

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

**A multi-site analysis of random error
in tower-based measurements of carbon and energy fluxes**

Andrew D. Richardson^{1,*}, David Y. Hollinger², George G. Burba³, Kenneth J. Davis⁴,
Lawrence B. Flanagan⁵, Gabriel G. Katul⁶, J. William Munger⁷, Daniel M. Ricciuto⁸,
Paul C. Stoy⁶, Andrew E. Suyker⁸, Shashi B. Verma⁸, and Steven C. Wofsy⁷

¹Complex Systems Research Center, University of New Hampshire, Morse Hall, 39 College Road,
Durham, NH 03824, USA.

²NE Research Station, USDA Forest Service, 271 Mast Road, Durham, NH 03824, USA.

³LI-COR Biosciences, Inc., 4421 Superior Street, Lincoln, NE 68504, USA.

⁴Department of Meteorology, Penn State University, 512 Walker Building, University Park, PA 16802,
USA.

⁵Department of Biological Sciences, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta,
Canada T1K 3M4.

⁶Nicholas School of the Environment and Earth Sciences, Duke University, Box 90328, Durham, NC
27708, USA.

⁷Division of Engineering and Applied Science/Department of Earth and Planetary Science, Harvard
University, Cambridge, MA 02138, USA.

⁸School of Natural Resources, University of Nebraska-Lincoln, P.O. Box 830728, Lincoln, NE 68583,
USA.

*Corresponding author. Tel 603 868 7654, e-mail andrew.richardson@unh.edu. Mailing address:
USDA Forest Service, 271 Mast Road, Durham, NH 03824 USA.

1 **Abstract**

2 Measured surface-atmosphere fluxes of energy (sensible heat, H , and latent heat,
3 LE) and CO₂ (FCO_2) represent the “true” flux plus or minus potential random and
4 systematic measurement errors. Here we use data from seven sites in the AmeriFlux
5 network, including five forested sites (two of which include “tall tower” instrumentation),
6 one grassland site, and one agricultural site, to conduct a cross-site analysis of random
7 flux error. Quantification of this uncertainty is a prerequisite to model-data synthesis
8 (data assimilation) and for defining confidence intervals on annual sums of net ecosystem
9 exchange or making statistically valid comparisons between measurements and model
10 predictions.

11 We differenced paired observations (separated by exactly 24 h, under similar
12 environmental conditions) to infer the characteristics of the random error in measured
13 fluxes. Random flux errors more closely follow a double-exponential (Laplace), rather
14 than a normal (Gaussian), distribution, and increase as a linear function of the magnitude
15 of the flux for all three scalar fluxes. Across sites, variation in the random error follows
16 consistent and robust patterns in relation to environmental variables. For example,
17 seasonal differences in the random error for H are small, in contrast to both LE and
18 FCO_2 , for which the random errors are roughly three-fold larger at the peak of the
19 growing season compared to the dormant season. Random errors also generally scale
20 with R_n (H and LE) and $PPFD$ (FCO_2). For FCO_2 (but not H or LE), the random error
21 decreases with increasing wind speed. Data from two sites suggest that FCO_2 random
22 error may be slightly smaller when a closed-path, rather than open-path, gas analyzer is
23 used.

1

2 **Keywords:** AmeriFlux, carbon, data assimilation, eddy covariance, flux, measurement
3 error, random error, uncertainty.

4

5 **1. Introduction**

6 Measurements of surface-atmosphere fluxes of carbon and energy at eddy
7 covariance sites around the world have provided important insight into how different
8 ecosystems function in relation to abiotic environmental forcings (Baldocchi et al., 2001).
9 However, there is a growing recognition within the eddy flux community that more
10 attention needs to be placed on quantifying the uncertainties inherent in these
11 measurements (Hollinger and Richardson, 2005). For example, in the context of
12 model–data fusion, Raupach et al. (2005) argue that “data uncertainties are as important
13 as data values themselves” because the specification of data uncertainties will affect not
14 only the uncertainty of the model, but also the model predictions themselves. Thus, since
15 eddy covariance data are increasingly being assimilated with terrestrial ecosystem models
16 (e.g. Braswell et al. 2005, Knorr and Kattge 2005, Raupach et al. 2005), a systematic
17 characterization of flux data uncertainties is needed.

18 The key issue is that although we want to know the actual flux, F , we really
19 measure $x = F + \delta + \varepsilon$, where δ is a random variable (random measurement error) whose
20 characteristics are generally unknown and ε is any systematic error. The random error is
21 therefore distinct from potential systematic errors due to incomplete spectral response,
22 lack of nocturnal mixing (u_*) or other factors. Here we focus on the random error, but
23 note that a complete description of total flux measurement error also requires a

1 quantification of the systematic error or bias (Goulden et al. 1996; Moncrieff et al.,
2 1996). This latter task seems to be especially difficult because if we knew the bias we
3 probably could correct it. Systematic errors, which cannot be evaluated with the approach
4 we use here, are discussed elsewhere (e.g., Baldocchi, 2003).

5 While previous authors have attempted to put bounds on the uncertainty
6 associated with annual sums of fluxes (e.g., net ecosystem exchange: Goulden et al.,
7 1996; Lee et al., 1999; Baldocchi et al., 2001; Griffis et al., 2003; Morgenstern et al.,
8 2004), surprisingly little is known about the measurement errors associated with the
9 turbulent fluxes computed for a single integration period. However, it is necessary to
10 know the characteristics of the random error, δ , to properly conduct a number of
11 advanced parameterization schemes for model fitting (e.g., maximum likelihood, van
12 Wijk and Bouten, 2002; data assimilation or model-data fusion, Knorr and Kattge, 2005;
13 Raupach et al., 2005; Williams et al., 2005; Gove and Hollinger, in press), to make
14 statistical comparisons between measurements and modeled predictions (validation), as
15 well as to accurately estimate confidence intervals on annual flux sums. Knowledge of
16 the random error has also been used in conjunction with Monte Carlo techniques to assess
17 the probability distribution of parameter estimates for models fit to eddy covariance data
18 (e.g. Hollinger and Richardson, 2005; Richardson and Hollinger, 2005).

19 The total flux measurement error is a composite of all error sources, including
20 errors associated with the measuring equipment, source (footprint) heterogeneity, and the
21 turbulent nature of the transport process (Moncrieff et al., 1996). If individual sources of
22 random error are quantified, they can be summed as the root mean square (Taylor, 1997).
23 Wesely and Hart (1985) derived an expression for flux uncertainty that considered

1 instrument noise as well as sampling error associated with turbulence. However, in the
2 meteorological literature, most researchers concentrate on the uncertainty of the turbulent
3 covariance, ignoring especially the additional contribution of flux source region
4 heterogeneity (Katul et al. 1999).

5 Important characteristics of δ include not only an estimate of the expected
6 magnitude of the random error (the standard deviation of the distribution, $\sigma(\delta)$, is
7 convenient in this regard), but also the higher order moments, such as skewness (is the
8 distribution symmetric) and kurtosis (how peaked is the distribution). It is also critical to
9 know how the random error covaries with environmental and ecosystem parameters.
10 Ideally, we would like to identify a probability density function (PDF) that characterizes
11 the random error. Commonly this distribution is assumed to be normal (Gaussian), but
12 there are many other PDFs (log normal, logistic, double-exponential, uniform, etc.), and
13 one or more of these may be a better fit than the normal distribution. In fact, it is
14 noteworthy that an analysis of the statistical properties of turbulence in the boundary
15 layer suggest that heat and momentum fluxes may be Gaussian for near-neutral
16 conditions but are non-Gaussian as the atmosphere becomes unstable (Chu et al., 1996).
17 Since many statistical analyses rely on the assumption of normality, it is essential to
18 know whether this assumption is met. This is because if δ does not follow a normal
19 distribution, and if the variance of δ is not constant across all observations y_i , then
20 optimization based on ordinary least squares minimization will not yield maximum
21 likelihood parameter estimates (Press et al., 1993).

22 Our objective in the present paper is to use data from sites within the AmeriFlux
23 network (38 measurement years, across a range of ecosystems: deciduous, coniferous,

1 and mixed forests, an agricultural site, and a grassland) to conduct a cross-site analysis of
2 the random errors associated with measured turbulent fluxes of energy (H and LE) and
3 CO₂ (FCO_2). We focus first on quantifying the magnitude and PDF of the random error at
4 these sites, and secondly on determining whether the characteristics of the random error
5 vary across sites or in relation to environmental parameters. Although other studies of
6 flux measurement error (Lenschow et al., 1994; Mann and Lenschow, 1994) have tended
7 to focus on the relative error, $\sigma(\delta)/|F|$, our emphasis here is on the expected magnitude of
8 the error, i.e., its standard deviation, $\sigma(\delta)$. There are two reasons for this. First, the
9 relative error is not a very useful (or well-defined) quantity as $F \rightarrow 0$, a situation that
10 occurs at least twice a day (for H , LE , and FCO_2) during the growing season, as well as
11 during the winter months at many sites (for LE and FCO_2). Second, maximum likelihood
12 estimation and various data assimilation techniques require knowledge of $\sigma(\delta)$, as it is an
13 essential component of the cost function to be minimized in the optimization. Raupach et
14 al. (2005) suggest that $1/\sigma(\delta)$ is a measure of confidence in the data, because data with
15 low $\sigma(\delta)$ are likely (but not necessarily, since δ is a random variable) to have smaller
16 errors than data with high $\sigma(\delta)$. In the maximum likelihood paradigm, the mismatch
17 between measured and modeled values is weighted by the estimated $[1/\sigma(\delta)]^n$ (where $n =$
18 2 in the case of weighted least squares), such that observations with high confidence
19 receive more weight than those with low confidence (Press et al., 1993).

20

1 **2. Method and Data**

2 *2.1 Uncertainty in turbulence*

3 Many authors have considered the random error in flux measurements, even if
 4 results of such analyses are not routinely reported. Finkelstein and Sims (2001) provide a
 5 recent and comprehensive review. They also improve on previous methods by showing
 6 how a numerical integration of raw (high frequency) data can correctly incorporate
 7 necessary lag and cross-correlation terms. To provide a conceptual framework, however,
 8 we return to the estimate for the relative error in an aircraft flux measurement developed
 9 by Lenschow et al. (1994) and Mann and Lenschow (1994), which is derived from the
 10 basic equations of turbulence:

11
$$\frac{\sigma_F(l)}{|F|} = \left(\frac{2\tau_l}{l} \right)^{0.5} \left(\frac{1+r_{wc}^2}{r_{wc}^2} \right)^{0.5} \left(1 - \frac{a}{z_i} \right) \quad [1]$$

12 Here, $\sigma_F(l)$ is the standard deviation of the random flux measurement error for a flight leg
 13 of length l , r_{wc} is the correlation coefficient between the vertical wind velocity w , and
 14 scalar c , τ_l is the integral lengthscale for c , a is the flight altitude of the aircraft, and z_i is
 15 the height of the boundary layer. For the surface (tower) approximation, we replace l with
 16 the averaging time, T , τ_l with τ_t , the integral timescale (integral of the auto-correlation
 17 function), take $\frac{a}{z_i} = 0$ (because for measurements near the surface, $a \ll z_i$), and note that

18 $r_{wc} = \frac{\overline{w'c'}}{\sigma_w \sigma_c}$ to give Eq. [2],

$$\frac{\sigma_F(t)}{|F|} = \left(\frac{2\tau_t}{T}\right)^{0.5} \left(\frac{1 + \left(\overline{w'c'}/\sigma_w\sigma_c\right)^2}{\left(\overline{w'c'}/\sigma_w\sigma_c\right)^2} \right)^{0.5} \quad [2]$$

2 (a) (b)

3 with primed quantities denoting departures from time averages (indicated by overbar).

4 This estimate is instructive because it separates out errors in the variance of the
 5 covariance (term *b*) from errors associated with the organization of turbulence into large
 6 eddies and a finite integration period (term *a*). The integral timescale, τ_t , is a measure of
 7 how long turbulence remains correlated with itself and signifies the scale of most flux
 8 transporting eddies, corresponding to the peak of the power spectral density in vertical
 9 velocity (Finnigan, 2000).

