Agricultural and Forest Meteorology 151 (2011) 394–405

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Estimating uncertainties in individual eddy covariance flux measurements: A comparison of methods and a proposed new method

D.P. Billesbach*

Dept. of Biological Systems Engineering, University of Nebraska, 157 L.W. Chase Hall, Lincoln, NE 68583, United States

ARTICLE INFO

Article history:
Received 15 July 2010
Received in revised form 1 December 2010
Accepted 2 December 2010

Keywords:
Eddy correlation
Covariance
Variance
Scalar fluxes
Flux uncertainties

ABSTRACT

We present a review of four methods that are currently and formerly used to estimate random flux measurement uncertainties in eddy covariance instrument systems, examining their theoretical basis. In our review, we illustrate a few of the strengths and weaknesses of each, paying particular attention to the explicit and implicit assumptions made. We then evaluate the practicality, ease of use, and relative performance for the methods reviewed. We also describe a new method of random flux uncertainty estimation that is designed to only be sensitive to random instrument noise, and compare its predictions with other methods on both real and synthetic data. We conclude that our new method is a good complement to the others examined. We also suggest that when our proposed technique is used with one or more of the other methods, new information is obtained about how contributions to the total uncertainty are distributed among their various causes.

1. Introduction

The eddy covariance method for measuring the transfer of energy and matter between terrestrial ecosystems and the atmosphere has been in use for several decades (Baldocchi et al., 1988; Lee et al., 2004). This ability to measure trace gas and energy exchange has been one of the major facilitators in advancing global climate science.

Since its inception, there have been significant advances in the instrumentation used with this method (Lee et al., 2004). Much progress has also been made in understanding the theoretical basis of eddy covariance measurements and in relating them to the quantification of energy, momentum, and mass transfer between ecosystems and the atmosphere (Webb et al., 1980; Moore, 1986; Massman, 2000; Massman and Lee, 2002; Finnigan et al., 2003; Finnigan, 2004; Lee et al., 2004). Additionally, the processes and practical aspects of converting raw instrumental data streams into high quality fluxes that accurately and precisely represent their ecosystems have undergone considerable evolution. Many of these advances have been incorporated into standardized system designs, adopted for major flux networks (AmeriFlux, FluxNet, FluxNet-Canada, CarboEuroFlux, AsiaFlux, KoFlux, NEON, etc.).

While much progress has been made, there is still some confusion and ambiguity in the terminology that we use. For example, in the past, we often have used the phrase "Webb–Pearman–Leuning corrections" to refer to the buoyancy components described by Webb et al. (1980). Nowadays, it is becoming more common to refer to them as the "Webb–Pearman–Leuning terms"; according them their correct status as legitimate components of the solution to the continuity equation. Similarly, we often use the terms "error" and "uncertainty" interchangeably. By its very nature, the term "error" connotes a situation that can or should be corrected, minimized, or otherwise accounted for. "Uncertainty", on the other hand, suggests a quantification of the precision of a measurement. Sources of uncertainties and errors can be traced to biogeochemical (source/sink), transport, and instrument factors (Businger, 1986) and an accounting of all possible causes is far beyond the scope of this paper. While these definitions leave ample ground for semantic argument, we will in this discussion, focus on the random uncertainties that contribute to the overall precision of a flux measurement and assume that (to the extent possible), true errors have been minimized.

Businger (1986) presented an analysis of the accuracy with which flux measurements could be made by several different methods including eddy covariance. He summarized ten different sources of measurement error and uncertainty. In general, however we can classify uncertainty sources as either systematic or random (Moncrieff et al., 1996; Lasslop et al., 2008). Systematic uncertainties are what we often call errors and, if not eliminated or corrected, will produce biased flux estimates with poor accuracy. Because of the nature of the eddy covariance method, these biases are sometimes (but not always) difficult if not impossible to estimate post-priori. They are, however the subject of significant pre-measurement attention (Lee et al., 2004). This is seen in

* Tel.: +1 402 472 7961.
E-mail address: dbillesbach1@unl.edu

0168-1923/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.agrformet.2010.12.001

elements of the experiment design (instrument height and relative placement, calibration/maintenance schedules, etc.) and in the data post-processing procedures (spectral corrections, block averaging times, etc.) (Lee et al., 2004; AmeriFlux, 2010).

Random uncertainties, on the other hand do not introduce a bias into the resulting flux, rather, they reduce our confidence that the reported number is the true value (Businger, 1986; Young, 1962) (low precision). We can further categorize the sources of random uncertainties in eddy covariance measurements as being either correlated (between the two data streams) or uncorrelated (Moncrieff et al., 1996; Lasslop et al., 2008).

While there have been theoretical and practical methods proposed to estimate measurement uncertainties (Lumley and Panofsky, 1964; Shurpali et al., 1993; Mann and Lenschow, 1994; Clement et al., 1995; Billesbach et al., 1998; Meyers et al., 1998; Finkelstein and Sims, 2001; Hollinger and Richardson, 2005; Richardson et al., 2006), it is unfortunate that many experimenters have not taken advantage of them. Often uncertainties are simply estimated as a constant fraction of the measurement (Knorr and Kattge, 2005). More recently though, Hollinger and Richardson (2005) and Richardson et al. (2006) have advocated a more robust and realistic reporting of flux uncertainties while describing a new method of estimating them.

In this paper, it is our intention to make a critical examination and an intercomparison of several techniques for estimating random uncertainties and to propose a new method that will assist in separating the usually combined effects of random instrument noise from total measurement uncertainty. It is our hope that this analysis will provide researchers a basis for choosing and implementing an appropriate method and to provide another tool for evaluation of instrument system performance.

2. Uncertainty estimation methods

Several methods have emerged for flux uncertainty estimation. One unfortunate feature of all methods is that they are vulnerable to contamination from systematic errors. This is especially true of instrument calibration errors. Another unfortunate aspect of some methods is their reliance on arbitrary parameter choices. An ideal method would be independent of any system wide biases and not be dependent on subjective parameter choices. In this section, we will describe four different uncertainty estimation methods that have been, or currently are being used.

The first method was originally developed by Mann and Lenschow (1994) and will be referred to as the M&L method. Traditionally, the variances of the individual flux components (var(w) and var(x)) had been used as a basis for uncertainty estimation (Lumley and Panofsky, 1964; Lenschow and Stankov, 1986; Lenschow et al., 1994; Mahrt, 1998). A simple formulation was presented by Mann and Lenschow (1994) that was originally derived for aircraft work (Lenschow et al., 1994) and was shown to be adaptable for tower-based fluxes by Hollinger and Richardson (2005). They postulated that Eq. (15) of Mann and Lenschow (1994) could be used for tower fluxes with appropriate substitutions:

$$\frac{\sigma_F(L)}{|F|} = \left(\frac{2\tau_f}{L}\right)^{0.5}\left(\frac{1+r_{wx}^2}{r_{wx}^2}\right)^{0.5}(1-az_*) \quad (1)$$

where

$$r_{wx} = \frac{\mathrm{cov}(w,x)}{\sigma_w\sigma_x} \quad \tau_f = \frac{z}{\bar{u}} \quad az_* = 0$$

In this relationship, F is the measured flux, σ_F is the uncertainty of the flux, τ_f is the integral time scale of the measurement, L is the length of the measurement interval in seconds, r_{wx} is the correlation coefficient between w and x, z_* is a scaling height (determined as the ratio of the measurement height to the height of the convective boundary layer), and a is a scaling factor relating the surface flux to the flux at the top of the convective boundary layer. The product az_* relates the height of the measurement to the height of the convective boundary layer. In the original Mann and Lenschow (1994) context of aircraft fluxes, this factor could be considerable however for most flux towers, it will be nearly zero.

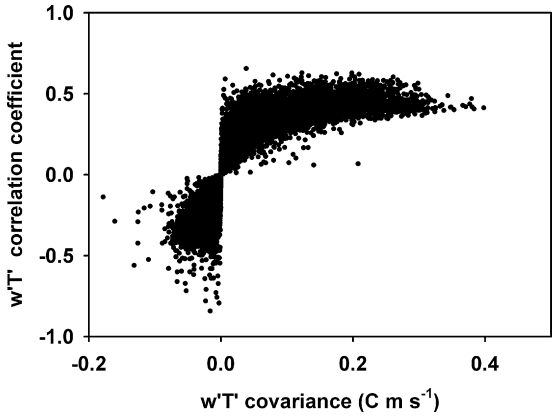

Fig. 1. Correlation coefficient plotted as a function of covariance [for cov(w, T)] over the entire SGP grassland data set.

The integral time scale τ_f can be estimated (under neutral stability) as the instrument height divided by the mean wind speed (Wyngaard, 1973). While this relationship is readily calculable, and uses well defined input parameters, it relies on implicit assumptions about the spectral and cospectral shapes and the atmospheric stability. Finkelstein and Sims (2001) suggested that a better approach would be to either fit a model to the autocorrelation function as suggested by Kaimal and Finnigan (1994) or to directly integrate the area under the autocorrelation curve. Their results however, show a relatively large and unexplained variation in values obtained for different covariances. For our comparison we will use the simpler approximation of Eq. (1).

