Random and systematic CO$_2$ flux sampling errors for tower measurements over forests in the convective boundary layer

Dean Vickers [a,*], Christoph Thomas [b], Beverly E. Law [b]

[a] College of Oceanic and Atmospheric Sciences, COAS Admin Building 104, Oregon State University, Corvallis OR 97331-5503, USA
[b] College of Forestry, Oregon State University, Corvallis, OR, USA

ARTICLE INFO

ABSTRACT

The perturbation timescale-dependences of the CO$_2$ flux and the random flux sampling error are evaluated from eddy-covariance tower observations in the mid-day convective boundary layer over mid-latitude conifer forests. The perturbation timescale is the timescale used in the standard Reynolds decomposition to define mean and perturbations quantities. The random error due to inadequate sampling of the turbulence is estimated using two different approaches (traditional and daily-differencing). A fixed record length of 3.6 h (dyadic timescale) is used for all results, where the record length is the timescale over which the products of perturbations are averaged (flux averaging timescale). Long multiple-hour records are required to evaluate the sampling errors.

When high temporal resolution flux estimates are of interest (e.g., sub-daily timescales), including incremental contributions to the flux from transport on timescales longer than 10 min cannot be justified based on the magnitude of the incremental increase in the random sampling error. That is, the additional flux obtained by increasing the perturbation timescale beyond 10 min is dominated by random sampling error. This result is supported by both the traditional and daily-differencing approaches. For a perturbation timescale of 30 min, the relative random error (random error divided by the flux) is 38% at the taller tower and 27% at the shorter tower, and increases with increasing perturbation timescale. The cost associated with reducing the random error by using the shorter 10 min perturbation timescale, compared to the standard practice of 30 min, is an increase in the systematic flux error from 3% to 7% (averaged over three sites). Such error, while systematic, may be small in comparison to other sources of uncertainty. The choice of the perturbation timescale, and the trade-offs between reducing systematic or random errors, may depend on the intended application of the flux data. When only longer term flux estimates (e.g., monthly or annual averages) are of interest the random sampling error tends to cancel because of the larger number of samples, and the perturbation timescale can be increased to further reduce the systematic flux error.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

The standard practice in the carbon flux community for computing eddy-covariance CO_2 fluxes from a sonic anemometer and a gas analyzer mounted on a tower is to decompose the time series of vertical velocity w and CO_2 mixing ratio c into mean and perturbation parts as,

$$w = \bar{w} + w' \tag{1}$$

$$c = \bar{c} + c' \tag{2}$$

where the overbar denotes block time averaging using a constant perturbation timescale of $\tau = 30$ min. The flux is then computed as $[w'c']$, where the brackets denote time averaging over the flux averaging timescale or record length τ_F of 30 min. The averaging associated with the overbar defines the perturbation timescale τ, and the averaging associated with the brackets defines the flux averaging timescale τ_F.

Choice of the perturbation timescale τ sets an upper-limit on the range of scales included in the computed flux (e.g., Friehe et al., 1991; Oncley et al., 1996; Howell and Sun, 1999; Vickers and Mahrt, 2003). The instrument response time, the sampling rate, the instrument path-length and instrument separation distance determine the lower limit of timescales included in the computed flux. The products of perturbations are time-averaged over the flux averaging timescale τ_F because individual estimates of $w'c'$ have very large scatter. In order to evaluate the random flux sampling error, it is normally necessary to use $\tau_F > \tau$.

Random and systematic flux sampling errors are directly linked to the choice of τ and τ_F (Lenschow et al., 1994). The systematic error is due to the failure to capture all of the largest transporting scales, typically leading to an underestimation of the flux. The random error is due to an inadequate sample of the main transporting eddies as a consequence of using too short a record length. For example, say there are large eddies on the 10-min timescale that make a significant contribution to the CO_2 flux. Then if τ_F is selected as 30 min, one would have at most only three (probably two) samples of such eddies, and two samples are insufficient to place much confidence in the flux. The systematic error can often be reduced by increasing τ, however, as τ increases the number of independent samples of the flux necessarily decreases (for a fixed τ_F), which can increase the random error. The random error can be reduced by increasing τ_F, which increases the number of independent samples of the flux, however, this reduces the temporal resolution of the flux estimates and also increases the probability of including nonstationary.

The choice of τ may be influenced by the goal of the particular research. When relating fluxes to the local mean wind shear and temperature stratification, as in similarity theory, the timescale τ would ideally include transport on all turbulence timescales and exclude all larger scale motions, which do not obey similarity theory. Including this larger scale transport has been shown to degrade similarity relationships because of the large random error associated with the poorly sampled larger scale motions (Smedman, 1988; Vickers and Mahrt, 2006a). On the other hand, for balancing surface energy budgets or computing scalar budgets (e.g., annual estimates of net ecosystem exchange of carbon, NEE) one might want to include fluxes on all scales regardless of their origins and regardless of the random error, which tends to cancel when the fluxes are averaged over long enough time periods (e.g., monthly to annual timescales).

Sakai et al. (2001), Malhi et al. (2002), Finnigan et al. (2003) and Sun et al. (2006) have investigated the effect on the computed CO_2 flux of increasing τ to include longer timescale motions (systematic error). Each of these studies found examples of significant CO_2 transport on timescales greater than 30 min and less than a few hours. Flux associated with such motions could be related to deep convection (boundary-layer scale eddies), large roll vortices and possibly local circulations due to topographical or surface heterogeneity, for example motions related to differential surface heating (Mahrt et al., 2001; Lee et al., 2004). However, purely stationary eddies associated with surface heterogeneity will not contribute to the calculated flux in any obvious way.

Sun et al. (2006) reported significant contributions to the CO_2 flux over forests from long timescale motions, but not over agricultural sites. Sakai et al. (2001) found that daytime fluxes of CO_2 at approximately 30 m agl can be underestimated by 10–40% by using the standard $\tau = 30$ min, and that the systematic error increased in weak winds. They found that including longer timescale motions in the fluxes of sensible and latent heat improved their energy-balance closure at two forested sites, and attributed the effect to passage of convective boundary-layer scale eddies past their tower. Malhi et al. (2002) found that extending the coordinate rotation and averaging period from 1 to 4 h improved the surface energy-balance closure over a 30-m canopy (Manaus site) in central Amazonia. Cava et al. (2008) found significant daytime CO_2 transport on timescales between 30 min and 2 h, and attributed it to large convective eddies.

Finnigan et al. (2003) reported systematic underestimation of surface exchange when using $\tau = 15$ min compared to longer timescales at three forested sites. They also explored the links between the choice of τ and the coordinate rotation of the 3-D wind components (sonic anemometer tilt correction), and showed that rotating the coordinate system individually for each record such that the record-mean vertical motion was zero had the effect of removing most of the long timescale contribution to the flux. They concluded that over tall canopies in flat terrain in convective conditions, or at hilly sites in near-neutral conditions, the scalar cospectra (sensible heat flux, latent heat flux and CO_2 flux) have much more energy at long timescales than classical surface-layer spectral forms would predict. They also found that the importance of mesoscale motions may be quite different at different sites, and probably depends on differences in climatology, canopy structure, topography and flux measurement height above ground.

