JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 106, NO. D4, PAGES 3503–3509, FEBRUARY 27, 2001



# Sampling error in eddy correlation flux measurements

Peter L. Finkelstein[1]

Atmospheric Sciences Modeling Division, Air Resources Laboratory, NOAA, Research Triangle Park, North Carolina

Pamela F. Sims

Federal Home Loan Mortgage Corporation, McLean, Virginia

**Abstract.** Sampling errors in eddy correlation flux measurements arise from the small number of large eddies that dominate the flux during typical sampling periods. Several methods to estimate sampling, or random error in flux measurements, have been published. These methods are compared to a more statistically rigorous method which calculates the variance of a covariance when the two variables in the covariance are auto- and cross-correlated. Comparisons are offered between the various methods. Compared to previously published methods, error estimates from this technique were 20 to 25% higher because of the incorporation of additional terms in the estimate of the variance. This new approach is then applied to define the random error component of representative eddy correlation flux measurements of momentum, sensible and latent heat, carbon dioxide, and ozone from five field studies, three over agricultural crops (corn, soybean, and pasture), and two from towers over forests (deciduous and mixed). The mean normalized error for each type of flux measurement over the five studies ranged from 12% for sensible heat flux to 31% for ozone flux. There were not large or significant differences between random errors for fluxes measured over crops versus those measured over forests. The effects of stability, flux magnitude, and wind speed on measurement error are discussed.

## 1. Introduction

Eddy correlation, also known as eddy covariance (EC), measurements of heat, momentum, and trace gas fluxes are frequently the most accurate and reliable way to measure exchange processes between the atmosphere and the land or water surface. In an EC measurement the flux is the covariance of the vertical velocity ($w$) with the state variable of interest ($c$); that is, flux is equal to $\overline{w'c'}$, where $c$ can be a scalar such as temperature, concentration of a gas, etc., or a vector such as horizontal wind velocity, and the prime denotes departure from the mean. The advent of more reliable and less expensive sonic anemometers, fast response instruments for temperature, water vapor, carbon dioxide, and other trace gases, and, not least of all, the ready availability of small, cheap, and powerful computers for data acquisition, has put the equipment to make good EC measurements within the reach of many researchers.

Like any other complex measurement, EC measurements can be subject to significant bias and random errors. However, unlike many measurements, there are no straightforward ways to calibrate or audit flux measurements in the field. Aside from the obvious issues of instrument calibration and data handling, sources of error include possibly invalid assumptions inherent in the EC method about the state of the atmosphere, and errors that are a by-product of real-world sampling strategies.

*Businger* [1986], in a review of the accuracy with which the flux of trace gas measurements can be made, listed and explained several possible sources of EC error, including the following: (1) errors caused by the buoyancy effects of heat and water vapor, (2) sampling error, (3) error due to limited response time of the sensors, (4) error due to separation of the sensors, (5) error due to random noise in the system, (6) errors due to entrainment, advection, and nonstationarity in the concentration or wind fields, (7) error due to inadequate or excessive height of the sampler over the surface, (8) errors due to inadequate fetch, (9) errors due to deliquescence of aerosol particles, and (10) error due to flow distortion caused by the sampling system or tower. *Mahrt* [1998] also discussed these and other possible sources of errors in flux measurements from towers.

This paper deals only with the second of these sources of error, sampling error. Error due to sampling cannot be eliminated by experimental design, but is a consequence of the limited number of independent samples that contribute substantially to a flux during any fixed sampling period, and of the various autocorrelations and cross correlations in the two covariates. *Dyer and Bradley* [1982] noted the highly variable nature of flux measurements in the International Turbulence Comparison Experiment (ITCE), indicating the need for statistically meaningful samples to achieve repeatable flux profile relationships. *Shaw et al.* [1983] was one of the first to quantify the infrequent energy-containing eddies that contribute to a flux measurement. Longer sampling periods increase the number of independent samples and may reduce sampling error, but longer sampling times frequently lead to other problems including lack of stationarity in the atmosphere.

In field programs carefully designed to avoid the other errors noted by *Businger* [1986], sampling error will remain as one of the largest sources of uncertainty. It is desirable to

[1]On assignment to National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, North Carolina.



Paper number 2000JD900731.
0148-0227/01/2000JD900731$09.00

quantify the contribution of this error, in order to answer questions such as whether or not any two flux measurements are significantly different from each other, whether a particular flux measurement is significantly different from zero, and whether a flux measurement differs significantly from a model result. This paper addresses the need to determine the confidence limits of an individual or series of EC flux measurements.

