

An Evaluation of Calibration Techniques for In Situ Carbon Dioxide Measurements Using a Programmable Portable Trace-Gas Measuring System

SEAN P. BURNS,* ANTHONY C. DELANY,[†] JIELUN SUN, BRITTON B. STEPHENS,
STEVEN P. ONCLEY, GORDON D. MACLEAN, AND STEVEN R. SEMMER

National Center for Atmospheric Research,[&] Boulder, Colorado

JOEL SCHRÖTER[#] AND JOHANNES RUPPERT[@]

Department of Micrometeorology, University of Bayreuth, Bayreuth, Germany

(Manuscript received 11 October 2007, in final form 4 June 2008)

ABSTRACT

The construction and deployment of a portable trace-gas measurement system (TGaMS) is described. The air-collection system (dubbed HYDRA) collects air samples from 18 different locations and was connected to either one or two LI-COR LI-7000 gas analyzers to measure CO₂. An in situ “field calibration” method, that uses four calibration gases with an uncertainty on the order of $\pm 0.1 \mu\text{mol mol}^{-1}$ relative to the WMO CO₂ mole fraction scale, revealed CO₂ output from the LI-7000 had a slightly nonlinear relationship relative to the CO₂ concentration of the calibration gases. The sensitivity of the field-calibrated CO₂ to different forms of the field-calibration equation is investigated. To evaluate TGaMS performance, CO₂ from collocated inlets, portable gas cylinders, and nearby independent CO₂ instruments are compared. Results are as follows: 1) CO₂ measurements from HYDRA multiple inlets are feasible with a reproducibility of $\pm 0.4 \mu\text{mol mol}^{-1}$ (based on the standard deviation of the CO₂ difference between collocated inlets when HYDRA was operating with two LI-7000s); 2) CO₂ differences among the various field-calibration equations were on the order of $\pm 0.3 \mu\text{mol mol}^{-1}$; and 3) comparison of midday hourly CO₂ measurements at 30 m AGL between TGaMS and an independent high-accuracy CO₂ measurement system (within 300 m of TGaMS) had a median difference and standard deviation of $0.04 \pm 0.81 \mu\text{mol mol}^{-1}$ over two months.

1. Introduction

Measuring trace gases to understand and to monitor the earth for climate change has become an important and common task within the biogeoscience community

(Hansen et al. 2007). Within the past several decades the World Meteorological Organization (WMO) and other organizations around the world have invested a considerable amount of effort examining the accuracy of trace-gas measurements. Since the late 1970s, the WMO has organized 13 meetings of the so-called Experts on Carbon Dioxide Concentration to discuss CO₂ and trace-gas-measurement-related topics (Worthy and Huang 2005; Miller 2006). One of the recommendations by this group is that interlaboratory comparisons of CO₂ on the global network be on the order of $\pm 0.1 \mu\text{mol mol}^{-1}$. Historically these discussions have focused on the accuracy of long-term CO₂ measurements (Bakwin et al. 1995; Trivett and Köhler 1999) and air sampling with flasks; however, recent discussions have included recommendations for in situ CO₂ measurements (Miller 2006) as well as examples of continuously measuring, multi-inlet systems using a single sensor (e.g., Manning 2005; Stephens et al. 2006).

* Additional affiliation: Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, Colorado.

[†] Retired.

[#] Current affiliation: Meteorology and Air Quality Group, Wageningen University, Netherlands.

[@] Current affiliation: Environmental Measuring Department, Research Institute of the Cement Industry, Duesseldorf, Germany.

& The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Sean P. Burns, NCAR/MMM/TIIMES, P.O. Box 3000, Boulder, CO 80307-3000.
E-mail: sean@ucar.edu

Atmospheric CO₂ measurements typically have one of three goals: 1) to make highly accurate measurements of CO₂ that can be used in long-term CO₂ monitoring or regional CO₂ budget studies, 2) to measure CO₂ fluctuations and calculate CO₂ fluxes using the eddy-correlation method (Baldochchi et al. 1988; Berger et al. 2001; Massman and Lee 2002), or 3) to measure CO₂ differences from multiple inlets using a single infrared gas analyzer (IRGA) with excellent reproducibility to measure changes in local CO₂ gradients. Each measurement goal requires different emphasis on the instrumentation as well as processing techniques. For example, calculating the CO₂ flux requires a fast-response instrument, high flow rates (for a closed-path gas analyzer), and good short-term repeatability, whereas a study of temporal changes of mean CO₂ concentration over multiple years requires a high-accuracy sensor or a methodology to correct for sensor drift over time (e.g., by the use of calibration gases). The term “CO₂ accuracy” refers to the accuracy of a CO₂ measurement relative to the WMO CO₂ mole fraction scale (hereafter WMO CO₂ Scale) that is currently maintained at the National Oceanic and Atmospheric Administration (NOAA) Global Monitoring Division (GMD) as described by Zhao and Tans (2006). Because accuracy is a qualitative expression, we will use the so-called combined standard uncertainty to show the degree of accuracy (expressed as ± 1 standard deviation such that there is a 67% probability the true value is within the uncertainty range). Definitions of “reproducibility” and “repeatability” are given toward the end of this section.

One use of multi-inlet CO₂-measurement systems that has recently received a lot of attention in the ecosystem flux community is determining the effect of horizontal CO₂ advection on the local CO₂ budget (e.g., Aubinet et al. 2003; Staebler and Fitzjarrald 2004; Sun et al. 2007; Yi et al. 2008, among many others). Horizontal advection is typically determined by estimating the horizontal divergence or convergence of CO₂ through the “sides” of an imaginary control volume around a tower that measures the vertical CO₂ flux. Because estimation of CO₂ advection requires accurate measurements of spatial CO₂ gradients, reproducibility of the mean CO₂ differences among inlets is much more important than a high-accuracy CO₂ measurement. To study the problem of horizontal CO₂ advection, a portable trace-gas measurement system was developed at the National Center for Atmospheric Research (NCAR) Earth Observing Laboratory (EOL). The NCAR trace-gas measurement system (TGaMS) comprises an 18 inlet air-sampling system (hereafter “HYDRA”) connected to either one or two analysis systems that house a LI-COR LI-7000 CO₂-H₂O non-

dispersive infrared (NDIR) dual-cell IRGA (LI-COR Inc. 2005). TGaMS has been deployed in several field campaigns to study the horizontal transport of CO₂ in a Colorado subalpine forest: the 2002 pilot study (NIWOT02) and the 2004 Carbon in the Mountains Experiment (CME04). TGaMS was also used in conjunction with another CO₂ measuring system (based on a LI-COR LI-820 IRGA) to measure vertical and horizontal CO₂ gradients in the Wavelet Detection and Atmospheric Turbulent Exchange Measurements 2003 (WALDATEM-2003) experiment that studied the turbulent exchange processes in a tall spruce forest at the FLUXNET station Waldstein-Weidenbrunnen in northeastern Bavaria (Thomas and Foken 2007; Thomas et al. 2004). During the winters of 2003–04 and 2004–05 TGaMS was used in the UNDERSNOW projects to study gradients of CO₂ within the snowpack of a subalpine forest (Monson et al. 2006a,b). The use of TGaMS in NIWOT02 and CME04 is the focus of our study.

The LI-7000 determines the CO₂ concentration of an air sample by comparing absorption within the “sample” cell to absorption within the “reference” cell that contains either CO₂-free air or air of constant CO₂ concentration. It is well known that CO₂ concentration measured by an IRGA depends on cell temperature and pressure following the ideal gas law (McDermitt et al. 1993). The LI-COR factory calibration uses a fifth-order polynomial to take into account cell temperature and pressure changes, and the LI-7000 is specified to have a nominal error on the order of $\pm 1\%$ over a range 0–3000 $\mu\text{mol mol}^{-1}$ (LI-COR Inc. 2005). Researchers that desire high-accuracy CO₂ measurements often control the temperature and pressure within the sample cell in order to minimize any imperfections in the factory calibration, as well as reducing errors due to the thermal micro-expansion–contraction of the optical bench and temperature effects on the internal electronics (D. Anderson 2007, personal communication). Another way to improve the IRGA CO₂-measurement accuracy, is by an in situ “field calibration”; achieved when air with varying CO₂ concentrations (so-called calibration gases) are sequentially passed through the sample cell (Trivett and Köhler 1999). Trivett and Köhler (1999) noted that a linear expression could be applied to correct the IRGA with the calibration gases, but the true calibration curve is nonlinear, which requires three or four calibration gases to characterize. Zhao et al. (1997) described a technique for processing data from a single-inlet CO₂-measuring system that used four calibration gases, a LI-COR IRGA (model LI-6251), and a second-order polynomial to convert the raw CO₂ voltage into CO₂ concentrations. They estimated the uncertainty of their system to be better than $\pm 0.2 \mu\text{mol mol}^{-1}$ by

considering the residuals in the second-order polynomial fit and errors in the CO₂ concentration of the calibration gas. Following the methodology presented by Zhao et al. (1997), other researchers have used similar data-processing techniques for their CO₂-measuring systems (e.g., Haszpra et al. 2001; Daube et al. 2002; Stephens et al. 2006). Ocheltree and Loescher (2007) used a LI-7000 and calibration gases prepared by NOAA/GMD to determine a third-order polynomial calibration curve over a range of 330–500 $\mu\text{mol mol}^{-1}$ in the laboratory. In the field, they calculated the CO₂ concentration from the raw CO₂ output by the LI-7000 (using the laboratory-determined coefficients) and an in situ linear field calibration using two calibration gases. Ocheltree and Loescher estimated an uncertainty of $\pm 0.6 \mu\text{mol mol}^{-1}$ for CO₂ by considering the following: residuals in the laboratory-determined third-order polynomial, uncertainty in the calibration gases (relative to the WMO scale), and the digital-to-analog conversion of the CO₂ data.

If a CO₂ measurement does not require high accuracy, then LI-COR IRGAs can be used with two calibration gases (or one calibration gas and one CO₂-free gas) and the LI-COR factory-determined calibration polynomial (Xu et al. 1999; Mölder et al. 2000; Monson et al. 2002; Liang et al. 2003). Most of these studies are more interested in reproducibility of CO₂ differences so accuracy relative to the WMO CO₂ Scale is not a primary concern.

Our study follows the International Organization for Standardization (ISO) definitions for accuracy, comparability, repeatability, and reproducibility (ISO/IEC 1995, 2007). “Comparability” refers to the ability to compare a measurand between two independent measuring systems using a common reference or scale (in our case the measurand is CO₂, and the common reference is the WMO CO₂ Scale). Comparability is a qualitative expression and we use empirical comparisons to express the degree of comparability—unless otherwise noted, this is done with a mean or median difference plus or minus the standard deviation of the difference. These comparison statistics will usually be influenced by real differences in the measured data, as well as any systematic and random measurement errors. Precision is defined as “the closeness of agreement between independent test results obtained under stipulated conditions” and is typically expressed by an estimate of the dispersion of the measurement, such as the variance or standard deviation (Taylor and Kuyatt 1994). Precision is called repeatability if the environmental conditions during the measurement do not change; otherwise it is labeled “reproducibility.” In general, precision without any qualifier should be taken

as a synonym for repeatability (Taylor and Kuyatt 1994). For our study, reproducibility refers to the ability of TGaMS to maintain consistent CO₂-difference measurements among the inlets as the environmental conditions change.

TGaMS was designed for studying horizontal advection—this requires excellent reproducibility among inlets, but not necessarily high accuracy. Previous advection studies around individual towers with a single IRGA and long tubing have proven successful (Aubinet et al. 2003; Staebler and Fitzjarrald 2004), but as these studies expand to “carbon-shed” or regional scales then the use of long tubing becomes impractical, and CO₂ measurements at the $\pm 0.1 \mu\text{mol mol}^{-1}$ uncertainty level suggested by WMO becomes necessary in order to calculate CO₂ differences from independent CO₂-measuring systems. One of the goals of our study is to determine whether TGaMS should be considered a high-reproducibility or high-accuracy CO₂-measuring system.

2. Instrumentation details

TGaMS consists of an air-sampling system, HYDRA, that collects air from 18 inlets (Fig. 1), and an analysis system, which is housed in a separate weatherproof box (Fig. 2). HYDRA and the analysis system are wired together for communication and power sharing. TGaMS was designed to be portable with a total weight of about 70 kg (Fig. 3, top-right photo). For our projects TGaMS had access to line power, and power consumption by TGaMS is on the order of 0.4 kW (primarily due to the pumps). A general description of TGaMS is given below, details specific to CME04 and NIWOT02 are in section 3.

Each of the HYDRA inlet lines was covered by a 1- μm Gelman filter (part #9967-008, LI-COR, Inc.) and air was drawn through a length of composite aluminum-polyethylene tubing (Type 1300 tubing, Synflex, inner diameter = 0.625 cm) that varied between 5 and 300 m for our experiments. The flow rate of the sample air from each inlet was typically set to 2 L min⁻¹ (lpm) by a valve (model MFVB6, OMEGA Engineering, Inc.), and monitored with a rotameter flowmeter (FL-1800 Series, OMEGA Engineering, Inc.). Upon leaving the Synflex tubing, the air sample entered a buffer volume that was a 2-L mason jar packed in foam at the bottom of the HYDRA box (Fig. 3, middle photo). The buffer volumes were fitted with stainless steel lids that had two 0.95-cm and one 0.635-cm bored-out Swagelok bulkhead fittings (models 600-61 and 400-61, Swagelok Co.) inserted through them. Tubing to one of the 0.95-cm fittings was from the inlet, while the other fitting connected the suction manifold. The 0.635-cm bulkhead fitting was for

