N_2O around 330 ppb and the spectroscopic correction value b=0.39 (the value for AR-CW-QCL), the spectroscopic correction can be a few tenths of nmol m^{-2} s^{-1} during mid-day, which is of the order of the flux magnitude. We used all auxiliary data available to investigate the possible reasons for the systematic differences, but found no explaining variable or reason. In particular, no systematic variation of the residual between AR-P-QCL and AR-CW-QCL fluxes was found over wide range of latent heat fluxes from -20 to 250 W m^{-2}. This proves that the dilution and spectroscopic corrections were properly accounted for. In addition, larger spectroscopic correction would not explain systematic difference observed during the first period only.

Thus the reasons for flux underestimation by AR-P-QCL during the period I are not known and we suggest that extreme care should be exercised during the long-term measurement campaigns both with N_2O and H_2O calibrations due to the strong impact of the latter on the N_2O flux through spectroscopic and dilution corrections.

A comment should be made regarding the observation level used in the study. When RCG was grown high, the measurement level was only about 0.5 m above the canopy top. The

measurements within the roughness sublayer can be disturbed in terms of several statistics, but the impact can be expected revealed more in spectral shapes than in integral statistics. The spectra obtained for N_2O (Fig. 3 and 5) were dominated by white noise over wider (CS-TDL) or narrow (AR-CW-QCL) frequency ranges depending on the instrument in question. The temperature spectra were similarly affected by the noise but only at the high frequency end of spectra and we believe that not evidencing the canopy impact on spectral shapes. We checked also the spectra for vertical wind speed (not shown). The spectra exhibited smooth and consistent shapes, without the particular impact of the canopy foliage on spectral forms usually observed inside canopy. Launiainen et al. (2007) studied the turbulence statistics and spectral shapes within pine forest canopy. They did not observe deviation of spectral shapes above canopy at height z/h = 1.47 (h being the canopy height) from the atmospheric surface layer forms, within the crown space (z/h = 0.78) the spectra deviated only slightly from the above-canopy forms. Within the trunk space (z/h = 0.4) the spectra were distorted due to the drag imposed by the canopy elements. This supports that the spectra measured close to but above canopy are weakly affected by the canopy presence. Thus we do not expect that the relatively low observation level biases the overall N_2O flux level and that the comparison of instrumentation is affected. Also the effect on the instrumental noise and flux random uncertainty analysis is expected to be very limited through the influence on the co-variance functions. The positive impact of the close positioning of the system could be its higher sensitivity in detecting the low fluxes through higher concentration fluctuations expected (more) close to the source level.

Important characteristics of the instruments for performing the EC measurements are the response time and the noise level. The response times for CS-TDL, AR-CW-QCL and AR-P-QCL flux measurements systems were determined to be 0.12 and 0.07 and 0.08 seconds, respectively. The main factors affecting the response time of the closed-path EC system are the damping of fluctuations in the sampling line and the instrumental response. Since the flow rate of CS-TDL system was higher, it can be concluded that the response characteristics of other two instruments are superior. The response time of the EC system including LGR-CW-QCL was larger due to the laminar tube flow regime, but the instrumental response was not determined based on the current field measurements.

In order to understand drivers of exchange and inferring the broad average fluxes such as seasonal or annual sums by using some gap-filling methodologies it is important that the

exchange at shorter time scale can be distinguished from random variation. Therefore understanding of the random errors is important when working with low fluxes as is frequently the case with N_2O. At half-hour averaging time scale the flux estimates for AR-CW-QCL and LGR-CW-QCL instruments were very well correlated and showed good correspondence. Apart from high N_2O fluxes exceeding a few nmol m^{-2} s^{-1} during the high emission period, CS-TDL was not able to resolve the emission fluxes at half-hourly time scale. Therefore one can conclude that CS-TDL is not suitable for measuring such low fluxes if the aim is to resolve fluxes at hourly time scale and not the daily or longer averages.

Aerodyne AR-CW-QCL had the lowest noise level (around 0.12 ppb at 10 Hz sampling rate) compared to Los Gatos LGR-CW-QCL instrument (std of noise 0.60 ppb) and has therefore advantage in resolving low fluxes over short averaging periods. The noise level of AR-P-QCL was comparable to LGR-CW-QCL instrument but the old generation instrument Campbell CS-TDL suffered clearly from higher noise level (around 2 ppb). Huang et al. (2014) reported for the instrument similar to AR-CW-QCL the precision 0.066 ppb for 10 Hz. The value obtained by us was higher roughly by a factor of two. According to manufacturer the precision of LGR-CW-QCL is 0.1 ppb at 1 Hz averaging; at 10 Hz this would correspond to 0.32 ppb. We have determined again a median value roughly twice higher than this. Kroon et al. (2007) reported for the instrument similar to AR-P-QCL the precision value 0.5 ppb $Hz^{-1/2}$ (equivalent to 1.6 ppb at 10 Hz), whereas Neftel et al. (2007) and Eugster et al. (2007) report 0.3 ppb $Hz^{-1/2}$ (equivalent to 0.95 ppb at 10 Hz). Pihlatie et al. (2005) and Wang et al. (2013) report as the noise of instrument CS-TDL 1 ppb and 1.5 ppb (at 10 Hz), respectively. Under field conditions the instrumental noise can be somewhat higher compared to laboratory conditions where the instrumental characteristics are typically studied. Also the estimation method from the field records where the turbulent variation is superimposed by the instrumental noise can introduce some uncertainty. In summary, the observed instrumental noise characteristics for instruments compare well with the results reported by others and are useful in characterising instrumental performance.

The flux errors due to instrumental noise for the observation conditions prevailing at the site were determined to be around 10^{-2} nmol m^{-2} s^{-1} for AR-CW-QCL, $4x10^{-2}$ nmol m^{-2} s^{-1} for LGR-CW-QCL and AR-P-QCL and 0.15 nmol m^{-2} s^{-1} for CS-TDL. Based on half-hour as well as long-term flux comparison, the best correspondence was observed between the systems with new generation instruments AR-CW-QCL and LGR-CW-QCL, of which the

former has the advantage in detecting lower fluxes at half-hourly averaging basis (lower noise level).

The signal noise of the anemometer used by the UH (USA1 by METEK) was determined to be 0.037 m s^{-1} at 10 Hz sampling frequency for vertical wind speed component. The noise level of the anemometer employed by the UEF was similar. The flux error due to anemometer's noise for the observation conditions prevailing at the site during the period DOY 206-271 (the period for the statistics presented in Fig. 6) were determined to be around $2x10^{-3}$ nmol m^{-2} s^{-1} (the median value). This was much less than the respective flux error around 10^{-2} nmol m^{-2} s^{-1} for the instrument AR-CW-QCL, which had the lowest noise level 0.012 ppb (median value) of all instruments compared. Therefore the assumption that the anemometer's noise affects flux detection much less than the gas analysers was well justified.

The chamber techniques are widely used to measure the soil N_2O exchange. The traditional way to perform chamber measurements is to determine the gas concentration at several time moments during the chamber operation (called deployment time DT). In such data collection the sources of uncertainty are the imprecesion related to gas sampling (either manual or automatic) as well as instrumental uncertainty (e.g. Venterea et al., 2009), leading to a measurement precision which is called a detection limit of chamber based flux measurement system. Neftel et al. (2007) report a flux detection limit of about 0.23 nmol m^{-2} s^{-1} for their chamber system with DT of 10 min and the concentration sampling interval of 1 min. The measurement cycle of the system was however two hours. Wang et al. (2013) found for their automatic and manual chamber systems detection limits of about 5 μg m^{-2} h^{-1} (0.05 nmol m^{-2} s^{-1}) for hourly DT. Their instrument precision was high, around 0.4% relative to ambient N_2O concentration. By using the methodology and scaled results presented by Parkin et al. (2012), we estimated the flux detection limit of a chamber system with assumed chamber height of 0.5 m, the area of 0.25 m^2, deployment time 30 min and instrumental precision as high as 0.1% to be 0.03 nmol m^{-2} s^{-1}. It has to be noted that the flux detection limit of the chamber systems depends on several factors such as the type of the chamber and respective sampling method, the precision of the instrument, chamber dimensions and operation time (DT). Nevertheless, the obtained result is well comparable with the EC systems. The random error of N_2O fluxes for 30 min averaging time for the instrument with lowest noise, the AR-CW-QCL instrument, was found to be 0.036 nmol m^{-2} s^{-1} (the median value). Note that here we compare the flux detection limit of the chamber based systems (which accounts for all

possible sources of uncertainty) with the total stochastic error of the EC fluxes. The results are of the same magnitude.