## 2. Variance Estimation

Sampling error can be expressed as the variance of the EC measurement, that is, the variance of the covariance. The ensemble variance $\sigma^2$ about the ensemble mean of some variable $c$ should go to zero as the averaging time $T$ approaches infinity. *Lumley and Panofsky* [1964] express this as

$$\sigma^2 \cong \frac{2\overline{c'^2}\tau}{T} \rightarrow 0 \qquad \text{as} \quad T \rightarrow \infty, \tag{1}$$

where $\tau$ is the integral timescale for $c$. This assumes $\tau$ can be defined for the appropriate scale of measurement. They define the percent error of measurement of the variable $c$ as $a = \sigma/\bar{c}$, where $a$ is the coefficient of variation. They show that the averaging time needed to achieve a desired accuracy $a$ may be expressed as

$$T \cong \frac{2\overline{c'^2}\tau}{\bar{c}^2 a^2}. \tag{2}$$

*Wyngaard* [1973] points out that the integral timescale can be approximated by $\tau = l/U$, where $l$ is the size of the dominant eddy and $U$ is the mean wind speed. For neutral conditions, $l$ can be approximated by the instrument height $z$. For covariances, Wyngaard has demonstrated that the averaging times can be expressed as

$$T_{\overline{uw}} \approx \frac{z}{a^2 U}\left[\frac{\overline{(u'w')^2}}{u_*^4} - 1\right]$$
$$T_{\overline{w\theta}} \approx \frac{z}{a^2 U}\left[\frac{\overline{(w'\theta')^2}}{u_*^2 T_*^2} - 1\right], \tag{3}$$

where $u_*$ is the surface friction velocity, $T_*$ is the turbulent temperature scale, and $\overline{u'w'}$ and $\overline{w'\theta'}$ are momentum and sensible heat fluxes, respectively.

From fast response data measured by *Haugen et al.* [1971] in the Kansas experiment, *Wyngaard* [1973] estimated that in a neutral atmosphere, longer averaging times are required for covariances of heat and momentum than for variances of $\theta$ and $w$. For the unstable case ($z/L \simeq -1$, where $L$ is the Monin-Obukhov length) the error in momentum stress should be about 3 times that of heat flux for the same $T$. On the basis of the observed data the momentum stress error increased for greater instability, but heat flux error did not. He estimates that the bracketed terms in (3) are approximately equal to 20 for both momentum and heat in the neutral case and are about 100 and 12, respectively, in an unstable atmosphere. In the stable case the bracketed expression is about 5 times higher than for the neutral case, or about 100, but $l$ will be smaller.

*Sreenivasan et al.* [1978] estimated the accuracy of higher moments of $u$, $w$, $\theta$, and $q$ (water vapor), and the accuracy of the covariances $\overline{u'w'}$, $\overline{w'\theta'}$, and $\overline{w'q'}$ from data taken over water at a height of 5 m under near-neutral conditions. They used an approach similar to that of *Lumley and Panofsky* [1964], estimating the appropriate integral timescales from the autocorrelation functions which were obtained from inverse Fourier transforms of the spectral density functions. They also derived the integral timescales using assumptions about the shape of the probability density functions (PDF) for the different variables. Comparisons of the two sets of integral timescales showed approximately 30% difference for $\overline{w'u'}$ and $\overline{w'\theta'}$, and 70% difference for $\overline{w'q'}$. Using their measured integral timescales, they found from their data that

$$\varepsilon^2 = \frac{\alpha z}{UT} \tag{4}$$

was a good fit, where $\varepsilon^2$ is the mean square relative error, $z$ is the height of measurement, $U$ is the wind speed, $T$ is the duration of the measurement, and $\alpha$ is an empirical constant. Note that these measurements were made at one height. From their measurements they calculated that $\alpha$ was about 30 for momentum flux, 44 for moisture flux, and 64 for heat flux.

The problem of the variance of a flux measurement has been considered more recently by *Lenschow et al.* [1993a, 1993b] and *Mann and Lenschow* [1994]. Their development differs somewhat from that of *Lumley and Panofsky* [1964] in that it defines the more standard error variance of the central moment of the time series (i.e., the mean is removed). They also defined the integral timescale by assuming that the autocorrelation function can be modeled by an exponential. The resulting relative error is

$$\frac{\sigma_F(T)}{|F|} = \left(\frac{2\tau_F}{T}\right)^{1/2}\left(\frac{1 + r_{wc}^2}{r_{wc}^2}\right)^{1/2}, \tag{5}$$

where $\tau_F$ is the integral timescale of the flux $F$ and $r_{wc}$ is the correlation coefficient of $w$ and $c$. Looking at aircraft data, they saw large scatter in the results, perhaps due to nonstationarity or perhaps, they hypothesized, due to low values of $r$. They showed that including estimates of skewness helps to get better estimate of the error variance.

*Wesely and Hart* [1985] derived an expression for the variance of flux measurements which broke up the variance into an instrument noise as well as sampling error, and estimated each separately. They showed that

$$\frac{\sigma_{\overline{w'c'}}}{F_c} = \frac{(1 - R_{wc}^2 + b^2)^{1/2}\sigma_w\sigma_c}{\left(\frac{T\bar{u}}{z}\right)^{-1/2} F_c}, \tag{6}$$

where

$$b^2 = \frac{\overline{(w'^2)'(c'^2)'}}{\sigma_w^2\sigma_c^2}. \tag{7}$$

Their expression for sampling error can be shown to be equivalent to that of *Wyngaard* [1973].

