Intercomparison of Nine Micrometeorological Stations during the BEAREX08 Field Campaign

JOSEPH G. ALFIERI,* WILLIAM P. KUSTAS,* JOHN H. PRUEGER,+ LAWRENCE E. HIPPS,# JOSÉ L. CHÁVEZ,@ ANDREW N. FRENCH,& AND STEVEN R. EVETT**

* *Hydrology and Remote Sensing Laboratory, USDA-ARS, Beltsville, Maryland*
+ *National Laboratory for Agriculture and the Environment, USDA-ARS, Ames, Iowa*
# *Department of Plants, Soils, and Biometeorology, Utah State University, Logan, Utah*
@ *Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, Colorado*
& *Arid-Land Agricultural Research Center, USDA-ARS, Maricopa, Arizona*
** *Conservation and Production Laboratory, USDA-ARS, Bushland, Texas*

(Manuscript received 29 July 2010, in final form 11 March 2011)

ABSTRACT

Land–atmosphere interactions play a critical role in regulating numerous meteorological, hydrological, and environmental processes. Investigating these processes often requires multiple measurement sites representing a range of surface conditions. Before these measurements can be compared, however, it is imperative that the differences among the instrumentation systems are fully characterized. Using data collected as a part of the 2008 Bushland Evapotranspiration and Agricultural Remote Sensing Experiment (BEAREX08), measurements from nine collocated eddy covariance (EC) systems were compared with the twofold objective of 1) characterizing the interinstrument variation in the measurements, and 2) quantifying the measurement uncertainty associated with each system. Focusing on the three turbulent fluxes (heat, water vapor, and carbon dioxide), this study evaluated the measurement uncertainty using multiple techniques. The results of the analyses indicated that there could be substantial variability in the uncertainty estimates because of the advective conditions that characterized the study site during the afternoon and evening hours. However, when the analysis was limited to nonadvective, quasi-normal conditions, the response of the nine EC stations were remarkably similar. For the daytime period, both the method of Hollinger and Richardson and the method of Mann and Lenschow indicated that the uncertainty in the measurements of sensible heat, latent heat, and carbon dioxide flux were approximately 13 W m^{-2}, 27 W m^{-2}, and 0.10 mg m^{-2} s^{-1}, respectively. Based on the results of this study, it is clear that advection can greatly increase the uncertainty associated with EC flux measurements. Since these conditions, as well as other phenomena that could impact the measurement uncertainty, are often intermittent, it may be beneficial to conduct uncertainty analyses on an ongoing basis.

1. Introduction

The land–atmosphere interface lies at the nexus of the complex web of interconnections and feedbacks linking biogeophysical and biogeochemical processes. As a result, the exchange of mass, energy, and momentum between the land surface and the atmosphere plays a critical role in regulating numerous meteorological, hydrological, and ecological processes. Thus, as pointed out by French et al. (2005), among others, accurately describing land–atmosphere interactions and their role in subsequent processes is essential for a broad array of applications with significant social, economic, and environmental impacts. These applications range from managing water and other natural resources (Neale et al. 2005; Gowda et al. 2008) to predicting agricultural productivity (Steduto et al. 2007; Ko and Piccinni 2009) and carbon sequestration (Svejcar et al. 2008; Alfieri et al. 2009a) to forecasting weather and climate (Chen et al. 2007).

Surface fluxes can vary substantially both spatially and temporally due to localized differences in both surface and atmospheric conditions. For example, Alfieri et al. (2007) found that while evapotranspiration (ET) during drought is strongly linked to soil moisture content, water availability is only one of several key environmental factors that influence the moisture flux over time. Therefore,

Corresponding author address: Joseph G. Alfieri, HRSL, USDA-ARS, 10300 Baltimore Avenue, Building 007, Room 104, BARC-W, Beltsville, MD 20705.
E-mail: joe.alfieri@ars.usda.gov

DOI: 10.1175/2011JTECH1514.1

© 2011 American Meteorological Society

measurements of surface fluxes collected across the continuum of surface and atmospheric conditions are requisite both for investigating the underlying mechanisms of land–atmosphere exchange processes and for developing, parameterizing, and evaluating numerical models and remote sensing–based products (e.g., Cosh et al. 2004; Anderson et al. 2005; LeMone et al. 2008; Alfieri et al. 2009c). For example, measurements were collected at 10 representative grassland and cropland sites during the 2002 International H2O Project (LeMone et al. 2007). Similarly, the 12 micrometeorological stations deployed during the 2002 Soil Moisture–Atmospheric Coupling Experiment (SMACEX) were installed in both corn and soybean fields (Kustas et al. 2005). Finally, a recent study by Blanken et al. (2009) compared measurements from alpine tundra and subalpine forest. On a larger scale, numerous regional measurement networks, such as AmeriFlux (Baldocchi et al. 2001) and ChinaFLUX (Yu et al. 2006), have been developed to collect long-term measurements across the full range of environmental conditions and ecosystems types. A core assumption of these networks is that measurements from the individual field sites are representative of the larger ecosystem (Baldocchi 2008; Chasmer et al. 2008).

To correctly interpret or compare measurements collected at different locations or times, it is imperative that the total uncertainty associated with the measurements be fully characterized (Prueger et al. 2005; Meek et al. 2005; Myklebust et al. 2008). Characterizing the uncertainty, however, is a difficult task because of the complexity of collecting surface flux measurements and the many potential sources of error. Businger (1986), for example, lists 10 potential sources of uncertainty in turbulent flux measurements ranging from limitations in the sensor response and violations of the theoretical underpinnings of measurement techniques to random noise. Each of these sources of uncertainty contributes to either the random uncertainty or the systematic error. As described by Billesbach (2011), random uncertainty reduces the precision of the measurement and, thereby, confidence that the measured value represents the true value, while systematic errors reduce the accuracy of the measurement by introducing a bias. Although this definition is useful, some care is needed in its application because not all sources of systematic error have a constant effect; these sources of error, which Moncrieff et al. (1996) refer to as selective systematic errors, are due to intermittent sources. Example sources of random error include the variability and heterogeneity of the measurement source area (Katul et al. 1999) and errors associated with the sensors themselves (Hollinger and Richardson 2005). An example of a source of systematic error is the undermeasurement of the turbulent fluxes during the night because of insufficient turbulent mixing (Mahrt 1998).

The total uncertainty is the combination of the systematic error and random uncertainty. Because of the potential impacts of measurement uncertainty and error on both research and applied activities, a number of methods have been developed to estimate these quantities for surface flux measurements. One of the first of these is the method built on the fundamental principles of turbulent transport by Mann and Lenschow (1994) to determine the uncertainty of airborne flux measurements. This method was later modified by Hollinger and Richardson (2005) for tower measurements collected using the eddy covariance (EC) method. In the same paper, they also proposed an alternate method for estimating uncertainty based on the differences in the measured flux from a pair of EC systems. Most recently, Billesbach (2011) discusses a ''random shuffle'' method that is unique because it quantifies only the random uncertainty.

Using a combination of these and other statistical techniques, the nine EC systems used in the Bushland Evapotranspiration and Agricultural Remote Sensing Experiment (BEAREX08) were evaluated in order to ascertain both the agreement among the measurement systems and the total uncertainty associated with each. Specifically, the study focused on three quantities: the sensible heat flux (H), latent heat flux (λE), and carbon dioxide flux (F_c). This analysis using the data from BEAREX08 represents a unique opportunity to characterize the uncertainty of a significant number of EC systems at once and to do so using measurements collected under strongly advective conditions. As such, this study will not only provide a better understanding of the uncertainty associated with the measurements collected during this one field campaign, it will also provide valuable guidance for understanding the uncertainty of the EC measurements collected in many arid and semiarid regions where advective conditions are commonplace.

The following section provides an overview of the BEAREX08 field campaign along with a description of the field site and data collection procedures. Section 3 discusses the statistical analysis methods, while the fourth section discusses the results of the analysis. The final section discusses the conclusions drawn from this study.

2. The 2008 Bushland Evapotranspiration and Agricultural Remote Sensing Experiment

a. *Overview of the field campaign*

BEAREX08 was conducted from June through August 2008 at the United States Department of Agriculture-Agriculture Research Service (USDA-ARS) Conservation

and Production Research Laboratory (CPRL) near Bushland, Texas (35.183°N, 102.100°W). The overarching goal of the field campaign was to investigate improved methods for monitoring and characterizing surface energy, moisture, and carbon fluxes using remote sensing in irrigated agricultural environments that have significant regional advection. To accomplish this goal, a suite of instrumentation was deployed including nine EC micrometeorological stations that collected data over irrigated and dryland cotton, grassland, and bare soil sites. Other instrumentation included airborne flux and remote sensing platforms; a tethered sonde system; a network of soil moisture, heat flux and temperature sensors; ground-based remote sensing instruments; and four large precision weighing lysimeters.

b. Site description

Prior to field deployment, the EC systems were collocated with an average separation distance of 3 m along the northeastern edge of an irrigated (center pivot) wheat field (Fig. 1) so that the instrument response of the EC stations could be compared. The wheat field was approximately 900 m east to west and 450 m north to south. Because of the configuration of the instrumentation systems, the minimum fetch for any system was nearly 200 m. In comparison, the flux footprint, which was calculated using the method of Schuepp et al. (1990), indicated that the source area of 90% of the measured flux at nominally 2 m was within 35 m of the micrometeorological station. At the time of the intercomparison, the height of the wheat canopy varied between 30 and 40 cm.

The 4-day period from 23 May [day of year (doy) 144] through 26 May (doy 147) 2008 was selected for this analysis because it was characterized by mostly clear-sky conditions without precipitation or synoptic weather activity. The prevailing wind during this period was from the southwest, minimizing the potential for data contamination due to flow distortion caused by air passing through the boom and mast of the measurement systems and maximizing upwind fetch of the wheat crop. Additionally, there were strongly advective conditions during the afternoon, particularly on 25 May. As a result of the lateral transport of warm, dry air across the study site during these advective periods, λE was enhanced and H was directed downward.

c. Data collection

Each of the nine EC micrometeorological systems was equipped with a sonic anemometer (CSAT-3, Campbell Scientific, Logan, Utah) to measure the orthogonal wind velocity components. The sonic anemometers were either new or were factory calibrated prior to the field campaign. A fine-wire thermocouple (FW05, Campbell Scientific) measured air temperature and an open-path infrared gas analyzer (LI-7500, Li-COR Biosciences, Lincoln, Nebraska) measured water vapor and carbon dioxide concentration.[1] These measurements were collected at a nominal measurement height of 2.25 m AGL and a frequency of 20 Hz. The instruments were mounted facing southwest (225°). A zero and span calibration of each of the gas analyzers was conducted immediately before the field campaign. Additional instruments included a combined humidity and temperature sensor (HMP45, Vaisala, Helsinki, Finland) and, in the case of five of the systems (stations 1–5), a four-component net radiometer (CNR-1, Kipp and Zonen, Delft, Netherlands). Stations B5 and B6 were equipped with Q*7 net radiometers (Radiation Energy Balance Systems, Bellevue, Washington).