In this study we followed the methodology proposed by Mauder et al. (2013) in quantification of the random errors in EC fluxes, i.e. the stochastic error and the error due to instrumental noise in flux. The relative random errors obtained in our study were much larger than the respective errors reported by Mauder et al. (2013) for CO_2 measurements, evidencing that the importance of random errors depends on the trace gas of interest via instrumental precision and the flux magnitude ratio. Kroon et al. (2010a) focus on the evaluation of the EC flux measurements of CH_4 and N_2O specifically. They observed over a dairy farm site the fluxes in the range of 15 to 110 ng N m^{-2} s^{-1} (0.5 to 4 nmol m^{-2} s^{-1}), which they classified from low to high flux classes. They performed calibration of the instrument similar to our AR-P-QCL weekly and considered the respective uncertainty random over longer periods of time. Kroon et al. (2010a) reported the average daily and monthly flux relative uncertainties of 31 and 7%, respectively. In our study the N_2O fluxes were typically much smaller (excluding the fertilisation episode), around 0.1 to 0.3 nmol m^{-2} s^{-1}. We measured with the similar instrument 36% lower fluxes than obtained by AR-CW-QCL over the period DOY 110-181 and 13% lower fluxes than obtained by two new generation instruments over the period DOY 206-271. Evidently our measurements performance was affected by unidentified error source being systematic in nature. In evaluation of the annual balances of CH_4 and N_2O fluxes over managed fen meadow Kroon et al. (2010b) made an assumption that the uncertainty in EC fluxes was random and was neglected in evaluation of long term averages. In our results this assumption was violated and we suggest that all possible systematic error sources should be considered very carefully in planning, implementing and evaluating the flux measurements of trace gases.

In analysing the random errors of the fluxes Kroon et al. (2010a) assumed that the flux error due to instrumental precision in concentration measurement was negligible. We observed that this was not necessarily the case for N_2O when low flux levels were measured and demonstrated that the method originally proposed by Lenchow et al. (2000) to determine instrumental noise variance worked well in the field conditions over a long period of time.

5 Conclusions

The new instruments based on continuous wave quantum cascade lasers, AR-CW-QCL and LGR-CW-QCL, were stable throughout of the campaign in terms of determination of absolute concentrations as well as obtaining very close cumulative fluxes.

The older instruments CS-TDL and AR-P-QCL measured systematically different fluxes over subperiods of the campaign up to 29% and -36%, respectively, compared to the new instruments based on CW-QCL-s, whereas the systematic differences did not prevail throughout the campaign. The reasons for the systematic differences were not identified. We suggest that special emphasis should be on the instrumental stability and correcting procedures that can affect systematically the accuracy of measured fluxes when conducting long-term measurements of prevailingly low fluxes.

The lowest noise level was determined for AR-CW-QCL (0.12 ppb at 10 Hz sampling rate) and the highest for the old generation instrument CS-TDL (precision 2 ppb at 10 Hz sampling rate). During the period DOY 206-272, when all instruments were operational, the lower quantile/median/upper quantile statistics of the fluxes measured by AR-CW-QCL instrument were 0.008/0.11/0.31 nmol m^{-2} s^{-1} as.

The random errors of fluxes originate from the stochastic nature of turbulence (one-point sampling over limited time interval). Additionally, the instrumental noise contributes to the random flux error. The median values for flux errors during the period DOY 206-272 (error due to instrumental noise / the total error) were detected for the instruments as follows: for CS-TDL 0.155/0.255, AR-CW-QCL 0.010/0.036, LGR-CW-QCL 0.046/0.065, and AR-P-QCL 0.031/0.068 nmol m^{-2} s^{-1}. These error statistics indicate that (i) the major component of the flux random error source is the instrumental noise, and (ii) the flux errors for CS-TDL are dominantly larger than the flux magnitude and only in case of AR-CW-QCL the flux error due to instrumental noise can be said to be much smaller than the typical flux value.

The following fractions of fluxes were smaller than the stochastic flux error: in case of CS-TDL 47%, AR-CW-QCL 15%, LGR-CW-QCL 28%, and AR-P-QCL 30%. We conclude that apart from AR-CW-QCL large fraction of the fluxes were within the error magnitude of single half-hour observations.

With the new generation analyzers based on continuous-wave QCL-s N_2O fluxes can be measured with the EC at locations where the fluxes are small, well below the detection limit of older instruments (CS-TDL for instance). According to our analysis the new instruments enable to attain the flux precision as good as the precision of the modern chamber systems. Thus the new instruments open up the possibility to study N_2O exchange at new ecosystems, broadening the scientific perspectives.

Acknowledgements

This work was supported by the Academy of Finland (project No. 118780 and 127456). ICOS (271878), ICOS-Finland (281255) and ICOS-ERIC (281250), DEFROST Nordic Centre of Excellence and InGOS EU are gratefully acknowledged for funding this work. This work was also supported by institutional research funding (IUT20-11) of the Estonian Ministry of Education and Research. The UEF part of the research work was supported by the funding from the UEF infrastructure funding, Academy of Finland FidiPro program (PIs - Profs Pertti Martikainen and Seppo Kellomäki) and the Ministry of Agriculture and Forestry, Finland.

References

Aubinet, M., Grelle, A., Ibrom, A., Rannik, Ü., Moncrieff, J., Foken, T., Kowalski, A.S., Martin, P.H., Berbigier, P., Bernhofer, Ch., Clement, R., Elbers, J., Granier, A., Grünwald, T., Morgenstern K., Pilegaard K., Rebmann C., Snijders W., Valentini R. and Vesala, T.: Estimates of the annual net carbon and water exchange of European forests: the EUROFLUX methodology, Advances Ecol, Res., 30, 113-175, 2000.

Christensen, S., Ambus, P., Arah, J. R., Clayton, H., Galle, B., Griffith, D. W. T., Hargreaves, K. J., Klemedtsson, L., Lind, A. M., Maag, M., Scott, A., Skiba, U., Smith, K. A., Welling, M., and Wienhold, F. G.: Nitrous oxide emissions from an agricultural field: comparison between measurements by flux chamber and micrometeorological techniques, Atmos. Environ., 30(24), 4183-4190, 1996.

Eugster, W., Zeyer, K., Zeeman, M., Michna, P., Zingg, A., Buchmann, N., and Emmenegger, L.: Methodical study of nitrous oxide eddy covariance measurements using quantum cascade laser spectrometery over a Swiss forest, Biogeosciences, 4, 927-939, 2007.

Famulari, D., Nemitz, E., Di Marco, C., Phillips, G.J., Thomas, R., House, E. and Fowler, D.: Eddy-Covariance measurements of nitrous oxide fluxes above a city, Agric. For. Meteorol. 150, 786-793, 2010.

FAO, World reference base for soil resources 2006. World soil resources reports 103, Rome, Italy, 2006.

Finkelstein, P. L. and Sims, P. F.: Sampling error in eddy correlation flux measurements, J. Geophys. Res., 106, NO. D4, 3503-3509, 2001.

Finnigan, J. J., Clement, R., Malhi, Y., Leuning, R., and Cleugh, H. A.: A re-evaluation of longterm flux measurement techniques – Part I: Averaging and coordinate rotation, Bound.-Layer Meteorol., 107, 1–48, 2003.

Flechard, C., Neftel, A., Jocher, M., and Amman, C.: Bi-directional soil-atmosphere N2O exchange over two mown grassland systems with contrasting management practices, Global Change Biology, 11, 2114-2127, 2005.

Foken, T., Leuning, R., Oncley, S. R., Mauder, M., and Aubinet, M.: Corrections and Data Quality Control, in Eddy Covariance. A Practical Guide to Measurement and Data Analysis, Springer Science+Business Media B.V., 85-131, DOI 10.1007/978-94-007-2351-1, 2012.

Foken, T. and Wichura, B.: Tools for quality assessment of surface-based flux measurements, Agr. Forest Meteorol., 78, 83–105, 1996.

Horst, T.W.: A simple formula for attenuation of eddy fluxes measured with first-order-response scalar sensors, Bound.-Layer Meteorol., 82: 219-233, 1997.

