Atmos. Meas. Tech. Discuss., 6, 1461–1553, 2013
www.atmos-meas-tech-discuss.net/6/1461/2013/
doi:10.5194/amtd-6-1461-2013

Atmospheric Measurement Techniques Discussions

CO_2, CO and CH_4 measurements from the NOAA Earth System Research Laboratory's Tall Tower Greenhouse Gas Observing Network: instrumentation, uncertainty analysis and recommendations for future high-accuracy greenhouse gas monitoring efforts

A. E. Andrews[1], J. D. Kofler[1,2], M. E. Trudeau[1,3], J. C. Williams[1,4], D. H. Neff[1,2], K. A. Masarie[1], D. Y. Chao[1,2], D. R. Kitzis[1,2], P. C. Novelli[1], C. L. Zhao[1,2], E. J. Dlugokencky[1], P. M. Lang[1], M. J. Crotwell[1,2], M. L. Fischer[5], M. J. Parker[6,7], J. T. Lee[8], D. D. Baumann[9], A. R. Desai[10], C. O. Stanier[11], S. F. J. de Wekker[12], D. E. Wolfe[1], J. W. Munger[13], and P. P. Tans[1]

[1]NOAA Earth System Research Laboratory, Boulder, CO, USA
[2]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA
[3]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA
[4]Science and Technology Corporation, Boulder, CO, USA
[5]Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[6]Savannah River National Laboratory, Aiken, SC, USA
[7]Savannah River Nuclear Solutions, LLC, Aiken, SC, USA
[8]Department of Plant, Soil and Environmental Science, University of Maine, Orono, ME USA
[9]Institute for Applied Ecosystem Studies, US Forest Service Northern Research Station, Rhinelander, WI, USA
[10]Atmospheric and Oceanic Sciences Department, University of Wisconsin, Madison, WI, USA
[11]Department of Chemical and Biochemical Engineering, University of Iowa, Iowa City, IA, USA
[12]Department of Environmental Sciences, University of Virginia, Charlottesville, USA
[13]School of Engineering and Applied Science, Harvard University, Cambridge, MA, USA

Received: 5 December 2012 – Accepted: 10 January 2013 – Published: 8 February 2013

Correspondence to: A. E. Andrews (arlyn.andrews@noaa.gov),
J. D. Kofler (jonathan.kofler@noaa.gov)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

A robust in situ CO_2 and CO analysis system has been developed and deployed at eight sites in the NOAA Earth System Research Laboratory's (ESRL) Tall Tower Greenhouse Gas Observing Network. The network uses very tall (> 300 m) television and radio transmitter towers that provide a convenient platform for mid-boundary layer trace gas sampling. Each analyzer has three sample inlets for profile sampling, and a complete vertical profile is obtained every 15 min. The instrument suite at one site has been augmented with a cavity ring-down spectrometer for measuring CO_2 and CH_4. The long-term stability of the systems in the field is typically better than 0.1 ppm for CO_2, 6 ppb for CO, and 0.5 ppb for CH_4, as determined from repeated standard gas measurements. The instrumentation is fully automated and includes sensors for measuring a variety of status parameters, such as temperatures, pressures and flow rates that are inputs for automated alerts and quality control algorithms. These algorithms provide detailed and time-dependent uncertainty estimates for all of the gases and could be adapted to other species or analysis systems. The design emphasizes use of off the shelf parts and modularity to facilitate network operations and ease of maintenance. The systems report high-quality data with > 93 % uptime. Recurrent problems and limitations of the current system are discussed along with general recommendations for high accuracy trace-gas monitoring. The network is a key component of the North American Carbon Program and a useful model for future research-grade operational greenhouse gas monitoring efforts.

1 Introduction

Increased concern about rising greenhouse gas concentrations has already motivated many nations to begin regulating carbon emissions. Accurate measurements of atmospheric carbon dioxide and other species can provide an objective basis for verifying reported emissions at regional to continental scales (10^4–10^6 km^2) (Committee on

Methods for Estimating Greenhouse Gas Emissions and National Research Council, 2010). Accumulation of CO_2 in the atmosphere is the result of anthropogenic emissions, but only about half of the emitted CO_2 remains in the atmosphere. The other half is absorbed by the oceans and the terrestrial biosphere in roughly equal amounts (e.g. Le Quéré et al., 2009). Net biological CO_2 uptake is the small residual of large opposing uptake and emission fluxes at diurnal as well as seasonal scales. Data records with very high precision and long-term stability are therefore needed to resolve the net annual flux.

Here, we describe an automated, robust and high-precision analysis system for routine unattended monitoring of atmospheric CO_2, CO and CH_4 from tall towers and a set of algorithms for estimating time-dependent uncertainties. CO_2 is the principal anthropogenic greenhouse gas, and mixing ratio measurements of its abundance are sensitive to upwind fluxes, including fossil fuel emissions as well as uptake by and emissions from vegetation and soils. CO measurements contribute to the interpretation of CO_2 data by helping to identify and quantify pollution episodes and biomass burning. CH_4 is a potent greenhouse gas, with important anthropogenic sources from agriculture, fossil fuel exploitation, landfills, wastewater treatment and natural biological sources from wetlands. Atmospheric data records of sufficient quality, density and duration have the potential to greatly advance understanding of the processes and reservoirs that dominate the budgets of these and other greenhouse gases on timescales of decades to centuries.

NOAA's Earth System Research Laboratory (ESRL) has been working to build a network of tall tower CO_2 measurement sites since the early 1990s (Bakwin et al., 1998). Observations from tall towers are unique because measurements at several heights along the tower describe the vertical gradient, which reflects the relative influence of remote and local sources. Measurements obtained from sampling levels above ~ 100 m are minimally impacted by nearby vegetation and other local emissions. Gloor et al. (2001) estimated that a typical sampling footprint for a tall tower is of order 10^6 km^2 or roughly 1/10th of the area of the contiguous United States. The towers frequently

penetrate the shallow nighttime boundary layer, in which case measurements from the highest levels are decoupled from the surface. Seasonal, day-to-day, and diurnal variability of CO_2 observed at a tall tower site can be very large. For example, Miles et al. (2012) analyzed data from a temporary installation of ~ 100 m towers in an agricultural region and showed that short-term variations of 10 ppm (parts per million dry air mole fraction) or more are common. Even though daily and seasonal variations may be large, high-precision stable measurements of CO_2 are needed to quantify year-to-year changes in carbon fluxes. Net carbon uptake by ecosystems results from the small difference between large uptake fluxes driven by photosynthesis and large emission fluxes from heterotrophic and autotrophic respiration. Background values of CO_2 are relatively high (currently ~ 390 ppm) and vary with latitude, altitude and time, so signals from individual sources are rapidly diluted, becoming faint.

The North American Carbon Program (NACP) Plan (Wofsy and Harriss, 2002) describes an observing network that would enable ongoing carbon flux estimates with coast-to-coast coverage at the regional scale. The proposed network would resolve spatial differences among regions roughly the size of New England, the Midwest cornbelt, the mid-Atlantic, or the southeast US at temporal scales of months to seasons. The plan calls for thirty sites with surface monitoring from towers, along with bi-weekly aircraft sampling. A substantially larger network would be needed in order to monitor carbon emissions on a state-by-state or city-by-city basis.

Under the NACP, a new in situ CO_2/CO analysis system was developed for the NOAA Tall Tower Greenhouse Gas Observing Network and the network expanded from three sites to seven that are equipped with in situ analyzers (Table 1). The instrument suite at one site has been augmented with a CO_2/CH_4 Cavity Ring-Down Spectrometer (CRDS). The towers are typically television or FM radio transmitter towers that are > 300 m in height and enable trace gas measurements that are representative of the planetary boundary layer (one site, AMT, uses a 107 m cellular telephone tower). There is also one short tower complex terrain site (SNP) located on a mountain ridge in Shenandoah National Park that was established in collaboration with the University of

Virginia (Lee et al., 2012). Complex terrain sites are needed to fill gaps in the monitoring network over mountainous regions, where tall broadcast towers are uncommon, but the representativeness of these sites can be difficult to determine due to complicated meteorological conditions (Brooks et al., 2012; Pillai et al., 2011; Lee et al., 2012).

We have also developed robust quality control algorithms to estimate uncertainties for individual measurements of CO_2, CO and CH_4 that facilitate quality control and allow for rapid identification and diagnosis of problems. Typical analytical uncertainty for CO_2 is < 0.1 ppm, CO < 6 ppb, and CH_4 < 0.5 ppb. In the case of CO_2 and CH_4, atmospheric variability is generally much larger than the analytical uncertainty of the measurements, but the CO analyzer is not sufficiently precise to resolve variability on timescales < 1 h.

In addition to NOAA's efforts, Environment Canada operates twelve greenhouse gas monitoring sites with towers that range in height from 20 to 105 m (Worthy et al., 2003), and an eight-site European tall tower network was recently constructed under the CHIOTTO project (e.g. Vermeulen et al., 2011; Popa et al., 2010; Thompson et al., 2009). These and other data provide the basis for prototype CO_2 data assimilation systems like NOAA's CarbonTracker (Peters et al., 2007; carbontracker.noaa.gov), which provides annually updated weekly estimates of carbon fluxes for a variety of ecosystems and oceans at $1^\circ \times 1^\circ$ resolution over North America and 2° latitude $\times 3^\circ$ longitude globally. CarbonTracker and other models are able to capture much of the synoptic scale variability observed at continental sites (e.g. Law et al., 2008), but the spatial resolution for which carbon fluxes can be determined depends sensitively on the density of the measurement network. Many regions remain under-constrained, and the current North American network falls short of the NACP recommended sampling density. Further expansion of the North American and European greenhouse gas monitoring networks is needed and could be accomplished by a variety of government, university and private sector institutions. Care must be taken to ensure that data from various independently operated networks are compatible, and measurement protocols must be clearly defined.

The purpose of this paper is to describe the CO_2/CO/CH_4 analytical system and data processing with enough detail so that other researchers seeking to make high precision measurements of CO_2 and related gases can replicate relevant components. Although CO_2 analyzers have evolved over the past several years, the gas handling and temperature control techniques described here are generically useful, as is the novel methodology for estimating time-varying uncertainties. Recurrent problems and limitations of the systems are discussed, along with potential improvements and recommendations for future greenhouse gas monitoring efforts. Furthermore, this paper serves as a reference for the data collected from the NOAA ESRL Tall Tower Greenhouse Gas Observing Network from 2006 to present, which are available at ftp://ftp.cmdl.noaa.gov/ccg/towers/. The CO_2 data have been used in several recent continental-scale and regional-scale studies of the North American carbon budget (Gourdji et al., 2012; Schuh, 2012; Lauvaux et al., 2012; Miles et al., 2012; Lee et al., 2012). The WGC CH_4 dataset was the primary record for two regional-scale analyses of CH_4 emissions in California (Zhao et al., 2009; Jeong et al., 2012), and the CO record has been used to evaluate new retrievals from the MOPITT satellite (Deeter et al., 2012).

2 Instrumentation

Starting in 2004, we developed and deployed an updated system for monitoring CO_2 and CO at NOAA tall tower sampling locations. The design is similar to the original CO_2 sampling equipment that was deployed at the NOAA ITN (now discontinued) and LEF tall tower sites (Bakwin et al., 1998; Zhao et al., 1997), and later at the WKT and AMT sites, but with modifications to minimize sensitivity to environmental conditions (such as room temperature) and to simplify maintenance of a larger network. The system is modular, so that a module with a component in need of repair can be quickly replaced with a spare, minimizing downtime and data gaps. Component-level repairs can be done in the laboratory, rather than on site, which keeps costs down and facilitates

quality control. Temperature stabilization enables high-precision measurements to be made with reduced use of expensive calibration gases. All of the major components are easily replaceable commercial off-the-shelf parts, and the modularity allows for new technology to be easily incorporated.

We designed the CO_2/CO analysis system during 2004–2005, and since then, new CO_2 and multi-species analyzers using cavity enhanced absorption spectroscopy techniques such as Cavity Ringdown Spectroscopy (Crosson, 2008) and Off-Axis Integrated Cavity Output Spectroscopy (O'Keefe et al., 1999) have become commercially available. These new analyzers have demonstrated improved off-the-shelf stability compared to the Licor Li-7000 CO_2 analyzer that is the core of the tall tower system (e.g. Richardson et al., 2012; Welp et al., 2012). In 2007, we integrated a Picarro Cavity Ringdown CO_2/CH_4/H_2O analyzer into the system for deployment at the WGC tall tower site (Zhao et al., 2009) as described in Sect. 2.11. The precision and accuracy of the Licor and Picarro CO_2 measurements at WGC are comparable. However, the Licor requires more frequent calibration than the Picarro analyzer as discussed in Sect. 6.3.1.

The CO_2/CO/CH_4 trace gas analysis system was developed according to the following design objectives: (1) ability to deliver high quality CO_2, CO, and CH_4 data; target long-term (year-to-year) cross-site precisions were 0.1 ppm for CO_2, 10 ppb for CO, and 1 ppb for CH_4. (2) Ease of maintenance. (3) Comprehensive monitoring of system parameters for quality control purposes. (4) Insensitivity to environment (e.g. room temperature, humidity, and atmospheric pressure) to minimize use of calibration gas. Figure 1 is a schematic diagram of the analysis system, which occupies a standard instrument rack (48.3 cm × 59.7 cm × 198.1 cm), not including the Picarro analyzer. The individual modules are described below and more detailed schematics are shown in Fig. 2. Many quality-control parameters, such as pressures, flow rates, and temperatures, are recorded in addition to the CO_2 and CO data. Table 2 lists the most important signals included in the data stream. Photographs of the equipment and installations are provided in the Supplement. Most of the towers are equipped with

meteorological sensors, but discussion of the meteorological measurement system is beyond the scope of this paper.

2.1 Sample tubing

At each site, the CO_2/CO analyzer is housed at the base of the tower in a building or portable laboratory built in a trailer or modified sea container. Air is drawn down the tower through three sampling lines (1.27 cm OD Synflex 1300 tubing, wall thickness = 1.57 mm). Three sample inlets are nominally positioned at 30, 100 and ≥ 300 m (as high as practical on a particular tower). Tubing is affixed to the tower using long UV-resistant plastic cable ties or stainless steel hose clamps at 1 m intervals. Tubes are run along tower legs and protected whenever possible to minimize wind-related vibration and stress and exposure to falling ice. Long horizontal runs and low points in the tubing are undesirable. However, at some sites these features cannot be entirely avoided. When possible, we install three tubes to each level so that a separate automated flask-sampling unit and the in situ system can be installed on separate lines with one spare line. Each line (including the spare) has a high-surface area PTFE 0.2 µm filter capsule (6711-7502, Whatman) on the inlet. Inlet filters occasionally become encased in ice or saturated with water during foggy conditions or following heavy rain. Under these conditions flow through the tubes is impeded or even entirely prevented. Flow generally returns to previous levels within a few days.

Each in situ sampling line has a dedicated pump and is continuously flushed at a typical flow rate of 5 to 9 standard liters per minute (slm, corresponding to flow rate at $T = 0\,^{\circ}C$, $P = 1013$ hPa), which corresponds to a residence time of 4 to 7 min in a 500 m synflex tube. The pressure drop in a 500 m sample tube is estimated to be ∼ 44(65) hPa with a Reynolds number of 889 (1333) for a flow rate of 5 (9) slm and depends strongly on tubing diameter. The tubes are checked for leaks at the time of installation by capping the inlet and pulling a vacuum on the tube, and the test is repeated whenever the inlet filters are replaced, ideally once per year or when climbers are on the tower for another repair. The final pressure achieved during the pump-down

is typically < 200 hPa. We use a shut-off valve to isolate the evacuated tube from the test pump and monitor the extent to which the capped line will hold the vacuum.

2.2 Power

DC power for the instrument components is provided by a power supply with 12 V (75 W), ±15 V (75 W each) and two 24 V (200 W each) output modules (Mini-Megapak MM5-15699, Vicor). This power supply was selected for its compact size, robustness, and low noise (ripple). The pumps and some of the temperature control equipment are powered through relays (SDM-CD16AC, Campbell Scientific) so that they can be shut down remotely or automatically restarted if necessary. An uninterruptable power supply (UPS) protects against short duration power outages and power surges (Eaton 9130, 1.5 KVA rackmount).

2.3 Pumps

Air from the sampling lines enters the pump enclosure through a set of 7 µm filters (Swagelok SS-4F-7). The filters are intended to protect the pumps and downstream components from particulates, in the event that the sample tubing is breached. Each of the three sampling lines has a dedicated pump (MPU1763-N828-6.05 115 V/60 Hz, KNF Neuberger) that compresses the air (Fig. 2a). Pump outlet pressures are set to 69 kPa (10 psi) above ambient using a back pressure regulator (GH30XTHMXXXB, Conoflow ITT Industries) and monitored using inexpensive electronic pressure transducers (68075-44 Cole Parmer, 0–25 psig). Excess flow is vented through the back pressure regulator and measured with an electronic mass airflow sensor (AWM5102VN, Honeywell). A fourth (exhaust) pump in the box pulls a vacuum (∼ 250 hPa) on the combined output from the CO_2 and CO analyzers. The exhaust pump enlarges the pressure gradients across the CO and CO_2 analyzers to provide improved pressure and flow control and also improves performance of the Nafion® (registered trade name of E.I. DuPont de Nemours) driers as described in Sect. 2.4. The

pressure upstream of the exhaust pump inlet is measured with a ±103.4 kPa (15 psi) transducer (68075-32, Cole Parmer) to monitor the exhaust pump performance. For convenience, this sensor is mounted in the Nafion drier enclosure. We have used pumps with either Viton® (a.k.a. FPM) or EPDM diaphragms. A fan mounted in a cutout on the side of the enclosure provides cooling. Note that the pumps are factory-equipped with automatic shut-off to prevent overheating (maximum recommended ambient temperature is 40 °C). Air temperature monitored in the interior of the pump enclosure does not typically exceed 35 °C.

The pump assembly is refurbished approximately once per year. Pump diaphragms are replaced, and pumps are tested for compression and vacuum. The "bypass" flow is the portion of the flow that is vented through the back pressure regulator (i.e. equivalent to the total flow minus that portion which is periodically delivered to the analyzers), and provides an indicator of pump performance. In addition to scheduled maintenance, pump units are recalled anytime the flow delivered by the sample pumps drops suddenly or if the total flow (sample plus bypass) drops below 4 slm or if the exhaust line pressure rises unexpectedly or rises above 400 hPa. However, we have found that pumps with torn diaphragms often deliver adequate flow rates and backpressure, but usually will not generate a vacuum. Any leakage of air across a torn sample pump diaphragm will contaminate the sample airstream, and must be avoided. Pumps in the field can be tested for leaks by simply capping the inlet and checking whether the flow drops to zero. Future versions of the system will likely include electronic shut-off valves upstream of the pumps so that this check can be automated. All connections to the pump box are made with quick-connect fittings, so that the entire unit can be easily replaced on site by a minimally trained technician.