Even in stationary periods, an exponential decay model for the autocorrelation function for typical meteorological variables is frequently a poor model, as will be shown in section 3.2. Note that in (3) the auto- and cross-covariance terms between the two time series are disregarded, while (5) includes $\tau_F$ instead of the cross covariance and does not consider the auto covariance. Disregarding these terms can, and frequently will, result in an underestimate of the variance. An approach that does not ignore these terms, or make any a priori assump-

**Table 1.** Integral Timescales $\tau$ of the Indicated Covariate Autocorrelation Function[a]

| Julian Day | Local Time | $w'U'$ | $w'T'$ | $w'O_3'$ | $w'H_2O'$ | $w'CO_2'$ |
|---|---|---|---|---|---|---|
| 203 | 1600 | 0.9 | | 0.5 | 1 | 0.9 |
| 208 | 0830 | 5.1 | 2.8 | 1.8 | 4.3 | 3.4 |
| 208 | 0930 | 2 | 0.1 | 1.8 | 2.8 | 3.2 |
| 208 | 1430 | 1.1 | 0.6 | 0.7 | 0.9 | 0.9 |
| 208 | 1500 | 1 | 1 | 0.6 | 1 | 1 |
| 208 | 1600 | 0.6 | 0.8 | 0.3 | 0.7 | 0.7 |
| 212 | 1230 | 1 | 0.5 | 0.4 | 0.8 | 0.6 |
| 212 | 1330 | 0.8 | 0.8 | 0.5 | 0.8 | 0.8 |
| 212 | 1500 | 0.6 | 0.4 | 0.4 | 0.8 | 0.8 |
| 212 | 1530 | 0.7 | 0.8 | 0.5 | 1 | 1.1 |

[a]Timescales are in seconds.

tions about the data, the boundary layer, or the autocorrelation function, is desirable and is presented below.

The following mathematically rigorous expression for the variance of a covariance, from a theorem developed by *Fuller* [1976], presented by *Bendat and Piersol* [1966], and used by us in the work of *Meyers et al.* [1998], which includes the auto- and cross-covariance terms for atmospheric fluxes, meets these criteria. (The notation follows that of Fuller, with $\hat{\gamma}_{x,y} = \overline{x'y'}$.)

$$\text{var}\,(\hat{\gamma}_{x,y}) = \frac{1}{n}\left[\sum_{p=-m}^{m} \hat{\gamma}_{x,x}(p)\hat{\gamma}_{y,y}(p) + \sum_{p=-m}^{m} \hat{\gamma}_{x,y}(p)\hat{\gamma}_{y,x}(p)\right], \tag{8}$$

where $\hat{\gamma}_{x,x}$ is the variance of $x$ estimated from data, $\hat{\gamma}_{x,y}$ is the estimated covariance of $x$ and $y$; where $x$ and $y$ are measured variables (e.g., vertical velocity and concentration of ozone), $n$ is the number of samples in the data set (i.e., 18,000 for 30 min of 10 Hz data), and $m$ is a number of samples sufficiently large to capture the integral timescale. Summing from $-m$ to $m$ is done for computational convenience and is equivalent to the sum from $-n$ to $n$ assuming that from $m$ to $n$ the correlations are approximately zero. In practice we have used an $m$ of 200 (20 s). Tests on sample data sets showed that the results varied by only 1–2% for $m$ between 100 and 400. At times longer than 400 the results do not change as long as there is no time trend in the data. The variance of the covariance estimate for each sample time series, (8), may be evaluated by computing sample estimates for the auto covariance, $\hat{\gamma}_{x,x}$, and cross covariance, $\hat{\gamma}_{x,y}$, of lag $h$, using

$$\hat{\gamma}_{x,x}(h) = \hat{\gamma}_{x,x}(-h) = \frac{1}{n}\sum_{t=1}^{n-h} (X_t - \bar{X})(X_{t+h} - \bar{X}) \tag{9}$$

and

$$\hat{\gamma}_{x,y}(h) = \hat{\gamma}_{y,x}(-h) = \frac{1}{n}\sum_{t=1}^{n-h} (X_t - \bar{X})(Y_{t+h} - \bar{Y}). \tag{10}$$

## 3. Comparison of Sampling Error Estimation Techniques

Three analytical techniques for estimating sampling error will be compared: that developed from *Wyngaard* [1973] in (3) (denoted as Wyn), *Mann and Lenschow* [1994] (M-L) in (5), and that presented above in this work (F-S) in (8)–(10). A comparison of estimates using the three approaches with a small sample of data follows.

### 3.1. Data for the Comparison

Data for this comparison were taken from a field study conducted near Plymouth, North Carolina, in the summer of 1996. Measurements of wind speed, temperature, water vapor, carbon dioxide, and ozone were taken with fast response sensors and recorded at 10 Hz. The instruments, or their intakes, were mounted 5 m above the ground on a tower located in a large, flat, soybean field. The fetch was approximately 1 km in a 180° arc centered pointing south. Details of the instrumentation system, calibration, quality control techniques, and data handling are presented by *Meyers et al.* [1998]. Ten periods, each one-half hour long, with somewhat different meteorology and fluxes, were selected for this comparison. The day and time of each example case are given in Table 1.