The correlation coefficient, r_{wx} can be influenced by non-instrumental factors (Kaimal and Finnigan, 1994). Fig. 1 shows the relationship between cov(w, T) and r_{wT} for our Southern Great Plains (SGP) data set. Clearly, there is a functional relationship between them. This relationship would appear to impose a limit on the minimum value of the correlation coefficient which is dependent on the covariance. This in turn implies a biophysical limitation on the maximum covariance uncertainty. Because of this, the method will not uniquely describe the instrument system uncertainty; rather it will describe the instrument uncertainty combined with prevailing atmospheric conditions. In addition, examination of Eq. (1) shows that in the limit of perfect correlation ($r_{wx} = 1$), the second term on the right approaches its minimum value of only $\sqrt{2}$. This in turn predicts a minimum uncertainty (for given wind speeds and instrument heights) that is independent of the measurement system.

Because other flux terms and corrections (Webb et al., 1980; Moore, 1986; Massman, 2000; Lee et al., 2004) are derived from measured quantities like wind speed, air temperature, or even covariances, they will all have unique uncertainties associated with them. Since uncertainties in general, are not simply additive (see Appendix A and Young, 1962), the estimation method described above as well as the others described below should only be applied to the raw covariance (or flux). The individual uncertainties (associated with the covariance, corrections, and other terms) may then be combined as outlined in Appendix A. Only this total uncertainty should be associated with the fully corrected, final flux value.

A second method was employed by Verma and Billesbach (denoted the V&B method) in the early and mid 1990s for estimates of "minimum detectable fluxes" from tunable diode laser absorption spectroscopic (TDLAS) measurements of methane fluxes from various wetland ecosystems (Shurpali et al., 1993; Clement et al., 1995; Billesbach et al., 1998). In this technique, a tank of compressed gas (with a fixed methane concentration) was connected to the inlet of the TDLAS (a closed path gas analyzer), and the eddy covariance data collection software was run for several averaging periods. In principle, the sample gas stream from the tank contained no (or minimal) methane density fluctuations. Thus any flux computed from the raw data was assumed to be a measurement of the system uncertainty. Because this technique does not draw concentration samples from the ecosystem under investigation, it was thought to not mix valid biophysical processes with instrumental noise. There are however a number of drawbacks to this scheme. First, it is only applicable to closed-path gas analyzers. Second, there is a possibility of "contamination" from environmental effects. For example, external heating and cooling of the sample tube running from the tank to the analyzer can induce low frequency density changes that may be correlated with the vertical wind. Most significantly though, data taking must be interrupted to employ this method. In practice it was used sparingly and preferentially during times when flux data representative of the ecosystem were unlikely to be obtained (i.e. bad wind directions, low wind speeds, low turbulence, etc.). This had the potential to introduce a bias into uncertainty estimates if there were any correlations with relevant environmental factors. Additionally, since these measurements cannot be made simultaneously with real data, it must be assumed that the uncertainties obtained from this procedure vary slowly with time, and are representative of the system for intervals as long as weeks or months.

A third method is described by Hollinger and Richardson (2005) (Richardson et al., 2006; Richardson et al., 2008; Lasslop et al., 2008) (denoted the H&R method). This method assumes that two flux measurements (separated in either space or time) are identical and that their difference contains information about the measurement uncertainty. The method can use data from a single flux tower, or from two towers that overlook identical (but non-overlapping) fetches.

The paired tower approach assumes that the fetches and instruments are identical, and thus the difference between simultaneously measured fluxes contains information only about the random measurement uncertainty. This difference (multiplied by the square root of two) is assigned to the uncertainty of those particular flux measurements. To smooth out the variations in these estimates, a statistical moment such as the standard deviation about the mean value (or a similar parameter, appropriate to the particular distribution) of an ensemble of measurements can be constructed to represent the random measurement uncertainty for each instrument system during a given time interval.

A variation of this approach uses data from a single tower, and forms flux differences between periods where the state variables for the flux (the drivers) are identical or nearly identical. To minimize the effects of weaker, non-considered drivers, the differences are usually confined to the proceeding or following 24-h period. These flux differences are then interpreted and treated in the same manner as the two tower approach.

This method makes no apriori assumptions about the spectral or cospectral shapes of the data, and it is applicable to both open-path and closed-path instrument systems. Like the M&L method, it can be employed without interrupting normal data collection. There are however drawbacks to the method. In the two tower approach, besides the obvious problem of requiring two identical fetches and instrument systems, it is implicitly assumed that the two flux systems are essentially identical and that no mechanical or calibration biases exist between them. While this may be an achievable goal, its (ongoing) verification adds complexity to the overall site operation and maintenance. More importantly, there is the explicit assumption that both flux systems are and will be viewing identically functioning footprints. Given that this is an unverifiable condition, the uncertainty values obtained would have to be interpreted as a combination of the measurement uncertainty and the spatial footprint variability. In practical terms, when two flux towers are located in close proximity to each other, it is usually done to observe two distinct ecosystems or manipulations, making the identical fetch condition problematic. In addition, because the uncertainty estimates are obtained from averages over significant time intervals (days, weeks, or months), we again must assume (as in the V&B method) that they are slowly changing.

Similar arguments can be made in the single tower approach. We must assume that the two flux periods chosen for a single uncertainty value are well characterized by some implicit model, that all of the relevant parameters of the model are used to accept or reject the data, and that any errors or uncertainty in the ancillary instrumentation (defining the acceptance or binning parameters) are small enough to not cause false inclusions or exclusions of data. Also we note that there is no guarantee that an uncertainty estimate will be generated for each flux period. Uncertainty values are produced only when the parameter selection criteria have been met and when both periods yield good flux measurements as defined by the usual quality control criteria (Lee et al., 2004). Finally, there is no objective method to define the acceptance criteria used to select the state parameters defining the flux bins. If the flux is overly sensitive to these parameters, they may have a pronounced effect on the uncertainty estimates and their values will represent a subjective input to the results. Once again, we must interpret uncertainty values obtained from this calculation as the combination of temporal footprint variability and measurement uncertainty.

Finally we consider a method used by Meyers et al. (1998) and described by Finkelstein and Sims (2001) (denoted the F&S method). This approach is based on the direct statistical calculation of the variance of the covariance. The method is similar to the Mann and Lenschow approach, but it is a more direct calculation, and does not rely on knowledge of (or assumptions about) the spectral and cospectral shapes. Instead, a series of variably time-lagged variances and covariances (the auto and cross covariance functions) are calculated and summed. This improves on prior methods by considering all terms contributing to the variance while others had previously assumed some to be negligible (Wyngaard, 1973; Lumley and Panofsky, 1964; Mann and Lenschow, 1994). An initial drawback to the F&S method was the high computational cost associated with the many sums involved. Nowadays, this objection is mostly irrelevant. Another advantage is that like the M&L method, we can guarantee the generation of an uncertainty estimate for each flux period, eliminating the need to assume that the uncertainties vary slowly in time.

All of the above methods calculate the total uncertainty in eddy covariance flux measurements. Examination reveals that none effectively separate the two main components (sampling error and random instrument noise; Businger, 1986) that contribute to the total uncertainty. Independent knowledge of these components could yield useful information. The instrumental noise can serve as a generic diagnostic for flux measurement systems that would complement, but not rely on proprietary and non-standard diagnostic information generated by specific instruments. The contribution due to sampling error may be used to diagnose problems caused by sub-optimal experiment design. Since it is often assumed that fractional uncertainties may be simply added (see for example Goulden et al., 1996), we could, if we knew the instrument component and the total fractional uncertainty derive the sampling fractional uncertainty. While it is often assumed that the instrumental noise is

a minor component of the total flux uncertainty for CO_2 or energy fluxes, it is not necessarily true for systems measuring fluxes of minor trace gases (such as CH_4, N_2O, CO, O_3, $^{13}CO_2$, etc.).

3. Proposed method

In light of the above, we propose a simple method for estimating the contribution of random instrument noise to the total uncertainty in flux measurements which we call the "random shuffle" (denoted as the RS) method. We suggest that this contribution to the uncertainty of a covariance (or flux) value can be obtained by re-calculating the covariance after one of the variables has been randomly "time-shuffled". The idea is that this shuffle will remove the covariance between biophysical (source/sink) and transport mechanisms, leaving only the random "accidental" correlations attributed mostly to instrument noise. In essence, this amounts to forming the covariance with randomly selected values from the probability density functions.

As in the previous methods, this relies on several basic assumptions. For the most part however, they are assumptions that have already been made in the flux calculation. First, we assume that the covariance of two random number sets will be zero. Second, we assume that by randomly shuffling one of the covariance elements in time, we will theoretically force the correlation coefficient for the averaging interval to zero. We note, however, that there are some instances when noise may be correlated between two instruments such as power line electrical noise (if time coherence is maintained between the source of the noise and both instruments) or debris passing through a sonic anemometer averaging path (applies to sensible heat and momentum fluxes). These cases will, unfortunately be missed by this method. We also assume that the length of the block averaging period is sufficiently long to accurately represent the lowest significant frequencies contributing to the covariance. This can be verified if the peak in the circular correlation versus time shift plot (Chahuneau et al., 1989) is narrow compared to the width of the averaging period. It can also be verified by forming ogive plots of the covariance, looking for asymptotic behavior (Lee et al., 2004; Berger et al., 2001). If these conditions are not met, the flux itself will be underestimated, and this then constitutes a biasing error. We also assume (as do all of the other methods) that the instrument system has been carefully constructed and calibrated such that any systematic instrumental errors have been eliminated or minimized. It is worth noting that this technique makes no assumptions about the ecosystem being sampled, the representativeness of the footprint, or the spectral and cospectral shapes of the quantities being measured. In as much as the block averaging time has been correctly chosen (to include all relevant low frequencies), this technique will yield valid random noise estimates even when the measured fluxes are non-representative of the desired fetch. This is a feature that it shares with the F&S method. Other advantages of this technique include its post-priori nature. Like the M&L, H&R, and F&S methods, the errors are estimated in post-processing, and data collection is not interrupted for a "special" run. Unlike the H&R method, we accumulate system uncertainty information for every averaging period and for every calculated flux.