It is expected that the long timescale contribution to the flux would preferentially be missed (systematic error) by the standard practice in weak winds. As the wind speed decreases, one would need to increase τ in order to retain the same range of spatial scales of motion in the flux. This potential effect may or may not be realized depending on what scales of vertical motion are present and whether or not they are correlated with variations in CO_2.

Previous studies suggest that contributions to the vertical velocity variance from longer timescale motions (the larger eddies) may be suppressed near the surface, where the large eddies are distorted by the mean wind shear and the blocking action of the surface (Caughey and Readings, 1975; Deardorff and Willis, 1985; Stull, Chapter 11, 1990; Hunt and Carlotti, 2001). Using Doppler lidar data, Lothon et al. (2006) observed that the turbulent eddies were "squashed" as they approached the top of the convective boundary layer, thus changing the anisotropy (Kristensen et al., 1989). A similar effect occurs as the large eddies approach the ground surface. The implication is that CO_2 transport at long timescales may be suppressed closer to the surface, however, isolating this potential effect from other effects with the datasets available in this study is problematic. Multiple levels of CO_2 flux measurements above the canopy on the same tower may be required.

Any estimate of the random error must assume a stationary time series, although real atmospheric time series are never strictly stationary (Lenschow et al., 1994). One could argue that it is not possible to evaluate the true random error because the atmosphere is never purely stationary for a long enough period. Small changes in environmental conditions are always occurring, and may influence the flux in some systematic but subtle way. Presumably if we had perfect measurements of the fluxes and all the forcing variables and a perfectly homogeneous site, what we call random error might be nearly eliminated with sufficiently long records. As a result, estimates of the random flux sampling error are always subject to interpretation because the true random error cannot be estimated from real atmospheric time series.

Different approaches have been proposed for estimating random flux sampling errors from observations. Lenschow et al. (1994), following Lumley and Panofsky (1964), developed theoretical expressions for the random flux error. An important general result of their study is that while the systematic flux error can be nearly eliminated by increasing the perturbation timescale (or aircraft flight track), the random error may still be significant. That is, in practice it is much more difficult to reduce the random error to acceptable levels than to reduce the systematic error. Howell and Sun (1999) developed an approach based on the scale-dependence of the flux and the flux sampling error to determine a Reynolds turbulence cut-off timescale. Contributions to the flux from scales larger than the cut-off scale are dominated by random sampling error and are therefore unreliable. Their study employed long 7-h data records in the stable boundary layer and focused on the fluxes of heat and momentum over a grass field.

Hollinger and Richardson (2005) employed two eddy-covariance towers separated by 775 m to obtain two simultaneous estimates of the CO_2 flux. This approach is not feasible for most flux sites, and it may contaminate the estimates with systematic differences related to instrumentation (e.g., flow distortion) and spatial variability, especially if the measurements are made in the roughness sublayer. Richardson et al. (2006) proposed a methodology for estimating random errors where CO_2 flux measurements made at the same local time on two successive days at one tower (daily-differencing) are used as analogues of the two-tower method. Such measurement pairs were considered valid only when differences in air temperature, solar radiation and wind speed were within prescribed thresholds. A potential problem with this method is that the daily-difference in the flux may be influenced by changes in environmental conditions during the intervening 24-h period between the measurements (e.g., rain showers, frost). In addition, the daily-differences in environmental conditions vary within the constraints of the threshold criteria. As the threshold criteria are made more strict, the sample size decreases.

The goal of the current study is to document the scale-dependence of the observed random and systematic sampling errors for CO_2 fluxes in the convective boundary layer over forests. The errors considered here are those due to inadequate sampling of the turbulence. The scale-dependences are studied using an orthogonal decomposition technique. The scale-dependence of the random flux error is estimated using the method of Howell and Sun (1999), and with the daily-differencing method of Richardson et al. (2006) after modification to include the scale-dependence. Estimates are also given for the scale-dependence of the total relative random flux sampling error. Such information on the scale-dependence of the systematic and random flux errors can form the basis of a selection criteria for τ and τ_F depending on the particular application of the fluxes.

2. Data

2.1. MPINE site

The first dataset analyzed is from a 90-year-old mature ponderosa pine forest in central Oregon, U.S.A. (44.451 N latitude, 121.558 W longitude, 1255 m elevation) during 2004 through 2005. The canopy extends from 10 to 16 m above ground, while the understory consists of scattered 1-m tall shrubs (bitterbrush and manzanita). The canopy is open such that the forest floor and understory receive direct sunlight during the day and are mostly open to the sky at night, leading to a large diurnal cycle of the subcanopy stability (unstable during the day and stable at night). The leaf area index (LAI) ranges from 3.1 to 3.4 during the growing season. The stand density is 325 trees ha^{-1} and the soil is sandy loam. Although the site is located on a relatively flat saddle region about 600 m across, it is surrounded by complex terrain. Imagery indicates that this age class of ponderosa pine is prevalent for several kilometers in all directions.

2.2. YPINE site

The second dataset analyzed is from a 3-m tall, young ponderosa pine ecosystem with seasonal grass understory in central Oregon, U.S.A. (44.315 N latitude, 121.608 W longitude, 1005 m elevation) during 2004 and 2005. The site was clear cut and replanted by the U.S. Forest Service in 1987. The stand density is 260 trees ha^{-1} and the peak LAI was 0.61 and 0.71 in 2004 and 2005, respectively. A heterogeneous stand of older (20–100 years) and taller (5–30 m) ponderosa pines surround the site on three sides. There is a large clearing 500 m north of the tower, and an anomalously wet region (Cold

Springs) 2 km southwest of the tower. The terrain slopes upward at about 2% to the west and southwest, and slopes weakly downward (<1%) or is flat in other directions. Compared to the mature pine site, the young site is more heterogeneous in terms of vegetation type and more homogeneous in terms of terrain height.

2.3. MFIR site

The third dataset analyzed is from a 33-year-old mature Douglas-fir forest located in the coast range of western Oregon, U.S.A. (44.646N latitude, 123.551W longitude, 310 m elevation) during 2006 and 2007. The average canopy height is 26 m. The vertical structure of the vegetation canopy consists of a dense understory composed mainly of Salal in the subcanopy below 1.5 m agl and the main tree crown space extending from 15 to 26 m agl separated by a clear bole space. The canopy is very dense with a measured LAI of 9.4.

The vertical stratification of the canopy and the high LAI lead to a reversed static stability regime in the crown space and the subcanopy. During the day, the crown space and air above the canopy are commonly unstably stratified with the maximum potential temperature observed at approximately 20 m agl, and a stably stratified subcanopy and bole space. At night with weak winds, the crown space and air above the canopy are mostly stably stratified while the stratification of the subcanopy and the bole space is neutral or slightly unstable with a minimum potential temperature at 20 m agl. Higher wind speeds at night lead to isothermal conditions from the ground to the top of the tower at 40 m. The site is surrounded by complex terrain with a flat saddle located to the northeast of the tower at a distance of 600 m. The fetch for the prevailing wind direction is located in relatively flat terrain with a homogeneous age class of Douglas-fir.