10

11 *2.2 Length scales and mechanism of large eddy production*

12 Much of the pioneering work done in describing atmospheric turbulence, the
 13 “Kansas experiments” (e.g., Businger et al., 1971; Wyngaard et al., 1971), was carried
 14 out over short wheat stubble, a situation that can be described as rough boundary layer
 15 turbulence where the boundary is the ground surface. One of the key results of this work
 16 was to provide strong experimental support for Monin-Obukhov similarity (or scaling)
 17 theory in the surface layer. A consequence of this scaling is that the integral lengthscale
 18 will be related to z/\bar{u} where z is the measurement height and \bar{u} is the mean wind speed. In
 19 the roughness sublayer (where z is within several times the canopy height h), however,
 20 turbulent statistics and large-eddy structure observed over forests appear different from
 21 those observed in the surface layer (Finnigan, 2000). Raupach et al. (1996) postulated

1 that in these situations the roughness sublayer was more akin to a mixing layer than a
 2 surface layer. Hydrodynamic instabilities in a mixing layer lead to the production of
 3 large, coherent eddies in the near-surface region (~ 2 to $5h$). This theory predicts that if
 4 shear at canopy height exceeds some threshold level, instabilities will trigger
 5 self-sustaining Kelvin-Helmholtz (KH) waves with streamwise wavelength Λ , and that
 6 $\Lambda \propto h$. Note, however, that if \bar{u} is sufficiently small, then KH instabilities will not be
 7 induced. Mixing layer theory is thus distinct from traditional surface layer theory in that
 8 the characteristic lengthscale is related to canopy height rather than measurement height.
 9 Katul et al. (1998) investigated scaling over a pine and a hardwood forest and found that
 10 the spectral peaks in vertical velocity and co-spectral peaks in scalars are well
 11 represented by the reciprocal of I_w , the vertical velocity integral time scale:

$$12 \quad \frac{1}{I_w} \approx 3.3 \frac{\bar{u}}{h}$$

13 This relationship also appears not to be sensitive to atmospheric stability.

14 Whether the scaling of large eddies is controlled by measurement height or
 15 vegetation height, the number of events passing a measurement system remains
 16 $\propto T \times \bar{u} / h_\tau$, where T is the flux averaging time period, \bar{u} is the mean horizontal velocity
 17 over T , and h_τ is the appropriate height measure for the integral timescale. Thus, if h_τ is
 18 relatively large and \bar{u} is small (perhaps conditions over moist tropical forest), the problem
 19 of adequately sampling a few large coherent eddies at a short integration period should
 20 add appreciably to the random error. When h_τ is smaller and \bar{u} greater (e.g., over crops or
 21 in the Great Plains) the many smaller turbulent structures will add little additional

1 uncertainty. An initial conclusion to be drawn from this analysis is that to reduce random
2 flux error, the integration period needs to increase as vegetation height increases because
3 of the larger size eddies, and also as wind speed drops.

4 The micrometeorological methods described previously (e.g. Eq. [2]) generally
5 require an estimate of τ_i , the integral timescale, as well as knowledge of the flux variance
6 and covariances. There are three reasons why these approaches to flux error
7 quantification are less than ideal. First, the timescale may depend upon canopy height,
8 measurement height, or some other factor depending upon wind speed, canopy
9 characteristics, and stability (Katul et al., 1998; Wesson et al., 2003; Poggi et al., 2004).
10 Second, the estimate of the random error is not independent of the flux measurement
11 itself; rather, both measurement and error estimate are based on the same variances and
12 covariances. Third, these methods do not give any insight into the distributional
13 properties (e.g., PDF) of the random error; as shown by Richardson and Hollinger (2005),
14 in a maximum-likelihood fitting paradigm, model parameters extracted from flux data
15 vary depending on the assumed PDF of δ because the optimization criterion is different.

16

17 *2.3 Repeated sampling method*

18 Finkelstein and Sims (2001) suggested that the random flux measurement error
19 could be characterized if multiple independent observations were made in one place.
20 Hollinger et al. (2004) and Hollinger and Richardson (2005) used simultaneous
21 measurements (X_1, X_2) from two towers separated by ≈ 775 m at the Howland Forest
22 AmeriFlux site to estimate the characteristics of δ . This analysis was based on the
23 assumption that the δ_i at each tower are independent and identically distributed.

1 Assume we have two simultaneous measurements of the same quantity F :

$$2 \quad x_1 = F + \delta_1 \quad [3a]$$

$$3 \quad x_2 = F + \delta_2 \quad [3b]$$

4 where δ_i is a random variable with variance $\sigma^2(\delta)$. We can quantify the random
 5 error in the measured values (x_1, x_2) by determining $\sigma(\delta)$. The variance of the difference
 6 $(x_1 - x_2)$ is given by:

$$7 \quad \sigma^2(x_1 - x_2) = \sigma^2(x_1) + \sigma^2(x_2) + 2\text{cov}(x_1, x_2) \quad [4]$$

8 Since δ_1 and δ_2 are independent and identically distributed:

$$9 \quad \sigma^2(x_1) = \sigma^2(x_2) = \sigma^2(\delta) \quad [5a]$$

$$10 \quad \text{cov}(x_1, x_2) = 0 \quad [5b]$$

11 By re-arranging [4] and substituting [5a,b], we obtain an expression (Eq. [6]) for
 12 $\sigma(\delta)$ that requires only multiple realizations (i.e., repeated over time) of the paired
 13 measurements (x_1, x_2) :

$$14 \quad \sigma(\delta) = \frac{\sigma(x_1 - x_2)}{\sqrt{2}} \quad [6]$$

15 Although results of the two-tower analysis were in reasonable agreement with
 16 predictions of the Mann and Lenschow (1994) sampling error model based on turbulence
 17 statistics (Hollinger and Richardson, 2005), there are few eddy covariance sites around
 18 the world where two appropriately distanced towers are simultaneously measuring fluxes
 19 from two independent patches of similar vegetation. Hollinger and Richardson (2005)
 20 developed an alternative method (“daily-differencing approach”) that would enable the
 21 estimation of $\sigma(\delta)$ even when researchers do not have a second tower.

1 In this approach we trade time for space, and use flux measurements made on two
2 successive days at one tower as analogues of the simultaneous two-tower paired
3 measurements described above. A measurement pair is considered valid only if both
4 measurements were made under “equivalent” environmental conditions, defined here as
5 at the same time of day (to minimize diurnal effects) and under nearly identical
6 environmental conditions (mean half-hourly *PPFD* within 75 $\mu\text{mol m}^{-2} \text{s}^{-1}$, air
7 temperature within 3 °C, and wind speed within 1 m s^{-1}). These criteria were chosen to
8 balance two conflicting requirements: 1) sufficiently similar environmental conditions
9 that the difference between the measured fluxes can be attributed to random error and not
10 differences in forcing variables; and 2) a large enough set of measurement pairs to
11 accurately characterize the PDF of the random error. We found that these rather stringent
12 requirements are frequently not met, so the sample size in one year for the daily-
13 differencing method is considerably smaller than for the two-tower method. We
14 considered including what appeared to be “equivalent conditions” at time lags longer than
15 one day, but as the lag between measurements increases, so does the risk of
16 non-stationarity in the physiological processes (e.g., seasonal trends in leaf area), which
17 will increase the estimated flux error. Although other abiotic factors, such as vapor
18 pressure deficit (VPD) or soil moisture, vary over time and also exert controls on
19 forest-atmosphere fluxes, we found that imposing additional selection criteria (e.g., VPD
20 within 0.1 kPa and soil moisture within 0.01 % volume) resulted in an $\approx 80\%$ decrease in
21 the number of measurement pairs, but only a $\approx 10\text{-}15\%$ decrease in the estimated error for
22 *FCO₂* at Duke. In heterogeneous landscapes, it may also be necessary to impose a wind
23 direction criterion, though this would likely cause a dramatic reduction in the number of

1 paired measurements with which to estimate $\sigma(\delta)$. For example, at the Harvard forest, the
2 estimated FCO_2 error was only about 10% lower (with no appreciable change in H or LE
3 error), and the data set considerably smaller, when daily-differenced measurement pairs
4 were excluded if the mean half-hourly wind directions differed by more than $\pm 15^\circ$.

5 Results from the daily-differencing approach have been shown to compare
6 favorably with random flux error estimates derived using the two-tower approach
7 (Hollinger and Richardson, 2005). For example, comparison of $\sigma(\delta)$ versus R_n
8 relationships suggested that the daily-differencing approach leads to an overestimation of
9 H random error by about 20 W m^{-2} , and an overestimation of LE random error by about
10 20–25%, compared to the two-tower approach. The $\sigma(\delta)$ versus wind speed relationship
11 suggested that FCO_2 random error is overestimated by about 20–25% compared to the
12 two-tower approach. The estimates of random flux error that we present here should
13 therefore be considered conservative “upper limits”.

14 Data for the present analysis were obtained for seven eddy covariance sites within
15 the AmeriFlux network (Table 1), representing a diverse range of ecosystems (deciduous,
16 coniferous, mixed, temperate and boreal forests; an agricultural site; and a grassland) and
17 instrument configurations (measurement heights from 3 m to 396 m, with data from both
18 closed- and open-path gas analyzers). For most sites, at least 6 or more years of data are
19 available. The Howland-Argyle tower, for which only a single year of data is available, is
20 included because it is a “tall tower” (instruments at 55 m on a cell-phone tower) site
21 located ≈ 20 km south-east of the Howland-Main tower, in central Maine. At the Duke
22 site, a switch was made from a closed-path (1998-2000) to open-path analyzer (May 1,
23 2001-2003) mid-way through the data record. The WLEF “tall tower” site has

1 instruments mounted at three different heights (30, 122 and 396 m) on a single 447 m
2 high television transmitter tower in northern Wisconsin. The Nebraska site offers a
3 comparison between two agricultural crops (soybeans, 2002, and maize, 2003), whereas
4 at the Lethbridge grassland site, the data record can be divided into a low productivity
5 drought period (1999-2001) and a more productive non-drought period (2002-2004).

6 The height of vegetation and measurement systems mean that measurements at
7 most of the sites (Duke, Harvard, Howland-Main, Nebraska, and lowest level of WLEF)
8 are in the roughness sublayer and thus subject to mixing layer scaling. However,
9 Howland-Argyle and the middle level of WLEF are transitional between mixing and
10 surface layer scaling, and Lethbridge should be considered in the surface layer ($z > 5h$).
11 The top level of WLEF (396 m) is interesting as it should be considered in the daytime
12 boundary layer where the timescale depends upon boundary layer thickness. At night the
13 upper levels of WLEF are frequently above the boundary layer (Davis et al., 2003) and
14 thus cannot evaluate the surface flux.

15 Quality control, flux corrections, and data editing were left to the site PIs, except
16 that for consistency across all sites a standard $u_* = 0.25$ threshold was applied during
17 nocturnal ($PPFD < 5 \mu\text{mol m}^{-2} \text{s}^{-1}$) periods. Site-specific procedures are described
18 elsewhere (Harvard, Barford et al., 2001; Howland: Hollinger et al., 1999, 2004; Duke,
19 Stoy et al., 2005; WLEF, Berger et al., 2001; Davis et al., 2003; Nebraska, Suyker et al.,
20 2004; Lethbridge, Flanagan et al., 2002). Note that estimates of FCO_2 random error are
21 calculated using measurements of the turbulent flux at instrument height z , rather than
22 storage-adjusted estimates of the net ecosystem exchange. Following
23 micrometeorological convention, a flux into the ecosystem is defined as negative.

1

2 *2.4 Estimation of distribution parameters for the PDF of the random error*

3 Previous work (Hollinger and Richardson, 2005) found that the probability
 4 distribution of random flux errors was better described by a double-exponential, or
 5 Laplace, distribution than a normal, or Gaussian, distribution. Unlike the Cauchy
 6 distribution, which has a superficially similar shape, the moments of a
 7 double-exponential distribution are well-defined, and the single distribution parameter
 8 (the scale parameter, β) is easily determined. A double-exponential distribution with
 9 mean zero has the following probability distribution function:

$$10 \quad f(x) = e^{-|x|/\beta} / 2\beta \quad [7]$$

11 The double-exponential distribution has a standard deviation of $\sigma = (\sqrt{2}) \beta$. An
 12 unbiased estimator for β is:

$$13 \quad \hat{\beta} = \frac{\sum_{i=1}^N |x_i - \bar{x}|}{N} \quad [8]$$

14 The double-exponential distribution is characterized by a more pronounced
 15 central peak ($|x| < 0.5\sigma$), and longer tails ($|x| > 2.3\sigma$), than a normal distribution (Fig. 1).
 16 Furthermore, whereas $\pm 1\sigma$ encompasses 68% of a normal distribution, the figure is 76%
 17 for a double exponential distribution (cf. $\pm 2\sigma = 95\%$ of a normal distribution, 94% of a
 18 double-exponential distribution).