Since there are a very large number of possible randomly shuffled combinations, in practice we can calculate several RS uncertainty estimates for each period, and evaluate them based upon traditional statistical parameters (mean, variance, etc.) to assure us that the random shuffle has not accidentally landed on a partially correlated state. Because we obtain uncertainty estimates for every flux averaging period, we do not have to apply the statistical treatment of the H&R method to these values, however we may if desired.

3.1. Proposed method: algorithm and implementation

The basic mathematical operation used in eddy covariance data processing is summarized in Eq. (2).

$$\mathrm{cov}(w,x) = \frac{1}{N}\sum_{i=1}^{N} w'(t_i)x'(t_i) \tag{2}$$

(where primes denote deviations from the mean). In this relation, w represents the vertical component of the wind velocity, x is the target entity of the covariance, and N is the number of measurements contained in the block averaging period. To accurately represent the flux, it is assumed that the time indices have been adjusted (or that the data streams have been time shifted relative to each other) to maximize the covariance (Chahuneau et al., 1989). In our proposed method, we assume that the contribution to the total uncertainty from random instrument noise will be the value of the covariance when the correlation coefficient has been minimized and that this condition is best met when we randomize the time-order of one of the above sequences. We may then write the covariance uncertainty as:

$$\mathrm{cov}_{unc}(w,x) = \frac{1}{N}\sum_{i,j=1}^{N} w'(t_i)x'(t_j) \tag{3}$$

where the time index j has been randomized with respect to the index i. The details of the actual shuffle can vary. In our case, we used a harvested sequence of random numbers of length equal to the number of data points in the averaging interval. These were scaled to integers ranging from 1 to the number of data points in the interval. This set was then used to determine address swaps for an index array. This procedure guarantees complete randomization in a single pass, and that each value will be used exactly once. Other algorithms applicable to the theory of card shuffling would also suffice. One possible concern is the quality of the sequence of random numbers used in the index (time) shuffle. Given that most flux calculations will use averaging times of around 30 min and operate on 10 Hz or 20 Hz data, the periods of most random number generators will be more than adequate. A second concern is the potential "leakage" of spectral information from the unshuffled component into the covariance. We performed several tests where both components from real-world data were shuffled and the results were compared to single component shuffles. In all cases, we observed no discernable differences between them. We therefore completed the remainder of our analyses using only a single component shuffle.

Because each random shuffle generates a unique configuration for the averaging period, we can statistically combine multiple repetitions to create our uncertainty estimate. This approach has the advantage of minimizing the effect that a single, partially correlated state would have on the reported uncertainty. In practical terms, this procedure adds surprisingly little complication to the overall post-processing program. In our case, this was implemented with less than two dozen additional lines of Fortran code.

4. Experiment

We can test and compare these five uncertainty estimation techniques by applying them to various data sets. For the M&L, H&R, F&S, and RS methods, we used a period of 299 days from a set of tower flux data collected in 2005 in a Southern Great Plains (SGP) grassland environment, and a synthetic data set of length four days which was designed to produce no calculated flux. To examine the V&B method, we used methane flux data obtained in 1994 as part of the BOREAS project.

The SGP data were from an experiment that measured ecosystem fluxes at the USDA-ARS Grazinglands Research Laboratory

near El Reno, OK (35°33′N, 98°02′W, 423 m AMS). The site was a 33 ha ungrazed tallgrass prairie pasture with no more than a 1% slope to the south. Eddy covariance data were measured and collected with a system described by Billesbach et al. (2004). In summary, the flux system consisted of a Gill WindMaster Pro (Gill Instruments Ltd., Lymington, UK) (pre-2006 version) sonic anemometer/thermometer, a LiCor LI-7500 CO_2/H_2O open path IRGA (LiCor Biosciences, Lincoln, NE) and a small single board computer (replacing the laptop described in Billesbach et al. (2004)). Data were collected continuously at 10 Hz and were post-processed by a companion program which had been modified to calculate the desired uncertainty estimates. This program had been previously tested and certified using the AmeriFlux "Gold" data files.

The BOREAS data were obtained at the Southern Study Area Fen (SSA FEN) site (53°49′N, 104°37′W, 485 m AMS) located in central Saskatchewan, between June and October of 1994. The instrumentation considered in our analysis consisted of an ATI SATI/3SX 3 dimensional sonic anemometer/thermometer (Applied Technologies, Longmont, CO), a modified LiCor LI-6262 closed path CO_2/H_2O IRGA (LiCor Biosciences, Lincoln, NE), and a Unisearch EM-50 Tunable Diode Laser Absorption Spectrometer closed path CH_4 analyzer (TDLAS) (Unisearch Associates, Toronto, Ontario). The data collection system and instrument configurations have been previously described (Shurpali et al., 1993; Clement et al., 1995; Billesbach et al., 1998).

The synthetic data set was generated by a computer program written in Compaq Visual Fortran V6.6c (Hewlett Packard-USA, Houston, TX), using the built-in random number generator. After initializing the generator, a matrix (4 columns by 864,000 rows) of random numbers was harvested. This represented a full day of 10 Hz data. The four columns simulate the output of a sonic anemometer/thermometer (u, v, w, T). All three wind velocity components were generated to make this set structurally compatible with the existing post-processing program where a 2-axis coordinate rotation is first performed on the wind speed data. The individual columns were scaled to approximate typical fluctuations of u, v, and w wind speed components as well as sonic temperature. The scaled data were then written to our standard raw flux data file format. This was repeated four times to yield our four day, synthetic test data set. Because of the nature of this data set, we assume that the correlation coefficients are zero and that proper processing should yield zero fluxes. This leads us to expect that calculated uncertainties should only be the residuals of numerical round off and accidental correlations which are assumed to be similar to random instrument noise.

M&L uncertainties were calculated from the variances and covariances computed in the normal coarse of data post-processing using Eq. (1). This was done by a short routine added to our standard post-processing program.

A routine was added to our post-processing software to calculate RS uncertainties as described above. The random number sequences used to shuffle the data were again generated by the algorithm internal to Compaq Visual Fortran V6.6c. We tested the convergence properties of this procedure by examining the averages of ten trials of various numbers of repetitions for a single averaging period of our SGP data. The results (shown in Fig. 2) for sensible heat flux ($\mathrm{cov}(w, T)$) indicate that ten repetitions provide both good convergence as well as acceptable computational efficiency. We therefore used ten repetitions for the remainder of our analyses.

The H&R uncertainty estimates were calculated by a macro applied to post-processed data in an Excel-2003 (Microsoft, Redmond, OR) spreadsheet. After performing a sensitivity analysis, we used the state variables and tolerances suggested by Hollinger and Richardson (2005) for CO_2 fluxes, to either accept or reject the covariance measured 24 h after the target time. These variables and their limits were: $\Delta\mathrm{PPFD} \leq 75\ \mu\mathrm{mol\ m^{-2}\ s^{-1}}$, $\Delta T_{\mathrm{air}} \leq 3\ ^{\circ}\mathrm{C}$, and $\Delta u \leq 1\ \mathrm{m\ s^{-1}}$. If these conditions were not met, −9999.0 was recorded as the covariance uncertainty. If they were met, and if both of the covariances had no other problems (as determined by our normal QA/QC processing), the absolute value of their difference was divided by the square root of two and assigned to the uncertainty value for the earlier covariance of the pair. Hollinger and Richardson (2005) examined the statistical distribution of these differences and found that they were not Gaussian, and suggested using an alternate description of the distribution width ($\sqrt{2}\beta$), even though the difference between this metric and the standard deviation was often small. Because our similar calculations for other methods involved variances and standard deviations, we, on the other hand, used the standard deviation in our analysis of the H&R method.

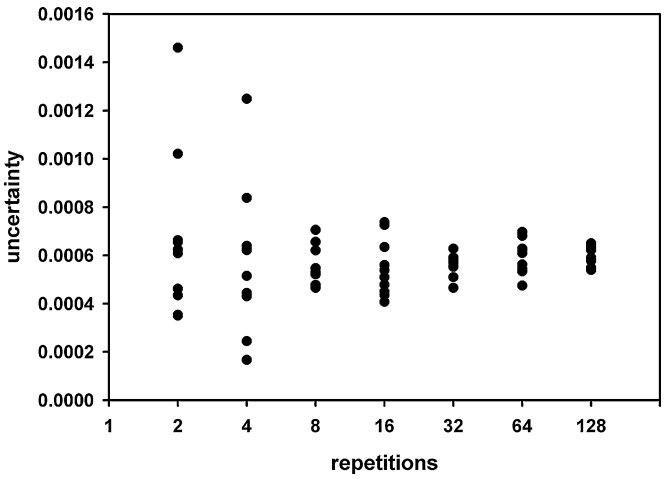

Fig. 2. RS (random shuffle) uncertainty convergence test. Mean calculated uncertainty for 10 trials of n repetitions.

The V&B uncertainty estimates were culled from the BOREAS data set. These consisted of periods where a compressed air tank was connected (through appropriate pressure and flow apparatus) to the TDLAS while normal data collection took place. The raw data from this experiment was re-processed for analysis by the programs described above which also produced M&L and RS uncertainty estimates.