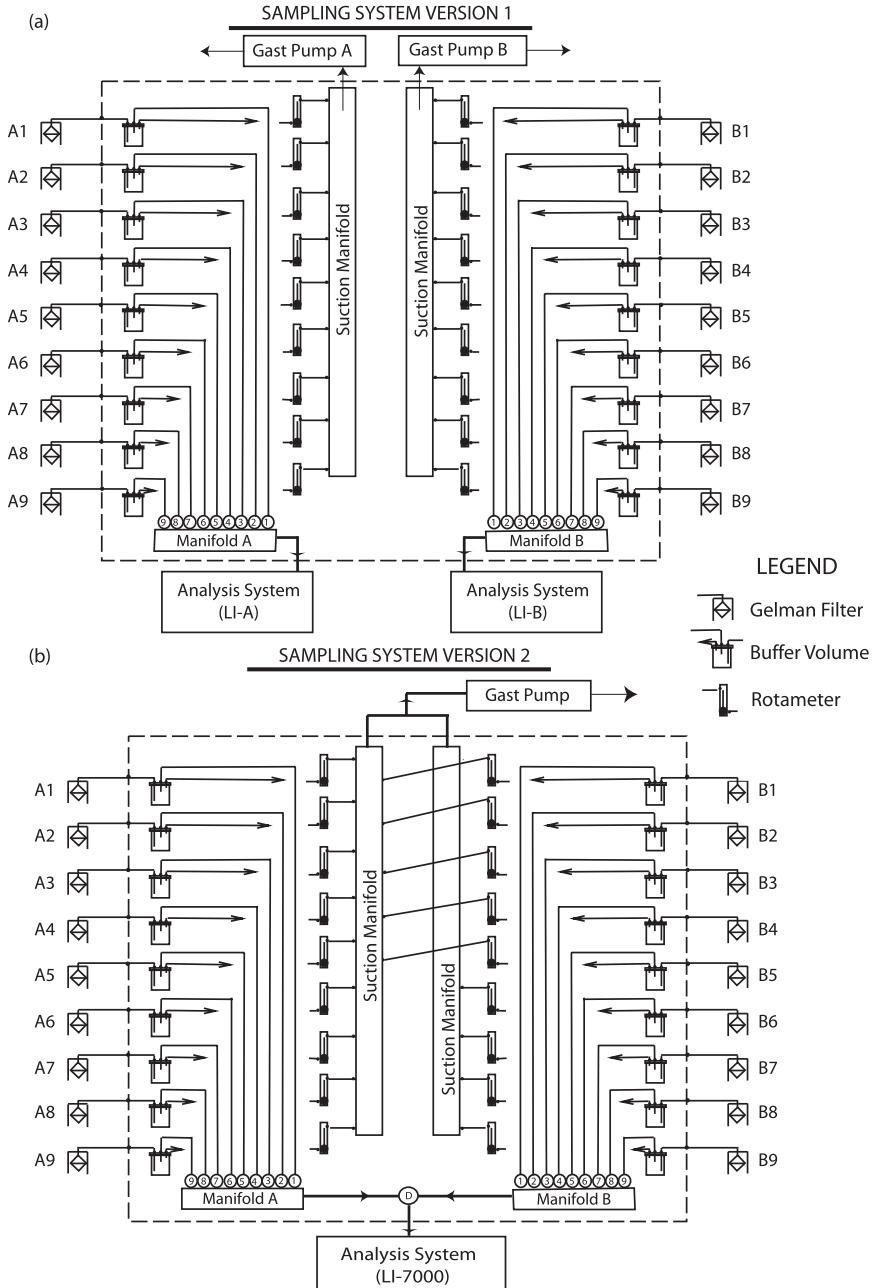


FIG. 1. Schematic of the TGaMS 18 inlet air-collection system (HYDRA) connected to (a) two or (b) one analysis system. In (b), only one analysis system is used by introducing a three-way solenoid valve (labeled D near bottom of panel), and only one Gast pump is used for the suction manifolds. The dashed line represents the HYDRA wooden box (60 cm long \times 55 cm wide \times 90 cm tall). For details on the analysis system see Fig. 2.

the tube that led to the analysis system. The tubing ends within the buffer volumes were arranged to maximize mixing as shown in Fig. 1. The suction manifolds were short lengths of PVC tubing (\sim 70 cm long) with Swagelok fittings tapped along the length of the tube to connect to the buffer volumes. One of the capped ends of

each suction manifold was drilled out and equipped with a barbed connector and flexible tubing to attach a vacuum pump (model RAA-V210-EB, Gast Manufacturing, Inc.) that created a sufficient pressure drop for drawing the air samples into HYDRA. By wiring the power supply for the Gast pumps into the control system, the

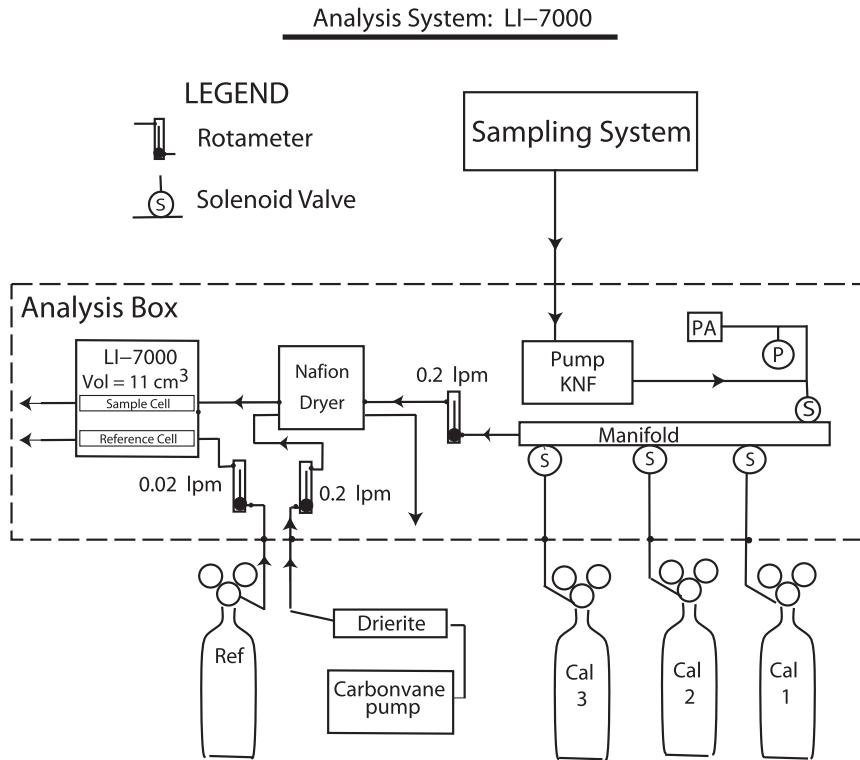


FIG. 2. The TGaMS analysis system with a LI-COR LI-7000 CO₂ gas analyzer. The schematic shown here is with three calibration cylinders (for NIWOT02). For CME04, four calibration cylinders were used. The dashed line represents the border of the weatherproof box that houses the analysis system. For details on the HYDRA sampling system, see Fig. 1. Here P is a pressure gauge and PA is a miniature sequence valve (model S-5300-10, Airtrol Components, Inc.).

pumps can be turned on and off by the TGaMS software to allow for periodic flushing of the buffer volumes (useful for measuring CO₂ within snow). Whenever the Gast pumps are turned off, a “snapshot” of the CO₂ concentration field is captured within the buffer volumes and can subsequently be sampled by the LI-7000. If one analysis system is used in TGaMS, a three-way solenoid valve can be used to connect the manifolds A and B as shown in Fig. 1b.

Air samples were drawn from the buffer volumes into the analysis system by a KNF diaphragm pump (model UN89, KNF Neuberger, Inc.) that was housed within the analysis system (Fig. 2). After the air sample was expelled by the KNF pump it passed through a normally open solenoid valve (model HS02M1H00V5, Numatics, Inc.) into a manifold. Whenever this valve was closed, one of the solenoid valves for the calibration gases was opened so that the manifold contained either an air sample or one of the calibration gases. Positive pressure within the manifold was created by either the KNF pump or the two-stage regulator attached to the calibration gas cylinders. The flow rate within the analysis system was controlled by an OMEGA valve and ro-

tameter located upstream of the LI-7000 sample cell and typically set to 0.2 lpm. Before the air sample reached the LI-7000 sample cell it was dried in a section of Nafion counterflow gas-drying tubing (model MD-110, Perma Pure LLC). Dry air for the Nafion tubing was created by an additional pump that pushed ambient air through a canister of DRIERITE dessicant. The flow rate of the dry air within the Nafion tubing was controlled by a second valve and rotameter within the analysis system, and was set to match the air sample flow rate (Fig. 2). A third valve and rotameter controlled the flow rate in the LI-7000 reference cell. Air exiting the LI-7000 reference and sample cells were exhausted at atmospheric pressure.

For our experiments, the LI-7000 was not temperature or pressure controlled; however, to improve the CO₂ accuracy and reproducibility, an in situ field calibration was achieved by periodically passing a series of calibration gases through the sample cell. A summary of the calibration gases used in NIWOT02 and CME04 is in Table 1. An accurate, rigorously tested calibration gas (i.e., gases with CO₂ uncertainty on the order of $\pm 0.1 \mu\text{mol mol}^{-1}$ or better) is the foundation of any

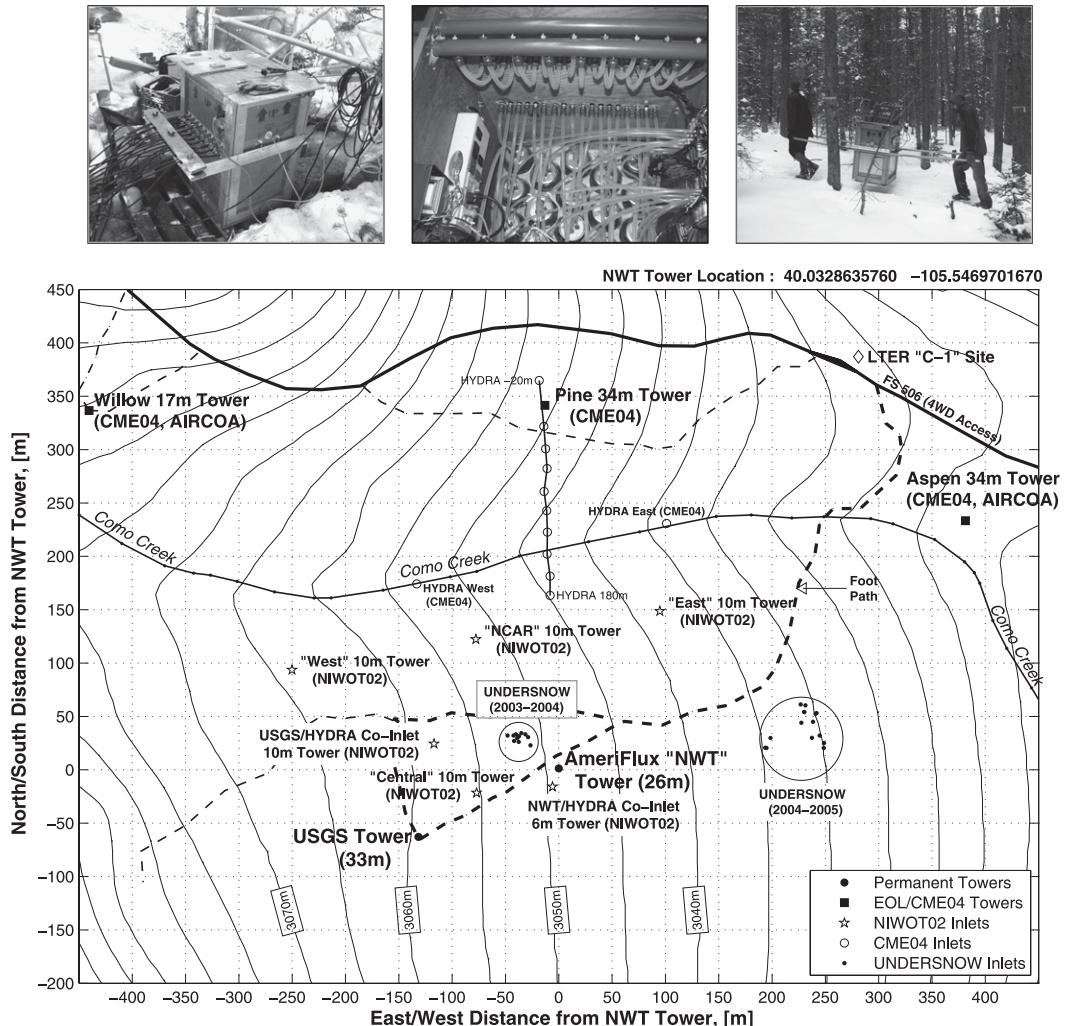


FIG. 3. Site topography and tower layout at the NWT site. The Long-Term Ecological Research (LTER) C-1 site is a long-term climate station within the Niwot Ridge Biosphere Reserve. Elevation contours at 5-m intervals are from the USGS 7.5-min DEM. The Como Creek data are from the “Hydrography–Streams” data available from the Boulder County GIS Web page (see online at <http://www.co.boulder.co.us/gis/>). All tower locations are differentially corrected measurements with a Trimble GPS Pathfinder Pro XL system. The photos show (left) outside of HYDRA box, (middle) inside of HYDRA box, and (right) HYDRA being transported through the forest.

high-accuracy CO₂ measurement. In our study we focus on the sensitivity of the CO₂ field calibration procedure and equations on the calculated CO₂ with the assumption that the CO₂ concentration assigned to each calibration gas has no systematic error. In reality there is, of course, some uncertainty in the concentration value assigned to each calibration gas. Zhao and Tans (2006) provided a detailed description of how the concentrations are assigned to calibration gases, and estimated an uncertainty of $\pm 0.071 \mu\text{mol mol}^{-1}$ relative to the WMO CO₂ Scale for the “working standards” (calibration gases that are prepared by NOAA/GMD and used in the field with CO₂-measuring systems). Because the CO₂ con-

centration of the calibration gases used in NIWOT02 and CME04 were calculated with different methods, this information is discussed in section 3.

The coordinator of TGaMS was a computer-controlled microprocessor [16-channel B1 (digital) Optomux Protocol Brain Board, Opto 22, Inc.] that was bolted to the inside of the wooden HYDRA box. The control board energized solid state optically-isolated OAC5 relays that determined which buffer volume was sampled within HYDRA, as well as controlling whether the LI-7000 measured a calibration gas or an air sample within the analysis system. To easily determine the state of the Opto 22, LEDs were connected to the relays and set

TABLE 1. Details of the NCAR CO₂-measuring system (TGAoMS) during the NIWOT02 and CME04 projects. (For an overview of the NIWOT02 and CME04 projects visit the NCAR EOL Web page at <http://www.eol.ucar.edu/rtf/projects/>.)

Project	NIWOT02	CME04
Dates	3–30 Sep 2002	21 Jul–5 Oct 2004
No. of LI-7000s	2	1
Reference cell CO ₂ concentration	323.16 ± 0.11 (4)	363.23 ± 0.02 (6)
Calibration gas CO ₂ concentration ^a	283.38 ± 0.11 (4), 360.06 ± 0.04 (4), 419.52 ± 0.10 (4)	335.09 ± 0.02 (11), 349.51 ± 0.03 (8), 398.23 ± 0.02 (8), 417.96 ± 0.02 (6)
Calibration gas uncertainty	±2 ^b $\mu\text{mol mol}^{-1}$	±0.1 $\mu\text{mol mol}^{-1}$
Calibration frequency	Hourly	Either hourly or every 2.5 h
Sampling strategy ^c	Continuous flow	Continuous flow
Flushing	CAL GASES: 215 s	CAL GASES: 116 or 40 s
Time ^d	AIR SAMPLES: 55 s	AIR SAMPLES: 90 or 40 s
Sampling time ^e	25 s	10 s

^a Calibration gas CO₂ concentrations all have units $\mu\text{mol mol}^{-1}$ and include plus or minus the standard deviation of the *N*-independent samples (*N* is shown in parentheses). For CME04 the mean from the pre- and postproject values determined at the NCAR O₂/CO₂ calibration facility (see online at <http://www.eol.ucar.edu/~stephens/CALFAC>) were used.

^b The uncertainty estimate for NIWOT02 gases is a rough estimate based on repeatability considerations and CO₂ field comparisons.

^c A “continuous flow” sampling strategy refers to whether or not the pump for the inlets (Gast) is running at the same time the air is sampled from the buffer volumes by the KNF pump.

^d “Flushing time” is the amount of time allowed to flush out the system before statistics are calculated. “CAL GASES” are the times used when calibration gases are measured, while “AIR SAMPLES” are the times used for atmospheric air samples. An example calibration period is shown in Fig. 7, where the flushing time is 215 s and the sampling time is 25 s.

^e “Sampling time” is the length of the time used to calculate the mean CO₂ values (and other statistics). This is equivalent to the sampling time period shown in Fig. 7.

into the outside of the HYDRA box. The software used to communicate with the Opto 22 control board and also collect the 1-Hz serial output from the LI-7000, was the UNIX-based NDAQ software designed by the NCAR EOL Integrated Surface Flux Facility (ISFF) group. The NDAQ software has been implemented on several platforms and is available on request from ISFF.