19

20 *2.5 Analysis of random error scaling relationships*

21 We interpret our results by considering the Mann and Lenschow (1994) model
 22 described in equations [1] and [2] and the integral timescale estimate. As previously

1 discussed, either traditional or mixing layer scaling theory suggests that the integral
 2 timescale depends on the dimensionless ratio $\frac{h_\tau}{\bar{u}T}$, where h_τ is the appropriate height
 3 measure for the integral timescale, \bar{u} is the mean wind speed at the measurement height,
 4 and T is the sample period length. From equation [2], we expect the standard deviation of
 5 the random flux error, $\sigma(\delta)$, to scale as a function of the product of the absolute value of
 6 the mean flux ($|\bar{F}|$) and the square root of this dimensionless ratio:

$$7 \quad \sigma \propto |\bar{F}| \sqrt{\frac{h_\tau}{\bar{u}T}} \quad [9]$$

8 In our analysis of scaling relationships, we omit the three tall-tower data sets
 9 (Howland-Argyle, and the 122 m and 396 m instruments on the WLEF tower) for which
 10 mixing layer scaling may not be appropriate. Furthermore, assessment of the effect of h_τ
 11 and T on $\sigma(\delta)$ is difficult given the small data set, seasonal variation in vegetation height
 12 at some sites (e.g., Nebraska), and the fact that h and z co-vary with other site
 13 characteristics. Re-analysis of the raw data from a single tower may be the best way to
 14 examine the dependence of $\sigma(\delta)$ on T . We focus instead on the scaling of the random
 15 error with \bar{F} and \bar{u} . This is done first by calculating $\sigma(\delta)$ for each site across all possible
 16 $\bar{F} \times \bar{u}$ bins, and then conducting an analysis of variance on the resulting data set, with
 17 “ \bar{F} bin” and “ \bar{u} bin” as ANOVA factors. Factors are considered significant at $P \leq 0.05$.
 18

1 **3. Results**

2 *3.1 Statistical properties of the inferred random error*

3 The daily-differenced paired fluxes $((x_1-x_2)/\sqrt{2})$ indicate that the inferred random
4 flux error, δ , has, as expected, a mean value close to zero (results for Harvard Forest are
5 shown in Table 2; similar data for Howland-Main are found in Hollinger and Richardson,
6 2005). The standard deviation of the flux differences varies among H , LE , and FCO_2 , and
7 in relation to environmental factors, e.g., time of year or time of day. The distribution of
8 the flux differences is, for the most part, symmetric, because the skewness is close to
9 zero, but the distribution is more strongly peaked than a normal distribution, because the
10 kurtosis is generally ≥ 3 . Under certain conditions (e.g., $R_n > 400$ for H and LE) the
11 distribution is much less peaked (kurtosis $\approx 2-3$) than under other conditions (e.g., $R_n <$
12 100 , kurtosis = 51 for H , 24 for LE). Results from the other sites (not shown) are similar,
13 and the patterns of variation in relation to environmental factors are consistent across
14 sites. However, the standard deviation of the flux differences varies among sites,
15 especially for LE and FCO_2 (Table 3, see below).

16 At all sites, and for each of H , LE , and FCO_2 , the distribution of the flux
17 differences, and hence δ , is more closely approximated by a double-exponential, rather
18 than a normal, distribution (results for Harvard, Howland-Argyle and Lethbridge shown
19 in Fig. 2; at other sites, the shape of the distribution is similar and varies only in scale).
20 The distribution of the flux differences is strikingly similar at Harvard (Fig. 2) and
21 Howland-Main (Hollinger and Richardson, 2005): at these two forested sites, the canopy
22 height is similar and mean wind speeds are comparable.

1 A double-exponential distribution is leptokurtic in that it has a tighter central peak
2 than a normal distribution (Fig. 1). In Fig. 3, 1:1 comparison (cumulative expected vs.
3 observed) plots are shown for double-exponential and normal probability distributions,
4 using data from Harvard Forest as an example. Compared to a normal distribution, the
5 double-exponential distribution is clearly a better approximation to the observed
6 distribution of the flux differences. Within the probability range ≈ 0.05 – ≈ 0.95 (note that
7 the range is slightly wider for H , and slightly narrower for LE), the observed distribution
8 of the flux differences coincides with that of a double-exponential distribution. The
9 tendency for both distributions to diverge from the 1:1 line at both low (<0.01) and high
10 (>0.99) cumulative probabilities is indicative of the fact that the tails of both distributions
11 are much shorter than what is actually observed for the flux differences. To put this
12 another way, extreme flux outliers occur with far greater frequency than would be
13 expected under either of these two standard probability distributions. Although not
14 shown, cumulative probability plots from other tower sites were very similar, and
15 exhibited a characteristic divergence from the 1:1 line at very low and very high
16 cumulative probabilities.

17

18 *3.2 Characterizing the distribution*

19 From here onwards, we use the standard deviation of the flux differences (i.e.,
20 $\sigma(\delta)$ from Eq. [6]) to characterize the distribution of the random flux measurement error.
21 For a double-exponential distribution with scale parameter β (Eq. [7, 8]), $\sigma(\delta)$ is simply
22 calculated as $(\sqrt{2})\beta$. Estimates of $\sigma(\delta)$ for H , LE and FCO_2 are summarized in Table 3
23 for the sites included in the present study; the previously published “two-tower” estimates

1 for Howland-Main are included for comparative purposes. The overall random error in H
2 tends to be somewhat larger than the overall random error in LE , but somewhat smaller
3 than the random error in LE during the May to mid-October (JD 122-295) “growing
4 season”. The random error in FCO_2 is larger during the day than at night, and larger
5 during the growing season than the rest of the year. These patterns are quite consistent
6 across sites. The random error in H varies little among sites, whereas the random error in
7 LE is markedly lower at Lethbridge than any of the other sites. Random errors for H
8 fluxes are comparable at Harvard, Howland-Main and Duke. However, the Duke LE
9 random error (twice as large as at Howland-Main for JD 122-295) and FCO_2 random
10 error (40% larger during the day, twice as large during the night) are considerably larger
11 than at these other two forest sites. Note that the magnitude of FCO_2 at Duke is generally
12 larger than at Harvard or Howland-Main (Law et al. 2002), but LE is similar among these
13 three forested sites (Wilson et al. 2002).

14 At the WLEF tower, random errors in LE and FCO_2 (but not H) increase with
15 measurement height. The FCO_2 (but not H or LE) random error is consistently larger for
16 the maize crop than the soybean crop at the Nebraska site. Finally, during the more
17 productive years (2002-2004) at Lethbridge, LE and FCO_2 (but not H) random errors are
18 larger than during the drought years (1999-2001) at the same site.

19

20 *3.3 Random error in relation to flux magnitude and system characteristics*

21 As suggested by the above results, and Eq. [2], the random flux error scales with
22 the flux magnitude. For the forested sites (Howland-Main, Harvard, Duke and WLEF 30
23 m), the flux magnitude (“ \bar{F} bin”) accounts for 64% of the variance in FCO_2 random

1 error, with an additional 10% accounted for by wind speed (“ \bar{u} bin”); other factors
 2 (including, for example, h_r and T) excluded from the ANOVA model account for the
 3 remaining 26%. Similarly, at both the Nebraska agricultural site and the Lethbridge
 4 grassland site, both \bar{F} and \bar{u} account for a significant amount (agricultural site: \bar{F} bin =
 5 50%, \bar{u} bin = 29%; grassland: \bar{F} bin = 74%, \bar{u} bin = 16%; all $P \leq 0.001$) of variation in
 6 FCO_2 random error.

7 For H and LE , the flux magnitude generally accounts for 50-75% of the variation
 8 in $\sigma(\delta)$ ($P \leq 0.001$ for each of the forested, Nebraska, and Lethbridge sites). However, for
 9 H and LE , there is no dependence of the random error on \bar{u} at any of the sites (LE : $P =$
 10 0.38, $P = 0.86$, and $P = 0.33$ at the forested, Nebraska, and Lethbridge sites, respectively;
 11 H : $P = 0.08$, $P = 0.46$, and $P = 0.61$, in the same order).

12 The dependence of FCO_2 random error on windspeed varies somewhat according
 13 to vegetation type (Fig. 4). At the Lethbridge site, where high wind speeds (up to 16 m/s)
 14 are common, there is virtually no change in the FCO_2 random error ($\approx 0.30 \mu\text{mol m}^{-2} \text{s}^{-1}$)
 15 at wind speeds of 5 m/s or more. With this exception, the windspeed- FCO_2 random error
 16 relationship is reasonably well approximated by a curve of the form $g(\bar{u}) = \frac{a}{(\bar{u})^b}$, where a
 17 = 2.7 ± 0.2 (mean ± 1 s.e.), $b = 0.36 \pm 0.07$ (forested sites, Fig. 4A), $a = 1.0 \pm 0.1$, $b =$
 18 0.43 ± 0.05 (Lethbridge, Fig. 4B), and $a = 4.2 \pm 0.2$, $b = 0.62 \pm 0.05$ (Nebraska, Fig. 4C).

19 For both energy and CO₂ fluxes, the relationship between flux magnitude and
 20 random flux error is linear, as illustrated for the forested and grassland sites in Fig. 5. The
 21 random error does not $\rightarrow 0$ as $\bar{F} \rightarrow 0$ (as would be predicted on the basis of Eq. [1]) for
 22 any of the three fluxes. Thus there appears to be an underlying base uncertainty that is

1 present regardless of the size of the flux. One implication of this is that the relative error
 2 tends to become smaller as the magnitude of the flux becomes larger. For $\bar{F} \geq 0$, both H
 3 (Fig. 5A, B) and LE (Fig. 5C, D) random errors increase more rapidly with increases in
 4 the flux magnitude at the forested sites compared to the Lethbridge grassland (Table 4).
 5 Furthermore, at the forested sites, the FCO_2 random error increases by $0.63 \pm 0.09 \mu\text{mol}$
 6 $\text{m}^{-2} \text{s}^{-1}$ for every $1.0 \mu\text{mol m}^{-2} \text{s}^{-1}$ increase in $|\bar{F}|$ for $\bar{F} \geq 0$ (nocturnal efflux), but by only
 7 $0.19 \pm 0.02 \mu\text{mol m}^{-2} \text{s}^{-1}$ for every $1.0 \mu\text{mol m}^{-2} \text{s}^{-1}$ increase in $|\bar{F}|$ for $\bar{F} \leq 0$ (daytime
 8 uptake) (Fig. 5E). Similarly, the slope of the relationship is less steep for $\bar{F} \leq 0$ than for
 9 $\bar{F} \geq 0$ at the grassland (Fig. 5F). This is probably related to the more intermittent nature
 10 of nocturnal turbulence compared to daytime turbulence (Fitzjarrald and Moore, 1990).
 11 Under nocturnal conditions, external factors such as passage of clouds may enhance the
 12 intermittency of the fluxes (Cava et al., 2004). (Our estimated β should not be construed
 13 as a measure of intermittency: note that although the slope of flux magnitude vs. random
 14 flux error relationship is steeper at night than during the day, the mean β is higher during
 15 the day than at night because the fluxes are generally larger during the day; see Table 3).

16 From the above analysis, it would appear that differences among sites in the
 17 estimated random flux error (Table 3) can be principally attributed to cross-site
 18 differences in the mean flux magnitude, with secondary effects related to vegetation type
 19 and wind speed (and, possibly, h_τ and T).

20

21 *3.4 Seasonal patterns in the flux uncertainty*

22 Because of the way in which the random flux error generally scales with the flux
 23 magnitude, the random error varies seasonally (Fig. 6). At all sites, the random error in H

1 is relatively constant ($\approx 20 \text{ W m}^{-2}$) throughout the year, reaching a maximum (23.7 ± 2.0
2 W m^{-2}) in April and a minimum ($17.4 \pm 1.1 \text{ W m}^{-2}$) in August (Fig. 6A). By comparison,
3 *LE* random error is generally $< 5 \text{ W m}^{-2}$ during the winter months, and $> 15 \text{ W m}^{-2}$ from
4 May–September (Fig. 6A).