Huang, H., Wang, J., Hui, D., Miller, D. R., Bhattarai, S., Dennis, S., Smart, D., Sammis, T., and Reddy, K. C.: Nitrous oxide emissions from a commercial cornfield (Zea mays) measured using the eddy-covariance technique, Atmos. Chem. Phys. Discuss., 14, 20417–20460, doi:10.5194/acpd-14-20417-2014, 2014.

Järvi, L., Nordbo, A., Rannik, Ü., Haapanala, S., Riikonen, A., Mammarella, I., Pihlatie, M. and Vesala, T.: Urban nitrous oxide fluxes measured using the eddy-covariance technique in Helsinki, Finland. Boreal Env. Res., 19 (suppl. B), 108–121, 2014.

Kaimal, J. C., and Finnigan, J. J.: Atmospheric Boundary Layer Flows. Their Structure and Measurement, Oxford University Press, New York, 1994.

Kaimal, J. C., Wyngaard, J. C., Izumi, Y., and Cot`e, O. R.: Spectral characteristics of surface layer turbulence, Q. J. Roy. Meteor. Soc., 98, 563–589, 1972.

Korhonen, J. F. J., Pihlatie, M., Pumpanen, J., Aaltonen, H., Hari, P., Levula, J., Kieloaho, A.-J., Nikinmaa, E., Vesala, T., and Ilvesniemi, H.: Nitrogen balance of a boreal Scots pine forest, Biogeosciences, 10, 1083-1095, doi:10.5194/bg-10-1083-2013, 2013.

Kroon, P.S., Hensen, A., Jonker, H.J.J., Ouwersloot, H.G., Vermeulen, A.T., Bosveld, F.C.: Uncertainties in eddy covariance flux measurements assessed from CH4 and N2O

observations, Agricultural and Forest Meteorology, 150, 806–816, doi:10.1016/j.agrformet.2009.08.008, 2010a.

Kroon, P.S., Hensen, A., Jonker, H.J.J., Zahniser, M.S., van't Veen, W.H., Vermeulen, A.T.: Suitability of quantum cascade laser spectroscopy for CH4 and N2O eddy covariance flux measurements, Biogeosciences, 4, 715-728, 2007.

Kroon, P . S ., Schrier - Uijl, A. P., Hensen, A. , Veenendaal, E. M., and Jonker, H. J. J.: Annual balances of CH_4 and N_2O from a managed fen meadow using eddy covariance flux measurements, European Journal of Soil Science, 61, 773–784, 2010b.

Launiainen, S., Vesala, T., Mölder, M., Mammarella, I., Smolander, S., Rannik, Ü., Kolari, P., Hari, P., Lindroth, A., and Katul, G.G.: Vertical variability and effect of stability on turbulence characteristics down to the floor of a pine forest, Tellus B, 59, 919–936, DOI: 10.1111/j.1600-0889.2007.00313.x, 2007.

Laville, P., Henault, C., Renault, P., Cellier, P., Oriol, A., Devis, X., Flura, D., and Germon, J. C.: Field comparison of nitrous oxide emission measurements using micrometeorological and chamber methods, Agronomie, 17, 375-388, 1997.

Lee, B. H., Wood, E. C., Zahniser, M. S., McManus, J. B., Nelson, D. D., Herndon, S. C., Santoni, G. W., Wofsy, S. C., and Munger, J. W.: Simultaneous measurements of atmospheric HONO and NO_2 via absorption spectroscopy using tunable mid-infrared continuous-wave quantum cascade lasers, Appl. Phys. B, 102:417-423, 2011.

Lenschow, D.H. and Kristensen, L.: Uncorrelated noise in turbulence measurements, J. Atm. Oceanic Technol., 2, 68-81, 1985.

Lenschow, D.H., Mann, J. and Kristensen, L.: How long is long enough when measuring fluxes and other turbulence statistics?, J. Atm. Oceanic Technol., 18, 661-673, 1994.

Lenschow, D., Wulfmeyer, V. and Senff, C.: Measuring second- through fourth-order moments in noisy data, J. Atmos. Ocean. Technol., 17, 1330-1347, 2000.

Mammarella, I., Launiainen, S., Gronholm, T., Keronen, P., Pumpanen, J., Rannik, Ü., Vesala, T.: Relative humidity effect on the high frequency attenuation of water vapour flux measured by a closed-path eddy covariance system, J. Atmos. Ocean. Tech., 26(9), 1856-1866, 2009.

Mammarella, I., Werle, P., Pihlatie, M., Eugster, W., Haapanala, S., Kiese, R., Markkanen, T., Rannik, Ü., and Vesala, T.: A case study of eddy covariance flux of N_2O measured within forest ecosystems: quality control and flux error analysis, Biogeosciences, 7, 427-440, 2010.

Massman, W.: A simple method for estimating frequency response corrections for eddy covariance systems, Agr. Forest Meteorol., 104, 185–198, 2000.

Mauder, M., Cuntz, M., Druee, C., Graf, A., Rebmann, C., Schmid, H. P., Schmidt, M. and Steinbrecher, R.: A strategy for quality and uncertainty assessment of long-term eddy covariance measurements, Agric. For. Meteorol., 169, 122-135, doi:10.1016/j.agrformet.2012.09.006, 2013.

McManus, J.B., Nelson, D.D., Shorter, J.H., Jiménez, R., Herndon, S., Saleska, S., and Zahniser, M.S.: A high precision pulsed QCL spectrometer for measurements of stable isotopes of carbon dioxide, J. Modern Optics, 52, 2309-2321, 2005.

Moncrieff, J. B., Massheder, J. M., de Bruin, H., Elbers, J., Friborg, T., Heusinkveld, B., Kabat, P., Scott, S., Soegaard, H., and Verhoef, A., : A system to measure surface fluxes of momentum, sensible heat, water vapour and carbon dioxide, J. Hydrol., 188-189, 589-611, 1997.

Moore, C. J.,: Frequency response corrections for eddy correlation systems, Boundary-Layer Meteorol., 37, 17-35, 1986.

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, Agr. Forest Meteorol., 150, 775–785, 2010.

Neftel, A., Flechard, C., Ammann, C., Conen, F., Emmenegger, L. and Zeyer, K.: Experimental assessment of N_2O background fluxes in grassland systems, Tellus, 59B, 470–482, 2007.

Nelson, D. D., Shorter, J. H., McManus, J. B., and Zahniser, M. S.: Sub-part-per-billion detection of nitric oxide in air using a thermoelectrically cooled mid-infrared quantum cascade laser spectrometer, Appl. Phys. B, 75, 343-350, 2002.

Nicolini, G., Castaldi, S., Fratini, G., Valentini, R.: A literature overview of micrometeorological CH_4 and N_2O flux measurements in terrestrial ecosystems, Atmospheric Environment, 81, 311-319, doi: 10.1016/j.atmosenv.2013.09.030, 2013.

Parkin, T.B., Venterea, R. T., and Hargreaves, S. K.: Calculating the Detection Limits of Chamber-based Soil Greenhouse Gas Flux Measurements, Journal of Environmental Quality, 41, 705–715, doi:10.2134/jeq2011.0394, 2012.

Peltola, O., Hensen, A., Helfter, C., Belelli Marchesini, L., Bosveld, F. C., van den Bulk, W. C. M., Elbers, J. A., Haapanala, S., Holst, J., Laurila, T., Lindroth, A., Nemitz, E., Röckmann, T., Vermeulen, A. T., and Mammarella, I.: Evaluating the performance of

commonly used gas analysers for methane eddy covariance flux measurements: the InGOS inter-comparison field experiment, Biogeosciences, 11, 3163-3186, doi:10.5194/bg-11-3163-2014, 2014.

Pihlatie, M., Rinne, J., Ambus, P., Pilegaard, K., Dorsey, J.R., Rannik, Ü, Markkanen, T., Launiainen, S., and Vesala, T. : Nitrous oxide emissions from a beech forest floor measured by eddy covariance and soil enclosure techniques, Biogeosciences, 2, 377-387, 2005.

Pihlatie, M., Kiese, R., Brüggemann, N., Butterbach-Bahl, K., Kieloaho, A.-J., Laurila, T., Lohila, A., Mammarella, I., Minkkinen, K., Penttilä, T., Schönborn, J., and Vesala, T.: Greenhouse gas fluxes in a drained peatland forest, Biogeosciences, 7, 1715-1727, doi:10.5194/bg-7-1715-2010, 2010.