3. Field deployments

Most TGAoMS deployments have taken place near the Niwot Ridge Forest AmeriFlux site (NWT; Fig. 3, more information available online at <http://public.ornl.gov/ameriflux/>) where measurements of vertical profiles of mean CO₂ with a LI-COR LI-6251, and CO₂ flux at 21.5 m with a LI-COR LI-6262 have been made since November 1998 (Monson et al. 2002; Turnipseed et al. 2002). The NWT LI-CORs are field calibrated with two gases—a “span” gas that has a CO₂ concentration of around 400 $\mu\text{mol mol}^{-1}$, and one with ultra-high purity Nitrogen (the N₂ gas is also used in the reference cell of each of the NWT LI-CORs). The span gas concentration is determined with a LI-6251 and a NOAA/GMD “secondary standard” gas cylinder in a trailer at the site; N₂ is first passed through the sample cell, followed by the GMD secondary standard gas, followed by the gas with the unknown concentration. The procedure is repeated

several times to assign a concentration to the span gas. In the field, NWT CO₂ is calculated from the factory calibration combined with in situ linear coefficients from field calibrations that occur every 4 h (Monson et al. 2002). Another tall tower making multiyear measurements at the Niwot Ridge forest site is managed by the U.S. Geological Survey (USGS). The USGS system has a LI-7000 located between the USGS and NWT towers and inlets located on both the USGS and NWT towers; the USGS calibration gas concentrations were determined in the laboratory using a NOAA/GMD standard gas and a LI-COR LI-6262 (Yi et al. 2008). The CO₂ data from both the NWT and USGS systems are used in section 5b.

Several different sampling strategies have been employed with TGAoMS. A typical sampling sequence from NIWOT02 (with two LI-7000s) and two different sampling strategies from CME04 (with one LI-7000) are shown in Fig. 4. The initial CME04 sampling strategy (middle panel) was chosen for sequential sampling of the inlets. On 10 September 2004 the sampling scheme was modified (bottom panel) so that the spatial location of the inlet was prioritized (starting at the southern inlets and ending at the northern ones). It was preferable to reprogram the inlet-sampling software rather than physically moving the inlets or reconfiguring the tubing at the HYDRA box.

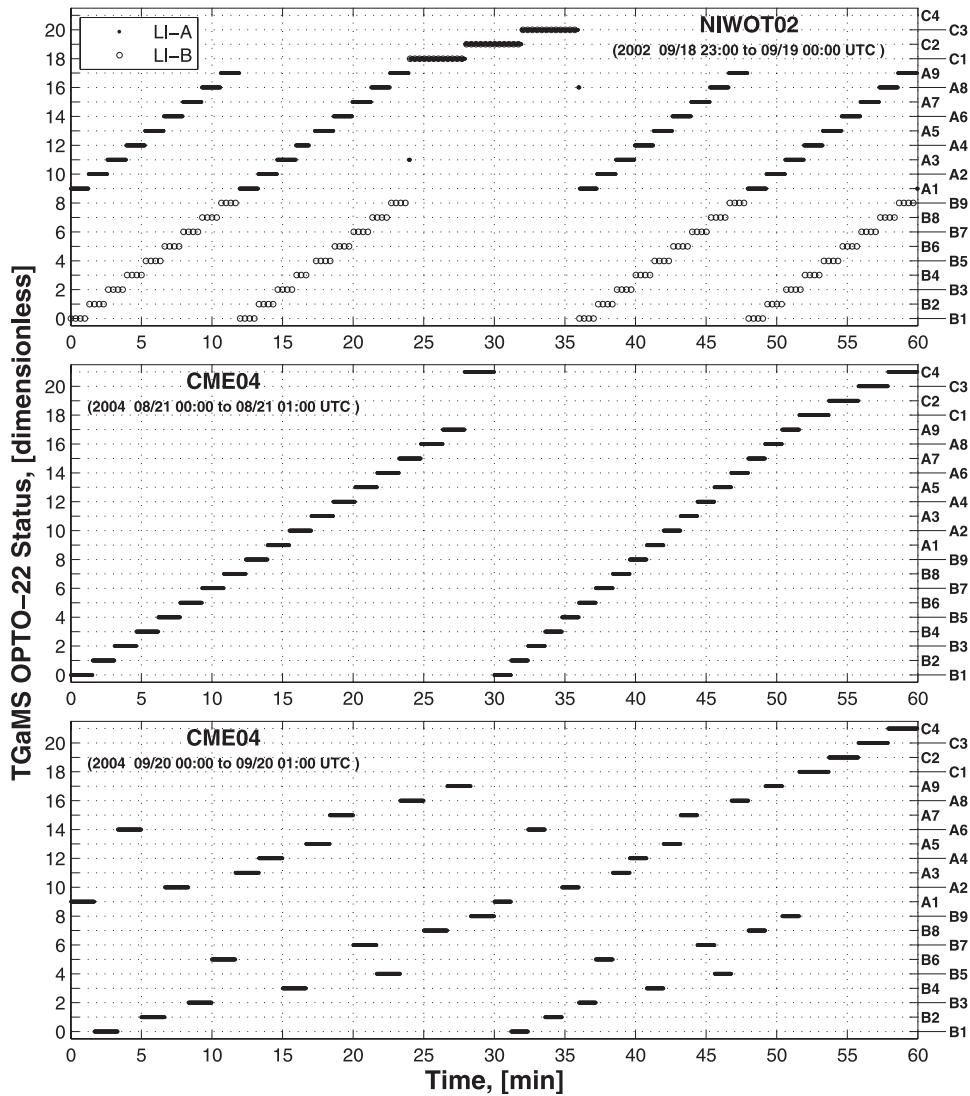


FIG. 4. 1-h time periods showing examples of the different sampling strategies used during (top) NIWOT02, (middle) CME04 in August, and (bottom) CME04 in September. Values of 0–17 correspond to air inlets being sampled and values of 18–21 are calibration gases, as indicated by the alphanumeric designations on the right-hand side of the figure. For clarity, not every 1-s data sample is shown. Note that during NIWOT02 there were occasional data dropouts by the data system that are visible.

a. NIWOT02 experiment

In September 2002 the first field deployment of TGaMS took place at the NWT site as part of an effort to study the effect of horizontal transport of CO₂ on the overall CO₂ budget within a forested mountainous region (Sun et al. 2007). During NIWOT02 the 18 HYDRA inlets were spread over a 300 m × 300 m section of the forest at locations shown in Fig. 3 and at various heights above the ground listed in Table 2.

Two LI-7000 gas analyzers (designated LI-A and LI-B) were running in parallel to increase the time-space res-

olution of the CO₂ measurements. The two identical analysis systems (Fig. 2) were connected to HYDRA as shown in Fig. 1a. Tees were used to connect the same calibration gases to both LI-A and LI-B, and the Opto 22 control board activated the valves in both analysis systems simultaneously (i.e., A1 and B1, A2 and B2, etc., were sampled at the same time). Figure 5a shows 12 h of the raw CO₂ data from both LI-A and LI-B with a 12-min calibration time period (4 min for each gas) that appears in the time series as “steplike” changes in the CO₂ values. Atmospheric inlets were each sampled for 80 s. All 18 inlets were sampled every 12 min, and a full cycle,

TABLE 2. The NIWOT02 HYDRA inlet ID values, locations (see Fig. 1), heights, tubing lengths, and estimated air sample travel times.

HYDRA ID	Inlet location	Inlet height (m)	Tubing length (m)	Delay time (s)
A1*	NCAR	1	5	25
A2	NCAR	6	10	35
A3	NCAR	3	7	30
A4*	NCAR	10	14	38
A5	Central	6	157	120**
A6	Central	3	154	115**
A7	Central	1	151	110
A8	USGS North	1	96	60
A9	NWT	1	181	125
B1	NCAR	1	5	25
B2	West	10	173	133**
B3	West	6	169	130**
B4	West	3	166	125**
B5	West	1	163	120**
B6	East	10	171	133**
B7	East	6	167	130**
B8	East	3	164	125**
B9	East	1	161	120

* Prior to 5 Sep 2002, the 10-m inlet was connected to A1 and the 1-m inlet was connected to A4.

** The delay times for these inlets were calculated based on the relationship between delay time and tubing length from the inlets with directly measured delay times.

consisting of a calibration followed by 4 cycles of atmospheric sampling, took 1 h. For comparison and evaluation of the CO₂ measured by TGaMS, 1-m AGL collocated inlets were placed at the NWT, USGS, and NCAR towers (the A1 and B1 inlets were collocated at the NCAR tower and sampled at the same time as shown in Fig. 5).

The CO₂ concentration of the three NIWOT02 calibration gases (Table 1) were estimated using the same method as that used for the NWT calibration gas (as described above). The uncertainty of this method is probably on the order of $\pm 2 \mu\text{mol mol}^{-1}$ relative to the WMO CO₂ Scale.

b. CME04 experiment

The CME04 experiment took place during summer 2004 in the same forest as the NIWOT02 experiment, but with expanded coverage from three additional above-canopy towers that included the north side of Como Creek [see Burns et al. (2006) for an overview of the CME04 instrumentation]. TGaMS was located at the “Pine” tower with inlets running in a north–south line across Como Creek, vertically up the Pine tower, and along Como Creek about 100 m southeast and southwest of the tower (Fig. 3). Because only a single LI-7000 was available to use with TGaMS, manifolds A and B were connected to each other by a three-way

distribution valve (Fig. 1b). A fourth calibration gas was added to TGaMS by attaching a fifth solenoid valve to the analysis system manifold. The two other CME04 towers (“Aspen” and “Willow”) were equipped with Autonomous Inexpensive Robust CO₂ Analyzers (AIRCOA) that were designed and assembled at the NCAR/EOL Research Aviation Facility (Stephens et al. 2006).

In an effort to minimize inconsistencies between systems, the concentrations of the calibration gases used by NWT, TGaMS, and the AIRCOAs were determined at the NCAR O₂/CO₂ calibration facility (see online at <http://www.eol.ucar.edu/~stephens/CALFAC>) using a Siemens Ultramat-6F NDIR gas analyzer along with 6 primary standards (with CO₂ concentrations determined by both NOAA/GMD and the Scripps Institution of Oceanography) and 12 secondary standards. Six of the secondary gases were used to assign CO₂ concentrations to cylinders with unknown concentrations. The cylinders are all housed horizontally in an insulated enclosure. The system uses active pressure and flow control (MKS Instruments, Wilmington, Massachusetts) to maintain a constant flow rate of 100 standard cubic centimeters per minute (sccm). The air is dried to a dewpoint of -95°C before being measured. The system samples each cylinder for 7.5 min, such that one pass through six known and six unknown cylinders takes 1.5 h. This is repeated 6–8 times in a pyramidal pattern so that the sampling order is varied (Komhyr et al. 1985). The CO₂ concentration of the unknown cylinder is then determined using a third-order polynomial fit of the six known CO₂ concentrations. The assigned CO₂ concentration of the air within the unknown cylinders was reproducible to $\pm 0.05 \mu\text{mol mol}^{-1}$ with an estimated uncertainty of $\pm 0.1 \mu\text{mol mol}^{-1}$ relative to the WMO CO₂ Scale. A postproject calibration of each CME04 cylinder was performed to look for any significant changes to the CO₂ concentrations (none were found).

4. Data-processing details

A list of practical considerations and possible problems associated with running continuous-measuring CO₂ systems can be found in Trivett and Köhler (1999), Miller (2006, his recommendation R9.3) and Stephens et al. (2006, their Table 1). For our study the focus is on the following items: the selection of appropriate CO₂ calibration gases, the form of the equation used to field calibrate the CO₂ output by the LI-7000, how the calibration coefficients are applied to the data between calibration periods, and the travel time of the air sample from the inlet to the buffer volume. Each of the TGaMS data-processing steps are described in detail below.

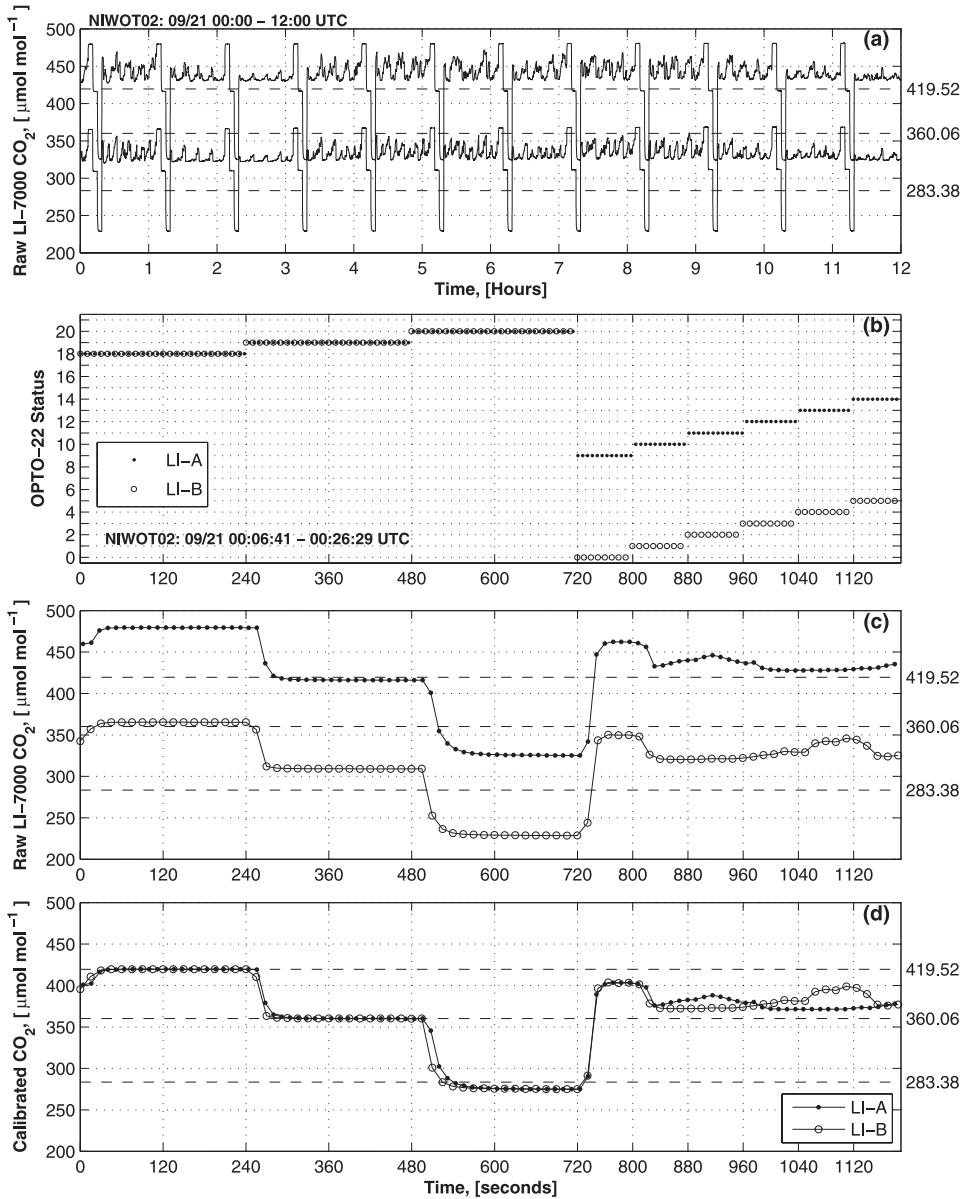


FIG. 5. (a) 12-h time series of LI-7000 LI-A and LI-B raw CO₂ data from NIWOT02. 20-min time series of (b) the Opto 22 control board status signal, (c) LI-A and LI-B raw CO₂ data, and (d) LI-A and LI-B field-calibrated CO₂ data. Note that the calibration gas goes to both LI-A and LI-B at the same time. Between $t = 720$ – 800 s, LI-A and LI-B are sampling from collocated 1-m inlets at the NCAR tower. For clarity not every 1-s data sample is shown; the concentration of the calibration gases are shown as dashed horizontal lines and given to the right of each panel.

a. LI-7000 settings and the field-calibration equation

The LI-7000 instrumentation manual and LI-COR online documentation provide a description of the LI-COR factory calibration and how to apply a field calibration (McDermitt 1997). The LI-7000 uses digital filtering with a range of selectable averaging times (from 0 to 20 s) that reduce the noise in the CO₂ data. If no

filtering is used, the peak-to-peak noise level is around $1.1 \mu\text{mol mol}^{-1}$ (at $370 \mu\text{mol mol}^{-1}$), whereas an averaging time of 20 s reduces the peak-to-peak noise to $0.039 \mu\text{mol mol}^{-1}$. For our experiments the averaging period was set to 1 s, which has a LI-COR-specified peak-to-peak noise level of $0.17 \mu\text{mol mol}^{-1}$. The 1-s averaging time was chosen as a trade-off between response time and noise reduction (for a 1-s averaging

period the delay time is less than 1 s compared to a 10-s delay for a 20-s averaging period).