### 3.2. Integral Timescales

The M-L approach requires the estimation of the integral timescale of the covariates, $\tau_{w'c'} = \tau_F$. They suggest that $\tau$ can be estimated from the spectra of the flux; however, this can be quite imprecise because of its noisy nature from any one half-hour observation period. *Kaimal and Finnigan* [1994] suggest estimating the integral timescale from a plot of the autocorrelation function, assuming that it can be estimated as

$$\rho(t) = e^{-t/\tau}, \tag{11}$$

where $t$ is the time step. This model did not, however, give a satisfactory fit for many of the data used in this study. Figure 1 is an example plot of an autocorrelation function for $w'u'$, with the best fit model from (11), and a best fit model for an exponential with $t^{1/2}$ as given in (12),

$$\rho(t) = e^{-\sqrt{t/b}}. \tag{12}$$

In looking at several cases from Table 1, using the more general form $\rho = \exp(-t^a/b)$, where $a$ is a fitting parameter, it was found that the range of $a$ was between 0.4 and 0.7. Thus 0.5 is a reasonable estimate for the model if raw data are not available. However, a much simpler approach is to integrate the area under the autocorrelation curve, and not depend on a model of the autocorrelation function. Since the computed autocorrelation functions are an estimate of the ensemble mean, as in Figure 1, they frequently remain at a constant small positive value for an extended period of time. This analysis

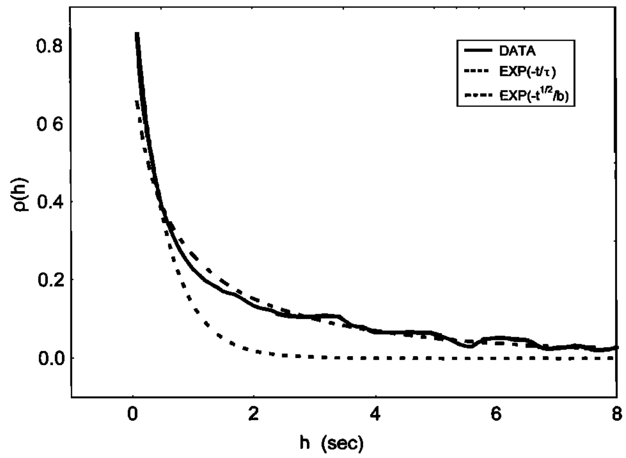


**Figure 1.** An example of the autocorrelation function $\rho(h)$ versus lag time $h$ for $w'u'$; showing the data, the best fit to exp $(-h/\tau)$, and exp $(-\sqrt{h}/a)$. For this example, $\tau$ = 0.5, and $a$ = 0.75.

truncated the integration when the range, defined as a small approximately constant, or stationary value, is reached. This imprecision introduces little error into the process. The calculated integral timescales for each of the 10 example cases and 5 fluxes are given in Table 1. Note that $\tau_{w,O_3}$ is usually shortest, while $\tau_{w,U}$, $\tau_{w,H_2O}$, and $\tau_{w,CO_2}$ are notably longer. No reason for these differences are apparent at this time.

### 3.3. Comparison of Methods for Estimating Sampling Error

Using the integral timescale for the covariances from Table 1 in the M-L method, and calculating directly from the raw, 10 Hz data for the Wyn and F-S methods, Figure 2 illustrates the mean fractional normalized sampling error of the 10 test cases for the 5 fluxes and 3 methods. In each case the square root of the variance of the covariance, or the error estimate, has been divided by the flux, or total covariance, to give a normalized fractional flux sampling error (FFSE). The mean FFSE tends to fall in the 10 to 20% range. The Wyn and M-L methods are approximately equivalent, with an average over all measurements of close to 11%, and less than the estimates using the F-S method with an average of approximately 14%. This is not unexpected, since the F-S approach takes into account the effect of the autocorrelation and cross correlation. These are not accounted for by the other techniques which inherently assume independence between variables and no autocorrelation. Autocorrelation and cross correlation usually increase the variance. Comparing the average levels of error across all fluxes for these examples, the autocorrelation and cross-correlation terms account for approximately 25% of the sampling error. This will obviously vary from case to case.

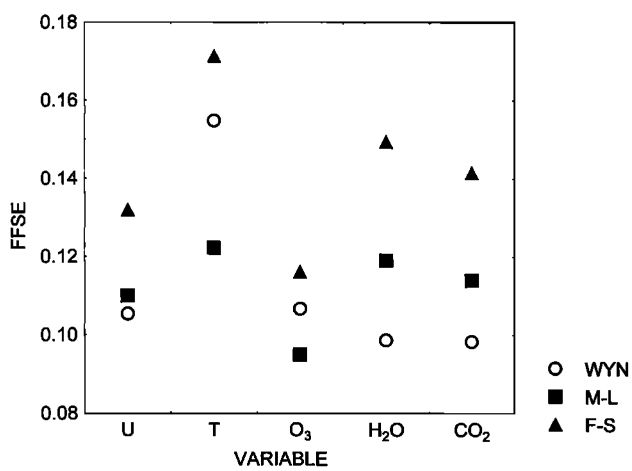


**Figure 2.** Mean value of the fractional flux sampling error (FFSE) for the models of *Wyngaard* [1973] (Wyn), *Mann and Lenschow* [1994] (M-L), and this work (Finkelstein and Sims (F-S)) for flux measurements of momentum $U$, sensible heat $T$, ozone ($O_3$), latent heat ($H_2O$), and carbon dioxide ($CO_2$). Each point is the average of the error from 10 sample cases as given in Table 1.

