Intercomparison and assessment of turbulent and physiological exchange parameters of grassland

E. Nemitz¹, K. J. Hargreaves¹, A. Neftel²,*, B. Loubet³, P. Cellier³, J. R. Dorsey⁴,**, M. Flynn¹**, A. Hensen⁵, T. Weidinger⁶, R. Meszaros⁶, L. Horvath⁷, U. Dämmgen⁸, C. Frühaufr, F. J. Löpmeier⁴, M. W. Gallagher⁴, and M. A. Sutton¹

¹Atmospheric Sciences, Centre for Ecology and Hydrology (CEH) Edinburgh, Bush Estate, Penicuik, Midlothian, EH26 0QB, UK
²Agroscope Reckenholz-Tänikon Research Station ART, 8046 Zürich, Switzerland
³Institut National de la Recherche Agronomique (INRA), UMR Environnement et Grandes Cultures, Thiverval-Grignon, 78850, France
⁴School of Earth, Atmospheric and Environmental Sciences, University of Manchester, Simon Building, Oxford Road, Manchester, M13 9PL, UK
⁵Energy research Centre of the Netherlands (ECN), Postbus 1, 1755 ZG Petten, The Netherlands
⁶Eötvös Loránd University, Department of Meteorology, 1117 Budapest, Hungary
⁷Institute of Atmospheric Physics P. O. Box 39 1675 Budapest, Hungary

Received: 22 September 2008 – Accepted: 11 October 2008 – Published: 8 January 2009

Correspondence to: E. Nemitz (en@ceh.ac.uk)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Commonly, the micrometeorological parameters that underline the calculations of surface atmosphere exchange fluxes (e.g. friction velocity and sensible heat flux) and parameters used to model exchange fluxes with SVAT-type parameterisations (e.g. latent heat flux and canopy temperature) are measured with a single set of instrumentation and are analysed with a single methodology. This paper evaluates uncertainties in these measurements with a single instrument, by comparing the independent results from nine different institutes during the international GRAMINAE integrated field experiment over agricultural grassland near Braunschweig, Lower Saxony, Germany. The paper discusses uncertainties in measuring friction velocity, sensible and latent heat fluxes, canopy temperature and investigates the energy balance closure at this site. Although individual 15-min flux calculations show a large variability between the instruments, when averaged over the campaign, fluxes agree within 2% for momentum and 11% for sensible heat. However, the spread in estimates of latent heat flux ($\lambda E$) is larger, with standard deviations of averages of 18%. While the dataset averaged over the different instruments fails to close the energy budget by 30%, if the largest turbulent fluxes are considered, near perfect energy closure can be achieved, suggesting that most techniques underestimate $\lambda E$ in particular. The uncertainty in $\lambda E$ feeds results in an uncertainty in the bulk stomatal resistance, which further adds to the uncertainties in the estimation of the canopy temperature that controls the exchange. The paper demonstrated how a consensus dataset was derived, which is used by the individual investigators to calculate fluxes and drive their models.

1 Introduction

When measuring surface/atmosphere exchange fluxes of trace constituents at the canopy scale, usually one single set of instrumentation is used to provide the micrometeorological information necessary for the calculation of canopy scale fluxes. The same is true for the measurement of parameters that are used to drive parameterisations and models to predict the exchange, usually in the form of soil-vegetation-atmosphere transport (SVAT) models. Key parameters are wind speed ($u$), friction velocity ($u^*$) and the sensible heat flux ($H$) for the calculation of fluxes, while the parameterisations require input of photosynthetically active radiation (PAR) or solar radiation ($S_t$), air temperature ($T_a$), canopy temperature ($T_c$) and relative humidity (RH).

This paper utilises measurements made during the GRAMINAE Integrated Experiment at Braunschweig, Germany, to investigate the effect of differences between approaches and uncertainties in the results, using an array of instrumentation operated and analysed by a number of independent institutes. The main aim of the experiment was to investigate the dynamics of ammonia exchange between grassland and the atmosphere, as described in detail in accompanying papers (Sutton et al., 2008a). The flux analysis techniques were deliberately not standardised, although all groups involved have extensive experience in the application of eddy-covariance techniques. Pure instrument comparisons have been presented elsewhere (e.g. Dyer et al., 1982; Tsvang et al., 1985; Fritschen et al., 1992; Christen et al., 2000). Instead, this paper focuses on the differences that may be expected to be introduced by a combination of differences in instrumentation, chosen measurement height and analysis protocols, as they would be applied by individual groups in real applications.

The measurements included fluxes of momentum and sensible heat made with a total of 10 independent ultrasonic anemometers, operated by 9 different institutes from 5 different countries and analysed according to their respective protocols. Variability in the results is discussed, together with the strengths and weaknesses of the different techniques and, as a quality control, the closure of the energy budget is explored. The paper also compares different ways to establish the leaf temperature that controls biogenic emissions and investigates the propagation of errors into the parameterisation of bulk stomatal resistance at this site.

The paper describes how the measurements were analysed to provide a “consensus” micrometeorological dataset on which all participants could base their individual
analyses (Meszaros et al., 2008; Sutton et al., 2008b; Burkhardt et al., 2008; Loubet et al., 2009; Milford et al., 2008; Meszaros et al., 2008; Sutton et al., 2008b; Personne et al., 2008). At the same time, the intercomparison of the paper provides the basis to assess uncertainties in the measurement of turbulent exchange parameters, which is particularly relevant to the more usual interpretation of measured fluxes where only one set of sensors is available.

2 Theory

2.1 Eddy-covariance approach for measuring turbulent exchange fluxes

Several micrometeorological approaches are available to measure fluxes of momentum and heat at the canopy scale. The two approaches used here are the aerodynamic gradient method (AGM) and the eddy-covariance (EC) technique, which have extensively been described in the literature (Sutton et al., 1993).

Eddy-covariance measures the flux \( F_\chi \) of a scalar \( \chi \) directly as the covariance

\[
F_\chi = \overline{w'\chi'} = \overline{w\chi} - \overline{w} \overline{\chi}
\]

where \( w' \) and \( \chi' \) are the instantaneous deviations about the mean, of the vertical wind velocity (m s\(^{-1}\)) and the scalar, respectively. For measurements above homogeneous flat terrain, \( \overline{w} \) is expected to be zero and a non-zero value is usually attributed to a misalignment of the wind sensor. Therefore, a co-ordinate rotation is performed by all groups taking part in the Braunschweig experiment, to align \( u \) with the mean wind.

For this study, momentum flux \( (\tau) \), sensible heat flux \( (H) \) and latent heat flux \( \lambda E \) (W m\(^{-2}\)) were derived directly from the eddy covariance measurements using equations equivalent to Eq. (1):

\[
\tau = \rho w' u' \quad (2)
\]

\[
H = \rho c_p w' T' \quad (3)
\]

\[
\lambda E = \lambda \rho \epsilon P \overline{w' e'} \quad (4)
\]

where \( \rho \) is the density of air (kg m\(^{-3}\)), \( c_p \) is the heat capacity of air (J g\(^{-1}\) K\(^{-1}\)), \( \lambda \) is the latent heat of evaporation of water (J kg\(^{-1}\)), \( \epsilon \) is the ratio of the molecular weights of water and air (=0.622) and \( P \) is atmospheric pressure (kPa).

The friction velocity \( (u^*) \) may be calculated from the turbulence measurements as:

\[
u^* = \sqrt{-\frac{\tau}{\rho}} = \sqrt{-\overline{u'w'}} \quad (5)
\]

or

\[
u^* = \sqrt{\left(\overline{u'w'}\right)^2 + \left(\overline{v'w'}\right)^2}, \quad (6)
\]

both of which are used by different institutes. In atmospheric turbulence, the covariance between the stream-wise wind component \( (u) \) and the horizontal cross-wind component \( (v) \) is expected to be small. In addition to the previously described co-ordinate rotation around two axes, a third rotation was used here by individual groups to set this covariance to zero (Aubinet et al., 2000).