F&S uncertainties were calculated for all data sets using an implementation of the algorithm described by Finkelstein and Sims (2001). The limits of the sums representing the auto and cross covariance terms should (in theory) run from plus to minus infinity. In practice, however we can (and must) truncate them to a value where additional terms only contribute minimally. Finkelstein and Sims (2001) suggested that for typical atmospheric data, a time delay of 10–40 s was sufficient. Since this amounts to an arbitrary parameter (like the selection parameters in the H&R method), we tested its effect on the results (convergence) by calculating uncertainties for offset times that ranged from 1 to 120 s. This was done for $\mathrm{cov}(w, T)$, $\mathrm{cov}(w, Q)$, $\mathrm{cov}(w, C)$, and $\mathrm{cov}(w, u)$ using data from DOY 178 of the SGP data set. The test was run for one half hour flux period when fluxes were large and one half hour period when fluxes were minimal. Like the M&L and RS analyses, this code was incorporated into our normal data processing programs. F&S uncertainty estimates were also generated for the BOREAS data set, but they revealed nothing unique and are not reported here.

As mentioned above, we performed a simple sensitivity test on the H&R technique in which we calculated uncertainties for five different values (centered around the nominal values used by Hollinger and Richardson (2005)) of each of the three state variables (125 different combinations) to observe their effects on the uncer-

Table 1
Parameter values used for the H&R (Hollinger and Richardson) sensitivity test.

Value	ΔPAR (μmol m^{-2} s^{-1})	ΔT (°C)	Δu (m s^{-1})
Minimum	25.0	1.0	0.5
	50.0	2.0	2.0
Nominal	75.0	3.0	1.0
	100.0	4.0	3.0
Maximum	125.0	5.0	4.0

tainty estimates. This was done for fluxes of sensible heat (H), latent heat (LE), CO_2, and vertical momentum transfer (u^*). The salient points of the analysis are contained in the maximum, nominal, and minimum parameter values; therefore for brevity and clarity, the other results are not reported.

5. Results

The values of the parameters used in the H&R sensitivity test are listed in Table 1 and a summary of the results is contained in Table 2. Several features emerge from this data. First, while the chosen parameter values may be suitable for CO_2 fluxes cov(w, C), they are not optimized for others such as latent heat cov(w, Q) or friction velocity cov(w, u). This is demonstrated by the wide range between minimum and maximum estimated uncertainties for these other covariances, especially for cov(w, u). The second feature concerns the actual number of uncertainty estimates generated. For the nominal parameter values, only about 17% of the measured fluxes (on average) received uncertainty estimates. For stricter conditions, this rate dropped to 4% while for looser conditions it increased to 42%. Finally, we see that the relative range of uncertainties varies between 18% and 845% of the minimum value. For comparative purposes, we used Hollinger and Richardson's (2005) nominal parameter values for the remainder of our analyses.

While we tested the convergence of the F&S time delay parameter for all of the covariances discussed above, we found similar results for them all. Fig. 3 shows the relative uncertainty (unc(t)/unc$_{max}$) for the high flux period of cov(w, T). All cases (all covariances for both high and low flux periods) showed over 85% convergence at 20 s maximum offset, confirming the recommendations of Finkelstein and Sims (2001). We therefore performed the rest of our F&S calculations using this value.

Fig. 4 shows uncertainty estimates in cov(w, T) for the entire SGP data set. We note that the M&L and the F&S methods both appear to be positively correlated with the covariance and seem to be of similar magnitudes as the H&R values. As expected, the magnitudes of the RS uncertainties are lower than any of the others by a factor of approximately two.

Table 3 lists the RMS averaged annual and seasonal values and uncertainties (cov(w, x)/unc(w, x)) for this data set for four key covariances, cov(w, C), cov(w, T), cov(w, Q), and cov(w, u) (which represent CO_2, sensible heat, latent heat, and momentum fluxes). These examples represent cases of positive and negative correlations from the same instrument (cov(w, T) and cov(w, u)) and positive and negative correlations from different instruments (cov(w, Q) and cov(w, C)). The values generally reflect the trends shown in Fig. 3 although we see that the H&R uncertainties are consistently larger than the others.

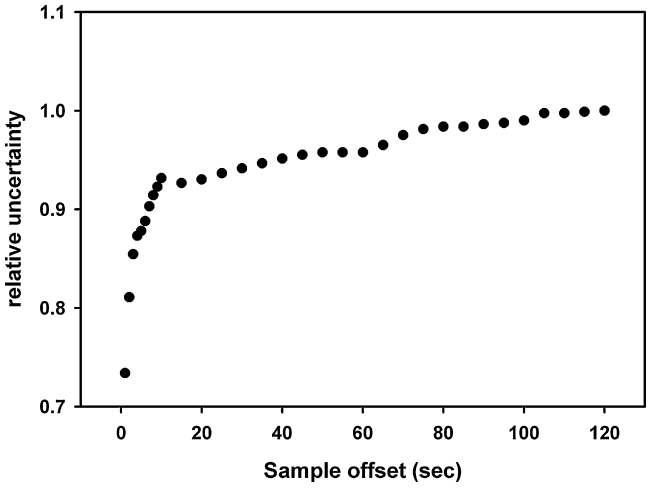

Fig. 3. Relative uncertainty of the F&S (Finkelstein and Sims) method as a function of sample time offset. The data shown are for cov(w, T) from DOY 178, 16:30 of the SGP grassland data set.

Fig. 5a and b shows covariances and uncertainty estimates for a selected four day period from the SGP data set. Because the number of valid H&R uncertainty estimates was small, we plotted their RMS average over the four day period (dashed line). Again, we see the same relative size trends that were evident in Fig. 4. We also note that the RS, M&L, and F&S methods all appear to be positively correlated with the absolute magnitude of the respective covariance.

Fig. 6 presents our analysis of the synthetic data set. Since the M&L method uses the correlation coefficient and because this data set was constructed to have no correlations, we assumed that M&L uncertainties would have limited validity and thus did not calculate them. In Fig. 6a, we note that the covariance is indeed nearly zero. We also see that while the uncertainty estimates for the H&R and the RS method are more similar than they were in Fig. 5a, the H&R values still appear to be slightly larger and show more scatter. Fig. 6b on the other hand shows that the F&S and RS methods seem to be converging, as there is little scatter, and their magnitudes are very nearly the same. Fig. 7 shows the relationships between the F&S and the RS (top panel) and between the H&R and the RS (bottom panel) methods. We note that while the F&S values appear strongly correlated with the RS ones, the H&R values do not.

Fig. 8 shows the results of uncertainty calculations for methane fluxes using the BOREAS data set. This was the only data tested that used a closed path gas analyzer and was therefore the only one where the V&B method was applied. The top panel shows the raw covariance between fluctuations in vertical wind speed and methane density. The bottom panel shows the RS, M&L, and the V&B uncertainties. Because only a limited number of V&B uncertainties were available, the RMS average of their values is plotted as a dashed line as we did previously for H&R estimates. Again, we see a strong positive correlation between the M&L uncertainty and the covariance. We also see that the RMS value of the V&B method appears to lie close to the yearly average for the M&L method.

Table 2
H&R (Hollinger and Richardson) sensitivity test results. The table lists covariances for the minimum, nominal, and maximum parameter values (see Table 1). Also listed, are the number of valid uncertainties obtained out of a possible 14,351 flux periods.

	$w'C'$ (mg m^{-2} s^{-1})	Num.	$w'T'$ (C m s^{-1})	Num.	$w'Q'$ (g m^{-2} s^{-1})	Num.	$w'u$ (m^2 s^{-2})	Num.
Min	0.1229	483	0.0121	751	0.0071	712	0.0293	386
Nom	0.1298	1872	0.0150	3065	0.0095	2843	0.0471	1738
Max	0.1446	4525	0.0207	7885	0.0117	7148	0.2768	4770

