Response to Referee 1

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.

For the systematic differences in fluxes see the answer under 4. Flux intercomparison.

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 acknowledge the comments by the referee and take the opportunity to improve the manuscript by discussing the results in the context of literature. In the revised MS we address the methodology of the analysis of instrumental characteristics as well as performance under demanding conditions of very low flux levels. The main findings will be summarised. In addition to the main areas of potential improvement, the referee has asked many interesting questions which we are happy to address and include in the revised MS when we consider it appropriate.

Some details:

The site description:

Please mention the amount and type of fertiliser.

The applied fertilizer was N-P-K-S fertilizer containing 76 kg N ha^{-1}, ammonium nitrate (NO_3-N : NH_4-N = 47:53). This will be added in the revised MS.

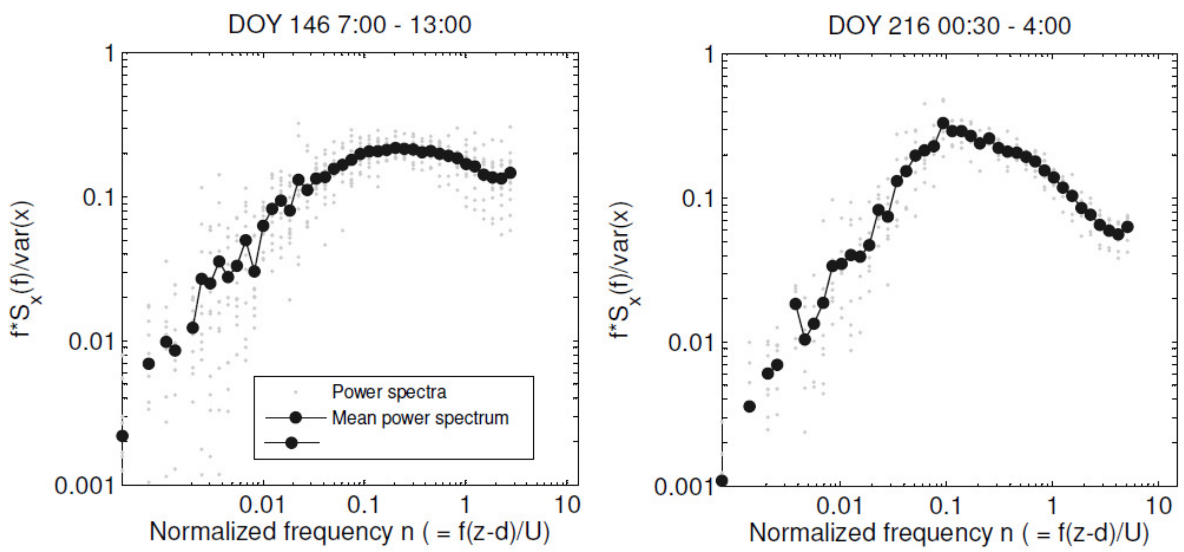
The measurement setup:

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.

The reported measurements heights and the canopy heights were correct. Indeed, when RCG was grown high, the measurement level was only about 0.5 m above the canopy top. We agree with the referee that 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. We do not expect that the relatively low observation level biases the overall N_2O flux level, in addition it is unlikely that the comparison of instrumentation is affected.

