Atmos. Meas. Tech. Discuss., 3, C2390–C2406, 2011
www.atmos-meas-tech-discuss.net/3/C2390/2011/

Atmospheric Measurement Techniques Discussions

Interactive comment on "CO₂-gradient measurements using a parallel multi-analyzer setup" *by* L. Siebicke et al.

L. Siebicke et al.

lukas.siebicke@uni-bayreuth.de

Received and published: 4 January 2011

Final Author Comments to Anonymous Referee #2

Lukas Siebicke

4 January 2011

On behalf of all co-authors I would like to thank the referee for the positive feedback about ideas presented in this publication and for suggestions provided for improvements. Below is our response to the referee's comments.

1 General comments

1.1 " * Though I think ..."

We agree that it is always good to know the source of inter-instrument bias and to avoid or eliminate it in the system design and by conventional calibration. However, there will almost always be remaining bias, and it depends on the particular situation whether the magnitude of the remaining bias is significant for the application or not. The presented statistical calibration should be evaluated independently because its functioning does not depend on the magnitude of the remaining inter-instrument bias after conventional calibration. Furthermore, statistical calibration might be useful in situations where conventional calibration is not possible or not sufficiently accurate.

We have reflected those changes in the conclusions in the revised version. We have also clarified that additional applications mentioned in the conclusions are suggestions for further research.

1.2 " * In my initial reading ..."

The referee's comment addresses pressure differences between individual analyzers and between calibration and measurement periods. First of all we would like to cite our response to a referee's comment during the initial evaluation of the manuscript here to make it available to everyone in the discussion:

"We absolutely share your surprise at this point. However, there are plausible explanations for significant bias even after calibration relative to measured standards. We explain the remaining bias as different conditions during standard measurements and during measurements of calibration standards (as noted in the text). Because each analyzer was connected to a 75 m long sample tube it was unfeasible to apply the reference gas at the end of the sample tube in front of the filter where normal samples are drawn from. This would have required many long additional tubes for the distribution of the reference gas which would have introduced their own uncertainties. That means, the effect of the sample tube and the filter at the inlet is excluded during calibration. The lines and filters can cause pressure differences between analyzers during normal measurements due to geometric differences, temperature differences or differences of filter resistance. We also suspect that geometric differences of the sample lines of the analyzers will affect pressure and therefore concentration readings. Even though flow rates were individually controlled for each line, they were set for normal measurement conditions, meaning that the flow rate in a sample line with a pressure drop different from other lines could change its flow rate during calibration and thus change the pressure in the sample cell. Individual mass flow controllers for each line should improve this issue. However, having said all that, we believe that the statistical

calibration should be largely independent of the magnitude of inter-instrument bias as long as the latter is reasonably constant between individual statistical calibrations."

In the following we would like to give some more answers to specific questions in this referee comment.

"...I mentioned my surprise at how large the differences (order of 25 umol/mol) were between the various IRGAs ..." and "...what is the magnitude of the pressure differences to cause such large concentration differences?...": yes, concentration differences on the order of 25 μmol mol^{-1} might be explained by pressure differences (pressure during calibration versus pressure during measurement periods) according to the following simple calculation example: the LI-6262 cell pressure during calibration was 77.1 kPa but only 68.8 kPa during the following measurement period. This is a difference of 8.3 kPa. Assuming the sample concentration was 380 μmol mol^{-1}, this pressure difference would be equivalent to approximately 45 μmol mol^{-1} ((380 μmol mol^{-1}/68.8 kPa)*8.3 kPa = 45.8 μmol mol^{-1}). So, from that perspective, 25 μmol mol^{-1} can well be explained by 45 μmol mol^{-1}. However, the pressure difference of 8.3 kPa is the difference in pressure during calibration versus pressure during measurement periods and not the pressure difference between individual analyzers during measurements. Unfortunately, not all analyzers were equipped with a pressure sensor. Three of the LI-6262 analyzers did have their own sensor and there was an additional sensor installed at the manifold just downstream of all analyzers outlets, the reading of which was close to the pressure in the sample cells thanks to short tubing connections and a large diameter of the manifold. That means with the given data it is impossible to answer what the inter-instrument pressure differences were. This also answers the next question: " Can the pressure differences be accounted for and used to correct the biases in the concentration data?" No, this is not possible in this case. Not even the comparison of the existing pressure measurements will help much to quantify inter-instrument differences because all existing sensors were installed in LI-6262 and none in the other types, whereas the sample line geometry and (related pressure differ-

ences) differ more between different analyzer types used. “What could be done to the measurement system to improve or eliminate this problem in the future?”: Pressure differences and possibly other sources of bias can be minimized if all analyzers are exactly of the same type, if all sample lines have individual automatic mass flow controllers and all analyzers have their own pressure sensor. Using identical analyzers will also help to minimize inter-instrument temperature bias, which is an answer to the question “Or, is it more than just pressure differences?”.