3. Methods

3.1. Flux measurements

Eddy-covariance measurements of the CO_2 flux were collected using a Campbell Scientific CSAT3 sonic anemometer and an open-path LICOR-7500 gas analyzer at all three sites. Instruments were deployed at 31 m agl (or about twice the canopy height) at the MPINE site, at 12 m agl (or about four times the canopy height) at the YPINE site and at 40 m agl (or about 1.5 times the canopy height) at the Douglas-fir site. The $(z - h)$ values are 15, 9 and 14 m for MPINE, YPINE, and MFIR respectively, where z is the measurement height and h is the canopy height. Raw time series from the CSAT and LICOR-7500 were recorded at a sampling rate of 10 Hz over the pine forests and 20 Hz over the fir forest. The difference between 10 and 20 Hz sampling has no influence on the flux sampling errors studied here. As part of the quality control process (Section 3.3), spikes are removed from the fast-response time series of the winds and sonic temperature from the CSAT3, and the humidity and CO_2 from the LICOR-7500 using the iterative method described in Vickers and Mahrt (1997).

A potential source of measurement uncertainty is that at the longer timescales examined here, the vertical velocity fluctuations approach very small values, and may be too small to accurately measure, possibly because they may be of comparable magnitude to errors associated with flow distortion. Vickers and Mahrt (2006b) showed that the mean vertical motion computed from a network of towers measuring the vertical profile of the mean horizontal wind and the mass continuity equation at the YPINE site had much less scatter compared to that based on a single 3-D sonic anemometer located in the center of the network. They suggested that the sign and amplitude of the vertical motion from mass continuity appeared more plausible. Very small vertical velocity fluctuations, such as those on timescales of 30 min or more, may be difficult to accurately measure.

3.2. Long data records

To evaluate the flux at longer timescales, longer data records must be used, however, increasing the record length also increases the chance of including nonstationarity. Unfortunately, atmospheric flows are characterized by motions that simultaneously vary on a variety of scales. As a result, some motion usually appears on scales just larger than the largest transporting scales, automatically causing some degree of nonstationarity. Detrending is undesirable because it does not satisfy Reynolds averaging and it may remove real flux.

To study the convective boundary layer, a record length of $\tau_F = 3.6$ h from approximately 11 to 15 LST during May through August was chosen as a compromise to obtain longer records yet minimize nonstationarity and retain reasonable temporal resolution. The record length must be an integer power of two (dyadic timescale) for the orthogonal decomposition method (Section 3.5). The 11–15 LST period is chosen because it occurs well after the nonstationarity associated with the morning transition and well before the evening transition. Only records with strong surface heating where the average upward buoyancy flux exceeds 50 W m^{-2} are used. After all quality control checks (Section 3.3), the final dataset includes 200 records at MPINE and 201 records at YPINE during 2004 and 2005, and 114 records at the MFIR site during 2006 and 2007. For the MFIR site, the records cover the period from 12 to 16 LST because of the limited number of good quality records for 11–15 LST due to moisture on the LICOR-7500 from heavy morning dew.

3.3. Quality control

The quality control testing applied to the fast response time series collected by the sonic anemometer and the LICOR-7500 is designed to detect only instrument problems, and does not eliminate any data based on criteria associated with stationarity, footprint or agreement with similarity theory. The revised software package (the original package was described in Vickers and Mahrt (1997)) applies tests to each variable to detect: (a) too many spikes (all spikes are subsequently removed prior to further testing), (b) a value outside a defined valid range, (c) unusually large skewness or kurtosis, (d) large discontinuities detected using the Haar transform, and (e) a variance that is outside a defined valid range (a variance that is either too small or too large is flagged). The tests are applied to individual windows of width 10 min that move sequentially

through the records. Variables that are flagged by any of the tests are plotted for visual inspection and either confirmed as instrument problems (and subsequently removed from all further analysis) or identified as unusual but plausible data (and retained for further analysis). The majority of instrument problems for the daytime data at the pine and fir forests are associated with moisture on the anemometer or the LICOR.

3.4. Tilt correction

A tilt correction based on the long-term average wind direction dependence of the tilt angle is applied to the fast-response wind components (Paw U et al., 2000; Mahrt et al., 2000; Feigenwinter et al., 2004). The tilt correction (or coordinate rotation) does not force the mean vertical motion to zero for individual records. The planar-fit approach (Wilczak et al., 2001) is an idealized case of the approach used here. For perfectly planar topography and a perfectly levelled anemometer, or, for perfectly flat terrain and a vertically tilted anemometer, the tilt correction method used here reduces to the planar fit method. The same coordinate rotation method is applied for all flux calculations regardless of the value of τ. Any small long-term mean (e.g., annual) vertical motion is removed prior to all other tilt correction calculations to account for a potential zero offset in the vertical motion from the anemometer.

3.5. Flux decomposition

The method used for computing the flux and for examining the timescale-dependence of the flux is based on multi-resolution decomposition (Howell and Mahrt, 1997; Vickers and Mahrt, 2003; Acevedo et al., 2006; van den Kroonenberg and Bange, 2007). Multiresolution analysis applied to time series decomposes the record into simple unweighted averages on dyadic timescales and represents the simplest possible orthogonal decomposition. In a dyadic sequence, each increasing scale is twice as large as the previous smaller one (e.g., 2 samples, 4, 8, etc.). Unlike Fourier cospectra, multiresolution decomposition satisfies Reynolds averaging at all scales and does not assume periodicity. Fourier cospectra tend to be shifted to larger scales because "local" multi-resolution cospectra respond to event widths, whereas the "global" Fourier cospectra are influenced by the time between events (Howell and Mahrt, 1997).

The flux is calculated for perturbation timescale τ as the sum of the orthogonal multiresolution modes from the smallest resolvable scale (2 data points) up to scale τ. This summation of modes to obtain the flux (or covariance) is analogous to integration of the cospectral density to obtain the covariance using Fourier decomposition. Products of perturbations are averaged over the entire record length (3.6 h) for all values of τ tested. Thus, the highest temporal resolution of the flux resolved here is 3.6 h.

The long timescale motions have been referred to as "low-frequency" motions in the carbon flux community and in some previous literature in this area. However, we note that use of the term frequency implies periodicity, and while Fourier analysis relies on periodicity our decomposition method does not.