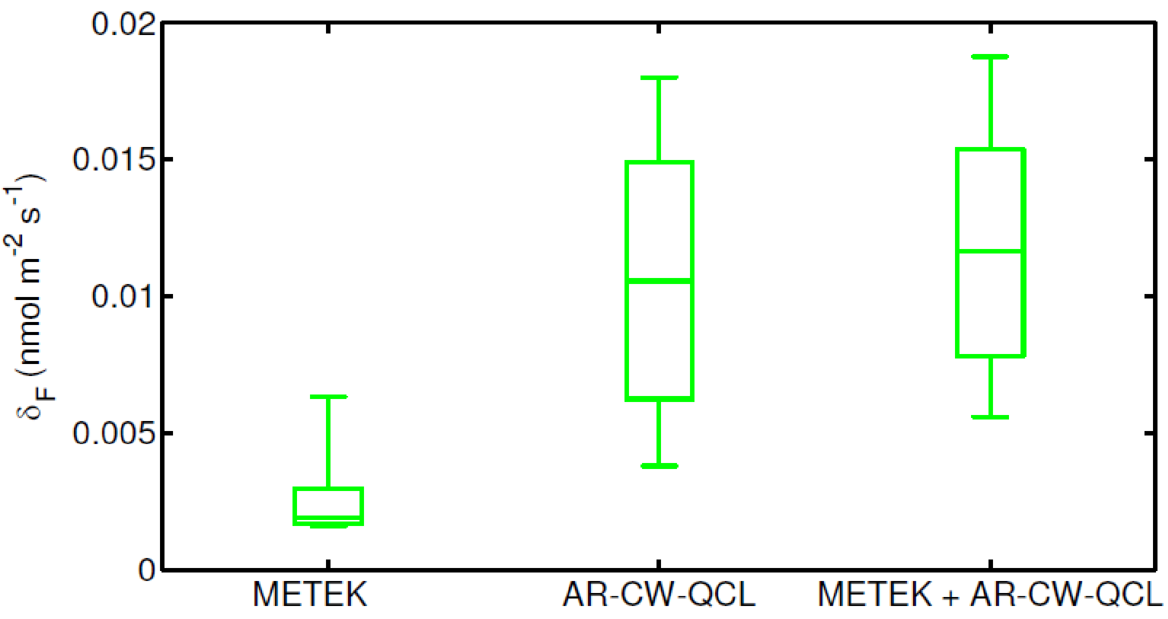
Below we present turbulent spectra as determined for vertical wind speed w. The spectra exhibit smooth and consistent shapes, without the particular impacts of the canopy foliage on spectral forms usually observed inside canopy. At the high frequency end of the spectra the impact of the noise on spectral shape can be observed. We do not associate this noise range in the spectral plot with any disturbance to the spectra related to close-to-canopy measurement level. We believe that the temperature spectra presented in the MS are at high frequency end similarly affected by the noise and not evidencing the canopy impact on spectral shapes.

We would like to refer also to the study by Launiainen et al. (2007) who 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 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 = 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. In conclusion, the sensor noise analysis is not affected at all (only the std of vertical wind and not the auto-covariance function being the input for calculation) and the effect on the random error 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 is however its higher sensitivity in detecting the low fluxes through higher concentration fluctuations expected (more) close to the source level.



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.

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 w. The noise level of the anemometer employed by the UEF was similar. This noise level was used to illustrate the resulting flux error together with the observations obtained by the instrument AR-CW-QCL, which had the lowest noise level 0.012 ppb (median value), see the figure below. The left bars represent the flux error statistics (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) due to anemometer noise only. In the middle is the error statistics due to the AR-CW-QCL signal random noise and the right statistics combine both. It can be concluded that our assumption was well justified for the instrument with the lowest noise and is therefore valid also for other instruments.



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. The UH system was run by using USA-1 anemometer by METEK GmbH, together with three N_2O instruments: model TGA100A by Campbell Scientific Inc., model CW-TILDAS-CS by Aerodyne Research Inc., and model N2O/CO-23d by Los Gatos Research Inc.

The UEF system was incorporating the sonic anemometer model R3-50 by Gill Instruments, Ltd. (lines 23-24 page 11752) and the N_2O analyser model QC-TILDAS-76-CS by Aerodyne Research Inc. The respective description will be improved to avoid confusion.

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.

This instrument was run with low flow rate due to existing setup from earlier experiments. The flow rate and resulting tube flow regime affects the response time of the respective system and does not allow comparing directly the instrumental response times. We admit that in the MS.

According to the manufacturer the instrumental performance is unaffected by changes in pressure.

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

More detailed answer given below.

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.