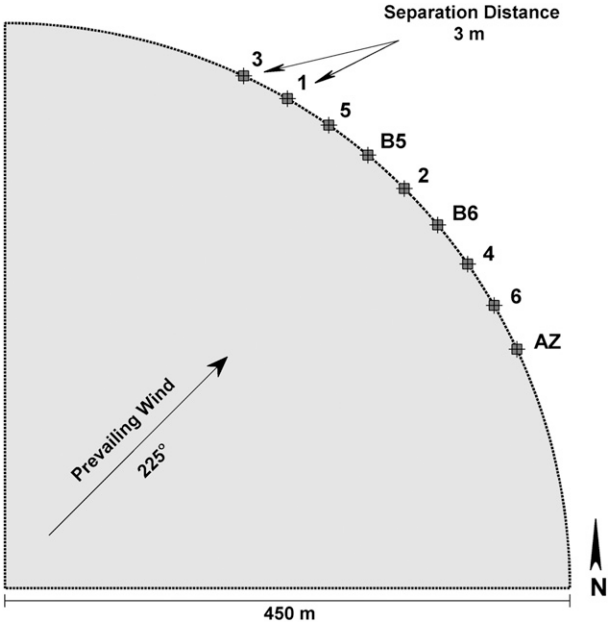

FIG. 1. A simple schematic showing the approximate location of each eddy covariance station along the perimeter of the wheat field.

The flux data were postprocessed using the full complement of standard corrections and adjustments. Nonphysical values and outliers were first removed without replacement from the high frequency (20 Hz) data using a moving window algorithm based on the method outlined by Højstrup (1993). The threshold for identifying a high frequency data point as an outlier was 4.5 standard deviations outside the window mean. In total for

[1] Trade and company names are given solely for the purpose of providing specific information and does not imply a recommendation or endorsement by the USDA.

the 4-day study period, of the nearly 7 000 000 high frequency measurements collected for each wind velocity component and scalar quantity, the despiking algorithm removed 9, 554, 114, and 31 measurements for vertical wind speed, carbon dioxide density, water vapor density, and air temperature, respectively. Next, a two-dimensional rotation was applied to the wind velocity components (u, v, and w) so that the coordinate system was aligned into the prevailing wind direction (Tanner and Thurtell 1969; Kaimal and Finnigan 1994). Third, the data were corrected for sensor displacement and frequency response attenuation (Massman 2000; Massman and Lee 2002). Finally, 1-h block average turbulent fluxes were calculated. These fluxes were then corrected for the effects of heat and water vapor density (Webb et al. 1980; Leuning 2004). The air temperature from the sonic anemometer was also corrected for humidity effects according to Schotanus et al. (1983).

3. Statistical methods

a. Variography

Variography is well-established geostatistical technique that has been proven to be an effective means of estimating spatial variability for a broad range of applications. The implementation described here parallels Alfieri et al. (2009b) who used variography to quantify the spatial variability in airborne flux measurements. Briefly, assuming the underlying spatial processes—in this case, turbulent exchange between the surface and the atmosphere at each of the measurement locations—exhibits second-order stationarity, so the correlation between the measurements at each location can be characterized using the semivariance according to

$$\gamma(h) = \tau^2 + \sigma^2[1 - C(h)], \tag{1}$$

where γ is the sermivariance, h is the separation distance between locations, τ^2 is the nonspatial component of the variance, σ^2 is the spatial component of the variance, and $C(h)$ is a valid covariance function. From this relationship, it is evident that relationship between the measurements is independent of location when σ^2 is zero. Second-order stationarity implies that the spatial process has a constant mean and a covariance that depends only on the distance between locations (Schabenberger and Gotway 2005). A valid covariance function is positive definite and fulfills the assumption that proximal locations are more strongly related than distal ones (Isaaks and Srivastava 1989).

b. Concordance correlation coefficient

Although the Pearson correlation coefficient (ρ) is commonly used to ascertain the agreement between measurements, the statistic measures only the precision of the measurements; it does not consider the effects of measurement accuracy (Lin et al. 2002). The concordance correlation coefficient (ρ_c) was introduced by Lin (1989, 1992) to assess the agreement between measurements collected by two methods in terms of both the precision and accuracy of the measurements. In brief, ρ_c is the product of ρ, which reflects the precision of the measurements, and a second coefficient (χ_a), which is indicative of the measurement accuracy. It can be expressed as

$$\rho_c = \rho\chi_a = \rho\frac{2\sigma_i\sigma_j}{\sigma_i^2 + \sigma_j^2 + (\mu_i - \mu_j)^2}, \tag{2}$$

where σ is the standard deviation, σ^2 is the variance, μ is the mean, and i and j are indices indicating the measurement method. The statistic was later extended by Barnhart et al. (2002) to determine the agreement between multiple measurements. This overall concordance correlation coefficient (ρ_o) can be determined according to

$$\rho_o = \frac{2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\rho_{i,j}\sigma_i\sigma_j}{(N-1)\sum_{i=1}^{N}\sigma_i^2 + \sum_{i=1}^{N-1}\sum_{j=i+1}^{N}(\mu_i - \mu_j)^2}. \tag{3}$$

c. Generation of the reference fluxes

To characterize the distribution of the fluxes, reference datasets were generated according to

$$\hat{F}_t = \frac{1}{M}\sum_{m=1}^{M}F_{t,m}, \tag{4}$$

where $\hat{F}_t$ is the reference value for a given flux (F) at time t, M is the number of measurements collected at time t, and $F_{t,m}$ is the flux measurement from the mth EC system at time t. The interinstrument variability ($\sigma_{\text{I-I}}$) was also estimated for each time step in terms of the standard deviation of the flux measurements from all of the EC systems for a given time:

$$\sigma_{\text{I-I}_t} = \left[\frac{1}{M-1}\sum_{m=1}^{M}(F_{t,m} - \hat{F}_t)^2\right]^{1/2}, \tag{5}$$

where $\sigma_{\text{I-I}_t}$ is the standard deviation of the measurements for time t. Then, $\sigma_{\text{I-I}}$ for a given time period was calculated as

$$\sigma_{\text{I-I}} = \left[\frac{1}{N}\sum_{t=1}^{N}\sigma^2_{\text{I-I}_t}\right]^{1/2}. \tag{6}$$

The interinstrument coefficient of variability ($\text{CV}_{\text{I-I}}$), which provides a measure of the variability among the measurements collected during a given time period relative to their mean, was also calculated as the mean of the $\text{CV}_{\text{I-I}}$ for each time step ($\text{CV}_{\text{I-I}_t}$) during the period. The interinstrument coefficient of variability was calculated for each time step according to

$$\text{CV}_{\text{I-I}_t} = \frac{\sigma_{\text{I-I}_t}}{|\hat{F}_t|}. \tag{7}$$

d. Uncertainty estimate of Hollinger and Richardson

This method, which will be referred to as the HR method hereafter, for estimating the uncertainty associated with EC measurements of the turbulent fluxes was first proposed by Hollinger and Richardson (2005). The method derives the uncertainty based on the difference the paired measurements collected independently by two collocated EC systems. Following Hollinger and Richardson (2005) and Richardson et al. (2006), the uncertainty of the flux measurements collected independently by two different sensors (F_i and F_j, respectively) can be decomposed into the true flux ($\dot{F}$) and a pair of error terms (η_i and η_j) such that

$$F_i = \dot{F} + \eta_i, \tag{8a}$$

$$F_j = \dot{F} + \eta_j. \tag{8b}$$

If η_i and η_j are assumed to be independent random variables drawn from identically the same distribution with a mean of zero and standard deviation of $\sigma(\eta)$, the expected difference between F_i and F_j is zero. The variance of the difference between F_i and F_j is equal to the variance of the difference between the error terms, that is, $\sigma^2(\eta_i - \eta_j)$. In this case, the uncertainty of the two measurement systems can be characterized by estimating $\sigma(\eta)$ according to

$$\sigma(\eta) = \frac{1}{\sqrt{2}}\sigma(\delta_{i,j}), \tag{9}$$

where $\delta_{i,j}$ is the difference between F_i and F_j, and $\sigma(\delta_{i,j})$ is the variance of that difference.