3.6. Random sampling error

Following Howell and Sun (1999), the scale-dependence of the random flux sampling error modes E for a single data record of length 2^M data points is given by

$$E(\tau) \simeq t(N_s(\tau), \alpha)\sigma_F(\tau)N_s^{-1/2}(\tau) \tag{3}$$

where timescale τ is equal to 2^m data points (dyadic scales), t is the Student's t-distribution, $N_s = 2^{M-m}$ is the number of independent samples of length 2^m contained within the record of length 2^M, and σ_F^2 is the observed variance of the N_s estimates of the flux. This development assumes that the N_s values of the flux are random samples from a Student's t-distribution. The expression for E is approximate because this assumption is not always met. The assumption may break down for records with significant nonstationarity where some of the variability may be associated with trend. In such case, Eq. (3) may over-estimate the true random error. The parameter α is determined such that the probability that the true record-averaged flux is contained within the interval $[F - E, F + E]$ is two times α minus one, where $F(\tau)$ is the mean flux for the record. Selecting a probability of 0.80, gives $\alpha = 0.9$ (Von Mises, 1964). For a large number of samples, the Student's t-distribution is close to the normal (Gaussian) distribution. The random error for timescale τ is obtained by summing the error modes E from the smallest resolved scale up to scale τ.

Applying the daily-differencing method (Section 1) to each individual flux mode rather than the flux, gives the scale-dependence of the random error modes as

$$R(\tau) \simeq (\sqrt{2})\beta(\tau) \tag{4}$$

$$\beta(\tau) = N^{-1}\sum_{i=1}^{N}|x_i(\tau) - \bar{x}(\tau)| \tag{5}$$

where R is the random error mode, β is the single parameter of a double exponential distribution with standard deviation R (Richardson et al., 2006), N is the number of observed daily-differences, x_i are the CO_2 flux mode daily-differences, and $\bar{x}$ is the mean of all the x_i. For a double exponential distribution, the probability that the true flux is inside the interval $[F - R, F + R]$ is 0.76, which is approximately the same confidence level applied in the formulation of E. The expression for R is of unknown accuracy because an unknown fraction of the daily-difference may be associated with differences in forcing variables, and not random variability. Day-to-day changes in environmental conditions in the daily-differencing method are analogous to nonstationarity across the record in the traditional approach in that both may result in over-estimation of the random error. Note that while E in Eq. (3) is practically guaranteed to increase with increasing τ because N_s decreases with τ, the random error mode estimate R in Eq. (4) has no such term. There are an equal number of samples N of each mode in the daily-differencing method. The random error modes R can be summed to obtain the random error.

Two sets of criteria are initially chosen to select the days used for daily-differencing. The first set requires the daily-differences in air temperature, PAR (photosynthetically active radiation), wind speed and soil moisture content to be less

than 2 °C, 100 μmol m^{-2} s^{-1}, 2 m s^{-1} and 0.02 m^3 m^{-3}, respectively. The second set of criteria relax the temperature and radiation difference thresholds to 3 ° C and 150 μmol m^{-2} s^{-1}. For both sets we require mostly clear skies (PAR greater than 1000 μmol m^{-2} s^{-1}) during May through August from 11 to 15 LST with strong surface heating. With these criteria, we find 32 (47) daily-differences at MPINE and 34 (49) daily-differences at YPINE for the first (second) set of criteria.

The scale-dependence of the relative random error (RRE, random error divided by the mean flux) is calculated as follows using the error mode estimates E. First, we note that the magnitude of the random error is linearly related to the magnitude of the mean flux, as observed by Richardson et al. (2006) and others, and as predicted by Mann and Lenschow (1994). This implies that the random error divided by the mean flux is approximately constant. The RRE is estimated for a given timescale τ by regressing the individual estimates of the random error (sum of the random error modes up to timescale τ) on the individual estimates of the flux (sum of the flux modes up to timescale τ). The slope from the regression is an estimate of RRE for that particular timescale. The linear relationship is found to apply to all modes, although the slope varies with timescale τ. This method is preferred to averaging the ratio of the random error to the flux because such an estimate is not well behaved as the mean flux approaches zero.

4. Results

The composite perturbation timescale-dependence of the CO_2 modes, vertical velocity modes and the CO_2 flux modes for the convective boundary layer are shown in Figs. 1–3 . The compositing is done over all records (Sections 2 and 3). The modes can be interpreted as the average change in the variance (or flux) associated with including variability (or transport) on the next larger dyadic timescale. Recall that the variance (or flux) for timescale τ is obtained by summing the modes from the smallest resolved timescale up to τ.

The peaks in the vertical velocity modes occur at $\tau = 13$, 6 and 26 s for MPINE, YPINE and MFIR, respectively (Fig. 4). These differences are consistent with the relative differences in measurement height z at the three sites, and with differences in $(z - h)$. The timescale associated with the peak in the vertical velocity modes increases as the measurement height (or $z - h$) increases because the larger eddies are preferentially suppressed by the presence of the ground (Section 1). The peaks in the CO_2 flux modes occur at $\tau = 51$, 26 and 51 s for MPINE, YPINE and MFIR, respectively (Fig. 4). These timescales are longer than those associated with the peak in the vertical velocity modes in part because the CO_2 modes increase with increasing timescale (Figs. 1–3). The timescale corresponding to the peak CO_2 flux mode is smallest at the YPINE site, which also has the lowest measurement level and smallest $(z - h)$ value.

4.1. Systematic error

The composite CO_2 flux modes are small at timescales longer than 30 min, suggesting that little additional flux is at
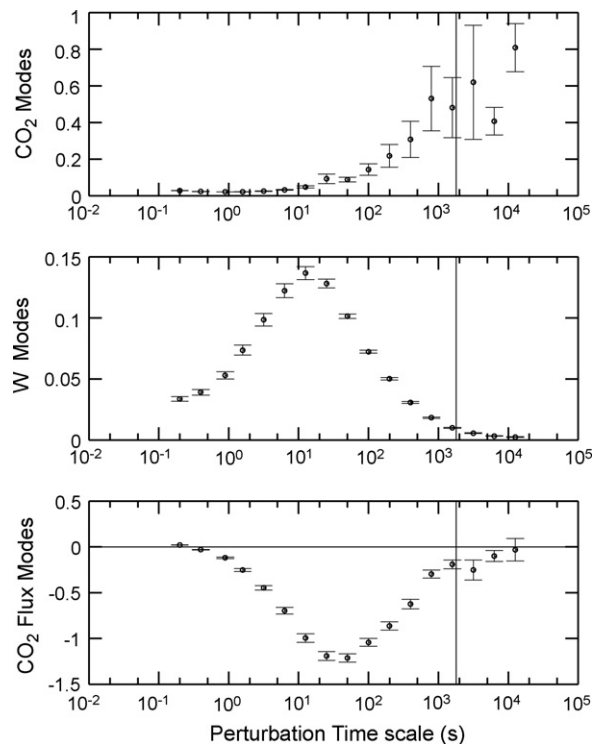

Fig. 1 – Composite timescale-dependence of the CO_2 modes (ppm^2), vertical velocity modes (m^2 s^{-2}) and the CO_2 flux modes (μ mol CO_2 m^{-2} s^{-1}) for the MPINE site for 200 4-h daytime (11–15 LST) convective boundary-layer records. Error bars show plus and minus one standard error due to variations between records. A negative flux indicates downward transport of CO_2 (net uptake by the ecosystem). The vertical line denotes $\tau = 30$ min.

scales between 30 min and a few hours after averaging over many records. This implies that the large variability in CO_2 on timescales greater than 30 min (Figs. 1–3) is generally not well correlated with vertical motion, and that such variability might be more related to horizontal transport of CO_2 than vertical transport of CO_2. However, for individual records, there can be large flux at scales longer than 30 min, and it can be of either sign (Fig. 5). We will show in the next section that the contribution to an individual 3.6-h average flux estimate from motions on timescales greater than about 10 min is primarily random.