Data quality assurance was performed and it was presented in the beginning of Section 3 Results (will be moved to methods section in the revised MS). Here it will be repeated.

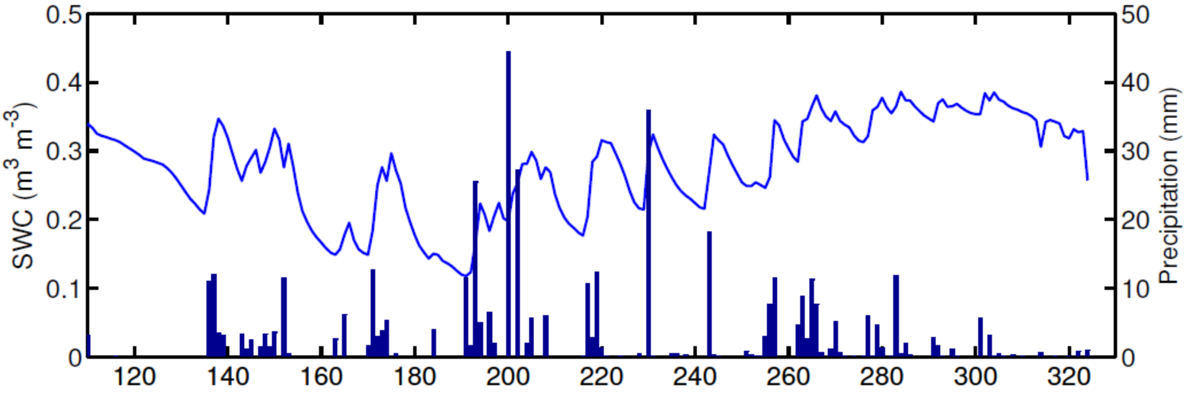
Prior to analysis data quality screening was performed. The measurements corresponding to wind direction interval 50-110° were excluded as possibly affected by instrumental cabin. In addition, quality screening was performed according to Vickers and Mahrt (1997) by applying the following statistics and 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 was performed according to H_2O concentration statistics for LGR-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.

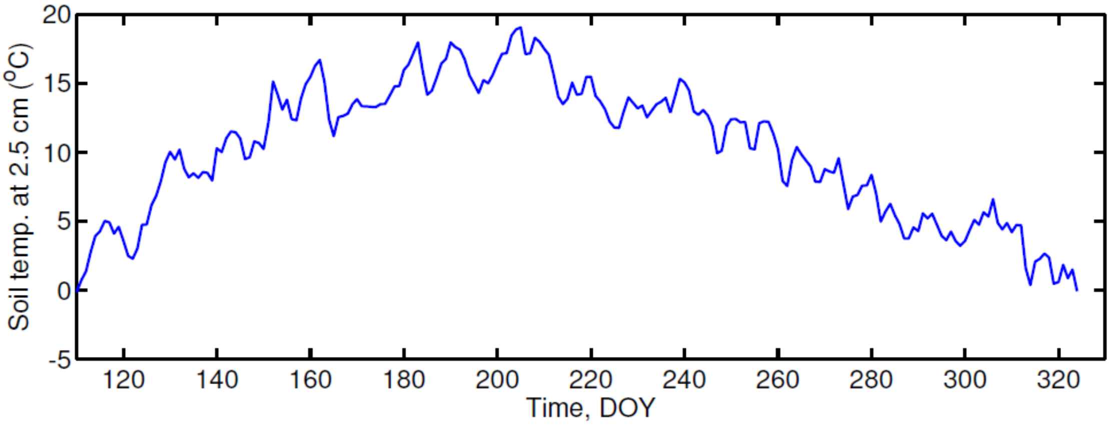
In summary, we used the quality criteria that should exclude the system malfunctioning as well as unphysical and/or unusual occasions. No quality screening for stationarity was performed as the focus of the study was the instrumental intercomparison, which is not affected by inclusion of data that would have classified as non-stationary according to respective criteria.

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.