Based on the earlier work of Hollinger and Richardson (2005) and Richardson et al. (2006), the distribution of the error estimates is not expected to be Gaussian. Rather, it is expected to have a Laplace (double exponential) distribution. The probability density function of this has the form

$$f(x) = \frac{\exp\left(-\dfrac{|x-\mu|}{\beta}\right)}{2\beta}, \tag{10}$$

where μ is the mean and β is the scaling parameter defined as

$$\beta = \frac{1}{N}\sum_{i=1}^{N}|x_i - \mu|. \tag{11}$$

The variance of the Laplace distribution is defined as

$$\sigma^2 = 2\beta^2. \tag{12}$$

In the case of this analysis, there are multiple uncertainty estimates associated with each EC station. These estimates were aggregated to a single value according to

$$\sigma_i(\eta) = \frac{1}{\sqrt{2}}\left\{\frac{1}{N(J-1)-1}\left[\sum_{j\neq i}^{J}(N-1)\sigma^2(\delta_{i,j}) + \sum_{j\neq i}^{J}N\overline{\delta_{i,j}}^2 - N(J-1)\overline{\delta}_i^2\right]\right\}^{1/2}, \tag{13}$$

where i indicates the EC system associated with the aggregate value, j is index indicating each of the J total EC systems, N is the number of measurement periods, $\overline{\delta_{i,j}}$ is the mean difference for the measurements from the ith and jth EC system, and $\overline{\delta}_i$ is the mean difference between the measurements from the ith EC system and all other EC system. It is calculated as

$$\overline{\delta}_i = \frac{1}{NJ}\sum_{j\neq i}^{J}N\overline{\delta_{i,j}}. \tag{14}$$

e. Partitioning of the variance

After determining the best-fit line via ordinary least squares (OLS) linear regression, the total error associated with a given EC station can be partitioned between its systematic and random component as described by Willmott (1982). Each of these errors can be related as follows:

$$\sigma_T^2 = \sigma_S^2 + \sigma_R^2, \tag{15}$$

where σ_T^2, σ_S^2, and σ_R^2 are the total, systematic, and random errors, respectively. They are defined as

$$\sigma_T^2 = \sigma^2(F - \dot{F}), \tag{16a}$$

$$\sigma_S^2 = \sigma^2(\hat{F} - \dot{F}), \tag{16b}$$

$$\sigma_R^2 = \sigma^2(F - \hat{F}), \tag{16c}$$

where $\dot{F}$ is the true flux, F is the measured flux, and $\hat{F}$ is the predicted flux from the linear regression.

f. Uncertainty estimate of Mann and Lenschow

This method, which will be referred to as the ML method hereafter, was derived from the governing principles of turbulent transport to estimate the uncertainty in the flux by considering the joint relationship between vertical airflow and the scalar quantity (e.g., temperature) of interest. The method was originally developed for application to airborne flux measurements (Lenschow et al. 1994; Mann and Lenschow 1994) but was later modified for use with EC towers by Hollinger and Richardson (2005) and Richardson et al. (2006). In this modified form, the uncertainty of the flux measurements is estimated according to

$$\sigma = |F|\left[\frac{2t(1 + \rho_{w,x}^2)}{L\rho_{w,x}^2}\right]^{1/2}\left(1 - a\frac{z_m}{z_i}\right), \tag{17}$$

where σ is the measurement uncertainty, F is the measured flux, t is the integral time scale of the measurement, L is the length of the measurement period, $\rho_{w,x}$ is the correlation coefficient between the vertical wind speed (w) and the scalar quantity of interest (x), a is an empirical coefficient relating the flux at the top of the convective boundary layer to the flux at the surface, z_m is the measurement height, and z_i is the height of the convective boundary layer. Although more complex methods have been suggested by Finkelstein and Sims (2001), the integral time scale is estimated here as $t = z_m/\overline{u}$, where $\overline{u}$ is the mean wind speed. This simple approach has been used successfully in numerous past studies (e.g., Hollinger and Richardson 2005; Billesbach 2011). During the day, the measurement height of the surface flux stations are much less than the convective boundary layer height, so the final term on the right-hand side can be neglected. The same assumption cannot be made for the overnight or early morning hours when the convective boundary layer is shallow or has collapsed altogether (LeMone et al. 2002).

4. Results and discussion

a. Spatial analysis

The EC systems were located within 24 m of one another and the upwind wheat field was both level and visually uniform. As a result, variability among the measurements due to the relative locations of the sensor systems was not expected. To confirm this, however, a variography analysis was conducted for each of the turbulent fluxes.

For H, a comparison of the mean τ^2 (42.6 $W^2\ m^{-4}$) and mean σ^2 (0.9 $W^2\ m^{-4}$) suggests that the relative locations of the sensor systems did not contribute significantly to the variability in the measurements. If the contribution to the total variability is considered in relative terms, spatial variability accounted for between 0% and 6% of the total variability with a median value of 0% and an interquartile range of 1.9%. (Because of the highly skewed distribution of the spatial contribution—the contribution was less than 2% for nearly 80% of the measurements—the median and interquartile range are used to provide a more robust estimate of the central tendency and dispersion of the data). In the case of λE, the mean τ^2 and σ^2 were 1.3 and 143 $W^2\ m^{-4}$, respectively. The relative contribution of the spatial component to the total variability ranged between 0% and 7.1% with a median value of 0% and an interquartile range of 1.3%. Finally, for F_c, the mean τ^2 and σ^2 were 0.0042 and 0 $mg^2\ s^{-2}\ m^{-4}$, respectively.

The analysis also suggests good agreement between the measurements from the nine EC systems. The total variability among the measurements of H averaged 6.6 $W\ m^{-2}$ while the total variability among the measurements of λE and F_c were 12.0 $W\ m^{-2}$ and 0.065 $mg\ s^{-1}\ m^{-2}$, respectively.

b. Measurement agreement

To more rigorously ascertain the agreement between the nine EC systems, ρ_o was calculated for each of the turbulent fluxes: H, λE, and F_c. The coefficient was calculated not only for the whole of the study period, but also for the daytime, overnight, and transitional periods. The daytime period is defined as the period from 0900 to 1900 central standard time (CST), the overnight period is the period from 2100 to 0700 CST, and the transitional period is defined as the periods from 0700 to 0900 CST and 1900 and 2100 CST.

With values ranging between 0.90 and 0.97, ρ_o indicates strong agreement among the nine EC systems for all of the turbulent fluxes when the full intercomparison period was considered. However, as can be seen in Table 1, the strength of this agreement varied depending on the time of day. This is particularly evident in the case of F_c, which had ρ_o of 0.67 for the overnight period. The weaker agreement during the overnight period likely was due to the calm conditions and a lack of turbulent mixing that typified nocturnal conditions during the intercomparison. During the overnight period, the maximum

TABLE 1. For each of the turbulent fluxes, the overall concordance correlation coefficients (ρ_o) for both the full intercomparison period and subsets of data collected during differing times of day are shown.

Time period	H	λE	F_c
Full	0.95	0.97	0.90
Daytime	0.95	0.92	0.87
Overnight	0.91	0.86	0.67
Transitional	0.98	0.90	0.84

wind speed was 3.1 m s^{-1} and the mean wind speed was 2.2 m s^{-1} while the maximum friction velocity (u_*) was 0.38 m s^{-1} and the mean was 0.21 m s^{-1}. Similarly, the mean Monin–Obukhov stability parameter was 0.11. This is in agreement with numerous studies (e.g., Goulden et al. 1996; Blanken et al. 1998; Baldocchi 2003; Alfieri et al. 2009a) that have shown that the measurement of surface fluxes, and especially F_c, can be problematic under stable nighttime conditions when turbulent intensity is low. If only those measurements collected when u_* exceeded 0.18 m s^{-1}—typically, a threshold of between 0.10 and 0.20 m s^{-1} is used to ensure sufficient turbulent intensity (Novik et al. 2004)—are used to calculate ρ_o during the overnight period, the agreement among the measurement of F_c increases somewhat to 0.78.

c. Evaluation of the distribution of the fluxes

To further investigate the differences among the measurements during differing times of day, the relationship between $\sigma_{\mathrm{I-I}}$ and both time of day and the magnitude of the flux was evaluated beginning with H. Based on the nonparametric Kruskal–Wallace test (Burt and Barber 1996) at a 95% confidence level, the measured values of H from all of the EC stations, as well as the reference H (H_{ref}), which is the average of all nine EC systems, had statistically equivalent Gaussian distributions with a mean of −28 W m^{-2} and standard deviation of 60 W m^{-2}. Although the associated means and standard deviations changed, the distributions of each of the subsets remained Gaussian when the data were partitioned into the day, overnight, and transitional periods (Table 2).

Although $\sigma_{\mathrm{I-I}}$ for H was less than 10 W m^{-2} regardless of the time of day (Table 2), which again indicates strong agreement among the measurements from the nine EC systems, there was a clear diurnal pattern with greatest variability among the nine EC systems occurring during the daytime and transitional periods (Figs. 2c,d). Since the flux measurements with magnitude near zero tended to occur during these periods, this suggests that $\sigma_{\mathrm{I-I}}$ for H varies in proportion with the magnitude of H. A scatterplot of $\sigma_{\mathrm{I-I}}$ as a function of the absolute value of H_{ref} confirms this (Fig. 2d). If the $\mathrm{CV}_{\mathrm{I-I}}$ for H is considered instead, the diurnal variability is less pronounced (Figs. 2f,g). Additionally, $\mathrm{CV}_{\mathrm{I-I}}$ appears to decrease asymptotically with the increasing magnitude of H_{ref} (Fig. 2h). None of these relationships, however, are as clearly defined for H as they are for the other turbulent fluxes. This is possibly because of the advective conditions at the study site, which tended to suppress H especially during the afternoon (Figs. 2a,b).

While the Kruskal–Wallace test at a 95% confidence level indicated that both the individual measurements of λE and the reference latent heat flux (λE_{ref}) shared the same distribution, the distribution of λE was more complex than the distribution of H. The distribution of λE had a bimodal distribution that can best be described as a composite of two distinct Gaussian distributions. The first distribution represents the overnight and transitional periods and has a mean of 43 W m^{-2} and a standard deviation of 35 W m^{-2}. The second represents the

TABLE 2. Summary statistics show the variability both of the reference dataset and among the nine EC systems for different time periods.