2.4 Driers

Air exiting each sample line pump is passed through one channel of a four-channel compressor chiller (02K1044A EC-4-G, M&C Products) to remove the bulk of the water vapor (Fig. 2a). The chiller is configured with four separate glass traps (one for each

sample intake line plus one spare). Each channel has a dedicated peristaltic pump to remove liquid effluent from the trap. The peristaltic pumps require routine maintenance so we reconfigure the set of four pumps as a single modular unit that can easily be removed by a non-skilled technician and returned to our laboratory for service. The temperature of the cooling element is maintained at a setpoint of 1.6 °C. The sample air pressure in the condensers is ∼ 68.9 kPa (10 psi) above ambient, which enables drying to a 1013 hPa dew point that is lower than the cooling element temperature. The temperature of the airstream exiting the chillers is a function of the flow rate and therefore varies with pump performance. We tested the chiller performance with a flow of 6 slm, using a mixture of dry and saturated water to vary the input water content from approximately 0.8 % to 2.8 % (mole fraction), corresponding to a dew point range of 3.9 to 23 °C at 1013 hPa. The moisture content of the output airstream over this range was nearly invariant at 4400 ± 180 ppm, equivalent to a dew point of −3.8 °C at 1013 hPa. Liquid alarm sensors (03E4100 KS2, M&C Products) on each intake line close relays (FA1.4, M&C Products) to disable the pumps if liquid water breaks through the chiller.

A further level of protection against liquid water infiltration is provided by PTFE filter membranes that are relatively impermeable to water (TF-200 PTFE 0.2 micron filters and model 1235 47 mm filter holders, Pall Life Sciences). Laboratory tests determined that an upstream pressure of 69 kPa (10 psi) is required to push liquid water through the filters. Saturated PTFE filters can block all airflow and then they do not dry. Thus, we expect that the filter membranes would need to be replaced after coming into contact with liquid water, although this has not happened. Early units used polycarbonate filter holders, which are substantially less expensive, but we switched to aluminum filter holders after several of the polycarbonate units cracked during shipping. The PTFE filters are installed downstream of the liquid alarm sensors, but they are housed in the pump box for easy access and so that they can be routinely replaced when pump maintenance is performed. We originally used 2 µm filters but recently found that fine black particles were present on the back side of the filters and in many downstream

components, including in the sample manifold and the bypass flow meters. We suspect that the pump diaphragms are shedding fine particles that are smaller than 2 µm.

Single-strand Nafion membrane driers (MD-110-144P-4, Perma Pure LLC) are used in self-purge configuration (Fig. 2b) to further reduce the sample dew point to approximately −36 °C (at 1013 hPa) for the CO_2 channel as indicated by laboratory and field tests using the WGC G-1301 Picarro $CO_2/CH_4/H_2O$ analyzer. The sample dew point for the CO channel is approximately −34 °C (at 1013 hPa), as indicated by a separate dew point sensor (DMT 142, Vaisala) immediately downstream of the sample cell. A single 3.66 m drier is used on the CO_2 channel, which has a flow rate of 250 standard cubic cm per minute (sccm). Two 3.66 m Nafion driers are used in series for the CO channel, which has a higher flow rate of 600 sccm. The effectiveness of Nafion membrane driers depends on the relative flow rates and partial pressures through the sample and purge tubing. The exhaust pump reduces the pressure on the purge side of the Nafion driers, resulting in a faster volume flow rate and improved drying. A non-hazardous desiccant (Drierite, WA Hammond; part number 27070 includes a canister with Swagelok fittings) is used to remove residual water from the analyzer exhaust before it enters the purge housing. The lifetime of the desiccant is several years given the extremely low water content of the analyzer exhaust. Nafion is more effective at cooler temperatures, and rapid temperature changes can produce large changes in the water content of the sample airstream. We therefore house the Nafion driers in an insulated enclosure equipped with a thermoelectric cooler (SD6C-HCAF-AARG, Watlow). The box temperature is maintained at ∼ 20 °C, and the enclosure is purged with a few sccm of dry air from a cylinder to prevent condensation.

The sample line pressure in the Nafion driers is not actively controlled. Instead, all sample pumps and calibration gas regulators are manually adjusted to deliver approximately the same pressure. At some sites, we have begun monitoring the line pressure at the exit of the Nafion drier on the CO_2 channel and have noted that calibration curve residuals are smaller when the pressures are carefully adjusted. Future versions of

the system may therefore include active pressure regulation upstream of the Nafion membrane driers (Welp et al., 2012).

Calibration gases are introduced upstream of the Nafion membrane driers. The Nafion membrane acts as a reservoir for water and is normally equilibrated with the pre-dried (chilled) sample air. The dry calibration gases are humidified as they pass through the Nafion driers and emerge with a dew point that is equal to that for dried atmospheric sample air. Differences in water content among samples from different inlet heights and the calibration gases should be < 100 ppm (mole fraction) of water to avoid artifacts associated with the so-called "dilution effect", which is the difference in mole fraction when computed with respect to dry versus wet air. Water vapor differences among samples and standards can also cause spectral artifacts related to line interference or pressure broadening. Our system produces differences between atmospheric samples and calibration standards that are < 10 ppm H_2O. All of the CO_2, CO, and CH_4 measurements are reported as dry air mole fractions (e.g. χ_{CO_2}).

2.5 Sample/calibration selection manifolds

Atmospheric samples from the three inlet lines are selected through a solenoid valve manifold (Fig. 2c). Two three-way valves are plumbed in series to minimize dead volumes. A 2-way valve at the end of the chain is used as a shut-off valve for the third inlet channel (see diagram). Calibration gases are selected using a similar manifold comprised of solenoid valves that are rated to 689.5 kPa (100 psi) inlet pressure (009-0933-900 (3-way) and 009-0631-900 (2-way), Parker Hannifin Pneutronics Division). We initially used large orifice Teflon valves (203-3414-215 (3-way) and 203-1414-215 (2-way), Galtek) for atmospheric sampling, but we encountered problems with cross-port leaks as described in Sect. 3.2 below and now use the same valves that are used in the calibration manifold. Solenoid valves were chosen instead of a multi-position (stream-selection) valve to increase reliability. We tested one system with a multi-position valve (10-position ECMT, Valco) with the expectation that the multi-position valve would have less dead volume than the solenoid valve manifolds, but the response

time after transitions between calibration gases was not improved, suggesting that other components dominate flushing and equilibration in this system.

2.6 CO_2 analyzer

CO_2 is measured using a non-dispersive infrared gas analyzer (Licor Li-7000 CO_2/H_2O). The analyzer is housed in a temperature-controlled enclosure to minimize sensitivity to variations in room temperature (Fig. 2d). CO_2 mole fractions reported by the Licor analyzer are temperature-compensated, but some sensitivity remains that can cause errors as large as a few tenths of a ppm of CO_2 in a typical operating environment. The temperature in our CO_2 analyzer enclosure is maintained 10–15 °C above the room temperature so that the Li-7000 internal temperatures typically fall in the range from 37 ° to 40 °C. Although the setpoint varies from site to site, the cell temperature for each unit is normally controlled to within 0.1 °C (see Sect. 2.8 for more details about temperature control). Flow through the sample cell of the Li-7000 is actively regulated upstream of the cell (1179A52CS1BV, MKS Instruments). The reference flow passes through a needle valve (4171-1505, Matheson) upstream of the analyzer, and sample and reference flows are joined downstream of the analyzer to ensure that the sample and reference cell pressures are nearly equal. A pressure controller (640A13TS1V22V, MKS) downstream of the junction actively regulates the pressure to 1066 hPa (800 Torr). Typical sample and reference flow settings are 250 sccm and 10 sccm, respectively. The reference flow rate is measured downstream of the analyzer (AWM3150V, Honeywell). The difference in sample and reference flows does result in a small but invariant difference in pressure across the cells. The flow and pressure controllers are sensitive to ambient temperature and are therefore housed inside the temperature-controlled area. The H_2O channel of the Li-7000 analyzer is used to continuously monitor the performance of the drying system. The absolute H_2O measurement from the Li-7000 is not accurate at the very low humidity levels achieved by our drying system (e.g. the analyzers may be offset by 500 ppm or more and frequently report negative H_2O mixing ratios for our dried sample airstream). However, the gain

of the H_2O channel is robust and able to reliably indicate changes in drier performance over time or differences in water content among calibration and sample gases. In-line filters (SS-4F-7, Swagelok) are mounted on the sample and reference inlets of the Li-7000 to prevent accidental introduction of debris. A pressure relief valve is plumbed between the sample outlet of the Li-7000 and the pressure controller (2391243-26-9, Tavco) to protect the Licor analyzer from accidental over pressurization. We occasionally have problems with the Tavco valve releasing unintentionally. It can be remotely reseated by sending a command to open the downstream pressure control valve. The resulting reduced pressure reseats the Tavco valve.

2.7 CO analyzer

CO is measured using gas filter correlation (48C Trace Level, Thermo Electron Corporation). The factory-installed internal pump is removed from the analyzer, and flow (1179A23CS1BV, MKS Instruments) and pressure (640A13TS1V22V, MKS) controllers are installed in that space, upstream and downstream of the sample cell, respectively (Fig. 2e). Cell pressure is maintained at 1066 hPa (800 Torr). Sample flow is controlled at 600 sccm. The factory-installed internal pressure and flow sensors and the heaters on the sample cell are disconnected. We do not use the optional zero and span solenoids available from the manufacturer. Instead, calibration and sample gases are introduced using an external gas selection manifold as described above. A dew point sensor (DMT 142, Vaisala) is plumbed downstream of the sample cell and mounted inside the analyzer. An in-line filter (SS-4F-7) is mounted on the inlet of the flow controller to protect downstream components from accidental introduction of debris. We use a catalytic reagent (Sofnocat 423, O. C. Lugo) to scrub CO from ambient air. The scrubbed air is measured at least twice per hour to track baseline drift in the analyzer. Frequent checks of the baseline drift are needed to achieve high precision (~ 3 ppb for a 2-min average) with this analyzer. The scrubber is a stainless steel tube (0.5″ OD × 12″ L) filled with the catalyst and with a glass wool plug and stainless steel mesh at each end. The tube is mounted at a slight angle from horizontal to prevent

unfilled spaces that might develop as a result of gravitational settling of the catalyst. The sample flow is periodically diverted through the scrubber by simultaneously switching two 3-way solenoid valves on either end of the scrubber (203-3414-215 (3-way) and 203-1414-215 (2-way), Galtek). Note that in this application, there is no problem with cross-port leaks across the Galtek valves, since there is no significant pressure gradient across the valve ports. For convenience, the scrubber and solenoid valves are mounted in the enclosure with the Nafion driers, outside of the temperature-controlled region. Lab tests indicate that a much smaller CO scrubber volume would perform equally well. However, there is no penalty (other than cost) for using an excessively large scrubber volume. Response time, for example is not affected, since all air exiting the scrubber is free from CO.

2.8 Temperature control

Both the CO_2 and CO analyzers are carefully temperature-controlled to a setpoint several degrees above the room temperature. The setpoint is site-specific, but usually falls in the range 37 °C to 40 °C. The Li-7000 CO_2 analyzer can operate at temperatures up to 50 °C (although serial communications may be unreliable above 45 °C), and the 48C TL CO analyzer has a specified operating temperature up to 45 °C. The CO_2 analyzer is housed in a rack-mounted aluminum chassis box (48.3 cm × 17.8 cm × 55.9 cm) along with its pressure and flow controllers. The CO analyzer is available from the manufacturer in a rack-mount configuration, so no separate enclosure is required. A small temperature controller unit (CT325PD2C1, Minco) is mounted inside each enclosure that drives six Kapton® (registered trade name of E.I. DuPont and Nemours) tape heaters (HK5340R58.9L36B, Minco), which are distributed evenly over the interior surface of the boxes, including the lid. The control temperature is measured with a 4-wire platinum RTD (S665 PDZ40A (D)). We performed several experiments to determine the optimal location for the control and obtained the best results when the sensor element was suspended in the air near the center of the enclosure. The temperature controllers are inexpensive and easy to use. However, we had several unexplained failures that

required replacement of the Minco temperature controller unit and found that reliability improved when used with a solid-state relay driver (e.g. Crydom, MPDCD-3), but with some degradation of temperature stability. Each box is wrapped with a single layer of Aramid fabric insulation (MC8-4596B 48″, Tex Tech). A small fan mounted inside each enclosure provides air circulation. The CO_2 enclosure is mounted above the CO analyzer in a standard instrument rack, with a gap of approximately 1 cm between the boxes. A scroll fan is used to circulate air between the boxes to prevent overheating. The variability in the CO_2 assembly is typically < 0.2 °C (1σ), and the CO_2 analyzer temperature is typically stable to ~ 0.05 °C.

We have had persistent problems maintaining the temperature control at one site (Walnut Grove, CA; WGC) due to wide extremes in room temperature, which cannot be accommodated with a seasonally invariant temperature setpoint. At that site, the equipment is housed in the antennae's transmitter building and we do not have direct control of the room temperature. Rather than repeatedly adjust the setpoint temperature, which must be done manually and is not easily accessible, we let the analyzer temperatures float (Fig. 3a). The insulation causes the analyzer temperatures to vary slowly enough that we can effectively correct for baseline drift using frequent measurements of one of the standard gases (for CO_2) or the scrubbed ambient air (for CO), along with an empirically determined relationship between the internal analyzer temperature and analyzer baseline (Fig. 3b–d) that is described in more detail in Sect. 5.1.1. For CO_2, our implementation is a relatively expensive solution, in that it requires frequent use of well-calibrated standard gases. The calibration frequency at WGC is approximately twice that now used at other sites. Fortunately, WGC is also equipped with a Picarro CO_2/CH_4 CRDS as described in Sect. 2.11 that is insensitive to typical room temperature variations, and we plan to rely primarily on that sensor going forward so that we can reduce calibration gas use at that site.

2.9 Data acquisition and control

A datalogger (CR-10X-ST-MA-NC, Campbell Scientific) with accessories is used for all data acquisition and control functions. All engineering and trace gas data are recorded every 30 s. We wanted a simple, commercially available, robust operating system, and the ability to take advantage of evolving technology for communications and data storage. In addition to the datalogger, other components include: two multiplexer boards (AM16/32A-ST-SW, Campbell Scientific), relay modules (SDMCD16AC, Campbell Scientific), an analog output module (SDM-AO4-SW), and a serial communications module (SDM-SIO4, Campbell Scientific). We designed custom printed circuit boards to simplify connections to the datalogger's wiring panel. The CR-10X datalogger has been discontinued, and we plan to eventually upgrade the systems to use the replacement CR-1000 model or another more modern data acquisition system. The CR-1000 has improved serial communications and much larger data storage capacity.

An onsite PC laptop is used for remote access by cellular modem or Digital Subscriber Line (DSL) and for data storage. The PC runs a Windows operating system, software to communicate with the datalogger (Loggernet, Campbell Scientific), and remote administrator software (Radmin). During the first few years of operation, most sites were accessed by analog modem, and limited bandwidth dictated that the data transfer (via ftp upload to a server at our laboratory) occur only once per day. Although we now have improved communications, we have not changed the data retrieval schedule. The PC time is synchronized to a time server every 15 min using commercially available software (Dimension 4, Thinking Man Software), which also logs differences due to PC clock drift. The PC time is uploaded to the datalogger daily. PC clock drifts are of order seconds per day, and become significant if uncorrected over periods of weeks or more (e.g. if Internet communications are lost or PC time updates fail).

Most of the engineering data are differential analog signals, but serial communications are used to retrieve data from the Li-7000 CO_2 analyzer and from the Thermo Electron 48CTL CO analyzer. Serial communications with the datalogger are somewhat

inefficient and limit the speed at which we are able to interrogate the sensors. We configured the datalogger program to run on a 5-s interval, in order to allow adequate time for serial polling and response. To compensate for the low sampling frequency, we use the built-in averaging capabilities of the CO_2 and CO analyzers. The Li-7000 CO_2 analyzer is set to report a 5-s average. The 48CTL CO analyzer, which is noisier, is set to report a 30-s average, and thus the 5-s samples recorded by the datalogger are not independent. The 5-s measurements are then aggregated to 30-s averages and stored in the datalogger's memory along with the corresponding standard deviations. The data are downloaded to the local PC every minute and a program (Baler, Campbell Scientific) running on the PC bins the data into hourly average files. The datalogger memory can store approximately two days worth of data, which provides some protection against communication interruptions or PC failures.

2.10 Standard gases and related components

A total of nine calibration gases are currently used for the CO_2/CO analysis system as described below in Sect. 4.2. We use high purity two-stage nickel-plated brass regulators with low internal volume (51-41C-590, Scott Specialty). The tank pressure gauge on each of these regulators is replaced with an electronic pressure transducer (68075-56, Cole-Parmer). The transducer is protected from rapid pressure changes by a flow restrictor (SS-4-SRA-2-EG, Swagelok). The tank pressure signals provide a measure of gas use that is tracked and used to identify tanks with slow leaks, which most often occur at the CGA connection between the tank and the regulator. Quick connect fittings with automatic shut-off (SS-QC4-D-200 and SSQC4-B2PM, Swagelok) are installed on the outlet of each regulator, so that on-site technicians can easily purge the regulator when a new cylinder is attached. Purging the regulator consists of opening the tank valve, quickly shutting it, allowing the fresh gas to sit in the regulator for a few minutes, then draining the gas through the regulator outlet. This process is repeated three times each time a new cylinder is attached. Purging the regulator minimizes the introduction of room air into the calibration lines and protects the gas in the new cylinder from

backward diffusion of room air or residual air from the previous cylinder. Clean stainless steel tubing is used for the calibration lines (3.18 mm (0.125 inch) OD, wall thickness 0.07 mm; SS-T2-S-028-20, Swagelok), which are typically a few meters long. The tube specifications reflect tradeoffs between minimizing volume and providing robust connections (i.e. connections to 3.18 mm tubing are generally more robust than connections to 1.59 mm tubing). An in-line 7 µm filter (SS-4F-7, Swagelok) is installed at the point where the calibration line enters the manifold to provide protection against introduction of particulate matter such as metal fragments from the plumbing connections or dust from the room.

2.11 Integration of CRDS CO_2/CH_4 analyzer

The WGC installation includes a Picarro G-1301 CRDS for measuring CO_2, CH_4 and water vapor. The Picarro analyzer is plumbed in parallel with the Licor CO_2 analyzer (Fig. 1). A 2 µm stainless steel filter (Swagelok, SS-2F-2) is installed on the inlet and a needle valve is used to restrict flow through the analyzer to approximately 80 sccm, which is adequate for flushing the sample cell during the five-minute sampling interval while minimizing calibration gas consumption. Note that a higher flow rate would be desirable for a stand-alone installation in order to flush upstream tubing and regulators in a reasonable time when switching among calibration and sample modes.

