**Biogeosciences
Discussions**

Interactive
Comment

Interactive comment on "Intercomparison and assessment of turbulent and physiological exchange parameters of grassland" *by* E. Nemitz et al.

E. Nemitz et al.

Received and published: 30 July 2009

General remarks

We thank the anonymous Referee for their careful reading of the manuscript, for the overall positive response, and for their constructive suggestions which have helped to improve the manuscript further. In the following we have responded to the individual more critical points raised by the Referee.

It should be noted that both Referees mainly comment on the comparison of the eddy-covariance results and heat fluxes, while neither of them comment on the comparison of canopy temperatures and the resistances derived here. This indicates that both Referees come from the research community addressing mainly fluxes of momentum,

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

heat and (presumably) CO2. By contrast, as the title indicates, this paper is distinct from more traditional intercomparison papers of micrometeorological instrumentation or flux calculation approaches and focuses on an analysis of the uncertainties in deriving input parameters for gradient flux calculations and SVAT modelling of reactive trace gases such as NH3. Although we agree that more details were needed throughout the manuscript to interpret differences between measurements, some of the suggestions of the referees would redirect the focus of this publication.

Specific responses *A so-called consensus data set is derived from these synchronous measurements as basis for the other studies within the GRAMINAE project. This part of the paper is not of great interest for a wider community unless the focus is shifted towards the methodology how to create such a consensus data set.*

While clearly a substantial scientific piece of work in its own right, this paper nevertheless provides important background for the interpretation of the different companion papers. The responses we have received on some of these companion papers imply that the illustration of how the best estimate was derived is very well important for understanding various quality control issues regarding the other papers. In fact, some referees requested to provide more detail on this in the companion papers. On balance we feel that this section is useful.

The appeal to a wider community outside GRAMINAE should also be considered when formulating the conclusions. The results of the intercomparison are certainly interesting. However, this reviewer does not agree with the approach of choosing always the largest flux estimate of all ten systems in order to close the energy budget. This reviewer does not believe that this one sensor combination measured the latent heat flux accurately while all the other systems underestimated the true flux.

We agree with this Referee that it is not very likely that the largest fluxes are all the correct ones and that a non-closure of 20% is in agreement with previous measurements (this was already stated in the original manuscript). We only meant to illustrate that,

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

if the largest fluxes were used, the energy balance could be closed. This implies that the non-closure is within the uncertainties of the various measurements. Clearly, we have given the wrong impression overall in the manuscript and have re-phrased various statements to clarify this in the revised manuscript. We have also expanded the conclusions, taking on board further results of new analysis prompted by the comments of this Referee.

It is commonly known that the calibration of KH20 krypton hygrometers is not very stable in time (e.g. Mauder et al., 2006; Mauder et al., 2007c). Therefore, it would be very interesting to know about the KH20 calibration procedures applied. Ideally, these instruments should be calibrated both before and after the field deployment. Moreover, frequent/daily cleaning of the optical windows is very important, since they are hygroscopic and are prone to scaling effects.

We are not quite sure what calibration the Referee refers to, and the Mauder et al. (2006) is also not specific on this point. The measurements do not rely on the calibration for absolute measurements of water vapour. Instead, the relative measurement of the co-variance (basically the exchange velocity) was combined with absolute measurements of the specific humidity from slow response sensors to derive absolute fluxes. This had not been described in the manuscript as we assumed that this is standard practice, and has now been clarified. The windows were indeed cleaned every day during this intensive field campaign.

In general, more information about the methodology would be helpful to analyse the reasons for disagreements between sensors, e.g. what were the order and the distance between the measurements systems, what kind of post-processing steps were applied and in which order.

This is addressed in more detail in response to the Referee's specific comments below. In particular, the different corrections have been added as a new Table 2.

Most of the theory part presented in section 2 is text book knowledge and can be

trimmed or omitted.

We disagree for the reasons already stated in the reply to a similar comment by Referee 1. In fact, the apparent misunderstanding how the KH20 was used to calculate latent heat fluxes indicates that, if at all, more information is needed.

