Final Technical Report: Interagency Agreement DE-IA02-96ER62274

The AmeriFlux Network of long-term CO₂ flux measurement stations: Methodology and Intercomparability

D. Y. Hollinger,¹ and R.. S. Evans¹

DOE Patent Clearance Granted

Mark P. Dvorscak
(630) 252-2393

Date

E-mail: mark.dvorscak@ch.doe.gov
Office of Intellectual Property Law
DOE Chicago Operations Office

¹ USDA Forest Service, 271 Mast Rd, Durham, NH 03824, USA

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Abstract.

A portable flux measurement system has been used within the AmeriFlux network of CO₂ flux measurement stations to enhance the comparability of data collected across the network. This system is used to assess environmental measurements as well as sensible heat (H), latent energy (LE) and CO₂ fluxes. We report results of comparisons made at 17 AmeriFlux sites as well as 3 FLUXNET sites (2 in Europe and 1 in New Zealand) and one Agricultural Research Service pasture site.

No systematic biases were observed in a comparison between portable system and site H , LE or CO₂ flux values although there were biases observed between the portable system and individual site temperature and photosynthetically active photon flux (PPFD) values. Not surprisingly, agreements between the portable system and site systems were better against systems that used the same instruments as the roving system. An analysis of results suggests that if values from two stations exceed 26% for H , 35% for LE and 32% for CO₂ flux they are likely to be significant. Methods for improving the intercomparability of the network are also discussed.

Keywords: eddy covariance, carbon exchange, long-term measurements.

Introduction.

AmeriFlux was formed in 1996 when 11 operating carbon dioxide flux measurement stations combined to form a network for coordinating research and comparing observations across a range of climates and

Fig. 1. Location of AmeriFlux sites (January, 2001).

ecosystems. The network now consists of more than 40 stations that stretch from the tundra of Alaska to the Cerrado of Brazil (Fig. 1 and Table 1). The stations represent most of climate zones in the Americas (Fig. 2). Twenty-one of the operational sites are forested, two are shrublands, six are grasslands, two are bogs or wetlands, three are croplands, and three are located in tundra. Several sites have ceased operation. Network, site details, and data can be found at <http://public.ornl.gov/ameriflux/>.

One of the central goals of AmeriFlux has been to ensure that results from the various stations can be readily compared.

Because all of the original stations did not use the same brand of equipment and employed a diversity of software for reducing data and calculating fluxes, it was decided that the best way to insure comparability of results was to establish a portable reference system that would travel to the various sites to assess the quality of the measurements. In practice, a research technician travels with the reference system to a site, installs the flux and meteorological instrumentation alongside that presently operating at the site and makes measurements for 3-5 days. Site and calibration system data are then statistically compared, and if any discrepancies are noted, the AmeriFlux and reference system researchers work together to resolve problems. In this way subtle, and rarely, gross errors in sensor or flux

Fig. 2. Climate zones of AmeriFlux network sites.

system operation have been detected and corrected. Examples of subtle errors in micrometeorological sensor operation detected by this methodology include incorrect measurements of air temperature caused by blocked aspirator fans and hysteresis in photosynthetically active photon flux density (PPFD) readings caused by tilted sensors. Some errors in flux readings have been traced to incorrect calculations or analyzer calibration problems. The portable system approach has thus been of value not only in establishing the intercomparability of stations, but also in improving it.

The approach of the AmeriFlux network is to focus on intercomparability and traceability to standards rather than prescribed instrumentation or procedures. This philosophy is distinct from that of EUROFLUX where all stations used identical sonic anemometers and gas analyzers and similar suites of processing software (Aubinet et al. 2000). The AmeriFlux approach was partly a pragmatic response to the diversity of equipment in use at various sites when AmeriFlux was established and partly an acknowledgement that a range of methodologies and instrumentation might be more suitable given the diversity of sites.

The AmeriFlux portable flux system was constructed in early 1997 and by the end of 2000 had been to 29 sites, including 2 sites in the EUROFLUX network and one site each in Australia and New Zealand. A comparison was also made with a bowen-ratio CO₂ flux system near Temple, Texas. Several of the comparisons failed because of damage to the roving system in transit or unsuitable weather during the visit. In other instances we await data analysis by the site team. We discuss here results based on comparisons made at 21 of these sites.

Methods

A useful AmeriFlux reference system had to meet the following criteria; (1) the ability to provide accurate measurements (precision and accuracy should exceed that of the permanent stations), (2) long-term stability, (3) rugged and lightweight to facilitate transport, (4) quick and flexible installation and removal. We describe below the measurement system built to meet these objectives. It consists of accurately calibrated meteorological sensors (net radiation, PPFD, air temperature), sonic anemometer, closed-path infrared gas analyzer (and associated pump), data acquisition system, and portable computer.

Portable measurement system environmental sensors- Net radiation is measured with a Radiation and Energy Balance Systems (REBS, Seattle, WA, USA) model Q*7.1 net radiometer and recorded by a Campbell Scientific model 21X datalogger (Campbell Scientific, Logan, UT, USA). We employ a REBS model RV2 ventilator to reduce errors associated with cooling of the domes by variable wind speeds. Calibration of the field instrument is verified by comparison to an identical laboratory instrument several times a year over short grass. Although there is no international standard for net radiometers, based on calculations using traceable pyranometers, REBS suggests that the accuracy of the Q*7.1 net radiometer instrument is better than 5%. According to the manufacturer, the intercomparability of 2 of these instruments when both are in good condition is ~1%. In our comparisons between the laboratory and field instrument, agreement is typically within 0.5%.