The pressure of the sample airstream exiting the Nafion drier assembly is ~ 68.9 kPa (10 psi) above ambient. The Picarro cell pressure is controlled at 186 hPa and the cell temperature is maintained at 45 °C. Robust pressure and temperature control are off-the-shelf features of Picarro analyzers, and additional pressure/temperature control is unnecessary. The Picarro H_2O channel reports the humidity of the dried sample airstream. The typical value is ~ 0.013 % (mole fraction), corresponding to a dewpoint of −39.8 °C at 1013 hPa. There is no discernable difference in water content between dried ambient air and the humidified standards. The Picarro analyzer is not equipped with a flow sensor, so an external sensor (AWM 3100V, Honeywell) is installed on the outlet. Exhaust from the Picarro is captured and combined with the exhaust from

the Li-7000 and used to purge the Nafion drier. The Picarro analyzer has a dedicated computer for data acquisition and control. However, to simplify post-processing, we use the Campbell Scientific serial communications data acquisition system to integrate key Picarro output fields into our primary data stream.

The Picarro analyzer was deployed in Fall 2007 and was among the first commercially available CO_2/CH_4 Picarro units to be installed at a field site. The stability of the analyzer and the reliability of the H_2O corrections to CO_2 and CH_4 were initially unknown. Our configuration was largely driven by convenience, so that standards and gas-handling could be shared between the Picarro and the Li-7000. Recent studies have shown that Picarro measurements of CO_2 and CH_4 can reliably be corrected for water vapor effects (Chen et al., 2010; Rella et al., 2012), and the analyzer has demonstrated remarkable stability over nearly five years of operation as will be described in more detail in Sect. 6.3.1.

3 Reliability

The hardware has generally been very reliable, and most sites have reported valid CO_2 and CO data for > 93 % of days since installation. However, certain recurrent problems have been encountered as described below.

3.1 Automated alerts

An important feature of the post-processing software is that it provides daily summaries of errors or anomalies that are emailed to the personnel who are responsible for monitoring system problems. The content of the summaries has evolved over time as new failure modes have been encountered. Alerts are generated if fewer than expected data files are transferred, if the file sizes are smaller than normal, or if signals are outside of the expected range. Calibration gas cylinder pressures and usage are tracked and

used to compute a regularly updated estimated replacement date for each cylinder. Some additional cases are discussed in Sect. 3.2.

Certain failures result in automatic flagging of the data. For example, fatal flags are assigned when flow through one or more of the analyzers is lost. Loss of flow may occur for all levels if there is a systematic problem or for a single intake when a pump fails or a liquid alarm sensor is triggered. New automatic flagging algorithms are developed when a new failure mode is discovered or for cases where manual flagging would be overly tedious. Automated flagging reduces the likelihood of human error associated with data entry, however some manual flagging is unavoidable, e.g. if work is being done while the system is running.

Approximately 50 plots are created nightly for each site that display measured χ_{CO_2}, χ_{CO} and χ_{CH_4} and detailed uncertainty information (see Sect. 5.2), along with important engineering signals and other diagnostics. The plots are viewed using a custom interface accessible via Internet browser. Plots for all sites are reviewed at least twice per week and whenever an automated alert is generated. The plots enable rapid diagnosis of problems, and are archived so that we can easily review data and diagnostics for any data range.

Additional alerts are available on an hourly basis. One data record per hour containing all instrument singles is uploaded from the site computer to a server, and an email is generated if any of the key parameters fall outside a specified range. Errors such as pump failures, power outages, and losses of communication can therefore be detected within one or two hours.

3.2 Notable or recurrent problems

3.2.1 Cross-port leaks and relay-failures

Within a month after deployment at WGC in Fall 2007, we noticed unusual patterns appearing in the sample line bypass flow signals. Investigation revealed that air was leaking across the ports of the PTFE solenoid valves (Galtek, 203-3414-215) in the

sampling manifold so that air reaching the analyzers was a mixture from different intake heights. Mounting screws securing the valves to the floor of the enclosure had been over-tightened, distorting the valve base. Upon reviewing the specifications for the valves, we opted to replace the sample valves in all of our systems with steel solenoid valves that are identical to those used for the calibration manifold.

The original intent of the bypass flow sensors was to monitor the performance of the pumps, but after the WGC sample solenoid cross-port leak incident, we implemented a "flow accounting" algorithm that has detected subsequent valve-switching failures. We have encountered recurring problems with valve switching that have affected at least four sites and that worsened over time. At AMT sample solenoids have intermittently failed to switch, and at AMT, SNP and WGC, a similar problem has affected the CO zeroing solenoids. At WKT, two of the CO_2 calibration solenoids intermittently failed. Evidence suggests electronics problems that may be internal to the Campbell Scientific relay module used to drive the valves (SDM-CD16) or perhaps faulty electrical connections elsewhere, rather than defective valves. Calibration and CO zeroing valve failures were easy to detect based on calibration residuals, since assigned standard values did not correspond to the air that was being sampled. However, flow accounting based on analyzer and sample-line bypass flows is needed to detect valve-switching failures in the sampling manifold and to flag the affected data.

3.2.2 Sampling line leaks

Contamination resulting from leaks in the sampling lines is often difficult to detect. Leaks within the field laboratory rarely develop spontaneously, but we have occasionally lost data because of failure to properly tighten one or more connections during an installation or maintenance/repair visit. When possible, we test for leaking fittings by placing a few pounds of dry ice near the system for several hours or overnight while monitoring the measured CO_2 signal. Care must be taken to avoid exposing personnel to dangerous levels of CO_2. It is useful to have an inexpensive handheld CO_2 monitor when performing these tests to ensure that ambient CO_2 levels are safe (< 5000 ppm

for 8 h time weighted average exposure, US Occupational Safety and Health Administration Permissible Exposure Limit). We have also developed a leak-checking apparatus consisting of a hand pump with an electronic pressure gauge that can be used to check whether a section of plumbing holds a vacuum. After pumping down the line, a valve between the pump and the gauge is closed, and the pressure is tracked for several minutes or longer.

Tubing on the tower can be damaged by falling ice, high winds, or fatigue at the points where it is secured to the tower. Once a leak has developed, rainwater can infiltrate the tubing and freeze/thaw cycles can cause additional damage. The first indication of a leak is often a liquid alarm signal after heavy rain. Other times, severe tubing damage was visible from the ground. The start date of a leak on the tower is usually difficult or impossible to determine, and unless the tube was severed, the sampled air would have been a mixture from two or more heights. The impact on the data depends on the vertical gradient of the gas being measured, so data collected during well-mixed periods may be minimally affected. We have successfully worked with tower climbers to repair damaged tubing, and vacuum leak checks are performed whenever we have climbers on a tower. One effective method for finding leaks is to pressurize the line with a pump, so that climbers can hear the air hissing out. We plan to develop an automated system to enable routine leak checking, where a large-orifice remotely actuated and normally open valve will be installed on the tower at the sample inlet, and the valve will be periodically closed to check the vacuum created by the sample pump. Modern radio modems provide extremely reliable communication with equipment mounted high on the towers, and AC power is generally available on the towers.

3.2.3 Temperature control

Room temperature at some sites exhibits strong seasonality and is outside of our control at sites where the equipment is located in the tower's transmitter building. It has proven difficult to find a single setpoint for the temperature controllers that will work at all sites under all conditions. Unfortunately, the setpoint potentiometers for the

temperature controllers are located inside the CO_2 and CO analyzer assemblies and are difficult to access. Ideally, we would be able to adjust the setpoint temperatures remotely, or at least install an adjustment dial on the outside of the enclosures so that local technicians could easily make adjustments when needed.

4 Sampling and calibration

4.1 Sampling sequence and system modes

The system operates on a twenty-four hour measurement cycle (Fig. 4). Calibration and atmospheric sampling modes are described in Table 3, and each mode is typically run for a 5-min interval. The CO and CO_2 analyzers can be calibrated independently, but during atmospheric sampling, they draw air from the same sampling inlet. The following discussion is centered on CO_2, which is the primary species of interest, but generally applies to CO except as noted.

The five-minute sampling interval allows for full equilibration even for system mode transitions corresponding to large χ_{CO_2} differences that can occur during calibration or during sampling when there is a large vertical gradient across the measurement heights. The precision of the Li-7000 analyzer (~ 0.02 ppm for a 30-s average) is sufficient to resolve the gradual approach to equilibration that is proportional to $\Delta\chi_{CO_2}$. Data from a typical summer day are shown in Fig. 5. Vertical gradients during the daytime are small, when heating of the surface causes vigorous turbulent mixing within the planetary boundary layer. During the night, CO_2 and CO emitted at the surface accumulates in the shallow stable layer. At the WBI site in Iowa, where vegetation and soils have high nighttime respiration rates, we have observed nighttime differences of > 60 ppm CO_2 between the 31 m and 379 m inlets. We also observe strong vertical gradients at some sites during winter, when temperature inversions are present and pollutants accumulate near the surface.

When switching between system modes, we allow three minutes for the system to flush and report data corresponding to the final two minutes of each five-minute sampling interval. Lab experiments and field calibration data show that χ_{CO_2} errors associated with incomplete equilibration are < 0.05 ppm for a three minute flushing interval when the difference between successive samples is < 60 ppm. Several times per day, the CO_2 analyzer dwells on the highest sampling height while the CO analyzer baseline is measured using scrubbed ambient air from that inlet. In such cases, where the CO_2 mode does not change between successive sampling intervals, we report all of the available data for CO_2 (i.e. there is no need to discard the first three minutes of the second five-minute interval).

The five-minute sampling interval limits the temporal resolution of our analyzer to no more than four three-intake profiles per hour. With a three-minute flushing time, we therefore report data corresponding to eight minutes out of every hour for each intake height, which limits our ability to confidently compute e.g. hourly or afternoon averages. We actually get slightly fewer than four profiles per hour, since the sampling cycle is interrupted for calibrations. At certain times of day, the temporal variability of CO_2 at a single height is considerably larger than the uncertainty resulting from incomplete equilibration within the five-minute sampling interval. Clearly, it would be advantageous to reduce the flushing time in order to increase the temporal resolution of the measurements, but doing so would adversely impact the CO measurements, which require a longer equilibration time. We have also considered adding integrating volumes to the sampling lines to achieve more representative sampling (Winderlich et al., 2010) or modifying the sampling sequence to spend most of the time dwelling on the highest intake, since we typically use the vertical gradient information primarily to identify periods with vigorous vertical mixing.

4.2 Calibration

4.2.1 CO_2

Four standards (CO2C1, CO2C2, CO2C3, and CO2C4) are used to calibrate the response curve for the Li-7000. The approximate CO_2 values for the standards are given in Table 3. A fifth cylinder (CO2REF) supplies gas to the Li-7000 reference cell, and the concentration is chosen to approximately match the CO2C2 standard, which is used to monitor the Li-7000 baseline. A sixth calibration standard, the "target" (CO2TGT), is measured independently to monitor the stability of the instrument.

We initially used a CO_2 calibration sequence consisting of a full response curve calibration four times per day and a baseline check approximately once per hour. The CO2TGT tank was measured four times per day with two of the measurements adjacent to full calibrations, and the other two measurements temporally distant from both full calibrations and baseline checks. After ~ 2 yr of operation, we gained confidence in the stability of the system and have reduced the frequency of full calibrations to twice per day with baseline checks every two hours. The CO2TGT tank is still measured four times per day, but now all target measurements are temporally distant from calibrations and baseline checks. We recently modified the sequence to run on a 23-h interval so that the timing of the calibrations and target measurements drifts throughout the day and covers a full diurnal cycle over the course of approximately 10 days. At WGC, we continue to use the original four hourly calibration sequence because of temperature control problems described in Sect. 2.8.

During experiments with a prototype system, we attempted to run true differential zero measurements to monitor the Li-7000 baseline, where the gas cylinder used to supply gas to the reference cell was routed through a "T" fitting so that we could periodically send the gas simultaneously through the Li-7000's sample and reference cells. We found that this setup, which has been widely used, apparently disrupts internal regulator and/or calibration line pressures and disturbs the measured CO_2 value. Recovery from this type of perturbation exceeded ten minutes. We settled on

a "pseudo-differential zero" measurement technique using the CO2C2 standard along with a separate CO2REF standard to monitor the Li-7000 baseline. The concentration of the CO2REF standard is within a few ppm of the CO2C2 standard so that baseline drift can be reliably distinguished from gain changes.

The Li-7000 analyzer signal is not inherently linear, but output is available that has been linearized according to a fifth-order polynomial with unit-specific calibration coefficients determined by the manufacturer. The linearization algorithm relies on a user-specified reference concentration, and this value must be accurate to within a few ppm to avoid significant deviations from linearity. CO2REF cylinders must be replaced several times per year, so for simplicity, we typically use an approximate value of 380 ppm, while actual values may range from 377 to 383 ppm. We were not initially confident that the linearized output would meet the accuracy requirements, so we developed a conservative calibration strategy using four standard gases to enable use of a higher-order polynomial calibration curve if needed. We also store the raw detector signals from the Li-7000 sample and reference cells (i.e. Li-7000 "CO2A W" and "CO2B W" signals and corresponding values for H_2O) so that we can recover a signal that is comparable to the analog output from earlier Licor CO_2 analyzer models (such as the Li-6251). We found that the Li-7000's linearized output is quite reliable. The linearized output benefits from additional internal signal averaging and consequently demonstrates improved precision compared to the raw signals. We continue to use four standard gases, since the cylinders last several years and the additional information is useful when problems arise.

4.2.2 CO

For the CO calibration, we use two standards (COC1 and COC2) and a target standard (COTGT). The baseline is tracked by measuring scrubbed ambient air with 0 ppb CO (COZER). The baseline is measured every 30–40 min, and the other calibration standards are measured approximately every 23 h. The COTGT is measured four times

per day and COTGT measurements are not typically adjacent to baseline checks or calibrations.

Note that for the Thermo Electron 48C TL CO analyzer, which is based on a gas filter correlation technique, it is important to use CO calibration standards that are made with a balance of whole air. In particular, we learned that the standards must have near-ambient CO_2 content. The absence of CO_2 in the standard gases will artificially raise the baseline of the analyzer due to spectral interference. In gas filter correlation, spectral interference can produce either a positive or negative response. We determined empirically that a change in the CO_2 content from 380 to 0 ppm corresponds to a change in the CO analyzer baseline that is equivalent to +20 ppb CO. The artifact was present and had consistent magnitude in all units tested (> 5 units tested to date) and is independent of the CO concentration. CO_2 interference leads to significantly biased calibrations when standards are employed that do not contain near-ambient levels of CO_2. In addition, variations in ambient CO_2 can cause CO measurement errors as large as 5 ppb (for a 100 ppm CO_2 variation) when scrubbed ambient air is used to track the analyzer baseline. However, if nearly simultaneous ambient CO_2 measurements are available, then the CO analyzer baseline could be corrected for CO_2 interference, although we have not yet implemented this. Water vapor is another interferent, and the Nafion humidifies the standards to the same level as the ambient samples are dried. The humidity of the CO sample airstream is monitored using a dew point sensor (Vaisala, DRYCAP DMT 142) installed immediately downstream of the sample cell (Fig. 2e).

4.3 Traceability of calibration standards

NOAA ESRL is responsible for maintaining the World Meteorological Organization's (WMO) mole fraction calibration scales for CO_2, CO, and CH_4. Details for each of these gases are described below.

– *CO_2:* the uncertainty of the WMO CO_2 X2007 calibration scale is estimated to be 0.07 ppm, and the scale is internally consistent to better than 0.03 ppm (Zhao and Tans, 2006). Typical standard deviations for repeated calibration measurements of our field standards separated by several days or longer are ≤ 0.02 ppm (1σ). The primary calibration scale covers the range 250 ppm to 520 ppm. CO_2 standards in the range 500–1000 ppm are calibrated manometrically with accuracy better than 0.1 ppm. We used high CO_2 standards (650 and 700 ppm) to characterize the linearity of the Licor and WGC Picarro analyzers.

– *CO:* CO results are referenced to the WMO CO X2004 scale (WMO, 2010), which ranges from 30 to 500 ppb by mole fraction. It's uncertainty is ∼ 0.7 %. Repeatability (1σ) of cylinder calibrations is < 0.1 ppb and the reproducibility over several years is ∼ 1 ppb (WMO, 2010).

– *CH_4:* CH_4 measurements are available from the Picarro CRDS analyzer at WGC. The absolute uncertainty of the WMO CH_4 calibration scale is estimated to be 0.2 %, the repeatability of cylinder calibrations is typically 1 ppb (1σ), and the reproducibility in the ambient range is typically < 0.5 ppb. The scale ranges from 300 to 2600 ppb (Dlugokencky et al., 2005), but will be extended to 5700 ppb within a year.

Field standards are calibrated in the laboratory relative to WMO scales before and after deployment. For CO_2, pre- and post-deployment calibrations are available for 177 tanks since 2004. The mean difference was 0.02 ± 0.05 ppm (post minus pre) and 14 cylinders had absolute differences > 0.1 ppm. For the 59 CO standards with pre- and post-deployment calibration data, the mean difference was 3.2 ± 2.6 ppb. The distribution is strongly skewed toward positive values, with six cylinders drifting up by more than 5 ppb and 2 drifting more than 10 ppb over their lifetime. Unfortunately, post-deployment calibrations were not performed for ∼ 29 CO standards prior to 2010 due to a shortage of cylinders and recurring instrument problems in the calibration laboratory. When there is a significant difference between pre- and post-deployment calibrations

or when post-deployment calibration data are missing, we evaluate field calibration residuals to decide whether to use the pre-deployment, post-deployment, mean, or a time-dependent drift-corrected value. To date, CH_4 standards have not received post-deployment calibrations, and field calibration residuals have not indicated any problems. CH_4 standards to not generally drift, but we plan to perform post-deployment calibrations from now on.

5 Post-processing

5.1 Algorithms for calculating χ_{CO_2} and χ_{CO} and χ_{CH_4}

Daily data are retrieved as a set of 24 hourly files. Data are stored with 30-s temporal resolution, and the timestamp corresponds to the end of the 30-s interval. Average values and standard deviations for each 30-s interval are recorded for the CO_2 and CO analyzer signals.

The data array contains a system mode indicator (**SYSMODE**) for each gas that is used within the datalogger program to set the position of valves in the calibration and sample manifolds. The data array also contains a counter (**INTERVAL**) that is used to track how many 30-s intervals have elapsed since the **SYSMODE** was last switched. Thus, during a typical five-minute sampling period, **INTERVAL** values range from 1 to 10. Higher counter values occur when the sampling sequence contains back-to-back occurrences of the same mode. This happens when either the CO or CO_2 analyzer enters a calibration mode, while the other analyzer continues to sample ambient air with no inlet height change and for variants of the sampling program that are designed to dwell on a particular intake height for longer than 5 min. **SYSMODE** and **INTERVAL** are used in post-processing to separate data from different calibration and sample modes and to filter data immediately following a sampling mode transition.

The post-processing code was developed in the numerical analysis package R and has been translated to IDL and improved for operational use. The post-processing

software operates on three days of data because calibration data from the previous and subsequent days are needed to compute the most accurate χ_{CO_2} values and uncertainties for the central day.