Time period	Full	Daytime	Overnight	Transitional
	Sensible heat flux			
Period mean (W m^{-2})	−28	−18	−38	−30
Period standard deviation (W m^{-2})	60	81	24	62
Interinstrument standard deviation (W m^{-2})	7.2	8.2	4.8	9.4
Interinstrument coefficient of variability	12%	15%	8%	14%
	Latent heat flux			
Period mean (W m^{-2})	179	368	24	92
Period standard deviation (W m^{-2})	177	105	17	60
Interinstrument standard deviation (W m^{-2})	22	32	5.6	17.3
Interinstrument coefficient of variability	13%	9%	16%	16%
	Carbon dioxide flux			
Period mean (mg s^{-1} m^{-2})	−0.09	−0.42	0.20	−0.02
Period standard deviation (mg s^{-1} m^{-2})	0.35	0.25	0.11	0.31
Interinstrument standard deviation (mg s^{-1} m^{-2})	0.07	0.06	0.06	0.09
Interinstrument coefficient of variability	30%	25%	32%	37%

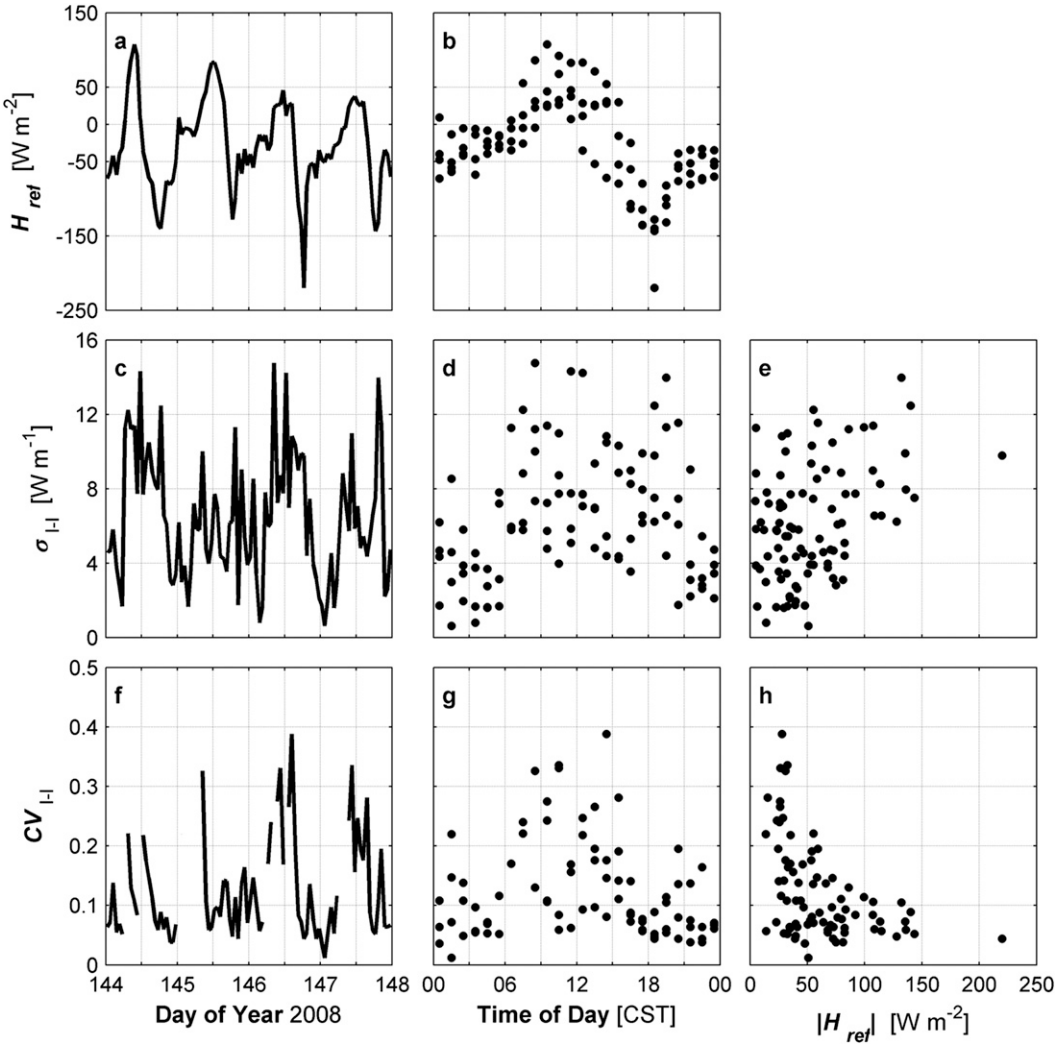

FIG. 2. The turbulent flux, as well as σ_{I-I} and CV_{I-I}, associated with H are shown as a function of (a),(c),(f) time, (b),(d),(g) time of day, and (e),(h) the magnitude of H_{ref}.

daytime period and has a mean of 368 W m^{-2} and a standard deviation of 105 W m^{-2}.

In contrast to H, the diurnal pattern of both λE and σ_{I-I} for λE are well defined (Figs. 3a–d). The greatest σ_{I-I} tended to occur during the day when λE_{ref} was greatest while the lowest variability among the measurements occurred overnight (Table 2) when the magnitude of the flux was near zero. This again indicates that the variability among the measurements from the nine EC systems is proportional to the magnitude of the flux. As can be seen in Fig. 3e, the relationship between σ_{I-I} and the magnitude of λE_{ref} is linear when the scatter due to measurements collected under highly advective conditions of 25 May is ignored. In contrast, CV_{I-I} for λE shows the opposite pattern with the lowest relative variability tending to occur during the day (Figs. 3f,g).

When CV_{I-I} is plotted as a function of the magnitude of λE_{ref}, the resulting curve is asymptotic and approaches 0.04 when the magnitude of the flux is large (Fig. 3h). One interpretation of the curve of CV_{I-I} as a function of the magnitude of the flux is in terms of the relative uncertainty of the measurements. In this light, the curve suggests that one could reasonably place more confidence in the daytime measurements than those collected overnight. The curve also suggests that at least part of the variability in the flux measurements is caused by factors that are unaffected by the changes in environmental conditions over time.

As with H_{ref} and λE_{ref}, the Kruskal–Wallace test at a 95% confidence level indicated that both the individual measurements of F_c and the reference carbon dioxide flux ($F_{c\ ref}$) shared the same distribution. The

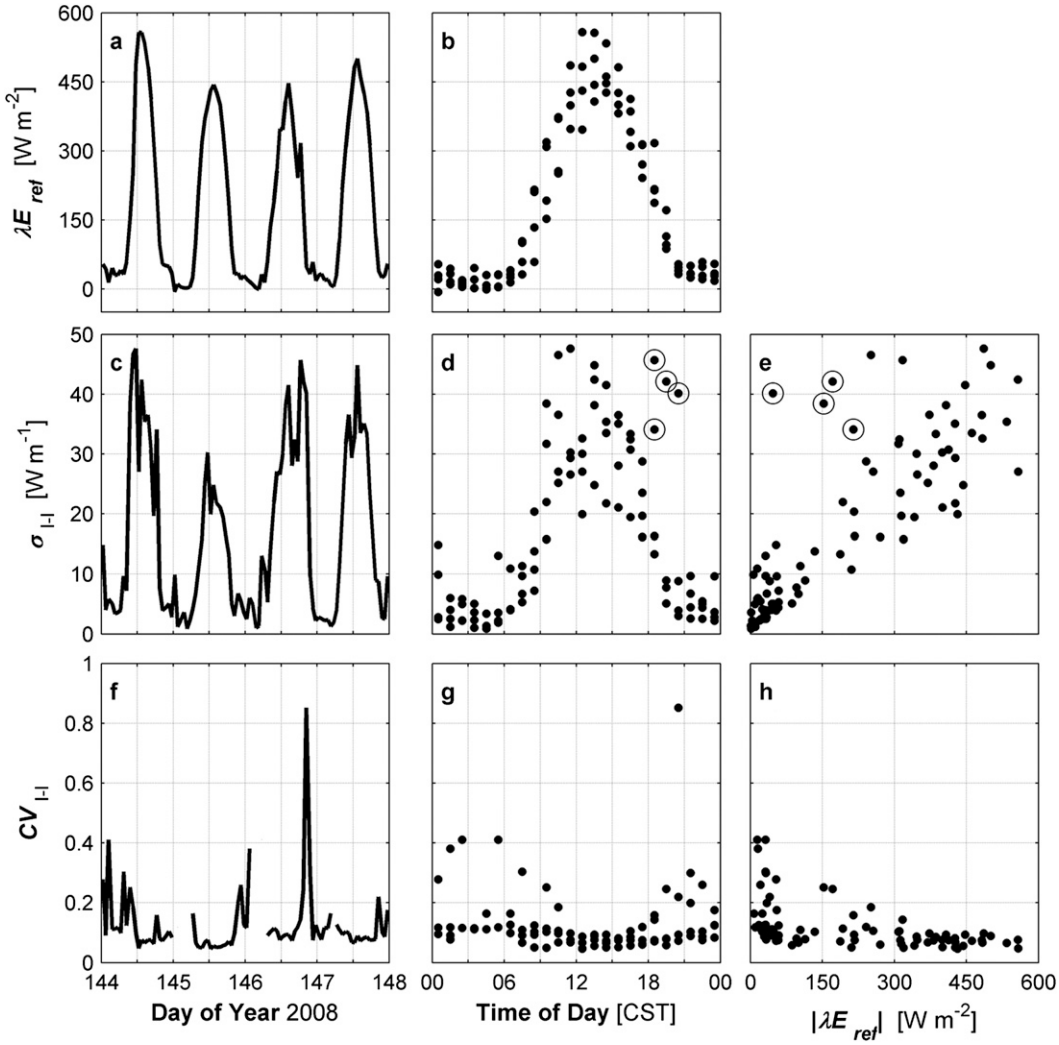

FIG. 3. The turbulent flux, as well as $\sigma_{\text{I-I}}$ and $\text{CV}_{\text{I-I}}$, associated with λE are shown as a function of (a),(c),(f) time, (b),(d),(g) time of day, and (e),(h) the magnitude of λE_{ref}. (d),(e) The circled points indicate the measurements collected during the highly advective conditions in the afternoon and evening of day of year 146.

distribution of $F_{c\ \text{ref}}$ was multimodal distinct Gaussian distributions representing the daytime, overnight, and transitional periods (Table 2).