Photosynthetically active photon flux density (PPFD) is measured with a model LI-190SA (LiCor, Inc., Lincoln, NB, USA) quantum sensor and model 2290 milivolt adapter to convert the current output to a voltage. This PPFD sensor is compared several times annually against two identical quantum sensors stored in our laboratory, and recalibrated if necessary to retain agreement. For these comparisons, the field and reference sensors are leveled and placed side-by-side on a frame outdoors on a cloud-free day. The current output of the 2 reference sensors is read sequentially by a high precision digital multimeter (model 2001, Keithley Instruments, Cleveland, OH, USA) and compared to the output from the field sensor recorded by a 21X datalogger. The agreement of the laboratory and field sensors at midday in the summer after a calibration is typically better than ~0.2%. However, LiCor specifies the absolute accuracy of the Li-190SA as 5%, based on traceability to a standard light source.

Air temperature is measured with a 100 ohm platinum resistance thermometer (model RTD-810, Omega, Inc., Stamford, CT, USA) used in a 4-wire configuration with a 0-5V linearizing module (model OM5-IP4-100C, Omega) and read by a Campbell model 21X datalogger. The sensor is compared with an identical sensor that together with the laboratory grade Keithley multimeter is calibrated annually at a standards laboratory. Drift and accuracy in the field sensor are assessed by comparison at the ice (0.00 °C) and gallium points (29.7646 °C) in our laboratory. These comparisons indicate that the absolute

accuracy of the field temperature measurement is better than 0.05 °C over the 0-30 °C range. The temperature sensor is mounted in a model 076B-1 (Met One) aspirated radiation shield.

Portable Flux System - Comparisons between sites are carried out by a portable system that measures carbon dioxide, water vapor, heat, and momentum fluxes using the eddy covariance technique (Baldocchi *et al.* 1988). The reference flux system is based on model SATI/3K (Kaimal probe with the electronics in the support arm) 3-axis sonic anemometer (Applied Technologies, Inc., Boulder, CO, USA) and model LI-6262 fast response CO₂/H₂O infrared gas analyzers (LiCor, Inc.). Data from the sonic anemometer are read at 10 Hz through an RS-232 port on a laptop computer. The raw analog outputs of the gas analyzer (CO₂, H₂O, cell temperature, and cell pressure) are recorded at 10 Hz using a model 200 DAQBOOK data acquisition system with model DBK18 low-pass filter card (Omega, Inc.) interfaced to the computer via the parallel port. The data acquisition system is configured for 16 bits over the 0-4 volts to maximize resolution (1 bit resolution ~ 0.01 μmol CO₂) and the low-pass filter card configured as 3-pole Butterworth filters with a corner frequency set at 30 Hz to reduce AC line noise. Data are stored on the computer internal hard drive and archived on writeable CD's.

Air is ducted to the infrared gas analyzer through an inlet funnel/filter (0.45 micron pore size) combination, either 30 or 60 m (depending on tower height) of 3.2 mm diameter high-density polyethylene tubing, 1 m of water-jacketed stainless steel tubing, a 1.0 micron gelman filter, and the infrared gas analyzer analysis cell at a mass flow rate of 6.5 l/min, regulated by a mass flow controller (model AFC2600, Aalborg, Orangeburg, NY, USA). Air pressure in the analyzer is typically about 85 KPa. A GAST model 1531-107B-G557X rotary vane vacuum pump (Ohlheiser, West Hartford, CT, USA) and 2-l ballast are used to minimize pressure fluctuations in the airstream. The gas analyzer, water jacket, and flow controller are mounted in an aluminum case where the temperature is controlled at 30 °C.

The 5th order polynomial relating gas analyzer CO₂ channel output voltage to CO₂ concentration was replaced with a new polynomial constructed with a suite of 6 precision CO₂ standards (NOAA CMDL, Boulder, CO, USA) spanning the range of 350 – 500 μmol mol⁻¹ CO₂. Before each field visit, the gas

analyzer was calibrated against a primary standard in the laboratory of $375.37 \mu\text{mol mol}^{-1}$. This calibration was carried out at a flow rate and analysis cell pressure that mimics field conditions (e.g. 6.5 l min^{-1} and 85 kPa). Comparison to other primary CO_2 standards under these conditions suggests that the absolute accuracy of the analyzer measurements is $\sim 0.2 \mu\text{mol mol}^{-1}$ before each field visit. The analyzer CO_2 calibration accuracy is checked after transport to each field site with 2 secondary standard gases shipped to each site prior to the visit. These secondary standards are commercially prepared aluminum cylinders that have been compared against our primary standards. Because of this secondary comparison, we estimate that our on-site gas analyzer CO_2 accuracy is $\sim 0.5 \mu\text{mol CO}_2 \text{ mol}^{-1} \text{ air}$. In the laboratory, the analyzer water vapor calibration is carried out at 2 points using dry air and air maintained at a dewpoint of 15°C by a LiCor model Li-610 water vapor generator. This calibration is carried out using a flow rate of $< 2 \text{ l min}^{-1}$ and at ambient pressure. Calibration is checked by comparing the water vapor concentration of the air in our laboratory drawn through the gas analyzer under “field” conditions (6.5 l min^{-1} and 85 kPa) against that recorded by a model M4 dew point hygrometer with model D-2 chilled mirror (General Eastern, Woburn, MA, USA) dewpoint hygrometer. Over a 10-minute period, the ventilation system of our building typically maintains the dewpoint of laboratory air to within the 0.1°C dewpoint accuracy of the hygrometer. Based on our calibrations, we expect an accuracy of $\sim 0.2 \text{ mmol H}_2\text{O mol}^{-1} \text{ air}$ in our water vapor measurements.