In case someone might get the impression that the pressure regime was badly controlled in the experiment, one should know that the pressure was well behaved in the sense that the pressure – although partly unknown – was quite constant (from one calibration to the next for individual analyzers and also between the different pressure sensors during one calibration).

Finally, it should be recalled that the functioning of the statistical calibration method, which is the focus of the publication, is largely independent from the discussion about pressure differences and the magnitude of related inter-instrument bias. However, if in future experiments a very accurate system with insignificant inter-instrument bias could be set up, this would help to evaluate the statistical calibration method independently and is therefore encouraged. After further verification of the statistical calibration method it might be applied in cases where inter-instrument bias can not be or is not eliminated due to resource limitations / intentional cost savings or if the type of instruments used does not allow to do so or possibly for post-processing existing data sets which do not have the desired inter-instrument accuracy.

1.3 “ * One possible ...”

From the referee’s comment it seems to us that the purpose of the LES study was not expressed clearly enough in the original version of the paper. In the revised version of the section dealing with LES, which was re-written to a large extend, we have hopefully

better conveyed the intention and added the following paragraph to the text:

“It is obvious that the Large Eddy Simulations presented here are an idealization and do not account for the complexity of the given forest site, particularly because they do not fully account for the forest canopy. However, we would like to stress that the purpose of the simulation is to test the idealized case of turbulent mixing given realistic physical values of scalar concentration gradients and a vertical source and sink distribution that does mimic sources at the forest floor and sinks in the forest canopy with respect to their vertical distribution and their intensity. Verifying and accepting the assumption made in Sec. 2.4.2 first for an idealized case is necessary before addressing measurements from the more complex forest setting. Whether conditions in the forest at any given time show sufficient mixing is not evaluated by LES but by the application of an empirical mixing index (see Sec. 2.6) which is based on measured data.”

We believe that the LES exercise is worthwhile for what it stands for (considering above made clarification about its purpose) because – although only a model – LES is very well suited to study the turbulent atmosphere whereas as of today LES has significant limitations regarding the applicability inside a forest canopy. Therefore, the statistical correction method presented relies on the mixing index to differentiate between sufficiently and not sufficiently turbulent conditions. In summary, the presented LES study does not prove that assumptions made are true at all times inside the forest but by verifying the assumptions for well mixed conditions it provides confidence that there are situations where the basic assumptions made are correct and we can proceed by applying other means to verify conditions inside the forest at any particular time. The presented mixing index was a convenient way for the current study to do this but there will certainly be other ways to achieve this, too.

1.4 " * I think some period ..."

This referee comment concerns inter-comparison between instruments by using co-located inlets. Although the original version of the publication, which was focusing on the statistical calibration method, was quite brief on that subject, this is still an important aspect, which we also addressed during the presented experiment. We have included a discussion about inter-instrument comparison (other than the statistical approach) in the revised version of the paper and we would like to share this discussion here to explain why we have included in the current paper what we have included and also to indicate possibilities for future experiments.

- Case A: In any sampling system with two or more inlets per analyzer, one of the inlets can be permanently or temporarily used for instrument inter-comparison. This is common practice for example in CO_2 profile measurements using more than one analyzer. It allows for determination of a constant or a time dependent bias between instruments. This technique was applied for vertical profile measurements conducted during the current experiment. Another example of a setup which determined a time dependent inter-instrument bias from co-located inlets is described in Sun et al. (2007). However, if one inlet is used for inter-comparison and one or more different inlets for actual measurements than the bias determined from the co-located instruments is instrument specific but not specific for the particular inlet, i.e. bias due to differences between sample tubes, filters, valves, and, depending on the system design, possibly also other system components such as individual pumps etc., can not be detected by this approach. The main point to note here which is relevant with respect to the current paper is that such systems which use two or more inlets do not measure continuously at one location.
- Case B: The current study presents a different approach in the sense that every

individual analyzer performs continuous measurements at one location using one inlet only. With only one inlet per analyzer, above described inter-instrument comparison is not possible. However, a modified approach to inter-instrument comparison might include the following:

- Option 1: With mobile inlets it is possible to place the inlets of several or all analyzers next to each other during inter-comparison periods and move them to their desired measurement location for actual measurements. Drawbacks: actual measurements are discontinuous, inlets need to be mobile which was not an option at the forest site presented in this paper, the number of inter-comparison periods is limited by the manual intervention necessary to move the inlets, moving inlets might introduce bias which can not be accounted for by this type of inter-comparison (bias e.g. due to differential heating of tubes in direct sunlight or shadow etc.). Mobile inlets might be considered for experiments at open sites (e.g. over grassland).
- Option 2: If all analyzers are positioned in a common location then co-located inlets can be created by disconnecting the (long) sample tubes leading to actual measurement locations, and have all analyzers sample the same air, connecting only short peaces of tubing and the necessary filters. This technique was applied during the current experiment. While easier to accomplish in a forest than Option 1, it has the following drawbacks: disconnecting long sample tubes including the terminal filters for inter-comparison creates a different pressure regime during actual measurements and during inter-comparison periods because of the pressure drop caused by the tubes and because filters used during measurements and during inter-comparison are likely not the same. Therefore the results of this inter-instrument comparison are not included in the current paper as they were not appropriate to evaluate the statistical calibration method described. Furthermore, similarly to Option 1, both actual measurements and inter-comparison periods are discontinuous in such a setup and the change between them is likely

to involve manual intervention, making calibration periods relatively rare and possibly insufficient for adequately addressing time dependent bias.

- Option 3: Create a hybrid between the systems described in Case A and in Case B. This might be achieved by taking the system as described in Case B with one analyzer per sample location but adding a second inlet to every analyzer for the purpose of inter-comparison (same type and length of tubing and same characteristics of other components like filters, valves etc. as in first sample line). All those second inlets might then be co-located for inter-instrument comparison. Using an automated system to regularly change between measurement periods (relatively long) and inter-comparison periods (relatively short) it should be possible to achieve semi-continuous measurements with such a system. Inter-instrument comparison does not need to involve all analyzers at the same time. Using subsets of all analyzers with various combinations avoids missing measurements at all sampling locations at the same time and might help to reduce negative impacts of gaps in the data during data analysis. A semi-continuous measurement system should be considered for further experiments. It might be combined with the statistical calibration method presented and should provide a basis to evaluate the latter method further.

2 Specific comments

2.1 " * p.4386, line 8 ..."

We have changed the wording here according to your suggestion.

2.2 " * p.4386, lines 14-16 ..."

The referee's comment addresses the applicability of presented method in a heterogeneous forest suggesting that "... the changes in the overlying canopy structure can result in different amounts of vertical mixing of CO2 at each location ...". Based on our experimental results we know that this is true at the given site. Moreover, we are so concerned about this important issue that we have devoted a separate manuscript to exactly this issue and had also referenced it in the original version of the current publication since it seemed relevant (see also the following reply concerning that publication). However, the good news concerning the applicability of the presented statistical calibration method in the forest is that the statistical method should largely be unaffected by different amounts of vertical mixing. This is because the method is based on probability density distributions. There can well be more vertical mixing at some locations compared to others without affecting the mode of the probability density distribution (the property that matters for the correction) as long as mixing between individual sampling locations (i.e. horizontal in our case) is dominant and produces the dominant mode in the probability density distribution, meaning the sub-canopy sampling locations have to be more frequently influenced by the common sub-canopy regime than by location specific vertical mixing with any particular magnitude of concentration perturbation. We would like to stress that the method is insensitive to any local concentration perturbation however strong it is as long as the particular value of the perturbation is not the most frequently observed one. In other words, only the frequency of perturbations with identical magnitude matters, not their magnitude itself. So, even if a sample location was frequently subjected to vertical mixing or to particular local sources or sinks this would only be harmful to the method if those peculiarities managed to produce the same concentration perturbation more often than the measurement point samples the general sub-canopy concentration regime. Considering those arguments the method seems to be more robust than one might expect at first (assuming that an appropriate mixing threshold is applied), and experimental data confirm this, too. Having said that,

we still encourage research towards identification of conditions which allow for statistical calibration and others that don't (using the "mixing index" or other measures).

2.3 " * p.4386, line 17 ..."

Although we had included the reference to Siebicke et al. (2011) in the original version of the current paper because it seemed relevant (see discussion in previous comment), following the referee's suggestion we have now deleted it from the revised version, thus conforming to more strict citation rules. Concerning the reference to Foken et al. (2011), although more appropriate but not yet available, we have replaced it with Serafimovich et al. (2008), which is ISSN referenced and currently available.