The time series of soil water content (SWC), precipitation and soil temperature at 2.5 cm depth are presented below. The soil temperature had increasing trend until about DOY 205 (24 July, 2011) and since August declining seasonal trend. SWC was 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 average air temperatures and RH, together with wind speed and sensible and latent heat fluxes, were presented in Table 5. The figure below will be included in the revised MS.





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.

Discussion of results and conclusions are addressed below.

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.

Indeed, several flux random error estimates have been used in the literature.

For example, Wienhold et al. (1995) defined an error estimate, calculating the standard deviation of cross-covariance function over the lag interval far from the maximum. They called the error estimate as the detection limit of the flux.

Billesbach et al. (2011) calculated the flux error estimate over product of vertical wind speed and concentration fluctuations, randomly re-distributing one of the series. The method was called as the “shuffling method” and the authors proposed that the method was designed to only be sensitive to random instrument noise.

The error estimates with very clear physical meaning are the random error due to instrumental noise only, and the total error resulting from stochastic nature of turbulence due to limited sampling in time and/or in space. We use these methods as defined by eq. (3) given in the manuscript and the formula for total random error as proposed by Finkelstein and Sims (2001).

Note that the error estimate as given by eq. (3) and defined by Wienhold et al. (1995) does not rigorously define the same flux error estimate. Neither is the error estimate proposed by Billesbach et al. (2001) equivalent to the estimate by eq. (3) and as defined by Wienhold et al. (1995). However, detailed discussion of the different formulations of random flux error estimates is out of the scope of the current MS.

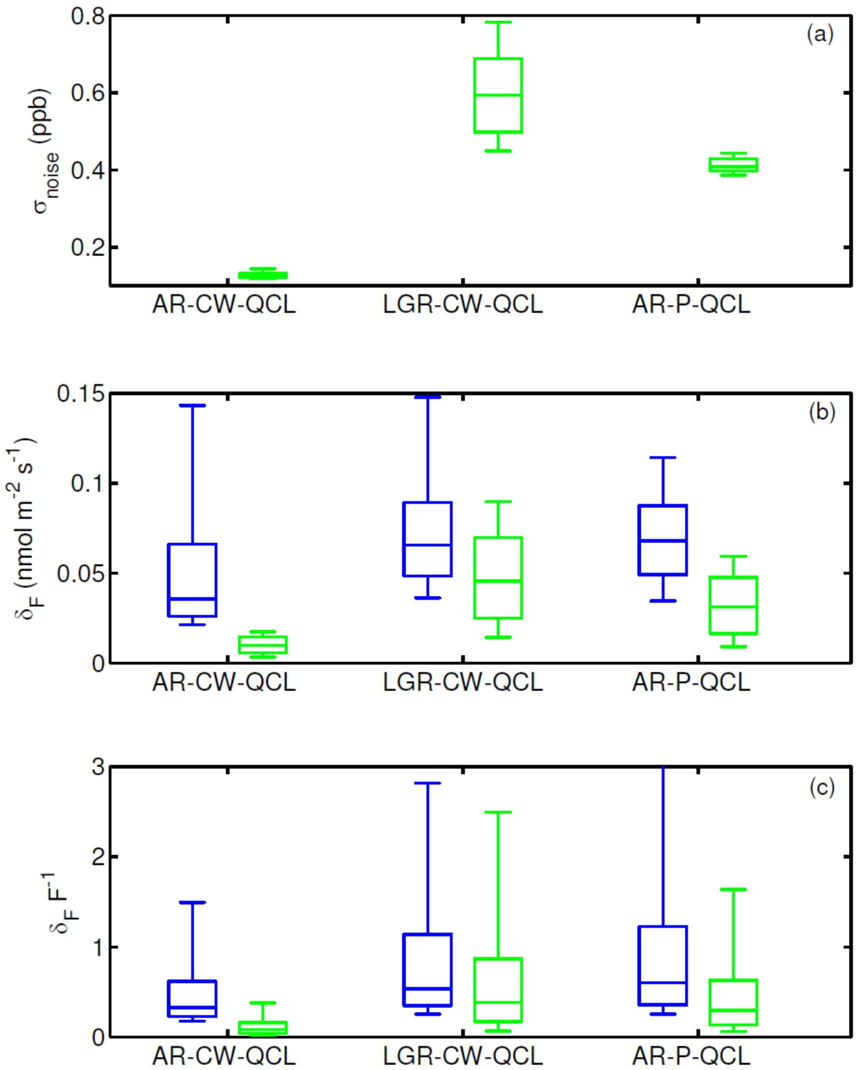
As the method to estimate the total flux error, we use the formulation by Finkelstein and Sims (2001) that defines the error as one standard deviation of the flux error. It is derived through the variance of the distribution of the individual flux realization around the ensemble mean (e.g. Lenschow et al., 1994). We commented that the same physical quantity could be estimated by several approaches, for example using the correlation-function approach (as done here) or by integration of spectra in frequency domain. There are also other approaches as discussed and analysed by Rannik et al. (2009). We will provide the reference but not repeat the details of that study more.