5.1.1 CO_2

We use the linearized, pressure-, H_2O-, and temperature-corrected differential CO_2 mole fraction signal reported by the Licor. We define $\boldsymbol{s}$ to be the vector of individual 30-s average analyzer signals $\boldsymbol{s}_i$ for all times t_i (grey curve in Fig. 4). The Licor baseline drift is tracked using repeated measurements of the CO2C2 standard, which is measured every 1–2 h. In post-processing, we extract a vector of analyzer baseline measurements, $\boldsymbol{s}_b$ at times t_b (times when **SYSMODE** = "CO2C2" and **INTERVAL** = 10) and linearly interpolate over time to create a continuous baseline timeseries b (red lines in Fig. 4). The baseline is subtracted from the raw data time series $\boldsymbol{s}$ to obtain the drift-corrected signal, $\boldsymbol{s}'$.

Drift-corrected values corresponding to standard gas measurements are extracted and interpolated to all times t_i (note, the CO2C2 standard is used to track the baseline, so $\boldsymbol{s}_{c2} = \boldsymbol{s}_b$ and $\boldsymbol{s}'_{c2} = 0$). A first-order (linear) calibration curve is computed from the interpolated calibration values for each time using a simple least-squares regression algorithm and the fit coefficients are stored in an array. The fit coefficients are applied to $\boldsymbol{s}'$ to compute χ_{CO_2} for all data.

For cases where a significant correlation exists between analyzer temperature and the baseline signal $\boldsymbol{s}_b$, we have the option to enable a temperature-dependent baseline algorithm as illustrated in Fig. 3. In that case, the slope from the baseline : temperature relationship (Fig. 3b) is applied to the difference between the measured analyzer temperature (red curve in Fig. 3c) and the analyzer temperature extracted at t_b and interpolated to all times t_i (black symbols and connecting lines in Fig. 3c). The resulting temperature-dependent baseline correction is added to the usual time-interpolated baseline (black symbols and connecting lines in Fig. 3d) to generate a continuous representation of the analyzer baseline (red curve in Fig. 3d).

In the case of the Picarro analyzer at WGC, no baseline is subtracted from the raw data. An average linear calibration curve is computed for each day that includes all calibration data within the 3-day post-processing window. The **SYSMODES** are the same as for the Licor.

5.1.2 CO

Because the CO analyzer signal is relatively noisy, we use a 2-min mean to smooth the CO analyzer output before computing the baseline. As for CO_2, the analyzer's baseline response $\boldsymbol{s}_b$ is linearly interpolated in time between successive zero measurements to generate a continuous baseline time series b. The continuous baseline is subtracted from the raw analyzer output $\boldsymbol{s}$ to compute $\boldsymbol{s}'$. We have found that the CO analyzer gain is quite stable, so $\boldsymbol{s}'$ values for each standard are averaged across the three-day post-processing window to minimize the impact of analyzer noise before computing calibration coefficients. Linear calibration coefficients are computed from a regression that includes the baseline measurement ($\boldsymbol{s}'_b = 0$; $\chi_{CO} = 0$ ppb) and 3-day average measurements from the COC1 and COC2 standards, $\boldsymbol{s}'_{1ave}$ and $\boldsymbol{s}'_{2ave}$. The resulting single set of fit coefficients is applied to the baseline-corrected analyzer time series $\boldsymbol{s}'$ at its native 30-s resolution. If either of the reference gas cylinders (COC1 or COC2) are replaced during the 3-day post-processing window, then separate average $\boldsymbol{s}'$ values are computed for each cylinder, and values for all cylinders are included in the regression.

5.1.3 CH_4

The WGC Picarro is plumbed in parallel with the Licor, so all standard gases are common for the two analyzers. Initially, each of the CO_2 calibration gases was also calibrated for CH_4. However, after ~ 1.5 yr, we gained confidence in the stability of the Picarro and stopped calibrating the CO2C2 standard for CH_4, since that cylinder is replaced frequently and CH_4 cylinder calibrations are time-consuming. All of the other calibration standards, including the target standard are calibrated for CH_4, and

the linear calibration coefficients are determined using three standards instead of four. Otherwise, the post-processing is the same as described above for WGC Picarro CO_2.

5.2 Estimated uncertainty

We have developed a set of algorithms to provide plausible time-varying uncertainty estimates for individual CO_2, CO and CH_4 measurements. The uncertainty algorithms separately and quantitatively track the major sources of error affecting the measurements. For applications like inverse modeling to estimate CO_2 surface fluxes, the most important considerations are long-term repeatability and compatibility of measurements. That is, we need to understand the extent to which we can confidently interpret differences among measurements made at the same site and within and across networks from hour-to-hour, month-to-month and year-to-year.

We separately report three aspects of the measurement error: (1) uncertainty related to the calibration scale as described above; (2) time-dependent analytical uncertainty for each measurement system; and (3) the standard deviation of each 30-s measurement, which reflects both instrument noise and atmospheric variability. These uncertainty estimates are not independent, and cannot be simply combined into a single value. For example, errors in assigned values for the standards contribute to calibration curve fit residuals, and in some cases instrument precision dominates the standard deviation of repeated measurements, while in other cases, real atmospheric variability dominates. The relevance of these various metrics depends on the nature of the application. In practice, the most important source of error when interpreting data is model-representation error, i.e. the extent to which a model with finite resolution can be expected to simulate point measurements. Many studies use hourly or afternoon-average data, and since the system switches between different sampling heights, the data are only quasi-continuous, with valid measurements from one of three sampling heights < 13 % of the time. The standard deviation of the available measurements (typically 4 per hour per intake height) gives an indication of the variability, and atmospheric conditions tend to persist for several hours.

We have not included the absolute uncertainty of the calibration scales in the analytical uncertainty estimates described below. However, if measurements from different laboratories or programs are combined for a particular analysis, then any calibration-scale differences must be taken into account. NOAA ESRL participates in ongoing standard gas and real air comparisons with a number of laboratories (Masarie et al., 2001; WMO, 2011, pages 207–211). The absolute accuracy and internal consistency of the calibration scales is discussed above in Sect. 4.3. We typically use four standards to determine the first-order calibration polynomial for the CO_2 detector, two for CO and three for CH_4. Separate target tanks are also measured for each gas. The "target residuals" (i.e. the measured minus assigned values) provide an independent measure of the extent to which we are able to propagate the WMO scales within our own laboratory and at our field sites.

5.2.1 Time-dependent analytical uncertainty estimates

The total analytical uncertainty is estimated as the quadrature sum of seven terms described below. The analytical uncertainty is meant to represent the extent to which year-to-year and site-to-site differences can be confidently interpreted. Typical values for each term are given in Table 4. In the description of the algorithms, we have used χ_{CO_2} for simplicity, but unless otherwise specified, the same algorithms apply to χ_{CO} and χ_{CH_4}.

Analyzer short-term precision, u_p

Uncertainty related to the analyzer precision, u_p is estimated by interpolating the 30-s standard deviation of baseline measurements, s_b, to all times t_i (Fig. 6). The short-term signal-to-noise ratio for the Li-7000 CO_2 analyzer is extremely high, and the 30-s standard deviation for the calibration standards is typically better than 0.03 ppm, while the variability during atmospheric sampling is rarely < 0.2 ppm and often > 1 ppm. In contrast, short-term analyzer noise is one of the dominant factors limiting the

overall precision of the CO measurements, with typical 30-s standard deviations of approximately 5 ppb for reference gas and baseline measurements, which is comparable to the variability observed for ambient air. The analyzer short-term precision reflects random analyzer errors and is dependent on the averaging interval (i.e. 30 s). As discussed at the end of this section, we therefore report u_p separately from the other uncertainty components described below, which result from systematic measurement errors.

Analyzer baseline-drift uncertainty, u_b

Unresolved temporal variations in the analyzer baseline are another source of error estimated here as follows (Fig. 6):

1. A set of alternate realizations of the continuous baseline are created where individual baseline measurements $\boldsymbol{s}_b$ have been omitted. This results in three realizations of the baseline for each time t (i.e. the original baseline including all available $\boldsymbol{s}_b$ and the cases where the bracketing $\boldsymbol{s}_b$ values have been excluded).
2. The standard deviation across the three unique realizations of the baseline, σ_b, is calculated for each time in t.
3. A time-varying weighting function is applied to σ_b so that the baseline-drift uncertainty, u_b is zero at times t_b and equal to σ_b halfway between successive baseline measurements.

This approach provides a reasonable measure of baseline-drift uncertainty in the absence of high-frequency baseline variations that are not captured by $\boldsymbol{s}_b$. We do not expect high-frequency CO_2 or CO baseline variations when analyzer temperatures and pressures are well controlled or slowly varying. Data are screened for the presence of strong correlation between analyzer temperature and $\boldsymbol{s}_b$ as described in Sect. 5.1.1. In cases where a temperature-dependent baseline correction is enabled, an additional term is included to represent the uncertainty in the baseline : temperature regression.

The target standard measurements also help to detect unresolved baseline variations as described in Sect. 6.1. Target measurements are only sensitive to baseline-drift errors when not directly adjacent to baseline calibrations and are most useful if distributed throughout the day.

Calibration curve fitting and extrapolation errors, u_f and u_{ex}

Within the χ_{CO_2} range spanned by the standard gases (i.e. typically 350 ppm to 460 ppm for CO_2, 0 to 350 ppb for CO, and 1650–2000 for CH_4), the fit uncertainty (u_f) is well represented by the 68 % prediction interval for the least-squares regression. An empirically determined extrapolation uncertainty (u_{ex}) is applied for values outside of the calibrated range. For Li-7000 measurements of CO_2, a conservative value of 0.005 ppm uncertainty per ppm outside the calibrated range was determined from laboratory tests of three different analyzers. The range of the WMO CO scale was only recently extended from ~ 350 ppb to 1000 ppb, and we have not yet measured high χ_{H_2O} standards on one of our systems. Instead, we rely on the linearity specification for the analyzer (1 % full scale for the values less than 1 ppm) to estimate the extrapolation error. The maximum deviation from linearity (10 ppb) is assigned for all values > 500 ppb. Values above 1000 ppb are exceedingly rare at our sites, and occur only during fires and strong inversions.

Various common measures of fit uncertainty are shown in Fig. 7 for a typical CO_2 calibration curve along with the estimated extrapolation uncertainty, u_{ex}. Calibration residuals for the standard gases corresponding to the three-day period from which this example was taken are shown in Fig. 8 for both CO_2 and CO. The 67 % prediction interval was selected to represent the 1σ curve fitting uncertainty, u_f, because the values are typically large enough to encompass ~ 67 % of the CO_2 and CO target standard residuals. We do not include the target standard measurements in the uncertainty calculation or the calibration curve, and instead use the target information as an independent check on the estimated analytical uncertainty.

Equilibration of standard gases, u_{stdeq}

For CO_2, the calibration coefficients are computed using the final 30-s average measurement for each of the four calibration standards (INTERVAL = 10). For CO, the calibration polynomial is computed using the 3-day average value computed from the final two minutes of individual standard gas and zero measurements (INTERVAL ≥ 7). To check whether standard gas values have stabilized, we re-compute the calibration coefficients using the analyzer values shifted earlier by one 30-s interval (i.e. INTERVAL = 9 for CO_2 and 6 ≤ INTERVAL ≤ 9 for CO). The new calibration coefficients are applied to the data and the difference between the result and the original χ_{CO_2} values is used to represent the calibration equilibration uncertainty.

Equilibration of sample gases, u_{smpeq}

When switching between inlets at different sampling heights, the time required for full equilibration depends on the vertical χ_{CO_2} gradient. We use the routine field calibration data to derive an empirical function describing the mean approach to equilibrium when switching between sampling modes (Fig. 9). The approach to equilibrium is not captured by a single time constant. We use calibration data within the three-day post-processing window to create an empirical function like the heavy solid curve in Fig. 9 that is applied to atmospheric data taking into account the measured χ_{CO_2} difference between successive modes. The approach to equilibrium is consistent across all of the standard gases and errors are small after 3-min of flushing time (< 0.1 ppm for a 100 ppm χ_{CO_2} difference for INTERVAL ≥ 7).

"Dilution" errors associated with varying humidity, u_{wv}

In the case of absorption measurements, the presence of water vapor can introduce errors by direct spectral interference, by changing the pressure broadening of the absorption lines, or by "dilution". Our system minimizes humidity differences between

the sample airstream and standards by passing sample and standard gas through Nafion membrane driers. The sample air is dried, while standard gases are humidified. We do not explicitly consider errors related to spectral interference or line broadening, but since we monitor the humidity of the CO_2 and CO airstreams, it is trivial to track the magnitude of any dilution errors, which for Li-7000 measurements of CO_2 are the largest source of error related to water vapor.

Dilution errors result from humidity differences between samples and standards. For CO_2, which is reported in parts per million, the dilution error can be computed as follows:

$$u_{wv} = \chi_{CO_2}\left(\frac{10^6}{10^6 + \Delta\chi_{H_2O}}\right) - \chi_{CO_2}, \quad (1)$$

where $\Delta\chi_{H_2O}$ is the difference between the measured χ_{H_2O} at a particular time and χ_{H_2O} measured during calibration modes and interpolated to all times t_i (red symbols and connecting lines in Fig. 10a and b). A 250 ppm χ_{H_2O} difference between samples and standards would cause a 0.1 ppm dilution error for χ_{CO_2} of 400 ppm. Note that the absolute water vapor content is not important, which is fortunate, because the Li-7000 water calibration is unreliable at very low humidity as indicated by the negative values of χ_{H_2O} in Fig. 10a and b. Laboratory tests indicated that the Li-7000 is able to detect relative changes in H_2O with sufficient accuracy for estimating u_{wv}.

Figure 10a shows a high-humidity case when counter-flow to the purge side of the Nafion driers was lost due to the unintentional release of a pressure relief valve. Even under these conditions of high-humidity the dilution error is small as long as humidity differences between sample air and standard gases are minimal. Residuals for CO_2 standards were < 0.06 ppm for this period, which was similar to adjacent periods where the drying system was operating normally.

5.2.2 Other sources of uncertainty

There are some potential sources of error that cannot be reliably detected from our available engineering data in an automated way. Two examples are (1) contamination related to the long sampling lines, pumps, and chillers that are upstream of where the calibration gases are injected and that are exposed to ambient humidity, temperature, and pressure, and (2) undetected leaks of room air or ambient air from a lower altitude into the sample airstream. We have relied on laboratory tests, field diagnostics, and comparison with independent measurements to assess the likely impact of these errors. Many independent tests over a wide range of conditions have been performed and are described in Sect. 6.

5.2.3 Propagation of uncertainty

The analyzer short-term precision is a random error, and therefore should be reduced by averaging according to the square root of the number of observations. The other terms (u_b, u_f, u_{ex}, u_{stdeq}, u_{smpeq}, and u_{wv}) are systematic errors that are not reduced by averaging. We separately report random and systematic uncertainty so that these components can be correctly propagated when computing average values from the data (e.g. hourly or afternoon averages). As described in Sect. 2.9, the analysis system operates on a 5-s cycle, and 30-s average data are archived. We also report the 30-s standard deviation corresponding to each measured value, which reflects both analyzer precision and atmospheric variability. For CO_2 and CH_4, the random uncertainty is often much less than the standard deviation, and we can interpret the difference as a measure of atmospheric variability. However, for CO, the random uncertainty is typically comparable to the standard deviation of the measurements and the analyzer cannot resolve variability at the sub-hourly scale.

6 Performance

To date, we have deployed eight of these systems to field sites. The locations are listed in Table 1 along with installation dates. All of the systems have been in the field for ≥ 4 yr.

6.1 Calibration residuals and target measurements

The calibration curve residuals and target tank measurements from WKT for nearly six years are shown in Fig. 11. The residuals for individual tanks are obviously not randomly distributed around zero, and time-dependent biases approaching 0.1 ppm are seen for some cylinders. We use the linearized, temperature- and water-corrected CO_2 signal from the Licor, and apply a first-order calibration polynomial as described in Sect. 3.2. The residuals are not significantly improved by adding a quadratic coefficient, and it is possible that some of the residuals result from imperfect delivery of standard gases to the analyzer or imperfect linearization, e.g. if the CO2REF concentration is substantially different from the specified value or if the Li-7000 H_2O calibration is especially poor. Errors in the assigned values for the reference gases also contribute, but are generally small. The pattern of residuals may or may not change when a tank is replaced, and sometimes the residuals vary slowly, in a manner that suggests that the CO_2 concentration of one or more of the cylinders might be unstable, which can occur e.g. in the case of a slow and/or temperature-dependent leak. Final calibration data are not available for the CO2C3 standard that was installed in Fall 2009, and the residuals near the end of the record may improve when post-deployment calibration data become available.

The frequency of calibrations was reduced in early 2009 as described in Sect. 4.2. Figure 11a shows that the CO2C2 standard was replaced as often as three times per year. A damaged pressure sensor on the CO2C2 regulator went undetected for nearly two years and indicated lower than actual pressure, leading several cylinders to be replaced prematurely. The typical lifetime of a CO2C2 cylinder is 6 months. Target

tanks last 1.2 yr, and CO2C1, CO2C3, CO2C4 last 2.3 yr. The stability of the residuals over timescales of days to weeks suggests that the Licor calibration frequency could be further reduced to something like once per 23 h, which would double the lifetime of CO2C1, CO2C3 and CO2C4. Licor baseline drift monitoring should continue at the current frequency of once per 2 h but could be done with air from an uncalibrated cylinder or a zero-air generator. Whenever possible, we try to avoid replacing more than one standard at a time, so that any unusually large residuals can be unambiguously attributed to a particular cylinder.

The residuals of the target measurements are shown along with estimated analytical uncertainty in Fig. 11b. The target measurements provide an independent measure of the analytical uncertainty excluding any errors resulting from inlet components that are upstream of the calibration manifold (e.g. intake filters, sampling lines, pumps, and condensers). The target residuals and analytical uncertainty agree well for most of the record. Although not shown, consistency between target residuals and estimated analytical uncertainty is also achieved for other sites and for CO and CH_4 measurements. For CO_2 measurements at WKT and other sites, the analytical uncertainty is dominated on long timescales by the fit uncertainty, u_f, but other uncertainty terms may contribute for particular cases.

During the first six months of operation, target residuals were unacceptably noisy and were greatly improved after a site visit in March 2007, when the output flow from the air conditioner in the trailer was directed away from the analyzer. The estimated uncertainty tracks the root-mean-square target residuals very well over the entire record, but the uncertainty of a subset of the individual measurements is too low early in the record. The original calibration sequence had one target tank measurement per day that was directly adjacent to a full calibration, and three others that were deliberately distant from full calibrations and baseline checks. For data collected prior to March 2007, the mean of the target measurements adjacent to full calibrations was 0.02 ± 0.05, while for other target measurements was -0.04 ± 0.08 ppm. The Licor CO2C2 measurements were not strongly correlated with room temperature or analyzer temperature, so the

only indication of a problem was provided by those target measurements that were not adjacent to baseline checks or full calibrations. This example illustrates the utility of having multiple target measurements distributed throughout the day that are temporally separated from other reference gas measurements for detecting problems that are not otherwise apparent and that may depend on the diurnal variation of room temperature.

6.2 Laboratory and field tests

We have evaluated the un-calibrated system components (e.g. inlet components, pumps and chillers) in the laboratory and with field studies at the BAO tower. Results from three tests are described below. In two cases, we observed small (≤ 0.15 ppm for χ_{CO_2}) possible biases but were not able to conclusively determine whether these observed differences reflect artifacts related to the inlet components or stemming from the test setup. We therefore conclude that 0.2 ppm is a conservative upper limit for biases affecting χ_{CO_2} measurements from this system.