As was the case with λE, the diurnal pattern in F_c and $\sigma_{\text{I-I}}$ for F_c are well defined (Figs. 4a–d). In this case, however, the greatest variability among the measurements occurred during the transitional period when the magnitude of F_c was small. Indeed, $\sigma_{\text{I-I}}$ was the same (0.06 mg s^{-1} m^{-2}) during the daytime and overnight periods. This is also evident in the scatterplot of $\sigma_{\text{I-I}}$ as a function of the magnitude of F_c (Fig. 4e), which shows substantial variability in $\sigma_{\text{I-I}}$ when F_c is small and has nearly constant variability when the magnitude of F_c is greater than 0.5 mg s^{-1} m^{-2}. The relative variability again shows asymptotic behavior and approaches 0.03 when the magnitude of F_c is large. The strongly advective conditions on doy 146 did not appear to have an effect on F_c.

d. *Assessment of uncertainty using the method of Hollinger and Richardson*

In an effort to ascertain the uncertainty associated with each of the nine EC stations, the method described by Hollinger and Richardson (2005) and Richardson et al. (2006) was applied to all possible pairs of EC systems. In addition, the technique was applied to both the full study period and the subsets based on time of day. As can be seen in Table 3, which shows, as an example, the results of the analysis for H when the full intercomparison period was considered, the uncertainty estimates varied depending on the combination of EC systems evaluated. For that data, the uncertainty estimates

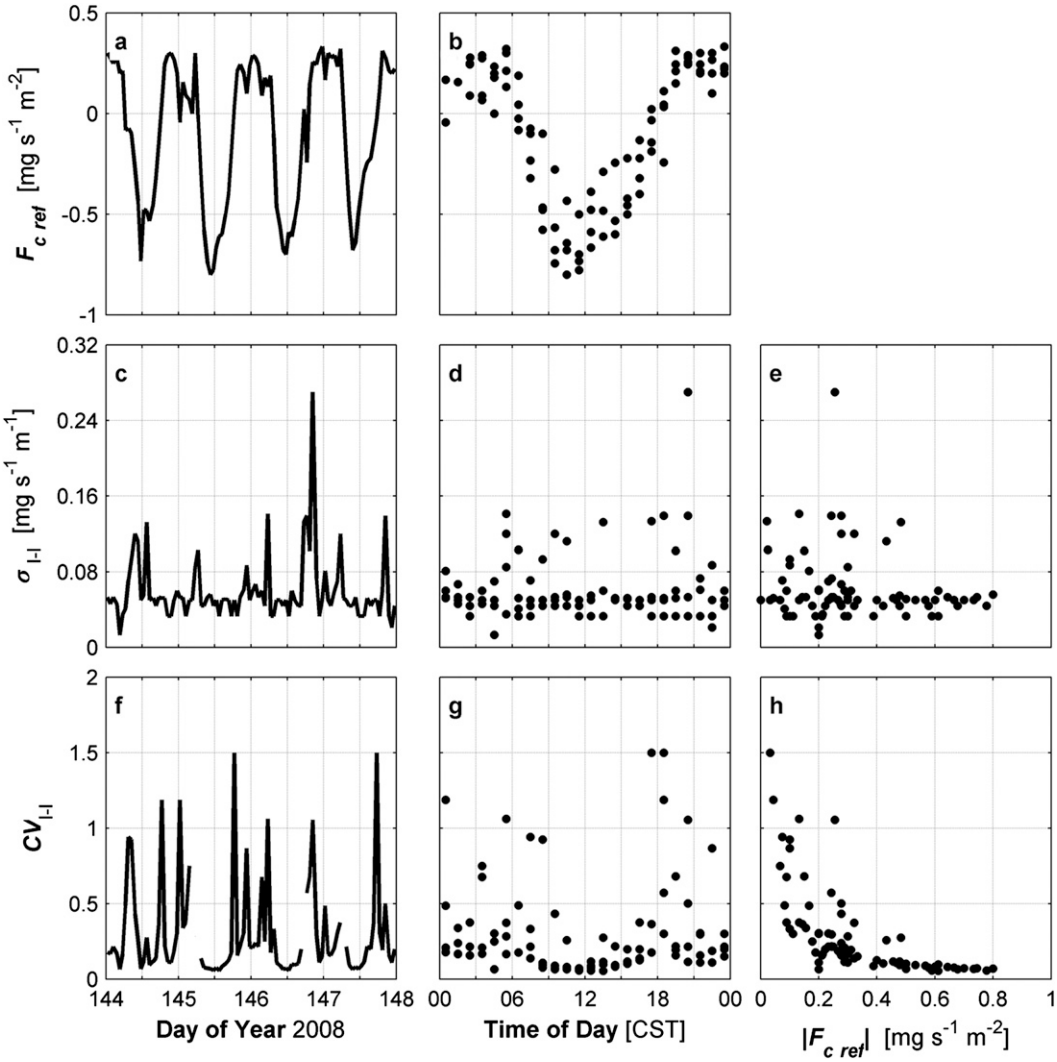

FIG. 4. The turbulent flux, as well as $\sigma_{\text{I-I}}$ and $\text{CV}_{\text{I-I}}$, associated with F_c are shown as a function of (a),(c),(f) time, (b),(d),(g) time of day, and (e),(h) the magnitude of $F_{c\ \text{ref}}$.

spanned between 6.6 and 12 W m^{-2}. While this range is fairly small, it (in conjunction with the nonzero mean differences) does suggest that the underlying assumption of this method that the measurement uncertainty is equal for all of the EC systems may not be valid for these data.

If the underlying assumption of the HR method holds true, one would expect the uncertainty estimates associated with each EC system to share the same distribution. That being the case, the validity of this assumption can be assessed by testing whether the uncertainty estimates for each EC station have the same distribution. This was done by again using the Kruskal–Wallis test. When the test was applied the uncertainty estimates for H when the full intercomparison period was considered, it indicated the distributions were the same at the 95% confidence level. When the Kruskal–Wallis test was applied to the uncertainty estimates for the daytime period, the same results were found. The distribution of uncertainty estimates for H during the overnight and transitional periods, however, were found to differ among the nine EC system. In the case of the overnight period, it was found that the distribution of the uncertainty estimates for station AZ differed in a statistically significant manner from the distributions of the other EC systems. Specifically, the mean uncertainty estimate for station AZ was significantly higher than for the other stations (Table 4). Statistically significant differences in the distributions of the uncertainty estimates of H were also found in the case of the transitional period; during this time period, the distribution of estimates from station 6 also differed from the others.

TABLE 3. The uncertainty estimates calculated with the method of Hollinger and Richardson (2005) and Richardson et al. (2006) using the measurements of H for the full intercomparison period are shown.

EC station	1	2	3	4	5	6	AZ	B5	B6
1		8.9	8.6	9.0	6.6	10	10	7.2	9.0
2	8.9		7.1	7.4	9.9	8.3	10	8.1	7.8
3	8.6	7.1		8.8	9.2	8.7	12	8.9	8.5
4	9.0	7.4	8.8		9.1	9.5	9.6	9.3	9.3
5	6.6	9.9	9.2	9.1		10	8.8	7.5	10
6	10	8.3	8.7	9.5	10		11	9.8	9.3
AZ	10	10	12	9.6	8.8	11		9.7	9.7
B5	7.2	8.1	8.9	9.3	7.5	9.8	9.7		8.4
B6	9.0	7.8	8.5	9.3	10	9.3	9.7	8.4	
Aggregate	9.2	8.8	9.4	9.8	9.4	10	12	8.8	9.3

By applying the Kruskal–Wallis test to the uncertainty estimates for λE, statistically significant differences at the 95% confidence level were found in the distributions of the uncertainty estimates for the whole intercomparison period for stations 6 and 3 as well as the daytime for station 3, and overnight and transitional periods for station AZ. By applying the Kruskal–Wallis test to the uncertainty estimates for F_c, statistically significant differences at the 95% confidence level were found in the distributions of the uncertainty estimates for the overnight for station 2 and transitional periods for station B5 (Table 4).

e. Regression analysis

Since there was a degree of ambiguity among the uncertainty estimates calculated using the HR method, a pair of additional analyses was conducted. As was discussed earlier, strong advection occurred during the afternoon and early evening. Also, as was discussed earlier, the u_* measurements during the overnight period indicated that there were intermittent periods of low turbulent intensity. As a result, a bias could have been introduced into the measurements of some, but not necessarily all, of the EC measurements. The presence of a bias would cause both nonzero mean differences and variability in the uncertainty estimates seen with the HR method.

To determine if bias affected the measurements, the measured flux from each EC system was regressed against the reference flux described earlier and the variance of the residuals was partitioned between its systematic and

TABLE 4. A summary of the aggregated uncertainty estimates calculated for each of the turbulent fluxes using the method of Hollinger and Richardson (2005) and Richardson et al. (2006) is provided. Bold font indicates that statistically significant differences in the distribution of the uncertainty estimates were found at the EC station.

Time period	Statistic	EC station 1	2	3	4	5	6	AZ	B5	B6	Aggregate of all stations
					Sensible heat flux						
Full	Uncertainty	9.2	8.8	9.4	9.8	9.4	10	12	8.8	9.3	9.8
	Mean difference	0.9	−0.1	1.0	−0.6	−0.2	−0.2	1.9	−2.6	−4.3	−0.5
Daytime	Uncertainty	13	12	14	13	13	13	14	11	11	13
	Mean difference	0.8	0.6	2.8	−1.0	−0.6	−1.0	−1.9	−4.7	−7.7	−1.1
Overnight	Uncertainty	6.0	6.3	5.2	5.9	6.2	6.2	**7.5**	6.0	5.4	6.1
	Mean difference	0.9	−1.3	0.5	−0.3	0.0	0.2	**1.3**	0.1	0.5	0.2
Transitional	Uncertainty	9.4	9.3	11	9.6	11	**13**	**14**	11	9.0	11
	Mean difference	0.8	0.9	−2.2	0.0	0.3	**0.7**	**3.5**	−4.0	0.3	0.0
					Latent heat flux						
Full	Uncertainty	24	22	**31**	23	26	**50**	25	22	23	29
	Mean difference	−6.3	0.0	**7.8**	1.9	3.4	**7.9**	−6.1	−1.9	−7.8	−0.1
Daytime	Uncertainty	37	36	**50**	38	40	**77**	36	35	35	46
	Mean difference	−12	1.0	**17**	4.6	7.4	**18**	−9.4	−6.4	−13	0.8
Overnight	Uncertainty	5.5	5.9	5.0	5.6	5.3	**7.6**	**6.1**	5.8	5.8	6.0
	Mean difference	0.9	−0.9	−0.2	0.3	0.3	**0.9**	**−0.9**	0.2	−0.3	0.0
Transitional	Uncertainty	19	17	19	17	17	**30**	**27**	20	16	21
	Mean difference	−9.8	0.2	2.8	−0.1	0.8	**3.1**	**−9.2**	4.0	−7.4	−1.7
					Carbon dioxide flux						
Full	Uncertainty	0.08	0.08	0.08	0.07	0.08	0.07	0.09	0.08	0.08	0.08
	Mean difference	0.00	−0.02	0.00	0.00	0.01	−0.01	−0.01	−0.02	0.00	−0.01
Daytime	Uncertainty	0.08	0.08	0.09	0.08	0.08	0.08	0.09	0.09	0.08	0.08
	Mean difference	0.00	0.00	−0.02	−0.01	−0.01	−0.02	−0.02	−0.02	−0.02	−0.02
Overnight	Uncertainty	0.06	**0.08**	0.06	0.06	0.07	0.06	0.07	0.06	0.06	0.06
	Mean difference	0.00	**0.00**	0.00	0.00	0.00	0.00	−0.01	0.00	−0.01	0.00
Transitional	Uncertainty	0.11	0.14	0.11	0.11	0.10	0.15	0.16	**0.18**	0.15	0.14
	Mean difference	0.04	−0.04	0.00	0.00	0.00	0.00	0.01	**−0.04**	0.03	0.00

TABLE 5. A summary of the regression analysis showing the partition of the variance between systematic and random components, and the percent contribution of each. Bold font indicates that the contribution of systematic error exceeded 20%.