Some of the more recent literature about sensor and software intercomparisons is not considered (Mauder et al., 2006; Mauder et al., 2007c; Mauder et al., 2007b; Meek et al., 2005; Högström and Smedman, 2004).

This literature has been incorporated into the revised discussion.

The structure of the paper is sometimes confusing.

We have taken on board the specific comments on this subject voiced below.

p.244, l.16-19: It is in deed interesting to compare not only the sensors alone but the entire measurement set-up and data analysis. However, in order to explain the differences, it is important to give more information about the differences in the methodology applied, and maybe analyse the impact of the data-processing separately, as has been done by Mauder et al. (2008; 2007c) for example.

We have added a table of the corrections applied to the individual systems to the manuscript and also a new Fig. 8a showing the average magnitude of the different corrections, some of which were applied by individual groups, while other were not (Table 2). A further analysis (e.g. analysing all datasets with a common software) is outside the scope of this paper as indicated in the introduction to this response.

p.244, l.22-24: The energy balance closure alone is not a very good measure for data quality of eddy covariance measurements. There are more fundamental tests for eddy covariance measurements available (Foken and Wichura, 1996;Vickers and Mahrt, 1997;Foken et al., 2004)

We fully agree and have not made a statement that it should ever be used as a sole

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

quality check. We have merely stated that it is used as one quality control element. The common use of this approach is now also backed up elsewhere with references, in response to another comment below. A test for stationarity is described elsewhere in the manuscript.

p.245, l.3-6: Please explain more clearly the motivation for this study not only for a readers within the GRAMINAE project but for a wider community.

We have expanded the text on the motivation in response to this referee.

Section 2 Theory: Most of this section is text book knowledge and can be omitted. A few sentences and references would be sufficient.

See response to a similar point raised by Reviewer 1.

Section 3.1 Field site: A paper has to be readable by itself. Therefore it is not enough to refer to other papers for description of the measurement site, set-ups, operation, measurements periods, participating research groups and abbreviations etc. Such a description does not need to be very extensive, but the basic information for understanding the results should be given here.

This has been added in response to Reviewer 1.

p.250, l.22-26: There is no need to mention measurements if the results are not reported in this paper. However, it might have been interesting to show these results since the gradient method was used in this project to measure trace gas fluxes.

We agree that the assessment of the gradient method against EC in this paper provides a useful link between the micromet intercomparison presented here and the trace gas flux papers. We have therefore added this assessment to the manuscript. The description of the gradient technique and measurements of wind speed and temperature gradients has therefore been expanded. By contrast, we have deleted the references to the gas analysers and the hygrometer gradient system, which did not perform well.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

p.251, l.20,21: Cospectral distributions not only depend on the surface roughness but also very much on the measurement height. Could you show some cospectral analysis or ogives to discuss the validity of the 15 min averaging time?

Turbulence scales with $(z-d)/u$, where u is affected by both surface roughness (affecting the gradient in u) and again the measurement height, and is further modified by stability. Comparing our measurements over grassland ($z-d = 2$ m; $z0 = 0.05$ m) with typical values of forest ($z-d = 10$ m; $z0 = 0.8$ m), under neutral conditions, for a given windspeed well above the surface (e.g., $z-d = 50$ m) of 6 m/s, the averaging period of 15 minutes over the short vegetation would be equivalent to an averaging period of 65 minutes in the forest situation. We have confirmed the levelling out of the ogives at low frequencies for CEH 1 and state this in the revised manuscript, but we do not think that the results warrant a further figure.

p.252, l.14-17: It is an interesting idea to use the median of all measurement to calculate the regressions. To this reviewer's knowledge, this is a novel approach, which makes sense. Usually one well-tested instrument is used as a reference.

We did not feel that we could single out one instrument (and setup and analysis system) as the reference and wanted to give all systems equal weight.

p.254, l.22-26: How was this averaging of 1-min data to 15-min averages done? What assumptions did you have to make? There is a precise formula available for this averaging, and you only need information about the number of measurements per time interval, no further assumptions are required (e.g. Mauder and Foken, 2004).