6.2.1 Laboratory tests of wetted versus $MgClO_4$-dried air

The lab is equipped with a high-volume dynamic dilution system, where high CO_2 air from a cylinder is diluted in an integrating volume with CO_2 free air. The integrating volume has an internal fan and provides a large volume (> 20 slm) of air with well-mixed and slowly varying χ_{CO_2} at super-ambient pressure (134.5 kPa). We developed a setup for testing the sample inlets under wet conditions, where air from the dilution system was routed through a bubbler and then split into three separate airstreams, two of which were routed to sample inlet ports and the third was passed through a $MgClO_4$ trap and into the CO2TGT port on the calibration manifold. The setup included a bypass for the bubbler, so that we could test for differences using either dry or wetted air. We found that air sampled through the two inlet ports consistently agreed to within 0.01 ppm and that dry air sampled through the $MgClO_4$ and the CO2TGT port was repeatedly found to have a positive offset of approximately 0.04 ppm with respect to the

other two levels. When wetted air was sampled, agreement between the two sample streams was unchanged, but the difference between the sample airstreams and the CO2TGT airstream changed sign and varied from −0.04 ppm to −0.15 ppm (CO2TGT port minus sample ports). The smaller offset was seen when a stream-selection valve (ECMT, Valco) was used instead of the typical solenoid manifold, and the larger offset was seen on a system with aluminum solenoid valves in the calibration manifold. We now use steel solenoid valves in the calibration manifold, which perform similarly to the stream selection valve.

6.2.2 Laboratory tests of humid ambient air sampled through long and short inlet tubes

The integrating volume apparatus described above has also been used to smooth high-frequency variations when sampling ambient air during system tests. We evaluated the impact of sampling through a long inlet tube under ambient conditions with moderate humidity levels (χ_{H_2O} ~ 0.75 %, 30 % RH, 18 °C, 833 hPa). Measured differences were < 0.02 ppm for CO_2 during well-mixed midday conditions when sampling outdoor air from the integrating-volume through a short inlet tube compared to sampling through a 76 m coil of 1.27 cm (0.5 inch) OD Synflex tubing. The mean χ_{CO_2} over a 5-h sampling window was 395.59 ± 0.25 (SD) ppm. During that period, there were sixteen or more independent measurements from each inlet, so the standard error of the mean value was ~ 0.06 ppm. Adjacent nighttime periods had higher variability, but also showed no significant differences across the sampling inlets.

6.2.3 Tank air sampled through BAO inlet tubes

The Boulder Atmospheric Observatory tall tower is a research platform equipped with two elevators. Two reference gas cylinders (assigned values χ_{CO_2} = 371.59 and 401.89 ppm and χ_{CO} = 115.7 and 133.5 ppb) were taken to the top of the tower (300 m)

and sampled through the tower inlet tubes. The measured minus assigned values were −0.13 ppm and −0.11 ppm for χ_{CO_2} and were −3.2 and +0.86 ppb for χ_{CO}.

6.3 Comparison with independent measurements

6.3.1 Ongoing co-located flask sampling

The tall tower sites are equipped with automated flask sampling systems that normally collect daily or alternate day mid-afternoon samples. Routine automated comparison of flask and in situ sampling tracks the level of measurement compatibility within our own laboratory and is useful for identifying experimental problems such as sampling line leaks in either system. A very useful feature of the flask sampling strategy is that we analyze each flask for about 50 compounds, including greenhouse gases, isotopic composition of CO_2, hydrocarbons and halocarbons. This results in a wealth of data that can be used in the interpretation of observed patterns in the major greenhouse gases, enabling at least partial attribution to specific sources/processes.

The flask sampling units are known as Programmable Flask Packages (PFPs), and each unit contains twelve individual 0.7 L borosilicate glass flasks, each with a valve on both ends. The valve manifold is stainless steel, and valves are glass with Teflon O-rings. A data logging and control system provide the interface for controlling the sampling valves and for storing sample flush volumes and fill pressures during sampling, along with system status and a time stamp. Pumps for the PFP sampling are housed in a separate Programmable Compressor Package (PCP). The PFP/PCP system was originally designed to operate on an airplane over a wide range of altitudes. A detailed description of the PFP and associated components is beyond the scope of the current paper and will appear in a separate publication. A schematic diagram of the PFP sampling system is included in the Supplement, along with a photograph showing the interior of a PFP.

To date, we have stayed close to the original PFP/PCP design, to maximize consistency of data and logistics with our laboratory's aircraft-sampling program (e.g. Karion

et al., 2012) but we hope to eventually modify the flask sampling apparatus for optimal performance at tower sites. Tower-specific modifications that have already been implemented include (1) a separate pump assembly to continuously flush the long sample tubes at about 4 slm, (2) a pressure sensor and flow meter on the sampling line, (3) an optional A/C power supply to replace the batteries and trickle charger that are normally used to provide power for aircraft sampling, (4) a datalogger and cellular modem to trigger samples and record line pressure and flow though the flush-pump.

The tower PFP sampling strategy has evolved over time, especially during 2006–2008, and is subject to logistical constraints particular to individual sites. Generally, flasks are sampled in pairs at approximately 1400 LST. The PFP units were not originally designed for parallel sampling, so until recently paired flasks were filled sequentially, with a typical time difference of 3 to 5 min. True-paired sampling began at BAO in January 2011 and throughout the network in January 2012. At most sites, PFP samples are drawn from the highest sampling level on the tower through a dedicated inlet and sample tube. To provide a truly independent measurement, the PFP does not share a sampling tube with the in situ system, except when only one suitable tube is available. When a flask sample is triggered and the PCP pumps are enabled, the flow rate in the tube increases from the standby rate of about 2–4 slm to ∼ 15 slm, the combined flow from the PCP and flush pumps. The increased flow causes a length-dependent pressure drop in the tube (140–250 hPa). Pressure fluctuations perturb the equilibrium between sample air and the walls of the tubing, and flush times of 10 min or more are needed to adequately flush the longest sample lines after a new equilibrium is reached. In the current configuration, the PFP manifold is flushed with 80 L of air, and flasks are flushed with 70 L. For reference, a 500 m length of 1.17 cm OD Synflex 1300 tubing has a volume of 35 L. The total time for flushing and filling is ∼ 15 min, of which the fill time is less than one minute.

In situ and PFP flow rates vary from site to site and depend on pump performance, which may change over the lifetime of the pumps and with temperature. The in situ system currently switches among three sampling heights, so only quasi-continuous

data are available for a particular level. Accurate measurement of all flow rates would be needed to ensure synchronous sampling of the PFP and in situ systems. Instead, we have simply triggered the flask samples at a fixed time of day and compared the closest temporal match within a specified window.

Annual flask versus in situ comparisons for CO_2 and CO are summarized in Tables 5 and 6 for PFP samples with hourly in situ standard deviations of < 0.5 ppm for CO_2 and 10 ppb for CO. Table 7 shows PFP comparisons with the Picarro CO_2 and CH_4 data at WGC. PFP samples have not been collected at the Shenandoah National Park site because of logistical challenges. Agreement is reasonably good for CO_2 at the annual level, but for several sites/years falls short of our 0.1 ppm target, and PFP CO_2 measurements are systematically higher than the in situ values, especially in the most recent years. Results for WBI during 2009–2011 and SCT in 2010 are especially concerning. WKT shares a line with the in situ system, which may explain the smaller bias than for other sites. Prior to October 2007, samples at WKT were collected from the 122 m inlet, and have been collected from 457 m since that time. The LEF in situ system was upgraded in May 2009 and a separate PFP intake was installed to 396 m sampling. Prior to that time, PFP samples were collected from a shared intake at 244 m. A pair of flask samples is collected manually each week at LEF using a separate collection system that shares a line with the in situ system, and those show consistently good agreement with the in situ system before and after the May 2009 upgrade ("LEF Manual" in Table 5). The AMT in situ system was replaced in February 2009, and only a handful of samples are available for comparison with the old system, which was performing poorly near the end of its lifetime. CO comparisons are generally satisfactory, with most annual median differences ≤ 3 ppb. The CO standard deviations at AMT since 2009 are higher than for most sites because that CO analyzer is very noisy (typical $u_p > 8$ ppb). Agreement between in situ and PFP CH_4 measurements is ≤ 1 ppb for all years except for 2007.

Karion et al. (2012) also evaluate PFP versus in situ measurements for routine aircraft flights over Alaska from 2009–2011. They report PFP minus in situ values of

0.20 ± 0.37, when data are filtered to exclude periods of high variability (1σ variability in the continuous analyzer over the flask flush and fill time < 0.2 ppm), which is very consistent with our results for those years. Their values for CO and CH_4 were also consistent with our results. Their continuous analyzer is an un-dried, rigorously characterized Picarro analyzer. Stephens et al. (2011) compared PFP versus in situ results from a high altitude site (Niwot Ridge, CO) for August 2005–early 2011, and reported differences with comparable magnitude, but opposite sign (−0.17 ppm ± 0.38 ppm, n = 745). They also compare their in situ results with weekly manually sampled glass flasks as we do at LEF and they find similarly good agreement (−0.06 ± 0.30 ppm, n = 585), and Niwot Ridge PFP versus in situ comparisons since 2010 show a trend with relatively higher PFP values (B. Stephens, personal communication, 2012) that is consistent with the increasingly positive PFP minus in situ χ_{CO_2} values in Table 5 for the most recent years.

Figure 12a shows the time series of PFP minus in situ CO_2 differences from BAO for samples collected when the standard deviation of in situ data within a 1.25 h window was < 0.5 ppm. The mean (median) of the individual differences is 0.12 (0.07) ± 0.49 ppm (1σ), with 67 % of the absolute monthly mean differences < 0.19 ppm and 95 % < 0.47 ppm. Months with fewer than five individual comparisons are excluded. In late 2009, the monthly mean bias shifted from negative to positive. The mean value for December 2010 was 0.79 ppm, and the reason for this anomaly is unknown. We would like to achieve agreement between the PFP and in situ systems of better than 0.1 ppm. Several lines of evidence point toward biases in a significant subset of the PFP samples as the driver of PFP versus in situ differences. We are aggressively working to understand the differences and to improve the PFP sampling configuration, as described below.

6.3.2 Picarro : Licor comparison and intensive flask sampling at BAO

Atmospheric variability causes significant noise in the PFP versus in situ comparisons, limiting our ability to use them to detect sampling problems. Starting in September

2011, we configured the BAO in situ system to dwell on the 300 m intake and commenced a series of experiments to investigate strategies for improving in situ versus flask agreement. Leak checks were performed on both the PFP and in situ lines in June 2011. A Picarro analyzer was installed on the PFP intake from 9 September until 28 October 2011 and reported 2-s data. For the first several days, no PFP samples were collected to enable an uncomplicated comparison of the Licor and Picarro CO_2 measurements. A laboratory calibration and water correction were applied to the Picarro CO_2 data, but no field calibrations were performed. For comparison with the Licor, the Picarro data were smoothed using a 30-s running average and the time shifted by −71 s to account for differing flows in the separate intake lines. The median difference between the Licor and Picarro measurements was 0.04 ± 0.06 ppm for 9–12 September, as shown in Rella et al. (2012) and that level of agreement was typical of the entire period when the Picarro analyzer was online, during which the atmospheric water vapor mole fraction varied from 0.30 to 1.34 %. The remarkable agreement between the un-dried, minimally calibrated Picarro and the well-calibrated Licor measurements suggests that PFP versus in situ differences may be attributable to collection, storage or analysis problems with the PFPs.

PFP versus in situ agreement for 30 September–28 October 2011 is shown in Fig. 12b for the Licor and the Picarro analyzer. During this period, the PFP and the Picarro analyzer shared a common intake line, in order to test for sampling artifacts that might result from perturbing the pressures in the PFP sampling line. We found optimal agreement with the Licor 30-s measurements when the PFP time was shifted by −180 s to account for different flow rates in the separate sample inlets. The PFP time was not shifted for comparison with the Picarro analyzer, since they shared a common sampling line. Picarro and Licor data within 60 s of the flask-fill end time were averaged for comparison with PFP data. Differences were relatively insensitive to the width of the averaging window applied to the Licor data up to at least 2 min. The in situ standard deviation within the averaging window was used to filter periods with atmospheric variability. We did not apply a sophisticated weighting function because it was apparent

from careful consideration of the time series that agreement would not substantially improve. Of 35 comparisons, only 4 had 2-min standard deviations > 0.2 ppm and were excluded from the statistics. The mean (median) CO_2 difference between the Licor and the Picarro values corresponding to the PFP samples was 0.00 (0.00) ± 0.07 ppm (1σ); the PFP minus Licor difference was 0.16 (0.02) ± 0.4 ppm; and the PFP minus Picarro difference was 0.19 (0.05) ± 0.4 ppm. Hourly average Licor minus Picarro differences are also shown in Fig. 12b for hours where the standard deviations of both in situ analyzers were < 0.3 ppm. For the hourly data, the mean (median) CO_2 difference was 0.00 (0.00) ± 0.03 ppm (N = 193). PFP minus Picarro CH_4 differences are not shown, but exhibit good agreement with a mean (median) difference of −0.84 (−1.37)±0.12 ppb for the same subset of samples.

The consistency between the totally independent Picarro and Licor measurements shows that PFP versus Licor differences are not explained solely by differences in the intake line that might result from e.g. pressure fluctuations that would also be expected to also affect the Picarro data. These PFP samples were collected with extra-long flush settings, corresponding to > 7 volumes of the 300 m intake line, and the PCP firmware had to be modified to avoid timing out. All tower PCPs have now received the firmware update to enable longer flush times.

Through testing at BAO and regular deployment in the network, we have identified several PFP units that appear to be contaminated such that they repeatedly have CO_2 > 0.5–3 ppm higher than in situ data. Unfortunately, laboratory tests with dry air do not show higher CO_2 values in these contaminated PFPs. It is possible that individual flasks are also contaminated, but since PFP versus in situ field comparisons have been complicated by atmospheric variability, we do not have enough statistics to identify problems with individual flasks. To improve the comparisons, we have modified the in situ sampling sequence to dwell for approximately one hour on the appropriate level when PFP is being sampled.

A modified version of the PCP has recently been developed that includes an integrating volume and uses a variable flow rate to provide integrated sampling over

~ 1 h (Turnbull et al., 2012). Whether integrated versus grab sampling is appropriate for a particular application depends on several factors, especially proximity to emissions sources. CO_2 and CO quality control is just one aspect of the PFP sampling objectives, and we plan to continue with grab sampling until we are able to thoroughly evaluate an integrating sampler.

6.3.3 Long-term Picarro : Licor comparison at WGC

The agreement between the WGC Licor and Picarro CO_2 measurements is shown in Fig. 13 for 1–31 July 2011. This was a period where the room temperature was reasonably well controlled, but the level of agreement is representative of the entire 5-yr record. Differences between the analyzers during calibration measurements show no detectable bias and are normally distributed (Fig. 13b) with a standard deviation of 0.04 ppm. For the ambient air comparison, the data were filtered to exclude periods of high variability. Data with 30-s standard deviations > 0.3 ppm were excluded, corresponding to 32 % of the available observations. Since the analyzers share standard gases that span a wide range of CO_2 concentrations, it is not surprising that the bias is negligible. However, the post-processing for the two-analyzers differs significantly in that a time and/or temperature-dependent baseline is subtracted from the Licor data, and the first order calibration coefficients are temporally interpolated between 6-hourly calibration cycles, whereas no baseline is subtracted and a 3-day average first-order calibration curve is used for the Picarro.

The lifetime of the reference gases at WGC is shorter than at other sites because of the increased frequency of calibrations to compensate for poor temperature control and because the additional gas is used to calibrate the Picarro. The CO2C3, CO2C1, and CO2C4 cylinders are the longest-lived and typically last ~ 18 months. Figure 14 shows the un-calibrated Licor and Picarro data corresponding to repeated measurements of a single CO2C3 cylinder over 16.3 months. The standard deviation of the Picarro measurements is 0.05 ppm, whereas the Licor signal varies by ~ 20 ppm with discontinuities that correspond to Licor reference gas replacements and a power outage. We are able

to reliably correct for variations in the Licor signal with hourly baseline checks, as evidenced by the repeatability of our target tank measurements (see e.g. Fig. 11) and by the excellent agreement between the post-processed data from the Picarro and the Licor as shown in Fig. 13 and described above. However, the effort and expense associated with frequent calibrations and standard gas replacements is substantial.

The short-term precision of the Picarro analyzer (i.e. 30-s standard deviation) is 0.04 ppm, which is consistent with the range of values observed in Fig. 14a. There is a step-change in the Picarro signal of ~ 0.1 ppm that corresponds to a period in August 2011 when the CO2C2 standard was offline, resulting in no flow through the analyzer for the 5-min intervals immediately preceding the CO2C3 measurements. The values returned to their previous mean when CO2C2 flow was restored. The three gaps in the Picarro timeseries correspond to: (1) a period when the Picarro cell pressure did not return to the setpoint after a power failure, (2) a Picarro software problem when the counter of spectral scans exceeded 10^6 (out of range) that can be prevented by periodic rebooting of the Picarro computer, and (3) a failure of the power supply for the pumps and flow controllers.

6.3.4 Comparison with other continuous measurements and airborne measurements

We have had several opportunities to compare our CO_2 measurements with other sensors. Results are summarized in Table 8, some of which have been published previously. The experiments varied in duration, site, season, time averaging and filtering strategies to remove periods with high variability. Agreement is within the combined uncertainties of the measurements and close to the WMO recommendation for compatibility of independent measurements of 0.1 ppm (WMO, 2011), with the exception of the summertime WBI comparison with the Penn State University Cavity Ringdown Spectrometer during 2009. Details of that comparison are presented in Richardson et al. (2012), but the source of the 0.3 ppm difference is unknown and points to the difficulty of achieving the WMO goal. Possible contributors are small leaks in the sample

tubing or artifacts related to uncalibrated inlet components under conditions of high humidity. The BAO tower is a fantastic resource, where we have easy access to the sample tubing on the tower for frequent leak checks and the ability to install additional sensors any time. We have shown repeatedly at BAO that comparability of 0.1 ppm can be achieved with well-characterized, independent analysis systems. The only drawback of tests from BAO is that very high humidity is rare, whereas sites like LEF, WBI, WKT, and SCT routinely experience humidity levels up to 3 % in summer. The two WKT aircraft spiral comparisons with a well-calibrated analyzer on the NOAA P-3 occurred on days with χ_{H_2O} of ~ 1 %, which is relatively low for that area and season. The long-term stability of calibration residuals and target measurements demonstrates that the calibrated portions of the analysis system are insensitive to high ambient humidity. Laboratory tests with wetted air (described earlier in this section) showed that artifacts under controlled conditions were < 0.2 ppm and arguably < 0.1 ppm, but more work is needed to unequivocally demonstrate < 0.1 comparability in the field under conditions of extreme humidity.