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.

Logarithmic scale was used to accommodate wide range of values for different instruments. On linear scale the noise of LGR-CW-QCL is roughly larger by a factor of 1.5 when compared to AR-P-QCL. In panel (b) it can be seen that the flux error due to instrumental noise is respectively larger for LGR-CW-QCL. The contribution of the instrumental error to total flux error is smaller than the other factors (primarily stochastic nature of turbulence) and the difference cannot be observed between these two instruments. Thus, in our opinion there is no inconsistency in the results related to flux errors corresponding to different instruments.

According to manufacture 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 a median value roughly twice higher than this. Under field conditions the instrumental noise can be somewhat higher as well as the estimation from the field measurements (and not under laboratory conditions) can introduce some uncertainty. In any case, the results obtained for different instruments are well comparable and the observed instrumental noise characteristics for other three instruments compare well with the results reported in other studies.

For example, Kroon et al. (2007) report for the instrument similar to AR-P-QCL the precision value 0.5 ppb $Hz^{-1/2}$, which is equivalent to 1.58 ppb at 10 Hz. Huang et al. (2014) report 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 presumably again due to the method of determination of these errors from field data.



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 tried to use as long time period for similar conditions (stability, wind speed) as possible over several hours. Averaging over different periods in terms of stability and wind speed might lead to ”smearing” of spectra as the spectra do not necessary follow the idealised normalizations in frequency scale and it is also known that turbulent spectra differ for different stabilities. Thus, averaging spectra is in our opinion compromise between better statistics over longer period, being potentially biased due to normalization issue referred to above, or shorter period with more similar conditions but more uncertain estimates. We chose to use optimal averaging period as several hours over similar meteorological conditions.

The time constant estimation for H_2O was performed for dry conditions to avoid the influence of RH. The influence of RH on time constant estimation for H_2O was not discussed as not being the subject of the study.

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 agree with the referee that for many applications the systematic differences are more important than the random errors. Nevertheless, to gain insight into measurements and build confidence it is very important that the flux values are not overwhelmingly dominated by noise. According to our experience it is very difficult to gain confidence in the broad average fluxes if the dynamics of the exchange at shorter time scale cannot 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.

The only evident systematic flux error source that could affect performance of CS-TDL would be incomplete drying of sample air. If that would be the case, then the measurements would suffer from partial density and spectroscopic corrections. However, the water fluxes are dominantly upward, therefore such a correction would tend to increase the flux values, therefore increasing more the systematic difference relative to other instruments.

However, we can accept that the old generation instrument CS-TDL is not able to determine the accurate fluxes due to signal stability issues. The effect of the low frequency signal drifting (fringes) of this instrument was illustrated and discussed by Mammarella et al. (2010). But it is not clear why the instrument AR-P-QCL resulted in systematically lower flux values during the first and second periods. We have used all auxiliary data available to investigate the possible reasons for the systematic differences, but found no explaining variable or reason. 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 (Peltola et al., 2014). However, we looked at the dependence of the flux residual between AR-CW-QCL and AR-P-QCL and found no systematic variation 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. We state this in the revised MS.