The reviewer is obviously correct, that the 15-minute values could theoretically be obtained in a way that would not require further assumption. We have re-evaluated the way this averaging was performed. The original calculation was a straight 15 minute blockaverage of 1-minute values. Incidentally, the reasonable agreement with the consensus dataset indicates that the averaging period of 15 minutes was more than sufficiently long. In the revised manuscript, we have recalculate the ECN data, taking into

account the covariance between 1-minute values. Unfortunately, the original online calculation of 1-minute fluxes incorporated a co-ordinate rotation (based on 1 minute averages) and only covariances with w were stored, insufficient to undo this rotation. Thus we are not able to calculate precise 15-minute values and some uncertainties (and scatter) remain. Taking into account of the co-variance between 1-minute values has slightly improved the scatter and the regression on u*, but has worsen the results for H.

p.255, l.5-8: There is no need to show the results for both friction velocity and the momentum flux. Either Fig. 1 or Fig. 2 can be omitted.

We expected this comment. As stated in the manuscript, tau is the parameter actually measured, while u* is the parameter used (e.g. in the flux-gradient relationship). This is the reason why we decided to present both. In response to the Referee we have decided to omit the graphs for tau, but to retain the results in Table 3.

Section 4.3 Comparison of the sensible heat flux: In order to interpret the differences between the sonics it would be important to know if the corrections according to Schotanus et al. (1983) and Liu et al. (2001) have been applied, since the impact of this correction can be very different for different sonic types.

This information can now be found in the new Table 2.

p.256, l.20-27, Fig. 5: Why bother showing the DWD measurements at all if they are not comparable?

We have removed the Rn measurements from Fig. 5, but still refer to the DWD measurements since other measurements (e.g. St) were used to derive other parameters in the consensus datasets, which are insensitive to the underlying surface.

Section 4.6 Ground heat flux: p.257, l.2-9 belongs into the Methods section, only l.10-13 is results.

The text has been moved to the Methods Section.

p.257, l.20-27: What is the reasoning behind using the maximum turbulent flux and the minimum Rn? Is there any physical explanation or is it just convenience? What is the relation to commonly debated causes for a lack of energy balance closure (Culf et al., 2004).

See comments above. The numerical experiment conducted here shows that energy budget closure is within the uncertainty of our measurements, but we agree that there are other, more likely, factors which could cause non-closure and now more clearly relate our results to the literature.

p.258, l.5: How long was the sampling line and what was the flow rate? Could you estimate a potential error due to damping effects?

This detail has now been provided in Section 3.2 and an error estimation (based on the CEH setup) has been presented in the new Fig. 8b.

p.258, l.8-10: This reviewer does not believe this explanation for the 14% higher latent heat flux estimates. This reviewer suspects that the calibration of the UMIST KH20 was erroneous. When and how was this KH20 calibrated? How often were the optical windows cleaned? By the way, this belongs to section 5 Discussion and not into Results.

We have clarified that we do not use the KH20 as an absolute sensor in the revised manuscript (see above). We have double-checked that the difference in the UMIST KH20 heat flux was not due to the absolute HUMITTER sensor used reading differently.

p.259, l.12,21: Terminology: An infrared thermometer is not called pyranometer but pyrometer. A pyranometer measures shortwave radiation. (also p.267, l.6)

Corrected.

p.260, l.6-9: What is the basis for your statement that T(z'o) and e(z'o) are robust parameters?

The emphasis of the statement was that these parameters are not robust under conditions of low turbulence; the extrapolation of T and e to z0'; according to Eqs. 18 and 19 becomes increasingly uncertain under low turbulence conditions, when Ra and Rb (i.e. gradients) are large. We have rephrased this in the revised manuscript.

p.261, l.18-20: This sentence belongs to section 1 Introduction. You might find the turbulence intercomparisons presented by Mauder et al. (2007c;2006) also interesting. Effects due to differences in post-processing methods are also analysed there.

These results are now cited in the discussion.

p.262, l.7-8: Conclusions belong into the Conclusions section.