2.4 " * p.4386, lines 20-25 ..."

The referee suggested: "your data set could be used to evaluate how much information is lost in the "sequential" approach (by simulating a "sequential" CO_2 measurement and comparing the results to your measurements made with the "parallel" approach)". This is a very good idea. We had actually already done exactly the suggested analysis and included it in a recently submitted publication. Given the referee's comment about not yet published papers, we refrain from citing it here but it can be cited in the final version of the current paper if it should happen to be available by that time. To give a brief summary of the results of this exercise we can say here that there are significant differences between 30-min mean values computed for any one sample location by the "continuous approach" and by the simulated "discontinuous approach". We interpret the differences mainly as a consequence of the skewness of frequency distributions of concentrations (and concentration perturbations), with different skewness observed at the individual sampling locations. Experimental data about frequency distributions of concentration perturbations are presented in Fig. 7 and Fig. 8 of the current paper,

showing that frequency distributions indeed differ between individual sample points within the same time interval. More distributions (frequency distributions of absolute concentrations and frequency distributions of concentration perturbations) are given in Siebicke et al. (2011) (please refer to the previous comment for this citation). Both theoretical considerations and our results show that discontinuously sampling a time series which has a highly skewed distribution will most likely lead to wrong estimates of the mean. Furthermore, those errors in 30-min mean values were found to be significant in relation to the magnitude of the gradients to be observed. This finding is a strong incentive to work towards the application of continuous measurements for the observation of gradients where small errors in mean values at the individual sampling locations can have significant impact on gradients and thus on derived fluxes.

2.5 " * p.4388, lines 7-8 ..."

The complete hill is covered with spruce forest. With "upper section" we were referring to the location of the site. The tree height we indicated applies to the trees within the footprint of the measurements. We have improved the text here to be more precise. Details about vegetation structure and the heterogeneity of the forest within the footprint is given in above mentioned publication Siebicke et al. (2011). In summary, on the spatial scale which is typical for the evaluation of flux footprints (hundreds of meters, e.g. transitions between forest and clearings etc.) the forest is quite homogeneous within the relevant footprint. However, on a smaller spatial scale (meters to tens of meters) there is, not surprisingly, a lot of heterogeneity as presented in mentioned study in detail.

2.6 " * p.4389, lines 5-7 ..."

Yes, the calibration gases were sampled every 4 hours by all ten analyzers at the same time. We have included this in the text of the revised version.

2.7 " * p.4392, lines 10-14 ..."

It is correct that the median uses the "SORTED" observations which was missing in the text. Thanks for pointing this out. We have added this to the text and differentiated between the index of the sorted list and the index referring to the sampling locations in the notation of Eq. 1. The calculations done are not affected because they have been using the sorted values already.

2.8 " * p.4395, line 20 ..."

Concerning the referee's question about characteristics of the convective boundary-layer: "a heat flux of 0.01 K m/s seems pretty weak...but it must be enough to develop a convective boundary layer?": yes, a heat flux of 0.01 $\mathrm{K\,m\,s^{-1}}$ is relatively small but is enough to develop a convective boundary-layer. To reflect that we have added information about the Obukhov-length to the text of the revised version: "At the bottom boundary of the model domain a near-surface heat flux of 0.01 $\mathrm{K\,m\,s^{-1}}$ was prescribed, so that a convective boundary layer with a Obukhov-length in the range between -40 and -50 m developed with time."

2.9 " * p.4396, lines 15-20 ..."

The correlation was determined between the central sampling point (M5) and the most distant point equipped with the same analyzer model (which was M6 for the along

slope transect but could just as well have been M7 because both locations have the same distance to M5, and which was M8 for the across slope transect, i.e. M9 was too close to M5 for the correlation to be representative for points further away). If the more distant points M5 and M8 are well correlated than it is likely that the closer points M5 and M9, (which are along the same transect as M5 and M8), are just as correlated. "How did the along slope correlations compare to the cross-slope ones?" We did some comparisons of along and across slope correlations in the sub-canopy in the context of some other work, according to which the across slope transect was generally more correlated. However, one should keep in mind that the calculation of the mixing index with the given selection of sampling points is just a convenient solution to the problem. It might seem more general to base the mixing index on the correlation between any possible combination of sample point pairs. Unfortunately, the computational resources needed to achieve this were prohibitive. If there is a statistical method to quantify the correlation of all sampling points of the field which is more appropriate than the mixing index presented, we would be interested to here about it. And yes, "the TF issue" is addressed in the discussion.