2.2 The aerodynamic gradient approach for measuring turbulent exchange fluxes

Eddy-covariance approaches can only applied for compounds for which fast-response sensors are available for measurement at a frequency for several Hz. For many highly reactive compounds such sensors do not generally exist, and here alternative, parameterised techniques are applied, which can utilise slow response measurements. Fluxes may be calculated as

\[
F_\chi = -u^* \chi^* \quad (7)
\]
where $u^*$ and $\chi^*$ may be derived from time-averaged gradient measurements, using the aerodynamic flux-gradient relationships (e.g. Flechard and Fowler, 1998):

$$u^* = -k \frac{du}{d[\ln(z - d) - \Psi_M \frac{(z-d)}{L}]}$$  (8)

and

$$\chi^* = -k \frac{d\chi}{d[\ln(z - d) - \Psi_H \frac{(z-d)}{L}]}$$  (9)

Note that in the literature the aerodynamic gradient approach is more often introduced in terms of a local gradient ($d\chi/dz$) of the logarithmic profile, or the differences between two heights ($\chi_2 - \chi_1$)/(z$_2 - z_1$). However, we present the approach in the (mathematically identical) form of a linear gradient (Eq. 8), as this can more easily be derived from measurements at more than 2 heights, by linear regression. In Eqs. (8) and (9), $k$ is von Karman’s constant (0.41) and $\chi$ is the mean scalar concentration at height ($z - d$), $z$ is the height above the ground, $d$ is the zero-plane displacement height, and $\Psi_M$ and $\Psi_H$ are the dimensionless integrated stability correction terms for momentum and heat, which can be calculated from the height and atmospheric stability as parameterised through the Monin-Obukhov length ($L$):

$$L = -\frac{u^* \rho c_p T}{kgH}$$  (10)

where $g$ is the acceleration due to gravity (m s$^{-2}$). Various formulations for calculating these stability corrections have been presented in the literature. In practice, a hybrid approach is often used, where $u^*$ in Eq. (7) is derived by ultrasonic anemometry, while $\chi^*$ is derived from averaged concentration profiles according Eq. (9).

2.3 Resistance analogy

For the purposes of determining the processes controlling the exchange of scalars such as ammonia, ozone, sulphur dioxide and nitrogen oxides, it is necessary to calculate the resistances to turbulent exchange. In the case of consistently deposited species it is often assumed that the concentration of the scalar at the absorbing surface is zero such that

$$R_t(z - d) = R_a(z - d) + R_b + R_c$$  (11)

where $R_t$ is the total resistance to transfer, $R_a$ is the aerodynamic resistance, $R_b$ is the laminar boundary-layer resistance close to the surface of the leaves and $R_c$ is the canopy resistance. The aerodynamic resistance, $R_a$, at $(z-d)=1$ m is obtained from Garland (1977):

$$R_a = \left( \frac{\rho c_p T}{kgH} \right) \frac{u(1)}{u^*} - \frac{\psi_h \left( \frac{1}{L} \right) - \psi_m \left( \frac{1}{L} \right)}{k u^*}$$  (12)

where the second r.h.s. term is zero in neutral and stable conditions. For the calculation of $R_b$, Owen and Thompson (1963) used the relationship

$$R_b = \frac{1}{(Bu^*)^{-1}}$$  (13)

where $B$, the sub-layer Stanton number was defined by Garland (1977) as

$$B^{-1} = 1.45Re^{0.24}Sc^{0.8}$$  (14)

Here, the roughness Reynold's number, $Re_*$, is given by

$$Re_* = \frac{z_0 u^*}{\nu}$$  (15)

and the Schmidt number, $Sc$, by

$$Sc = \frac{\nu}{D}$$  (16)
where $\nu$ is the kinematic viscosity of air ($m^2 s^{-1}$), $D$ is the diffusion coefficient of the scalar of interest ($m^2 s^{-1}$). There are a number of alternative approaches to calculate the sub-layer Stanton number (Wesely and Hicks, 1977; Sutton et al., 1993), but in practice the differences for $R_b$ are small for short vegetation. It should be noted that $R_b$ is specific for each chemical species, due to differences in $D$.

For chemical species that are exchanged with the plant through the leaf stomata, but not with the soil or leaf cuticles, $R_c$ may be substituted by the bulk stomatal resistance ($R_{sb}$). In other cases, where stomatal exchange is only one of several exchange pathways, $R_c$ may often be represented by a resistance network which contains $R_{sb}$ (Shuttleworth and Wallace, 1985; Sutton et al., 1998; Nemitz et al., 2001). For water vapour, if it is assumed that over a transpiring canopy with dry leaf surfaces, the bulk of the latent heat flux is transported via the stomates, then it is possible to calculate a bulk stomatal resistance, $R_{sb}$, from vapour pressure at the leaf surface, $e(z'_0)$ and saturated vapour pressure at the leaf surface temperature, $e_s(T(z'_0))$ as:

$$ R_{sb} = \frac{e_s(T(z'_0)) - e(z'_0)}{E} \quad (17) $$

The surface values can be calculated for a notional mean height of the canopy exchange ($z'_0$), from the values at a reference height ($z_{ref}$) and the turbulent fluxes, assuming the canopy to act as a big leaf:

$$ T(z'_0) = T(z_{ref}) + \frac{H}{\rho c_p} (R_a(z_{ref}) + R_{b,H}) \quad (18) $$

and

$$ e(z'_0) = e(z_{ref}) + E (R_a(z_{ref}) + R_{b,H,O}) \quad (19) $$

### 3 Methods

#### 3.1 Field site

A full description of the field site, measurement periods and site management may be found elsewhere (Sutton et al., 2008a). The field site was a *Lolium perenne* dominated agricultural grassland, which was cut 10 days into the 27 day measurement period (19 May to 15 June 2000), from 0.7 to 0.07 m canopy height, and which grew to 0.35 m by the end of the campaign. A large array of micrometeorological equipment was deployed over the canopy by several groups from different European research institutes. The bulk of this equipment was placed at “Site 1” (Loubet et al., 2009); in practice the sensors were distributed along a roughly north-south axis covered a distance of about 100 m along a transect through the field. The available fetch was approximately 300 m to the west and east of Site 1, 200 m to the south and 50 to 100 m to the north. A further, smaller array of instruments was located at “Site 2”, approximately 250 m east of Site 1 and close to the eastern edge of the field, which was bounded to the east by a deciduous shelterbelt approximately 8 m tall. The participating research groups and the abbreviations used for each have been presented elsewhere (Hensen et al., 2008).

#### 3.2 Instrumentation deployment

The measurements analysed here were made at nine eddy flux towers, all of which were equipped with an ultrasonic anemometer to measure fluxes of momentum and sensible heat. Only one of these eddy towers was operated at Site 2, while four towers also measured fluxes of latent heat (Table 1).

In addition to the eddy covariance measurements reported here, momentum fluxes and sensible heat fluxes were also derived from 2 wind-speed gradients (using cup anemometers) and 3 temperature profiles (using fine thermocouples). As these measurements showed larger variability than the eddy-covariance measurements, it was decided to exclude these from the analysis. Latent heat fluxes were also measured...
with the dewpoint hygrometer profile of a Bowen ratio system (Campbell Scientific). These measurements were rejected as the response time of the hygrometer was found to be insufficient for the switching frequency between the two heights, despite having followed the manufacturer’s guidelines. Several setups, including the nearby station of the German Weather Service (DWD), included measurements of solar radiation ($S_t$), photosynthetically active radiation (PAR) or net radiation ($R_n$), as well as absolute temperature and relative humidity. Here, the measurements of $R_n$ are compared to investigate the effect of uncertainties in $R_n$ on the energy budget closure. It should be noted that most groups also deployed a range of chemical analysers for gases and aerosols at each location, but these are described in the accompanying papers within this issue (Hensen et al., 2008b; Meszaros et al., 2008; Milford et al., 2008; Nemitz et al., 2008).