Pirinen, P., Simola, H., Aalto, J., Kaukoranta, J., Karlsson, P. and Ruuhela, R.: Tilastoja Suomen ilmastosta 1981 - 2010, Finnish Meteorological Institute, Helsinki, 2012.

Provencal, R., Gupta, M., Owano, T. G., Baer, D. S., Ricci, K. N., O'Keefe, A., and Podolske, J. R., Cavity-enhanced quantum-cascade laser-based instrument for carbon monoxide measurements, Appl. Optics, 44, 6712-6717, 2005.

Rannik, Ü., Kolari, P., Vesala, T., and Hari, P.: Uncertainties in measurement and modelling of net ecosystem exchange of a forest ecosystem at different time scales, Agric. Forest. Meteorol., 138, 244-257, 2006.

Rannik, Ü., Mammarella, I., Aalto, P., Keronen, P., Vesala, T., Kulmala, M., Long-term particle flux observations Part I: Uncertainties and time-average statistics. Atmospheric Environment, 43, 3431-3439, 2009.

Rannik, Ü. and Vesala, T.: Autoregressive filtering versus linear detrending in estimation of fluxes by the eddy covariance method, Bound.-Layer Meteorol., 91, 259–280, 1999.

Scanlon, T. M., and Kiely, G.: Ecosystem-scale measurements of nitrous oxide fluxes for an intensely grazed, fertilized grassland, Geophys. Res. Letter, 30 (16), 1852, doi:10.1029/2003GL017454, 2003.

Schotanus, P., Nieuwstadt, F. T. M., and Debruin, H. A. R.: Temperature-measurement with a sonic anemometer and its application to heat and moisture fluxes, Bound.-Layer Meteorol., 26, 81–93, 1983.

Smith, K. A., Clayton, H., Arah, J. R. M., Christensen, S., Ambus, P., Fowler, D., Hargreaves, K. J., Skiba, U., Harris, G. W., Wienhold, F. G., Klemedtsson, L., and Galle, B.: Micrometeorological and chamber methods for measurement of nitrous oxide fluxes

between soils and the atmosphere: Overview and conclusions, J. Geophys. Res., 99, D8 16541-16548, 1994.

Venterea, R.T., Spokas, K.A., Baker, J.M.: Accuracy and Precision Analysis of Chamber-Based Nitrous Oxide Gas Flux Estimates, Soil Sci. Soc. Am. J., 73, 1087-1093, doi:10.2136/sssaj2008.0307, 2009.

Vickers, D. and Mahrt, L.: Quality control and flux sampling problems for tower and aircraft data, J. Atmos. Ocean. Tech., 14, 512–526, 1997.

Wang, K, Zheng, X., Pihlatie, M., Vesala, T., Liu, C., Haapanala, S., Mammarella, I., Rannik, Ü., Liu, H.: Comparison between static chamber and tunable diode laser-based eddy covariance techniques for measuring nitrous oxide fluxes from a cotton field, Agric. Forest Meteorol., 171–172, 9–19, 2013.

Werle, P., Muecke, R. and Slemr, F.: The limits of signal averaging in atmospheric trace gas monitoring by tunable diode-laser absorption spectroscopy, Appl. Phys. B 57, 131-139, 1993.

Wienhold, F. G., Frahm, H., and Harris, G. W.: Measurements of N2O fluxes from fertilized grassland using a fast response tunable diode laser spectrometer, J. Geophys. Res., 99, D8 16557-16567, 1994.

Zahniser, M. S., Nelson, D. D., McManus, J. B., Hern-don, S. C., Wood, E. C., Shorter, J. H., Lee, B. H., Santoni, G. W., Jimenez, R., Daube, B. C., Park, S., Kort, E. A., and Wofsy, S. C.: Infrared QC laser applications to field measurements of atmospheric trace gas sources and sinks in environmental research: enhanced capabilities using continuous wave QCLs, Proc. SPIE, 7222, doi: 10.1117/12.815172, 2009.

Table 1. Instrumental characteristics. Experimental precision values are based on flux measurements during the period DOY 206-271 (period II). TDL – Tunable Diode Laser; CW-QCL - Continuous-Wave Quantum Cascade Laser; P-QCL – Pulsed QCL.

Instrument model	**TGA100A**	**CW-TILDAS-CS**	**N2O/CO-23d**	**QC-TILDAS-76-CS**
Manufacturer	Campbell Scientific Inc.	Aerodyne Research Inc.	Los Gatos Research Inc.	Aerodyne Research Inc.
Acronym used in current study	CS-TDL	AR-CW-QCL	LGR-CW-QCL	AR-P-QCL
Measured species	N_2O	N_2O, H_2O, CO	N_2O, H_2O, CO	N_2O, CO_2, H_2O
Sample cell volume (ml)	480	500		500 (76 m path length)
Sample cell pressure (hPa)	50	53	117	53
Spectroscopic correction coefficient b	0.00 (drier used in sampling line)	0.39	0.00 (built-in correction by the instrument)	0.0235
Precision, 10 Hz noise std, $P_{10}/P_{50}/P_{90}$ this study (ppb)	1.89/1.98/2.1	0.12/0.12/0.14	0.46/0.60/0.78	0.43/0.46/0.51

Table 2. Eddy covariance measurements setup, flux calculation and quality screening parameters

Instrument	**CS-TDL**	**AR-CW-QCL**	**LGR-CW-QCL**	**AR-P-QCL**
Sampling height (m)	2.2/2.4	2.2/2.4	2.4	2.0/2.5
Horizontal separation[1] (m)	0.05	0.05	0.07	0.1
Tube inner diameter (mm)	4	4	8	4
Tube length (m)	17.8	16	16	8.5
Flow rate (LPM)	17	13.2	11.6	13.5
Lag time from tube flow (s)	0.79	0.91	4.2	0.48
Lag time window used in flux calculation (s)	1.0+-0.0	1.0 +-0.0	1.0+-0.0[2]	1.0+-0.8[3]
Time constant used in spectral corrections (s)	0.12	0.07	0.26	0.15

[1]Refers to separation of the sampling inlet from the center position of the sonic anemometer. Vertical separation was 0.1 m for all instruments.

[2]Prior to flux calculation concentration records of LGR-CW-QCL were syncronised with AR-CW-QCL outputs.

[3]The lag time window was used to determine the lag time for CO_2, which was assigned as the lag time for N_2O.

Table 3. Statistics of spectral corrections of fluxes as % of raw uncorrected fluxes: lower percentile/median/upper percentile. Based on flux measurements during the period DOY 206-271 (period II) and data classified as qualified (Table 4). Day time was defined by the elevation of sun higher than zero and night time lower than zero, respectively. Statistics were derived for data when measurements were available for all four instruments.

	CS-TDL	**AR-CW-QCL**	**LGR-CW-QCL**	**AR-P-QCL**
All data	4.0/6.2/10.2	2.4/3.6/6.0	6.9/12.3/20.0	4.5/7.3/14.8
Daytime data	4.0/6.1/9.8	2.6/3.6/5.8	6.9/12.0/18.5	4.5/6.9/10.5
Night data	3.6/6.3/11.3	2.2/3.6/6.4	6.7/12.9/22.3	4.5/7.7/20.2

Table 4. Average fluxes (nmol m^{-2} s^{-1}) ± random error of the average. Period I DOY 110-181 (20.04-30.06.2011), Period II DOY 206-271 (25.07-28.09.2011), Period III DOY 272 – 324 (29.09-20.11.2011). % data available represents the fraction of half-hour periods when data from all 3 (or 4) instruments was available (data from wind direction interval 50-110° excluded), relative to full time period length. Averaging of fluxes for each instrument was performed only for data if measurements were available for all instruments used in respective period. No gap filling was used.

	% data available	**% data qualified (out of available)**	**# 30 min periods averaged**	**CS-TDL**	**AR-CW-QCL**	**LGR-CW-QCL**	**AR-P-QCL**
Period I	69.2	75.2	1797	0.931 ±0.018	0.870 ±0.009		0.560 ±0.011
Period II	55.0	79.4	1383	0.183 ±0.010	0.146 ±0.006	0.138 ±0.007	0.124 ±0.003
Period III	61.4	78.2	1220		0.067 ±0.002	0.057 ±0.002	0.058 ±0.003

Table 5. Average micrometeorological conditions during the experimental periods. Period I DOY 110-181 (20.04-30.06.2011), Period II DOY 206-271 (25.07-28.09.2011), Period III DOY 272 – 324 (29.09-20.11.2011). Day time was defined by the elevation of sun higher than zero and night time lower than zero, respectively. Average latent heat fluxes were determined from IRGA measurements.

