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

Anonymous Referee #2

Received and published: 12 February 2009

General comments

This paper assesses the quality of turbulence measurements on the basis of an intercomparison experiments, where ten different measurement systems were operated side-by-side in the field independently by nine different institutes, so that not only the instrumentation but the combination of instrumentation, operation and data post-processing can be analysed. Such intercomparison experiments are very important tools for the assessment of the uncertainty of flux measurements. Other intercomparison experiments have been conducted before also with the same sensors, but this study is still interesting since it allows verifying the results of former intercompar-
isons and it includes the different analysis methods that are unique for the participating institutes. 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. 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. 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. 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. Most of the theory part presented in section 2 is text book knowledge and can be trimmed or omitted. 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). The use of the English language is mostly appropriate. The structure of the paper is sometimes confusing. In summary, this reviewer recommends that major revisions are required before this paper can be accepted.

Specific comments

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.

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)

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.

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

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.

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.

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?

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.
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).

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.

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.

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

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

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).

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?

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.

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)

p.260, l.6-9: What is the basis for your statement that $T(z'0)$ and $e(z'0)$ are robust parameters?

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.

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

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.

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).

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.

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.
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.

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

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.

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.

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.

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).

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.

Section 6 Conclusions: It could very well be that the calibration of the UMIST KH20 is off and this system is overestimating 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., 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?

Technical comments

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

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

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

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

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

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

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

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

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

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

References


2007b.


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