Several of the sonic anemometers formed part of relaxed eddy accumulation (REA) systems for NH$_3$ (CEH REA; ECN and FAL-IUL) (Hensen et al., 2008b). All groups calculated averaged data every 15 min, and clocks were synchronised to UTC (GMT) (local time minus two hours). The comparatively short averaging period was chosen because it was felt that the high time-resolution would maximize the information on NH$_3$ exchange processes. The frequency at which the spectral density functions peak increases linearly with measurement height. It was therefore estimated that the 15 min calculations at a height of about 2 m over the smooth grassland vegetation is at least comparable to an averaging time of 30 min over forest (Kaimal and Finnigan, 1994). The validity of this estimate is discussed below.

Slow sensors such as the different components of the gradient systems were recorded on data loggers (Model 21X, Campbell Scientific), while all fast data were recorded on PCs. With the exception of INRA and CEH 2, who used the commercial logging and analysis software Edisol 2.0 (Moncrieff et al., 1997), all institutes applied their own logging and analysis code (listed in Table 1).

### 3.3 Data analysis

The first stage of data analysis was performed by the individual research groups and involved filtering of the 15-min flux data to remove periods of instrument calibration, instrument malfunction or power failure. This coarsely filtered data was then drawn together and subjected to the following filtering procedure: the exact position of each instrument mast in relation to the other masts, mobile laboratories and other obstructions to the fetch was determined and all flux data falling within obstructed sectors were removed from that individual dataset. Where more than one group measured an individual parameter, the median of each of wind direction ($dd$), $u^*$, $H$, and $\lambda E$ from the eddy covariance systems, together with $S_t$, $R_n$ and PAR were then calculated and carried forward in the analysis to allow the validity of each individual dataset to be assessed by comparison with the median data.

In the case of $dd$ this was performed by a simple inspection of the time-series plots to confirm that no gross alignment errors were evident. The assessment of the extent to which each individual dataset was representative of the “consensus” dataset (cf. Sect. 4.10), consisted of performing a least-squares linear regression of the individual dataset on the median dataset.

In addition, values of the estimated length of fetch available for the wind direction observed during each 15 min period, and the cumulative normalised footprint (CNF after Kormann and Meixner, 2001) function were calculated and included in the consensus dataset. Flags were also provided for each 15 min value to indicate whether the measurements were in any way compromised by field conditions, thus allowing individual groups to filter the data according to their specific needs. Specifically, unsuitable micrometeorological conditions were defined as occurring under any of the following conditions: $u$ (1 m) $<$ 0.8 m s$^{-1}$; $-5 \, m < L < +5 \, m$; CNF $<$ 67% within the fetch. Non-stationarity flags were calculated following Dutaur et al. (1998) and Nemitz et
al. (2002), by defining $I(T)$ such that:

$$I(T) = \frac{1}{T} \int_0^T w' \chi' \, dt; \quad 0 \leq T \leq \tau. \quad (20)$$

The value of $I(T)$ was regressed on $T$ for each 15 min averaging period and the standard deviation of the regression line ($\sigma_f$), used to calculate the relative stationarity coefficient ($\zeta$) as

$$\zeta = \frac{2\sigma_f}{w' \chi'} \quad (21)$$

and periods of instationarity were defined as $\zeta>1.2$.

A consensus time-series of the zero plane displacement height ($d$) was derived from comparison of eddy-covariance results with the profile measured with cup anemometers (Vector Instruments) at 6 heights. For periods of near neutral stability ($\Psi_M \approx 0$) the value of $d$ in Eq. (8) was adjusted until the gradient estimate of $u^*$ matched the consensus value. This exercise was repeated for periods of varying $h_c$, to develop a relationship between $d$ and $h_c$, which was then used to derived a continuous time series of $d$ (shown in Fig. 8a).

4 Results

4.1 Initial data reduction

The first and second stages of data analysis (data filtering by institutes and filtering in relation to bad wind sectors) resulted in a reduction of the quantity of suitable flux data to between 52 and 82% at the individual measurement sites (Table 2). This reduction in data was a reflection principally of the degree of obstruction the individual masts experienced, rather than any inherent unreliability in individual systems.

In the following sections the different estimates are compared against a consensus dataset derived for Site 1. This was calculated as the average of those instruments that were deemed to provide equally reliable measurements for this site, as described in more detail below (Sect. 4.10).

4.2 Comparison of momentum fluxes

The comparison of the analysis of surface stress, or momentum flux ($\tau$) is presented in Fig. 1a to g. This indicates that with the exception of ECN (Fig. 1c), the average values of $\tau$ for each individual mast at Site 1 lie within $+7.9\%$ and $-7.7\%$ of the median value (as derived as the deviation of the slope from unity) an agreement judged to be very encouraging in view of the relatively large spatial distribution of masts in the field and the diverse nature of the anemometry, measurement height and eddy covariance software employed. While the median standard deviation between measurements for each 15 min period lies at $27.1\%$, averaged over the campaign, the standard deviation decreases to $2.0\%$ (Table 3). This indicates that differences are due to spatial and temporal fluctuations in the turbulence, rather than systematic differences.

Although the ECN data showed a discrepancy of $-16.1\%$ compared to the median, inspection of Fig. 1c shows that the least-squares regression was skewed by a relatively small number of scattered data points at low $u^*$ values and that the bulk of the data points lie along the 1:1 line. It was therefore decided to retain the ECN data within the consensus dataset for $\tau$ and $u^*$. The ECN data were taken as part of the ECN REA system and its data acquisition was not optimized for eddy-covariance application. Thus, although the system calculated the parameters needed for the REA calculations online, over suitable averaging periods, eddy-covariance results were stored every minute and had to be averaged in post-processing to provide 15-min values. Here additional assumptions had to be made to estimate the contribution of eddies in the frequency range between 1 and 15 min, explaining the higher variability.

The eddy covariance system at Site 2 (CEH EC2) (Fig. 1i) also produced values within 8% of the consensus dataset, but in this case the data were excluded from the...
consensus calculation on the basis that the spatial separation was in excess of 100 m and the mast was relatively close to the shelter belt at the eastern end of the field, although easterly winds were removed from the CEH EC2 dataset, when filtering for obstructed wind sectors.

While $\tau$ is the fundamental parameter, the parameter needed for the flux calculation is actually the friction velocity, for which the equivalent correlation plots are shown in Fig. 2. Due to the close link between $u^*$ and $\tau$, the scatter plots for $u^*$ show similar features.

4.3 Comparison of sensible heat flux

All sensible heat fluxes were calculated using the individual ultrasonic anemometers calculation of temperature based on the speed of sound in air. The results of the regression analysis are presented in Figs. 3a to h for Site 1, and in Fig. 3i for the single instrument at Site 2. For the majority of the instruments the discrepancy in the slope of the regression against the median value of $H$ lay in the range +5.3 and −6.9%, while the intercept, was less than 2 W m$^{-2}$, indicating how consistently the transition from unstable to stable conditions was measured. The exception to this rule was the ECN results, which again showed considerable scatter, for the reason described in the previous section. No systematic differences were found between different anemometer types.

4.4 Latent heat flux

Latent heat fluxes were measured solely at Site 1 using two open-path sensors of CEH and UMIST (Fig. 4a and d) and two closed-path sensors of INRA and CEH (Fig. 4b and c). Details of the different instruments used are summarized in Table 1.

Agreement between the four instruments for latent heat flux was poorer than that for sensible heat or momentum flux, with the CEH open- and closed-path instruments (Fig. 4a and b) providing similar medium estimates, the INRA system a lower and the UMIST system an upper bound. Possible reasons are discussed below (Sect. 5.2).

4.5 Net radiation

During the GRAMINAE integrated experiment at Braunschweig, fluxes of ammonia and other trace gases were either calculated by eddy-covariance (fluxes of latent and sensible heat, momentum, ozone, particles), hybrid aerodynamic gradient techniques (with the $u^*$ taken from sonic anemometry, NH$_3$, acid gases) or relaxed eddy accumulation (NH$_3$). Hence, net radiation ($R_n$) was not needed for the flux calculations per se as it would be the case in Bowen ratio or modified Bowen ratio techniques. However, the accuracy with which $R_n$ can be measured is important for the interpretation of the energy balance closure at this site. In addition, $R_n$ was needed to drive some of the numerical models, which incorporated their own heat balance calculation (Personne et al., 2008).