	Temperature	**Air rel. humidity, %**	**Wind speed, m s^{-1}**	**Friction velocity, m s^{-1}**	**Sensible heat flux, W m^{-2}**	**Latent heat flux, W m^{-2}**
Day, I	11.6	62.9	2.21	0.28	27.5	78.9
Night, I	6.5	78.3	1.34	0.14	-20.2	8.1
Day, II	15.3	75.2	1.35	0.26	9.7	109.3
Night, II	11.2	90.3	1.06	0.17	-18.6	10.1
Day, III	6.1	85.0	1.46	0.29	-10.8	41.5
Night, III	4.8	90.6	1.21	0.23	-23.5	11.5

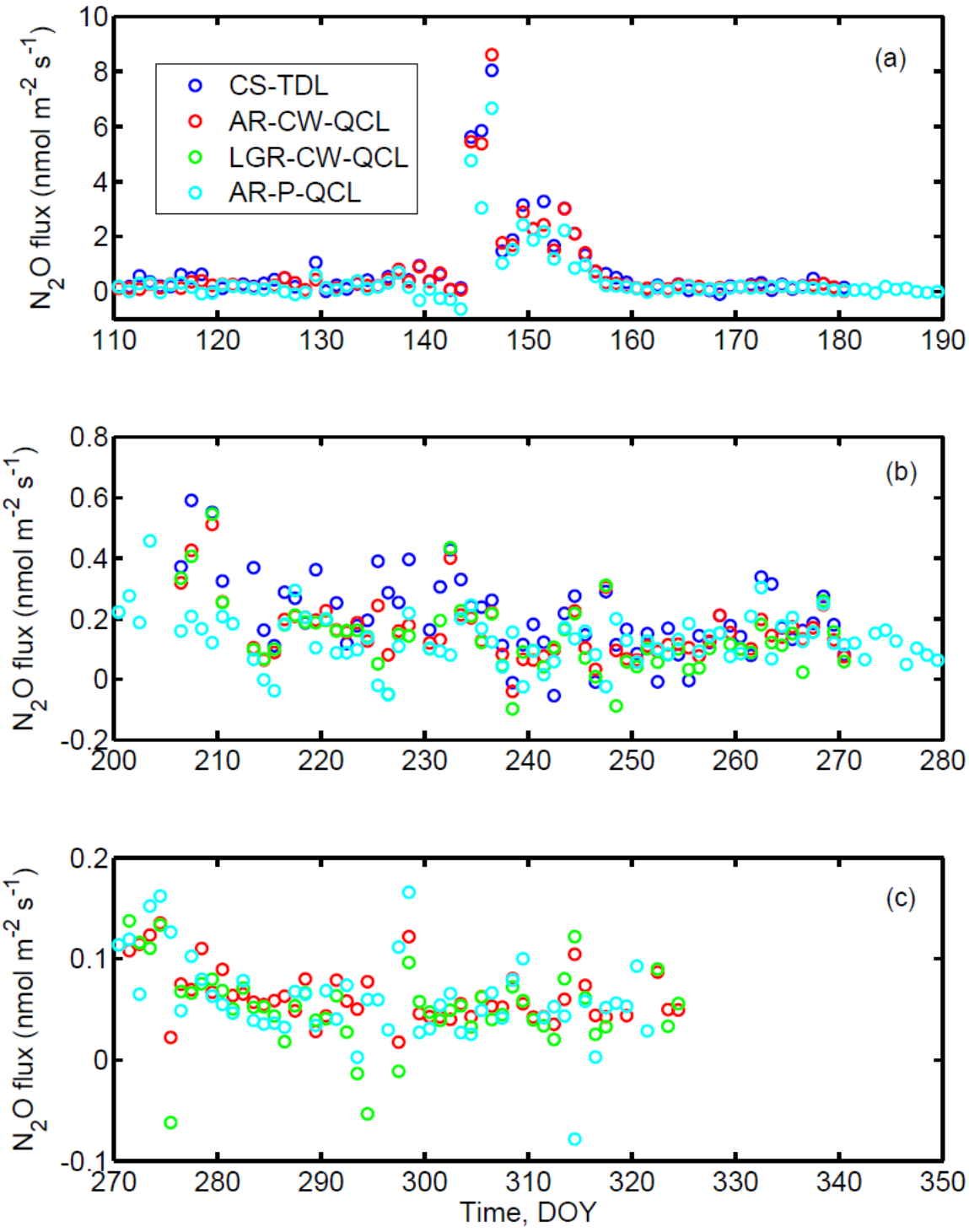

Figure 1. Daily average fluxes for four instruments containing period I DOY 110-181 (a), period II DOY 206-271 (b) and period III DOY 272-324 (c). No gap-filling was used in calculation of daily average fluxes.

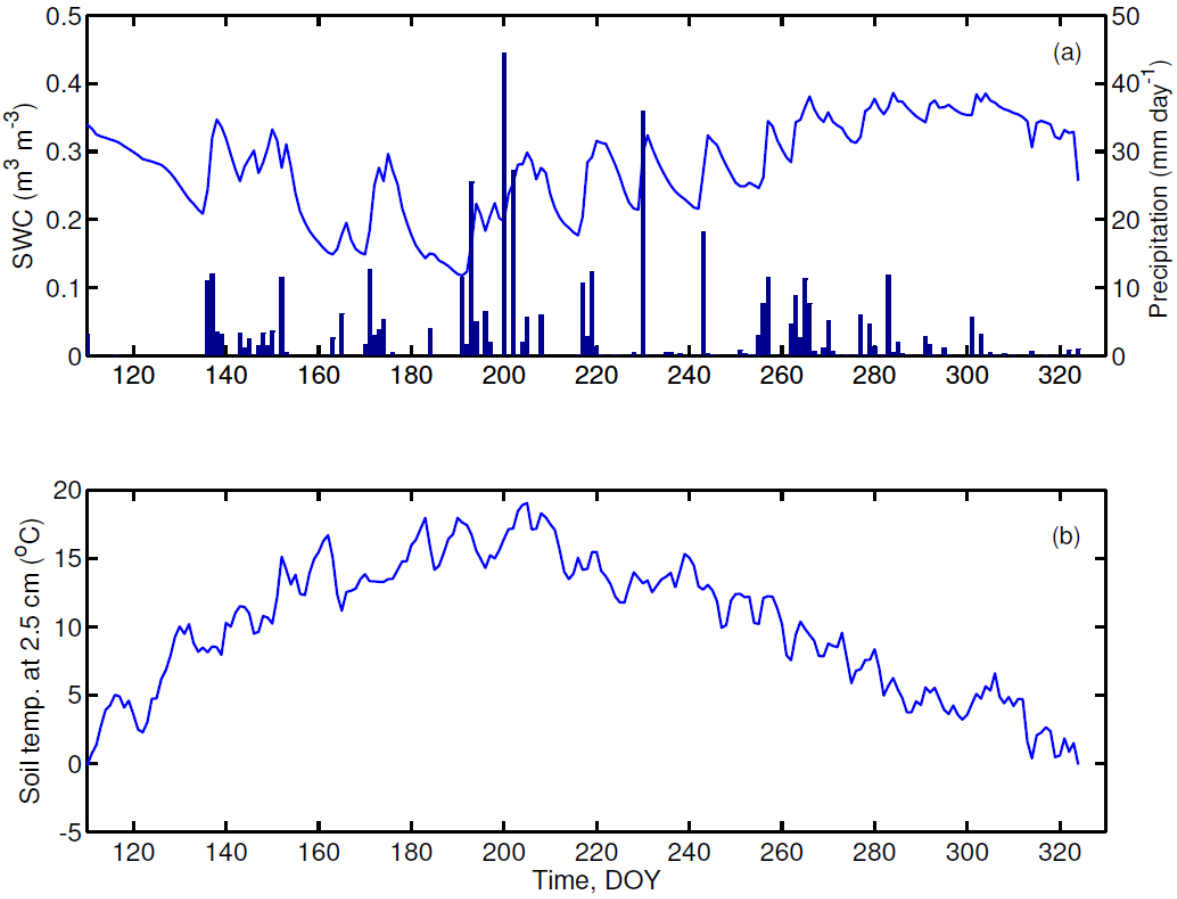

Figure 2. Soil water content (SWC) at 2.5 cm depth and precipitation (a) and soil temperature at 2.5 cm depth (b) during the measurement campaign.