The range of sampling error, even over this very limited set of example cases, is quite large, suggesting that care must be used in interpreting any one flux measurement. Because the F-S model accounts for factors not considered in the other approaches and should be more inclusive in the estimation of sampling error, it will be used in subsequent analyses of this study.

## 4. Characteristics of EC Sampling Error

EC sampling errors computed using 10 Hz data from half-hour sampling periods, taken during several recently concluded field experiments, are described and characterized, in the expectation that they will be illustrative of the sampling errors in flux measurements from similar field programs. They should represent typical values of EC sampling error at a variety of locations. Because field campaigns differ in size, instrumentation, meteorology, and other details, the observations reported here may not be applicable to other studies, but enough commonality exists to give one an idea, or rule of thumb, in estimating sampling error for EC field measurements for which no direct computation is possible.

### 4.1. Data

Representative samples of EC flux measurements were selected from five field programs. These field programs included studies over a corn field near Bondville, Illinois, a pasture in Sand Mountain, Alabama, a soybean field near Nashville, Kentucky, a deciduous forest near Kane, Pennsylvania, and a mixed (deciduous and coniferous) forest in the Sand Flats State Forest in the Adirondack region of New York. Each of these sites had uniform fetches of at least ½ km to usually more than 1 km in the sampling sector. At the three agricultural sites, flux measurements were taken at 4 to 5 m above the tops of the crops. These agricultural sites, and the flux instrumentation, data collection, and QA, are described by *Meyers et al.* [1998]. At the two forest sites the flux measurements were taken on top of a walk-up tower, 12–15 m above the canopies, which were about 22–25 m high. The forest sites and instrumentation are described by *Finkelstein et al.* [2000].

For this analysis, fluxes of momentum, sensible heat (made using virtual temperature), ozone, latent heat, and carbon dioxide are considered. The same set of instruments was used at each of the studies. The wind speed and virtual temperature were measured with an ATI sonic anemometer, the ozone was measured with a fast response ozone analyzer which employs the chemiluminescent reaction of ozone with eosin-Y dye [*Ray et al.*, 1986], and the water vapor and carbon dioxide concentrations were measured with a Licor LI-6262 infrared absorp-

**Table 2.** Mean Value and Statistics on Fractional Flux Sampling Error (FFSE) Averaged Over All Cases and the Five Sites

| Flux | Mean | Standard Deviation | Maximum | Minimum | Skewness | *N* |
|---|---|---|---|---|---|---|
| *U* | 0.19 | 0.14 | 1.01 | 0.02 | 2.92 | 114 |
| *T* | 0.12 | 0.05 | 0.31 | 0.04 | 1.62 | 94 |
| $O_3$ | 0.31 | 0.21 | 1.16 | 0.10 | 2.39 | 108 |
| $H_2O$ | 0.20 | 0.20 | 1.57 | 0.08 | 4.39 | 99 |
| $CO_2$ | 0.25 | 0.32 | 2.43 | 0.07 | 4.37 | 116 |

tion analyzer. Samples for trace gas analysis are drawn from very close to the sonic anemometer through a draft tube and delivered to the analyzers that are housed in an air-conditioned shelter on the ground. Instruments were calibrated once per day, and all data were screened very carefully.

Samples from each site were selected that had easily measurable fluxes without any significant trends or nonlinearities in concentration or other sampling anomalies, and that represented a variety of stabilities and magnitudes of flux. Because these samples were selected somewhat subjectively for "good" flux measurements that were not "too" small, some bias toward lower average FFSE probably has been introduced. This bias would result from the fact that small fluxes frequently occur during the night, due to increased stomatal and atmospheric resistance. The very small nighttime fluxes usually have larger sampling errors since the frequency of significant eddies that contribute to the flux transport decreases. The results below should be interpreted in that light.

### 4.2. Analysis and Results

The FFSEs for each type of flux measurement are given in Table 2, along with the number of observations, the minimum and maximum value, the standard deviation, and the skewness. Except for sensible heat flux sampling error, which was 12%, the mean sampling errors for the other measurements were in the 20–30% range. Minimum values were quite small for momentum and sensible heat flux, but were closer to 10% for the gases. Maximum values were over 100% for all but sensible heat. Scatter in the sampling error of flux measurements is considerable, ranging from ±5% for sensible heat to ±32% for $CO_2$.

To show the utility of normalizing the standard deviation of the covariance by the absolute value of the flux to produce the FFSE, Figure 3 is a plot of FFSE against the flux for sensible and latent heat. As can be seen, the error is proportional to the flux magnitude, as most of the points lie on a straight line with zero slope. For sensible heat flux there is very little scatter. For latent heat flux there is more, especially for small fluxes where a few points are quite large. The other fluxes behaved similarly. It is not unreasonable to expect that the average FFSE for a sample of flux measurements from a field study will well represent the random sampling error over the full range of flux measurements, excluding those near zero.