Why use field calibrations rather than the factory-calibrated CO₂ output from the LI-7000? The LI-COR factory calibration covers a CO₂ range of 0–3000 μmol mol⁻¹ and has been found to drift by several μmol mol⁻¹ depending on changes in cell temperature and pressure. By using our own calibration gases we can measure CO₂ with greater accuracy over the range of CO₂ concentrations specific to our project, and also take into account any environmental changes or other unforeseen factors (e.g., dust getting into a cell or drift in the electrical components) that might affect the factory calibration. The equation used to “field calibrate” the internally calculated CO₂ output by the LI-7000 can be expressed in a general form as

$$\text{CO}_2 = A_3(\text{CO}_2)_{\text{raw}}^3 + A_2(\text{CO}_2)_{\text{raw}}^2 + A_1(\text{CO}_2)_{\text{raw}} + A_0, \quad (1)$$

where (CO₂)_{raw} is the digital CO₂ output from the LI-7000, and A₃, A₂, A₁, and A₀ are empirical calibration coefficients determined from each field calibration time period by fitting (CO₂)_{raw} to the CO₂ concentration of the external calibration gases (Fig. 6). Although the relationship between (CO₂)_{raw} and the known values of the calibration gases may appear linear (Fig. 6a), its slight nonlinearity is shown by their differences (Fig. 6b). As a general guideline for CO₂ measurements, Miller (2006) recommends that the number of calibration gases (N) should exceed the order of the polynomial fit by 2 [e.g., if there are N = 4 calibration gases available, a second-order polynomial should be used for Eq. (1)]. In general, for N data points, a N – 2 order polynomial will not fit the data exactly so that differences, so-called residuals, exist between the data calculated using the polynomial equation and the N data points; a N – 1 order polynomial, however, will always fit N data points with residuals of zero (Devore 1987). Part of our study was to investigate the magnitude of the differences if a linear (N – 3), piecewise linear (N – 1), second-order polynomial (N – 2), or third-order polynomial (N – 1) form of Eq. (1) is used (Table 3).

b. Flushing time of the analysis system

Flushing the previous air sample (or calibration gas) from within the tubing, manifolds, and solenoid valves of the system depends on tubing length and flow rate, and needs to be checked for each unique setup. CO₂ concentrations should only be calculated after flushing has been achieved; therefore, we will designate the time period used for calculating CO₂ as the “sampling time period.” The flushing and sampling times used in

NIWOT02 and CME04 are listed in Table 1. A calibration time period from NIWOT02 is shown in Fig. 7 with the 25-s “sampling time period” marked by vertical lines and a linear fit of the CO₂ data versus time shown as a solid line. (Note that the solid line extends outside the sampling time period for reasons explained below.) If the slope of the “CO₂ versus time” line is near zero it indicates that the stale air has been flushed out. After 240 s of sampling the 283 μmol mol⁻¹ calibration gas, both LI-A and LI-B had a slope larger than –0.18 μmol mol⁻¹ min⁻¹ indicating that stale air was still present. Since the 283 μmol mol⁻¹ gas is at a much lower concentration than the other calibration gases, any flushing problem would be most obvious for this gas. It should also be noted that for each calibration gas, the peak-to-peak range of CO₂ within the sampling time period RNG_{CO₂} was at or below the LI-COR specified peak-to-peak noise of 0.17 μmol mol⁻¹.

Using CME04 data, the flushing time period was systematically investigated by examining changes in the CO₂ statistics as the flushing time was varied for fixed sampling time periods of 10 and 50 s (Fig. 8). Based on the median slope values (Fig. 8b), the vast majority of the air is flushed out by around 40 s; however, the standard deviation of the slope does not level off until a flushing time of around 60 s. In Fig. 8c it is noteworthy that the residuals of the second-order polynomial fit tend to decrease as the flushing time is increased (this was also observed for the other three CME04 calibration gases). This result, coupled with the long flushing time necessary for the 283 μmol mol⁻¹ gas in NIWOT02, indicate that some residual air remained within the system likely due to either slow flushing of dead volumes or surface effects of CO₂ absorbing and desorbing from the internal plumbing surfaces. These flushing issues add to the uncertainty of the CO₂ concentrations assigned to the calibration gases and are an important consideration in relation to the field calibration equation. To evaluate a system for the existence of incomplete flushing, a method recommended by Komhyr et al. (1985) and Trivett and Köhler (1999) is to run the calibrations gases in a “pyramidal” pattern where the sampling order of the calibration gases is constantly being altered (i.e., from highest to lowest concentration during one calibration, and then from lowest to highest concentration for the next, etc). A pyramidal calibration pattern was not used with TGaMS in CME04 (or any of the other TGaMS projects), but will be used in future field projects.

We found that the median of the CO₂ versus time slopes for the 50- and 10-s sampling time periods were similar (Fig. 8b); however, the standard deviation of the slopes is larger for the 10-s sampling time period, which is why slopes from 50-s periods are shown in Fig. 7.

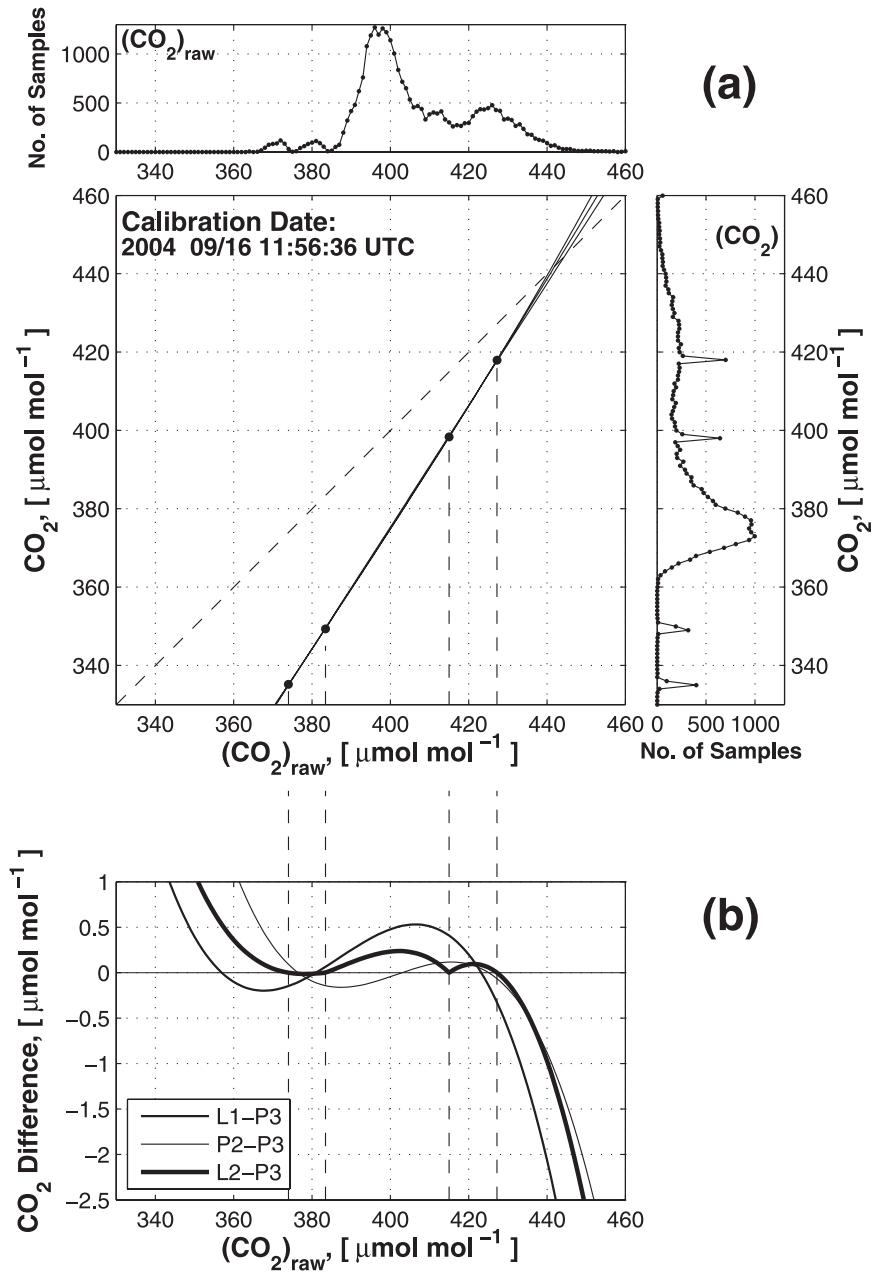


FIG. 6. An example in situ field calibration from CME04 (16 Sep 2004). Filled circles are the calibration points and the vertical dashed lines indicate the value of $(\text{CO}_2)_{\text{raw}}$ for each calibration gas. (a)(large panel) The linear (L1), second-order (P2), and third-order (P3) polynomial fits of the internally calculated CO₂ output by the LI-7000, $(\text{CO}_2)_{\text{raw}}$, to the calibration gas CO₂ concentrations are shown [e.g., see Eq. (1) in text]; also in (a) are the frequency distributions of (top) $(\text{CO}_2)_{\text{raw}}$ and (right) calibrated CO₂ data from all the calibration periods between 10 Aug and 10 Sep 2004. (b) The difference between L1, piecewise-linear (L2), and P2 fits relative to P3 from the same calibration period as in (a) are shown.

c. Determination of the field-calibration coefficients

The CO₂ concentrations of the lowest CME04 calibration gases (335 and 349 $\mu\text{mol mol}^{-1}$) were about 20

and 10 $\mu\text{mol mol}^{-1}$ below the lowest CO₂ concentration observed in the atmosphere and differed from each other by only 14 $\mu\text{mol mol}^{-1}$ (Fig. 6a). If a 375 $\mu\text{mol mol}^{-1}$ gas were included (rather than either the 335 or 349

TABLE 3. Different forms of the CO₂ field-calibration equation [Eq. (1)] that were examined in the postprocessing of the CME04 raw LI-7000 data, (CO₂)_{raw}.

Abbreviation	Equation form	Notes
P3	Third-order polynomial	A_3 – A_0 are determined at each calibration period. The coefficients are linearly interpolated between calibration periods.
P2_con	Second-order polynomial	$A_3 = 0$. A_2 – A_0 are determined at each calibration period. These coefficients are used until the subsequent calibration period.
P2	Second-order polynomial	Same as P2_con, but the A_2 – A_0 coefficients are linearly interpolated for the time between the calibration periods.
L1	Linear	$A_3 = A_2 = 0$. A_1 and A_0 are determined by a linear fit for each calibration time period using all four calibration gases.
L2	Piecewise linear	$A_3 = A_2 = 0$. A_1 and A_0 are determined by a linear fit for each calibration time period and between each pair of calibration gases. This method will result in a unique set of coefficients for each pair of calibration gases.

$\mu\text{mol mol}^{-1}$ gas), it would have created a more evenly spaced distribution among the calibration gases, and added a gas that was approximately the median CO₂ concentration of the air sampled throughout the project. It also would have also been better if the 418 $\mu\text{mol mol}^{-1}$ calibration gas was closer to 450 $\mu\text{mol mol}^{-1}$ so the highest calibration gas would have bracketed the largest observed CO₂ concentration.

Frequency distributions of the differences between the calibration gas concentration and the calculated CO₂ data from 893 CME04 calibration time periods are shown in Fig. 9. These distributions are the 1-Hz CO₂ data from the sampling time periods during the calibrations (i.e., the distributions include any peak-to-peak noise that exists in the CO₂ data). Not surprisingly, the linear calibration that spans all four calibration gases (L1 in Table 3) shows the largest residuals (on the order of $\pm 0.3 \mu\text{mol mol}^{-1}$) while the third-order polynomial and piecewise linear equations (P3 and L2 in Table 3) have distributions that are centered on zero with peak-to-peak noise between $\pm 0.1 \mu\text{mol mol}^{-1}$ (both P3 and L2 are fits of order $N - 1$ so it is expected that these distributions closely agree with each other, and are more “peaked” than the other distributions). Although the third-order polynomial fits the data exactly, it is highly nonlinear outside the range of the calibration points. With a $N - 2$ order fit (P2 in Table 3), the absolute value of the residuals in CME04 was around 0.1 $\mu\text{mol mol}^{-1}$. There are several potential sources of measurement error that cause the residuals: higher-order nonlinear response of the IRGA, uncertainty in the CO₂ concentration assigned to each calibration gas, dead volumes, and effects of pressure changes inside the cylinder and regulator that cause CO₂ to absorb–desorb from surfaces (Trivett and Köhler 1999). For a LI-7000 in the laboratory at NOAA/GMD, the residuals from 496 calibrations using a second-order calibration equation were found to be smaller than $\pm 0.02 \mu\text{mol mol}^{-1}$ with a root-mean-square value of 0.011 $\mu\text{mol mol}^{-1}$ [C.

L. Zhao 2007, personal communication; also see Zhao and Tans (2006)], which is a factor of 5 smaller than the residuals we found in the field with the TGaMS LI-7000. The reason for the discrepancy between the residuals with our LI-7000 and those from Zhao is not known, however, the IRGA setup in the laboratory likely had an advantage of minimal dead volumes and a slowly varying (or controlled) temperature environment. Although these features could be added to a field-based instrument, they were not used within the TGaMS system. To minimize dead volumes, a rotary valve system could be used rather than solenoid valves and manifolds; rotary valves, however, are more sensitive to dirt in the lines, tend to leak after extensive use, and are more expensive (Trivett and Köhler 1999). It is also possible to have a design that flushes the manifolds in a more efficient manner by drawing air from both sides of the manifold rather than from just one end (Bakwin et al. 1995; Stephens et al. 2006).