Ok, this is a question of terminology. Is the Reviewer happier, if we say "However, the analysis presented here provides strong evidence for the main factors causing differences in the measurements of the individual parameters."?

p.263, l.6: What do you mean by latent heat flux correction for the measurement of H? If you refer to the Schotanus correction, this procedure usually reduces the sensible heat flux by 10-15% (Mauder and Foken, 2006). It is very important, whether a group has applied this correction or not. What about corrections to the latent heat flux, such as the correction for density fluctuations (Webb, 1982) or spectral losses (Moore, 1986;Moncrieff et al., 1997;Eugster and Senn, 1995;Horst, 2000;Horst and Lenschow, 2009;Horst and Oncley, 2006), have they been applied? This is particularly important to know since the measurement height and the sensor separation was not the same for the different instruments.

All this information can now be found in the new Table 2.

p.263, l.10-15: Where in the literature did you find the statement that low frequency flux contributions average out over time? This is new to the reviewer (cf. Lee et al., 2004; Mauder et al., 2007a).

We have retracted this statement.

p.263, l.16-19: This error depends on the averaging procedure. As mentioned above, it is possible to compose 15-min averages for 1-min averages accurately without any additional uncertainty.

See response above.

p.263, l.26-30: It is possible to correct for the high-frequency losses due to path-length averaging and low measurement height (Moore, 1986;Horst and Oncley, 2006). Is the agreement better with this correction applied? In this reviewer's opinion, it is not useful to compare flux estimates with incomplete post-processing.

As acknowledged by this Referee and Referee 1 in their overall assessments the value of this intercomparison exercise is to compare the results of the different groups as derived with their individual measurement protocols. To address the point raised by the Referee here fully (e.g. by reprocessing all datasets in a common way) would redirect the objective of the paper. Comparisons of instrumentations and processing packages have been presented elsewhere (see numerous references pointed out by the referee). Both micrometeorological theory and their routine implementation by the different groups has moved on since the data were worked up and passed on to the groups involved in the companion papers after the field campaign in 2000. It is regrettable that not all corrections were applied by all groups. However, it is clearly informative to know the magnitude of the different corrections. Thus, to address the concern of the reviewer as much as feasible within the context of this study, we have added a section where we estimate the magnitude of the different corrections (Fig. 8b) and their effect on the energy budget closure.

p.264, l.1: This reviewer cannot follow this line of argumentation. Normally the lower measurement height of the FAL-IUL system should lead to a smaller footprint and reduce problems with spatial heterogeneity.

If there were spatial variabilities within the field, the smaller footprint of the FAL-IUL is more likely to have been affected by a non-representative flux than the larger footprints

of the higher measurements. The argument is the same as why micrometeorological flux measurements at the field scale provide a more robust estimate of the average flux than chamber measurements, which only enclose small areas. We are here not talking about the effect of heterogeneity on the validity of the eddy-covariance approach.

p.264, l.8-9: Why don't you present a list of the correction procedures applied by each group?

The new Table 2 lists the correction procedures applied by the different groups and Fig. 8b provides an estimate of the magnitude of the different corrections.

p.264, l.9-10: It is not appropriate to use the KH20 humidity measurement for as absolute measurements. This instrument is designed for measuring fluctuations. Wasn't there any slow response humidity sensors deployed? You could potentially also use the closed-path IRGA measurements as absolute measurement.

This is a misunderstanding. The latent heat fluxes were derived by combining the relative measurement from the KH20 with the absolute measurement from a slow response sensor as explained in a reply above and in the revised mansucript. We here used Vaisala HUMITTER probes rather than the IRGA measurements, to keep results fully independent and to use the combination of instruments usually used by the groups in situations where the IRGA is not available, in the spirit of this intercomparison. The statement the Referee refers to referred to differences in the two HUMITTER probes which could have been responsible for the differences in the latent heat fluxes (but are not). This has been made clearer in the revised manuscript.

p.264, l.25-26: Are the UMIST KH20 measurements really realistic and all the other instruments are underestimating? This reviewer thinks the underestimation of turbulent fluxes is due to flux contributions which cannot be captured using a point-measurement with 15-min averaging time. Attributing the energy balance residual to the latent heat flux alone may be misleading.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

We have expanded the discussion of this source of underestimation.