7 Recommendations

We have learned many lessons over the course of this work and have attempted here to concisely summarize the most critical in the form of recommendations. Many of these recommendations are already documented elsewhere (e.g. WMO, 2011, 2012) or are simply practical, and our experience further underscores their importance.

7.1 Modularity and automation

The modular design of our analytical system has greatly simplified maintenance and repair. Component level repairs are rarely if ever performed in the field. For eight field systems, we maintain one working system in the laboratory for testing components or proposed design changes, evaluating new gas analyzers, and other diagnostic testing

(e.g. attempting to replicate anomalies or suspected problems under controlled conditions). At least one complete set of spare modules is also needed. We have a few extra pump modules, since they require routine maintenance. The system should be entirely automated with minimal need for human attention and on-site diagnosis. Use of high-quality quick-connect fittings on reference gases and between modules minimizes or eliminates the need for trained technicians in the field. The control software should have a user-friendly interactive mode to enable remote troubleshooting, e.g. switching valves, power switching for certain components (e.g. pumps, heaters). It is convenient to have a separate system mode for manual operation (e.g. for troubleshooting remotely or during site visits), so that affected data can be automatically filtered.

7.2 Calibrations

Although modern CO_2, CO, and CH_4 spectrometers are extremely stable compared to the previous generation of analyzers, field calibrations are still needed to establish continuity and comparability within and across networks. Long-term stability of the analytical system is more important than the short-term precision (signal-to-noise ratio), since day-to-day, year-to-year, and site-to-site comparability is the relevant measure of uncertainty for data analysis. We recommend deploying any analyzer with two or more cylinders than required to generate a calibration curve. For example, an analyzer with a linear response should be deployed with at least four calibration standards, and an analyzer that requires only an offset correction should be deployed with at least three standards. This approach yields enough information to compute meaningful residuals from the calibration polynomial and one standard can be treated as a target that is not included in the regression. The standards should span the range of expected values and must be measured frequently enough to capture any temporal drift in the analyzer baseline or response. The measurement sequence should operate on a cycle such that calibrations capture any diurnal variation in analyzer response. Ideally the calibration cycle will have a period not equal to 24 h, so that gaps in the sampling do not always occur at the same time of day. Target standards should be measured so that they are not

temporally adjacent to full calibrations in order to maximize sensitivity to un-resolved analyzer drift. It would be useful to periodically calibrate the analyzer at a higher than normal frequency and/or to dwell on a reference gas long enough to detect high frequency variations in the analyzer sensitivity or baseline. Some types of analyzer may still require frequent baseline correction to provide useful data, which can be performed using inexpensive uncalibrated cylinders or, in some cases, a source of zero air. Standards that are used to generate instrument response curves should have the same composition of interferents as the sample air, and the isotopic composition of the calibrated species should be close to that of ambient air. Finally, whenever possible, multiple standards should not be replaced on the same day so that any problems related to improper installation or altered concentration can be unambiguously attributed to a particular cylinder.

7.3 Drying the sample airstream

There has been much debate about whether sample drying is necessary for CO_2 and CH_4 measurement systems using CRDS or other cavity enhanced spectroscopic techniques, since those methods potentially enable reliable correction for water vapor interference and dilution. The two lines of argument against drying are that it requires additional hardware that increases expense and complexity and that accurate water vapor measurements are intrinsically valuable. Sample drying is a requirement for our system because the water vapor correction for the Li-7000 analyzers is not sufficiently accurate or stable to meet our target precision for CO_2. Our experience demonstrates that, if needed, sample drying can be accomplished at a remote site with modest initial expense and minimal need for maintenance. By routing calibration gases through the Nafion dryer, we eliminate or render negligible any biases associated with CO_2 permeation across the membrane, as demonstrated by small calibration and target gas residuals. The upstream chiller and liquid alarm sensors ensure that the detectors and the Nafion dryer are not exposed to liquid water, which can cause swelling of the membrane and flow restriction, or to very high humidity, which may exacerbate

cross-membrane transport of CO_2. The humidification of standard gases to the same level as sample gas avoids abrupt transitions between dry standards and potentially humid ambient air that could result in long equilibration times or artifacts. Desiccant is consumed extremely slowly during normal operation, and replacement is needed only after many years (note that one site, WKT, has been operating for > 6 yr and desiccant has not been replaced). The only routine maintenance required is annual replacement of the peristaltic pump module, which involves a single quick-connect plumbing connection and a simple electronic connection. The pump module is returned to the laboratory for refurbishment, which simply involves replacing the compressible tubing and a few springs in the roller assembly.

7.4 Sample integrity and redundancy

Repeatability of target gas measurements is a useful measure of long-term analytical stability but is not sufficient to ensure the integrity of the data record. Comparison with totally independent data of comparable quality is the best measure of overall data uncertainty and provides redundancy to protect against gaps in the data record that can cause significant uncertainty in mean data, inferred trends, and estimated fluxes. Care must be taken to ensure that any components upstream of the point where calibration gases enter the sample line do not cause artifacts. For our system, this includes inlet filters, sample tubing, condensers and pumps. Testing should be done under a wide range of representative conditions and should be performed on aged as well as new components. Routine and preferably automated checks that inlet tubing is intact are necessary and could be simply achieved by installing remotely actuated valves at each intake and pressurizing or evacuating the lines. Wireless communication via radio modem provides reliable communication with tower-mounted sensors, and abundant power is readily available on all transmission towers.

7.5 Post-processing

Comprehensive status data for critical pressures, flow rates, and temperatures are necessary for detecting insidious problems such as cross-port leaks in the manifold. Automated alerts based on these data can provide near real time notification of a failure. Prior to the development of automated alerts for our system, problems sometimes went unnoticed for several days or occasionally much longer. Many times a problem can be solved remotely, such as when a pump fails to restart after a power outage. Other failures have been detected simply by monitoring the number and size of data files.

7.6 Estimating and reporting uncertainty

The time-dependent uncertainty algorithms that we have described adequately represent many of the main sources of error. It is inevitable that the analyzers deployed for long-term monitoring will experience periods of sub-optimal performance. Our algorithms facilitate quality control, and enable automatic filtering of data depending on the error tolerance for a particular application. The algorithms perform well for a variety of sensors with a wide range of performance specifications and characteristics and could be adapted for other modes of operation (e.g. un-dried CRDS) or for other analyzer types. Random and systematic uncertainties should be reported separately so that error estimates can be correctly propagated.

7.7 Tower height

It is difficult to justify the expense and complication associated with operating on very tall towers. During well-mixed periods, vertical gradients of CO_2 between 100 m and 400 m are typically < 0.1 ppm. At night, levels higher than 200 m are frequently decoupled from the surface, and vertical gradients frequently exceed 10 ppm. Under these conditions, the highest levels often sample remnants of the previous afternoons boundary layer from some distance upwind. Nighttime data are especially difficult to model

because of the steep vertical gradients near the surface and wind shear associated with the nocturnal jet. Tower lease, installation and maintenance costs are largely driven by height, and shorter towers are more abundant than very tall towers. Data from towers > 100 m a.g.l. would likely suffice for most carbon-budgeting applications with current models. Many studies rely primarily on afternoon data, and model residuals are generally much larger than 0.1 ppm. However, tall tower observations are extremely useful for evaluating the fidelity of boundary layer processes in models, especially when a full complement of meteorological measurements and additional trace-gas data are available. An effective strategy for carbon monitoring would be to maintain a small number of tall tower "super-sites" representative of a variety of environmental conditions, and a larger network of shorter tower installations with a simpler instrument suite.

7.8 Complementary measurements

Whenever possible, tower greenhouse gas measurements should be co-located with other observations that are useful for evaluating atmospheric transport models and that provide additional constraints on flux estimates. Measurements of meteorological parameters such as wind speed and direction, temperature and humidity should be included at two or more heights on the tower using high-quality and routinely calibrated sensors. Radiation shields for temperature and humidity probes should be mechanically aspirated and include flow or Hall effect sensors to verify adequate ventilation (French and May, 2004). Commercially available remote sensors such as microwave temperature profilers, pulsed Doppler Light Detection and Ranging (LiDAR) wind profilers, and laser ceilometers can provide detailed information about atmospheric structure and/or estimates of mixed layer height that are useful for evaluating model boundary layer parameterizations, especially when combined with other data that describe the surface energy budget, e.g. radiation and eddy covariance measurements. Additional trace gas measurements are also helpful. The North American Carbon Program Plan (Wofsy and Harriss, 2002) recommends bi-weekly aircraft profiles over surface monitoring sites. Solar occultation measurements from ground-based spectrometers such

as those used in the TCCON network (Wunch et al., 2011) along with co-located tower measurements and boundary layer height data would place strong constraints on estimates of surface fluxes. Tall tower eddy covariance measurements of CO_2 and H_2O fluxes (Berger et al., 2001) can potentially help to separate near- and far-field contributions to observed CO_2.

8 Conclusions

In situ measurement and communications technologies have improved dramatically over the last decade. For the first time, research-grade operational monitoring is feasible for CO_2, CH_4 and a growing suite of other important trace gases, but measurement requirements for future greenhouse gas monitoring efforts need to be carefully defined. Data records with high precision and long-term stability are needed to resolve annual mean gradients and trends of CO_2 and other greenhouse gases. Many processes that drive net biological and oceanic fluxes operate on timescales of decades to centuries, so multi-decade records are needed to diagnose the underlying mechanisms. If atmospheric data are to be used for emissions verification, or to inform policy more generally, then the data must be fully disclosed and documented with minimal delay. For both research and regulatory applications, the uncertainties must be well understood and thoroughly documented.

Several of the measurement comparisons described here meet the WMO recommended goal for compatibility of 0.1 ppm, but others fall short. Agreement better than 0.3 ppm is relatively easy to achieve, but is insufficient for emissions verification. Several research groups have developed and demonstrated robust detector calibration strategies that account for analyzer drifts and deliver records with long-term stability of calibration standard residuals and target measurements. The remaining challenges relate to sample integrity: are the sampling lines intact? Is the sample being modified en route to the detector? Is the sampling strategy adequate for capturing mean values over relevant timescales in the presence of typical variability? We have

outlined tractable solutions to address all of these issues and have shown that a network of high-quality sensors can be efficiently maintained. The analytical system and post-processing methods described here provide one model to inform future expanded monitoring efforts. The time-dependent uncertainty algorithms are flexible and robust and could be straightforwardly adapted to other species and analytical systems.

Supplementary material related to this article is available online at: http://www.atmos-meas-tech-discuss.net/6/1461/2013/amtd-6-1461-2013-supplement.pdf.

Acknowledgements. We thank Peter Bakwin for valuable advice based on his experiences as lead scientist for the NOAA tall tower network. Aris Legoretta and David (Zim) Sherman contributed to design and testing of the CO_2/CO system. Ron Teclaw helped to establish the LEF site and assisted with operations for more than a decade. Ranjan Muttiah, Joaquin Sanabria, and Onyango Okello provided support at WKT with cooperation from the Blackland Research and Extension Center. Local support at WBI has been provided by students from the University of Iowa, including Alicia Pettibone, Adam Beranek-Collins, Jameson Schoenfelder, Alex Bender, Robert Bullard, Ben Behrendt, and Andrew Hesselink, several of whom were sponsored by the NASA Iowa Space Grant Consortium. Site support at SNP has been provided by Temple Lee. Sonja Wolter, Jack Higgs, Doug Guenther, and Kelly Sours have contributed substantially to tower PFP sampling efforts. Colm Sweeney, Anna Karion, Huilin Chen and Tim Newberger have loaned Picarro CRDS analyzers for testing and have provided helpful advice especially regarding installation and calibration of Picarro analyzers. Anna Karion and Huilin Chen also provided helpful comments on this manuscript. We thank Andrew Crotwell for helpful comments on this manuscript and for advice especially regarding PFP analysis. Jeff Peischl and Tom Ryerson contributed data and results from comparisons with the NOAA P-3 CO_2 analyzer. Scott Richardson and Tasha Miles contributed time and effort to carry out and analyze data from the WBI CRDS comparison. Installation of equipment at AMT was made possible by a grant from the National Science Foundation Biocomplexity in the Environment Program (ATM-0221850). We thank Verizon Wireless for continuing support at AMT and Steve Wofsy and David Hollinger for assistance with establishing the AMT site. Installation of equipment at SCT was accomplished with funding provided by the DOE Office of Science – Terrestrial Carbon Processes program and performed under Contract No. DE-AC09-08SR22470. Roger Strand and others at Wisconsin Public Television have facilitated installation and maintenance of equipment at LEF. Deployment and operation of the Picarro analyzer at WGC is collaboration between NOAA ESRL and LBNL and has been supported by the California Energy Commission (CEC) Public Interest Environmental Research Program and the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under Contract No. DE-AC02-05CH11231. Ankur Desai's work at LEF has been supported by NSF Grant No. DEB-0845166.

References

Bakwin, P. S., Tans, P. P., Hurst, D. F., and Zhao, C. L.: Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program, Tellus B, 50, 401–415, 1998.

Berger, B. W., Davis, K. J., Yi, C. X., Bakwin, P. S., and Zhao, C. L.: Long-term carbon dioxide fluxes from a very tall tower in a northern forest: flux measurement methodology, J. Atmos. Ocean. Tech., 18, 529–542, 2001.

Brooks, B.-G. J., Desai, A. R., Stephens, B. B., Bowling, D. R., Burns, S. P., Watt, A. S., Heck, S. L., and Sweeney, C.: Assessing filtering of mountaintop CO_2 mole fractions for application to inverse models of biosphere-atmosphere carbon exchange, Atmos. Chem. Phys., 12, 2099–2115, doi:10.5194/acp-12-2099-2012, 2012.

Chen, H., Winderlich, J., Gerbig, C., Hoefer, A., Rella, C. W., Crosson, E. R., Van Pelt, A. D., Steinbach, J., Kolle, O., Beck, V., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Santoni, G. W., and Wofsy, S. C.: High-accuracy continuous airborne measurements of greenhouse gases (CO_2 and CH_4) using the cavity ring-down spectroscopy (CRDS) technique, Atmos. Meas. Tech., 3, 375–386, doi:10.5194/amt-3-375-2010, 2010.

Committee on Methods for Estimating Greenhouse Gas Emissions, and National Research Council: Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements, The National Academies Press, Washington, DC ,2010.

Crosson, E. R.: A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor, Appl. Phys. B, 92, 403–408, doi:10.1007/s00340-008-3135-y, 2008.

Dlugokencky, E. J., Myers, R. C., Lang, P. M., Masarie, K. A., Crotwell, A. M., Thoning, K. W., Hall, B. D., Elkins, J. W., and Steele, L. P.: Conversion of NOAA atmospheric dry air CH_4 mole fractions to a gravimetrically prepared standard scale, J. Geophys. Res.-Atmos., 110, D1803, doi:10.1029/2005jd006035, 2005.

French, B. and May, E.: USCRN Temperature Aspirated Sheild Modifications, NOAA Technical Note NCDC No. USCRN-04-06, available at: http://www1.ncdc.noaa.gov/pub/data/uscrn/documentation/program/technotes/TN04006AspiratedShieldMods.pdf (last access: 9 November 2004), 2004.

Gloor, M., Bakwin, P., Hurst, D., Lock, L., Draxler, R., and Tans, P.: What is the concentration footprint of a tall tower?, J. Geophys. Res.-Atmos., 106, 17831–17840, 2001.

Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J., Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and Michalak, A. M.: North American CO_2 exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, doi:10.5194/bg-9-457-2012, 2012.

Karion, A., Sweeney, C., Wolter, S., Newberger, T., Chen, H., Andrews, A., Kofler, J., Neff, D., and Tans, P.: Long-term greenhouse gas measurements from aircraft, Atmos. Meas. Tech. Discuss., 5, 7341–7382, doi:10.5194/amtd-5-7341-2012, 2012.

Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., Sweeney, C., Diaz, L. I., Martins, D., Shepson, P. B., and Davis, K. J.: Constraining the CO_2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmos. Chem. Phys., 12, 337–354, doi:10.5194/acp-12-337-2012, 2012.

Law, R. M., Peters, W., Rodenbeck, C., Aulagnier, C., Baker, I., Bergmann, D. J., Bousquet, P., Brandt, J., Bruhwiler, L., Cameron-Smith, P. J., Christensen, J. H., Delage, F., Denning, A. S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S. R., Kleist, J., Krol, M. C., Lin, S. J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Patra, P. K., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Takigawa, M., Vautard, R., Vermeulen, A. T., and Zhu, Z.: TransCom model simulations of hourly atmospheric CO_2: experimental overview and diurnal cycle results for 2002, Global Biogeochem. Cy., 22, Gb3009, doi:10.1029/2007gb003050, 2008.

Lee, T. R., De Wekker, S. F. J., Andrews, A. E., Kofler, J., and Williams, J.: Carbon dioxide variability during cold front passages and fair weather days at a forested mountaintop site, Atmos. Environ., 46, 405–416, doi:10.1016/j.atmosenv.2011.09.068, 2012.

Le Quéré, C., Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., Conway, T. J., Doney, S. C., Feely, R. A., Foster, P., Friedlingstein, P., Gurney, K., Houghton, R. A., House, J. I., Huntingford, C., Levy, P. E., Lomas, M. R., Majkut, J., Metzl, N., Ometto, J. P., Peters, G. P., Prentice, I. C., Randerson, J. T., Running, S. W., Sarmiento, J. L., Schuster, U., Sitch, S., Takahashi, T., Viovy, N., van der Werf, G. R., and Woodward, F. I.: Trends in the sources and sinks of carbon dioxide, Nat. Geosci., 2, 831–836, doi:10.1038/ngeo689, 2009.

Masarie, K. A., Langenfelds, R. L., Allison, C. E., Conway, T. J., Dlugokencky, E. J., Francey, R. J., Novelli, P. C., Steele, L. P., Tans, P. P., Vaughn, B., and White, J. W. C.: NOAA/CSIRO flask air intercomparison experiment: a strategy for directly assessing consistency among atmospheric measurements made by independent laboratories, J. Geophys. Res.-Atmos., 106, 20445–20464, doi:10.1029/2000jd000023, 2001.

Miles, N. L., Richardson, S. J., Davis, K. J., Lauvaux, T., Andrews, A. E., West, T. O., Bandaru, V., and Crosson, E. R.: Large amplitude spatial and temporal gradients in atmospheric boundary layer CO_2 mole fractions detected with a tower-based network in the US upper Midwest, J. Geophys. Res.-Biogeosciences, 117, G01019, doi:10.1029/2011jg001781, 2012.

Novelli, P. C., Elkins, J. W., and Steele, L. P.: The development and evaluation of a gravimetric reference scale for measurments of atmospheric carbon monoxide, J. Geophys. Res.-Atmos., 96, 13109–13121, doi:10.1029/91jd01108, 1991.

Novelli, P. C., Collins, J. E., Myers, R. C., Sachse, G. W., and Scheel, H. E.: Reevaluation of the NOAA-CMDL carbon monoxide reference scale and comparisons with co reference gases at NASA-LANGLEY and the Fraunhofer-Institut, J. Geophys. Res.-Atmos., 99, 12833–12839, doi:10.1029/94jd00314, 1994.

O'Keefe, A., Scherer, J. J., and Paul, J. B.: cw Integrated cavity output spectroscopy, Chem. Phys. Lett., 307, 343–349, 1999.