There is a weak inverse relationship between wind speed and the systematic error, defined for these purposes as the magnitude of the $\tau = 3.6$-h flux minus the $\tau = 27.3$-min flux. That is, the systematic error here is defined as the additional flux captured by including transport on timescales between approximately 30 min and 4 h. At the MPINE site, the systematic error is a minimum for the weakest wind speeds (<2 m s^{-1}), is a maximum for intermediate winds (2–4 m s^{-1}), and has intermediate values for the strongest winds (4–8 m s^{-1}). At the YPINE site, the systematic error is a maximum for the weakest wind speeds (<1 m s^{-1}), and is nearly invariant for wind speeds greater than 2 m s^{-1}. Excluding

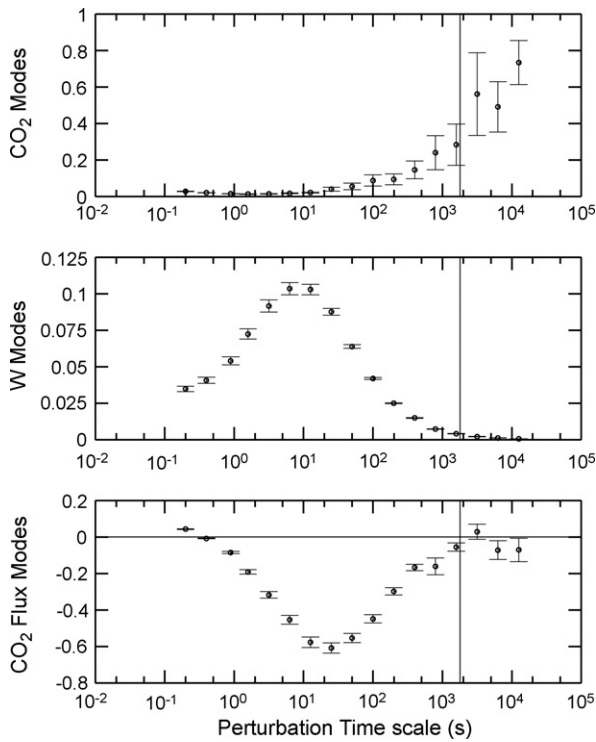Fig. 2 – Composite timescale-dependence of the CO_2 modes (ppm^2), vertical velocity modes (m^2 s^{-2}) and the CO_2 flux modes (μ mol CO_2 m^{-2} s^{-1}) for the YPINE site for 201 4-h daytime (11–15 LST) convective boundary-layer records.
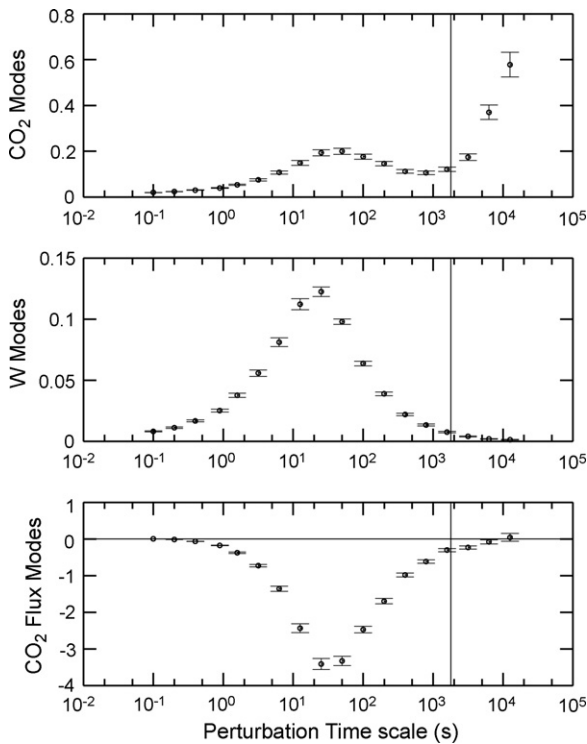
Fig. 3 – Composite timescale-dependence of the CO_2 modes (ppm^2), vertical velocity modes (m^2 s^{-2}) and the CO_2 flux modes (μ mol CO_2 m^{-2} s^{-1}) for the MFIR site for 114 4-h daytime (12–16 LST) convective boundary-layer records.
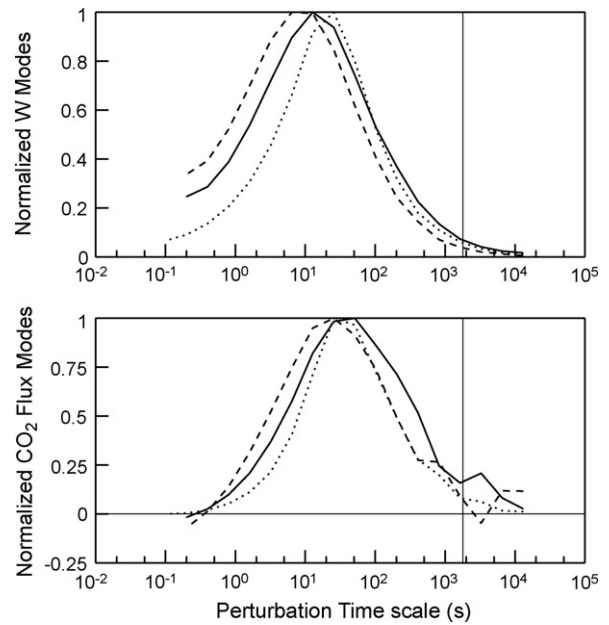

Fig. 4 – Normalized composite vertical velocity and CO_2 flux modes for the MPINE (solid), YPINE (dashed) and MFIR (dotted) sites. The modes are normalized such that the maximum mode is unity. The vertical line denotes $\tau = 30$ min.

winds less than 2 m s^{-1} at MPINE, the systematic CO_2 flux error tends to increase with decreasing wind speed, as observed in previous studies (Sakai et al., 2001).

The composite perturbation timescale-dependence of the CO_2 flux is shown in Fig. 6, where the flux has been normalized separately at each site such that the $\tau = 3.6$-h flux is unity. At each site, the flux increases rapidly with increasing perturbation timescale for τ less than about 10 min, after which the flux very slowly approaches its maximum value at the largest value of τ tested. A systematic flux error can be defined as the fractional flux deficit for a given τ compared to the flux obtained using the largest timescale tested (the record length of 3.6 h). For these purposes, the systematic error is defined to be zero for $\tau = 3.6$ h. The systematic error for $\tau = 27.3$ min ranges from 1.6% at MFIR to 4.5% at MPINE (Table 1). For $\tau = 54.6$ min, the systematic error ranges from 0.4% at MFIR to 3.5% at YPINE (Table 1). Interpolating from the dyadic timescales to nominal timescales, and averaging over the three sites, the systematic flux error decreases from 7% for $\tau = 10$ min, to 3% for $\tau = 30$ min, to 2% for $\tau = 60$ min.