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.

The calibration was done as described below.

In the beginning of the campaign the instruments CS-TDL and AR-CW-QCL were calibrated. The factory calibration of LGR-CW-QCL was checked but no deviation was observed within the uncertainty range of the calibration gases.

No calibration was performed during the campaign for the Continuous Wave analysers (AR_CW-QCL and LGR-CW-QCL). These analysers are very stable according to manufacturers' and also according to the current study. The concentrations measured by these two instruments were very consistent throughout the campaign and the slope (sensitivity) of the 30 min average concentration comparison did not deviate from unity by more than 5%. Also the absolute levels of the concentrations detected by these two instruments were very close.

The operating parameters of CS-TDL, 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. Nevertheless, the instrument is known for low frequency signal variation dominated by offset drifts (fringe effect), therefore accurate measurement of absolute concentration by this instrument over a long period of time cannot be expected.

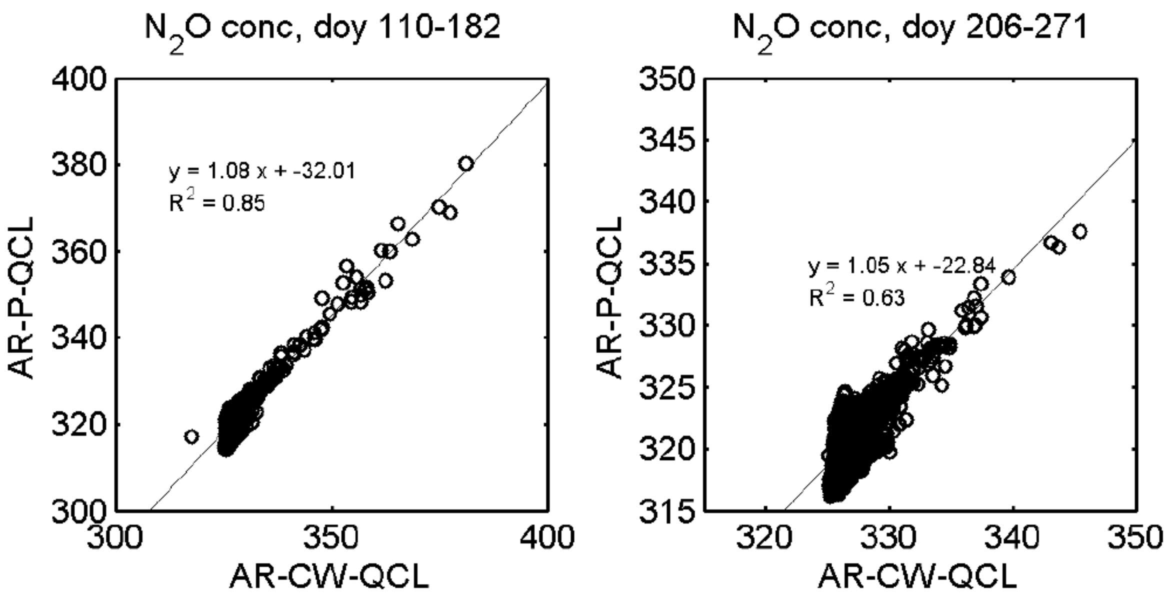
As stated in the MS, AR-P-QCL was calibrated every 2-3 weeks during summer with two standard gases 299 and 342 ppb. The calibration slope did not change by more than 7.6% throughout the campaign and maximum 6.1% between consecutive calibrations. Thus 6.1% can be considered as the maximum flux error arising from calibration accuracy of this instrument. Nevertheless, the correlation of the 30 min average concentration measured by this instrument as compared to AR-CW-QCL was not as good as among two Continuous Wave analysers letting to conclude that the instrument was not able to measure the absolute concentrations as accurately.