Table 3 gives the mean value of the FFSE for each measurement variable and site. Figure 4 shows the mean FFSE for each site and flux measurement with ±1 $\sigma$ error bars. No clear pattern emerges from this table and figure with regard to site differences. No site has consistently higher or lower sampling error for all flux measurements. Instead, the highest and lowest values of sampling error are scattered among each of the sites.

One hypothesis tested was that there would be a difference between the agricultural sites, in which the canopy is short, the surface is smooth, and the flux measurements are close to the surface, and the forest sites, in which the canopy is much higher, the roughness layer may be deeper, and the samplers are farther above the top of the canopy in what may or may not be a constant flux layer. We can see this difference to some extent for momentum and sensible heat fluxes. On the basis of the Newman-Keuls post hoc test [*Casella and Berger*, 1990] only the flux errors for sensible heat from the forest sites are significantly (5%) different from the agricultural sites. On average, the sampling errors are lower for the agricultural sites for momentum and sensible heat fluxes, but there are exceptions, such as Sand Mountain for momentum and Kane for sensible heat. On the other hand, the trace gas fluxes show no consistent pattern at all. When considering all fluxes, it would

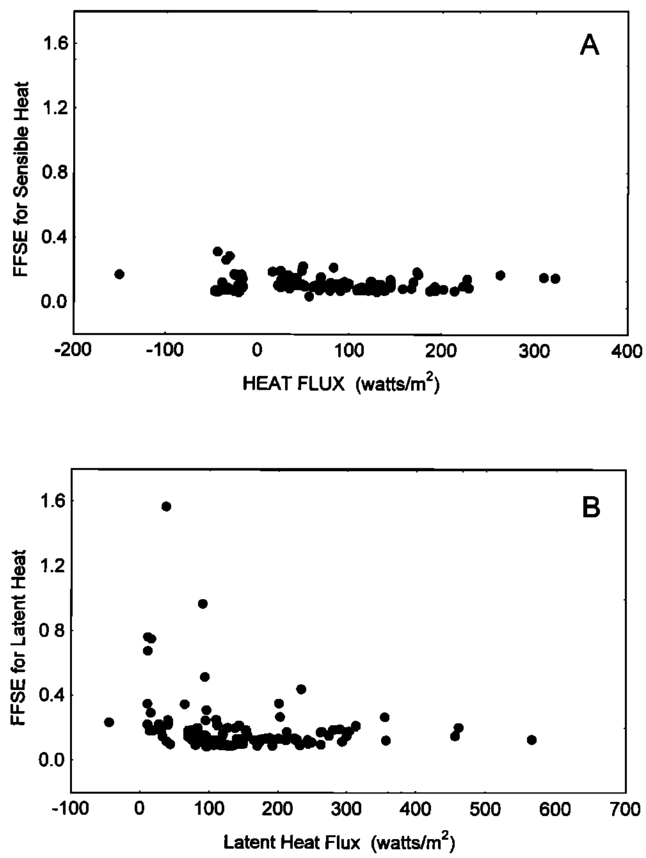


**Figure 3.** FFSE versus flux of (a) sensible heat and (b) latent heat for all sites. The almost constant value of FFSE throughout the range of flux indicates the utility of using the flux to normalize the error in covariance measurements.

**Table 3.** Mean Value of the Fractional Flux Sampling Error at Each Site for Each Measure[a]

| Site | $U$ | $T$ | $O_3$ | $H_2O$ | $CO_2$ |
|---|---|---|---|---|---|
| Sand Flat (f) | 0.24 | 0.20 | 0.23 | 0.23 | 0.31 |
| Kane (f) | 0.27 | 0.16 | 0.31 | 0.24 | 0.26 |
| Sand Mountain (ag) | 0.28 | 0.13 | 0.27 | 0.35 | 0.15 |
| Bondville (ag) | 0.13 | 0.11 | 0.32 | 0.14 | 0.26 |
| Nashville (ag) | 0.16 | 0.09 | 0.33 | 0.22 | 0.26 |

[a]Forest sites are denoted as (f), and agricultural sites are denoted as (ag).

seem to be hard to show from these data that flux measurements taken over the forest are inherently less accurate than those taken close to the ground over a smooth surface. This suggests that the eddy structure over the forests was not different enough from the agricultural sites to impact the accuracy of the flux measurements since sampling error is a function principally of the eddy structure and instrumentation, but the instrumentation was the same at all sites.

One could also speculate that there might be differences between fluxes that were directed toward the ground (e.g., momentum, ozone, and carbon dioxide) and those that were directed up (e.g., sensible and latent heat). However, the summaries in Tables 2 and 3 and Figure 4 show that this hypothesis is not supported.