Eddy Flux Calculations – Massman and Lee (2001) have recently reviewed the derivation of flux calculations and discussed corrections needed to account for imperfections in the measurement system. Here we simplify the continuity equation to a one-dimension situation and focus on comparing vertical wind-scalar covariances between the portable and *in situ* systems. Portable system covariances are computed with a running-mean removal technique based on a digital recursive filter with a time constant of 600 s (McMillen 1988). Two-angle coordinate rotations are performed on the wind data in the covariance calculations to remove the effects of instrument tilt or terrain irregularity on the airflow. Variations in air density (e.g. Webb *et al.* 1980) are removed by the 1 m water-jacketed stainless steel

portion of the air inlet line and via software using the simultaneous measurement of water vapor concentration. Covariances between CO₂ and the vertical wind velocity (w) and between H₂O and w are calculated by digitally lagging the w signal to account for the transit time of the air sample down the inlet tube (McMillen 1988). Calculations are made for lags bracketing the expected lag (based on flow/volume calculations) over each one-half hour, and the lag providing the maximum lag-covariance used in subsequent flux calculations. This approach fails when CO₂ or H₂O flux rates approach zero (evenings and early morning for CO₂, late night for H₂O) and lags from nearby suitable periods are used at these times. The timing of the maximum in the lagged covariances typically varies by < 2%.

The sensible heat flux (H , W m⁻²) is calculated as

$$H = \frac{M_a C_p}{V} \cdot (\overline{w'T'} - (3.2 \cdot 10^{-4} \cdot T_K \cdot \overline{w'q'})) \quad (1)$$

where M_a is the molecular weight of dry air (g mol⁻¹), C_p is the specific heat of dry air (J °K⁻¹ g⁻¹), V is the molar volume of air at the ambient temperature and pressure (m³ mol⁻¹), $\overline{w'T'}$ is the covariance between the vertical wind velocity and air temperature estimated from the sonic temperature (m s⁻¹ °K⁻¹), $\overline{w'q'}$ is the covariance between the vertical wind velocity and the atmospheric water vapor concentration (m s⁻¹ mol H₂O mol⁻¹ air), and T_K is the actual air temperature estimated from the sonic temperature according to Kaimal and Gaynor (1991).

The latent heat flux (LE , W m⁻²) is calculated as

$$LE = \lambda \frac{M_w \cdot \overline{w'q'}}{V} \quad (2)$$

where λ is the latent heat of vaporization (J g⁻¹), M_w is the molecular mass of water (g mol⁻¹), V is the molar volume of air (m³ mol⁻¹), and $\overline{w'q'}$ is the covariance between the vertical wind velocity and atmospheric water vapor content (m s⁻¹ mol H₂O mol⁻¹ air).

The CO₂ eddy flux (F , μmol CO₂ m⁻² s⁻¹) is calculated similarly,

$$F = \frac{\overline{w'c'}}{V} \quad (3)$$

where $\overline{w'c'}$ is the covariance between the vertical wind and atmospheric carbon dioxide content ($\text{m s}^{-1} \mu\text{mol CO}_2 \text{ mol}^{-1} \text{ air}$) and V is as before.

Eddy covariance systems underestimate fluxes because of insensitivity to low and high frequency components of the signal. High frequency losses result from a variety of factors including line averaging, separation distance between the sonic anemometer and scalar instrument, and, for closed-path systems, attenuation by the inlet tubing (Horst 1997, 2000; Massman 2000, 2001). For closed-path analyzers, the amount of high-frequency attenuation depends on tube length and diameter, and flow rate of air. Low frequency losses are caused by software detrending of the input signals (linear or high-pass recursive filters) and the calculation of fluxes over a finite time period (block averaging). The peak of the flux or cospectral energy is shifted to higher or lower frequencies depending upon windspeed (u), height (z), and stability (L) (Kaimal 1972). For the unstable (generally daytime) case, $f_x = \bar{u} n_x / z$ where f_x is the cospectral peak, and n_x is the normalized peak frequency. Horst (1997) suggests $n_x \approx 0.085$. This relationship indicates that loss of high frequency variation is of greatest concern for flux systems close to the ground ($< 10\text{m}$), especially at higher windspeeds. Conversely, for systems mounted above $\sim 30\text{m}$, low frequency losses generally predominate, especially at low windspeeds. For typical close-path flux systems, high frequency losses in configurations mounted within a few meters of the surface, or low frequency losses in systems on towers taller than 50m , can exceed 30%. Because of better high-frequency response, open-path IRGAs should be used with systems mounted within 5m of the ground.