Three of the four net radiometers were operated at Site 1, while the fourth was operated at the DWD compound, 200 m to the SW, over continuously short “standard” grassland. The net radiometers were typically mounted at a height of 2 m and their footprint is therefore very different to that of the turbulent flux measurements. The CEH and FRI radiometers in particular showed a very tight relationship, while the INRA instrument shows some more variability. The DWD radiometers reported significantly smaller values of $R_n$ (Fig. 5b). These were calculated as the difference of a measurement of total downward radiation and total upward radiation. Substitution of the measurement of total downward radiation by the sum of an alternative estimate of shortwave downward radiation ($S_I$) and long-wave downward radiation (both also from DWD) provided much better agreement ($R_n$ (DWD, alternative) = 1.019$R_n$ (consensus) + 18.2 W m$^{-2}$; $R^2=0.963$, not shown). However, since these alternative values were only reported at hourly resolution (rather than 15 min resolution) and since the management of the grass in the DWD measurement compound was different, the DWD estimates were excluded from the consensus dataset.
4.6 Ground heat flux

Ground heat fluxes at the soil surface ($G$) were derived with two semi-independent systems as part of the setups of CEH and FRI at Site 1. In both cases, soil heat fluxes were derived at a depth of 8 cm, from duplicate measurements with heat flux plates (Campbell Scientific). To this was added the heat storage in the top 8 cm, calculated from changes in soil temperature (averaged over measurements at 2 and 6 cm depth within each setup, by soil thermocouples, Campbell Scientific), continuous measurements of the soil water content at one single site (by INRA) and measurements of the bulk density (average of two independent measurements of 1.35 and 1.65 g cm$^{-3}$).

The scatter in the comparison between the estimates of $G$ of the two different systems (Fig. 6a) is dominated by the disagreement at times in the soil heat storage (Fig. 6c), while the soil heat fluxes agreed closely (Fig. 6b).

4.7 Closure of the energy balance

The closure in the energy balance at the site is an often used test to assess potential losses in the turbulent fluxes. In the ideal case, the net radiation (balance of up- and down-ward short and long-wave components) should balance the sum of heat flux into the soil at the soil surface ($G$), and the turbulent fluxes of sensible heat ($H$) and latent heat ($\lambda E$). With the consensus dataset approximately 80% energy balance closure is achieved (Fig. 7), which is typical in the range of the energy closure observed elsewhere (Laubach and Teichmann, 1999; Wilson et al., 2002; Oliphant et al., 2004). As the array of instrumentation provides alternative answers for all parameters that feed into the assessment of the energy balance, an alternative (maximum) estimate of the energy balance closure may be compiled by considering the maximum turbulent fluxes ($\lambda E$ from the UMIST KH20 and $H$ from the FAL Gill HS anemometer) and minimum $R_n$ (from INRA) measured during the campaign. With these extreme values almost full closure is achieved (cf. Fig. 7). The 21% improvement of the energy balance is largely due to the increase in $\lambda E$ (+20% compared with the consensus dataset), with smaller contributions from the increase in $H$ (+6%) and decrease in $R_n$ (-7%). By contrast, choosing a single of the two ground heat fluxes ($G$) improves the energy balance only very little, because $G$ is on average much smaller than the sum of $H$ and $\lambda E$.

The fact that the open-path KH20 sensor of the UMIST setup derived the largest $\lambda E$ may be taken as an indication that damping effects in the sampling line and due to the sensor response time are not fully compensated for in the analysis of the two closed path IRGA systems. However, the $\lambda E$ estimate from the CEH KH20 is also 14% lower than that of the UMIST, despite a similar sensor configuration. This may, in part, be due to the faster anemometer and improved A/D converter of the UMIST Gill HS sonic compared with the CEH Gill R2.

Interestingly, the largest $H$ was derived with the FAL setup, which was operated at the lowest measurement height, where turbulence should be faster. This would be consistent with low frequency losses at an averaging time of 15 min at the higher heights (where turbulence structures are larger).

4.8 Solar radiation and PAR

Solar radiation ($S_t$) or PAR is needed to parameterise the stomatal resistance needed for SVAT modelling. The comparison of the three measurements of $S_t$ (by CEH, FRI and DWD) was very encouraging. CEH and FRI estimates were on average within 3% of each other, with the DWD estimate showing good agreement overall, but a larger amount of scatter. This was probably due to the spatial separation reflecting changes in cloudiness at the averaging scale of 15 min. The INRA PAR sensor derived a quantum flux which was 22% higher than that measured by DWD. Hence it was decided to use the more robust estimates of $S_t$ for parameterisations.

4.9 Comparison of canopy temperature estimates

Canopy temperature critically controls the potential for vegetation to react as a source of certain trace compounds. For example, isoprene emissions are known to be closely...
linked to leaf temperature. Similarly, ammonia emission potentials (compensation points) represent the gas phase concentration in equilibrium with the liquid phase NH₄⁺ concentration and the pH in the leaf apoplast. This gas-phase concentration is therefore governed by the temperature dependence of the Henry and solubility equilibria and, at ambient temperature, approximately doubles every 5°C (Sutton et al., 2001). Thus for the correct parameterisation of the emission potential, an accurate estimate of the leaf surface temperature is paramount. We here compare three different ways of estimating leaf surface temperature:

1. A micrometeorological estimate of the average canopy temperature is calculated as the surface value of the temperature, following the big-leaf approach of Eq. (18).

2. An infrared radiation pyranometer (KT19.85, Heitronics GmbH, Wiesbaden) and

3. Fine thermocouple wires, mounted to the surface of leaves at different heights and senescence stages.

The intercomparison of the different measures of canopy temperature are presented in Fig. 8 alongside the best estimate of the air temperature at z−d=1 m. The graph contrasts two four day example periods before and after the cut of the grassland from 0.75 m, between which the position of the thermocouples was necessarily changed.

Before the cut the vertical profile of the temperature of the green leaves is linked to light interception and the measured temperature profile in the canopy air space (not shown). The pyranometer measurement closely follows the temperature of the green top leaves of the canopy. By contrast, the micromet estimate of \( T(z') \) is more closely related to the temperature of the lower leaves in the canopy (where the bulk of the biomass is located) (Herrmann et al., 2008). This estimate also shows the largest diurnal range and values which appear to be lower or higher than the temperature of any physical element measured by the thermocouples.

After the cut, all leaves are exposed to the incoming radiation and here the colour (reflectance) and ability for evapotranspiration appear to account for the large difference of the physical temperature between green leaves as well as yellow/brown and senescent leaves of typically 10 K during warm days. While the pyranometer measurement reflects the temperature of the green leaves only, the micrometeorological estimate is heavily influenced by the dry vegetation.

4.10 Estimates of bulk stomatal resistance

The bulk stomatal resistance \( R_{sb} \) may be calculated from \( \lambda E \) according to Eq. (17), during periods when (a) \( \lambda E \) is dominated by evapotranspiration (leaf surfaces dry) and (b) the calculation of the surface values of \( T(z') \) and \( e(z') \) is reasonably robust \( (R_s + R_b \ll \text{small, i.e. windy conditions}) \). Former parameterisations (Jarvis, 1976) have shown \( R_{sb} \) to vary with LAI, PAR (closely related to \( S_z \)), leaf water potential and relative humidity (or water vapour pressure deficit, VPD). Light availability is clearly the main driver for stomatal functioning. However, prolonged dry and warm periods during the Braunschweig experiment meant that drought stress also had to be taken into account, together with changes in LAI during the management of the grassland. While LAI was measured only sporadically throughout the campaign, canopy height \( (h_c) \) was continuously monitored. Hence, a relationship between LAI and \( h_c \) was derived which allowed a continuous time series of LAI to be constructed (Fig. 9a):

\[
\text{LAI} = 1.8899 \times \ln(h_c) + 5.8483
\]