5 Seasonal patterns in *FCO₂* random error also mimic the seasonal course in *NEE*;
6 the random error is small in the winter months, when fluxes are negligible, and increases
7 several-fold by July (Fig. 6B), when rates of photosynthetic uptake and soil respiration
8 are both near their annual maxima. The seasonal course of *FCO₂* random error at
9 Lethbridge during the drought years (1999-2001) contrasts with the seasonal course
10 during the more productive, non-drought, years (2002-2004): from June through
11 September, the random error during the drought years is about 50% lower (Fig. 6B),
12 presumably because of drought effects on both photosynthesis and respiration during the
13 growing season.

14 The random error tends to scale, in a manner that varies seasonally, with R_n (for *H*
15 and *LE*) and *PPFD* (for *FCO₂*) (Fig. 7). The scaling relationships with R_n and *PPFD* are
16 important because they can be used to estimate $\sigma(\delta)$ independently of the actual
17 measured flux (if the actual measured flux was used, in conjunction with the scaling
18 relationships presented in Table 4, for example, then the estimated $\sigma(\delta)$ would be
19 positively correlated with the actual, but unknown, measurement error: a random error
20 causing the net flux to be under-estimated would also result in under-estimation of $\sigma(\delta)$,
21 and a random error causing the net flux to be over-estimated would result in over-
22 estimation of $\sigma(\delta)$). We compare these relationships (summarized in Table 5) for the
23 forested and grassland sites, and for JD 122-295 (“growing season”) vs. the rest of the

1 year (“dormant season”). At the forested sites, but not the grassland site, the H random
2 error scales more steeply with R_n during the dormant season (Fig. 7A, B; Table 5). The
3 opposite appears to be true for LE random error, which scales more steeply with R_n
4 during the growing season at both forested and grassland sites (Fig. 7C, D; Table 5). The
5 difference in seasonal patterns between H and LE can be attributed to seasonal changes in
6 the energy balance. At the forested sites, FCO_2 random error (across the entire $PPFD$
7 range) is about twice as large during the growing season compared to the dormant season;
8 at the grassland site, the seasonal difference is closer to four-fold (Fig. 7E, F; Table 5).
9 The slope of the $PPFD-FCO_2$ random error relationship is steeper (Table 5) at the
10 forested sites than the grassland site for two reasons: first, because at a given $PPFD$,
11 FCO_2 is larger at the forested sites than the grassland site; and second, because for a
12 given FCO_2 bin the random flux error tends to be larger at the forested sites than the
13 grassland site (Fig. 5E, F).

14

15 *3.5 Differences between closed- and open- path gas analyzers*

16 CO₂ flux measurements made with an open-path (e.g., Li-Cor LI-7500) gas
17 analyzer must be adjusted for density effects due to concurrent H and LE fluxes
18 (Webb-Pearman-Leuning [WPL] correction, see Webb et al., 1980), and these corrections
19 can, under certain conditions (when H is large and FCO_2 is small, as in late winter or
20 over sparse canopies), be larger in magnitude than the actual flux being measured. Errors
21 in H and LE will also be propagated in the process of WPL correction. It has been
22 suggested (Hollinger and Richardson, 2005), therefore, that the open-path analyzer
23 measurements of FCO_2 may be noisier or less reliable than those made using a

1 closed-path analyzer (e.g., Li-Cor LI-6262). Data from the Nebraska site, where
2 simultaneous measurements using both an open- and closed-path analyzer were made in
3 2002 and 2003, allow investigation of this issue.

4 Across all observations, for the soybean crop the FCO_2 random error is larger (by
5 $\approx 10\%$) with the open-path analyzer than the closed-path analyzer, whereas the reverse
6 (by $\approx 3\%$) is true for the maize crop (Table 6). But, regardless of crop, when the analysis
7 is limited to nocturnal periods, the random error is larger (by $>12\%$) for the open-path
8 analyzer than the closed-path analyzer (Table 6). However, these comparisons are
9 confounded to some degree by the fact that WPL-corrected open-path fluxes tend to be
10 smaller in magnitude than those measured with the closed-path system (by about 10-15%
11 for the soybean crop; by about 3-5% for the maize crop, except at night, when open-path
12 fluxes are 16% larger, see Table 6). Therefore, to account for this, we compare the
13 instruments using a measure of relative error ($R_{\sigma(\delta)} = \sigma(\delta) / \bar{F}$). These results (Table 6)
14 suggest that the relative random error is slightly lower (by $\approx 10\%$ or less) for the
15 closed-path analyzer; the difference is negligible during the day, but on the order of 15-
16 20% during the night.

17 At the Duke site, a closed-path analyzer was used for the first three years of
18 operation before being replaced with an open-path analyzer in May 2001. However, the
19 closed- versus open-path comparison is not as direct as at the Nebraska site, because
20 during this time, the height of the instruments above the canopy (and hence \bar{u} and the
21 measurement footprint) was also changed due to forest growth. Nevertheless, the FCO_2
22 random error for the open-path years (2001-2003) at Duke is about 20% higher than
23 during the closed-path years (1998-2000) (Table 3).

1 Therefore, in light of the Nebraska and Duke data, it would seem reasonable to
2 conclude that when an open-path gas analyzers is used, the random error in the measured
3 turbulent flux is probably somewhat larger than when a closed-path analyzers is used.

4

5 **4. Discussion**

6 *4.1 Implications for model fitting*

7 The analysis presented here demonstrates that the random error in tower-based
8 measurements of energy and CO₂ fluxes follows consistent patterns across sites in a
9 range of ecosystems. These robust results are in full agreement with data presented
10 previously for just the Howland-Main tower. The distribution of the random error is
11 better approximated by a double-exponential, rather than a normal, distribution. The
12 random flux error is also heteroscedastic, meaning that its variance is not constant. For H ,
13 LE , and FCO_2 , the standard deviation of the random flux error increases as a linear
14 function of the magnitude of the flux, as would be expected from theory. However, both
15 slope and intercept of these scaling relationship vary somewhat among sites, and
16 according to whether the flux is positive or negative (Fig. 5). Nevertheless, the similarity
17 of the characteristics of the random error at Harvard and Howland-Main, suggests that it
18 may be possible to identify model systems that could be used as a basis for estimating the
19 random errors at other sites that share comparable vegetation, meteorological, and
20 instrumentation characteristics.

21 Ordinary least squares fitting yields maximum likelihood parameter estimates
22 when the data meet the assumptions of normality and homoscedasticity. However, when
23 these assumptions are not met, other fitting methods should be used. Given the

1 double-exponential distribution of the random error in turbulent flux measurements,
2 maximum likelihood fitting should be based on minimizing the sum of the absolute,
3 rather than squared, deviations; since the random error is heteroscedastic, the absolute
4 deviations should further be weighted by the reciprocal of the estimated standard
5 deviation of this error (Press et al., 1993). A key difference between fitting by the least
6 squares criterion and the absolute deviation criterion is that with least squares, outliers
7 exert a much stronger influence fit, precisely because the deviations are squared. When
8 the sum of the absolute deviations is minimized, outliers, which may have no biological
9 significance, are not given undue weight. One area where the choice of fitting paradigm
10 is highly relevant is gap filling: Richardson and Hollinger (2005) report that when the
11 standard Howland gap-filling routine is implemented using the absolute deviation
12 criterion, the mean (1996–2002, ± 1 S.D.) annual *NEE* is boosted by 44 ± 9 g C m⁻² y⁻¹.
13 In percentage terms ($26 \pm 9\%$), this represents a substantial increase in the estimated
14 *NEE*.

15 Knowledge of the random errors in half-hourly flux measurements is critical for
16 evaluating the accumulated uncertainty in temporally-integrated (daily, monthly, annual)
17 fluxes. At the Howland site, Monte Carlo simulations (Richardson and Hollinger, 2005)
18 indicate that accumulated random error in measured (day + night) net CO₂ fluxes is about
19 ± 20 g C m⁻² y⁻¹ at 95% confidence. The accumulated random error due to gap filling
20 (given a particular gap filling model) adds a further ± 10 - 15 g C m⁻² y⁻¹, for a total
21 random error in (measured + filled) *NEE* of ± 25 g C m⁻² y⁻¹, or about 13% of the net
22 exchange. However, at sites with poorer data coverage, or larger *FCO*₂ random errors at
23 the half-hourly level, the annually integrated random error will be larger. Furthermore,

1 systematic errors will add additional uncertainty. At the Harvard site, Goulden et al.
2 (1996) estimated a 90% confidence interval due to sampling uncertainty of $\pm 30 \text{ g C m}^{-2}$
3 y^{-1} for annual *NEE*, compared with a total confidence interval (considering systematic
4 errors, sampling uncertainty, and under-estimation of nocturnal respiration) of -30 to $+80$
5 $\text{g C m}^{-2} \text{ y}^{-1}$.

6

7 *4.2 Interpretation of scaling relationships*

8 The fact that the magnitude of the flux is the primary factor driving the random
9 flux measurement error is in agreement with the Mann and Lenschow (1994) error model
10 based on turbulence statistics. However, whereas the Mann and Lenschow model predicts
11 that uncertainty of all fluxes should scale with $1/\sqrt{\bar{u}}$, our results indicate that this occurs
12 only for *FCO₂*. The exact cause for this discrepancy is unclear, although it may be related
13 to the location of the flux exchanging layer(s) within the ecosystem. Katul et al. (1999)
14 investigated the spatial variation in turbulence statistics from six towers over three days
15 in the Duke pine forest, in what can be considered the first direct assessment of random
16 flux measurement errors by multiple, independent measurements. Single-tower
17 measurements above the canopy were shown to represent horizontally averaged flow
18 statistics, i.e., the “canonical dynamics” of turbulent transport, and the scaling of
19 turbulent statistics was similar at each tower and thus not overly sensitive to location.
20 Relevant to the present study, Katul et al. found that the coefficient of variation (CV) of
21 $H < LE < FCO_2$, and concluded that the observed pattern resulted from the *H* exchange
22 occurring at the top of the canopy, whereas *LE* and *CO₂* are exchanged throughout the
23 canopy and are also influenced by stomatal resistance. Lai et al. (2000) demonstrated that

1 90% of the sensible heat flux occurs within the top third of the canopy, compared to 80%
2 within the top half of the canopy for latent heat. FCO_2 is an extreme case in that the
3 ground surface is frequently a CO₂ source while the canopy is a sink. Efficient mixing of
4 the entire canopy-understory-forest floor system may require particularly energetic and
5 less frequent eddies. For this multi-layered system, FCO_2 random error is therefore
6 expected to depend strongly on \bar{u} . In ecosystems with short or sparse canopies, the CO₂
7 exchange sites may be more appropriately thought of as a single layer, and FCO_2 random
8 error would, therefore, be less dependent on \bar{u} . In support of this hypothesis, the
9 $\bar{u}-FCO_2$ random error relationship is relatively flat across the entire range of wind
10 speeds at the Lethbridge grassland site where the canopy and ground layers are
11 essentially in immediate proximity.

12 Another explanation for the lack of relation between windspeed and H or LE
13 random error may be related to the distinct effect of the fluxes of these quantities on
14 atmospheric stability. Increasing fluxes of both scalars is associated with increasing
15 buoyancy, directly contributing to atmospheric mixing.

16 A practical consequence of the fact that FCO_2 random error increases
17 dramatically at low wind speeds for most sites is that windy sites are to be preferred
18 because this will lead to better sampling of the larger eddies which, over a forest, are
19 responsible for most of the turbulent transport (Raupach et al., 1996). Within the
20 roughness sublayer, random error likely decreases with z because \bar{u} increases with $z - h$.
21 Spatial (footprint) integration also is improved with increasing z . However, the results
22 from Howland-Argyle are instructive. At 55 m, the measurement height is near the top
23 (or out) of the roughness sublayer governed by mixing layer scaling, and random error is

1 increased: for example, daytime growing season FCO_2 random error at Howland-Argyle
2 has a standard deviation of $4.2 \mu\text{mol m}^{-2} \text{s}^{-1}$ ($n = 262$), compared with $3.3 \mu\text{mol m}^{-2} \text{s}^{-1}$ (n
3 $= 2924$) at Howland-Main. (Under low wind conditions, e.g., $\bar{u} \leq 2.0 \text{ m s}^{-1}$, this
4 difference is even more pronounced: $7.7 \mu\text{mol m}^{-2} \text{s}^{-1}$ ($n = 28$) at Howland-Argyle,
5 compared with $4.3 \mu\text{mol m}^{-2} \text{s}^{-1}$ ($n = 1119$) at Howland-Main). Moving out of a mixing
6 layer and into the surface layer where the length scale depends upon tower height results
7 in increasing random error. Similarly, based on the theory of Lenschow and Stankov
8 (1986), Berger et al. (2001) demonstrated that the relative error for H and FCO_2 at WLEF
9 increased with measurement height normalized by boundary layer height (z_i), up to, and
10 including, $z/z_i \approx 1$. To a certain extent, however, longer integration periods, which are
11 necessary to adequately capture the low frequency range of the cospectra, can
12 compensate for the height caused error increase (e.g., Berger et al., 2001; Malhi et al.,
13 2002).