Peischl, J., Ryerson, T. B., Holloway, J. S., Parrish, D. D., Trainer, M., Frost, G. J., Aikin, K. C., Brown, S. S., Dube, W. P., Stark, H., and Fehsenfeld, F. C.: A top-down analysis of emissions from selected Texas power plants during TexAQS 2000 and 2006, J. Geophys. Res.-Atmos., 115, D16303, doi:10.1029/2009jd013527, 2010.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, P. Natl. Acad. Sci. USA, 104, 18925–18930, 2007.

Pillai, D., Gerbig, C., Ahmadov, R., Rödenbeck, C., Kretschmer, R., Koch, T., Thompson, R., Neininger, B., and Lavrié, J. V.: High-resolution simulations of atmospheric CO_2 over complex terrain – representing the Ochsenkopf mountain tall tower, Atmos. Chem. Phys., 11, 7445–7464, doi:10.5194/acp-11-7445-2011, 2011.

Popa, M. E., Gloor, M., Manning, A. C., Jordan, A., Schultz, U., Haensel, F., Seifert, T., and Heimann, M.: Measurements of greenhouse gases and related tracers at Bialystok tall tower station in Poland, Atmos. Meas. Tech., 3, 407–427, doi:10.5194/amt-3-407-2010, 2010.

Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., Karion, A., Miles, N. L., Richardson, S. J., Steinbacher, M., Sweeney, C., Wastine, B., and Zellweger, C.: High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air, Atmos. Meas. Tech. Discuss., 5, 5823–5888, doi:10.5194/amtd-5-5823-2012, 2012.

Richardson, S. J., Miles, N. L., Davis, K. J., Crosson, E. R., Rella, C. W., and Andrews, A. E.: Field testing of cavity ring-down spectroscopy analyzers measuring carbon dioxide and water vapor, J. Atmos. Ocean. Tech., 29, 397–406, doi:10.1175/jtech-d-11-00063.1, 2012.

Schuh, A. E., Denning, A. S., Corbin, K. D., Baker, I. T., Uliasz, M., Parazoo, N., Andrews, A. E., and Worthy, D. E. J.: Corrigendum to "A regional high-resolution carbon flux inversion of North America for 2004" published in Biogeosciences, 7, 1625–1644, 2010, Biogeosciences, 7, 2245–2245, doi:10.5194/bg-7-2245-2010, 2010.

Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R. L., Steele, L. P., Francey, R. J., and Denning, A. S.: Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2, Science, 316, 1732–1735, doi:10.1126/science.1137004, 2007.

Stephens, B. B., Miles, N. L., Richardson, S. J., Watt, A. S., and Davis, K. J.: Atmospheric CO_2 monitoring with single-cell NDIR-based analyzers, Atmos. Meas. Tech., 4, 2737–2748, doi:10.5194/amt-4-2737-2011, 2011.

Thompson, R. L., Manning, A. C., Gloor, E., Schultz, U., Seifert, T., Hänsel, F., Jordan, A., and Heimann, M.: In-situ measurements of oxygen, carbon monoxide and greenhouse gases from Ochsenkopf tall tower in Germany, Atmos. Meas. Tech., 2, 573–591, doi:10.5194/amt-2-573-2009, 2009.

Turnbull, J., Guenther, D., Karion, A., Sweeney, C., Anderson, E., Andrews, A., Kofler, J., Miles, N., Newberger, T., Richardson, S., and Tans, P.: An integrated flask sample

collection system for greenhouse gas measurements, Atmos. Meas. Tech., 5, 2321–2327, doi:10.5194/amt-5-2321-2012, 2012.

Vermeulen, A. T., Hensen, A., Popa, M. E., van den Bulk, W. C. M., and Jongejan, P. A. C.: Greenhouse gas observations from Cabauw Tall Tower (1992–2010), Atmos. Meas. Tech., 4, 617–644, doi:10.5194/amt-4-617-2011, 2011.

Winderlich, J., Chen, H., Gerbig, C., Seifert, T., Kolle, O., Lavrič, J. V., Kaiser, C., Höfer, A., and Heimann, M.: Continuous low-maintenance CO_2/CH_4/H_2O measurements at the Zotino Tall Tower Observatory (ZOTTO) in Central Siberia, Atmos. Meas. Tech., 3, 1113–1128, doi:10.5194/amt-3-1113-2010, 2010.

Welp, L. R., Keeling, R. F., Weiss, R. F., Paplawsky, W., and Heckman, S.: Design and performance of a Nafion dryer for continuous operation at CO_2 and CH_4 air monitoring sites, Atmos. Meas. Tech. Discuss., 5, 5449–5468, doi:10.5194/amtd-5-5449-2012, 2012.

Worthy, D. E. J., Platt, A., Kessler, R., Ernst, M., Braga, R., and Racki, S.: The Canadian atmospheric carbon dioxide measurement program: measurement procedures, data quality and accuracy, In: Report of the 11th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracer Measurement Techniques, edited by: Toru, S. and Kazuto, S., Tokyo, Japan, September 2001, World Meteorological Organization Global Atmosphere Watch, 112–128, 2003.

Wunch, D., Toon, G. C., Blavier, J. F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The total carbon column observing network, Philos. T. Roy. Soc. A, 369, 2087–2112, doi:10.1098/rsta.2010.0240, 2011.

WMO: Guidelines for the measurement of atmospheric carbon monoxide, GAW 192, available at: http://www.wmo.int/pages/prog/arep/gaw/gaw-reports.html (last access: 14 January 2013), World Meteorological Organization, Geneva, Switzerland, 2010.

WMO: Report of the 15th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases, and Related Tracers Measurement Techniques, Jena, Germany, 7–10 September 2009, GAW 194, available at: http://www.wmo.int/pages/prog/arep/gaw/gaw-reports.html (last access: 14 January 2013), World, Meteorological Organization, Geneva, Switzerland, 2011.

Wofsy, S. C. and Harriss, R. C.: The North American Carbon Program Plan (NACP), Report of the NACP Committee of the US Carbon Cycle Science Program, US Global Change Research Program, Washington, DC, 2002.

Zhao, C. F., Andrews, A. E., Bianco, L., Eluszkiewicz, J., Hirsch, A., MacDonald, C., Nehrkorn, T., and Fischer, M. L.: Atmospheric inverse estimates of methane emissions from Central California, J. Geophys. Res.-Atmos., 114, D16302, doi:10.1029/2008jd011671, 2009.

Zhao, C. L. and Tans, P. P.: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, J. Geophys. Res.-Atmos., 111, D08s09, doi:10.1029/2005jd006003, 2006.

Zhao, C. L., Bakwin, P. S., and Tans, P. P.: A design for unattended monitoring of carbon dioxide on a very tall tower, J. Atmos. Ocean. Tech., 14, 1139–1145, 1997.

Table 1. Site information.

Site	Start date	Location	Lat.	Lon.	Surface elev. (masl)	Intake heights (magl)	Partners
LEF	Oct 1994 Upgrade May 2009	Park Falls, WI	45.9451	−90.2732	472	30, 122, 396, 11[a] 76[a], 244[a]	Penn State U of WI US Forest Service
WKT	Feb 2001 Upgrade May 2006	Moody, TX	31.3149	−97.3269	251	30, 122, 457, 9[a], 61[a], 244[a]	Blackland Research and Extension Center
BAO	May 2007	Erie, CO	40.0500	−105.0040	1584	22, 100, 300	
AMT	Sep 2003 Upgrade Feb 2009	Argyle, ME	45.0345	−68.6821	50	12, 30[b], 107	Harvard U of ME US Forest Service
WBI	Jul 2007	West Branch, IA	41.7248	−91.3529	241.7	31, 99, 379	U of IA
WGC	Sep 2007	Walnut Grove, CA	38.2650	−121.4911	0	30, 91, 483	Lawrence Berkeley National Laboratory
SCT	Aug 2008	Beech Island, SC	33.4057	−81.8334	115	30, 61, 305	Savannah River National Laboratory
SNP	Aug 2008	Shenandoah National Park, VA	38.6170	−78.3500	1008	5, 10, 17	U of VA

[a] Sampling at these heights was discontinued at time of upgrade. [b] Additional sampling level added at time of upgrade.

Table 2. Signal list.

Timestamp	Analyzer enclosure temperature
Analyzer signals (CO_2, CO, CH_4)	Analyzer internal temperature
Water content of sample flow (CO_2, CO, CH_4)	Room Temperature
Sample flow through each analyzer	Pump box temperature
Analyzer pressures	Chiller element temperature
Licor CO_2 Analyzer reference flow	Nafion box purge flow
Gas cylinder pressures	Combined analyzer exhaust pressure
Bypass flow from each sampling height	Manifold/valve position (SYSMODE)
Bypass back pressure for each sampling height	Liquid alarm status for each sampling height

Table 3. System modes.

	SYSMODE	Description	Approximate concentration	Interval (hours)
CO_2 Calibration	C1	Standard gas	350 ppm	12
	C2/ZER	Standard gas	380 ppm	1–2
	C3	Standard gas	410 ppm	12
	C4	Standard gas	460 ppm	12
	TGT	Standard gas	400 ppm	6
CO Calibration	C1	Standard gas	100 ppb	12
	C2	Standard gas	350 ppb	12
	TGT	Standard gas	220 ppb	6
	ZER	Scrubbed ambient air	0 ppb	0.5–0.75
Sampling	L1	Lowest inlet		0.25
	L2	Middle inlet		0.25
	L3	Highest inlet		0.25

Table 4. Typical values for uncertainty components (WGC July 2011).

SENSOR		CO_2 Licor ppm	CO_2 Picarro ppm	CO Thermo-electron ppb	CH_4 Picarro ppb
SHORT TERM	Median	0.01	0.04	2.1	0.26
	95th%ile	0.02	0.06	4.3	0.45
FIT	Median	0.11	0.10	1.3	0.24
	95th%ile	0.14	0.11	3.7	0.57
BASELINE	Median	0.01	N/A	0.5	N/A
	95th%ile	0.02		1.8	
CAL EQUIL	Median	0.01	0.00	0.7	0.05
	95th%ile	0.02	0.01	2.2	0.24
SMP EQUIL	Median	0.00	0.00	0.0	0.00
	95th%ile	0.01	0.01	1.0	0.08
H2O	Median	0.00	0.00	0.0	0.00
	95th%ile	0.00	0.00	0.0	0.01
EXTRAP	Median	0.00	0.00	0.0	0.01
	95th%ile	0.00	0.00	0.0	1.69
Total Analytical	Median	0.11	0.10	3.0	0.47
	95th%ile	0.14	0.12	5.4	1.78
30-s standard deviation	Median	0.30	0.18	2.7	0.84
	95th%ile	1.60	1.05	6.6	4.98
Target (Measured minus assigned)	50 % ± 1σ (n)	0.05 ± 0.03 (154)	0.06 ± 0.04 (155)	−3.3 ± 5.14 (89)	−0.21 ± 0.15 (155)
Calibration Scale	Accuracy	0.07 ppm		1 %	0.2 %
	Precision	0.03 ppm			1 ppb

Table 5. Annual summary of flask minus in situ χ_{CO_2}.

	Median ± standard deviation (number of samples) dry air mole fraction, ppm					
	2006	2007	2008	2009	2010	2011
LEF*	0.04 ± 0.5 (90)	−0.08 ± 0.4 (393)	−0.11 ± 0.3 (424)	0.04 ± 0.3 (184)		
				0.16 ± 0.3 (180)	0.15 ± 0.4 (318)	0.23 ± 0.5 (247)
LEF Manual	−0.08 ± 0.3 (46)	0.02 ± 0.2 (53)	−0.04 ± 0.3 (65)	0.04 ± 0.2 (46)	−0.13 ± 0.9 (56)	0.05 ± 0.2 (54)
WKT*	0.03 ± 0.3 (107)	−0.25 ± 0.8 (63)				
		0.03 ± 0.1 (23)	0.00 ± 0.3 (189)	0.06 ± 0.3 (180)	0.05 ± 0.4 (264)	0.11 ± 0.4 (257)
AMT*			0.4 ± 0.3 (23)			
				−0.04 ± 0.3 (94)	0.19 ± 0.4 (246)	0.15 ± 0.4 (235)
BAO		−0.01 ± 0.7 (31)	−0.07 ± 0.4 (215)	−0.09 ± 0.5 (211)	0.15 ± 0.5 (235)	0.18 ± 0.5 (217)
WBI		0.13 ± 0.4 (115)	0.10 ± 0.5 (198)	0.28 ± 0.4 (181)	0.29 ± 0.6 (270)	0.23 ± 0.5 (231)
WGC		0.15 ± 0.5 (31)	0.10 ± 0.4 (115)	0.13 ± 0.5 (160)	0.11 ± 0.5 (140)	0.15 ± 0.6 (64)
SCT				0.12 ± 0.4 (242)	0.27 ± 0.6 (275)	0.18 ± 0.4 (210)

* New rows within a site entry correspond to significant configuration changes as described in the text.

Table 6. Annual summary of flask minus in situ χ_{CO}.

	Median ± standard deviation (number of samples) dry air mole fraction, ppb					
	2006	2007	2008	2009	2010	2011
LEF				2.0 ± 2.9 (283)	2.0 ± 2.8 (404)	2.5 ± 2.9 (282)
LEF Manual				1.9 ± 9.5 (44)	3.7 ± 12.2 (80)	3.8 ± 8.0 (89)
WKT*	−0.8 ± 4.9 (171)	−1.2 ± 4.8 (263)				
		1.3 ± 4.1 (34)	1.6 ± 4.4 (297)	0.2 ± 6.9 (270)	2.9 ± 5.1 (379)	2.0 ± 5.2 (281)
AMT*			1.6 ± 2.9 (31)			
				1.8 ± 8.1 (111)	2.4 ± 7.3 (302)	Not yet available
BAO		−2.6 ± 13.1 (64)	−0.7 ± 6.3 (326)	−3.4 ± 7.1 (305)	−2.7 ± 6.4 (330)	−3.2 ± 7.1 (282)
WBI		−0.1 ± 4.0 (262)	1.0 ± 5.1 (311)	0.6 ± 4.9 (377)	1.0 ± 5.5 (460)	1.3 ± 5.7 (343)
WGC		0.7 ± 4.6 (62)	1.8 ± 4.6 (268)	2.5 ± 5.0 (288)	1.9 ± 8.2 (308)	3.1 ± 4.3 (151)
SCT		(151)	−0.8 ± 7.3 (18)	0.7 ± 5.5 (413)	0.9 ± 5.1 (418)	1.6 ± 5.0 (273)

* New rows within a site entry correspond to significant configuration changes as described in the text.

Table 7. WGC Picarro comparisons.

		Median ± standard deviation (number of samples) dry air mole fraction, ppb				
		2007	2008	2009	2010	2011
CO_2 (ppm)	Picarro minus PFP	0.17 ± 0.4 (32)	0.07 ± 0.4 (93)	0.16 ± 0.4 (150)	0.10 ± 0.5 (136)	0.16 ± 0.6 (72)
CH_4 (ppb)	Picarro minus PFP	3.1 ± 3.2 (31)	−1.1 ± 3.4 (134)	−0.3 ± 3.1 (188)	0.24 ± 2.8 (203)	−0.8 ± 3.1 (89)

Table 8. Summary of comparisons between NOAA ESRL Tall Tower Licor CO_2 and other CO_2 measurements.

Site	Date	Instruments	CO_2 Difference (other – NOAA TT)	Comparison type
BAO	9–12 Sep 2011	Picarro • 30 s	0.04 ± 0.06 (n = 6982)	Independent intake to 300 m (Rella et al., 2012 and Fig. 12)
	30 Sep–28 Oct 2011	• hourly, std < 0.3	0.0 ± 0.03 (n = 193)	
BAO	15–18 Nov 2007	P-3 Licor • 30-s, std < 0.5 (intermittent)	0.16 ± 0.20 (n = 118)	P-3 instrument on elevator
BAO	29 Jul–1 Aug 2008	P-3 Licor • 30-s, std < 0.5	0.04 ± 0.06 (n = 3130)	Shared intake line to 300 m
WBI	Jan–Oct 2010	PSU CRDS • 5-min • afternoon average • Jul/Aug 16:00–17:00	−0.12 ± 1.37 −0.13 ± 0.63 −0.33 ± 0.83	Separate intake to 99 m (Richardson et al., 2012)
WKT	13 Sep 2006 25 Sep 2006	P-3 Licor analyzer • 10 min (duration of spiral)	0.02 ± 0.17 −0.03 ± 0.23	Aircraft Spiral (Peischl et al., 2012)
BAO	1 Apr 2008	P-3 Licor analyzer • 20 min (duration of spiral)	0.01 ± 0.27	Aircraft Spiral (Peischl et al., 2012)

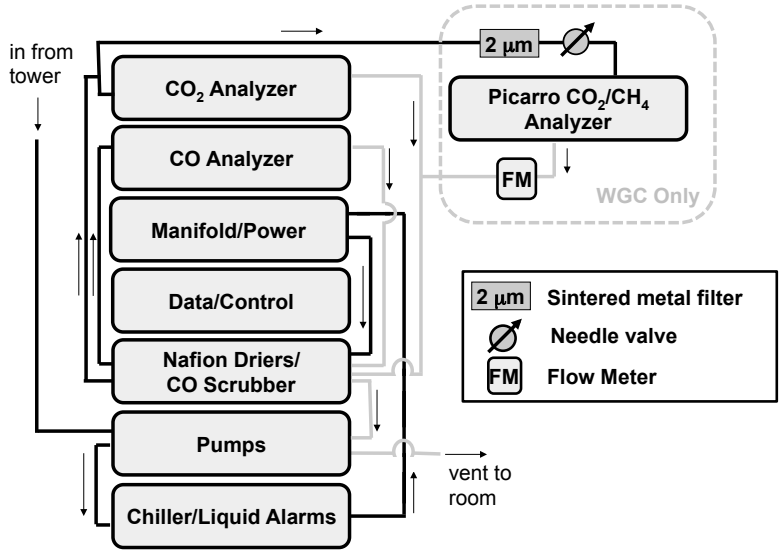

Fig. 1. Schematic diagram of the NOAA ESRL Tall Tower CO_2/CO analysis system.

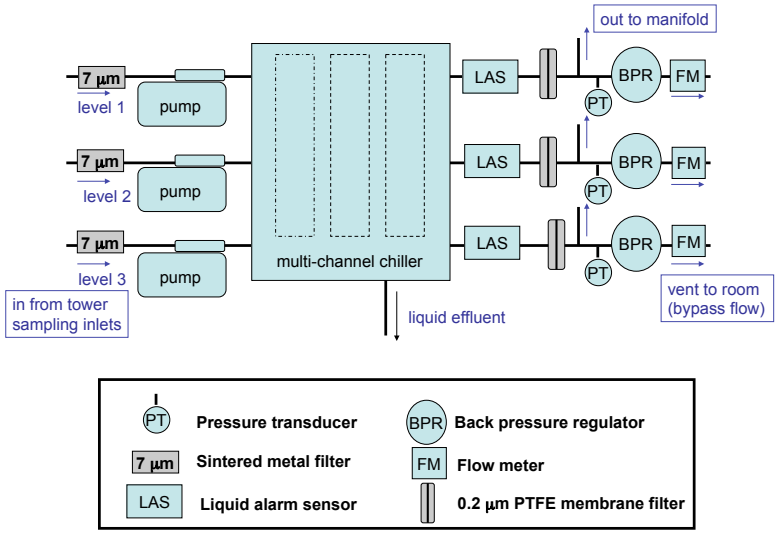

Fig. 2a. Pump and chiller components. The back pressure regulators, flow meters, and pressure transducers are shown here for convenience but are physically located in an enclosure with the sample/calibration manifold.