where LAI is in m² m⁻² and \( h_c \) is in m. The measurement derived estimate of \( R_{sb} \) is shown as circles in Fig. 9b. It clearly responds to the cut of the grass on 29 May. Although a parameterisation that ignores the water status (parameterised through VPD) can reproduce the measurement derived values of \( R_{sb} \) well on many days (Fig. 9b), it tends to under-estimate the \( R_{sb} \) on hot, dry days (e.g. 31 May–4 June). Inclusion of VPD into a parameterisation, based on the consensus data, leads to a much improved fit to the measurement derived values (based on Jarvis, 1976):

\[
R_{sb} = R_{sb, \text{min}} \left( 1 + \frac{b}{\max(0.01, \text{Stf})} \right) \frac{\text{LAI}_{\text{ref}}}{\text{LAI}}(1 - b_e \times \min(\text{VPD}, 2.5))^{-1}
\]
Here $R_{sb}$ is in $\text{m}^{-1}$, $S_t$ is in $\text{W} \text{m}^{-2}$ and VPD is in kPa. The fit parameters are $R_{sb,\text{min}} = 50 \text{ s m}^{-1}$, LAI$_\text{ref} = 5.18$, $b = 200 \text{ m}^2 \text{W}^{-1}$ and $b_a = 0.31 \text{ kPa}^{-1}$. As discussed in the previous section, analysis of the energy budget closure suggests that the larger UMIST $\lambda E$ may be a more accurate measure of the true evapotranspiration. Thus an alternative parameterisation of $R_{sb}$ was derived to fit the UMIST data, resulting in modified parameters of $R_{sb,\text{min}} = 30 \text{ s m}^{-1}$ and $b_a = 0.4 \text{ kPa}^{-1}$. The resulting resistances are typically 40 s m$^{-1}$ smaller during daytime, which is similar to the contribution of $R_s + R_b$ (Fig. 9c).

4.11 Generation of a consensus dataset

One of the reasons for the detailed intercomparison of the micrometeorological measurements was to produce a single, consensus dataset which all participants could use for further analysis of their individual measurements, such as the calculation of gas and particle fluxes and the parameterisations of models to reproduce the exchange. The measurements summarised in the consensus dataset were based on a 15 min mean for Site 1 and are summarised in Table 4.

5 Discussion

5.1 Sources of discrepancy in the estimates

Comparisons between ultrasonic anemometers have been presented in the literature (Dyer et al., 1982; Tsvang et al., 1985; Fritschen et al., 1992; Christen et al., 2000; Wieser et al., 2001). In these studies an attempt was generally made to keep all parameters and analysis procedures as similar as possible, to assess the effect of the instrumentation only. By contrast, this study deliberately compares the measurement results achieved by independent groups using their own instrumentation and analysis techniques, to assess typical uncertainties that would be expected by these differences.

Disagreement between individual sensors may generally be due to: (i) intrinsic differences in the instrumentation and sensor response times; (ii) differences in the mounting (e.g. potential interferences from gas inlets, REA; difference in turbulence scales at different heights); (iii) landscape heterogeneity (due to horizontal inhomogeneities and/or different footprint sizes associated with different measurement heights); (iv) statistical variations and (v) differences in the analysis procedures. The relative contribution of these factors is in general difficult to quantify. However, some important conclusions can be drawn from the analysis presented here.

The momentum fluxes (and the associated parameter $\nu$) shows significant variation between anemometers for each 15-min period, especially for low windspeeds. Averaged over the whole campaign, however, the different estimates are very close indeed, with a standard deviation of <1%, indicating that no biases are introduced by the instrumentation or the analysis techniques applied. The uncertainty in the momentum flux is dominated by spatial and temporal variability (which are conceptually similar, if Taylor's hypothesis is fulfilled). These findings are consistent with the study of Dämmgen et al. (2005), who operated an array of identical sonic anemometers, analysed with the same technique, to assess the averaging time required for the results to converge.

The standard deviation of sensible heat fluxes for each 15-min averaging period is on average 14.3 W m$^{-2}$, and here the campaign averages show similar variability ($10.8 \text{ W m}^{-2}$) (cf. Table 2). This indicates that there are systematic differences between flux towers. The sensible heat flux is derived from the speed of sound, averaged over the same volume as the momentum flux and, presumably, calculated with similar numerical routines as $\tau$. Hence, the reason for the small systematic differences is not immediately obvious. The way temperature is calculated from the speed of sound differs between anemometers. The Gill R1012 is known to have difficulties in measuring the speed of sound reliably at high windspeeds, as this is derived from only one transducer pair, whose distance may vary when the anemometer cage flexes under high windspeeds. This has been improved in the more recent models (such as the Gill HS) where all transducer pairs enter the calculation and distortions of the anemometer
cage will lead to compensating effects on the different transducers. Also, these newer
anemometers can now directly calculate the speed-of-sound temperature in the hard-
ware, while this calculation has to be performed off-line in the software for the R1012.
Indeed, the HS sonic anemometers of FAL and UMIST show a reduced amount of scat-
Fig. 3d and g), which was also observed in other studies (Christen et al., 2000).
Not all groups have applied the latent heat flux correction for the measurement of \( H \), as
latent heat fluxes were only measured as part of 4 of the 9 setups. However, assessment
of the biases between institutes (Fig. 3) does not reveal a consistent relationship
with anemometer model or latent heat flux correction.
In addition, some groups perform a high-pass filtering procedure on the raw data
(McMillen, 1988), to remove low frequency noise, while others have assumed that low
frequency variations contribute to the vertical turbulent flux and average out over time.
Both views can be supported by the literature (Finnigan et al., 2003, and references
therein). The former approach will tend to result in on average smaller fluxes and the
effect of this filter could indeed be larger on \( H \) than on \( \tau \).
The ECN data showed a large amount of scatter both for \( \tau \) and \( H \). As mentioned
before, the ECN REA setup recorded 1-min averages of the eddy-covariance results,
which had to be averaged in post-processing to derive their best estimate of the ex-
change parameters, which therefore shows higher uncertainty.
Interestingly, the FAL-IUL system derived one of the largest sensible heat fluxes
at the same time as it produced the smallest momentum flux. This instrument was
mounted at a considerably lower measurement height than the other systems (Table 1),
where the power spectrum of the turbulence is shifted towards higher frequencies. The
reason for this apparent inconsistency is not fully understood, but it may suggest that
momentum was on average carried by smaller and faster eddies than the heat flux.
Thus, the instruments operated at higher height may have suffered some low-frequency
flux loss of \( H \) due to the relatively short averaging time of 15 min, while the FAL-IUL sys-
tem may have suffered high-frequency loss of \( \tau \) due to the lower measurement height.
In addition, the flux footprint of the FAL-IUL would have been significantly smaller and
more affected by spatial heterogeneity. The reason for this lower measurement height
was that the FAL-IUL group wanted to test the setup as it was used back at their Swiss
GRAMINAE site, where the available fetch is more restricted than at the Braunschweig
site.
Significant differences were observed for the measurements of latent heat fluxes, with
the INRA system, based on an IRGA deriving a lower estimate and the UMIST system,
based on a Krypton Hygrometer, deriving the upper estimate. Possible reasons for the
disagreement are (a) differences in the flux losses in the setups and their correction
procedures and (b) differences in the absolute humidity measurement used for the cal-
culation of fluxes from the (not absolute) open path sensors. However, the absolute
humidities that were used for the flux calculations agree much more closely than the
fluxes and, unlike the fluxes, the UMIST system used slightly lower values than the
INRA system. It is therefore likely that flux losses and their treatment are the main
cause for the systematic differences. The Krypton hygrometer and IRGA operated
by CEH provided very similar results, indicating that the disagreement is not simply
a question of open vs. closed path sensors. The IRGA-based estimates differed pos-
sibly due to differences in the correction of flux losses. However, it is currently less

certain what causes the discrepancy between the two estimates based on the Krypton
hygrometers. This analysis should be similar to the calculation of sensible heat fluxes
which tended to be larger in the CEH setup than in the UMIST setup.