We performed also the comparison of the 30 min concentrations as measured by three instruments (excluding CS-TDL). The new instruments with Continuous Wave lasers were stable, the measured fluxes as well as concentrations were very close. The concentration statistics obtained for two measurement periods were correlated as following: period DOY 206-272 slope 0.94, intercept 19 ppb and $R^2 = 0.86$; period DOY 206-272 slope 1.04, intercept -13 ppb and $R^2 = 0.87$.

The comparison of concentrations between AR-P-QCL and AR-CW-QCL does not give as good statistics during the period DOY 206-272 (see below, concentrations presented in ppb): slope 1.05, intercept -23 ppb and $R^2 = 0.63$. The reasons for weaker concentrations correlation of AR-P-QCL with other instruments are not known. However, the concentration comparison presented here does not reveal that the calibration bias was the reason for the flux systematic difference for the instrument AR-P-QCL.

The analyser CS-TDL is known for its signal drifts and the absolute concentrations were not well determined. Therefore the concentration comparison cannot be used as the method for evaluation of the instrument's calibration impact on flux systematic bias.

The MS will be amended with the details of the calibration and the discussion of the potential impact of the sensitivity drift on cumulative fluxes.



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?

The noise level of CS-TDL was detected to be approx. 2 ppb, which translates into flux error of about 0.05 to 0.3 nmol m^{-2} s^{-1}. The total flux error, including the stochastic nature of turbulence, was within the range from 0.1 to 0.45 nmol m^{-2} s^{-1} (upper and lower quantiles of the distribution). 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 fluxes as measured by AR-CW-QCL less than this value). We observed that 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. Thus we conclude that the fluxes of CS-TDL, corresponding to close-to-zero fluxes as determined by AR-CW-QCL, are consistent with the flux error estimates. We will add this explanation into revised MS.

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.

The observed instrumental noise values and the flux errors will be compared and discussed with literature perspective (Pihlatie et al. (2005); Wang et al. (2013); Eugster et al. (2007); Kroon et al. (2007); Neftel et al. (2007); Neftel et al. (2010); Huang et al. (2014)).

The methodology will be discussed in the context of publications by Mauder et al. (2013), Kroon et al. (2010a) and Kroon et al. (2010b).

The ability of the state of the art EC systems to detect small N_2O exchange will be compared with that of the chamber flux detection limits (e.g. Venterea et al., 2009). As the chambers are widely employed for N_2O exchange measurement, this places our results in the perspective for N_2O 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.

Main conclusions of the study were following.

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

CS-TDL was not able to measure accurately the concentration in spite of weekly routine of checking and adjustment of the operating parameters (such as laser current and laser, housing and detector temperatures) and the shape and intensity of the absorption line. Also the concentrations and fluxes determined by AR-P-QCL instrument (calibrated every 2 to 3 weeks) deviated from the values determined by CW-QCL instruments. The reasons for systematic differences in fluxes determined by different instruments were not determined. We suggest that the special emphasis should be on the instrumental long-term stability and correcting procedures that can affect systematically the accuracy of detected fluxes when conducting long-term measurements of prevailingly low fluxes (relative to random flux errors).

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). It is important to analyse simultaneously the frequency response of the systems because the precision is improved (std of noise reduced) with instrumental averaging as $t^{-1/2}$. In the present study direct comparison between all instruments could not be done due to different flow setups.

During the period DOY 206-272, when all instruments were operational, the lower quantile/median/upper quantile statistics of the fluxes were 0.0083/0.11/0.31 nmol m^{-2} s^{-1} as measured by AR-CW-QCL instrument (with the lowest random errors).