If atmospheric stability impacts flux sampling error, as suggested by *Wyngaard* [1973], that effect is rather small, as can be seen in Figure 5. In all cases but heat flux there is variability with stability, but no strong pattern. Sampling errors for heat and ozone flux show almost no effect. Momentum and carbon dioxide fluxes seem to have a larger sampling error at extreme stability and instability, but there are few samples in these groups, so that those points should be interpreted with caution. Wyngaard noted that the sampling error should be 2 to 3 times greater for momentum flux than sensible heat flux when the atmosphere is unstable; that is, $z/L \simeq -1$. We note that sensible heat FFSE is indeed lower than that for the other fluxes by a factor of 2 to 4, but for all stabilities, and not just in the unstable region, although the differences are smaller between heat and momentum in near-neutral stability; that is, $z/L \simeq 0$.

The methods of *Wyngaard* [1973] and *Sreenivasan et al.* [1978], (3) and (4), would also suggest that FFSE would be inversely proportional to the square root of wind speed. The FFSE results of the F-S method, when plotted against $1/\sqrt{U}$, Figure 6, show that this hypothesis is not supported. Momentum flux has only a weak relation with the wind speed, while sensible heat has almost none. While not shown in the figure, this held true for the other fluxes as well.

## 5. Conclusions

The direct calculation of the variance of the covariance is a powerful and inclusive method for calculating the random sampling error in EC measurements. It takes into account sources of error not considered by previously published approaches and, as a consequence, tends to produce larger estimates for the error. For the approximately 100 samples taken from a variety of field programs over both low agricultural crops and forests, the normalized sampling error, FFSE, ranged from approximately 10% for sensible heat to 25–30% for trace gases. These rather large sampling errors in flux measurements should be kept in mind when considering problems related to closing the energy balance, $CO_2$ budgets, or comparing deposition model results to measurements.

Since there is no absolute standard or more precise method for measuring fluxes, these methods cannot be checked against a standard, but they could be evaluated if multiple independent observations were made in one place. We hope that data

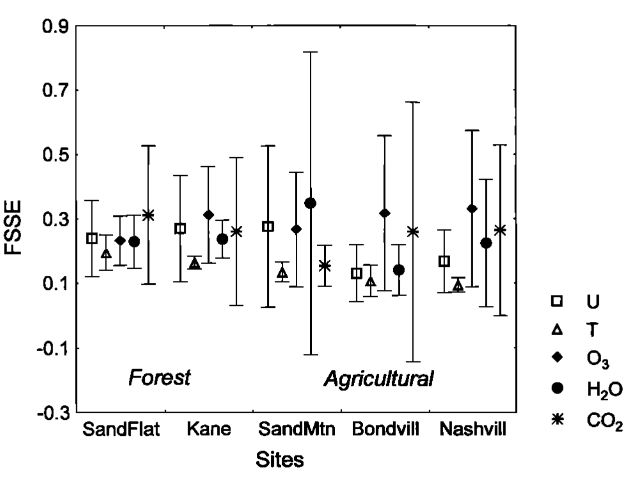


**Figure 4.** FFSE by site with ±1 σ error bars. The fluxes are as given in Figure 2. Sites that are in forests or agricultural lands are denoted.

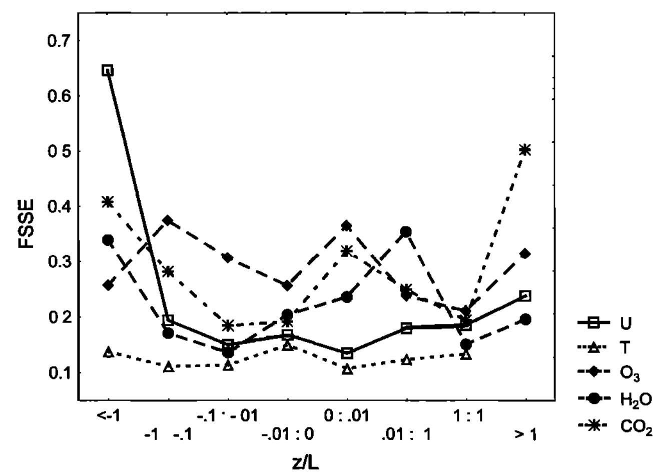


**Figure 5.** FFSE for various stability classes, as denoted by $z/L$. Data from all sites are combined. Values at the extreme ends have very few datum points making up the average and should be interpreted cautiously.

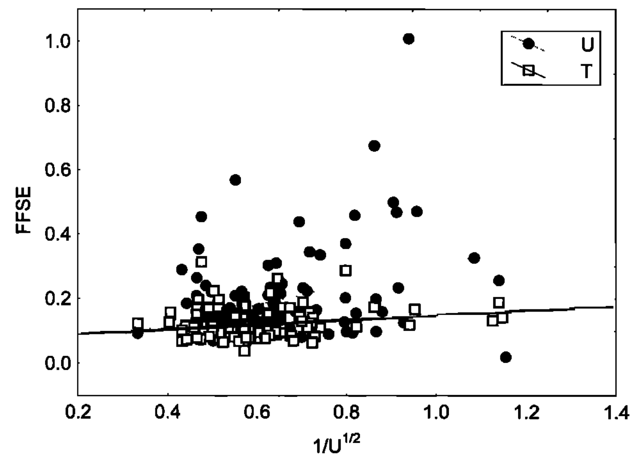


**Figure 6.** FFSE plotted against $1/\sqrt{U}$ for momentum and sensible heat fluxes. Data from all sites are combined. The least squares best linear fit to the data are shown.

from past or future studies which have such a design could be made available so that these methods could be tested.