Roving system CO_2 and water fluxes were corrected for high frequency losses following the procedure of Goulden *et al.* (1997). This involves online calculation of sensible heat fluxes calculated with sonic anemometer temperature signals that have been digitally filtered to match the frequency response of the CO_2 and H_2O channels. Specifically, half-hourly carbon dioxide and water vapor fluxes are multiplied by H/H_{f_l} where H is the heat flux calculated with the unfiltered sonic temperature and H_{f_l}

and H_{f2} are heat fluxes calculated with temperature signals filtered to match the high frequency cutoff of the IRGA channels. This approach has the advantage of not assuming any particular spectral shape but does assume cospectral similarity between the fluxes. To correct for low frequency losses in heat, water vapor, and CO₂ fluxes resulting from a running mean filter (600 s) and half-hourly block averaging, we used the Horst/Massman approach of calculating a transfer function based on stability and theoretical spectra (e.g. Horst 1997, 2000; Massman 2000, 2001; see also Moore 1986). The equations in Massman (2001, Table 1) were used but with p , the term accounting for high frequency corrections, set to zero.

Most sites did not provide spectrally-corrected data so we used the Massman (2001) approach when necessary.

Flux values in our data were excluded from further analysis if the wind speed was below 0.5 m s^{-1} , when sensor variance was high (Hollinger *et al.* 1995), or rain or snow was falling.

Statistical Comparisons – Pearson product moment correlation coefficients were calculated to assess the agreement between portable system and on-site sensors and fluxes. The slope and intercept of the relationships between the portable system (functioning as the “independent” variables) and on-site sensors and fluxes were calculated with reduced major axis (RMA) regressions (Hammer *et al.* 2001) because errors exist in measurements from both systems.

Results

Meteorological Sensors

Net radiation – The agreement between net radiation values measured by sensors at the permanent sites and the net radiometer of the roving system was very good. At an equivalent net radiation input of 500 W m^{-2} , the average difference was $\sim -3.5 \pm 4.9\%$ (mean and standard deviation, $n=20$). The differences were normally distributed (Fig. 3a), with all but one site agreeing to within 10% of the roving system.

Fig. 3. Percentage difference (except air temperature is absolute difference in $^\circ\text{C}$) between the portable system and site environmental measurements. (a) Net radiation, (b) photosynthetically active photon flux density, (c) air temperature.

Photosynthetically active Photon Flux Density (PPFD) – At all but one site, PPFD values recorded by the *in situ* sensors were less than those recorded by the roving system. On average, the difference between the roving and permanent sites (normalized to $2000 \mu\text{mol m}^{-2} \text{ s}^{-1}$) was $\sim 8.1 \pm 11.6 \%$ ($n=18$) and this difference is significantly different from zero. One interpretation of this difference is that the calibration of the roving sensor is incorrect. However, we have noticed a slow and steady decline in the output of PPFD sensors left exposed to the elements over many months, but never an increase in output. Because our sensor is checked against sensors that are stored in the laboratory, we believe that the observed average difference (Fig. 3b) is indicative of the aging effect in the field sensor population. This effect plus the significant outliers observed at 4 sites suggest that all environmental researchers should check their PPFD sensor calibration annually. This is an important consideration because

the errors observed here indicate that in some cases there will be large errors in coefficient values of models that fit CO_2 flux to PPFD.

Temperature – Comparisons between the roving system and the field sites indicate that field temperature measurements need to be improved. When normalized to 20°C , temperature errors of $>0.5^\circ \text{C}$ are found at one-half of the sites, with temperatures at 2 sites off by more than 3°C (Fig. 3c). An examination of the regressions between roving and site temperature measurements indicate that generally slopes are very close to 1 with low scatter (r^2 values usually in excess of 0.99) but that intercepts are often

significantly different from zero and skewed to positive values. Some of the sites rely on passive ventilation of their temperature sensors. Sensors in non-ventilated housings often record excess temperatures under sunny, low wind conditions.

Eddy Fluxes

Software comparison – We tested our flux calculation algorithms with data from the EUROFLUX “goldfiles” (Aubinet et al. 2000). We concatenated and time stamped the 19 individual half-hourly raw data files into one file (available on the AmeriFlux web site), calculated fluxes for H , LE , and CO_2 , and used linear regression analysis to compare our results to the mean EUROFLUX values obtained by sites using a running mean (recursive filter). We calculated our results prior to obtaining the mean EUROFLUX values. In our regression analyses the intercepts were never significantly different from zero, so all regressions were forced through zero and the comparisons made on the slopes. We also compare our values to those obtained by the Harvard group (Munger, pers. comm.). Slopes of the regressions with the EUROFLUX or Harvard values as the independent variable are shown in Table 2. In general, values from the AmeriFlux system were a few percent higher than the average from the EUROFLUX sites or the Harvard group. An exception was the LE values from the Harvard group, which were larger than either roving system or EUROFLUX results. In most cases, the differences were not significant (e.g. regression slopes were not significantly different from 1). Exceptions where fluxes calculated from the “goldfiles” with the AmeriFlux software were on average significantly greater ($P < 0.05$) than those obtained by the other groups were the comparison with the EUROFLUX LE and Harvard CO_2 flux (Table 2). We believe these slight differences result from the longer time constant (600 s) we use in our recursive mean filter that leads to less low frequency loss (Moore 1986, Massman 2000, Massman and Lee 2001).

Fig. 4. Mean percent differences between portable system and site measurements of sensible heat and CO₂ flux.

Field Comparisons - The mean differences

between the roving system and field site values for sensible heat flux (normalized to 500 W m^{-2}) and CO₂ flux (normalized to $-10 \mu\text{mol m}^{-2} \text{ s}^{-1}$) are shown in Fig. 4. On average, deviations of the H and CO₂ fluxes from roving system values were not significantly correlated. However, sites that had the greatest absolute deviation from the roving system in CO₂ flux also tended to have the largest deviations in water vapor flux ($r=0.46$, $P<0.05$). This is understandable since both measurements are typically

made by the same instrument at most sites and are thus susceptible to some common error modes (e.g. drift or noise in the analyzer).