We now consider the following question—is it better to fit the N calibration data points with a $N - 1$ or $N - 2$ order polynomial? If there were absolutely no systematic errors in the assigned calibration gas concentrations, the gas from each calibration cylinder was delivered cleanly (e.g., no errors due to incomplete flushing, etc.), and the true instrument response was not captured by a $N - 2$ order fit, then a $N - 1$ order fit could be considered. One might consider the residuals introduced by a $N - 2$ order fit to be a negative aspect of the calibration procedure because residuals are transferred to the final, calculated CO₂; however, rather than being detrimental, the magnitude of the residuals are actually a valuable indication of the “health” of a sampling system as well as an indication of the accuracy of the calibration gases. Because other IRGA systems have shown that the residuals can be very small [on the order of 0.02 $\mu\text{mol mol}^{-1}$ (Zhao and Tans 2006) or 0.07 $\mu\text{mol mol}^{-1}$ (van der Schoot et al. 2006)], the residuals on the order of 0.1 $\mu\text{mol mol}^{-1}$ from CME04 may be an indication of dead

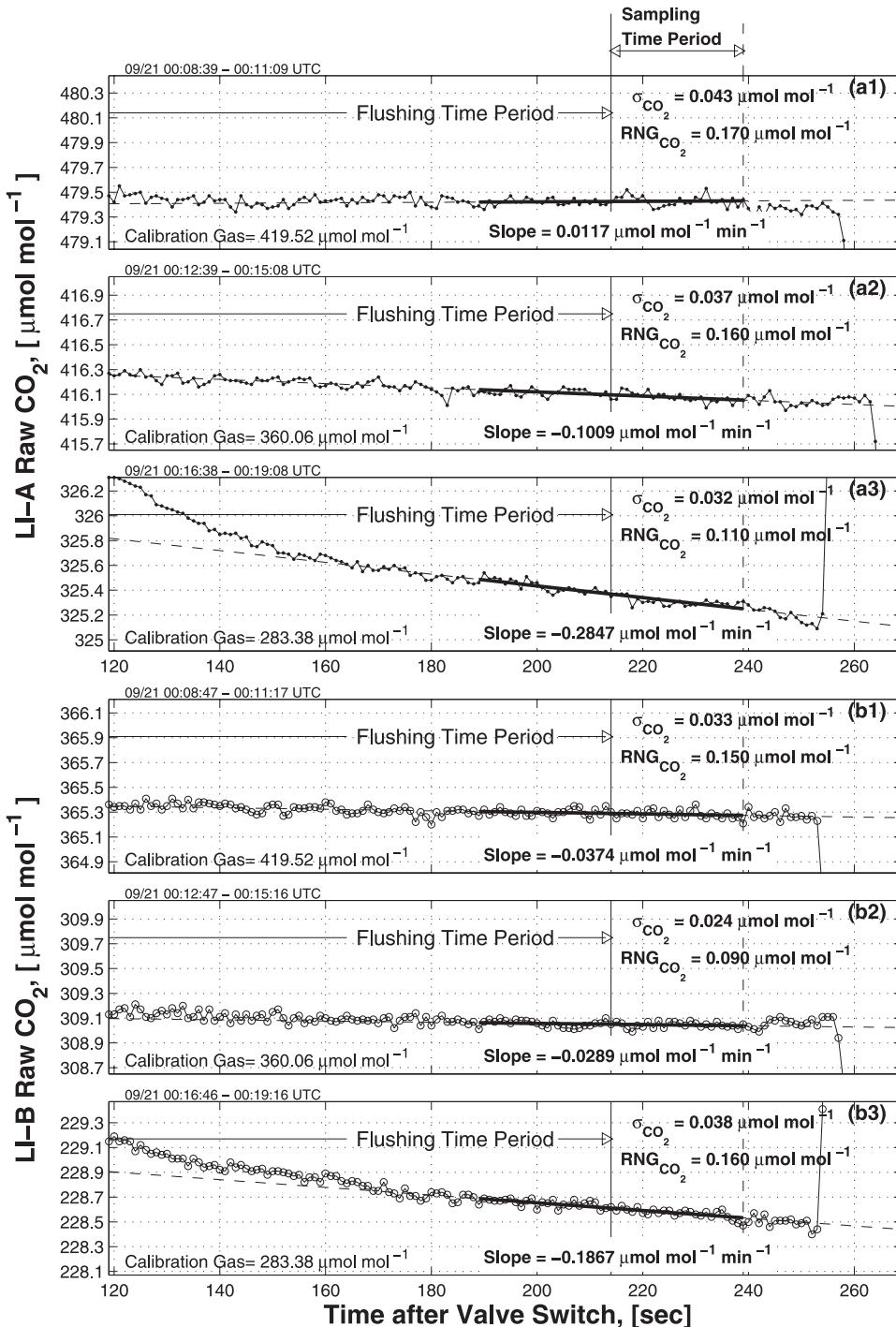


FIG. 7. Time series from NIWOT02 of the raw CO_2 output by (a1)–(a3) LI-A and (b1)–(b3) LI-B during time periods when the 419, 360, and 283 $\mu\text{mol mol}^{-1}$ calibration gases were being sampled (concentration of the calibration gas are indicated in the bottom left). The data used to calculate the mean CO_2 are indicated by the vertical lines and labeled, sampling time period. The standard deviation (σ) and peak-to-peak range (RNG) of the CO_2 from the sampling time period are shown within each panel along with the slope from a linear fit of CO_2 vs time (as indicated by the solid line). The flushing time period starts when the valve is switched (i.e., at $t = 0$). For NIWOT02 the flushing time was 215 s and the sampling time was 25 s (Table 1).

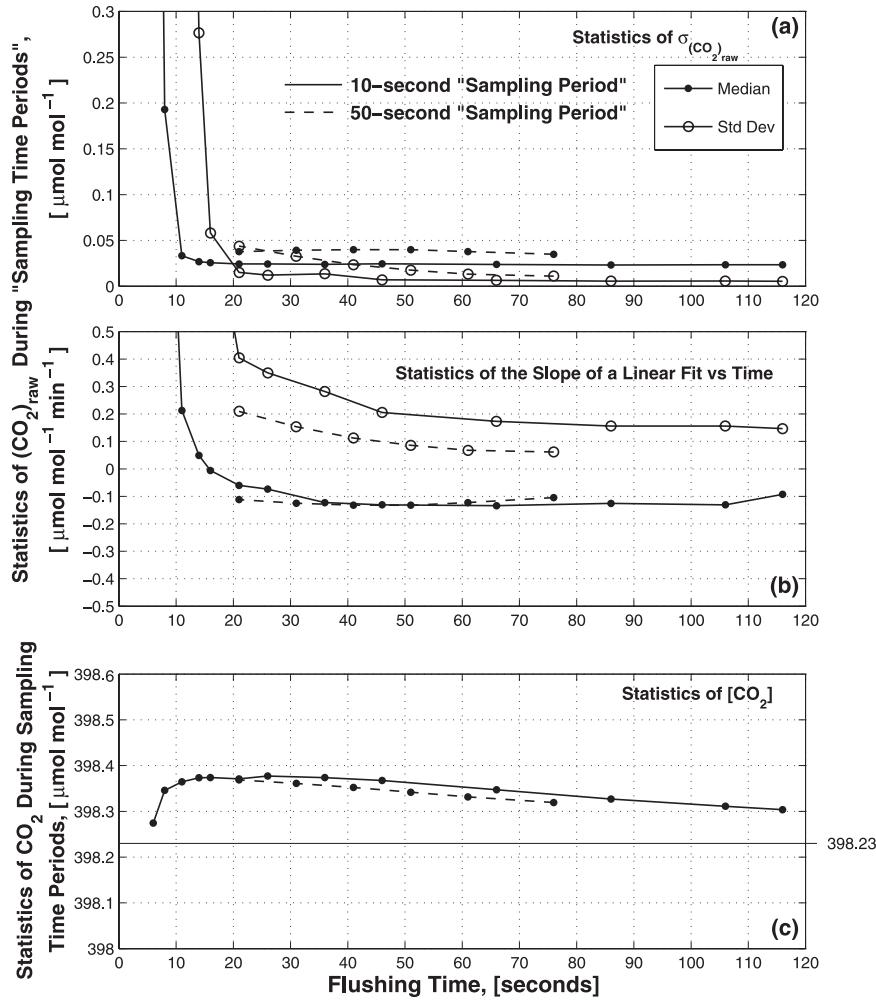


FIG. 8. Statistics (median and standard deviation) from 283 different sampling time periods for the $398 \mu\text{mol mol}^{-1}$ calibration gas as a function of flushing time for (a) the standard deviation (σ) of $(\text{CO}_2)_{\text{raw}}$ and (b) the slope of $(\text{CO}_2)_{\text{raw}}$ vs time within the sampling time period. (c) Median of the calibrated CO_2 data using a second-order polynomial (P2) form of the field-calibration equation. The solid and dashed lines indicate a 10- or 50-s sampling time period, respectively. Data shown are from CME04 between days 220 and 251; other calibration gases show similar results.

volumes affecting our ability to accurately measure the gas concentrations.

The CO_2 community recommends that independent monitoring of the CO_2 accuracy be performed by sampling from a long-term “surveillance” gas cylinder (i.e., a cylinder containing CO_2 in air that is not used in the calibration, but is sampled periodically by the system; Bakwin et al. 1995; Zhao et al. 1997; Trivett and Köhler 1999; Stephens et al. 2006; Miller 2006). The surveillance gas is supposed to last beyond the time period when the regular calibration gases are changed, to keep track of any changes that might be caused by changing a calibration gas or the introduction of a gas with a error in the assigned concentration value. The disadvantage

of the surveillance gas is that it only checks the calibration curve at one specific CO_2 concentration. There was no independent surveillance gas available with the TGAMS system so we are not able to evaluate this aspect of the field calibration.

For NIWOT02, the initial attempt at data processing used all three calibration gases and a linear form of Eq. (1); however, when using all three gases, the calculated CO_2 did not agree well with the $360 \mu\text{mol mol}^{-1}$ calibration gas. Because the 360 and $419 \mu\text{mol mol}^{-1}$ concentration gases nearly bracketed the range of CO_2 observations during the experiment, it was decided not to include the $283 \mu\text{mol mol}^{-1}$ concentration gas in the calibration. The time series of the linear coefficients [A_1

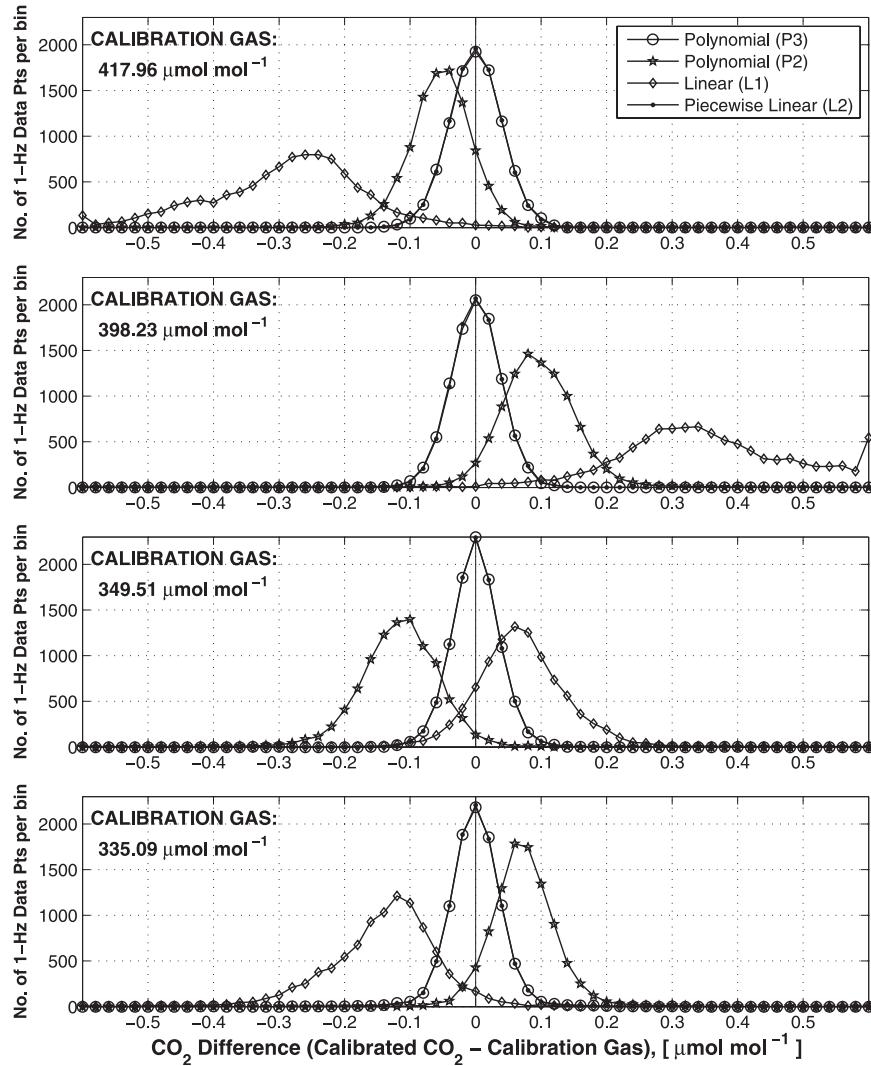


FIG. 9. Frequency distribution of the difference between the field-calibrated CO_2 data and the calibration gas concentration during the calibration gas sampling time periods. Data are from CME04 between 7 Aug and 4 Oct 2004. Different symbols represent different forms of Eq. (1) as detailed in Table 3: (a),(b) LI-A (a) A_1 and (b) A_2 , (c) P , (d) T , and (e),(f) LI-B (e) A_1 and (f) A_2 .

and A_0 in Eq. (1)] and changes in cell temperature T_{cell} and pressure P_{cell} are shown in Fig. 10. The T_{cell} experiences the same diurnal temperature changes as the atmosphere, but stays about 10°C warmer. The atmospheric pressure and P_{cell} closely follow each other because the pump is upstream of the sample cell. For LI-A, the changes in T_{cell} caused A_1 to vary (by less than 1%), and A_0 to change by around $2 \mu\text{mol mol}^{-1}$ (to compensate for the change in A_1). These variations in CO_2 with temperature are comparable to the quoted value of $\pm 0.3 \mu\text{mol mol}^{-1} \text{ }^\circ\text{C}^{-1}$ (LI-COR Inc. 2005). For LI-B, the connection between A_1 and A_0 was less obvious and both terms appear to have a slight drift. There was a

jump in the LI-B offset value that occurred around day 259 (16 September 2002). This was a day LI-A and LI-B were visited with a set of portable calibration gas cylinders (section 5b), and it is possible something caused the calibration to change slightly when the portable gas cylinder was attached and removed from LI-B. This is an example of why it is useful to use a field calibration rather than rely on the factory calibration.

d. Application of the field-calibration coefficients to LI-7000 data

Typical methods for extending the calibration coefficients to the time between the calibration periods