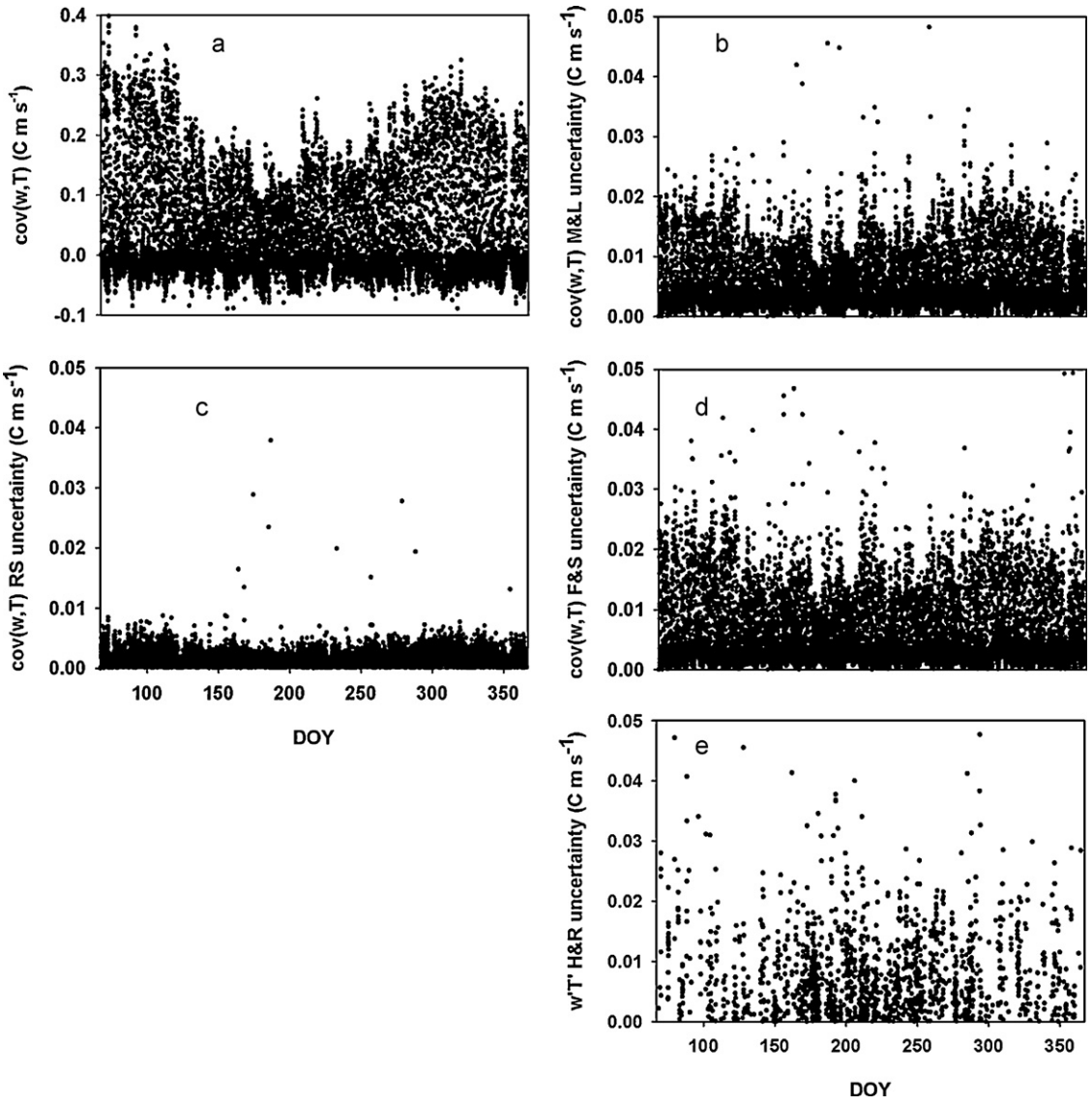

Fig. 4. Covariance and uncertainty estimates [cov(w, t)] for the full SGP grassland data set. Panel a is covariance, panel b is the M&L uncertainty, panel c is the RS uncertainty, panel d is the F&S uncertainty, and panel e is the H&R uncertainty.

Table 3
RMS averaged covariances and uncertainties (covariance/uncertainty) for the tested methods using the SGP grassland data set.

	$w'C'$ (mg m^{-2} s^{-1})	$w'T'$ (C m s^{-1})	$w'Q'$ (g m^{-2} s^{-1})	$w'u$ (m^2 s^{-2})
H&R				
All year	0.4349/0.0918	0.0517/0.0106	0.0421/0.0067	0.0974/0.0333
DOY 68–90	0.1709/0.0499	0.0638/0.0140	0.0089/0.0056	0.0837/0.0394
DOY 91–180	0.5144/0.0884	0.0488/0.0112	0.0528/0.0086	0.1042/0.0370
DOY 181–270	0.4807/0.1066	0.0433/0.0098	0.0434/0.0068	0.0796/0.0285
DOY 271–360	0.2254/0.0743	0.0671/0.0106	0.0141/0.0035	0.1117/0.0341
M&L				
All year	0.4349/0.0241	0.0517/0.0032	0.0421/0.0043	0.0974/0.0067
DOY 68–90	0.1709/0.0193	0.0638/0.0033	0.0089/0.0042	0.0837/0.0050
DOY 91–180	0.5144/0.0233	0.0488/0.0029	0.0528/0.0044	0.1042/0.0072
DOY 181–270	0.4807/0.0231	0.0433/0.0030	0.0434/0.0048	0.0796/0.0060
DOY 271–360	0.2254/0.0156	0.0671/0.0039	0.0141/0.0018	0.1117/0.0073
RS				
All year	0.4349/0.0063	0.0517/0.0008	0.0421/0.0009	0.0974/0.0022
DOY 68–90	0.1709/0.0047	0.0638/0.0009	0.0089/0.0007	0.0837/0.0016
DOY 91–180	0.5144/0.0065	0.0488/0.0008	0.0528/0.0011	0.1042/0.0025
DOY 181–270	0.4807/0.0064	0.0433/0.0007	0.0434/0.0010	0.0796/0.0019
DOY 271–360	0.2254/0.0040	0.0671/0.0010	0.0141/0.0004	0.1117/0.0024
F&S				
All year	0.4349/0.0313	0.0517/0.0038	0.0421/0.0060	0.0974/0.0095
DOY 68–90	0.1709/0.0238	0.0638/0.0043	0.0089/0.0055	0.0837/0.0073
DOY 91–180	0.5144/0.0295	0.0488/0.0040	0.0528/0.0056	0.1042/0.0106
DOY 181–270	0.4807/0.0341	0.0433/0.0024	0.0434/0.0073	0.0796/0.0083
DOY 271–360	0.2254/0.0207	0.0671/0.0047	0.0141/0.0025	0.1117/0.0101

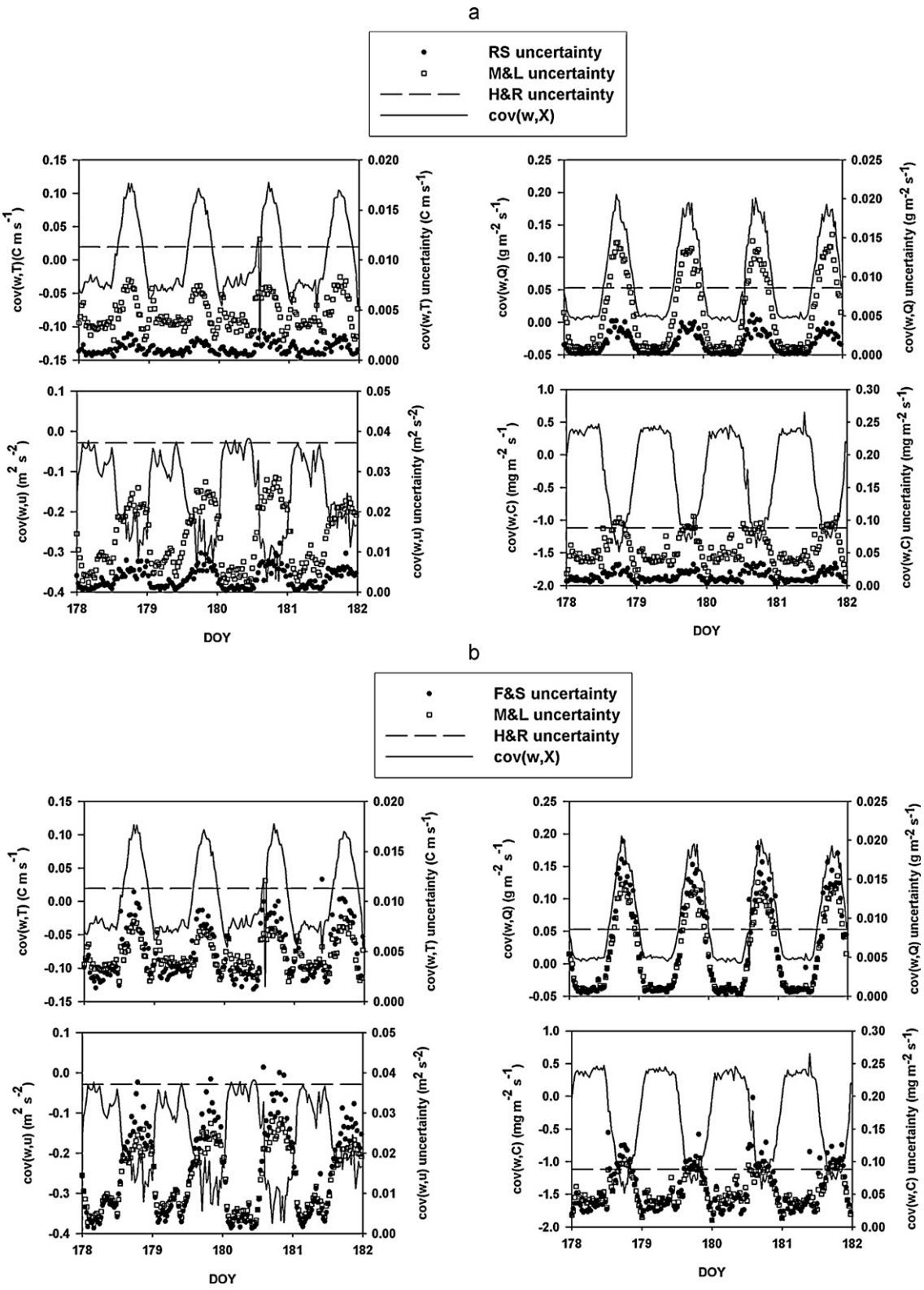

Fig. 5. Covariance and uncertainty estimates for four selected days of the SGP grassland data set. cov(w, T), cov(w, Q), cov(w, C), and cov(w, u) are shown clockwise, starting at the upper left. Panel a shows the covariance, RS uncertainty, M&L uncertainty, and the H&R uncertainty. Panel b shows the covariance, F&S uncertainty, M&L uncertainty, and the H&R uncertainty.