14

15 **5. Conclusion**

16 Results from seven eddy covariance tower sites in the AmeriFlux network have
17 been used to show that the PDF of the random flux measurement error in H , LE and
18 FCO_2 is approximated by a double-exponential distribution. This distribution has a much
19 tighter central peak than a normal distribution. The standard deviation of the random
20 error is not constant, but rather scales with the magnitude of the flux, and varies in
21 relation to other environmental parameters (e.g., wind speed for FCO_2). It should be
22 possible to apply these scaling relationships to other study sites with characteristics
23 similar to those used here (i.e., agricultural crops, grasslands, and temperate/boreal

1 forests). The exact relationships are probably different in tropical forests (very tall trees
2 and generally low wind speeds) compared to the forests studied here, but it is virtually
3 certain that even in such systems the random error will scale with the magnitude of the
4 flux, and follow a Laplace distribution. We note, however, that in non-ideal flux sites
5 (where factors such as topography, footprint heterogeneity, or fetch, may be problematic)
6 the total flux uncertainty may be dominated by systematic, rather than random, errors.

7 The broader implications of these results are two-fold. First, these results provide
8 a foundation for incorporating information about random flux errors in model-data
9 synthesis problems: correct specification of a cost function requires knowledge of this
10 uncertainty. Because the random error is non-normal and heteroscedastic (non-constant
11 variance), two of the assumptions underlying least squares optimization are violated.
12 Maximum likelihood estimation techniques, which make use of information about the
13 distribution of the random error, have been developed for the double-exponential case
14 with non-constant $\sigma(\delta)$, and are therefore preferable to least squares methods.

15 Second, these results can be used to estimate confidence intervals on fluxes at
16 various time scales; in conjunction with Monte Carlo methods, for example, the estimated
17 random error in gap-filled *NEE* can be evaluated at the annual time step (but note that
18 confidence intervals need to be calculated on a site-by-site basis since both the half-
19 hourly errors, and the distribution of data gaps, vary among sites). This is a required first
20 step before defensible, statistically-based comparisons can be made either across flux
21 tower sites, or between fluxes and biometric estimates of carbon sequestration.

22

1 Acknowledgments

2 This research was supported by the Office of Science (BER), U.S. Department of Energy,
3 through the National Institute for Global Environmental Change and TCP programs, the
4 USDA Forest Service Northern Global Change Program, and the Natural Sciences and
5 Engineering Research Council of Canada (NSERC). A.D.R. was supported by the Office
6 of Science (BER), U.S. Department of Energy, through Interagency Agreement No. DE-
7 AI02-00ER63028. Flux data from the different tower sites are available at
8 <http://public.ornl.gov/ameriflux/> subject to AmeriFlux “Fair-use” rules.

9

10 References

- 11 Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P.,
12 Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B.,
13 Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K.T., Pilegaard,
14 K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K. and Wofsy, S.,
15 2001. FLUXNET: A new tool to study the temporal and spatial variability of
16 ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Am.*
17 *Met. Soc.* 82: 2415-2434.
- 18 Baldocchi, D.D., 2003. Assessing the eddy covariance technique for evaluating carbon
19 dioxide exchange rates of ecosystems: past, present and future. *Global Change*
20 *Biol.* 9: 479-492.
- 21 Barford, C.C., Wofsy, S.C., Goulden, M.L., Munger, J.W., Pyle, E.H., Urbanski, S.P.,
22 Hutyrá, L., Saleska, S.R., Fitzjarrald, D. and Moore, K., 2001. Factors controlling
23 long- and short-term sequestration of atmospheric CO₂ in a mid-latitude forest.
24 *Science* 294: 1688-1691.
- 25 Berger, B.W., Davis, K.J., Yi, C.X., Bakwin, P.S. and Zhao, C.L., 2001. Long-term
26 carbon dioxide fluxes from a very tall tower in a northern forest: Flux
27 measurement methodology. *J. Atmos. Oceanic Technol.* 18: 529-542.
- 28 Braswell, B.H., Sacks, W.J., Linder, E. and Schimel, D.S., 2005. Estimating diurnal to
29 annual ecosystem parameters by synthesis of a carbon flux model with eddy
30 covariance net ecosystem exchange observations. *Global Change Biol.* 11: 335-
31 355.

- 1 Businger, J.A., 1986. Evaluation of the accuracy with which dry deposition can be
2 measured with current micrometeorological techniques. *J. Clim. Appl. Meteorol.*
3 25: 1100-1124.
- 4 Cava, D., Giostra, U., Siqueira, M. and Katul, G., 2004. Organised motion and radiative
5 perturbations in the nocturnal canopy sublayer above an even-aged pine forest.
6 *Bound.-Layer Meteorol.* 112: 129-157.
- 7 Chu, C.R., Parlange, M.B., Katul, G.G. and Albertson, J.D., 1996. Probability density
8 functions of turbulent velocity and temperature in the atmospheric surface layer.
9 *Water Resour. Res.* 32: 1681-1688.
- 10 Davis, K.J., Bakwin, P.S., Yi, C.X., Berger, B.W., Zhao, C.L., Teclaw, R.M. and
11 Isebrands, J.G., 2003. The annual cycles of CO₂ and H₂O exchange over a
12 northern mixed forest as observed from a very tall tower. *Global Change Biol.* 9:
13 1278-1293.
- 14 Finkelstein, P.L. and Sims, P.F., 2001. Sampling error in eddy correlation flux
15 measurements. *J. Geophys. Res.—Atmos.* 106: 3503-3509.
- 16 Finnigan, J., 2000. Turbulence in plant canopies. *Annu. Rev. Fluid Mech.* 32: 519-571.
- 17 Fitzjarrald, D.R. and Moore, K.E., 1990. Mechanisms of nocturnal exchange between the
18 rain-forest and the atmosphere. *J. Geophys. Res.—Atmos.* 95: 16839-16850.
- 19 Flanagan, L.B., Wever, L.A. and Carlson, P.J., 2002. Seasonal and interannual variation
20 in carbon dioxide exchange and carbon balance in a northern temperate grassland.
21 *Global Change Biol.* 8: 599-615.
- 22 Goulden, M.L., Munger, J.W., Fan, S.-M., Daube, B.C. and Wofsy, S.C., 1996.
23 Measurements of carbon sequestration by long-term eddy covariance: methods
24 and critical evaluation of accuracy. *Global Change Biol.* 2: 169-182.
- 25 Gove, J.H. and Hollinger, D.Y., 2006. Application of a dual unscented Kalman filter for
26 simultaneous state and parameter estimation in problems of surface-atmosphere
27 exchange. *J. Geophys. Res.—Atmos.*, in press.
- 28 Griffis, T.J., Black, T.A., Morgenstern, K., Barr, A.G., Nescic, Z., Drewitt, G.B.,
29 Gaumont-Guay, D. and McCaughey, J.H., 2003. Ecophysiological controls on the
30 carbon balances of three southern boreal forests. *Agric. For. Meteorol.* 117: 53-
31 71.
- 32 Hollinger, D.Y., Aber, J., Dail, B., Davidson, E.A., Goltz, S.M., Hughes, H., Leclerc, M.,
33 Lee, J.T., Richardson, A.D., Rodrigues, C., Scott, N.A., Varier, D. and Walsh, J.,
34 2004. Spatial and temporal variability in forest-atmosphere CO₂ exchange. *Global*
35 *Change Biol.* 10: 1689-1706.

- 1 Hollinger, D.Y., Goltz, S.M., Davidson, E.A., Lee, J.T., Tu, K. and Valentine, H.T.,
2 1999. Seasonal patterns and environmental control of carbon dioxide and water
3 vapour exchange in an ecotonal boreal forest. *Global Change Biol.* 5: 891-902.
- 4 Hollinger, D.Y. and Richardson, A.D., 2005. Uncertainty in eddy covariance
5 measurements and its application to physiological models. *Tree Physiol.* 25: 873-
6 885.
- 7 Katul, G., Hsieh, C.-I., Bowling, D., Clark, K., Shurpali, N., Turnipseed, A., Albertson,
8 J., Tu, K., Hollinger, D., Evans, B., Offerle, B., Anderson, D., Ellsworth, D.,
9 Vogel, C. and Oren, R., 1999. Spatial variability of turbulent fluxes in the
10 roughness sublayer of an even-aged pine forest. *Bound.-Layer Meteorol.* 93: 1-28.
- 11 Katul, G.G., Geron, C.D., Hsieh, C.-I., Vidakovic, B. and Guenther, A.B., 1998. Active
12 turbulence and scalar transport near the forest-atmosphere interface. *J. Appl.*
13 *Meteorol.* 37: 1533-1546.
- 14 Knorr, W. and Kattge, J., 2005. Inversion of terrestrial ecosystem model parameter values
15 against eddy covariance measurements by Monte Carlo sampling. *Global Change*
16 *Biol.* 11: 1-19.
- 17 Lai, C.T., Katul, G., Oren, R., Ellsworth, D. and Schafer, K., 2000. Modeling CO₂ and
18 water vapor turbulent flux distributions within a forest canopy. *J. Geophys.*
19 *Res.—Atmos.* 105: 26333-26351.
- 20 Law, B.E., Falge, E., Gu, L., Baldocchi, D.D., Bakwin, P., Berbigier, P., Davis, K.,
21 Dolman, A.J., Falk, M., Fuentes, J.D., Goldstein, A., Granier, A., Grelle, A.,
22 Hollinger, D., Janssens, I.A., Jarvis, P., Jensen, N.O., Katul, G., Mahli, Y.,
23 Matteucci, G., Meyers, T., Monson, R., Munger, W., Oechel, W., Olson, R.,
24 Pilegaard, K., Paw, K.T., Thorgeirsson, H., Valentini, R., Verma, S., Vesala, T.,
25 Wilson, K. and Wofsy, S., 2002. Environmental controls over carbon dioxide and
26 water vapor exchange of terrestrial vegetation. *Agric. For. Meteorol.* 113: 97-120.
- 27 Lee, X., Fuentes, J.D., Staebler, R.M. and Neumann, H.H., 1999. Long-term observation
28 of the atmospheric exchange of CO₂ with a temperate deciduous forest in southern
29 Ontario, Canada. *J. Geophys. Res.—Atmos.* 104: 15975-15984.
- 30 Lenschow, D.H., Mann, J. and Kristensen, L., 1994. How long is long enough when
31 measuring fluxes and other turbulence statistics? *J. Atmos. Oceanic Technol.* 11:
32 661-673.
- 33 Lenschow, D.H. and Stankov, B.B., 1986. Length scales in the convective boundary
34 layer. *J. Atmos. Sci.* 43: 1198-1209.