	Flux	Sensible heat flux					Latent heat flux					Carbon dioxide flux				
	Statistic	Std dev			%		Std dev			%		Std dev			%	
EC station	Time period	T	S	R	S	R	T	S	R	S	R	T	S	R	S	R
1	Full	8.3	2.9	7.8	12	88	14	5.3	13	14	86	0.08	0.00	0.08	0	100
	Daytime	11	4.7	10	18	82	23	9.6	21	17	83	0.08	0.01	0.08	1	99
	Overnight	5.7	1.1	5.6	4	96	4.9	0.6	4.9	3	97	0.07	0.02	0.07	9	91
	Transitional	7.1	2.9	6.5	16	84	15	6.8	13	20	80	**0.10**	**0.05**	**0.09**	**24**	**76**
2	Full	8.2	1.3	8.1	3	97	12	4.0	12	11	89	0.08	0.02	0.08	3	97
	Daytime	11	2.4	11	4	96	18	7.2	16	16	84	0.09	0.01	0.09	2	98
	Overnight	7.5	2.9	6.9	15	85	6.4	2.4	6.0	13	87	0.07	0.01	0.07	2	98
	Transitional	7.0	1.6	6.8	5	95	**13**	**6.1**	**11.8**	**21**	**79**	**0.11**	**0.05**	**0.10**	**22**	**78**
3	Full	11	5.0	10	20	80	**33**	**20**	**26**	**37**	**63**	0.07	0.01	0.07	1	99
	Daytime	**19**	**9.1**	**17**	**23**	**77**	**38**	**18**	**33**	**23**	**77**	0.11	0.03	0.10	3	97
	Overnight	4.8	1.6	4.5	11	89	3.9	0.6	3.8	2	98	0.05	0.01	0.05	6	94
	Transitional	9.5	4.1	8.6	18	82	**17**	**9.6**	**14**	**31**	**69**	0.07	0.03	0.06	18	82
4	Full	8.7	1.6	8.6	4	96	15	2.8	15	3	97	0.06	0.01	0.06	5	95
	Daytime	10	9.1	4.1	17	83	23	4.5	23	4	96	0.09	0.04	0.09	14	86
	Overnight	6.9	2.4	6.5	12	88	4.8	1.3	4.6	7	93	0.05	0.02	0.05	12	88
	Transitional	**7.1**	**4.0**	**5.9**	**32**	**68**	11	0.9	11	1	99	0.09	0.01	0.08	2	98
5	Full	8.9	2.5	8.6	8	92	**14**	**9.2**	**11**	**40**	**60**	0.07	0.01	0.07	1	99
	Daytime	**14**	**6.7**	**12**	**24**	**76**	**30**	**19**	**24**	**38**	**62**	0.08	0.01	0.08	1	99
	Overnight	4.5	1.7	4.0	17	83	4.5	1.6	4.2	13	87	0.07	0.01	0.07	4	96
	Transitional	**8.4**	**5.0**	**6.8**	**36**	**64**	**10**	**6.3**	**8.2**	**37**	**63**	0.06	0.00	0.06	0	100
6	Full	9.7	1.1	9.6	1	99	29	12	26	18	82	0.10	0.01	0.10	1	99
	Daytime	14	3.3	13	6	94	**43**	**24**	**36**	**31**	**69**	0.13	0.03	0.13	10	90
	Overnight	6.0	1.7	5.3	20	80	8.4	3.4	7.7	16	84	0.05	0.01	0.05	2	98
	Transitional	13	1.8	13	2	98	**32**	**18**	**26**	**33**	**67**	0.15	0.05	0.14	9	91
AZ	Full	10	4.5	8.9	20	80	15	4.1	14	8	92	0.10	0.01	0.10	1	99
	Daytime	**13**	**7.5**	**10**	**35**	**65**	25	11	23	19	81	0.13	0.01	0.13	1	99
	Overnight	8.6	3.4	7.9	16	84	8.8	1.8	8.6	4	96	0.08	0.02	0.07	7	93
	Transitional	**12**	**7.4**	**9.1**	**37**	**63**	**30**	**15**	**26**	**26**	**74**	0.17	0.03	0.16	3	97
B5	Full	8.4	1.4	8.3	3	97	18	0.9	18	1	99	0.09	0.02	0.09	3	97
	Daytime	9.1	1.6	8.9	4	96	24	8.8	22	14	86	0.08	0.02	0.08	8	92
	Overnight	6.8	2.9	6.1	18	82	7.3	2.4	6.9	11	89	0.05	0.00	0.05	1	99
	Transitional	**9.8**	**4.5**	**8.7**	**21**	**79**	22	8.2	21	13	87	0.21	0.06	0.20	9	91
B6	Full	9.4	2.8	9.2	10	90	**16**	**8.0**	**14**	**25**	**74**	0.16	0.00	0.16	0	100
	Daytime	**9.2**	**5.4**	**7.5**	**34**	**66**	**22**	**12**	**19**	**28**	**72**	0.21	0.02	0.21	1	99
	Overnight	4.9	0.9	4.8	4	96	7.8	3.0	7.2	14	86	0.05	0.02	0.05	15	85
	Transitional	**5.6**	**3.1**	**4.6**	**32**	**68**	**13**	**5.8**	**11**	**21**	**79**	0.14	0.03	0.14	4	96

random components. Assuming that the reference flux (average of the measurements from all nine EC systems) closely approximates the true flux, the residuals of the linear regression reflect the errors associated with the measurements from the individual systems. That being the case, not only would the distribution of the residuals be indicative of the distribution of the measurement errors, σ_T would be indicative of the total measurement error while σ_S and σ_R would be indicative a bias and uncertainty, respectively.

Not altogether unexpectedly, the analysis yielded σ_T values (Table 5) that were similar to the uncertainty estimates of the HR method. The analysis showed that σ_S accounted for as much as 40% of the total error but typically contributed much less (Table 5). On average, the systematic error accounted for 18% of σ_T for H, 17% of σ_T for λE, and 6% of σ_T for F_c.

A further comparison of the distribution of residuals (Fig. 5) showed them to have expected Laplace distributions with the occasional outlier. These outliers consistently represented the same measurement periods. For example, in the case of H, two outlying points were evident; the first of these represented the period from 1800 to 1900 CST on doy 144, while the second represented the period from 1800 to 1900 CST on doy 146. In the case of λE, four points were outliers representing the periods from 1800 to 1900 CST on doy 144, 1800 to 2000 CST on doy 146, and 1500 to 1600 CST on doy 147 were found. In the case of F_c, three points were outliers representing the period from 1800 to 2100 CST

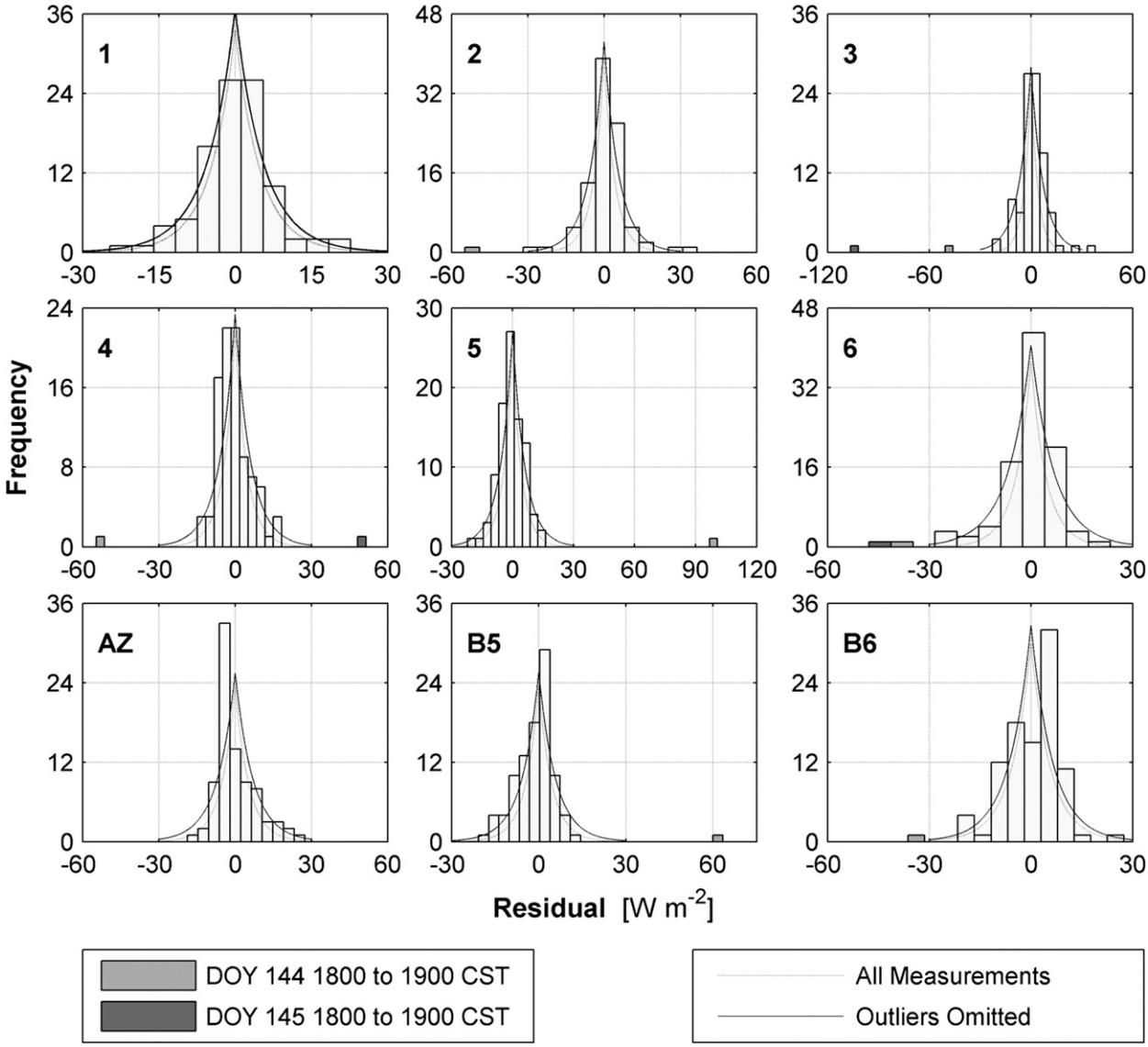

FIG. 5. The distribution of the residuals is shown for each of the nine EC systems along with it corresponding probability density function. The probability density function of the analysis when the outlying points were omitted is also shown.

on doy 146 were found. In all cases, the outliers occurred during the late afternoon and evening periods when advective conditions tended to be strongest. If the measurements corresponding to these time periods are omitted from the analysis, σ_T collapses to very nearly the same value (Table 6).