5.2 Energy balance closure
The consensus dataset fails to close the energy balance closure by about 20%, which
is well within the range reported by other authors (Laubach and Teichmann, 1999;
Wilson et al., 2002; Oliphant et al., 2004).
By selecting individual datasets full closure may be achieved, and this is largely due
to the larger \( \lambda E \) measured by the UMIST Krypton hygrometer. This closure suggests
that the UMIST \( \lambda E \) may in fact be the better estimate of the evapotranspiration and
suggests that about two thirds of the closure deficit of 20% may be due to deficiencies
in the correction of flux losses due to inadequate frequency response of the inlets and IRGAs used. This compares well with estimates of Oliphant et al. (2004), who attributed 16% to the same effect and concluded this error to be larger than heat storage within (forest) plant canopies.

5.3 Uncertainties in turbulent exchange in unreplicated measurements

The absence of systematic biases in the measurement of momentum fluxes is extremely encouraging for the calculation of surface exchange fluxes by the aerodynamic gradient technique, where $u^*$ is a key parameter, equally important as the measurement of the concentration profile itself. It implies that gradient flux estimates should be equally uncertain for each 15-min, but robust, if averaged over longer time-periods.

Figure 10 indicates what uncertainty may be expected for $u^*$ and $H$, when measuring with one unreplicated setup, as would be used in most studies. The uncertainty decreases with increasing value to 10% for $u^*$ values approaching 0.5 m s$^{-1}$ and 16% for $H$ values approaching 200 W m$^{-2}$. Hence, replicated measurements are most valuable when observing small fluxes.

There are several potential explanations: for example, there are constant absolute errors associated with the measurements (e.g. resolution of the analogue/digital converters), which make a larger relative contribution if the measured values are small. Christen et al. (2000) also reported enhanced inter-instrument variation in $u^*$ between R2 anemometers at $u^*<0.2$ m s$^{-1}$, indicating that the measurement accuracy of the wind speed makes an important contribution to the uncertainty of this anemometer.

In addition, the co-ordinate rotations become less certain under calm conditions. Planar fit rotations (Wilczak et al., 2001; Finnigan et al., 2003), which may overcome this problem, were not assessed as part of this study.

5.4 Uncertainties in the establishment and values of the consensus estimates

Spatial and temporal statistical variability has been identified as the main reason for the uncertainty in individual 15-min measurements of $u^*$ in particular. Thus, the compilation of a consensus $u^*$ based on 9 anemometers should have helped greatly in reducing the error of each 15 min measurement. The same holds true for other estimates that show random variability. For estimates that indicate clear systematic biases between setups, an individual (unbiased) measurement may in fact provide the more accurate answer than the consensus dataset. In particular, it is potentially possible that the UMIST measurement of $\lambda E$ is the most accurate measurement, as suggested by the assessment of the energy budget closure.

As statistical variability was found to be a major reason for the variability observed, the consensus dataset was calculated as the median of the different estimates rather than as the arithmetic mean. This accounts for the effect that turbulent parameters in the surface layer are log-normally distributed and it gives less weight to extreme outliers.

Figure 11 shows the time-series of an example period of the consensus values of $u$ (1 m), $u^*$, $T$ (1 m), $R_n$, $H$ and $\lambda E$, together with the standard errors as calculated from the statistical variation between the datapoints.

5.5 Uncertainties in parameters used for the parameterisation of exchange models

Stomatal resistances and leaf temperature are important drivers for the surface atmosphere exchange of many trace compounds. The uncertainty in $\lambda E$ has important implications for the calculation and parameterisation of the bulk stomatal resistance ($R_{sb}$). An increase in $\lambda E$ by 20% is shown to result in $R_{sb}$ which are 40 m s$^{-1}$ smaller during daytime, which is similar to the magnitude of the sum of $R_a$ and $R_v$. This implies that, during the day, uncertainties in the atmospheric resistances are of secondary importance.

Big-leaf approaches to derive bulk stomatal resistance from heat fluxes (e.g. by the
Penman-Monteith method) assume sensible and latent heat fluxes being driven by the same notional canopy temperature, \( T(z_0') \). By contrast, this may not the most appropriate temperature that governs the exchange of other trace gases such as VOCs and ammonia. A closer inspection of the temperature of different canopy elements reveals differences in leaf temperatures of up to \( \approx 10^\circ\text{C} \) during the day, and similar differences are found between the micromet estimate and a pyranometer measurement (Fig. 6). This variability in the temperature of individual surface elements has important influences on the parameterisation of trace gas exchange and the interpretation of ammonia exchange during the Braunschweig experiment: ammonia emission was observed not just after fertilisation, but also already after the cut, prior to fertilisation (Milford, 2004; Milford et al., 2008). Measurements of high ammonium concentrations in leaf litter suggest that the emission may originate from senescing plant material (Herrmann et al., 2008; Mattsson et al., 2008). The present analysis suggests that the micrometeorological estimate of the canopy temperature would tend to overestimate the day-time temperature of senescent material before the cut and underestimate this temperature after the cut.

In many situations, however, ammonia exchange is governed by the green foliage at the top of the canopy, the temperature of which appears to be overestimated by \( T(z_0') \). If stomatal compensation points derived from micrometeorological measurements of \( T(z_0') \) are used to estimate the ammonium concentration in the apoplast, a typical daytime overestimation of the real leaf temperature of 5\(^\circ\text{C} \) would underestimate ammonium concentrations by a factor of two. Similar effects would be expected where \( T(z_0') \) is used to derive temperature response curves for VOC emissions.

6 Conclusions

In this paper we have compared the results of micrometeorological measurements of turbulent exchange parameters, heat fluxes and parameters for modelling surface/atmosphere exchange fluxes, measured and analysed by independent laboratories, with the aim to assess typical uncertainties associated with difference in instrumentation and measurement practice. Of particular interest in the context of our study were parameters needed to calculate fluxes by the aerodynamic gradient technique and those required to model surface/atmosphere exchange of atmospheric ammonia. The results show that ultrasonic anemometry can be robustly applied to derive the key parameters (\( u^* \) and \( H \)) required to establish flux gradient relationships. Although \( u^* \) values of individual 15-min averaging periods can scatter significantly (median relative standard deviation of 13.8%), especially at low wind speeds, this variability averages out in time, leading to campaign averages with a standard deviation of only 0.7%.

Hence, the variability is caused by spatial and temporal variability of turbulence, rather than systematic differences in instrumentation or analysis techniques.

Larger uncertainties are associated with measurements of the latent heat flux (\( \lambda E \)), campaign averages of which showed a standard deviation of 17.8%. While the energy budget is only 70% closed using the “consensus” dataset averaged over all instruments that passed the quality criteria, the use of the largest measured \( \lambda E \) goes a long way in closing the energy balance. This would suggest that flux losses associated with \( \lambda E \) measurements remain a key reason for poor energy balance closure. These uncertainties propagate to a key parameter required to parameterise exchange fluxes, i.e. the stomatal resistance (which is derived from the latent heat fluxes), and adds to the uncertainty in leaf temperature estimates observed in this study.

Acknowledgements. These measurements were made in the context of the EU project “GRAM-INAIE”, while the final analysis was supported by the EU “NitroEurope IP” and national UK funding from the UK Department of Environment, Food and Rural Affairs (Defra), under the Acid Deposition Processes Project. The authors gratefully acknowledge the Institute of Agroecology of the German Federal Research Centre for Agriculture at Braunschweig-Völkenrode for access to the field site and provision of excellent site infrastructure.
References


Table 1. Summary of the instrumentation deployed on the eddy flux towers during the Braunschweig experiment.