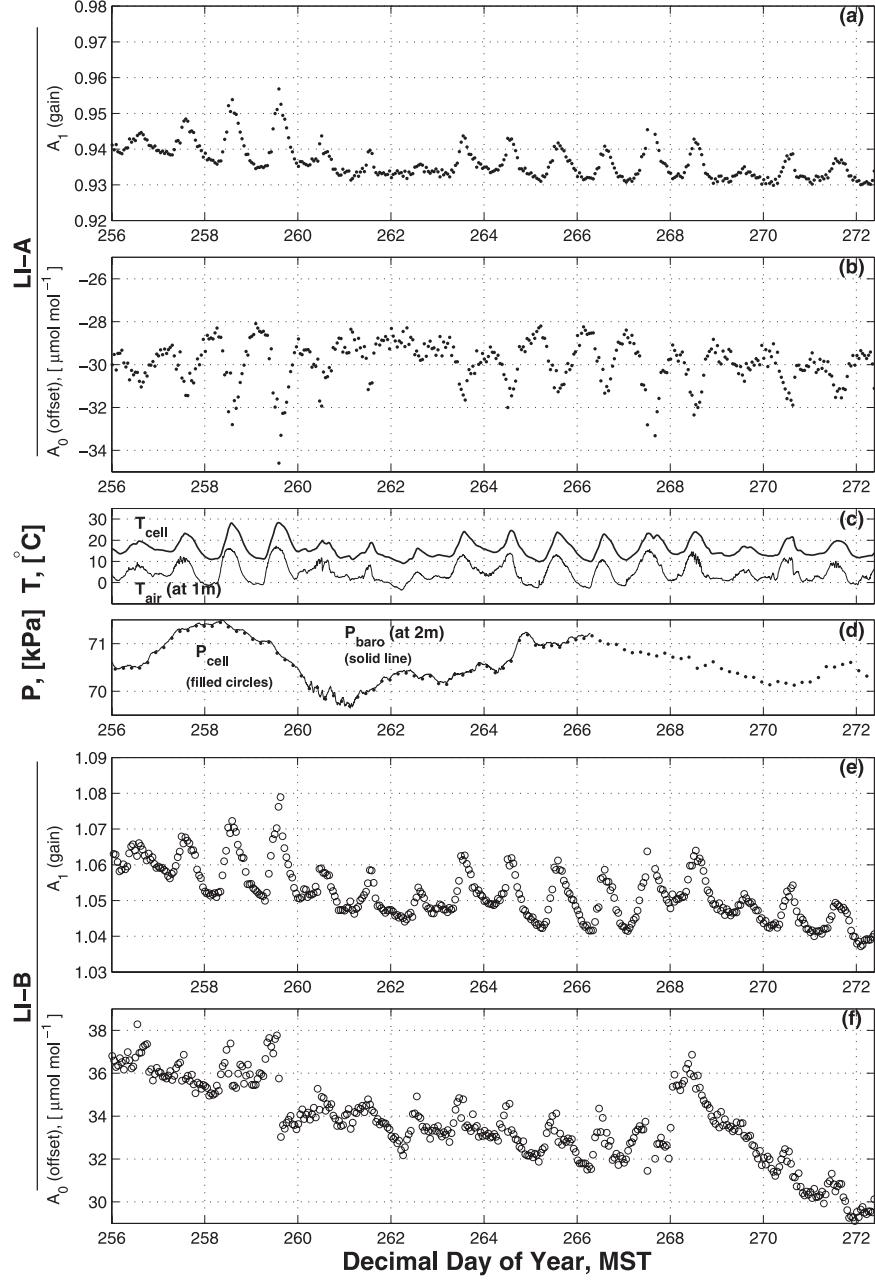


FIG. 10. Calculated calibration coefficients from NIWOT02 of the gain (A_1) and offset (A_0) from (a),(b) LI-A and (e),(f) LI-B. Each point represents the results of a linear calibration from one calibration period that occurred every hour. (c) The air temperature (T_{air}) and LI-A cell temperature (T_{cell}) and (d) the atmospheric pressure (P_{baro}) and LI-A cell pressure (P_{cell}). Note that second-order and third-order polynomials were not used in NIWOT02.

used in previous studies are linear interpolation of the coefficients (Zhao et al. 1997) or using the average between calibration periods (Monson et al. 2002). If there is not too much change in the calibration coefficients between calibrations, one could simply apply the coefficients from each calibration until the following data cycle. However, we found that during

NIWOT02 the drift in the calibration coefficients was too large to maintain an acceptable level of accuracy in the calibrated CO₂ data (to better than $\pm 1 \mu\text{mol mol}^{-1}$). To account for any change in the calibrations between calibration periods, the calibration coefficients were linearly interpolated and then applied to the (CO₂)_{raw} data.

e. Air sample travel time and data interpolation

The time for an air sample to travel from the inlet to the LI-7000 sample cell was determined by one person exhaling into the inlet at the command of the other person who monitored the LI-7000 with a stopwatch. The travel time of the pulse of CO₂-rich breath was recorded. Typical delay times were on the order of 70 s for a 100-m length of tubing, and the results from NIWOT02 are listed in Table 2. Rather than checking every inlet in NIWOT02 only several were checked and the rest of the delay times were estimated based on the relationship between tubing length and delay time. A similar procedure was performed for CME04 except every inlet was tested. The lag-time-check exercise serves as an excellent way to confirm the consistency among the inlets, the connection at the HYDRA box, and the software.

The final steps in the data processing are to take into account the different travel times of the CO₂-laden air for different tubing lengths, and then linearly interpolate these time-lag-corrected CO₂ data to a common time stamp. During periods of rapidly changing CO₂ the interpolation of the CO₂ data to a common 15-min time stamp tends to smooth out the high and low values in the original CO₂ time series. As long as the time period of the common time stamp is not too much larger than the time period between measured data samples, the interpolation should not add any significant systematic error to the measurements. However, it would be impossible to detect any high-frequency movement or change in the CO₂ moving past the array of inlets with a period on the order of 15 min or shorter.

5. Results

The effect of using different forms of the field-calibration equation (as described in section 4) and methods of data processing on the final calculated CO₂ data are presented within this section. We also present in situ CO₂ field comparisons that were found to be useful in evaluating data quality.

a. Effect of the field-calibration equation on the calculated CO₂

Based on the discussion in section 4c (e.g., Figs. 6 and 9) we already have some insight into how the CO₂ data calculated with the different calibration equations will compare with each other. Figure 11b shows that the CO₂ differences among the various calibration equations were within 0.2 $\mu\text{mol mol}^{-1}$ of each other near the calibration gas concentrations. However, at CO₂ con-

centrations much greater than the highest calibration gas (418 $\mu\text{mol mol}^{-1}$), the linear (L1) and third-order polynomial (P3) equations quickly diverge (with differences over 1 $\mu\text{mol mol}^{-1}$) from the CO₂ calculated with the piecewise linear (L2) and second-order polynomial (P2) equations. The divergence is due to the range of the CO₂ measurements not being properly bracketed by the calibration gases.

An overall summary of the CO₂ differences among the methods used in our study is shown by frequency distributions with the median difference and standard deviation of the difference for each combination listed in the top-left corner of Fig. 11c. Since most of the atmospheric CO₂ measurements were within the 370–390 $\mu\text{mol mol}^{-1}$ range, the results in the frequency distributions will be heavily weighted to the differences at this CO₂ concentration, and is why the divergence in P3 and L1 at large values of CO₂ do not have a large impact on the frequency distribution of the difference. The frequency distribution of the P2_con-L2 difference (standard deviation = 0.53 $\mu\text{mol mol}^{-1}$) is more spread out than the difference of P2-L2 (standard deviation = 0.31 $\mu\text{mol mol}^{-1}$), which indicates the importance of linearly interpolating the coefficients between adjacent calibration periods.

b. Empirical data quality tests

The CO₂ measured by TGaMS was evaluated for accuracy, reproducibility, and comparability by empirical CO₂ comparisons among a set of portable cylinders, collocated inlets attached to independent CO₂-measuring instruments, and CO₂-measuring systems that are located near each other. For CME04, the TGaMS CO₂ data used in these comparisons are from the P2 field-calibration equation (Table 3).

During NIWOT02 a set of four 6-L calibration cylinders were prepared and brought to each CO₂-measuring system involved in the project. The CO₂ concentrations assigned to each portable gas cylinder (353, 369, 392, and 416 $\mu\text{mol mol}^{-1}$) were determined by comparison to gases from NOAA/GMD, and selected to span a large portion of the CO₂ range measured in the field. The comparison between LI-B and the portable cylinders is the most consistent (Fig. 12). Results for LI-A were similar to LI-B with a slightly larger variability (primarily due to the comparison on 19 September, which should probably not be considered because someone breathed into the A1 and B1 inlets causing the CO₂ to briefly go over 20 000 $\mu\text{mol mol}^{-1}$ 10 min before the LI-A comparison, and 20 min before the LI-B comparison). Based on these few comparisons, the CO₂ measured by LI-A and LI-B was 1–2 $\mu\text{mol mol}^{-1}$ larger than the CO₂ concentration within the portable cylinders.

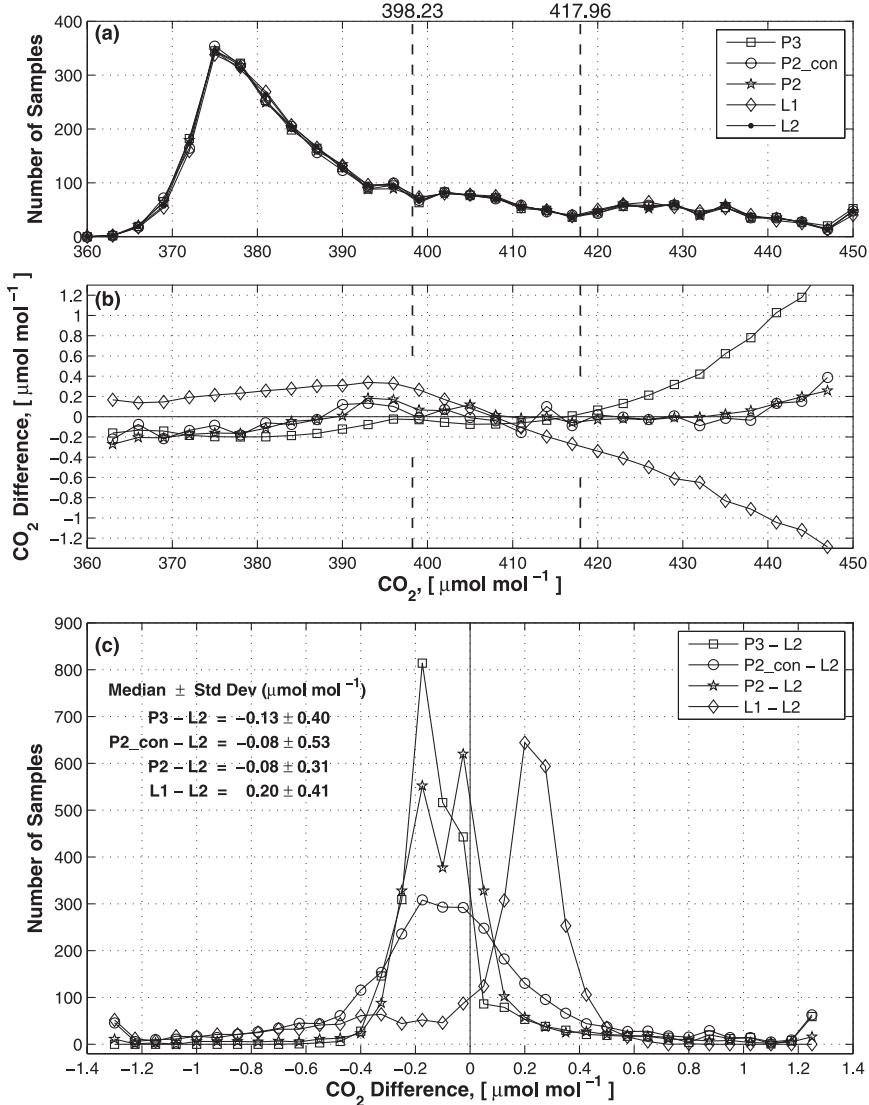


FIG. 11. Plots of (a) the frequency distribution of CO₂ calculated using different forms of the field-calibration equation (each form is identified in the legend and described in Table 3), (b) the difference in CO₂ calculated with each calibration equation [relative to the piecewise linear L2 form of Eq. (1)], and (c) the differences from (b) shown as a frequency distribution (the median difference \pm the standard deviation of the difference for each combination are shown in the upper left). The dashed vertical lines in (a) and (b) are the CO₂ concentration of the calibration gases (note that the two lower gases are beyond the x limits of the plot). These are CO₂ data from the 1-m inlet (A7) at the Pine tower during CME04 between days 220 and 280.

It is possible that a CO₂ bias was introduced by the method used to determine the CO₂ concentration of the TGaMS calibration gases (described in section 3a). Over the range of portable cylinder CO₂ concentrations, the CO₂ difference varies on the order of 0.3–0.4 $\mu\text{mol mol}^{-1}$, which is consistent with the magnitude of differences found in the evaluation of the different field-calibration equations in CME04 (recall that a linear field-calibration equation was used for NIWOT02). The

comparisons of the portable cylinders with NWT and USGS were more variable, perhaps due to the fact that the USGS and NWT systems used the LI-COR factory calibration and only two calibration gases (as described in section 3). Also, the calibrations of the NWT/USGS systems were performed less frequently than LI-A and LI-B (NWT calibrations were every 4 h, and USGS calibrations were every 2 h compared to every hour for TGaMS). On 18 September 2002 the calibration gas in

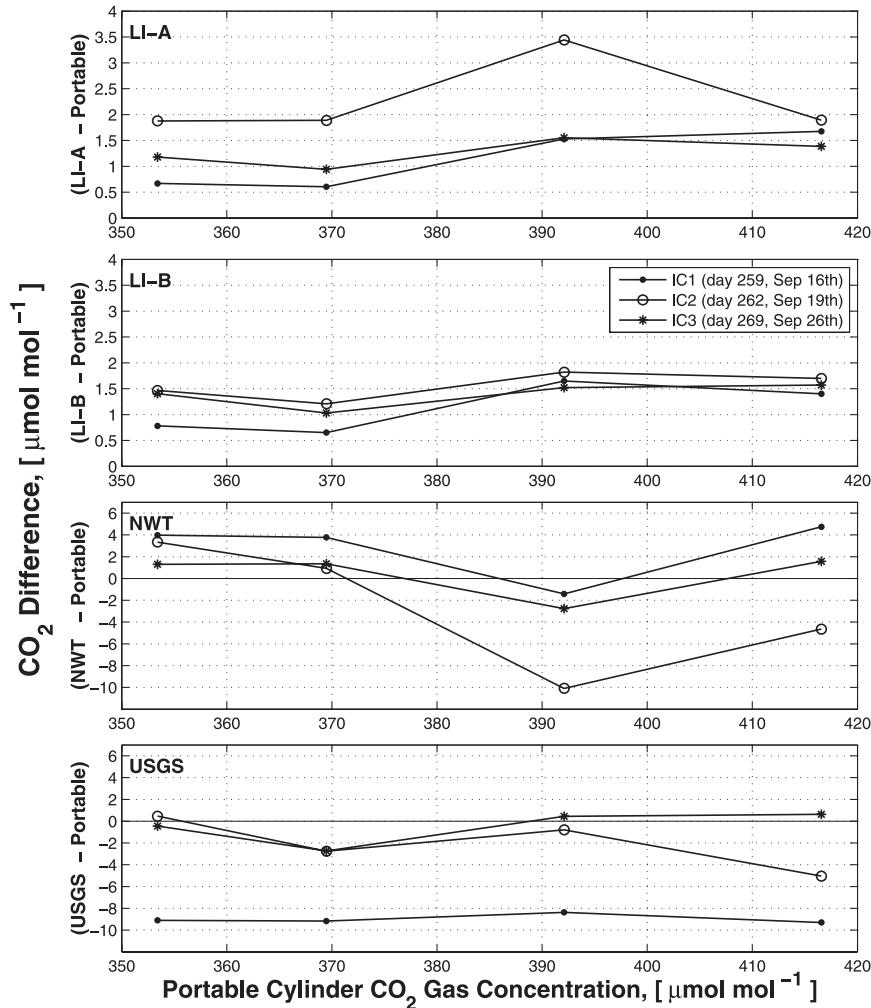


FIG. 12. Differences in CO₂ between each continuous CO₂-measuring system—(top to bottom) LI-A, LI-B, NWT, and USGS—and the four transportable gas concentrations during NIWOT02 (note that the range of the y axis for the top two panels differs from the bottom two panels). The legend in the second panel identifies the date of the comparison. The IC2 comparison on 19 September for LI-A was affected by someone exhaling into the LI-A and LI-B collocated 1-m inlet 10 min before attaching the portable cylinder to the LI-A system. Each CO₂-measuring system at the site was equipped with a three-way valve to allow the gas from the portable cylinders to be run through the system following the same pathway that the calibration gas would normally use (the line pressure with the portable calibration gases attached was matched as best as possible to that used by the calibration cylinders at each system).

the USGS system was changed, which is the reason the first comparison was so different than the later ones.