Energy balance closure reported by the 2 sites showing the largest sensible heat deviations from the roving system was poor ($< 50\%$) indicating serious problems with the measurements or subsequent data processing. These issues could not be resolved during or after the site visits so these sites (points '1' and '2' in Fig. 4) were excluded from further analysis. Similarly, there were known problems with the infrared gas analyzer (lack of stability) at the site with the largest deviation in CO₂ flux (point '3' in Fig. 4) and data from this site were also excluded from the analysis that follows. This sort of diagnosis of instrumentation problems and exclusion of suspect data is part of the analysis protocol of every site.

Sensible Heat Flux – There was good agreement between sensible heat flux values measured by the roving system and the visited tower sites (Fig. 5a). The mean difference in normalized heat fluxes of $0.05 \pm 9.7\%$ (mean and standard deviation, $n=17$) between the roving system and site instrumentation was not significantly different from zero. The differences between the roving system and site measurements were normally distributed. Surprisingly, we found slight but apparently significant differences in the agreement between roving system and site instrumentation depending on the sonic anemometer used

(Table 3). A one-way analysis of variance suggests that sites using Campbell CSAT-3 sonic anemometers reported slightly lower values of sensible heat flux ($F(2,13)=5.7$, $P<0.05$) than sites using Gill or ATI instrumentation. After observing this result, we carried out a comparison between the roving system and a near-new CSAT-3 from our own laboratory. The instruments were mounted about 0.5 m apart and pointed into the wind at a height of 29 m on a boom above a walk-up tower at the Howland AmeriFlux site for 6 days during July 2000. For this comparison, data from the CSAT and roving system ATI sonic anemometers were recorded by the same data acquisition system and processed in the same way. The wT covariances calculated from the CSAT were slightly less than those recorded from the ATI sonic ($-3.5\%\pm 1.0\%$, mean and 95% confidence level for the first 3 days when the wind was such that the CSAT was slightly upwind of the ATI and $-4.1\%\pm 2.2\%$ when the ATI was slightly upwind of the CSAT). The larger difference observed between the portable system and other CSATs during the site visits might result primarily from systematic differences in how these sites process data rather than the instruments themselves. It is important for all flux sites to evaluate the EUROFLUX “gold” files to assure consistency in calculations.

Latent Heat Flux – The average difference between normalized roving system and site measurements of – $0.8\pm 12.7\%$ (mean and standard deviation, $n=16$) was not significantly different from zero. The sample size for this comparison is 16 because the Temple Bowen Ratio system does not directly measure *LE*. An analysis of the frequency of the differences suggests results were normally distributed

Fig. 5. Percentage difference between the portable system and site flux measurements. (a) sensible heat, (b) latent heat, (c) CO₂ flux.

(Fig. 5b). There was no obvious explanation for the large positive outlier. There were no significant impacts of either the type of sonic anemometer or infrared gas analyzer (LiCor closedpath vs. NOAA openpath) on the agreement of site and roving system *LE* values (Table 3).

Carbon dioxide Flux – On average, there is very good agreement between half-hourly CO₂ flux

Fig. 6. Half-hourly CO₂ flux values from the AmeriFlux portable system compared to corresponding values from the visited sites.

values measured by the portable system and values from the *in situ* systems (Fig. 6). A regression using the reduced major axis technique suggests that the slope of this relationship is not significantly different from 1 ($P < 0.01$) and that the intercept is not significantly different from zero. Similarly, the average difference between the normalized roving system and site measurements of $1.5 \pm 11.3\%$ (mean and standard deviation, $n = 17$), was also not significantly different from zero. An examination of the frequency histogram (Fig. 5c) indicated that sites with

the largest deviations were using either openpath sensors (the low outlier) or the bowen-ratio approach (high outlier). The agreement of individual sites with roving system values was significantly worse for sites using the NOAA openpath analyzer than for sites using the LiCOR model 6262 closedpath instrument (Table 3). Surprisingly, the sites using the openpath instrument reported significantly lower CO₂ fluxes than the roving system.

Discussion.

One important result from these comparisons is that with the exception of the measurements of air temperature, no significant bias was observed between values recorded by the roving system and the mean of those recorded by the sites. This suggests that the individual sites (and portable system) are, in general, well calibrated to standards. In the case of air temperature, the positive bias may be partly explained by solar heating of the temperature housing. This cause of bias can be corrected by using a forced-ventilation radiation shield. Examination of diurnal data from other sites suggests that in some

cases the observed temperature biases are relatively fixed and do not vary with solar radiation. Since most temperature sensors are relatively stable, these kinds of bias are easy to remove by calibration (see e.g. Nicholas and White, 1994).

Intercomparability of the network – An important characteristic of a network is the degree to which results from one site are directly comparable to those recorded at another site. Put another way, what is the minimum difference between sets of stations that is significant at some specified level of confidence? With certain assumptions, the present data can address this question. First, we assume that the short-term comparisons with each station are indicative of the long-term relationship with the roving system. It is probably realistic to assume a representative sample of station performance, as sites were not known to carryout any extra calibrations prior to a visit from the roving system. In fact, differences between the roving system and individual sites probably sometimes decreased following a report of our observations and suggestions for reducing differences. In a comparison between stations, most interest will be in integrated values or functional relationships between exchange rates and environmental factors. For integrated values, we assume that the different sites utilize similar data filling procedures when they model missing data (e.g. Falge et al. 2000) and that there are no systematic biases between day and night performance at any of the sites (Moncrieff et al 1996). This last assumption is probably the weakest because nocturnal measurements at some sites are problematic and can lead to significant errors in integrated fluxes (Lee 1998). We return to this last point later.