- 1 Malhi, Y., Pegoraro, E., Nobre, A.D., Pereira, M.G.P., Grace, J., Culf, A.D. and Clement,
2 R., 2002. Energy and water dynamics of a central Amazonian rain forest. *J.*
3 *Geophys. Res.—Atmos.* 107: 8061, doi:10.1029/2001JD000623.
- 4 Mann, J. and Lenschow, D.H., 1994. Errors in airborne flux measurements. *J. Geophys.*
5 *Res.—Atmos.* 99: 14519-14526.
- 6 Moncrieff, J.B., Malhi, Y. and Leuning, R., 1996. The propagation of errors in long-term
7 measurements of land-atmosphere fluxes of carbon and water. *Global Change*
8 *Biol.* 2: 231-240.
- 9 Morgenstern, K., Black, T.A., Humphreys, E.R., Griffis, T.J., Drewitt, G.B., Cai, T.B.,
10 Nestic, Z., Spittlehouse, D.L. and Livingstone, N.J., 2004. Sensitivity and
11 uncertainty of the carbon balance of a Pacific Northwest Douglas-fir forest during
12 an El Niño–La Niña cycle. *Agric. For. Meteorol.* 123: 201-219.
- 13 Poggi, D., Porporato, A., Ridolfi, L., Albertson, J.D. and Katul, G.G., 2004. The effect of
14 vegetation density on canopy sub-layer turbulence. *Bound.-Layer Meteorol.* 111:
15 565-587.
- 16 Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P., 1993. Numerical
17 recipes in Fortran 77: The art of scientific computing. Cambridge UP, New York,
18 992 pp.
- 19 Raupach, M.R., Finnigan, J.J. and Brunet, Y., 1996. Coherent eddies and turbulence in
20 vegetation canopies: The mixing-layer analogy. *Boundary-Layer Meteorology* 78:
21 351-382.
- 22 Raupach, M.R., Rayner, P.J., Barrett, D.J., Defries, R.S., Heimann, M., Ojima, D.S.,
23 Quegan, S. and Schimullius, C.C., 2005. Model-data synthesis in terrestrial carbon
24 observation: methods, data requirements and data uncertainty specifications.
25 *Global Change Biol.* 11: 378-397.
- 26 Richardson, A.D. and Hollinger, D.Y., 2005. Statistical modeling of ecosystem
27 respiration using eddy covariance data: Maximum likelihood parameter
28 estimation, and Monte Carlo simulation of model and parameter uncertainty,
29 applied to three simple models. *Agric. For. Meteorol.* 131: 191-208.
- 30 Stoy, P.C., Katul, G.G., Siqueira, M.B.S., Juang, J., McCarthy, H.R., Kim, H.-S., Oishi,
31 A.C. and Oren, R., 2005. Identifying temporal differences in the net ecosystem
32 exchange of a pine and a hardwood forest—a wavelet analysis of ecosystem
33 models. *Tree Physiol.* 25: 887-902.
- 34 Suyker, A.E., Verma, S.B., Burba, G.G., Arkebauer, T.J., Walters, D.T. and Hubbard,
35 K.G., 2004. Growing season carbon dioxide exchange in irrigated and rainfed
36 maize. *Agric. For. Meteorol.* 124: 1-13.

- 1 Taylor, J.R., 1997. An introduction to error analysis. University Science Books,
2 Sausalito, CA, 327 pp.
- 3 van Wijk, M.T. and Bouten, W., 2002. Simulating daily and half-hourly fluxes of forest
4 carbon dioxide and water vapor exchange with a simple model of light and water
5 use. *Ecosystems* 5: 597-610. Webb, E.K., Pearman, G.I. and Leuning, R., 1980.
6 Correction of flux measurements for density effects due to heat and water-vapor
7 transfer. *Q. J. Roy. Meteorol. Soc.* 106: 85-100.
- 8 Wesely, M.L. and Hart, R.L., 1985. Variability of short term eddy-correlation estimates
9 of mass exchange, *The Forest-Atmosphere Interaction: Proceedings of the Forest
10 Environmental Measurements Conference, Oak Ridge, Tennessee, October 23-28,
11 1983.* D. Reidel, Norwell, MA, pp. 591-612.
- 12 Wesson, K., Katul, G.G. and Siqueira, M., 2003. Quantifying organization of
13 atmospheric turbulent eddy motion using nonlinear time series analysis. *Bound.-
14 Layer Meteorol.* 106: 507-525.
- 15 Williams, M., Schwarz, P.A., Law, B.E., Irvine, J. and Kurpius, M.R., 2005. An
16 improved analysis of forest carbon dynamics using data assimilation. *Global
17 Change Biol.* 11: 89-105.
- 18 Wilson, K.B., Baldocchi, D.D., Aubinet, M., Berbigier, P., Bernhofer, C., Dolman, H.,
19 Falge, E., Field, C., Goldstein, A., Granier, A., Grelle, A., Halldor, T., Hollinger,
20 D., Katul, G., Law, B.E., Lindroth, A., Meyers, T., Moncrieff, J., Monson, R.,
21 Oechel, W., Tenhunen, J., Valentini, R., Verma, S., Vesala, T. and Wofsy, S.,
22 2002. Energy partitioning between latent and sensible heat flux during the warm
23 season at FLUXNET sites. *Water Resour. Res.* 38: 1294-1305.
- 24 Wyngaard, J.C., Coté, O.R. and Izumi, Y., 1971. Local free convection, similarity, and
25 budgets of shear stress and heat flux. *J. Atmos. Sci.* 28: 1171-1182.
- 26

1 **Table 1.** Site information for AmeriFlux sites used in the error analysis.

2

Site Name	Lat.	Long.	Data years	Ecosystem type	Sonic height	Canopy height	\bar{u} (m s ⁻¹)
Howland-Main	45° 15'N	68° 44'W	1996-2002	Boreal transition	29 m	20 m	2.63
Howland-Argyle	45° 2'N	68° 41'W	2004	Boreal transition	55 m	15 m	4.12
Harvard	42° 32'N	72° 10'W	1991-2001	Temperate mixed	30 m	24 m	2.37
Duke	35° 59'N	79° 5'W	1998-2000	Temperate conifer	15 m	15 m	1.38
			2001-2003		20 m	18 m	1.61
WLEF	45° 57'N	90° 16'W	1997-2003	Mixed evergreen and deciduous	30 m	20 m	3.32
					122 m		6.28
					396 m		8.23
Nebraska	41° 6'N	96° 17'W	2002	Soybean field	3 m	0.8 m	3.79
			2003	Maize field	6 m	2.9 m	3.40
Lethbridge	49°43'N	112°56'W	1999-2001	Grassland, dry	6 m	18 cm ('01)	4.82
			2002-2004	Grassland, wet		34 cm ('02)	

3

4

1 **Table 2.** Statistical properties (first four moments) of the inferred random flux
 2 measurement error, δ , for Harvard Forest (1991-2001), across the entire year, during the
 3 growing season (days 122-295), and under different R_n and $PPFD$ conditions, for fluxes
 4 of energy, water and CO₂. Random errors inferred using the “daily-differencing”
 5 approach, where $\delta = (x_1 - x_2)/\sqrt{2}$, with x_1 and x_2 paired measurements separated by 24 h.
 6

Flux	# obs	Mean	Standard Deviation	Skewness	Kurtosis
H (W m ⁻²)	14,563	0.4	27.2	-0.5	36.1
JD 122–295	8,358	0.2	26.8	-1.0	40.2
$R_n > 400$	699	-8.0	46.5	0.0	3.3
$R_n < 100$	12,458	1.3	25.0	-0.6	51.5
LE (W m ⁻²)	12,053	-0.9	17.0	-0.4	29.1
JD 122–295	7,275	-1.2	20.2	-0.4	22.3
$R_n > 400$	579	-10.5	49.7	0.0	2.1
$R_n < 100$	10,261	-0.1	9.5	1.2	24.4
FCO_2 (μmol m ⁻² s ⁻¹)	13,471	0.0	2.1	0.3	15.2
JD 122–295	7,738	-0.1	2.5	0.2	10.2
$PPFD > 1000$	772	-0.3	2.9	0.0	8.4
Day	6,760	-0.1	2.2	0.5	15.0
Night	6,711	0.0	1.9	0.0	14.9

7

1 **Table 3.** Random error estimates (expressed as the standard deviation of a double-exponential distribution with scaling parameter β ,
 2 where $\alpha(\delta) = (\sqrt{2}) \beta$) for measured fluxes of energy (H , LE) and CO₂ (FCO_2) across the entire year, during the growing season (days
 3 122-295), and under different R_n and $PPFD$ conditions. Howland “2 tower” data from Hollinger and Richardson (2005), where random
 4 errors were estimated using simultaneous measurements from two flux towers separated by ≈ 775 m; all other errors estimated using
 5 the “daily-differencing” approach, as described in text.
 6
 7

 $(\sqrt{2}) \beta$ by site

Flux	Howl. 2 tower	Howl. (Main)	Howl. (Argyle)	Harv.	Duke 98-00	Duke 01-03	WLEF 30 m	WLEF 122 m	WLEF 396 m	Nebr. 2002	Nebr. 2003	Lethbr. 99-01	Lethbr 02-04
H ($W m^{-2}$)	19.5	24.6	23.3	24.1	21.0	22.5	15.7	19.9	15.4	16.1	16.0	18.6	18.4
JD 122–295	23.1	24.2	21.8	23.7	20.0	21.3	15.7	19.0	15.6	15.4	15.7	18.1	16.2
$R_n > 400$	56.9	67.8	72.2	49.1	35.2	40.6	41.0	49.3	59.5	22.7	23.6	36.1	35.3
$R_n < 100$	10.4	21.9	18.2	22.0	17.2	18.5	13.3	16.5	12.9	14.5	14.2	16.7	16.9
LE ($W m^{-2}$)	16.5	11.2	15.2	11.9	18.4	26.0	12.8	16.6	25.0	14.3	16.7	5.8	6.0
JD 122–295	26.6	15.6	17.5	14.6	27.5	31.0	17.2	25.9	33.6	15.8	17.4	7.4	8.7
$R_n > 400$	51.6	60.7	72.3	52.9	55.2	62.0	75.6	100.2	169.7	28.0	33.1	21.7	38.1
$R_n < 100$	7.0	8.3	9.8	7.8	6.7	14.4	7.3	9.7	15.6	11.1	12.9	3.9	3.4
FCO_2 ($\mu mol m^{-2} s^{-1}$)	1.5	2.0	2.4	1.8	2.9	3.7	1.7	1.6	2.3	1.9	2.4	0.4	0.6
JD 122–295	2.5	3.2	4.1	2.3	3.1	4.3	2.4	2.4	2.5	2.3	3.0	0.6	1.2
$PPFD > 1000$	2.5	3.8	3.4	2.8	4.1	4.9	2.1	3.2	5.6	1.8	2.8	0.7	1.4
Day	1.7	2.4	2.7	1.9	3.1	3.9	2.1	1.9	2.2	1.9	2.6	0.5	0.8
Night	0.9	1.5	1.9	1.7	2.3	3.1	1.1	1.1	2.5	1.8	1.1	0.3	0.2

8

1 **Table 4.** Random flux measurement error ($\sigma(\delta) = (\sqrt{2}) \beta$) scales linearly with the
 2 magnitude of the flux (\bar{F}), but the best-fit linear regressions differ for $\bar{F} \geq 0$ and $\bar{F} \leq 0$.

3	4	5	6	7	8	9
Flux	Veg. type	$\bar{F} \geq 0$		$\bar{F} \leq 0$		
5	<i>H</i>	Forested	19.7 (3.5) + 0.16 (0.01) <i>H</i>	10.0 (3.8) – 0.44 (0.07) <i>H</i>		
6		Grassland	17.3 (1.9) + 0.07 (0.01) <i>H</i>	13.3 (2.5) – 0.16 (0.04) <i>H</i>		
8	<i>LE</i>	Forested	15.3 (3.8) + 0.23 (0.02) <i>LE</i>	6.2 (1.0) – 1.42 (0.03) <i>LE</i>		
9		Grassland	8.1 (1.7) + 0.16 (0.01) <i>LE</i>	no data		
11	<i>FCO₂</i>	Forested	0.62 (0.73) + 0.63 (0.09) <i>FCO₂</i>	1.42 (0.31) – 0.19 (0.02) <i>FCO₂</i>		
12		Grassland	0.38 (0.25) + 0.30 (0.07) <i>FCO₂</i>	0.47 (0.18) – 0.12 (0.02) <i>FCO₂</i>		

13 **Note:** Standard errors for parameter estimates are in parentheses. All slope coefficients
 14 are significantly different from zero ($P \leq 0.01$).

15
 16

1 **Table 5.** The random flux measurement error ($\sigma(\delta) = (\sqrt{2}) \beta$) scales linearly with the
 2 magnitude of the driving variables (absolute value of net radiation, $|R_n|$, and $PPFD$) for
 3 H , LE and FCO_2 .

5	Flux	Veg. type	Growing season	Dormant season
6	H	Forested	16.3 (1.2) + 0.057 (0.003) $ R_n $	20.0 (1.3) + 0.077 (0.004) $ R_n $
7		Grassland	13.4 (1.4) + 0.044 (0.004) $ R_n $	17.6 (1.6) + 0.043 (0.005) $ R_n $
8	LE	Forested	17.7 (3.4) + 0.086 (0.009) $ R_n $	8.6 (3.7) + 0.039 (0.011) $ R_n $
10		Grassland	4.5 (1.2) + 0.052 (0.003) $ R_n $	3.0 (1.3) + 0.032 (0.005) $ R_n $
11	FCO_2	Forested	2.71 (0.14) + 0.75 (0.13) $\times 10^{-3} PPFD$	1.32 (0.14) + 0.87 (0.16) $\times 10^{-3} PPFD$
13		Grassland	0.77 (0.05) + 0.32 (0.09) $\times 10^{-3} PPFD$	0.18 (0.05) + 0.09 (0.05) $\times 10^{-3} PPFD$

14 **Note:** Standard errors for parameter estimates are in parentheses. All slope coefficients
 15 are significantly different from zero ($P \leq 0.01$), and all slope coefficients are significantly
 16 different between growing and dormant season ($P \leq 0.01$).