6. Discussion

Our analysis of the Mann and Lenschow (M&L) uncertainties suggests that, as expected, the measured covariances (or fluxes) have a large influence on the estimated uncertainties. This is primarily due to the dependence of the correlation coefficient on the covariance as illustrated in Fig. 1. Additionally, we see that in all cases, the M&L uncertainties are of similar magnitude to those esti-

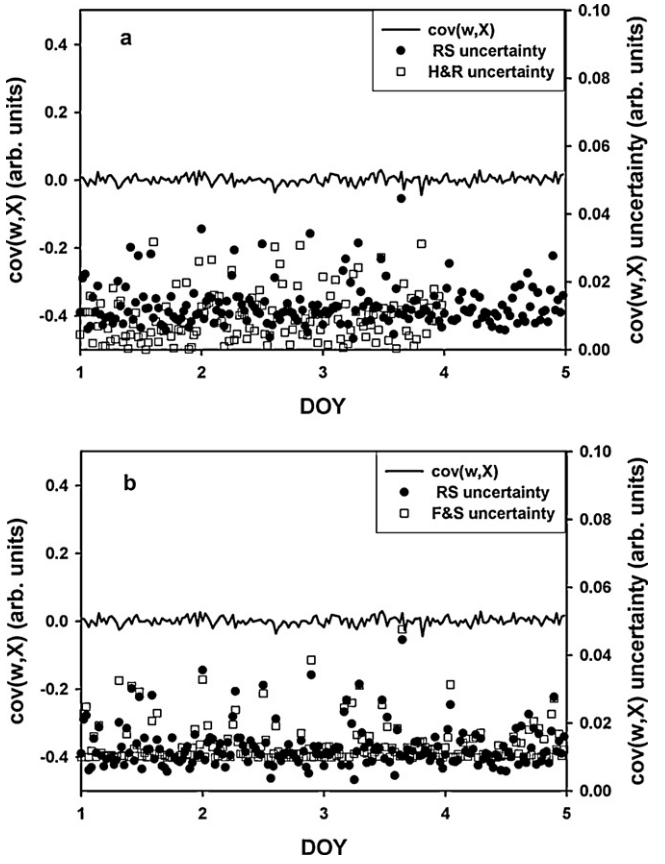

Fig. 6. Covariance and uncertainty estimates for the synthetic (zero flux) data set. The top panel shows the covariance, RS uncertainty, and the H&R uncertainty. The bottom panel shows the covariance, the RS uncertainty, and the F&S uncertainty.

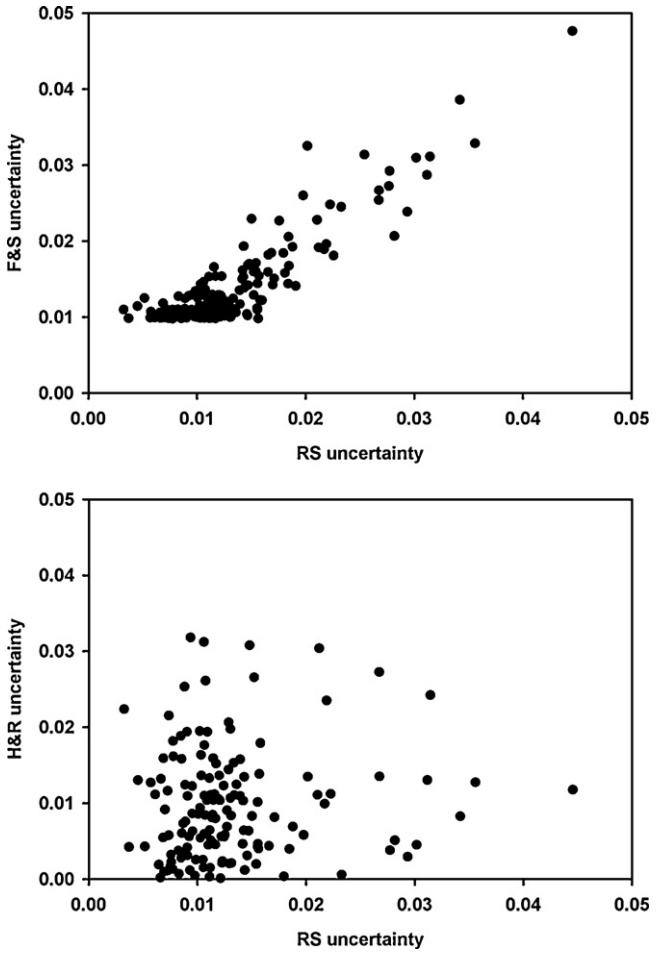

Fig. 7. Relationships between uncertainty estimates of the synthetic data set (zero flux) for the RS and F&S methods (top panel) and for the RS and H&R methods (bottom panel).

mated by the F&S method. This too is not unexpected as the M&L technique appears to be a simplification of the more rigorous F&S case.

The Verma and Billesbach (V&B) method (Fig. 8) generally predicts uncertainties that are on average similar to the ones obtained from the M&L method (and to the F&S method although the data are not shown). Unfortunately, there was not enough data available to determine whether this technique shows sensitivity to the covariance like the M&L method. Due to its disruptive nature (to normal data collection), this method can only be used sparingly in practice. In addition, since the method is only applicable to closed path gas analyzers, its utility is further limited. It is, however encouraging to find that the annual mean value obtained from this procedure is in rough agreement with uncertainties calculated by other methods.

The Hollinger and Richardson (H&R) method applied to real and synthetic data sets (Figs. 4–6 and Table 3) in general, produced uncertainty estimates that were larger than those derived from the M&L, F&S, or RS methods. The reasons for this are somewhat unclear, but we can speculate that they come from three sources. First, the two covariances used to generate the estimate may not be from identical fluxes. This is supported by the synthetic data shown in Fig. 6 where the H&R values are much closer to those obtained from other methods. Legitimately different fluxes from the two periods used to generate the H&R values would be a violation of the fundamental assumption of the method. Unfortunately, there is no practical way to either verify or nullify this assumption, and this has to be considered a weakness of the method. Second, the parameters used to select flux pairs and their values may not be ideal predictors of identical flux periods. This is especially true when applying parameters chosen for one flux to another. This limitation can be partially overcome by using specific variables for each flux. However, this does not address the problem that the underlying model (to which the parameters apply) may not be a universally good predictor of the flux. There is also no apriori method to choose the range of acceptable parameter values (or bin size) used to accept or reject covariance pairs. Finally, because it is less common to have acceptable flux values at night (due to low turbulence, low wind speeds, etc.) it may be more likely that H&R uncertainty estimates are generated during the day than they will be at night. Examination of Fig. 5a and b show that both M&L and F&S methods (and to a lesser extent, the RS method) exhibit strong diel trends with daytime uncertainties being larger than nighttime. This could result in a possible bias of H&R uncertainties toward the larger daytime values.

The different character of the H&R uncertainties (from the other methods) that is shown in Fig. 7 is difficult to explain. We assume that it must lie in how the methods fundamentally account for components of the uncertainties, either explicitly (in the case of RS and H&R) or implicitly (in the case of H&R). It would seem then that the RS and F&S methods are more closely related to each other than either of them is to the H&R method. This, in turn suggests that the H&R method generates different (and possibly complementary) information about the system than the other methods do. One could speculate that while the M&L and the F&S methods contain two components (instrument noise and sampling error), the H&R method contains an additional component related to fetch variability.

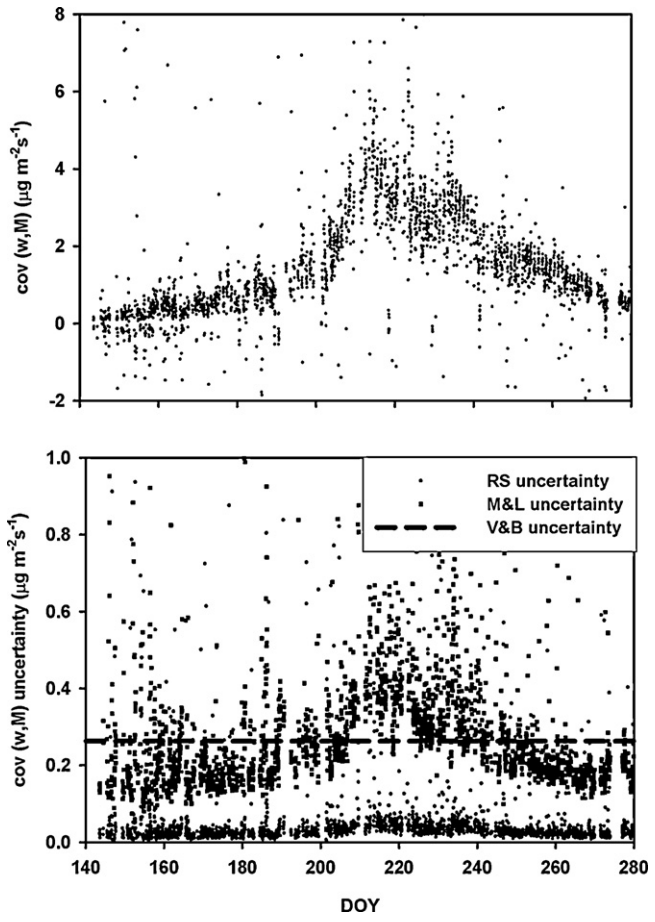

Fig. 8. Methane covariance (upper panel) and uncertainty estimates (lower panel) for the BOREAS data set.