<table>
<thead>
<tr>
<th>Short name</th>
<th>Site</th>
<th>Height(s) [m]</th>
<th>Logging/analysis software</th>
<th>Ultrasonic anemometer</th>
<th>H₂O sensor</th>
<th>Net radiation</th>
<th>Other parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEH 1</td>
<td>CEH</td>
<td>2.13</td>
<td>CEH EC software</td>
<td>Gill Solent 1012RA</td>
<td>IRGA (L-ConTok26); H₂O sensor</td>
<td>Rebo Q7</td>
<td>CD, flux, T gradient, RH, S₀, G, Soil T, U gradient, V gradient, wind direction (wind vane), soil moisture, rainfall, volumetric water content</td>
</tr>
<tr>
<td>CEH 2</td>
<td>CEH ECA</td>
<td>2.13</td>
<td>Edisol</td>
<td>Gill Solent 1012RA</td>
<td>IRGA (L-ConTok26); H₂O sensor</td>
<td>Rebo Q7</td>
<td>CD, flux, T gradient, RH, S₀, G, Soil T, U gradient, V gradient, wind direction (wind vane), soil moisture, rainfall, volumetric water content</td>
</tr>
<tr>
<td>CEH-REA</td>
<td>CEH-REAC</td>
<td>2.09</td>
<td>CEH EC software</td>
<td>Gill Solent 1012RA</td>
<td>IRGA (L-ConTok26); H₂O sensor</td>
<td>Rebo Q7</td>
<td>CD, flux, T gradient, RH, S₀, G, Soil T, U gradient, V gradient, wind direction (wind vane), soil moisture, rainfall, volumetric water content</td>
</tr>
<tr>
<td>DWD</td>
<td>near 1</td>
<td>various</td>
<td>N/A</td>
<td>N/A</td>
<td>Difference total up and down</td>
<td>DAQ</td>
<td>PAR, T, RH, U, wind direction (wind vane), precipitation, Lᵤ, Lᵥ, Lᵤₑ, Lᵥₑ</td>
</tr>
<tr>
<td>ECN</td>
<td>1</td>
<td>2.0</td>
<td>EC software</td>
<td>Gill Solent 1012R</td>
<td>IRGA (L-ConTok26); H₂O sensor</td>
<td>Rebo Q7</td>
<td>Leaf wetness (clip sensors), O₂ flux, O₂ concentration, S₀, G, Soil T, U gradient, V gradient, U profile</td>
</tr>
<tr>
<td>FAL/IUL</td>
<td>1</td>
<td>2.0</td>
<td>FAL/IUL software</td>
<td>Gill HS</td>
<td>IRGA (L-ConTok26); H₂O sensor</td>
<td>Rebo Q7</td>
<td>Leaf wetness (clip sensors), O₂ flux, O₂ concentration, S₀, G, Soil T, U gradient, V gradient, U profile</td>
</tr>
<tr>
<td>FRI</td>
<td>1</td>
<td>2.15 (2.00)</td>
<td>FRI software</td>
<td>Gill Solent 1012R</td>
<td>IRGA (L-ConTok26); H₂O sensor</td>
<td>Rebo Q7</td>
<td>Leaf wetness (clip sensors), O₂ flux, O₂ concentration, S₀, G, Soil T, U gradient, V gradient, U profile</td>
</tr>
<tr>
<td>INRA</td>
<td>1</td>
<td>2.04</td>
<td>Edisol</td>
<td>Gill Solent 1012R</td>
<td>IRGA (L-ConTok26); H₂O sensor</td>
<td>Rebo Q7</td>
<td>Leaf wetness (clip sensors), O₂ flux, O₂ concentration, S₀, G, Soil T, U gradient, V gradient, U profile</td>
</tr>
<tr>
<td>UMIST HS</td>
<td>1</td>
<td>2.02</td>
<td>UMIST software</td>
<td>Gill HS</td>
<td>IRGA (L-ConTok26); H₂O sensor</td>
<td>Rebo Q7</td>
<td>Leaf wetness (clip sensors), O₂ flux, O₂ concentration, S₀, G, Soil T, U gradient, V gradient, U profile</td>
</tr>
<tr>
<td>UMIST R2</td>
<td>1</td>
<td>2.86</td>
<td>CEH ASASP-x flux software</td>
<td>Gill Solent 1012R</td>
<td>IRGA (L-ConTok26); H₂O sensor</td>
<td>Rebo Q7</td>
<td>Leaf wetness (clip sensors), O₂ flux, O₂ concentration, S₀, G, Soil T, U gradient, V gradient, U profile</td>
</tr>
</tbody>
</table>
Table 2. Statistics of initial data reduction.

<table>
<thead>
<tr>
<th>EC Tower</th>
<th>Number of valid 15-min mean data points</th>
<th>% valid data</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEH 1</td>
<td>1797</td>
<td>78</td>
</tr>
<tr>
<td>CEH 2</td>
<td>1464</td>
<td>64</td>
</tr>
<tr>
<td>CEH-REA</td>
<td>1344</td>
<td>58</td>
</tr>
<tr>
<td>ECN</td>
<td>1637</td>
<td>71</td>
</tr>
<tr>
<td>FAL-IUL</td>
<td>1191</td>
<td>52</td>
</tr>
<tr>
<td>FRI</td>
<td>1419</td>
<td>62</td>
</tr>
<tr>
<td>INRA</td>
<td>1883</td>
<td>82</td>
</tr>
<tr>
<td>UMIST HS</td>
<td>1651</td>
<td>72</td>
</tr>
<tr>
<td>UMIST R2</td>
<td>1533</td>
<td>67</td>
</tr>
</tbody>
</table>

Table 3. Summary of the statistics of the measurements with the different setups.

<table>
<thead>
<tr>
<th></th>
<th>No of estimates</th>
<th>Mean Stdev</th>
<th>Median Rel. Stdev [%]</th>
<th>Rel Stdev of campaign averages [%]*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Momentum flux ($\tau$)</td>
<td>8</td>
<td>0.021 N m$^{-2}$</td>
<td>27.1</td>
<td>2.0</td>
</tr>
<tr>
<td>Friction velocity ($u^*$)</td>
<td>8</td>
<td>0.037 m s$^{-1}$</td>
<td>13.8</td>
<td>0.7</td>
</tr>
<tr>
<td>Sensible heat flux ($H$)</td>
<td>8</td>
<td>14.3 W m$^{-2}$</td>
<td>56.1</td>
<td>10.8</td>
</tr>
<tr>
<td>Latent heat flux ($\lambda E$)</td>
<td>4</td>
<td>20.6 W m$^{-2}$</td>
<td>25.1</td>
<td>17.8</td>
</tr>
<tr>
<td>Net radiation ($R_n$)</td>
<td>3</td>
<td>6.6 W m$^{-2}$</td>
<td>3.8</td>
<td>6.0</td>
</tr>
<tr>
<td>Solar radiation ($S_i$)</td>
<td>3</td>
<td>10.3 W m$^{-2}$</td>
<td>5.0</td>
<td>1.7</td>
</tr>
</tbody>
</table>