Table 6. Random flux measurement errors for FCO_2 fluxes measured using closed-path and open-path (WPL corrected for concurrent H and LE fluxes) gas analyzers at the Nebraska agricultural site, compared across the entire year, during the growing season (days 122-295), and under different $PPFD$ conditions, for both soybean and maize crops. \bar{F} is the mean measured flux, which tends to be somewhat smaller in absolute magnitude for the open-path system. The random error is compared both in terms of its absolute magnitude (standard deviation of the inferred random error, $\sigma(\delta)$) and its relative magnitude ($R_{\sigma(\delta)} = \sigma(\delta)/\bar{F}$). Estimates are calculated based only on measurement periods when data are available from both analyzers. Units for \bar{F} and $\sigma(\delta)$ are $\mu\text{mol m}^{-2} \text{s}^{-1}$; $R_{\sigma(\delta)}$ is a unitless ratio.

	n	<u>Closed-path IRGA</u>			<u>Open-path IRGA</u>		
		\bar{F}	$\sigma(\delta)$	$R_{\sigma(\delta)}$	\bar{F}	$\sigma(\delta)$	$R_{\sigma(\delta)}$
Soybeans							
All observations	599	-2.29	1.78	0.78	-1.96	1.62	0.83
JD 122-295	437	-3.27	2.06	0.63	-2.72	1.91	0.70
$PPFD > 1000$	206	-6.28	1.75	0.28	-5.27	1.49	0.28
Day ($PPFD \geq 5$)	530	-2.94	1.87	0.64	-2.54	1.66	0.65
Night ($PPFD < 5$)	69	2.70	1.13	0.42	2.48	1.27	0.51
Maize							
All observations	553	-4.62	2.29	0.49	-4.38	2.37	0.54
JD 122-295	401	-6.66	2.79	0.42	-6.28	2.64	0.42
$PPFD > 1000$	166	-14.37	2.82	0.20	13.85	2.62	0.19
Day ($PPFD \geq 5$)	470	-5.87	2.48	0.42	-5.66	2.52	0.45
Night ($PPFD < 5$)	83	2.44	1.18	0.48	2.84	1.57	0.55

1 **Figure captions**

2 **Fig. 1.** Probability distribution functions of a normal distribution and double-exponential
3 distribution compared. The x -axis is scaled in terms of the standard deviation, σ . Vertical
4 lines at $x = \pm 2.3\sigma$ and $x = \pm 0.5\sigma$ indicate the intersection points of the two distributions.
5 That is, the double-exponential distribution has a more pronounced central peak ($|x| <$
6 0.5σ) and much longer tails ($|x| > 2.3\sigma$). In panel (A), the y -axis is shown with a standard
7 linear scale; in (B), the y -axis scale is logarithmic (base 10), to better illustrate the very
8 long tails of the double-exponential distribution.

9

10 **Fig. 2.** Histograms depicting the frequency distribution of the inferred random flux
11 measurement error, δ , for energy (H and LE) and CO₂ (FCO_2) fluxes at three sites within
12 the AmeriFlux network. (A)–(C): Harvard Forest, MA; (D)–(F): Howland-Argyle, ME;
13 (G)–(I): Lethbridge, Alberta. The solid gray line depicts a normal distribution, whereas
14 the dotted black line shows a double-exponential distribution.

15

16 **Fig. 3.** 1:1 comparison (cumulative expected vs. observed) plots confirm that the
17 observed distribution of the inferred random flux measurement error, δ , is better
18 approximated by a double-exponential distribution than a normal distribution, because
19 the double-exponential distribution lies closest to the 1:1 line (the vertical lines in each
20 panel denote the range over which the double-exponential distribution coincides with the
21 observed distribution of δ). However, neither the normal nor the double-exponential
22 distribution adequately captures the very long tails of the observed distribution of δ . Data
23 are shown for the Harvard Forest, MA.

1

2 **Fig. 4.** Scaling of FCO_2 random flux measurement error with mean wind speed for three
3 vegetation types. (A) Forested sites; (B) grassland site; (C) agricultural site.

4

5 **Fig. 5.** Scaling of H , LE , and FCO_2 random flux measurement error with flux magnitude
6 for four forested sites (panels A, C, E) and one grassland site (panels B, D, F). Best-fit
7 linear regressions (fit separately for fluxes ≥ 0 and ≤ 0) are indicated.

8

9 **Fig. 6.** Seasonal course of the random flux measurement error (means across all sites;
10 error bars represent ± 1 s.e. of the mean). (A) H random error varies by less than $\approx 30\%$
11 across months, whereas LE random error increases at least four-fold between winter and
12 summer; (B) FCO_2 random error also follows a strong seasonal pattern. Lethbridge data
13 for unproductive drought years (99-01) and more productive non-drought years (02-04)
14 illustrate how the random error is larger when the flux itself is larger (i.e.,
15 June–September).

16

17 **Fig. 7.** Relationships between random flux measurement error and magnitude of driving
18 variables at four forested sites (panels A, C, E) and a grassland site (panels B, D, F). H
19 and LE random errors scale with $|R_n|$, and FCO_2 random error scales with $PPFD$. Best-fit
20 linear regressions fit separately for JD 122-295 (“growing season”) and the rest of the
21 year (“dormant season”).