We used a combinatorial approach to estimate the minimum differences between stations that were likely to be significant. We assume that the calibration differences between a site and the roving system represent fixed biases. We use the 95th percentile of all 2-station comparisons as our metric and assume that the potential difference between as yet unvisited sites and the roving system errors of other stations are similarly distributed. If differences between 2 stations exceed our 95 percentile values, we can assume that these differences are real.

Site net radiation values deviated from roving system measurements by the smallest margins of all variables. Based on our comparisons, 95% percent of the possible 2-station comparisons differed by less

than 15% (Table 4). For PPFD, in contrast, 95% of the possible 2-station comparisons differed by less than 44%, a result of the large positive and negative outliers (Fig. 3B) from 4 of the stations. This poor result compromises our ability to compare light response curves among the network because of uncertainties in the true PPFD values at some sites. For temperature, if we include all sensors, temperature differences between 2 stations must exceed 4.5 °C in order for us to be confident that they are significantly different. Interestingly, the two outliers are stations that do not use commercial sensors. Restricting the analysis only to stations utilizing commercial temperature sensors reduces the significant temperature difference to 2.1 °C.

These results for temperature and PPFD are disappointing and limit our ability to draw inferences about how trends in these variables influence CO₂ and other fluxes. My importantly, model parameterizations or assessments will also suffer if poor quality data are accepted uncritically. Our poor results for PPFD and air temperature also show the way to improved performance in comparisons between stations. We have tried to estimate threshold values that must be exceeded to have reasonable confidence that the differences are real. These estimates are based on no specific knowledge of the calibrations of sensors at these stations. If the specific stations supply calibration information, the precision of these comparisons can be greatly enhanced. For example, if stations recalibrate PPFD sensors on a regular basis, we could potentially discuss whether differences of 10% in PPFD recorded at 2 stations represents a significant difference. Another way for individual stations to increase precision (in addition to comparison with a known standard at frequent intervals) is to maintain duplicate sensors and to watch carefully for deviations in their relative performance.

A third way to reduce bias is by careful selection of sensor. For temperature measurements, platinum resistance thermometers are extremely stable and biases of <0.1 °C can be easily removed. With relatively little cost or effort, the intercomparability of site temperatures could be narrowed by an order of magnitude to less than 0.5 °C. For PPFD, there appears to be a significant decrease in output with time (aging) in silicon sensors. Thermopile-based instruments do not generally suffer from this effect to the same degree.

For fluxes, there is no independent standard, hence the development of the roving system. Our results suggest that the intercomparability of fluxes measured at different sites is moderately good. If values of two-station comparisons differ by more than 26% for H and 35% for LE and 32% for CO_2 flux, these differences are likely to be significant. As in the case of environmental sensors, specific knowledge can reduce the difference needed to attain confidence that values from stations are significantly different. In the case of CO_2 measurements, the stations with flux values that differed most from the reference system utilized either NOAA openpath IRGAs or the bowen-ratio technique. Eliminating these stations from the comparison suggests that if differences between stations in CO_2 flux exceed 23% they will be significant.

Conclusion

This exercise has revealed the need for constant attention to the details of operating long-term meteorological and flux measurement equipment. At ~10% of the sites visited serious calibration problems (disagreement between the site and portable system values of environmental data or fluxes of >25%) were observed. Many of these problems could be identified by the sites via an examination of the energy balance or other diagnostic. In some instances the problems were traced to calculation error rather than measurement error. Evaluating standardized raw data files can help reduce later calculation errors. The EUROFLUX “gold” files are one such standard and should be evaluated by all flux sites.

Sensors at a single station may exhibit a relatively fixed bias and yet provide good internal comparisons. However, such biases reduce the intercomparability of results between network stations and should be vigorously pursued and minimized. For environmental data this is relatively inexpensive and involves using inherently stable sensors such as platinum resistance thermometers rather than e.g. thermocouples, comparison to physical standards or other sensors maintained in the researcher’s laboratory, and regular calibration of sensors by their manufacturer or a standards laboratory. For flux systems, this requires calibration of individual sensors (IRGA, sonic anemometer, etc.) and on-going comparison with a reference system such as that described here.

An unanticipated result of the comparison program was the observation of systematic differences in

sensible heat or CO₂ fluxes depending upon the instrument manufacturer. The trends observed here should be examined with additional instrument samples. Also, flux values from systems utilizing new instruments (LI-7500, LI-7000) remain to be evaluated. It appears that routine intercomparison of flux and environmental instrumentation is warranted and should continue within and between flux measurement networks.

Acknowledgements.

The Department of Energy, Office of Biological and Environmental Research, Terrestrial Carbon Program, provided support to the USDA Forest Service for the AmeriFlux program of intercomparisons and other AmeriFlux Science Team Activities through Interagency Agreement # DE-AI02-96ER62274. AmeriFlux stations are supported within the US by the National Institute of Global Environmental Change (NIGEC) and other programs within the Department of Energy, the National Science Foundation, NASA, the USDA Forest Service, NOAA, the USGS, and the EPA. Support for other AmeriFlux stations comes from the Canadian National Environmental Research Council, and Embrapa, Brazil.