As seen in Figs. 4, 5b and 6, the F&S method performed consistently in all cases. The magnitudes of estimated uncertainties were very similar to those obtained from the M&L method. This is also reflected in the data presented in Table 3. When applied to our synthetic data set, the F&S uncertainty magnitudes were very similar to those obtained by the RS procedure. This observation was reinforced by the high degree of correlation between the two as shown in Fig. 7. This behavior was to be expected because the synthetic data were constructed to contain no correlated components, which is equivalent to eliminating all uncertainties except random instrument noise. While not presented here, we calculated F&S uncertainties for the BOREAS data set as well, and observed that the qualitative features remained the same as those observed in the SGP data.

The proposed method, random shuffle (RS) was applied to all of the data sets. Universally, this method yielded lower uncertainty estimates when compared to the other four (M&L, V&B, F&S, and H&R) except when applied to the synthetic data set which was specifically constructed for zero flux and to contain no spectral content. In this case, RS, M&L, F&S, and H&R estimates were all of similar magnitude. This was not unexpected, and suggests that the difference in uncertainties in the other cases were due to spectral components, derived from environmental processes, that remained present in the H&R, M&L, F&S, and the V&B methods. In the case of the M&L method, these components are contained in the correlation coefficient. In the case of the F&S method, they are contained in the auto and cross covariance terms. In the cases of the V&B and H&R methods, they enter through the normally calculated covariances. The larger, noisier uncertainty estimates derived from the H&R method, however are probably the result of taking the difference of non-equivalent fluxes caused by real ecosystem differences, representing an additional component of the total uncertainty that is not contained in the other methods. Again, we expected the RS values to be universally lower than those from the other methods because they only represent part of the total uncertainty (that which is due to uncorrelated instrument noise) while the other methods include uncertainty from all sources.

As mentioned above, when we applied the RS procedure to our synthetic data set, which was designed to have no flux or spectral components, all of the uncertainty estimation methods tended to give the same or very similar values. This, along with the observation that the RS method gives significantly lower results for real data reinforces the idea that the RS method captures only the non-environmentally influenced instrumental noise as well as the arithmetic artifacts of the calculations.

As we have previously stated, one reason for not estimating flux uncertainties analytically has traditionally been the high computational cost. While testing these methods with our standard post-processing programs, we do, in fact notice a distinct slow down. The difference however, is not objectionable. In subsequent trials, in one of our on-going projects, we note that a week of raw data from a single flux tower can be processed in less than 10 min while generating both RS and F&S uncertainty estimates. The benefits of analytical, defensible flux uncertainty estimates vastly outweigh this minor inconvenience.

7. Conclusions

These observations lead us to several conclusions. First, we believe that while the Verma and Billesbach (V&B) method produces credible results, it is of limited practical value. The sporadic nature of the measurements and the limited instrumental applicability restrict its use to special situations. Unfortunately in this study, we only had a limited number of examples to test. A targeted study of this method that could produce a more statistically robust sample would be informative.

Because the Hollinger and Richardson (H&R) method relies on a number of arbitrary parameters and on implicit mechanistic models, we believe that the results do not represent the same information as those obtained from other methods. The larger variation between individual uncertainty estimates is probably due to "leakage" of real flux components. We also observed that the character of this method seems to be somewhat different (even on synthetic data) than other analytical methods. As yet, we have not been able to definitively explain this, but it does suggest that the H&R method might give unique and complementary information. Despite these criticisms however, the technique did generally agree with the other methods tested and most surely has uses, especially when raw, high frequency data are unavailable, or when it is desirable to combine temporal or spatial aspects of the fetch variability in the total.

Upon reflection, it seems that the H&R technique shares some of its basis with the "Lookup Table" method for eddy covariance gap-filling as described by Falge et al. (2001). Both use a set of parameter windows to select acceptable data for their respective purposes. Carrying this line of thinking further gives an interesting extension to the H&R method whereby the "binning" parameters could be applied to all data within a multi-day window to select many potentially identical fluxes. A statistical analysis of the differences between these values and the target period could then be used to evaluate the uncertainty in the target flux. This variation would potentially produce a higher percentage of uncertainty estimates and minimize their scatter. On the other hand, a larger time window would require a more robust implicit model and the

inclusion of all applicable drivers to bin the data. This concept has been recently explored by Richardson et al. (2008) and by Moffat et al. (2007). One aspect of the H&R method that should always be accounted for, however is the potential bias caused by the possibly higher percentage of daytime values.

If, as we speculate, the H&R method contains an additional uncertainty component that is not contained in the other methods, we can imagine using a combination of these techniques to uniquely separate them.

The Mann and Lenschow (M&L) method is probably the simplest and least "costly" to employ since it uses (in its simplest form) components that are routinely calculated in the normal course of data post-processing. While it performed well in these tests, it still relies on assumptions about the cospectral properties of the underlying data, especially in the choice of the integral time scale. Like the H&R method, it is probably most useful for re-analyses of existing data sets where the raw data are no longer available.

The Finkelstein and Sims (F&S) method appears to be well suited to both new data analysis and re-analyses of old data sets where the raw data are available. This method agrees well with the M&L method and yields results similar to the H&R and even the V&B methods. Its main advantage is the lack of adjustable parameters. Only one parameter needs be set, which is not dependent on any biophysical models, and can be chosen by a simple convergence test. We have shown that (in the case of our data sets), the method is reasonably insensitive to the value of this parameter and great care in its selection is not needed.

Another benefit of both the F&S and M&L methods is the guarantee that they will produce uncertainty estimates for every flux or covariance period. This is in contrast to the H&R method where the existence of an uncertainty value depends on meeting the set of eco-atmospheric "binning" conditions and on having valid covariances available. The availability of mean uncertainty values is certainly less desirable than uncertainties that were uniquely derived for individual covariances. This is also true of the V&B method which also cannot produce unique uncertainties for each observation period.

The random shuffle (RS) method proposed here is seen to not be equivalent to the other methods reviewed. The tests suggest that the RS method measures only the instrument uncertainty while the other methods measure the total uncertainty. Separation of the total uncertainty into components can be of interest, however when one is evaluating the performance of instrument systems. This allows the normal instrument characterization to move out of the laboratory and into the field. It also provides valuable, real-time feedback about the health of the instrument system as a whole.

In short, we recommend that researchers calculate both the F&S and the RS uncertainties during post-processing of their data. The F&S values may be used to report standard errors for each flux measurement, and the RS values can serve as an instrument system quality assurance feedback.

We must remember, however that these covariance uncertainty estimates describe only the raw covariances are not the complete uncertainty (or error bar) that should be attached to flux values. Additional uncertainties associated with other components of the flux such as WPL terms (Webb et al., 1980) or frequency corrections (Moore, 1986; Massman, 2000, see also Massman and Clement, Ch. 4 in Lee et al., 2004) must be correctly propagated through the relevant relationships as illustrated in Appendix A.

Acknowledgments

This research was supported by the Office of Biological and Environmental Research of the U.S. Department of Energy, under Contract No. DE-AC02-05CH11231, as part of the Atmospheric Radiation Measurement Program (DOE-ARM). The author also wishes to express his gratitude to Dr. Shashi Verma for sharing his raw data from the BOREAS experiment. He also expresses his thanks to his colleagues at the Lawrence Berkeley National Laboratory, the USDA-ARS Grazing Lands Research Laboratory, and from the DOE-ARM Climate Research Facility for their help in obtaining the SGP dataset.

Appendix A.

The propagation of random uncertainties through a mathematical transformation is a straight forward although not always transparent procedure. The basics are described in Young (1962) and involve calculations of the relevant partial derivatives of the transformation equation. We begin by assuming that we wish to propagate uncertainties in the independent parameters through some transform given by:

$$f(x_1, x_2, \ldots) \tag{A.1}$$

Young (1962) then defines the uncertainty (or variance) of the transform in terms of the independent parameters x_i as:

$$\sigma_f^2 = \sum \left(\frac{\partial f}{\partial x_i} \right)^2 \sigma_{x_i}^2 \tag{A.2}$$

where σ_{xi} represents the uncertainty in parameter x_i. As an example of this process, we will explicitly derive the total uncertainty for CO_2 fluxes calculated from an open path IRGA with both WPL terms (Webb et al., 1980) and frequency corrections (Moore, 1986; Massman, 2000, see also Massman and Clement, Ch. 4 in Lee et al., 2004) incorporated. The complete flux is given by:

$$F_C = f_I \operatorname{cov}(w, \rho_C) + \rho_C \left[\frac{\mu}{\rho_d} f_I \operatorname{cov}(w, \rho_V) + \frac{(1+\mu\sigma)}{T} f_S \operatorname{cov}(w, T) \right]$$

$$\text{where} \quad \mu = \frac{M_d}{M_V} = 1.61 \quad \text{and} \quad \sigma = \frac{\rho_V}{\rho_d} \tag{A.3}$$

where w represents vertical wind speed, ρ_C is CO_2 density, ρ_d is dry air density, ρ_V is water vapor density, T is air temperature, M_d is the molecular weight of dry air, and M_V is the molecular weight of water. The terms f_I and f_S represent the frequency corrections for covariances calculated for the IRGA and the sonic anemometer respectively and may have come from the formulation of either Moore (1986), Massman (2000), or any other estimation. We see that there can be measurement uncertainties in all of the components except μ. If we then take the partial derivatives of Eq. (A.3) and substitute them into Eq. (A.2), we get:

$$\begin{aligned}
\sigma_{F_C}^2 = {} & f_I^2 \sigma_{\operatorname{cov}(w,\rho_C)}^2 + \left(\frac{\rho_C \mu}{\rho_d} f_I \right)^2 \sigma_{\operatorname{cov}(w,\rho_V)}^2 + \left(\frac{\rho_C (1+\mu\sigma)}{T} f_S \right)^2 \sigma_{\operatorname{cov}(w,T)}^2 \\
& + \left(\frac{\mu}{\rho_d} f_I \operatorname{cov}(w,\rho_V) + \frac{(1+\mu\sigma)}{T} f_S \operatorname{cov}(w,T) \right)^2 \sigma_{\rho_C}^2 + \left(\frac{\mu \rho_C}{\rho_d^2} \left(f_I \operatorname{cov}(w,\rho_V) + \frac{\rho_V f_S \operatorname{cov}(w,T)}{T} \right) \right)^2 \sigma_{\rho_d}^2 \\
& + \left(\frac{\mu \rho_C f_S \operatorname{cov}(w,T)}{\rho_d T} \right)^2 \sigma_{\rho_V}^2 + \left(\frac{\rho_C (1+\mu\sigma) f_S \operatorname{cov}(w,T)}{T^2} \right)^2 \sigma_T^2 + \left(\operatorname{cov}(w,C) + \frac{\mu \rho_C}{\rho_d} \operatorname{cov}(w,\rho_V) \right)^2 \sigma_{f_I}^2 \\
& + \left(\frac{\rho_C (1+\mu\sigma)}{T} \operatorname{cov}(w,T) \right)^2 \sigma_{f_S}^2
\end{aligned} \tag{A.4}$$

This rather daunting expression is the analytical form for the square of our final flux uncertainty. In practice, several of the terms may be negligible and thus ignored. For instance, the term inversely proportional to the square of the absolute temperature will probably be small. Of course it is always up to the user to determine how many and which terms to keep in this type of calculation. While implementation of this relationship may seem overly tedious, one must remember that once coded and tested, the calculation will be automatic. We also caution that the above example is not to be interpreted as having universal applicability. This is a simple example designed to illustrate the process of properly propagating independent uncertainties. If the users flux system, for example, employs of a closed path gas analyzer, the exact procedure will differ from the one outlined above. In practice, users should always assess their own systems in terms of the recommendations of Lee et al. (2004). In light of the benefits gained by having analytic uncertainties associated with every flux measurement, it seems well worth the effort to include these procedures in data post-processing.

References

AmeriFlux web site, measurement guidelines and standard operating procedures (2010). http://public.ornl.gov/ameriflux/sop.shtml.

Baldocchi, D., Hicks, B.B., Meyers, T.D., 1988. Measuring biosphere–atmosphere exchanges of biologically related gases with micrometeorological methods. Ecology 69, 1331–1340.

Berger, B.W., Davis, K.J., Yi, C., Bakwin, P.S., Zhao, C., 2001. Long-term carbon dioxide fluxes from a very tall tower in a Northern Forest. Part I. Flux measurement methodology. J. Atmos. Ocean. Technol. 18, 529–542.

Billesbach, D.P., Kim, J., Clement, R.J., Verma, S.B., Ullman, F.G., 1998. An intercomparison of two tunable diode laser spectrometers used for eddy correlation measurements of methane flux in a prairie wetland. J. Atmos. Ocean. Technol. 15, 197–206.

Billesbach, D.P., Fischer, M.L., Torn, M.S., Berry, J.A., 2004. A portable eddy covariance system for the measurement of ecosystem-atmosphere exchange of CO_2, water vapor, and energy. J. Atmos. Ocean. Technol. 21, 639–650.

Businger, J.A., 1986. Evaluation of the accuracy with which dry deposition can be measured with current micrometeorological techniques. J. Clim. Appl. Meteorol. 25, 1100–1124.

Chahuneau, F., Desjardins, R.L., Brach, E., Verdon, R., 1989. A micrometeorological facility for eddy flux measurements of CO_2 and H_2O. J. Atmos. Ocean. Technol. 6, 193–200.

Clement, R.J., Verma, S.B., Verry, E.S., 1995. Relating chamber measurements to eddy correlation measurements of methane flux. J. Geophys. Res. 100, 21047–21056.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N., Katul, G., Keronen, P., Kowalski, A., Lai, C., Law, B., Meyers, T., Moncrieff, J., Moors, E., Munger, J., Pilegaard, K., Rannik, U., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange agric. For. Meteorol. 107, 43–69.

Finkelstein, P.L., Sims, P.F., 2001. Sampling error in eddy correlation flux measurements. J. Geophys. Res. D4 106, 3503–3509.

Finnigan, J.J., Clements, R., Malhi, Y., Leuning, R., Cleugh, H.A., 2003. A re-evaluation of long-term flux measurement techniques. Part 1. Averaging and coordinate rotation. Boundary Layer Meteorol. 107, 1–48.

Finnigan, J.J., 2004. A re-evaluation of long-term flux measurement techniques. Part 2. Coordinate systems. Boundary Layer Meteorol. 113, 1–41.

Goulden, M.L., Munger, J.W., Fan, S.M., Daube, B.C., Wofsy, S.C., 1996. Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy. Glob. Change Biol. 2, 169–182.

Hollinger, D.Y., Richardson, A.D., 2005. Uncertainty in eddy covariance measurements and its application to physiological models. Tree Physiol. 25, 873–885.

Kaimal, J.C., Finnigan, J.J., 1994. Atmospheric Boundary Layer Flows: Their Structure and Measurement. Oxford University Press.

Knorr, W., Kattge, J., 2005. Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling. Glob. Change Biol. 11, 1333–1351.

Lasslop, G., Reichstein, M., Kattge, J., Papale, D., 2008. Influences of observation errors in eddy flux data on inverse model parameter estimation. Biogeosciences 5, 1311–1324.

Lee, X., Massman, W., Law, B. (Eds.), 2004. Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis. Klewer Academic Publishers.

Lenschow, D.H., Stankov, B.B., 1986. Length scales in the convective boundary layer. J. Atmos. Sci. 43, 1198–1209.

Lenschow, D.H., Mann, J., Kristensen, L., 1994. How long is long enough when measuring fluxes and other turbulence statistics? J. Atmos. Ocean. Technol. 11, 661–673.

Lumley, J.L., Panofsky, H.A., 1964. Structure of Atmospheric Turbulence. Wiley Interscience.

Mahrt, L., 1998. Flux sampling errors for aircraft and towers. J. Atmos. Ocean. Technol. 15, 416–429.

Mann, J., Lenschow, D.H., 1994. Errors in airborne flux measurements. J. Geophys. Res. 99, 14519–14526.

Massman, W.J., 2000. A simple method for estimating frequency response corrections for eddy covariance systems. Agric. For. Meteorol. 104, 185–198.

Massman, W.J., Lee, X., 2002. Eddy covariance flux corrections and uncertainties in long-term studies of carbon and energy exchanges. Agric. For. Meteorol. 113, 121–144.

Meyers, T.P., Finkelstein, P.L., Clarke, J.F., Ellestad, T.G., Sims, P.F., 1998. A multilayer model for inferring dry deposition using standard meteorological measurements. J. Geophys. Res. 103, 22645–22661.

Moffat, A.M., Papale, D., Reichstein, M., Hollinger, D.Y., Richardson, A.D., Barr, A.G., Beckstein, C., Braswell, B.H., Churkina, G., Desai, A.R., Falge, E., Grove, J.H., Heimann, M., Hui, D., Jarvis, A.J., Kattge, J., Noormets, A., Stauch, V.J., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agric. For. Meteorol. 147, 209–232.

Moncrieff, J., Malhi, Y., Leuning, R., 1996. The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water. Glob. Change Biol. 2, 231–240.

Moore, C.J., 1986. Frequency response corrections for eddy correlation systems. Boundary Layer Meteorol. 37, 17–35.

Richardson, A.D., Hollinger, D.Y., Burba, G.G., Davis, K.J., Flanagan, L.B., Katul, G.G., Munger, J.W., Ricciuto, D.M., Stoy, P.C., Suyker, A.E., Verma, S.B., Wofsy, S.C., 2006. A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes. Agric. For. Meteorol. 136, 1–18.

Richardson, A.D., Mahecha, M.D., Falge, E., Kattge, J., Moffat, A.M., Papale, D., Reichstein, M., Stauch, V.J., Braswell, B.H., Churkina, G., Kruijt, B., Hollinger, D.Y., 2008. Statistical properties of random CO_2 flux measurement uncertainty inferred from model residuals. Agric. For. Meteorol. 148, 38–50.

Shurpali, N.J., Verma, S.B., Clement, R.J., Billesbach, D.P., 1993. Seasonal distribution of methane flux in a Minnesota Peatland measured by eddy correlation. J. Geophys. Res. 98, 20649–20655.

Webb, E.K., Pearman, G.I., Leuning, R., 1980. Correction of flux measurements for density effects due to heat and water vapor transfer. Q. J. R. Meteorol. Soc. 106, 85–100.

Wyngaard, J.C., 1973. On surface layer turbulence. In: Haugen, D.A. (Ed.), Workshop on Micrometeorology. American Meteorological Society.

Young, H.D., 1962. Statistical Treatment of Experimental Data. McGraw Hill.