Given the above definition of systematic error, the standard practice leads to a 3% under-estimate of the long-term average daytime CO_2 flux (averaged over three sites). Such error may be small in comparison to other sources of uncertainty (e.g., a variable flux footprint coupled with surface heterogeneity and unknown effects due to advection). The systematic error reported here is caused by the failure to capture all of the largest transporting scales and is generally smaller than previously published estimates (Section 1).
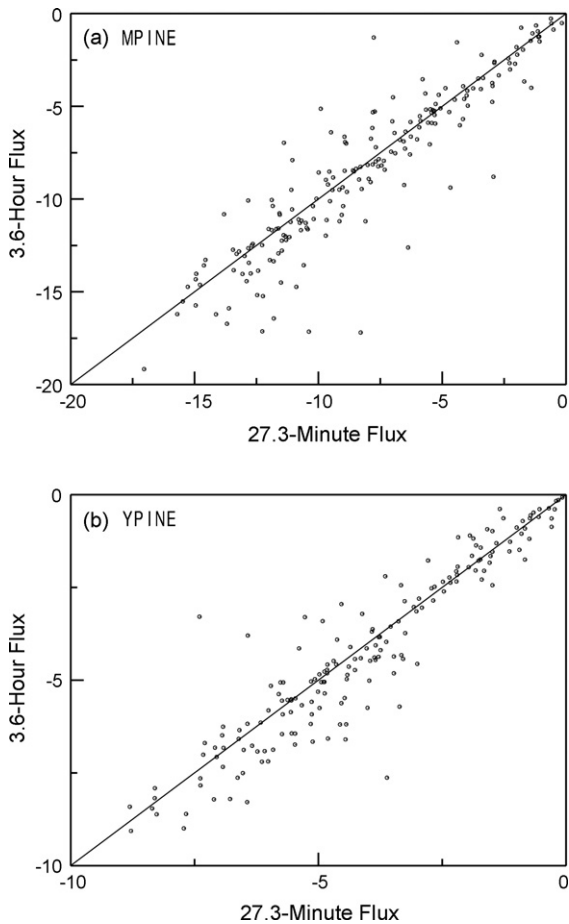
Fig. 5 – CO_2 flux estimates (μ mol CO_2 m^{-2} s^{-1}) for a perturbation timescale of 3.6 h compared to estimates for a perturbation timescale of 27.3 min at the (a) MPINE and (b) YPINE sites.

4.2. Random error

Our analysis of random errors focuses on the two pine forest sites which have much more data compared to the fir site. The incremental increase in the random error E exceeds the incremental increase in the flux as τ exceeds 10 min at MPINE
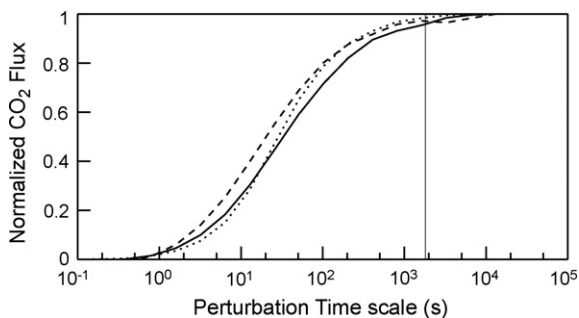

Fig. 6 – Composite timescale-dependence of the CO_2 flux normalized by the $\tau = 3.6$-h flux for the MPINE (solid), YPINE (dash) and MFIR (dotted) sites. The corresponding systematic error estimates are given in Table 1. The vertical line denotes $\tau = 30$ min.

Table 1 – Systematic CO_2 flux error (%) for dyadic perturbation timescales τ (min) for the mid-day convective boundary layer at the MPINE, YPINE and MFIR sites and for the average for the three sites				
τ (min)	MPINE	YPINE	MFIR	Average
3.4	17.9	12.3	12.0	14.1
6.8	10.4	8.2	6.8	8.5
13.6	6.8	4.1	3.1	4.7
27.3	4.5	2.7	1.6	3.0
54.6	1.5	3.5	0.4	1.8
109.2	0.3	1.7	0.1	0.7
218.5	0	0	0	0

The error is defined to be zero for $\tau = \tau_F = 3.6$ h. For reference, the magnitude of the composite net downward CO_2 flux for $\tau = \tau_F = 3.6$ h is approximately 3 at YPINE, 6 at MPINE and 12 at MFIR in μ mol CO_2 m^{-2} s^{-1}.

and 7 min at YPINE (Fig. 7). That is, any additional flux obtained by increasing τ beyond about 10 min (for example, in an attempt to reduce the systematic error), is dominated by random sampling error, and therefore little confidence can be placed in its value. The random error is due to an inadequate sample of the main transporting eddies, although nonstationarity complicates the interpretation (Section 1).

The random error modes based on the daily-differencing method are shown in Fig. 8. The estimated random flux error increases faster than the flux as τ exceeds 12 (10) min at MPINE and 4 (4) min at YPINE for the first (second) set of selection criteria (Section 3.6). Relaxing the criteria to obtain more samples of daily-differences has little influence on the scale-dependence of the random error for these data. While there are differences between the two methods of estimating
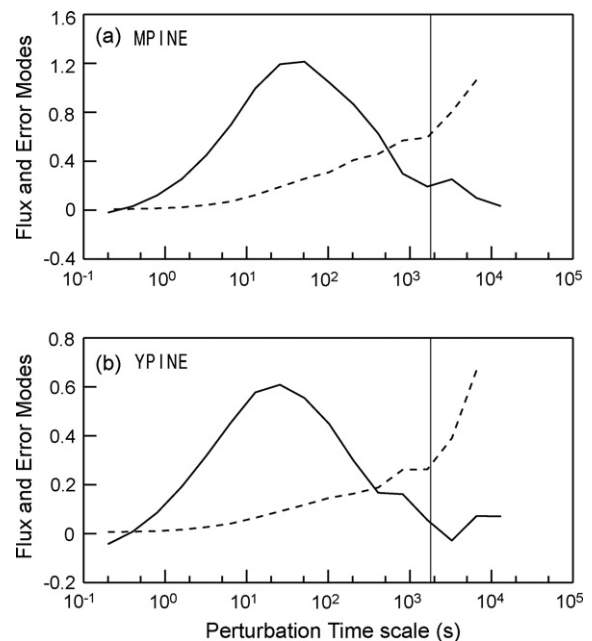