References.

- Aubinet M, Grelle A, Ibrom A, Rannik Ü, Moncrieff J, Foken T, Kowalski AS, Martin PH, Berbigier P, Bernhofer CH, Clement R, Elbers J, Granier A, Grünwald T, Morgenstern K, Pilegaard K, Rebmann C, Snijders W, Valentini R., Vesala T (2000) Estimates of the annual net carbon and water exchange of forests: The EUROFLUX methodology. *Advances in Ecological Research* **30**, 113-175.
- Baldocchi DD, Hicks BB, Meyers TP (1988) Measuring biosphere-atmosphere exchanges of biologically related gases with micrometeorological methods. *Ecology*, **69**, 1331-1340.
- Falge E, Baldocchi D, Olson RJ, Anthoni P, Aubinet M, Bernhofer C, Burba G, Ceulemans R, Clement R, Dolman H, Granier A, Gross P, Grünwald T, Hollinger D, Jensen N-O, Katul G, Keronen P, Kowalski A, Ta Lai C, Law BE, Meyers T, Moncrieff J, Moors E, Munger JW, Pilegaard K, Rannik Ü, Rebmann C, Suyker A, Tenhunen J, Tu K, Verma S, Vesala T, Wilson K, Wofsy S (2001) Gap Filling Strategies for Longterm Energy Flux Data Sets. *Agricultural and Forest Meteorology* **107**, 71-77.
- Goulden ML, Daube BC, Fan S-M, Sutton DJ, Bazzaz A, Munger JW, Wofsy SC (1997) Physiological responses of black spruce forest to weather. *Journal of Geophysical Research*, **102**, 28987-28996.
- Goulden ML, Munger JW, Fan S-M, Daube BC, Wofsy SC (1996) Measurements of carbon sequestration by long-term eddy covariance: Methods and a critical evaluation of accuracy. *Global Change Biology*, **2**, 169-182.
- Hammer, Ø, Harper, DAT, Ryan, PD (2001) PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica* **4(1)**, 9pp. http://palaeo-electronica.org/2001_1/past/issue1_01.htm
- Hollinger DY, Kelliher FM, Schulze E-D, Vygodskaya NN, Varlagin A, Milukova I, Byers JN, Sogachov A., Hunt JE, McSeveny TM, Kobac KI, Bauer G, and Arneth A (1995) Initial assessment of multi-scale measures of CO₂ and H₂O flux in the Siberian taiga. *Journal of Biogeography*, **22**, 425-431.
- Hollinger DY, Goltz SM, Davidson EA, Lee JT, Tu K, Valentine HT (1999) Seasonal patterns and environmental control of carbon dioxide and water vapor exchange in an ecotonal boreal forest. *Global Change Biology*, **5**, 891-902.
- Horst, TW (1997) A simple formula for attenuation of eddy fluxes measured with first-order-response scalar sensors. *Boundary-Layer Meteorology*, **82**, 219-233.
- Horst, TW (2000) On frequency response corrections for eddy covariance flux measurements. *Boundary-Layer Meteorology*, **94**, 517-520.

- Kaimal JC, Gaynor JE (1991) Another look at sonic thermometry. *Boundary-layer Meteorology*, **56**, 401-410.
- Kaimal JC, Wyngaard JC, Izumi Y, Cote OR (1972) Spectral characteristics of surface-layer turbulence. *Quarterly Journal of the Royal Meteorological Society* **98**, 563-589.
- Lee X (1998) On micrometeorological observations of surface-air exchange over tall vegetation. *Agricultural and Forest Meteorology* **91**, 39-49.
- McMillen RT (1988) An eddy correlation technique with extended applicability to non-simple terrain. *Boundary-Layer Meteorology*, **43**, 231-245.
- Massman, WJ (2000) A simple method for estimating frequency response corrections for eddy covariance systems. *Agricultural and Forest Meteorology* **104**, 185-198.
- Massman, WJ (2001) Reply to comment by Rannik on "A simple method for estimating frequency response corrections for eddy covariance systems". *Agricultural and Forest Meteorology* **107**, 247-251.
- Massman WJ, Lee X (2002) Eddy covariance flux corrections and uncertainties in long term studies of carbon and energy exchanges. *Agricultural and Forest Meteorology* **113**, 121-144.
- Moncrieff JB, Mahli Y, Leuning R (1996) The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water. *Global Change Biology* **2**, 231-240.
- Moore CJ (1986) Frequency response corrections for eddy correlation systems. *Boundary-layer Meteorology*, **37**, 17-35.
- Nicholas JV, White DR (1994) Traceable temperatures. An introduction to temperature measurement and calibration. John Wiley & Sons, Chichester, England.
- Webb EK, Pearman GI, and Leuning R (1980) Correction of flux measurements for density effects due to heat and water vapor transfer. *Quarterly Journal of the Royal Meteorological Society*, **106**, 85-100.

Table 1.

AmeriFlux Sites and other sites visited by the portable system (January, 2001). Bold indicates a comparison site. Some recently visited sites are not in the comparison described here.