* Averages were calculated only over those periods where all measurement systems were providing data.
Table 4. Summary of data present in the consensus dataset.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Units</th>
<th>Derived from</th>
</tr>
</thead>
<tbody>
<tr>
<td>(d)</td>
<td>Wind direction</td>
<td>degrees</td>
<td>CEH EC1, CEH REA 1, ECN, INRA1, UMIST</td>
</tr>
<tr>
<td>(u^*)</td>
<td>Friction velocity</td>
<td>m s^{-1}</td>
<td>CEH EC1, CEH REA 1, ECN, INRA1, UMIST</td>
</tr>
<tr>
<td>(H)</td>
<td>Sensible heat flux</td>
<td>W m^{-2}</td>
<td>CEH EC1, CEH REA 1, ECN, INRA1, UMIST</td>
</tr>
<tr>
<td>(\lambda E)</td>
<td>Latent heat flux</td>
<td>W m^{-2}</td>
<td>CEH EC1 (IRGA), CEH EC 1 (KH20), INRA1</td>
</tr>
<tr>
<td>(R_n)</td>
<td>Net radiation</td>
<td>W m^{-2}</td>
<td>CEH1 BR, FRI1</td>
</tr>
<tr>
<td>(G)</td>
<td>Soil Heat Flux</td>
<td>W m^{-2}</td>
<td>CEH1 BR, FRI1</td>
</tr>
<tr>
<td>(S_i)</td>
<td>Incident Solar Radiation</td>
<td>W m^{-2}</td>
<td>CEH1 BR, FRI1, DWD</td>
</tr>
<tr>
<td>(PAR)</td>
<td>Photosynthetically active</td>
<td></td>
<td>INRA1, DWD</td>
</tr>
<tr>
<td>(d)</td>
<td>Zero-plane displacement</td>
<td>m</td>
<td>derived from wind profile and consensus (u^*) and measured (h_c) at Site 1</td>
</tr>
<tr>
<td>(h_c)</td>
<td>canopy height</td>
<td>m</td>
<td>interpolated from daily measurements</td>
</tr>
<tr>
<td>(u (1 m))</td>
<td>Wind speed at (z-d=1 m)</td>
<td>m s^{-1}</td>
<td>CEH EC1, CEH REA 1, ECN, INRA1, UMIST</td>
</tr>
<tr>
<td>(L)</td>
<td>Monin-Obukhov stability</td>
<td>m</td>
<td>derived from (u^*, H, T (1 m)) approx</td>
</tr>
<tr>
<td>(z_0)</td>
<td>Surface roughness length</td>
<td>mm</td>
<td>derived from (u, u(1 m)) and (L)</td>
</tr>
<tr>
<td>(R_n (1 m))</td>
<td>Aerodynamic resistance</td>
<td>s m^{-1}</td>
<td>derived from (L, u^*, u)</td>
</tr>
<tr>
<td>(R_{NH3})</td>
<td>Sub-layer resistance for NH3</td>
<td>s m^{-1}</td>
<td>derived from (z_0, u, T(1 m)) approx</td>
</tr>
<tr>
<td>(R_{SO2})</td>
<td>Sub-layer resistance for SO2</td>
<td>s m^{-1}</td>
<td>derived from (z_0, u, T(1 m)) approx</td>
</tr>
<tr>
<td>(R_{CO2})</td>
<td>Sub-layer resistance for CO2</td>
<td>s m^{-1}</td>
<td>derived from (z_0, u, T(1 m)) approx</td>
</tr>
<tr>
<td>(R_{H2O})</td>
<td>Sub-layer resistance for H2O</td>
<td>s m^{-1}</td>
<td>derived from (z_0, u, T(1 m)) approx</td>
</tr>
<tr>
<td>(R_{O3})</td>
<td>Sub-layer resistance for O3</td>
<td>s m^{-1}</td>
<td>derived from (z_0, u, T(1 m)) approx</td>
</tr>
<tr>
<td>(T(1 m))</td>
<td>Temperature at (z-d=1 m)</td>
<td>°C</td>
<td>CEH1 BR, INRA1, UMIST CPC1, FRI1 BR, …</td>
</tr>
<tr>
<td>(e (1 m))</td>
<td>Water vapour pressure at (z-d=1 m)</td>
<td>kPa</td>
<td>CEH1 EC, INRA1, UM11 CPC</td>
</tr>
<tr>
<td>(T(z_0'))</td>
<td>Micromet estimate of leaf surface temperature</td>
<td>°C</td>
<td>derived from (H) and (T (1 m)), uses (R_n) for H2O</td>
</tr>
<tr>
<td>(T(\text{surf}))</td>
<td>Surface radiant temperature</td>
<td>°C</td>
<td>UMIST KT19 IR Pyranometer</td>
</tr>
<tr>
<td>(e(z_0'))</td>
<td>Water vapour pressure at leaf surface</td>
<td>kPa</td>
<td>derived from (\lambda E ) and (e (1 m)), uses (R_n) for H2O</td>
</tr>
<tr>
<td>RH(1 m)</td>
<td>Relative humidity at (z-d=1 m)</td>
<td>%</td>
<td>derived from (T (1 m)) and (e (1 m))</td>
</tr>
<tr>
<td>RH(z_0')</td>
<td>Relative humidity at the surface</td>
<td>%</td>
<td>derived from (T(z_0')) and (e(z_0'))</td>
</tr>
<tr>
<td>(R_{st})</td>
<td>Stomatal resistance from water vapour transfer</td>
<td>s m^{-1}</td>
<td>derived from (e(z_0'), \lambda E, T(z_0'))</td>
</tr>
<tr>
<td>(P)</td>
<td>Precipitation</td>
<td>mm (15 min)^{-1}</td>
<td>DWD</td>
</tr>
<tr>
<td>(f)</td>
<td>Fetch for centre of site 1</td>
<td>M</td>
<td>derived from digitized field map and wind direction</td>
</tr>
<tr>
<td>CNFgrad</td>
<td>Cumulative normalized footprint function for gradient measurements</td>
<td>%</td>
<td>calculated according to Kormann and Meixner (2001), using stability corrected average height of gradient systems (Haendel and Gruenhage, 1999)</td>
</tr>
<tr>
<td>CNFEC</td>
<td>Cumulative normalized footprint function for eddy-correlation measurements</td>
<td>%</td>
<td>calculated according to Kormann and Meixner (2001) for a fixed height of (z=2.1 m-d)</td>
</tr>
<tr>
<td>(P)</td>
<td>Atmospheric pressure</td>
<td>kPa</td>
<td>UMIST</td>
</tr>
<tr>
<td>VPD(1 m)</td>
<td>Vapour pressure deficit at (z-d=1 m)</td>
<td>kPa</td>
<td>derived from RH(1 m) and (e (1 m))</td>
</tr>
<tr>
<td>(M)</td>
<td>Poor micrometeorological condition flag</td>
<td>–</td>
<td>Set if (</td>
</tr>
<tr>
<td>(I_m)</td>
<td>Non-stationarity flag (momentum flux)</td>
<td>–</td>
<td>Calculated according to Eqs. (20) and (21)</td>
</tr>
</tbody>
</table>
Table 4. Continued.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Units</th>
<th>Derived from</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_H$</td>
<td>Non-stationarity flag</td>
<td>–</td>
<td>Calculated according to Eqs. (20) and (21)</td>
</tr>
<tr>
<td>$I_{LE}$</td>
<td>Non-stationarity flag</td>
<td>–</td>
<td>Calculated according to Eqs. (20) and (21)</td>
</tr>
<tr>
<td>$I_{CO_2}$</td>
<td>Non-stationarity flag</td>
<td>–</td>
<td>Calculated according to Eqs. (20) and (21)</td>
</tr>
<tr>
<td>$I_P$</td>
<td>Non-stationarity flag</td>
<td>–</td>
<td>Set if $P&gt;0$</td>
</tr>
</tbody>
</table>

Fig. 1. Regression analysis of individual momentum fluxes as a function of median momentum flux.
Fig. 2. Regression analysis of individual friction velocities as a function of median friction velocity.

Fig. 3. Regression analysis of individual sensible heat fluxes as a function of median sensible heat flux.
Fig. 4. Regression analysis of individual latent heat fluxes as a function of median latent heat flux.

Fig. 5. Regression analysis of individual measurements of net radiation ($R_n$) a function of the median value of $R_n$. 
Fig. 6. Regression analysis of the two measurements of ground heat flux ($G$) and its two components, the soil heat flux at 8 cm depth and the heat storage within the top 8 cm of the ground.

Fig. 7. Assessment of the energy closure using the consensus dataset (black squares and black regression line) and the combination of maximum turbulent fluxes and minimum net radiation (blue crosses and dark blue regression line).
Fig. 8. Comparison between different estimates of canopy temperature, comparing a pyrometer, a bulk value derived from the micrometeorological parameters and measurements by thermo couple on a range of leaf types. (a) before the cut, grass height 0.75 m; (b) after the cut, grass height 0.20 to 0.26 m.

Fig. 9. (a) Time-series of zero-plane displacement height ($d'$) and leaf area index (LAI). (b) Parameterisation(s) of $R_{sb}$ in relation to measurement derived values. (c) Breakdown of consensus total resistance into aerodynamic ($R_a$), laminar-sublayer resistance ($R_b$) and bulk stomatal resistance ($R_{sb}$).
Fig. 10. Average relative standard deviations between eddy flux towers for (a) friction velocity and (b) sensible heat flux, in relation to the absolute magnitude of the values observed.

Fig. 11. Example time-series of selected parameters of the consensus dataset, together with error ranges (+/− standard error).