Time series of 1-m CO₂ from the NIWOT02 collocated inlets shows that each CO₂-measuring system captured a similar diurnal CO₂ pattern; however, there were some significant differences between the various systems (Fig. 13). The median and standard deviation of the CO₂ difference between the A1 and B1 inlets was $-0.02 \pm 0.42 \mu\text{mol mol}^{-1}$ (Fig. 14e1), whereas the differences between A9 and NWT ($0.66 \pm 2.7 \mu\text{mol mol}^{-1}$), and A8 and USGS ($-3.99 \pm 1.7 \mu\text{mol mol}^{-1}$) were

larger and more variable. It is not too surprising that A1–B1 agree better than NWT–A9 and USGS–A8, since LI-A and LI-B used the same calibration gases, were situated near the inlets (less than 5 m away), and the CO₂ data were processed in a similar manner. A difficulty with this comparison was the long tubing lengths used to reach the NWT and USGS towers. The air sample within the long tubing undoubtedly underwent additional mixing that likely has a detrimental effect when trying to compare CO₂ on a point-by-point basis over fairly short (order of 30 min) time scales. CO₂

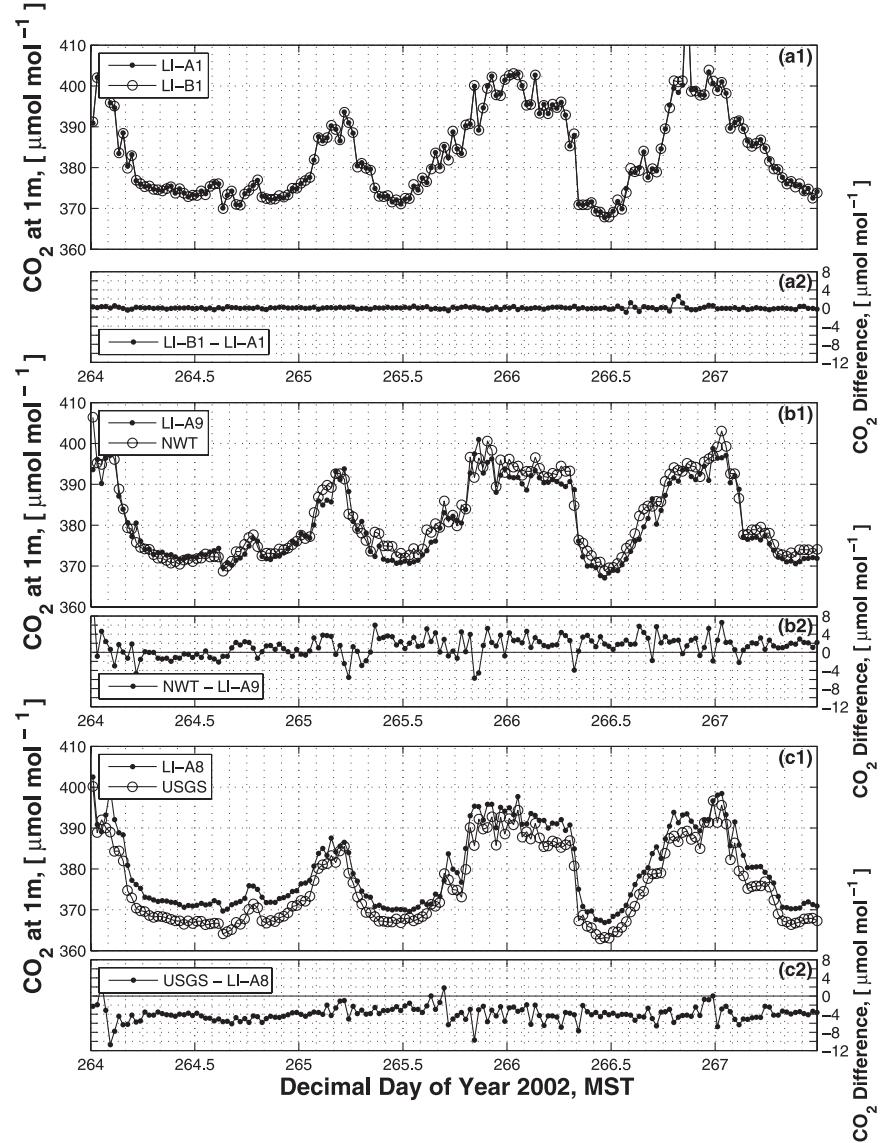
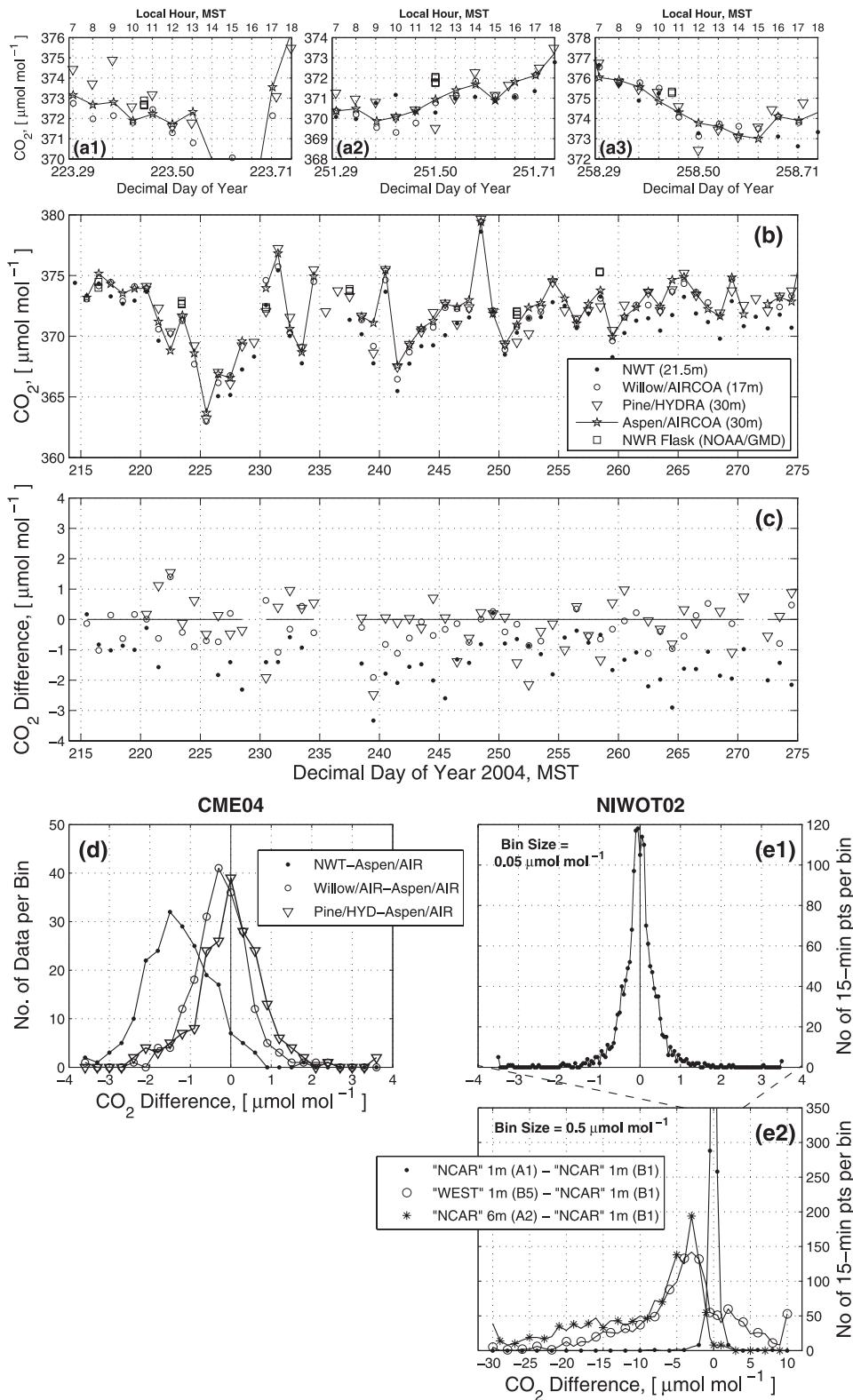


FIG. 13. Time series of half-hourly 1-m CO₂ for collocated HYDRA inlets between (a1),(a2) the two TGaMS LI-7000s (LI-A and LI-B); (b1),(b2) NWT tower and LI-A; and (c1),(c2) USGS North tower and LI-A. Each top panel shows the mean data and the bottom panel, the CO₂ difference between the measurements.

reproducibility achieved with TGaMS is demonstrated by comparing the 1-m CO₂ measured from inlet B1 at the NCAR tower with CO₂ from the collocated 1-m inlet (A1–B1), the 1-m inlet at the West tower (B5–B1), and the 6-m inlet at the NCAR tower (A2–B1) (Fig. 14e2). From these frequency distributions it can be observed that the CO₂ difference between the collocated inlets were all between $\pm 1 \mu\text{mol mol}^{-1}$, whereas the horizontal and vertical CO₂ differences observed in the environment ranged from -30 to over $10 \mu\text{mol mol}^{-1}$ and are what one would expect to find (e.g., CO₂ at 1 m was always greater than CO₂ at 6 m, and CO₂

upslope of the NCAR tower was usually smaller than CO₂ at the NCAR tower, but the horizontal difference was of either sign). A detailed examination of CO₂ in NIWOT02 can be found in Sun et al. (2007).

In CME04, there were no collocated inlets, so the midday 1-h median from the upper level at each tower is used to cross compare CO₂ from the various systems. Hourly CO₂ time series from the top of each tower along with NOAA/GMD CO₂ air samples collected with a flask system at the NWR site on Niwot Ridge (Conway et al. 1994) are shown in Figs. 14a1–a3. The NWR air-collection site is above tree line at 3523 m and



3.4 km to the northwest of the CME04 forest towers and is included to give an estimate of how much CO₂ spatial variability existed. The hourly panels show the temporal variability on a given day. A comparison of the midday CO₂ measured at the top of the NWT tower with a tunable diode laser absorption spectrometer and the NWR flask CO₂ has been presented previously by Bowling et al. (2005, their Fig. 9), who found that CO₂ differences during summer 2003 were on the order of 0–4 $\mu\text{mol mol}^{-1}$. For CME04, the NWR CO₂ flask measurements were generally within 1 $\mu\text{mol mol}^{-1}$ of the tower CO₂ measurements indicating the high degree of atmospheric mixing present within the convective daytime boundary layer; however, local differences among the CME04 towers, were also sometimes on the order of 1 $\mu\text{mol mol}^{-1}$. It should be noted that the NWR flask samples were sometimes collected closer to 1100 MST rather than 1200 MST, so the time series of NWR and the CME04 tower CO₂ in Fig. 14b do not always have a consistent time stamp. To get an empirical estimate of the degree of CO₂ comparability among the various systems, the time series and frequency distribution of the midday CO₂ differences are plotted in Figs. 14c,d, and the median differences and standard deviations are: $0.04 \pm 0.81 \mu\text{mol mol}^{-1}$ (Pine – Aspen), $-0.27 \pm 0.73 \mu\text{mol mol}^{-1}$ (Willow – Aspen), and $-1.33 \pm 0.94 \mu\text{mol mol}^{-1}$ (NWT – Aspen). It should be emphasized that these statistics include differences due to local meteorology and different underlying surfaces; the highest inlet at Willow was at 17 m surrounded by low shrubs, whereas the 30-m inlets at Pine and Aspen and the 21.5-m inlet at NWT were located above an 11-m-tall subalpine forest. The CO₂ measured by NWT was typically about 1–2 $\mu\text{mol mol}^{-1}$ lower than the other three systems and this systematic bias is probably due to the limitation of using two calibration gases and the LI-COR factory calibration.

6. Conclusions

The construction of a portable, 18-inlet, CO₂-measuring system (TGaMS) used in two different field projects was

described. The air-sampling system, called HYDRA, was attached to either one or two analysis systems that used a LI-COR model LI-7000 to analyze the CO₂ concentrations of the air samples. To improve the accuracy of the measured CO₂, *in situ* “field calibrations” (i.e., short time periods when a series of high-accuracy calibration gases were passed through the LI-7000 sample cell) were used by fitting the raw CO₂ data output from the LI-7000 to the assigned CO₂ concentration of the calibration gases. CO₂ calculated using a third-order polynomial, second-order polynomial, linear, and piecewise linear form of the field-calibration equation [Eq. (1)] were compared. For $N = 4$ calibration gases, the second-order polynomial had residuals of the fit that were on the order of $\pm 0.1 \mu\text{mol mol}^{-1}$, whereas a linear fit had residuals on the order of $\pm 0.3 \mu\text{mol mol}^{-1}$ (Fig. 9). Part of our study explored the possibility of using a third-order fit to eliminate the residuals. Although the CO₂ concentrations from a third-order fit were within 0.2 $\mu\text{mol mol}^{-1}$ of the CO₂ calculated from the second-order fit, using a third-order fit was rejected for the following reasons: 1) any errors in the CO₂ concentrations assigned to the calibration gases are fit exactly with a $N - 1$ fit, 2) any dead volume or incomplete flushing of the system could bias the assigned CO₂ concentration, 3) the highly nonlinear nature of the $N - 1$ order fit makes extrapolation of the field-calibration curve inadvisable, and 4) although the $N - 1$ order fit was fairly well behaved for the calibration gas concentrations used in CME04, certain combinations of gas concentrations can produce nonlinear results between the calibration points. Rather than eliminating the residuals of the fit, we concluded that residuals act as an indicator of system integrity and large residuals might indicate some problem within the system or with a particular calibration gas. A few changes to TGaMS (listed below) might help to reduce the magnitude of the TGaMS residuals in future experiments.