Given that these results strongly suggested that advective periods can impact and inflate the uncertainty estimates, the HR method was repeated omitting the measurements from those periods that occurred during advective periods. The revised estimates (Table 7) are much more consistent than were the initial uncertainty estimates calculated with the HR method. For example, the range in the uncertainty estimates for H for the full intercomparison period was reduced from 3.2 to 1 W m^{-2}. If only the transitional period were considered, the range in the uncertainty estimates for H was reduced from 5 to 2 W m^{-2}. Much more striking, however, was the impact of eliminating the measurements collected during strong advection on the uncertainty estimates of λE. In the case of the former, the range in the uncertainty estimates was reduced from 28 to 3.2 W m^{-2} when the full intercomparison period was considered, from 42 to 3.8 W m^{-2} when only the daytime period was considered, and from 14 to 9.0 W m^{-2} when only the transitional period was considered. Similarly, the range

TABLE 6. A summary of the σ_T of the residuals of the linear regression when the periods associated with the outliers were omitted.

Time period	EC station 1	2	3	4	5	6	AZ	B5	B6	Mean of all stations
					Sensible heat flux					
Full	7.0	7.0	6.6	6.3	6.6	7.1	7.9	6.9	7.5	7.0
Daytime	8.8	8.4	8.5	9.3	7.6	9.2	7.8	8.9	8.9	8.6
Transitional	6.3	7.3	7.7	6.9	6.9	7.3	7.7	7.7	7.8	7.3
					Latent heat flux					
Full	9.9	11	11	12	9.4	12	9.8	10	9.4	11
Daytime	15	16	15	16	15	16	16	16	14	16
Transitional	13	9.1	13	14	10	12	12	9.8	10	11
					Carbon dioxide flux					
Full	0.05	0.04	0.05	0.04	0.05	0.05	0.05	0.04	0.04	0.05
Daytime	0.04	0.04	0.04	0.03	0.04	0.04	0.05	0.04	0.04	0.04
Transitional	0.04	0.04	0.05	0.04	0.04	0.06	0.06	0.06	0.05	0.05

uncertainty estimates for F_c was reduced from 0.08 to 0.02 mg m^{-2} s^{-1} when only the measurements collected during the transitional period were considered.

With the exception of the estimates for the transitional period, which increased somewhat, the revised uncertainty estimates for H were little changed from the initial estimates. Similarly, the revised uncertainty estimates for F_c differ little from the initial estimates except during the transitional period where it decreased by half from 0.14 to 0.07 mg m^{-2} s^{-1}. In contrast, the uncertainty estimate for λE calculated using the daytime data decreased by 40%, or nearly 20 W m^{-2}. In terms of percent reduction, a similar decrease was seen for the error estimates during the transitional period, which decreased by 33% from 21 to 14 W m^{-2}. The uncertainty estimates for F_c decreased by 50% from 0.14 to 0.07 mg m^{-2} s^{-1} for the transitional period.

It is posited that the difference in the uncertainty estimates reflect the impacts of strongly advective conditions. The fundamental assumptions underlying the EC method are violated under these conditions since measured flux does not reflect only the vertical transport of heat and moisture, but may include a horizontal component as well. Although the spatial analysis did not confirm this, advection is likely to impact the EC systems differently depending on the path of the advected air

TABLE 7. A summary of the revise uncertainty estimates calculated for each of the turbulent fluxes using the method of Hollinger and Richardson (2005) and Richardson et al. (2006) is provided.

Time period	Statistic	EC station 1	2	3	4	5	6	AZ	B5	B6	Aggregate of all stations
						Sensible heat flux					
Full	Uncertainty	10	10	10	11	10	11	11	10	11	11
	Mean difference	0.7	−0.7	−0.2	−0.7	−0.5	−0.4	0.5	0.4	0.7	0
Daytime	Uncertainty	12	11	10	12	11	11	12	11	11	12
	Mean difference	1.0	−0.7	−0.3	−0.3	−0.7	0.8	0.2	−0.6	0.8	0
Transitional	Uncertainty	14	14	15	14	15	16	16	16	15	15
	Mean difference	0.5	0.4	−0.6	−0.4	0.0	−0.3	0.6	−0.7	0.5	0
						Latent heat flux					
Full	Uncertainty	18	18	21	18	18	20	19	18	18	19
	Mean difference	−0.6	0.5	0.8	0.6	−0.9	0.1	−0.4	−0.5	0.0	0.0
Daytime	Uncertainty	26	26	29	28	28	26	25	25	25	28
	Mean difference	−1.3	0.5	0.4	0.4	0.7	0.7	−0.9	−0.5	−0.2	0.0
Transitional	Uncertainty	14	11	14	13	13	20	13	11	11	14
	Mean difference	−0.5	−1.0	0.8	−0.5	0.3	0.8	0.0	−0.4	0.8	0.00
						Carbon dioxide flux					
Full	Uncertainty	0.09	0.09	0.08	0.08	0.09	0.08	0.09	0.08	0.09	0.09
	Mean difference	0.01	−0.01	−0.01	0.00	0.00	0.00	−0.01	0.01	−0.01	0.00
Daytime	Uncertainty	0.09	0.10	0.09	0.08	0.10	0.08	0.09	0.08	0.10	0.09
	Mean difference	0.02	−0.03	−0.01	0.01	0.01	0.01	−0.01	−0.02	0.01	0.00
Transitional	Uncertainty	0.06	0.06	0.08	0.06	0.07	0.08	0.08	0.06	0.07	0.07
	Mean difference	0.01	−0.01	0.01	0.01	0.00	0.00	−0.02	0.01	0.00	0.00

TABLE 8. A summary of the mean uncertainty estimates calculated for each of the turbulent fluxes using the modified form of the Mann and Lenschow method described by Hollinger and Richardson (2005) is provided.

Flux	EC station 1	2	3	4	5	6	AZ	B5	B6	Mean of all stations
H	13	12	13	12	13	15	12	14	15	13
λE	27	27	27	26	27	25	27	27	27	27
F_c	0.10	0.11	0.11	0.11	0.09	0.11	0.10	0.10	0.09	0.10

across the study site. For example, if the wind is southerly, it will have traveled over a greater portion of the field to reach the sites at the north end of the intercomparison, (e.g., stations 1 and 3) than those stations farther to the south. As a result the air parcel would be in greater equilibrium with the local surface conditions at the northern sites than those to the south.

f. Assessment of uncertainty using the method of Mann and Lenschow

To further evaluate the revised uncertainty estimates calculated using the HR method, a second analysis was conducted using the ML method. As can be seen from the summary of results in Table 8, the uncertainty estimates using this technique tended to agree closely among the nine EC systems. For example, the range in the uncertainty estimates for H was less than 3 W m^{-2} such that the mean uncertainty estimate was 13 W m^{-2}. This value agrees quite closely with the result using the HR method (12 W m^{-2}). The results of the ML method for λE (27 W m^{-2}) and F_c (0.10 mg m^{-2} s^{-1}) were also remarkably similar to the revised estimates using the HR method. As a result, the ML method not only reaffirms the revised values from the HR method, it also confirms the underlying assumptions of that technique—the measurement uncertainty from the nine EC stations shares the same distribution.

5. Conclusions

The result of this analysis showed the nine EC systems used during BEAREX08 tended to agree strongly with one another. They also show that the uncertainty associated with these systems were quite similar during quasi-normal (nonadvective) conditions. Under strongly advective conditions, however, the uncertainty estimates were typically both much higher and different from station to station. This indicates that the advection introduces significant uncertainty. In the case of F_c, the uncertainty during advective conditions was 50% greater than during quasi-normal conditions. Furthermore, since the revised HR uncertainty estimates, which were limited to quasi-normal conditions, tended to vary with time of day and the magnitude of the flux, it is likely that the amount of additional uncertainty introduced by advective conditions would also vary. While this could not be confirmed with the data that was collected during the intercomparison period, it is a reasonable hypothesis that could be tested in a future study.

More broadly, the results of this study underscore the need to be cognizant of ambient conditions when working with EC flux measurements, both when determining the measurement uncertainty and when using the measurement in subsequent analyses. While this study suggests that the nine EC systems responded in a similar fashion during quasi-normal conditions, it clearly demonstrated that strongly advective conditions can result in greater and more variable measurement uncertainty. Thus, while a single uncertainty assessment conducted as a part of an intercomparison study such as this one provides valuable baseline information, it might not capture the true uncertainty of the flux measurements under the full range of ambient conditions. To more fully understand the degree of uncertainty in a flux dataset, an ongoing assessment would be highly beneficial and should be conducted as an integral part of the postprocessing of eddy covariance data. This is particularly true for regional or global networks where local conditions can vary significantly from site to site. While it is not practical to deploy multiple EC systems at each site in order to use the HR method, the ML method provided equivalent uncertainty estimates and can be calculated from the high frequency data of a single tower. While this method is limited to the daytime period only, it could prove to be a useful technique to characterize the variability in the measurement uncertainty over time and changing environmental conditions.