The random errors of fluxes originate from the stochastic nature of turbulence (one-point sampling over limited time interval) as well as instrumental noise is contributing to the total flux error. The flux errors (median values 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 median 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 below the detection limit of the instrumentation (the flux error due to instrumental noise larger than the flux magnitude) and smaller than the total flux error (the flux error exceeding the flux magnitude): in case of CS-TDL 32/47 %, AR-CW-QCL 3.8/15 %, LGR-CW-QCL 23/28 %, and AR-P-QCL 16/30 %. 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). Thus the new instruments open up the possibility to study N_2O exchange at new ecosystems, broadening the scientific perspectives.

The summary of conclusions (perhaps with less quantitative information) will be included in the revised MS.

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.

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 as suggested. In addition, the expression for the flux error according to Finkelstein and Sims (2001) will be added for easier understanding of the error estimate used in the study.

(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 is given as follows according to Finkelstein and Sims (2001) :

$$\delta_F \approx \sqrt{\frac{1}{n}\left[\sum_{p=-m}^{m}\overline{w'w'}(p)\overline{c'c'}(p)+\sum_{p=-m}^{m}\overline{w'c'}(p)\overline{c'w'}(p)\right]} \quad (7),$$

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 have 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. We will amend the MS with this expression for the flux error.

(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.

σ_{noise} denotes the standard deviation of the signal variation due to the noise. P. 11757 line 5 it is incorrectly expressed. 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.

(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 the expression $\delta_{F,noise}=\frac{\sigma_w\sigma_{noise}}{\sqrt{fT}}$, where f denotes frequency and T the flux averaging period. The expression gives the std of the covariance error assuming variation in vertical wind speed as characterised by std σ_w and the noise in instruments signal as characterised by std σ_{noise}. As discussed previously, we have assumed that the contribution to the error due to noise by the anemometer's signal noise is negligible.

Assuming that there is no true turbulent variation of concentration and thus no 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 true fluxes smaller than the ones obtained from the expression above. Therefore we called this flux error as the error "giving essentially the detection limit of the flux that the system is able to measure", referring to the fact that the error is mainly determined by the instrumental noise characteristics. Note that the error depends also on the observation conditions through σ_w.

However, this is a loose definition based on the std of the flux error due to gas analyser's noise. It would be perhaps more rigorous to define the flux detection limit as 2 times $\delta_{F,noise}$, including 95% of variation range of error values due to noise only (assuming Gaussian distribution). Nevertheless, it serves as a useful quantity to compare different instruments and we will keep the same definition of the error due to the instrumental noise, adding the considerations leading to calling it as the flux detection limit.

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

Description of data availability and definition of respective periods as well as quality screening will be moved into methods section.

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

We revise the discussion by removing the repetition and adding the comparison with literature and the discussion of the uncertainty estimation of the fluxes.

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 will be inserted for reference.

References

Billesbach, D. P.: Estimating uncertainties in individual eddy covariance flux measurements: A comparison of methods and a proposed new method, Agr. Forest Meteorol., 151, 394–405, 2011.

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, 2014. doi:10.5194/acpd-14-20417-2014.

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, 2010. doi:10.1016/j.agrformet.2009.08.008.

Kroona, P . S ., Schrier - Uijl, A. P., Hensen, A. , Veenendaal, E. M., and Jonker, H. J. J.: Annual balances of CH4 and N2O from a managed fen meadow using eddy covariance flux measurements, European Journal of Soil Science, 61, 773–784, 2010.

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, 2007. DOI: 10.1111/j.1600-0889.2007.00313.x

Neftel, A., Ammann, C., Fischer, C., Spirig, C., Conen, F., Emmenegger, L., Tuzson, B., and Wahlen, S.: N2O exchange over managed grassland: application of a quantum cascade laser spectrometer for micrometeorological flux measurements, Agr. Forest Meteorol., 150, 775–785, 2010.

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, 2014. doi:10.5194/bg-11-3163-2014.

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.

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, 2009. doi:10.2136/sssaj2008.0307.

Wienhold, F. G., Welling, M. and Harris, G. W., Micrometeorological measurement and source region analysis of nitrous oxide fluxes from an agricultural soil, Atmospheric Environment, 29, 2219–2227, 1995.