Figure 10 is a very nice figure that summarizes clearly the results of the intercomparison. This result can be the centre piece for the discussion and conclusion.

We thank the Reviewer for this positive remark and have put a little more emphasis on this figure in the revised discussion. For example, we now explain the shape of the curve and derive an absolute and a relative contribution to the error.

p.265, l.14-18: This reviewer believes that the errors due to A/D conversion are minor. However, the stochastic nature of turbulence itself introduces a larger error that can only be reduced by choosing longer sampling intervals (Lenschow et al., 1994).

Unfortunately, in this case the choice of the A/D gain factor was unfortunate and may have added more uncertainty than in typical situations. We agree that the stochastic nature of the turbulence is a main reason for uncertainty in the measurements and this is clearly stated in the conclusions: the improvement of averaging over longer time-periods can be seen in the fact that 15-minute values of u* and tau vary significantly between sonic anemometers, while long-term averages agree closely. This is, however, not the case for IE, suggesting that there are systematic differences between the measurement setups. This is consistent with the results from other studies.

p.266, l.8-10: I agree it is potentially possible that the UMIST measurement of IE is the most accurate but this is very unlikely.

As an open path sensor, the KH20 is less dependent on corrections for high-frequency damping. However, the second KH20 derives fluxes of IE similar to the closed path sensors and data treatment was similar.

Section 6 Conclusions: It could very well be that the calibration of the UMIST KH20 is off and this system is overestimation the true IE. In recent literature, much of the lack of energy balance closure is attributed to longwave flux contributions, large stationary circulations, and turbulent organized structures (Finnigan et al., 2003; Kanda et al.,

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

2004; Inagaki et al., 2006; Mauder et al., 2006; Mauder et al., 2007a;Foken, 2008). These are methodological problems rather than instrumental problems. Therefore, it is "normal" to have an unclosed energy balance and an underestimation of the turbulent fluxes of 20-30% when the standard eddy covariance method is applied. Why should all other systems underestimate IE and only one system with a KH20 is correct? Based on the results of this intercomparison, what would be the recommendation for other experiments where only one set of instruments is available?

We agree that there are good reasons why the energy balance may not be closed and that low-frequency flux losses are likely to contribute. We were merely exploring whether all instrument combinations are implying non-closure and concluded that this is not the case. Since now flux loss corrections were applied to the IE closed path sensors, it would be quite possible that these somewhat underestimate the true IE. Contrary to this is the observation that the second KH20 sensor is more in line with the closed path sensors than with the UMIST KH20. With the data available we can only conclude that closure is within the measurement uncertainty between measurement systems. We have reworked the discussion to provide a more balanced view and discuss other reasons for non-closure in more detail.

p.244, l.21,22: Please write out numbers one to twelve.

Unlike in German (for example), the English convention is not so clear on this subject, but we are happy to spell out numbers one to twelve where it reads well.

p.245, l.12: Sutton et al. (1993) is probably not the best reference for the gradient method and the eddy covariance method.

We have removed this reference.

p.251, l.5: How far away was the DWD station?

This is now clarified through the inclusion of the new schematic of Fig. 1.

p.251, l.15: UTC and GMT are not identical, which one did you use?

UTC and GMT agree within a second and therefore identical within the accuracy relevant for this study. To avoid confusion, however, we have removed the reference to GMT.

p.253, l.1, Eq. 20: Please use unambiguous symbols. Doesn't T stand for temperature and Tau for the momentum flux?

New symbols have been introduced in the revised manuscript to avoid ambiguity.

p.253, l.13: Please correct: ... was used to derive a continuous ...

Corrected.

p.256, l.13-16: This belongs into the Methods section.

This text was modified as the DWD Rn measurement has been removed from the analysis. The remainder was moved to Section 3.2.

p.257, l.14-15: Please give references for this statement.

Two example references have been added.

p.262, l.19-20: Repetition of results in the discussion section should be avoided.

We disagree. The reference to individual values in Table 2 (now Table 3) is necessary to guide the reader through the discussion.

p.267, l.2: Please correct: ... this may not be the most ... Corrected.

Interactive comment on Biogeosciences Discuss., 6, 241, 2009.