SITE	Location	Vegetation	Institution
Barrow	Barrow, Alaska	tundra	San Diego
Blodgett Forest	Georgetown, CA	ponderosa pine	UC Berkeley
Bondville	Champaign, IL	soybean, corn	NOAA/ATDD
BOREAS N.	Thompson, Canada	boreal spruce forest	Harvard Univ
BOREAS S.	Prince Albert, Canada	boreal aspen forest	U B C
Camp Borden	Camp Borden, Canada	mixed forest	U Virginia
Cerrado	Sao Paulo State, Brazil	C4 grassland	U Sao Paulo
Duke FACE	Durham, NC	loblolly pine	Duke Univ.
Gainesville	Gainesville, FL	slash pine	Univ. Fla
Glacier Lakes	Glacier Lakes, WY	sub-alpine fir	USDA-FS
Happy Valley	north slope, Alaska	tundra	San Diego
Harvard Forest	Petersham, MA	deciduous forest	Harvard Univ
Howland	Howland, ME	coniferous forest	USDA-FS/U Maine
Irvine	Irvine, CA	marsh	UC Irvine
Jasper Ridge	Stanford, CA	grassland	Carnegie Inst.
Konza Prairie	Konza Prairie, KS	C4 prairie	Kansas State
La Selva	La Selva, Costa Rica	tropical forest	Florida U
Lethbridge	Lethbridge, Canada	C3/C4 grassland	U Lethbridge
Little Washita	Chickasha, OK	rangeland	NOAA/ATDD
Mer Bleue	Gloucester, Canada	Fen, wetlands	McGill
Metolius	Redmond, OR	Ponderosa pine	Oregon State
Morgan Monroe	Monroe, IN	deciduous forest	Indiana U
Niwot Ridge	Niwot Ridge, CO	sub-alpine fir	U Colorado
Park Falls	Park Falls, WI	mixed forest	NOAA/CMDL
Ponca City	Ponca City, OK	wheat	U Nebraska
SERC	Edgewater, MD	deciduous forest	Smithsonian
Shidler	Shidler, OK	tallgrass prairie	U Nebraska
Sky Oaks	southern CA	chaparral	San Diego
Sugar cane	Sao Paulo State, Brazil	Sugar cane	U Sao Paulo
UMBS	UMBS, MI	mixed forest	U Michigan
U-Pad	north slope, Alaska	tundra	San Diego
Vancouver Isl.	Vancouver, Canada	Douglas fir	U B C
Walker Branch	Oak Ridge, TN	oak-hickory forest	NOAA/ATDD
Walnut River	Smileyberg, KS	grazed pasture	Argonne Nat'l lab
Wind River	Wind River, WA	Douglas fir	UC Davis
Non-AmeriFlux sites:			
Norunda	Norunda, Sweden	coniferous	Swedish Lund U
Temple	Temple, TX	prairie	Blackland Res Center
Tumbarumba	Tumbarumba, Australia	Eucalyptus forest	CSIRO
Twizel	Twizel, New Zealand	grassland	Landcare Research

Table 2.

Relationship between the AmeriFlux roving system team calculations of fluxes from the EUROFLUX “gold files” and the EUROFLUX average (recursive filter detrending) or Harvard team values. The AmeriFlux value is the dependent variable in the relationship $Y = mX$.

Flux	slope	95% confidence intervals	r^2
<i>H</i> (vs. EUROFLUX)	1.0184	±0.0260	0.997
<i>H</i> (vs. Harvard)	1.0028	±0.0271	0.997
<i>LE</i> (vs. EUROFLUX)	1.0538 *	±0.0520	0.990
<i>LE</i> (vs. Harvard)	0.9693 *	±0.0261	0.997
<i>FCO₂</i> (vs. EUROFLUX)	1.0315	±0.0424	0.993
<i>FCO₂</i> (vs. Harvard)	1.0354 *	±0.0351	0.995

Asterisks indicate slopes that are significantly different from 1.

Table 3.

Mean difference between site and AmeriFlux roving system eddy covariance flux measurements by instrument.

Sonic Anemometers

Flux	Mean difference by instrument (%)			<i>P</i>	F(2,13)	LSD
	<u>ATI¹ (n=6)</u>	<u>Gill² (n=7)</u>	<u>CSAT3 (n=3)</u>			
<i>H</i>	4.9(a)	-1.2(a)	-12.2(b)	<0.05	5.67	9.4
<i>LE</i>	4.8(a)	-3.5(a)	-5.6(a)	ns	0.95	16.9
<i>FCO₂</i>	0.8(a)	2.1(a)	-4.6(a)	ns	1.25	14.6

1. 5 Kaimal probe and 1 model Sx instruments

2. 6 model R3/R3A and 1 model HS instruments

Infrared Gas Analyzers

Flux	Mean difference by instrument (%)		<i>P</i>	F(1,14)	LSD
	<u>LiCOR (n=13)</u>	<u>NOAA (n=3)</u>			
<i>H</i>	-0.0(a)	0.5(a)	ns	0.01	10.4
<i>LE</i>	-0.6(a)	-1.6(a)	ns	0.02	14.1
<i>FCO₂</i>	3.9(a)	-10.0(b)	<0.05	4.57	10.6

Table 4.

Comparison threshold values – if differences between a pair of stations exceed these values, the differences are likely to be real and not associated with calibration errors.

Variable	95 percentile difference	range	n
R _n	15%	23%	20
PPFD	44%	62%	18
T _{air} (all sensors)	4.5 °C	7.4 °C	20
T _{air} (commercial sensors)	2.1 °C	2.7 °C	18
H	26%	33%	17
LE	35%	50%	16
F _{CO2} (all)	32%	43%	17
F _{CO2} (closed-path EC)	23%	33%	14