Empirical comparisons were used to estimate how well TGaMS measured CO₂ during various field projects. In NIWOT02, a pair of side-by-side inlets (with one inlet attached to each LI-7000) was used to empirically



FIG. 14. Time series from CME04 of: (a1)–(a3) hourly CO₂ measurements for selected days when NOAA/GMD NWR flask measurements were collected (see text for NWR details), (b) midday (1130–1230 MST) median CO₂ measured at the top of each tower between 1 Aug to 30 Sep 2004, and (c) the difference in CO₂ between each tower (the 30-m CO₂ measured by the Aspen AIRCOA is used as the reference). The legend in (b) applies to (a1)–(a3) and (c). (d) The frequency distribution of the CO₂ difference among the CME04 towers is shown (based on 196 hourly median values between 1000 and 1500 MST). (e1) The frequency distribution of the CO₂ difference between the collocated 1-m inlets of LI-A and LI-B from NIWOT02; (e2) an expanded view of (e1), along with other inlet pairs as specified in the legend (these distributions are based on 1593 15-min CO₂ values collected over 17 days).

estimate the TGaMS CO₂ reproducibility; the CO₂ difference between the collocated inlets had a median and standard deviation of $-0.02 \pm 0.42 \mu\text{mol mol}^{-1}$ over 17 days (Fig. 14e). During NIWOT02 four portable cylinders with gas concentrations closely linked to the WMO CO₂ Scale were periodically sampled by TGaMS. These comparisons revealed a positive bias of $1\text{--}1.5 \mu\text{mol mol}^{-1}$ in the TGaMS CO₂ data, which is most likely due to the way CO₂ concentrations were assigned to the calibration gases in NIWOT02 (section 3a). Therefore, in NIWOT02 we conclude that TGaMS ran in a mode with acceptable reproducibility between inlets, but was not a high-accuracy system relative to the WMO CO₂ Scale.

In CME04, two high-accuracy CO₂-measuring AIRCOA systems (Stephens et al. 2006) were deployed within 400 m of TGaMS. The calibration gases used with TGaMS and AIRCOA were assigned CO₂ concentrations with an estimated uncertainty of $\pm 0.1 \mu\text{mol mol}^{-1}$ relative to the WMO CO₂ Scale by the NCAR O₂/CO₂ Calibration Facility. The empirical assessment of how well TGaMS measured CO₂ was based on midday, above-canopy CO₂ comparisons. The 30-m CO₂ measured by TGaMS (at Pine tower) and AIRCOA (at Aspen tower) had a median difference and standard deviation of $0.04 \pm 0.81 \mu\text{mol mol}^{-1}$ based on data between 1000 and 1500 MST over 2 months. This result includes real variability in the CO₂ difference, but is comparable to the differences measured between the two AIRCOA systems (Fig. 14d). Therefore, we conclude that TGaMS in CME04 had improved CO₂ accuracy compared to NIWOT02.

For others interested in making highly-accurate CO₂ measurements (relative to the WMO Scale) we recommend the following guidelines: 1) high-accuracy calibration gases must be used, 2) calibration gases should fully span the expected range of environmental CO₂ and be approximately equally spaced within the range, 3) a long-term “surveillance” gas should be used to independently evaluate the system performance, 4) flushing of the system should be checked by examining the slope of “CO₂ versus time” during the calibration periods, and 5) the order of sampling the calibration gases should be varied to detect any flushing issues. Hardware changes that would improve TGaMS are as follows: 1) active temperature control of the LI-7000; 2) minimize dead volumes by better flushing of the system (this can be achieved by a redesign of the manifold as well as using a higher flow rate for flushing); 3) minimize internal plumbing surfaces that absorb–desorb CO₂; 4) develop a systematic way to check for tubing damage (i.e., cracks or holes), as well as leaks in valves, rotameters, fittings, etc. (as an example of this, in CME04 chunks of dry ice were dropped into the HYDRA box

and revealed a leak in one of the internal connections for one inlet; such a method works for spot checks, but a more comprehensive leak-checking methodology should be developed); 5) measure the pressure, temperature, and humidity of the incoming air sample so that any blockage due to ice or debris can be detected; and 6) tubing should be insulated and/or heated to minimize condensation (or ice) formation on the inner walls of the tubing.

Acknowledgments. We thank Russ Monson and Thomas Foken for supporting experiments with TGaMS. Dean Anderson provided CO₂ data from the USGS tower. Pieter Tans, Cong Long Zhao, Thomas Conway, and Duane Kitzis from NOAA/GMD generously provided helpful advice, information about NOAA/GMD laboratory calibrations, and the NWR flask data. Dan Anderson from LI-COR supplied useful information. We gratefully acknowledge discussions and field help by Jeff Beauregard, Dave Bowling, Pete Casey, Jia Hu, Kurt Knudsen, Don Lenschow, Kirk Ranno, Kelly Sensecqua, Christoph Thomas, Andrew Turnipseed, Andy Watt, Mike Weintraub, and Chuixiang Yi. We also thank two anonymous reviewers for their helpful comments. This work was supported by the NCAR Director Opportunity Fund and NSF Grant EAR-0321918.

REFERENCES

- Aubinet, M., B. Heinesch, and M. Yernaux, 2003: Horizontal and vertical CO₂ advection in a sloping forest. *Bound.-Layer Meteor.*, **108**, 397–417.
- Bakwin, P. S., P. P. Tans, C. Zhao, W. Ussler, and E. Quesnell, 1995: Measurements of carbon-dioxide on a very tall tower. *Tellus*, **47B**, 535–549.
- Baldocchi, D. D., B. B. Hicks, and T. P. Meyers, 1988: Measuring biosphere-atmosphere exchanges of biologically related gases with micrometeorological methods. *Ecology*, **69**, 1331–1340.
- Berger, B. W., K. J. Davis, C. Yi, P. S. Bakwin, and C. Zhao, 2001: Long-term carbon dioxide fluxes from a very tall tower in a northern forest: Flux measurement methodology. *J. Atmos. Oceanic Technol.*, **18**, 529–542.
- Bowling, D. R., S. P. Burns, T. J. Conway, R. K. Monson, and J. W. C. White, 2005: Extensive observations of CO₂ carbon isotope content in and above a high-elevation subalpine forest. *Global Biogeochem. Cycles*, **19**, GB3023, doi:10.1029/2004GB002394.
- Burns, S. P., and Coauthors, 2006: Measurements of the diurnal cycle of temperature, humidity, wind, and carbon dioxide in a subalpine forest during the Carbon in the Mountains Experiment (CME04). Preprints, *17th Symp. on Boundary Layers and Turbulence/27th Conf. on Agricultural and Forest Meteorology*, San Diego, CA, Amer. Meteor. Soc., JP4.7. [Available online at <http://ams.confex.com/ams/pdfs/papers/110188.pdf>.]
- Conway, T. J., P. P. Tans, L. S. Waterman, K. K. Thoning, D. R. Kitzis, K. A. Masarie, and N. Zhang, 1994: Evidence for in-

- terannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network. *J. Geophys. Res.*, **99** (D11), 22 831–22 856.
- Daube, B. C., K. A. Boering, A. E. Andrews, and S. C. Wofsy, 2002: A high-precision fast-response airborne CO₂ analyzer for in situ sampling from the surface to the middle stratosphere. *J. Atmos. Oceanic Technol.*, **19**, 1532–1543.
- Devore, J. L., 1987: *Probability and Statistics for Engineering and the Sciences*. Brooks/Cole, 672 pp.
- Hansen, J., M. Sato, P. Kharecha, G. Russell, D. W. Lea, and M. Siddall, 2007: Climate change and trace gases. *Philos. Trans. Roy. Soc. London*, **365A**, 1925–1954, doi:10.1098/rsta.2007.2052.
- Haszpra, L., Z. Barcza, P. S. Bakwin, B. W. Berger, K. J. Davis, and T. Weidinger, 2001: Measuring system for the long-term monitoring of biosphere/atmosphere exchange of carbon dioxide. *J. Geophys. Res.*, **106**, 3057–3069.
- ISO/IEC, 1995: Guide to the expression of uncertainty in measurement (GUM). ISO/IEC Guide 98:1995, International Organization for Standardization/International Electrotechnical Commission, Geneva, Switzerland, 101 pp. [Available online at <http://www.iso.org/>.]
- , 2007: International vocabulary of metrology—Basic and general concepts and associated terms (VIM). ISO/IEC Guide 99:2007, (revises ISO Guide 99:1993), International Organization for Standardization/International Electrotechnical Commission, Geneva, Switzerland, 55 pp. [Available online at <http://www.iso.org/>.]
- Komhyr, W. D., T. B. Harris, and L. S. Waterman, 1985: Calibration of nondispersive infrared CO₂ analyzers with CO₂-in-air reference gases. *J. Atmos. Oceanic Technol.*, **2**, 82–88.
- Liang, N. S., G. Inoue, and Y. Fujinuma, 2003: A multichannel automated chamber system for continuous measurement of forest soil CO₂ efflux. *Tree Physiol.*, **23**, 825–832.
- LI-COR, Inc., 2005: LI-7000 CO₂/H₂O instruction manual. Publication 0009-120 (3rd printing), LI-COR, Inc., Lincoln, NE, 227 pp. [Available online at www.licor.com/.]
- Manning, A., 2005: A calibration and intercomparison scheme for continuous, multi-species measurements from a network of tall towers in Europe. *Proc. 12th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracers Measurement Techniques*, WMO Tech. Doc. 1275, GAW Rep. 161, Geneva, Switzerland, WMO, Global Atmosphere Watch Programme, 109–115. [Available online at <http://www.wmo.int/pages/prog/arep/gaw/gaw-reports.html>.]
- Massman, W. J., 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.
- McDermitt, D. K., 1997: Some recommendations for using LI-COR gas analyzers in eddy correlation measurements. Topics discussed at Ameriflux Workshop, 1996. Application Note 118, LI-COR Inc., Lincoln, NE, 7 pp. [Available online at <http://www.licor.com/>.]
- , J. M. Welles, and R. D. Eckles, 1993: Effects of temperature, pressure, and water vapor on gas phase infrared absorption by CO₂. Application Note 116, LI-COR Inc., Lincoln, NE, 5 pp. [Available online at <http://www.licor.com/>.]
- Miller, J.B., Ed., 2006: *Proceedings of the 13th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracers Measurement Techniques*. WMO Tech. Doc. 1359, GAW Rep. 168, Geneva, Switzerland, WMO, Global Atmosphere Watch Programme, 217 pp. [Available online at <http://www.wmo.int/pages/prog/arep/gaw/gaw-reports.html>.]
- Mölder, M., A. Lindroth, and S. Halldin, 2000: Water vapor, CO₂, and temperature profiles in and above a forest—Accuracy assessment of an unattended measurement system. *J. Atmos. Oceanic Technol.*, **17**, 417–425.
- Monson, R. K., A. A. Turnipseed, J. P. Sparks, P. C. Harley, L. E. Scott-Denton, K. Sparks, and T. E. Huxman, 2002: Carbon sequestration in a high-elevation, subalpine forest. *Global Change Biol.*, **8**, 459–478.
- , S. P. Burns, M. W. Williams, A. C. Delany, M. Weintraub, and D. A. Lipson, 2006a: The contribution of beneath-snow soil respiration to total ecosystem respiration in a high-elevation, subalpine forest. *Global Biogeochem. Cycles*, **20**, GB3030, doi:10.1029/2005GB002684.
- , D. L. Lipson, S. P. Burns, A. A. Turnipseed, A. C. Delany, M. W. Williams, and S. K. Schmidt, 2006b: Winter forest soil respiration controlled by climate and microbial community composition. *Nature*, **439**, 711–714, doi:10.1038/nature04555.
- Ocheltree, T. W., and H. W. Loescher, 2007: Design of the ameriflux portable eddy covariance system and uncertainty analysis of carbon measurements. *J. Atmos. Oceanic Technol.*, **24**, 1389–1406.
- Staebler, R. M., and D. R. Fitzjarrald, 2004: Observing subcanopy CO₂ advection. *Agric. For. Meteor.*, **122**, 139–156, doi:10.1016/j.agrformet.2003.09.011.
- Stephens, B. B., A. Watt, and G. Maclean, 2006: An autonomous inexpensive robust CO₂ analyzer (AIRCOA). *Proc. 13th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracers Measurement Techniques*, WMO Tech. Doc. 1359, GAW Rep. 168, Geneva, Switzerland, WMO, Global Atmosphere Watch Programme, 95–99. [Available online at <http://www.wmo.int/pages/prog/arep/gaw/gaw-reports.html>.]
- Sun, J., and Coauthors, 2007: CO₂ transport over complex terrain. *Agric. For. Meteor.*, **145**, 1–21, doi:10.1016/j.agrformet.2007.02.007.
- Taylor, B. N., and C. E. Kuyatt, 1994: Guidelines for evaluating and expressing the uncertainty of NIST measurement results. NIST Tech. Note 1297, National Institute of Standards and Technology, Gaithersburg, MD, 26 pp. [Available online at <http://www.nist.gov>.]
- Thomas, C., and T. Foken, 2007: Organised motion in a tall spruce canopy: Temporal scales, structure spacing and terrain effects. *Bound.-Layer Meteor.*, **122**, 123–147.
- , J. Ruppert, J. Lüers, J. Schröter, J. C. Mayer, and T. Bertolini, 2004: Documentation of the WALDATEM-2003 Experiment, April 28th to August 3rd 2003. Work Rep. 24, Department of Micrometeorology, University of Bayreuth, 46 pp.
- Trivett, N., and A. Köhler, 1999: Guide on sampling and analysis techniques for chemical constituents and physical properties in air and precipitation as applied at stations of the Global Atmosphere Watch. Part I: Carbon dioxide. Tech. Rep. WMO Tech. Doc. 980, Global Atmosphere Watch (GAW) Rep. 134, 63 pp.
- Turnipseed, A. A., P. D. Blanken, D. E. Anderson, and R. K. Monson, 2002: Energy budget above a high-elevation subalpine forest in complex topography. *Agric. For. Meteor.*, **110**, 177–201.
- van der Schoot, M. V., L. P. Steele, R. J. Francey, and D. A. Spencer, 2006: Performance of the LoFlo continuous CO₂ analyser: Baseline and urban air monitoring, diagnostic capability, and potential for enhanced CO₂ calibration. *Proc. 13th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracers Measurement Techniques*,

- WMO Tech. Doc. 1359, GAW Rep. 168, Geneva, Switzerland, WMO, Global Atmosphere Watch Programme, 90–94. [Available online at <http://www.wmo.int/pages/prog/arep/gaw/documents/gaw168.pdf>.]
- Worthy, D., and L. Huang, Eds., 2005: *Proceedings of the 12th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracers Measurement Techniques*. WMO Tech. Doc. 1275, GAW Rep. 161, Geneva, Switzerland, WMO, Global Atmosphere Watch Programme, 264 pp. [Available online at <http://www.wmo.int/pages/prog/arep/gaw/gaw-reports.html>.]
- Xu, L. K., A. A. Matista, and T. C. Hsiao, 1999: A technique for measuring CO₂ and water vapor profiles within and above plant canopies over short periods. *Agric. For. Meteor.*, **94**, 1–12.
- Yi, C., D. E. Anderson, A. A. Turnipseed, S. P. Burns, J. P. Sparks, D. I. Stannard, and R. K. Monson, 2008: The contribution of advective fluxes to net ecosystem CO₂ exchange in a high-elevation, subalpine forest. *Ecol. Appl.*, **18**, 1379–1390, doi:10.1890/06-0908.1.
- Zhao, C., and P. P. Tans, 2006: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air. *J. Geophys. Res.*, **111**, D08S09, doi:10.1029/2005JD006003.
- , P. S. Bakwin, and P. P. Tans, 1997: A design for unattended monitoring of carbon dioxide on a very tall tower. *J. Atmos. Oceanic Technol.*, **14**, 1139–1145.