2.10 " * p.4396, lines 21-22 ..."

The mixing index threshold MI_c was chosen as a fixed value for the complete analysis, i.e. MI_c did not vary over time. It was chosen based on the density distribution of all mixing indexes observed during the experiment which is given in Fig. 4a. We have clarified in the figure caption that the data of Fig. 4a are from all periods over the experiment. "... how was the MI_c chosen within the range? ...": the density distribution of MI shown in Fig. 4a shows a prominent peak at low values of MI. Those are obviously conditions with poor mixing, mostly at night time, so a threshold above which sufficient mixing can be expected needs to be at higher values of MI than this peak. And yes, above a certain value of MI_c such as approx. 0.12, the results are relatively insensitive to the choice of MI_c. Too high a value of MI_c means there are fewer periods allowing

for statistical calibration, thus making the system more prone to the effect of drifting instrument bias. As you correctly mention, this issue is addressed in Fig. 10 and in the discussion. The value for MI_c chosen here should not be taken as objective in the sense of a physical constant or anything similar but objective in the sense that it is reproducible and applies the same criteria for all periods of the experiment.

2.11 " * p.4403, line 5 ..."

Thank you for pointing out this typo.

2.12 " * p.4404, lines 17-19 ..."

We have changed the sentence on the basis of the referee's comment.

2.13 " * p.4404, line 11 ..."

Thank you for pointing out this typo.

2.14 " * Fig. 7 ..."

Yes, these distributions are "created from the 1-hz data (so there are 3600 samples to create each distribution)?"! We have added this to the figure caption.

2.15 " * Some Figures have ..."

We have applied your suggestions for the revised version, i.e. put the reference to subfigures always after the text. All subfigure numbers are now placed below the figures

using the template of the publisher. There is only one exception (Fig. 8), which has the subfigure reference inside the subfigures because this figure had to be produced as a combined panel to guarantee that the y-axis scaling is correct for the different subfigures.

2.15.1 General reply to this referee's comments:

some of the referee's comments address information related to but not included in the current paper. While this information is interesting and important as such, we had concentrated on outlining the statistical calibration method in this publication. We hope that with our additional explanations above and the changes in the revised version it is possible to follow the main concepts of the current publication until the other publications, which address the application of the method presented here, will also be available.

Thank you again for helping to improve this publication and for participating in the discussion about this very interesting topic.

References

Thomas Foken, Franz X. Meixner, Eva Falge, Cornelius Zetzsch, Andrei Serafimovich, Anika Bargsten, Thomas Behrendt, Tobias Biermann, Claudia Breuninger, Tobias Gerken, Martina Hunner, Lydia Lehmann-Pape, Korbinian Hens, Georg Jocher, Jürgen Kesselmeier, Johannes Lüers, Jens-Christopher Mayer, Alexander Moravek, Daniel Plake, Michael Riederer, Friederike Rütz, Stephanie Schier, Lukas Siebicke, Matthias Sörgel, Katharina Staudt, Ivonne Trebs, Anywhere Tsokankunku, Veronika Wolff, and Zhilin Zhu. Atmospheric transport and chemistry in forest ecosystems – overview of the EGER-project. *Agric. For. Meteorol.*, 2011. (to be submitted).

Andrei Serafimovich, Lukas Siebicke, Katharina Staudt, Johannes Lüers, Martina Hunner, Tobias Gerken, Stephanie Schier, Tobias Biermann, Friederike Rütz, Jannis von Buttlar, Michael Riederer, Eva Falge, Jens-Christopher Mayer, and Thomas Foken. Exchange pro-

cesses in mountainous regions (EGER) - documentation of the intensive observation period (IOP2) June, 1^{st} to July, 15^{th} 2008. Arbeitsergebnisse, 37, University of Bayreuth, Department of Micrometeorology, 2008. ISSN 1614-8916, 180 pp.

Lukas Siebicke, Andrei Serafimovich, and Thomas Foken. Linking CO_2-advection estimates to vegetation structure at a forest site. *Agric. For. Meteorol.*, 2011. (to be re-submitted in revised version).

Jielun Sun, Sean P. Burns, Anthony C. Delany, Steven P. Oncley, Andrew A. Turnipseed, Britton B. Stephens, Donald H. Lenschow, Margaret A. LeMone, Russell K. Monson, and Dean E. Anderson. CO_2 transport over complex terrain. *Agric. For. Meteorol.*, 145(1-2):1 – 21, 2007.