Figure 1

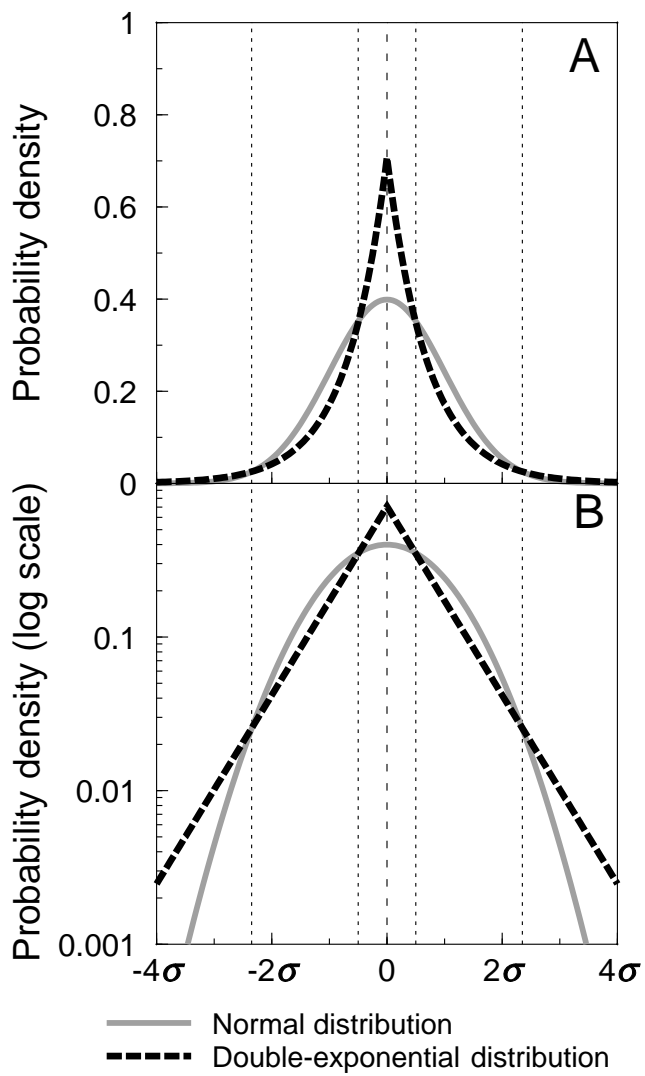


Figure 2

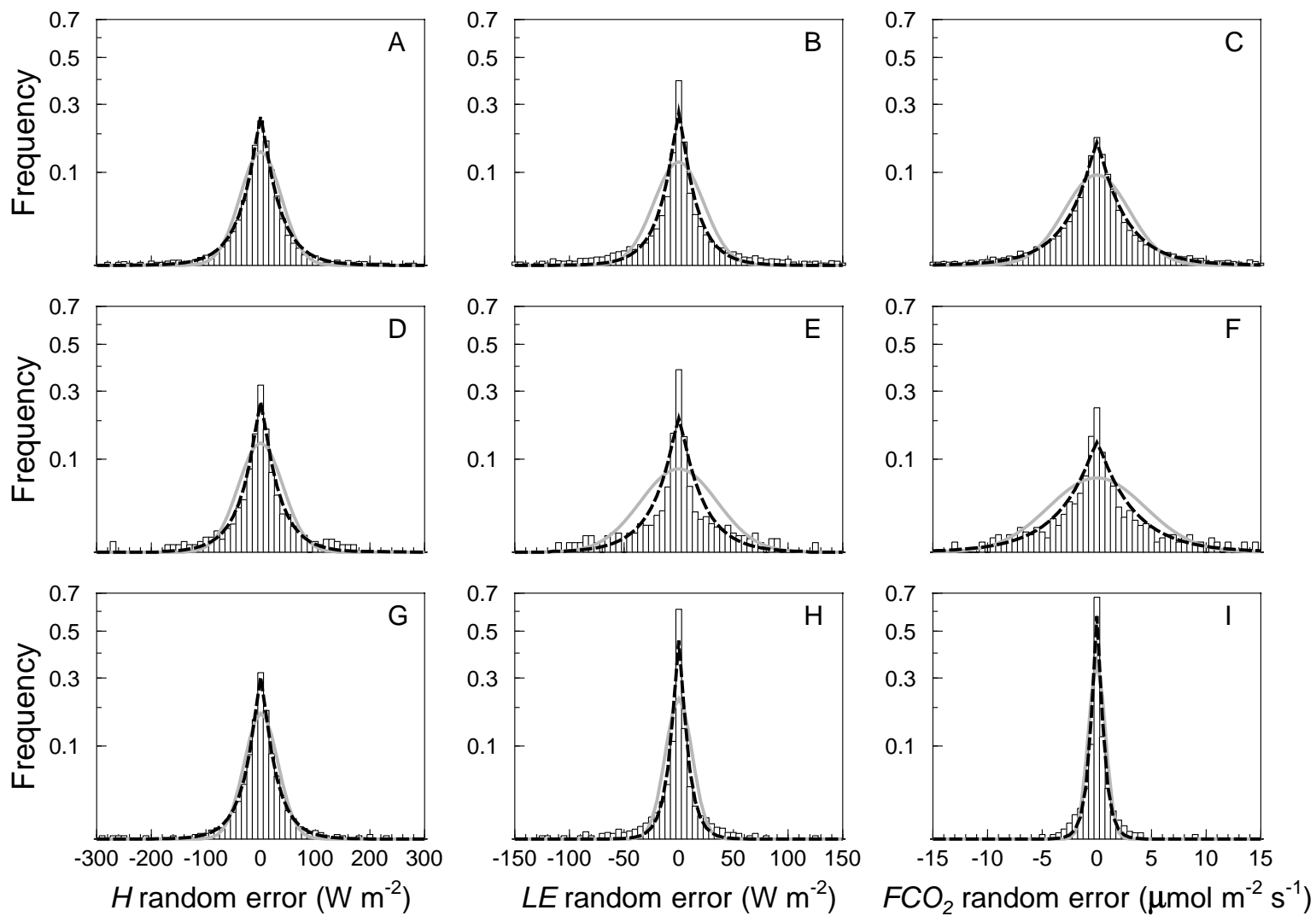


Figure 3

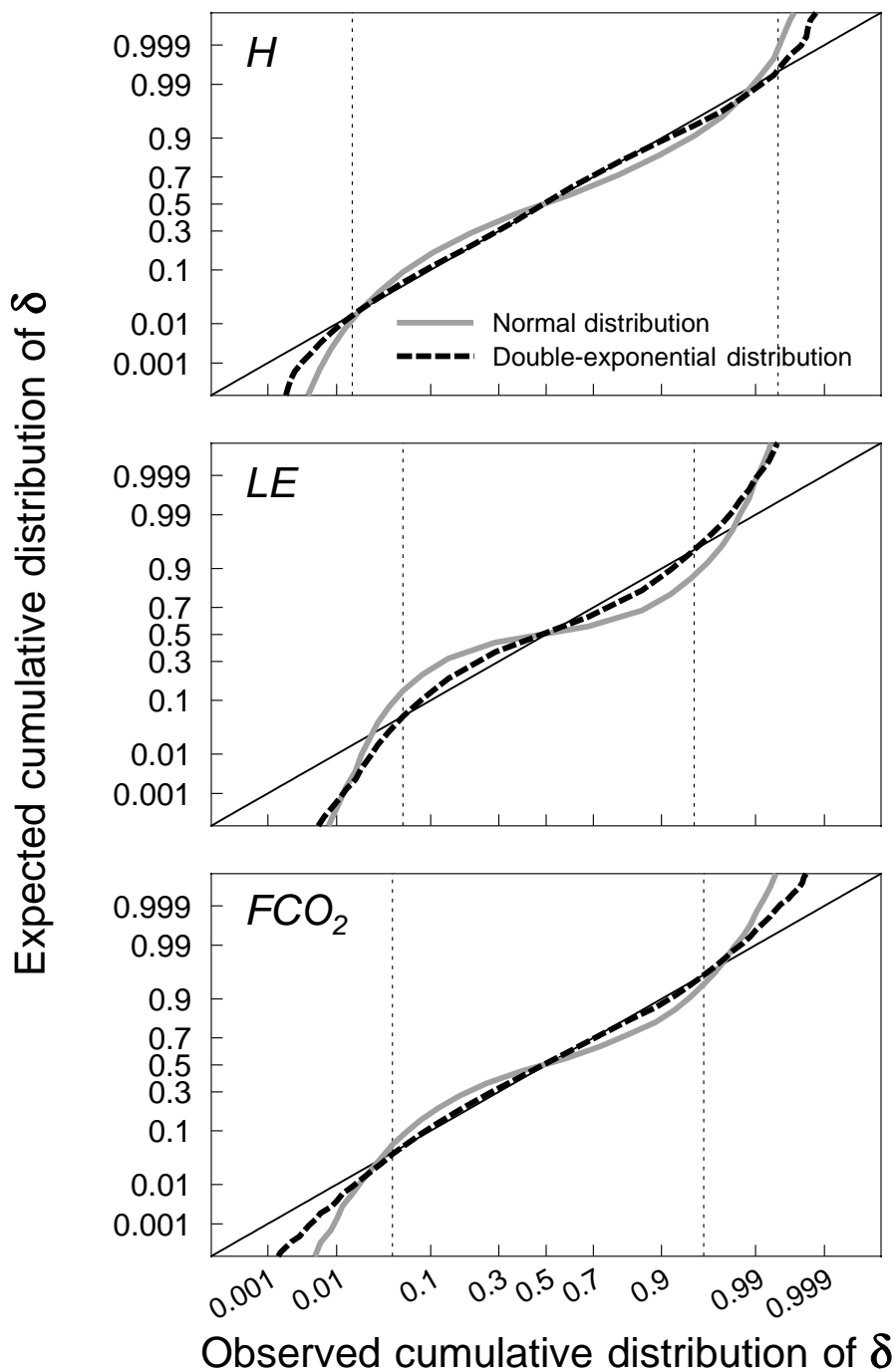


Figure 4

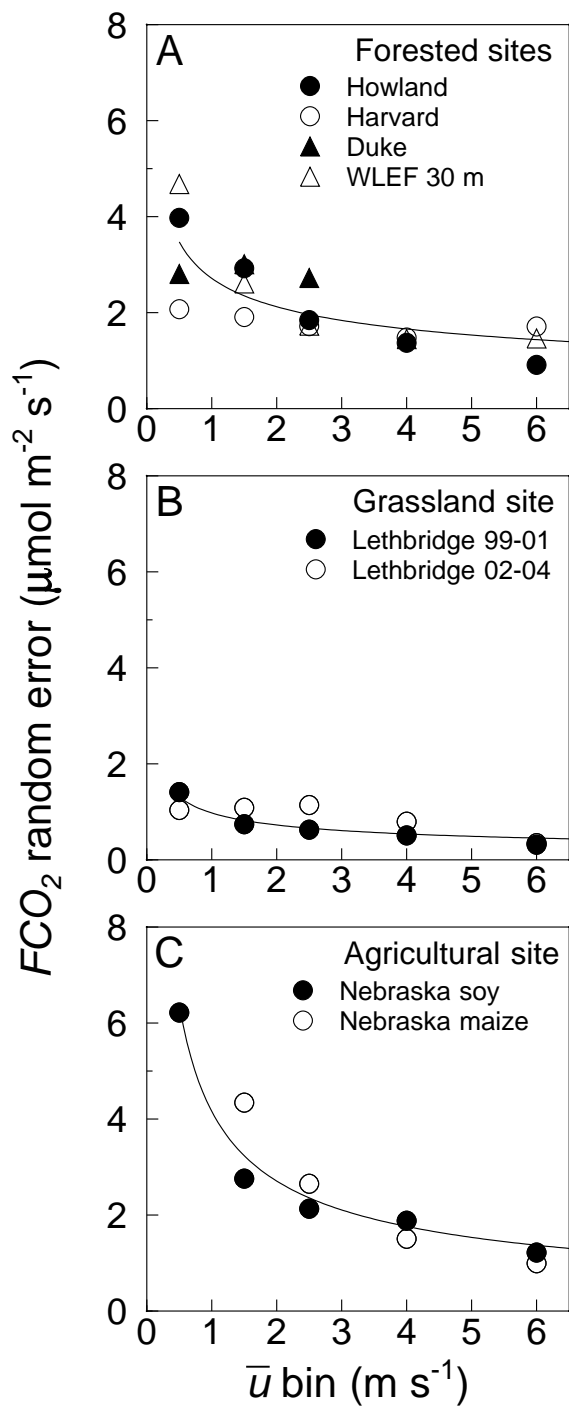


Figure 5

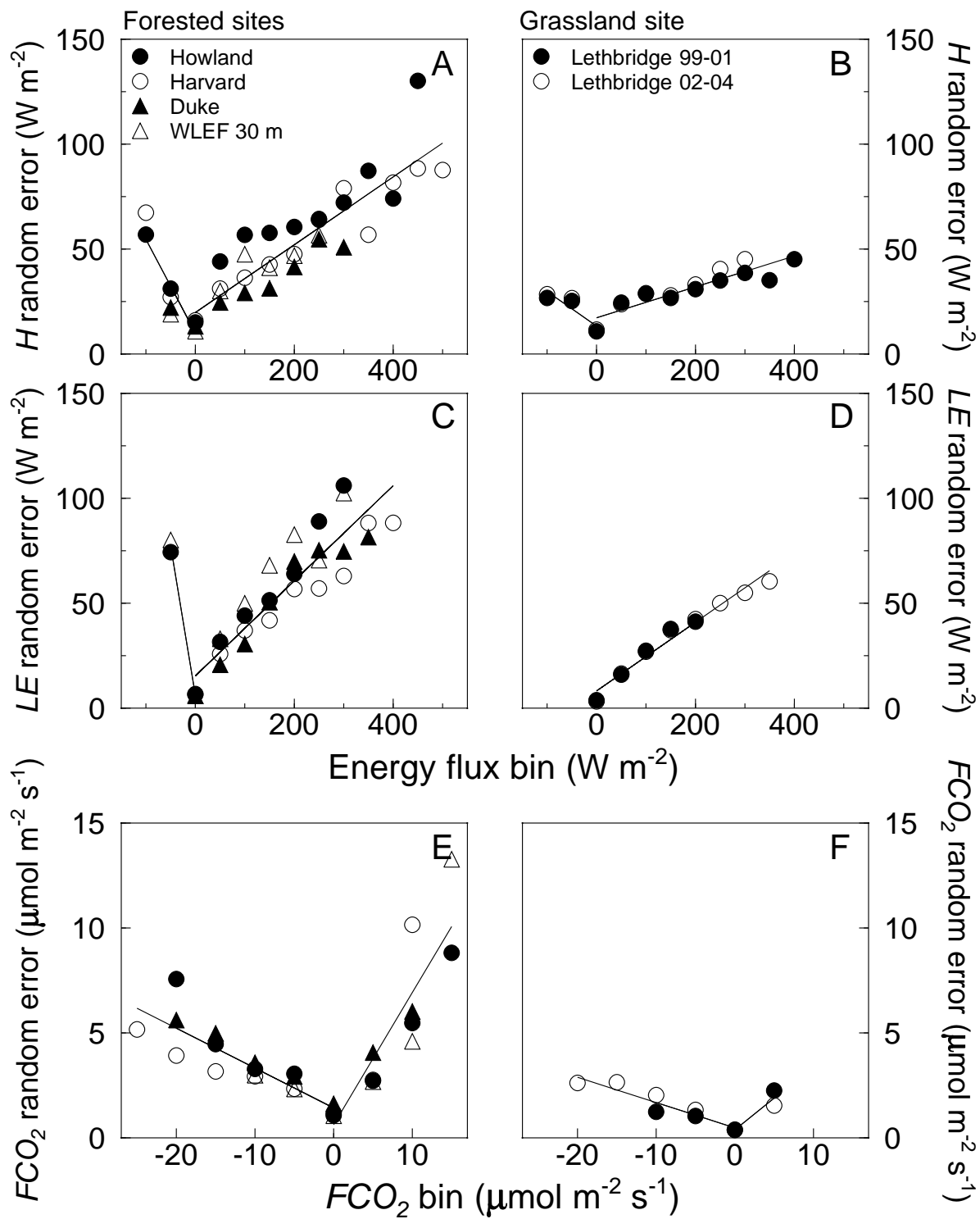


Figure 6

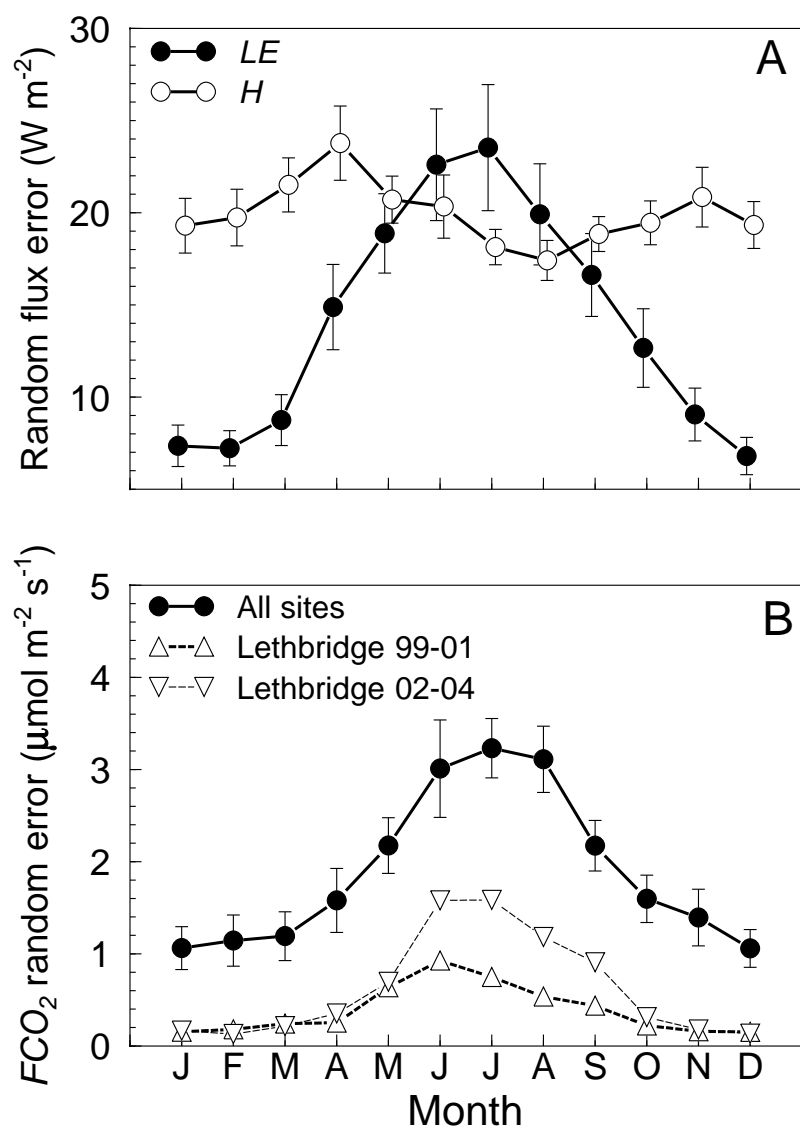


Figure 7