The autocorrelation function for covariances is frequently assumed to fall off as an inverse exponential of time, but it is better modeled as an inverse exponential of the square root of time. Using this approach, a better estimate of the integral timescales would be available. Integral timescales for the daytime test cases evaluated in section 3.2 are very short, ranging from approximately 0.5 to 3 s.

**Acknowledgments.** We would like to thank John F. Clarke, Thomas G. Ellestad, and Eric O. Hebert for their help in conducting the field studies, and Sharon LeDuc and David Holland for their very helpful comments on this paper. Our research was funded by the U.S. Environmental Protection Agency. This paper has been reviewed and approved for publication by both the Environmental Protection Agency and the National Oceanic and Atmospheric Administration. Mention of commercial products does not constitute an endorsement for use.

## References

Bendat, J. S., and A. G. Piersol, *Measurement and Analysis of Random Data*, John Wiley, New York, 1966.

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

Casella, G., and R. Berger, *Statistical Inference*, 650 pp., Wadsworth, Belmont, Calif., 1990.

Dyer, A. J., and E. F. Bradley, An alternative analysis of the flux-gradient relationships in the 1976 ITCE, *Boundary Layer Meteorol.*, *22*, 3–19, 1982.

Finkelstein, P. L., T. G. Ellestad, J. F. Clarke, T. P. Meyers, D. B. Schwede, E. O. Hebert, and J. A. Neal, Ozone and sulfur dioxide dry deposition to forests: Observations and model evaluation, *J. Geophys. Res.*, *105*, 15,365–15,377, 2000.

Fuller, W. A., *Introduction to Statistical Time Series*, 470 pp., John Wiley, New York, 1976.

Haugen, D. A., J. C. Kaimal, and E. F. Bradley, An experimental study of Reynolds stress and heat flux in the atmospheric surface layer, *Q. J. R. Meteorol. Soc.*, *97*, 168–180, 1971.

Kaimal, J. C., and J. J. Finnigan, *Atmospheric Boundary Layer Flows: Their Structure and Measurement*, 290 pp., Oxford Univ. Press, New York, 1994.

Lenschow, D. H., J. Mann, and L. Kristensen, How long is long enough when measuring fluxes and other turbulent statistics?, *NCAR Tech. Note NCAR/TN-389+STR*, 53 pp., Natl. Cent. for Atmos. Res., Boulder, Colo., 1993a.

Lenschow, D. H., J. Mann, and L. Kristensen, How long is long enough when measuring fluxes and other turbulent statistics?, *J. Atmos. Oceanic Technol.*, *11*, 661–673, 1993b.

Lumley, J. L., and H. A. Panofsky, *The Structure of Atmospheric Turbulence*, 239 pp., Wiley-Interscience, New York, 1964.

Mahrt, L., Flux sampling errors from aircraft and towers, *J. Atmos. Oceanic Technol.*, *15*, 416–429, 1998.

Mann, J., and D. H. Lenschow, Errors in airborne flux measurements, *J. Geophys. Res.*, *99*, 14,519–14,526, 1994.

Meyers, T. P., P. L. Finkelstein, J. F. Clarke, T. G. Ellestad, and P. F. Sims, A multilayer model for inferring dry deposition using standard meteorological measurements, *J. Geophys. Res.*, *103*, 22,645–22,661, 1998.

Ray, D., D. H. Stedman, and G. J. Wendel, Fast chemiluminescent method for measurement of ambient ozone, *Anal. Chem.*, *58*, 598–600, 1986.

Shaw, R. H., J. Tavangar, and D. P. Ward, Structure of the Reynolds stress in a canopy layer, *J. Appl. Meteorol.*, *22*, 1922–1931, 1983.

Sreenivasan, K. R., A. J. Chambers, and R. A. Antonia, Accuracy of moments of velocity and scalar fluctuations in the atmospheric surface layer, *Boundary Layer Meteorol.*, *14*, 341–359, 1978.

Wesely, M. L., and R. L. Hart, Variability of short term eddy-correlation estimates of mass exchange, in *The Forest-Atmosphere Interaction: Proceedings of the Forest Environmental Measurements Conference, Oak Ridge, Tennessee, October 23–28, 1983*, pp. 591–612, D. Reidel, Norwell, Mass., 1985.

Wyngaard, J. C., On surface layer turbulence, in *Workshop on Micrometeorology*, edited by D. A. Haugen, pp. 101–149, Am. Meteorol. Soc., Boston, Mass., 1973.

P. L. Finkelstein (corresponding author), National Exposure Research Laboratory, U.S. EPA, MD-80, Research Triangle Park, NC 27711. (finkelstein.peter@epa.gov)

P. F. Sims, Federal Home Loan Mortgage Corporation, 8200 Jones Branch Dr., McLean, VA 22102.

(Received June 1, 2000; revised November 1, 2000; accepted November 7, 2000.)