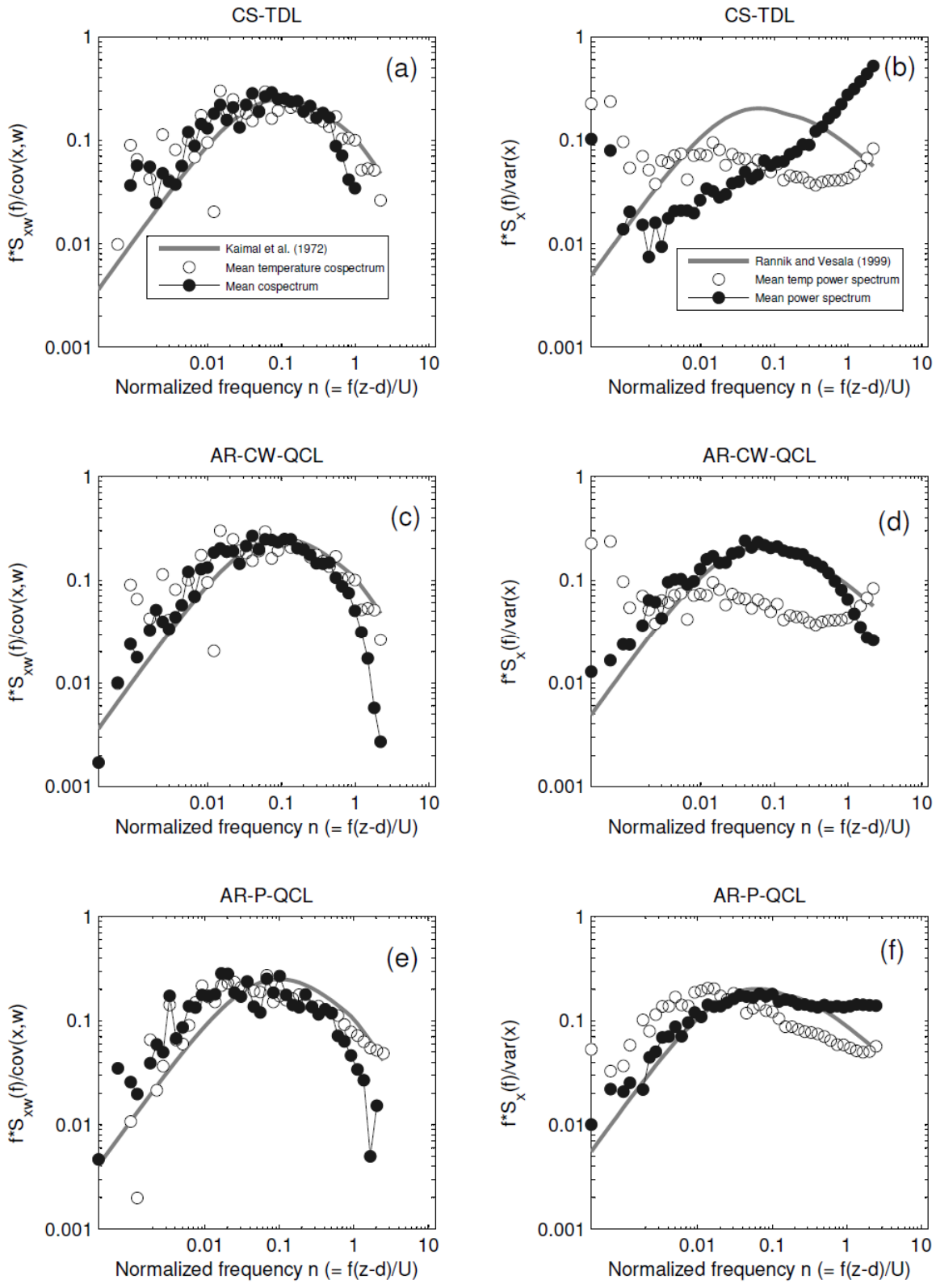

Figure 3. Normalised co-spectra (left panels) and spectra (right panels) of N_2O measurements by instruments CS-TDL (upper panels a, b), AR-CW-QCL (middle panels c, d) and AR-P-QCL (lower panels e, f) during the high flux period, DOY 146 (26.05.2011) 7:00 to 13:00 EET. The RCG crop was about 0.4 m tall during the given period.

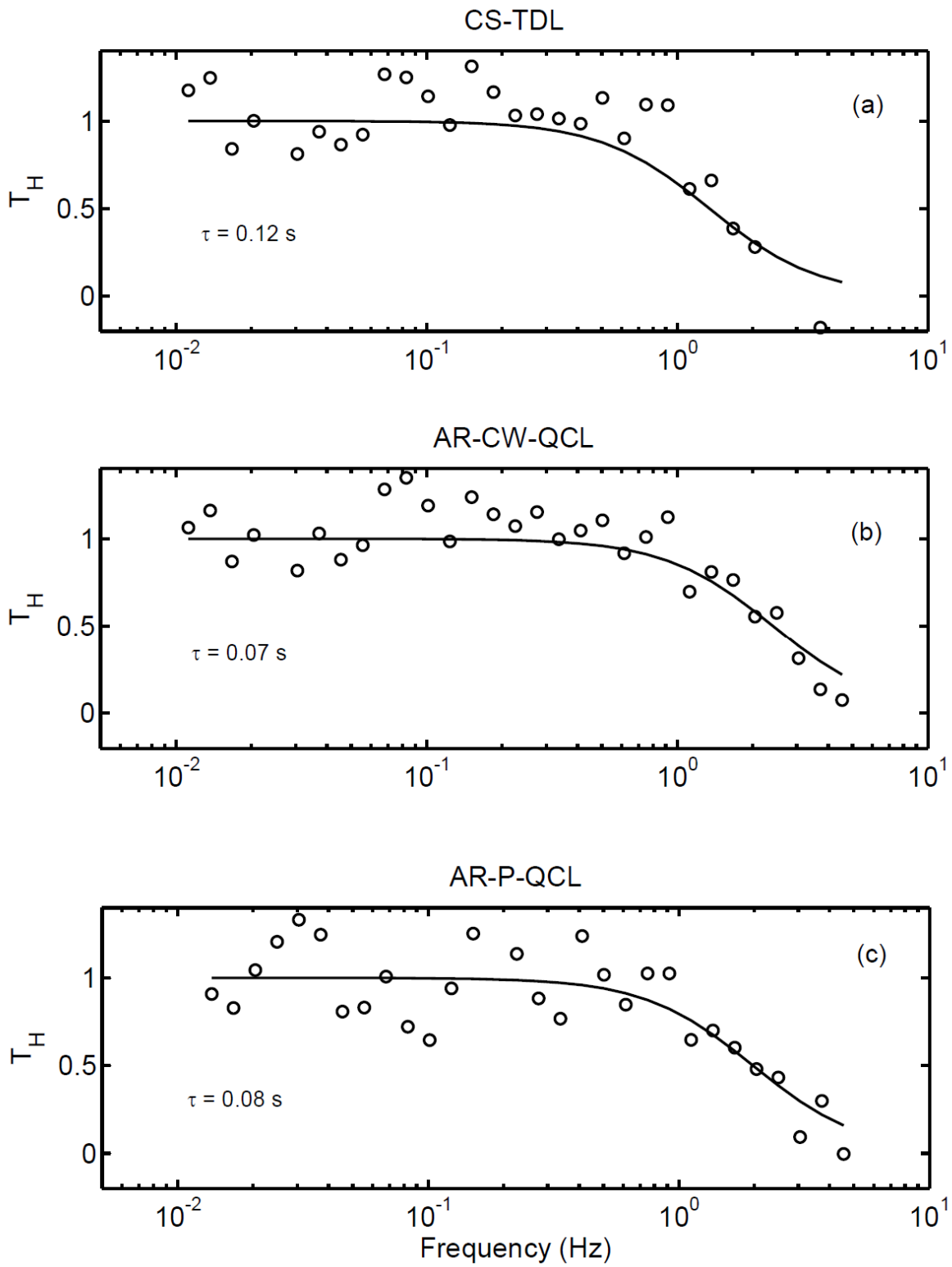

Figure 4. Co-spectral transfer functions derived for CS-TDL (a), AR-CW-QCL (b) and AR-P-QCL (c) from the temperature and N_2O co-spectra presented in Fig. 2.