Fig. 7 – Composite timescale-dependence of the CO_2 flux modes (μ mol CO_2 m^{-2} s^{-1}, times minus one, solid curve) and the random flux sampling error modes E (μ mol CO_2 m^{-2} s^{-1}, dashed, Eq. (3)) for the (a) MPINE, and (b) YPINE sites. The vertical line denotes $\tau = 30$ min.
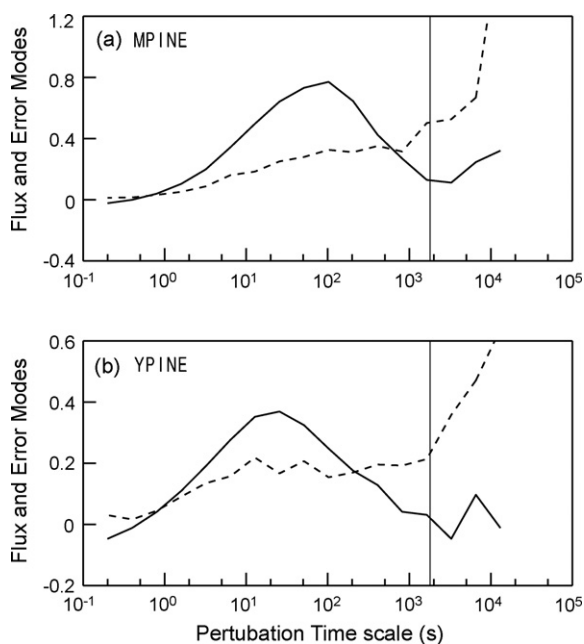
Fig. 8 – Composite timescale-dependence of the CO_2 flux modes (μ mol CO_2 m^{-2} s^{-1}, times minus one, solid curve) and the random flux sampling error modes R (μ mol CO_2 m^{-2} s^{-1}, daily-differencing method, dashed, Eq. (4)) for the (a) MPINE site and (b) YPINE site for the first set of criteria. The vertical line denotes $\tau = 30$ min.

random error (traditional and daily-differencing), and between the two sites, a consistent finding is that including transport on timescales exceeding about 10 min introduces more random error than flux.

As a sensitivity test, a turbulence-based criterion was added to the daily-differencing selection criteria. The strength of the turbulence may be important because it influences the transport of CO_2 and moisture away from the leaf boundary layer and up to the flux measurement height. Adding a further condition to the second set of criteria, where the difference in the standard deviation of vertical velocity is required to be less than 0.1 m s^{-1}, had only a minor influence on the results and reduced the sample size from 47 to 26 at MPINE and 49 to 21 at YPINE. As a further sensitivity test, a vapor pressure deficit
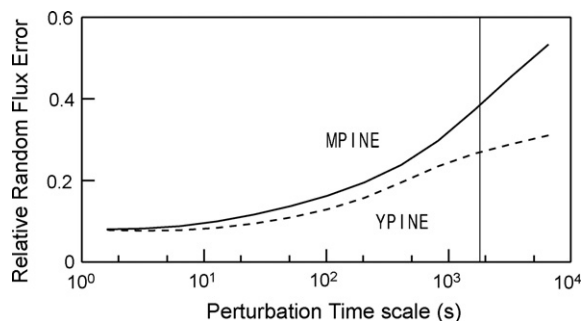

Fig. 9 – The composite timescale-dependence of the relative random CO_2 flux error (RRE, Section 3.6) for the MPINE (solid) and YPINE (dashed) sites. The vertical line denotes $\tau = 30$ min.

criterion was added to the second set of criteria with the turbulence-based condition. Additionally requiring the difference in vapor pressure deficit to be less than 0.5 kPa, reduced the sample size from 26 to 13 at MPINE and from 21 to 9 at YPINE, and did not significantly change the results compared to those in Fig. 8.

For the standard perturbation timescale of 30 min, but using 3.6-h records instead of standard 30-min records, the relative random flux error (RRE, Section 3.6) is estimated to be 38% at MPINE and 27% at YPINE (Fig. 9). Decreasing τ to 10 min decreases RRE to 26% at MPINE and 21% at YPINE. These RRE estimates would be much larger using standard 30-min records. In fact, the random error cannot even be evaluated with the traditional approach using the standard practice with $\tau = \tau_F$ because there is only a single independent sample of the flux. Recall that these estimates are for the convective boundary layer only, and that the RRE is typically much larger for the nocturnal stable boundary layer (e.g., Howell and Sun, 1999). For $\tau_F = 3.6$-h, one could use the RRE estimates (Fig. 9) to select τ if one defined an acceptable level of the relative random error.

The RRE increases faster with increasing perturbation timescale at MPINE compared to YPINE (Fig. 9). This could imply that the relative flux variability increases with measurement height (or $z - h$). A probable explanation is that the larger scale eddies are suppressed closer to the surface (Section 1), and it is the larger scale eddies which contribute heavily to the random variability because they are the most poorly sampled. However, other site-specific differences may also be important.

5. Conclusions

Flux sampling errors are directly linked to the choice of two timescales: τ and τ_F. Timescale τ defines the perturbation quantities and sets an upper-limit on the range of scales included in the flux. Products of perturbations are time-averaged over timescale τ_F because instantaneous estimates of the flux have very large scatter. To evaluate the random flux sampling error, it is normally necessary to have $\tau_F > \tau$. We used eddy-covariance tower data from the convective boundary layer at forested sites to examine the dependence of the CO_2 flux and the flux sampling errors on τ for a fixed record length of $\tau_F = 3.6$ h. The records studied are from 11 to 15 LST at the pine forest sites and 12 to 16 LST at the fir site to minimize nonstationarity associated with the morning and evening transitions.

Increasing τ to reduce the systematic error tends to increase the random error, and decreasing τ to reduce the random error increases the systematic error. The choice of τ and τ_F might be based on the relative importance of systematic and random errors for the particular application. For example, in canopy models and coupled climate-carbon models (e.g., SPA Williams et al., 2005), sub-daily (hourly) flux estimates are typically required for model validation or data assimilation (e.g., Williams et al., 2005). In such case, the random sampling errors can be large and need to be quantified. On the other hand, for biogeochemistry models that operate at daily timescales (e.g., Biome-BGC Law et al., 2003, 3PG Law et al.,

2000), and where comparisons between models and flux data are typically performed on monthly or annual timescales, the random sampling error is less important because it tends to cancel due to the large number of samples.

For sub-daily flux estimates, including contributions to the flux from timescales longer than about 10 min cannot be justified based on the random sampling error modes, even when using the long 3.6-h records. Any additional flux obtained by increasing τ beyond 10 min, for example, in an attempt to reduce the systematic flux error, is dominated by random sampling error. The random sampling error would be much larger for standard shorter records. Random sampling errors for the CO_2 flux in the mid-day convective boundary layer could be reduced compared to the current standard practice by increasing τ_F from 30 min to 4 h and by decreasing τ from 30 min to 10 min. However, the reduction in random error comes at the cost of reduced temporal resolution of the fluxes (4 h compared to 30 min), and an increase in the systematic flux error from about 3% to 7% (averaged over three sites).

When only longer term flux estimates (e.g., monthly or annual NEE) are of interest, and sub-daily temporal resolution is not required, τ can be increased to reduce the systematic error because the random error tends to cancel due to the large number of samples. The three sites studied here generally support the standard practice of using $\tau = 30$ min for long-term average fluxes. Increasing τ beyond 30 min has almost no effect on the systematic error (Table 1). The systematic CO_2 flux error may be larger than reported here at sites with deeper mid-day convective boundary layers and more well organized convection.