Acknowledgments. This project would not have been possible without the cooperation and assistance of the scientists and staff at the USDA, ARS, Conservation and Production Laboratory, Soil and Water Management Research Unit. The authors would also like to thank the three anonymous reviewers for their insightful comments.

REFERENCES

Alfieri, J. G., P. D. Blanken, D. N. Yates, and K. Steffen, 2007: Variability in the relative influence of the environmental drivers of evapotranspiration over heavily grazed pasture during drought. *J. Hydrometeor.,* **8,** 207–220.

——, ——, D. Smith, and J. Morgan, 2009a: Concerning the measurement and magnitude of heat, water vapor, and carbon dioxide exchange from a semiarid grassland. *J. Appl. Meteor. Climatol.,* **48,** 982–996.

——, D. Niyogi, H. Zhang, M. A. LeMone, and F. Chen, 2009b: Quantifying the spatial variability of surface fluxes using data from the 2002 International H2O Project. *Bound.-Layer Meteor.,* **133,** 323–341.

——, X. Xiao, D. Niyogi, R. A. Pielke Sr., F. Chen, and M. A. LeMone, 2009c: Satellite-based modeling of transpiration from the grasslands in the Southern Great Plains, USA. *Global Planet. Change,* **67,** 78–86.

Anderson, M. C., J. M. Norman, W. P. Kustas, F. Li, J. H. Prueger, and J. R. Mecikalski, 2005: Effects of vegetation clumping on two-source model estimates of surface energy fluxes from an agricultural landscape during SMACEX. *J. Hydrometeor.,* **6,** 892–909.

Baldocchi, D. D., 2003: Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: Past, present, and future. *Global Change Biol.,* **9,** 479–492.

——, 2008: Breathing of the terrestrial biosphere: Lessons learned from a global network of carbon dioxide flux measurement systems. *Aust. J. Bot.,* **56,** 1–26.

——, and Coauthors, 2001: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Amer. Meteor. Soc.,* **82,** 2415–2434.

Barnhart, H. X., M. Harber, and J. Song, 2002: Overall concordance coefficient for evaluating agreement among multiple observers. *Biometrics,* **58,** 1020–1027.

Billesbach, D. P., 2011: Estimating uncertainties in individual eddy-covariance flux measurements: A comparison of methods and a proposed new method. *Agric. For. Meteor.,* **151,** 394–405.

Blanken, P. D., and Coauthors, 1998: Turbulent flux measurements above and below the overstory of a boreal aspen forest. *Bound.-Layer Meteor.,* **89,** 109–140.

——, M. W. Williams, S. P. Burns, R. K. Monson, J. Knowles, K. Chowanski, and T. Ackerman, 2009: A comparison of water and carbon dioxide exchange at a windy alpine tundra and subalpine forest site near Niwot Ridge, Colorado. *Biogeochem.,* **95,** 61–76.

Burt, J. E., and G. M. Barber, 1996: *Elementary Statistics for Geographers*. Guilford Press, 640 pp.

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

Chasmer, L., N. Kljun, A. Barr, A. Black, C. Hopkinson, H. McCaughey, and P. Treitz, 2008: Influences of vegetation structure and elevation on CO_2 uptake in a mature jack pine forest in Saskatchewan, Canada. *Can. J. For. Res.,* **38,** 2746–2761.

Chen, F., and Coauthors, 2007: Description and evaluation of the characteristics of the NCAR high-resolution land data assimilation system. *J. Appl. Meteor. Climatol.,* **46,** 694–713.

Cosh, M. H., T. J. Jackson, R. Bindish, and J. H. Prueger, 2004: Watershed scale temporal and spatial stability and its role in evaluating satellite estimates. *Remote Sens. Environ.,* **92,** 427–435.

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

French, A. N., and Coauthors, 2005: Surface energy fluxes with the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER) at the Iowa 2002 SMACEX site (USA). *Remote Sens. Environ.,* **99,** 55–65.

Goulden, M. L., and Coauthors, 1996: Measurements of carbon storage by long-term eddy correlation: methods and a critical assessment of accuracy. *Global Change Biol.,* **2,** 169–182.

Gowda, P. H., J. L. Chavez, P. D. Colaizzi, S. R. Evett, T. A. Howell, and J. A. Tolk, 2008: ET mapping for agricultural water management: Present status and challenges. *Irrig. Sci.,* **26,** 223–237.

Højstrup, J., 1993: A statistical data screening procedure. *Meas. Sci. Technol.,* **4,** 153–157.

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

Isaaks, E. H., and R. M. Srivastava, 1989: *Applied Geostatistics*. Cambridge University Press, 561 pp.

Kaimal, J. C., and J. J. Finnigan, 1994: *Atmospheric Boundary Layer Flows*. Oxford University Press, 304 pp.

Katul, G. G., and Coauthors, 1999: Spatial variability of turbulent fluxes in the roughness sublayer of an even-aged pine forest. *Bound.-Layer Meteor.,* **93,** 1–28.

Ko, J., and G. Piccinni, 2009: Corn yield responses under crop evapotranspiration-based irrigation management. *Agric. Water Manag.,* **96,** 799–808.

Kustas, W. P., J. L. Hatfield, and J. H. Prueger, 2005: The Soil Moisture–Atmosphere Coupling Experiment (SMACEX): Background, hydrometeorological conditions, and preliminary findings. *J. Hydrometeor.,* **6,** 791–804.

LeMone, M. A., and Coauthors, 2002: CASES-97: Late-morning warming and moistening of the convective boundary layer over the Walnut River watershed. *Bound.-Layer Meteor.,* **104,** 1–52.

——, F. Chen, J. A. Alfieri, R. Cuenca, D. Niyogi, S. Kang, K. Davis, and P. D. Blanken, 2007: Surface, soil, and vegetation network during the International H2O Project 2002 field campaign. *Bull. Amer. Meteor. Soc.,* **88,** 65–81.

——, M. Tewari, F. Chen, J. G. Alfieri, and D. Niyogi, 2008: Evaluation of the Noah land surface model using data from a fair-weather IHOP_2002 day with heterogeneous surface fluxes. *Mon. Wea. Rev.,* **136,** 4915–4941.

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

Leuning, R., 2004: Measurements of trace gas fluxes in the atmosphere using eddy covariance: WPL corrections revisited. *Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis,* X. Lee, W. Massman, B. Law, Eds., Kluwer Academic, 119–132.

Lin, L., 1989: A concordance correlation coefficient to evaluate reproducibility. *Biometrics,* **45,** 255–268.

——, 1992: Assay validation using the concordance correlation coefficient. *Biometrics,* **48,** 599–604.

——, A. S. Hedayat, B. Sinha, and M. Yang, 2002: Statistical methods in assessing agreement: Models, issues, and tools. *J. Amer. Stat. Assoc.,* **97,** 257–270.

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

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

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

——, and X. Lee, 2002: Eddy covariance flux corrections and uncertainties in long term studies of carbon and energy exchanges. *Agric. For. Meteor.,* **113,** 121–144.

Meek, D. W., J. H. Prueger, W. P. Kustas, and J. L. Hatfield, 2005: Determining meaningful differences for SMACEX eddy covariance measurements. *J. Hydrometeor.,* **6,** 812–824.

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

Myklebust, M. C., L. E. Hipps, and R. J. Ryel, 2008: Comparison of eddy covariance, chamber, and gradient methods of measuring soil CO2 efflux in an annual semi-arid grass, *Bromus tectorum. Agric. For. Meteor.,* **148,** 1894–1907.

Neale, C. M. U., H. Jayanthi, and J. L. Wright, 2005: Irrigation water management using high resolution airborne remote sensing. *Irrig. Drain. Syst.,* **19,** 321–336.

Novick, K. A., P. C. Stoy, G. G. Katul, D. S. Ellsworth, M. B. S. Siqueira, J. Juang, and R. Oren, 2004: Carbon dioxide and water vapor exchange in a warm temperate grassland. *Oecologia,* **138,** 259–274.

Prueger, J. H., and Coauthors, 2005: Tower and aircraft eddy covariance measurements of water vapor, energy, and carbon dioxide flues during SMACEX. *J. Hydrometeor.,* **6,** 954–960.

Richardson, A. D., and Coauthors, 2006: A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes. *Agric. For. Meteor.,* **136,** 1–18.

Schabenberger, O., and C. A. Gotway, 2005: *Statistical Methods for Spatial Data Analysis*. Chapman & Hall, 516 pp.

Schotanus, E. K., F. T. M. Nieuwstadt, and H. A. R. de Bruin, 1983: Temperature measurement with a sonic anemometer and its application to heat and moisture flux. *Bound.-Layer Meteor.,* **26,** 81–93.

Schuepp, P. H., M. Y. Leclerc, J. I. MacPherson, and R. L. Desjardins, 1990: Footprint prediction of scalar fluxes from analytical solutions of the diffusion equation. *Bound.-Layer Meteor.,* **50,** 355–373.

Steduto, P., T. C. Hsiao, and E. Fereres, 2007: On the conservative behavior of biomass water productivity. *Irrig. Sci.,* **25,** 189–207.

Svejcar, T., and Coauthors, 2008: Carbon fluxes on North American rangelands. *Rangeland Ecol. Manag.,* **61,** 465–474.

Tanner, C. B., and G. W. Thurtell, 1969: Anemoclinometer measurements of Reynolds stress and heat transport in the atmospheric surface layer. Research and Development Tech. Rep. to US Army Electronic Command, ECOM 66-G22-F, Department of Soil Sciences, University of Wisconsin, Madison, Wisconsin, 199 pp.

Webb, E. K., G. L. Pearman, and R. Leuning, 1980: Correction of flux measurements foe density effects due to heat and water vapour transfer. *Quart. J. Roy. Meteor. Soc.,* **106,** 85–100.

Willmott, C. J., 1982: Some comments on the evaluation of model performance. *Bull. Amer. Meteor. Soc.,* **63,** 1309–1313.

Yu, G.-R., X.-F. Wen, X.-M. Sun, B. D. Tanner, X. Lee, and J.-Y. Chen, 2006: Overview of ChinaFLUX and evaluation of its eddy covariance measurements. *Agric. For. Meteor.,* **137,** 125–137.