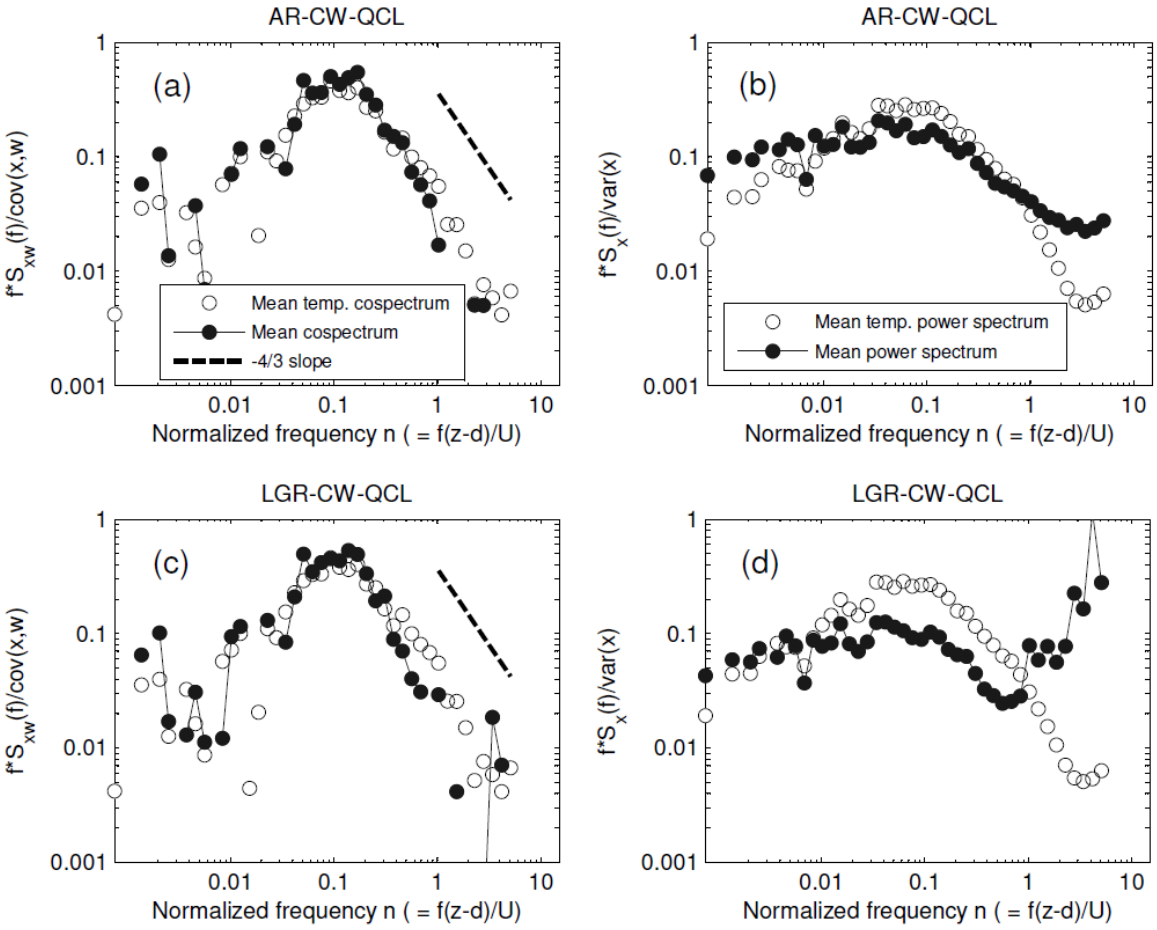

Figure 5. Normalised co-spectra (left panels) and spectra (right panels) of N_2O measurements by instruments AR-CW-QCL (upper panels a, b) and LGR-CW-QCL (lower panels c, d) during the period DOY 216 (04.08.2011) 00:30 to 4:00 EET. The RCG crop was about 1.8 m tall during the given period.

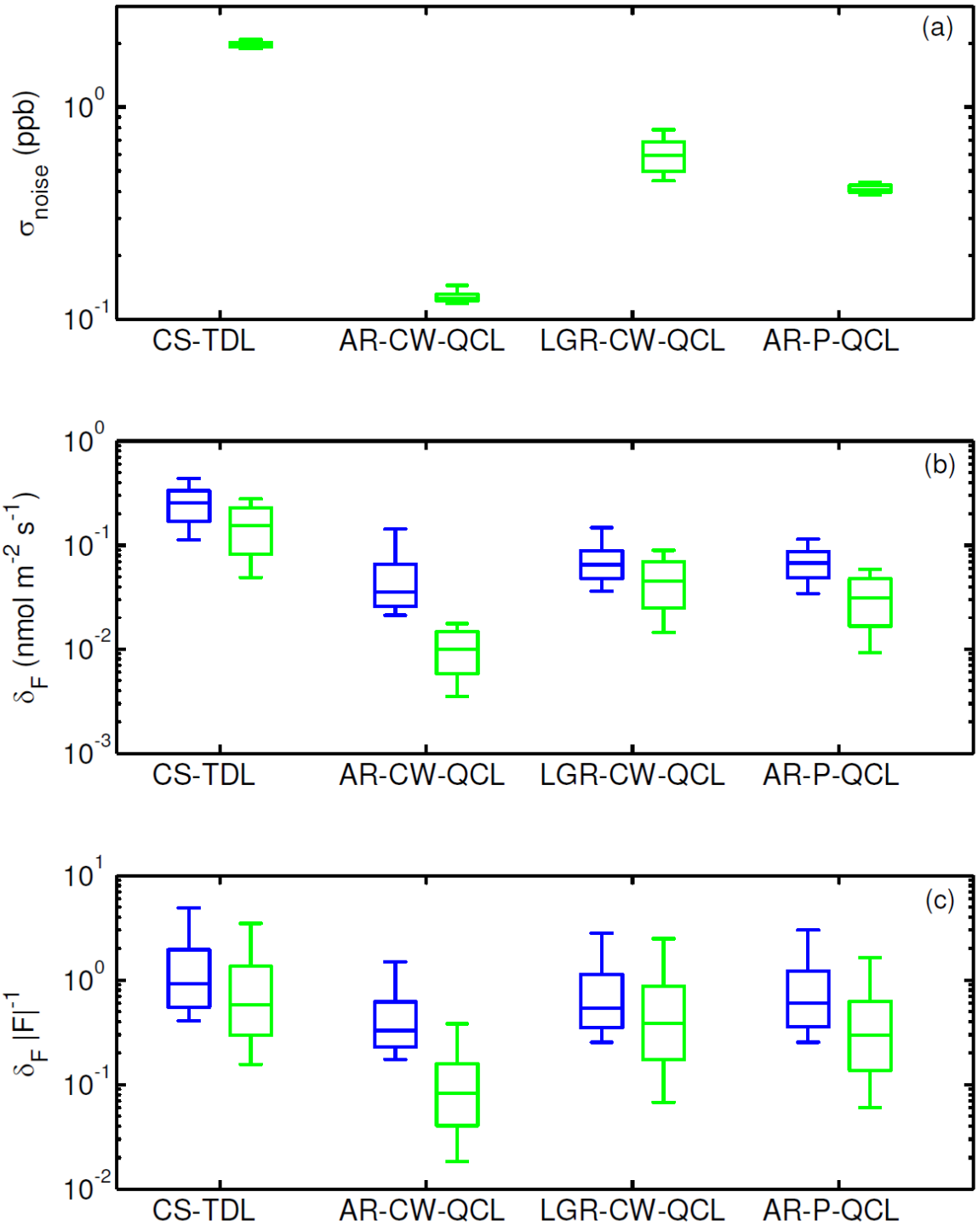

Figure 6. (a) Instrumental noise, presented as one standard deviation of the noise at 10 Hz frequency, (b) N_2O flux random error (blue) and flux random error due to instrumental noise (green) statistics; (c) the same as (b) but for relative fluxes. The boxplots present the lower and upper percentiles, quartiles and median values of the distributions. Based on flux measurements during the period DOY 206-271 (period II).