The three sites studied indicate that the underestimation of the mean flux using $\tau = 30$ min for the daytime convective boundary layer is 5% or less when averaged over all cases. This error may be minor compared to other problems, such as heterogeneity of the vegetation and topography, advection of CO_2 due to spatial variability in the wind field and/or the mean CO_2 concentration and roughness sublayer effects. While the error associated with using too small a value for τ is systematic presumably at all sites, errors associated with these other problems are probably site-specific and may or may not contribute significantly to systematic errors.

Acknowledgments

R E F E R E N C E S

Acevedo, O.C., Moraes, O.L.L., Degrazia, G.A., Medeiros, L.E., 2006. Intermittency and the exchange of scalars in the nocturnal surface layer. Boundary-Layer Meteor. 119, 41–55.

Caughey, S.J., Readings, C.J., 1975. Turbulent fluctuations in convective conditions. Quart. J. Roy. Meteor. Soc. 101, 537–542.

Cava, D., Contini, D., Donateo, A., Martano, P., 2008. Analysis of short-term closure of the surface energy budget above short vegetation. Agric. Forest Meteor. 148, 82–93.

Deardorff, J.W., Willis, G.E., 1985. Further results from a laboratory model of the convective planetary boundary layer. Boundary-Layer Meteor. 32, 205–236.

Feigenwinter, C., Bernhofer, C., Vogt, R., 2004. The influence of advection on the short-term CO_2 budget in and above a forest canopy. Boundary-Layer Meteor. 113, 201–224.

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

Friehe, C.A., Shaw, W.J., Rogers, D.P., Davidson, K.L., Large, W.G., Stage, S.A., Crescenti, G.H., Khalsa, S., Greenhut, G.K., Li, F., 1991. Air-sea fluxes and surface-layer turbulence around a sea surface temperature front. J. Geophys. Res. 96 (C5), 8593–8609.

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

Howell, J.F., Sun, J., 1999. Surface-layer fluxes in stable conditions. Boundary-Layer Meteor. 90, 495–520.

Howell, J.F., Mahrt, L., 1997. Multiresolution flux decomposition. Boundary-Layer Meteor. 83, 117–137.

Hunt, J.C.R., Carlotti, P., 2001. Statistical structure at the wall of the high Reynolds number turbulent boundary layer. Appl. Sci. Res. 66, 453–475.

Kristensen, L., Lenschow, D.H., Kirkegaard, P., Courtney, M., 1989. The spectral velocity tensor for homogeneous boundary layer turbulence. Boundary-Layer Meteor. 47, 149–193.

Law, B.E., Sun, O., Campbell, J., Van Tuyl, S., Thornton, P., 2003. Changes in carbon storage and fluxes in a chronosequence of ponderosa pine. Global Change Biol. 9, 510–524.

Law, B.E., Waring, R.H., Anthoni, P.M., Aber, J.D., 2000. Measurements of gross and net ecosystem productivity and water vapor exchange of a *Pinus ponderosa* ecosystem, and an evaluation of two generalized models. Global Change Biol. 6, 155–168.

Lee, X., Massman, W., Law, B., 2004. Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis. Kluwer Academic Publishers, Boston, p. 250.

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

Lothon, M., Lenschow, D.H., Mayor, S.D., 2006. Coherence and scale of vertical velocity in the convective boundary layer from a Doppler lidar. Boundary-Layer Meteor. 121, 521–536.

Lumley, J.L., Panofsky, H.A., 1964. The Structure of Atmospheric Turbulence. John Wiley & Sons, p. 239.

Mahrt, L., Lee, X., Black, A., Neumann, H., Staebler, R.M., 2000. Vertical mixing in a partially open canopy. Agric. Forest Meteor. 101, 67–78.

Mahrt, L., Moore, E., Vickers, D., Jensen, N.O., 2001. Dependence of turbulent and mesoscale velocity variances on scale and stability. J. Appl. Meteor. 40, 628–641.

Malhi, Y., Pegoraro, E., Nobre, A.D., Pereira, M.G., Grace, J., Culf, A.D., Clement, R., 2002. Energy and water dynamics of a central Amazonian rain forest. J. Geophys. Res. 107 (D20), 8061, doi:10.1029/2001JD000623.

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

Oncley, S.P., Friehe, C.A., LaRue, J.C., Businger, J.A., Itsweire, E.C., Chang, S.S., 1996. Surface-layer fluxes, profiles, and

turbulence measurements over uniform terrain under near-neutral conditions. J. Atmos. Sci. 53, 1029–1044.

Paw U, K.T., Baldocchi, D.D., Meyers, T.P., Wilson, K.B., 2000. Correction of eddy-covariance measurements incorporating both advective effects and density fluxes. Boundary-Layer Meteor. 97, 487–511.

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. Forest Meteor. 136, 1–18.

Sakai, R.K., Fitzjarrald, D.R., Moore, K.E., 2001. Importance of low-frequency contributions to the eddy fluxes observed over rough surfaces. J. Appl. Meteor. 40, 2178–2192.

Smedman, A.S., 1988. Observations of multi-level turbulence structure in a very stable atmospheric boundary layer. Boundary-Layer Meteor. 44, 231–253.

Stull, R.B., 1990. An Introduction to Boundary Layer Meteorology. Kluwer Academic Publishers, Boston, p. 666.

Sun, X.M., Zhu, Z.L., Wen, X.F., Yuan, G.F., Yu, G.R., 2006. The impact of averaging period on eddy fluxes observed at ChinaFLUX sites. Agric. Forest Meteor. 137, 188–193.

van den Kroonenberg, A., Bange, J., 2007. Turbulent flux calculation in the polar stable boundary layer: multiresolution flux decomposition and wavelet analysis. J. Geophys. Res. 112, D06112, doi:10.1029/2006JD007819.

Vickers, D., Mahrt, L., 2006a. A solution for flux contamination by mesoscale motions with very weak turbulence. Boundary-Layer Meteor. 118, 431–447.

Vickers, D., Mahrt, L., 2006b. Contrasting mean vertical motion from tilt correction methods and mass continuity. Agric. Forest Meteor. 138, 93–103.

Vickers, D., Mahrt, L., 2003. The cospectral gap and turbulent flux calculations. J. Atmos. Oceanic Technol. 20, 660–672.

Vickers, D., Mahrt, L., 1997. Quality control and flux sampling problems for tower and aircraft data. J. Atmos. Oceanic Technol. 14, 512–526.

Von Mises, R., 1964. Mathematical Theory of Probability and Statistics. Academic Press, New York, p. 694.

Wilczak, J.M., Oncley, S.P., Stage, S.A., 2001. Sonic anemometer tilt correction algorithms. Boundary-Layer Meteor. 99, 127–150.

Williams, M., Schwarz, P., Law, B.E., Irvine, J., Kurpius, M., 2005. An improved analysis of forest carbon dynamics using data assimilation. Global Change Biol. 11, 89–105.