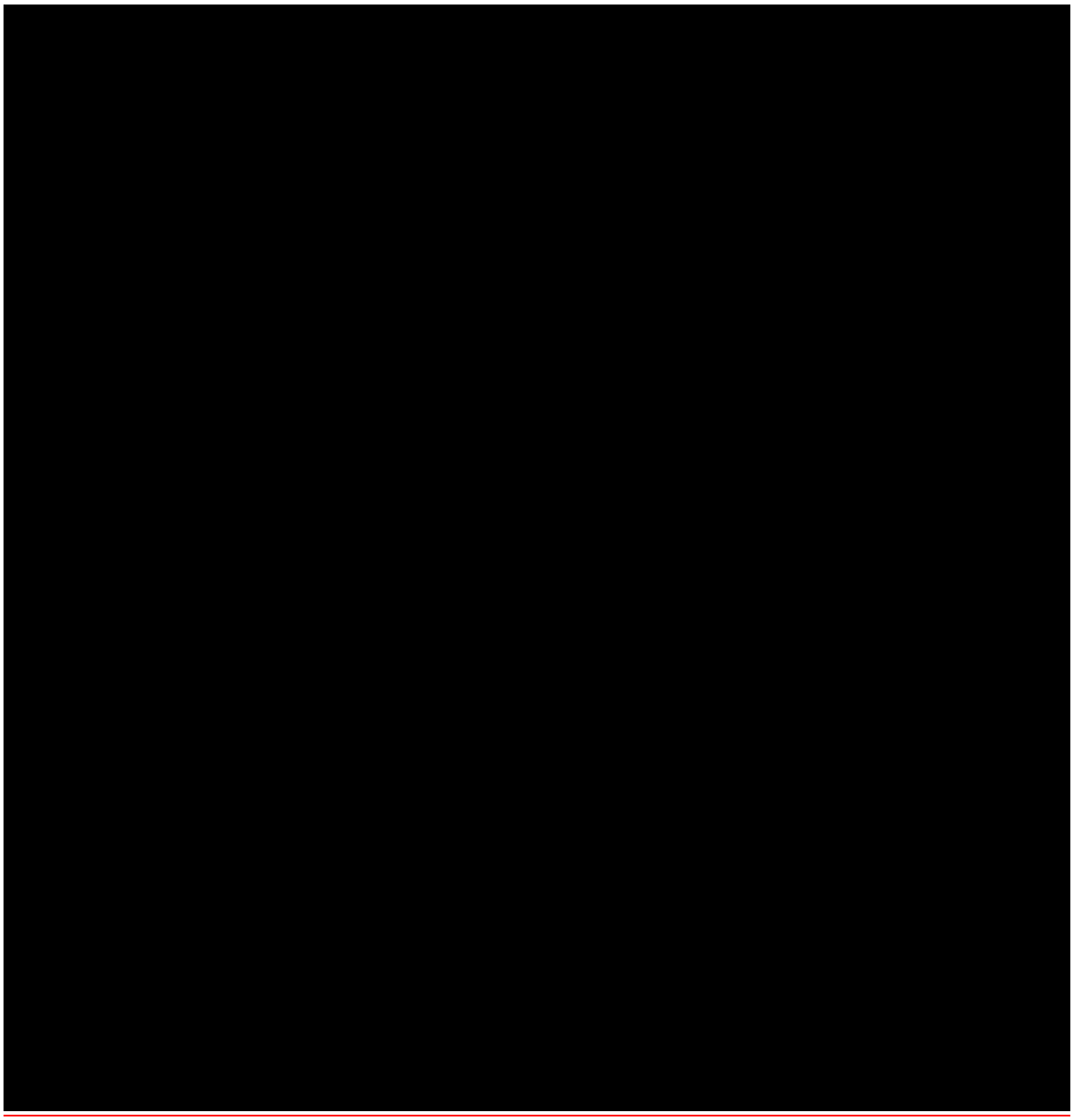

Figure 7. Correlation scatter plots of 30 min average N_2O fluxes (in nmol m^{-2} s^{-1}), as measured by CS-TDL and AR-P-QCL vs. AR-CW-QCL during the period I DOY 110-181 (upper panels a, b), and CS-TDL and LGR-CW-QCL vs. AR-CW-QCL during the period II DOY 206-271 (lower panels c, d). The lines present the linear fit with coefficients presented on the plots.

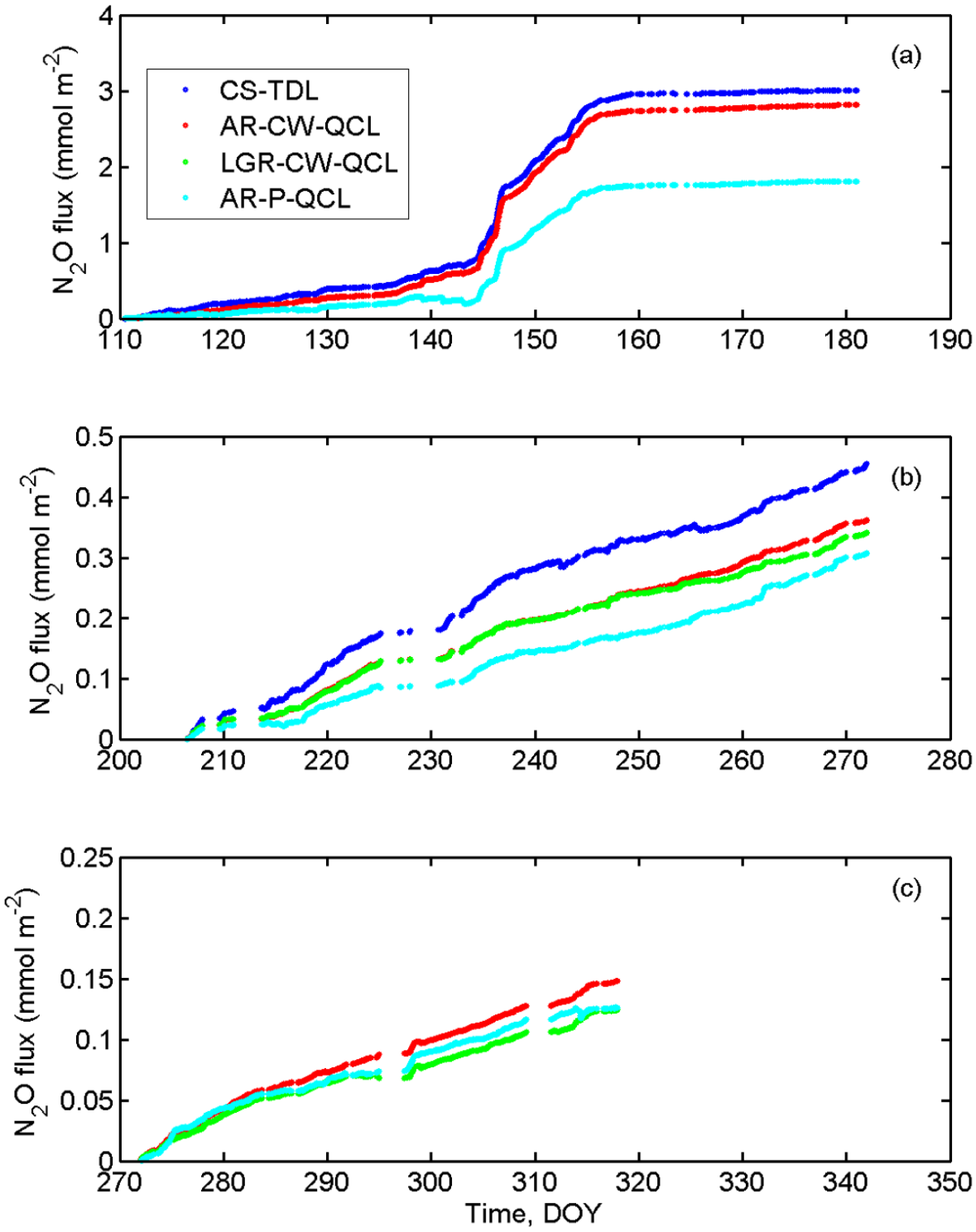

Figure 8. Cumulative sums of available flux data for three periods: upper panel (a) period I DOY 110-181 (20.04-30.06.2011), middle panel (b) period II DOY 206-271 (25.07-28.09.2011), lower panel (c) period III DOY 272 – 324 (29.09-20.11.2011). Accumulation of fluxes for each instrument was performed only for data if measurements were available for all instruments used in respective period. No gap filling was used.