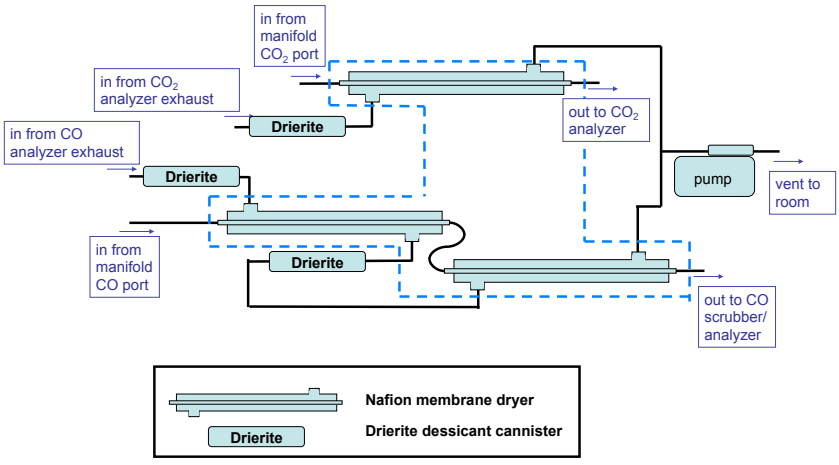

Fig. 2b. Nafion drier assembly. The exhaust pump is shown here for convenience, but is physically located in an enclosure with the inlet pumps. Components within the blue dashed lines are temperature controlled to ~ 20°C.

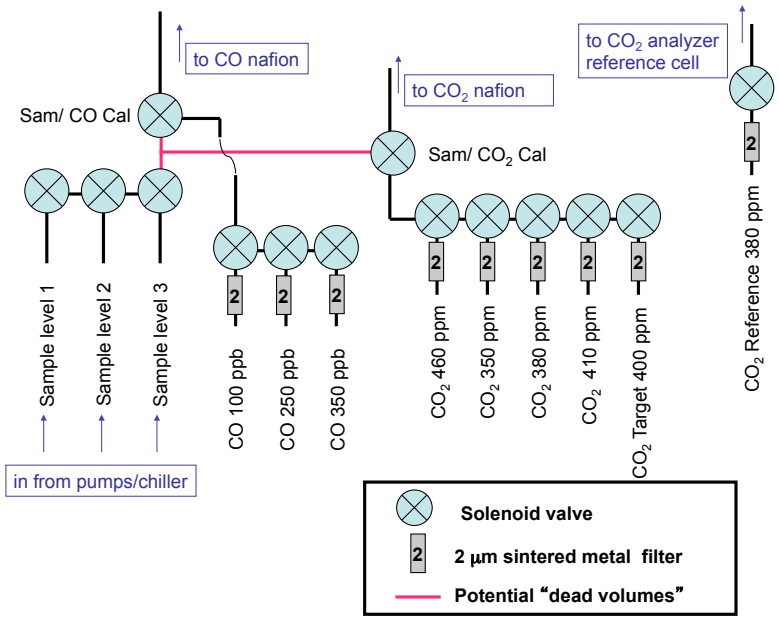

Fig. 2c. Calibration and sample selection manifold. Two-way versus three-way solenoids can be distinguished based on the number of connections indicated.

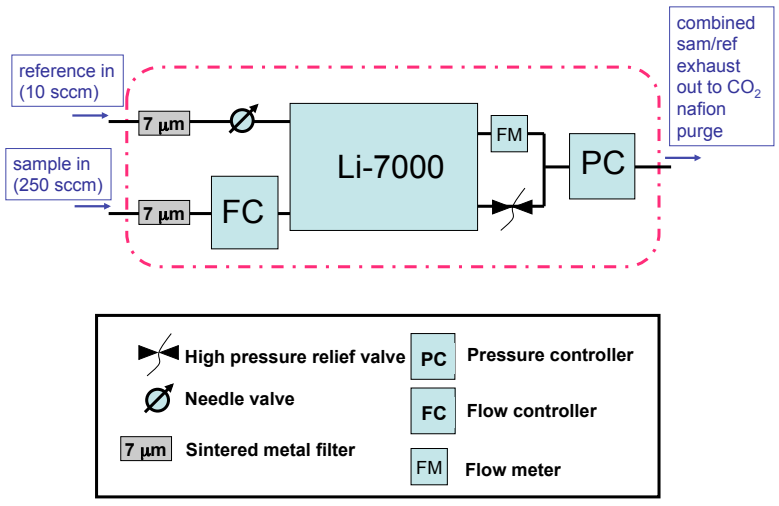

Fig. 2d. Licor CO_2 analyzer assembly. Components within the magenta dot-dashed box are temperature controlled 10–15 ˚C above typical room temperature.

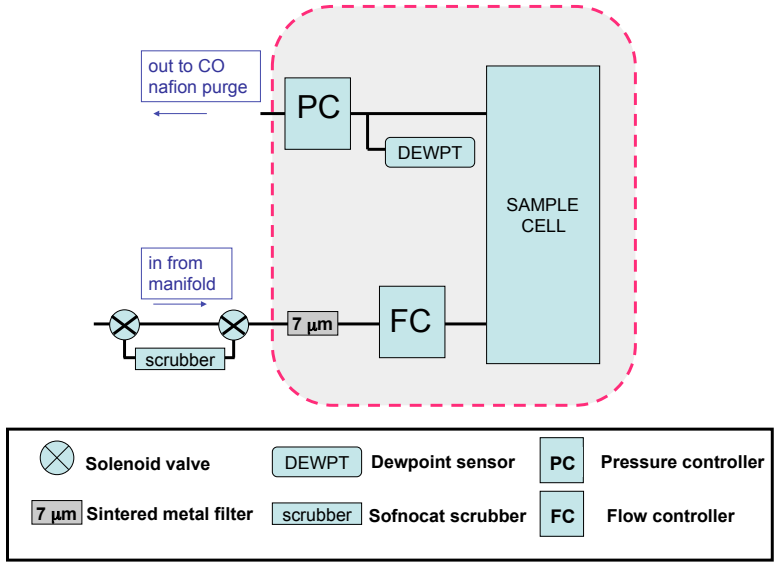

Fig. 2e. CO analyzer assembly. Components within the magenta dot-dashed box are temperature controlled 10–15 °C above typical room temperature. The Sofnocat scrubber is housed in a separate enclosure.

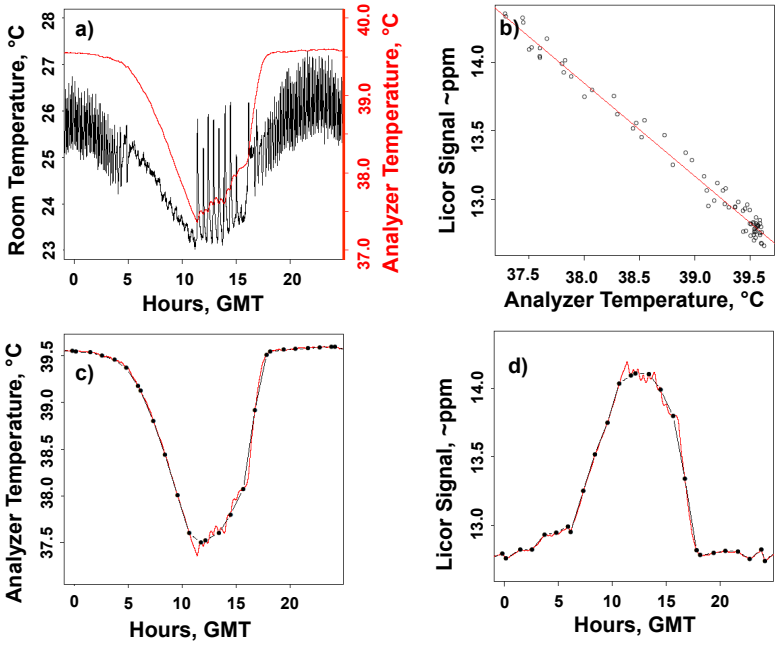

Fig. 3. Poor temperature control performance for the CO_2 analyzer at WGC on 7 August 2010. **(a)** Room temperature (black curve, left axis) and Licor cell temperature (red curve, right axis); **(b)** LicorCO2C2 (baseline) signal versus analyzer temperature; **(c)** Licor cell temperature (red curve) and the same temperature signal sampled corresponding to CO2C2 measurements (black filled circles) and interpolated to all times (black connecting lines); **(d)** CO2C2 measurements (black filled circles) interpolated linearly in time (black connecting lines) and estimated for all times using the slope from panel **(b)** multiplied by the difference between the black and red lines in panel **(c)**.

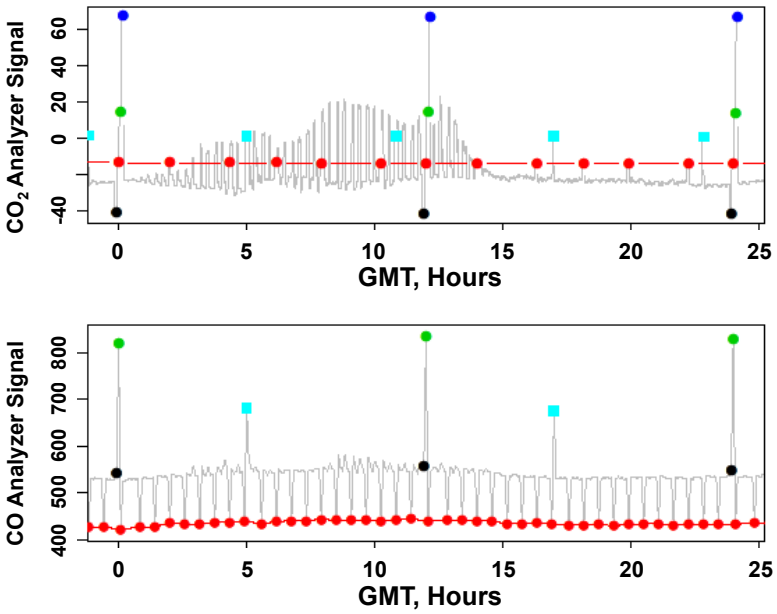

Fig. 4. Analyzer signals for a typical daily measurement cycle from the LEF tall tower site (Park Falls, WI; 6 August 2009) for **(a)** CO_2 and **(b)** CO. Filled circles correspond to calibration standards (CO_2) or scrubbed ambient air (CO) measurements that are used to track the analyzer baseline (red = CO2C2, COZER) and to determine the calibration polynomial (black = CO2C1, COC1; green = CO2C3, COC2; dark blue = CO2C4). Note that CO2C2 is typically also used in the calculation of the calibration polynomial. Red connecting lines show how a continuous estimate of the analyzer baseline is constructed by linearly interpolating between consecutively measured values. Target (CO2TGT, COTGT) measurements (cyan squares) are treated as unknowns and used to monitor system performance.

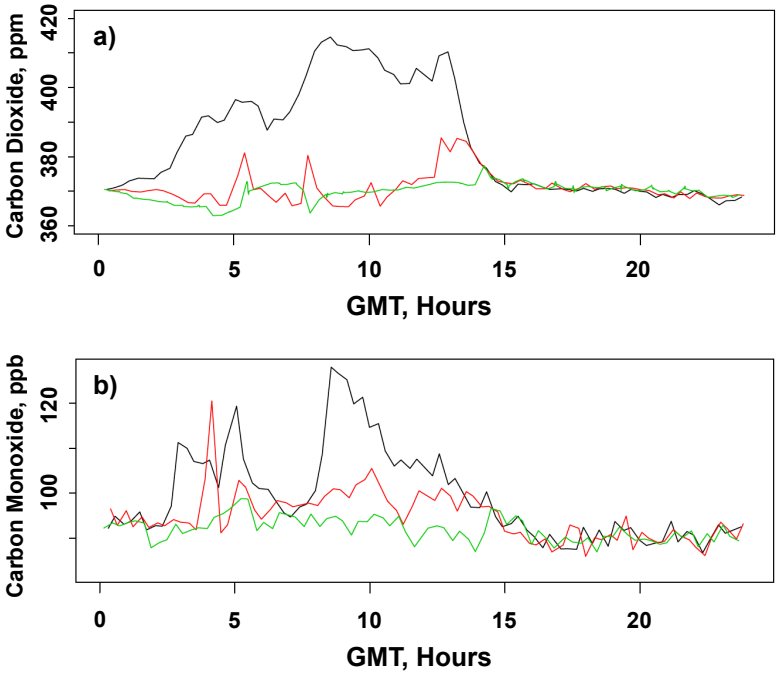

Fig. 5. (a) Time series of the observed χ_{CO_2} and **(b)** χ_{CO} computed from analyzer signals shown in Fig. 4 for the 30 m (black), 122 m (red), and 396 m (green) sampling heights of the LEF tall tower from 6 August 2009. Local Standard Time for the LEF site is 6 h behind GMT.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

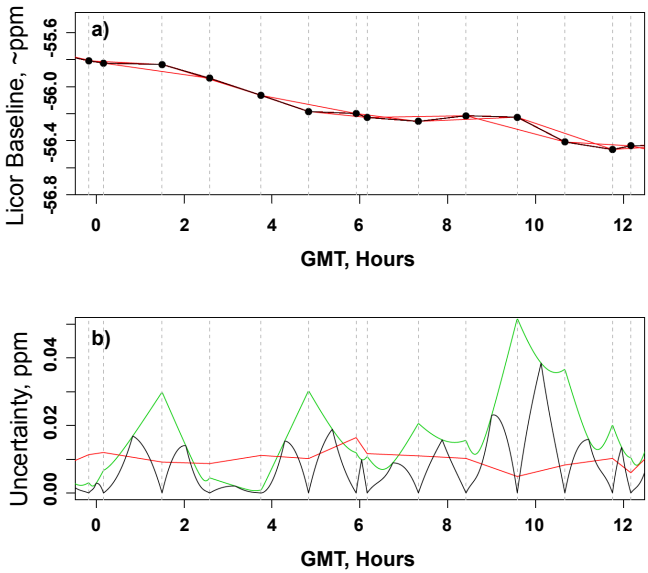

Fig. 6. (a) Measured Licor-7000 baseline signal s_b (black filled circles), linearly interpolated in time (black lines), and alternate realizations of the baseline obtained by leaving out individual baseline measurements (red lines). **(b)** The analyzer short term precision, u_p (red), defined as the time-interpolated 30-s standard deviation of the individual baseline measurements, the standard deviation computed across all realizations of the analyzer baseline (green), and the analyzer baseline drift uncertainty, u_b (black), which is the green curve weighted by a function that varies linearly from 0 at times t_b to 1 at times halfway between sequential baseline measurements.

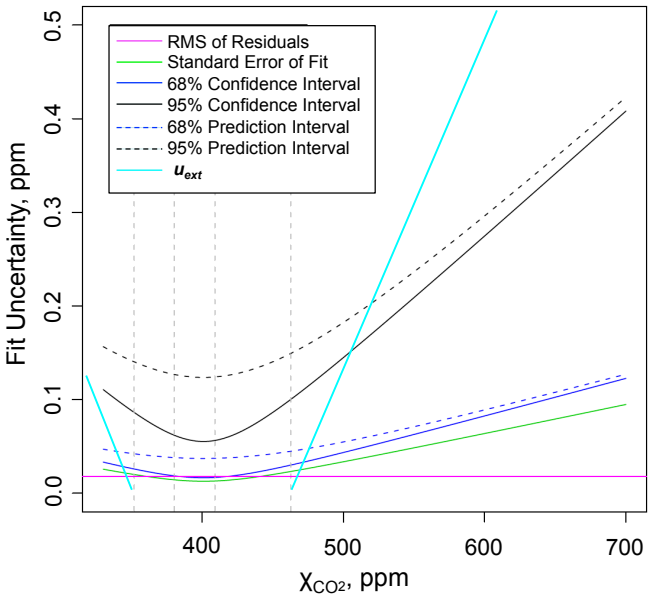

Fig. 7. Several different metrics for calibration curve fitting uncertainty corresponding to a typical field calibration of the Licor CO_2 analyzer. We use the 68 % prediction interval (blue dashed curve) to represent the curve fitting uncertainty, u_f. The extrapolation uncertainty, u_{ex}, (cyan) was determined empirically in the laboratory by measuring reference gases that were outside the calibrated range.

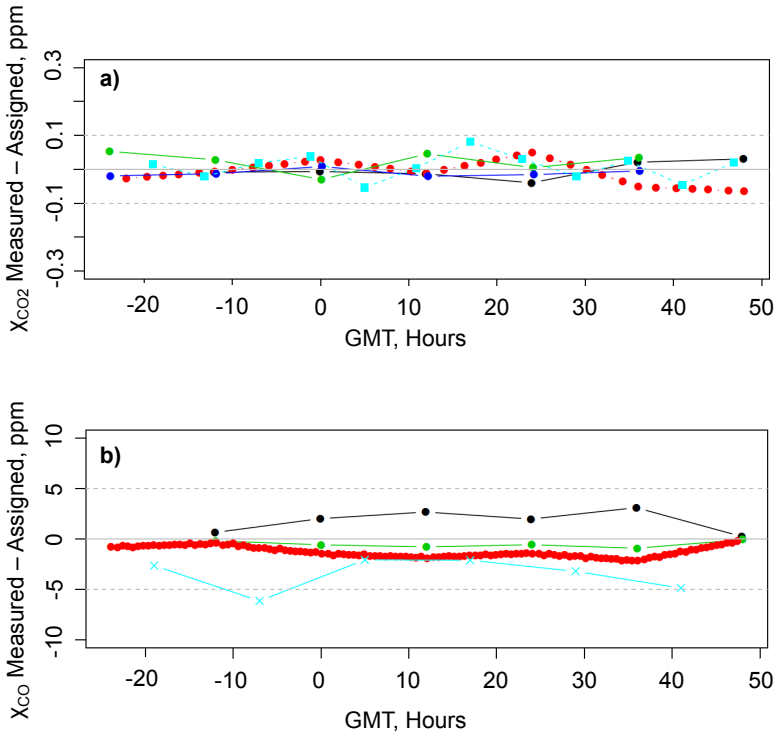

Fig. 8. Measured minus assigned values for **(a)** CO_2 standards (Black = CO2C1, Red = CO2C2, Green = CO2C3, Dark Blue = CO2C4, and Cyan = CO2TGT) and **(b)** CO standards (Red = COZER, Black = COC1, Green = COC2, Cyan = COTGT) for a three day period from the LEF tall tower where the central day is 6 August 2009.

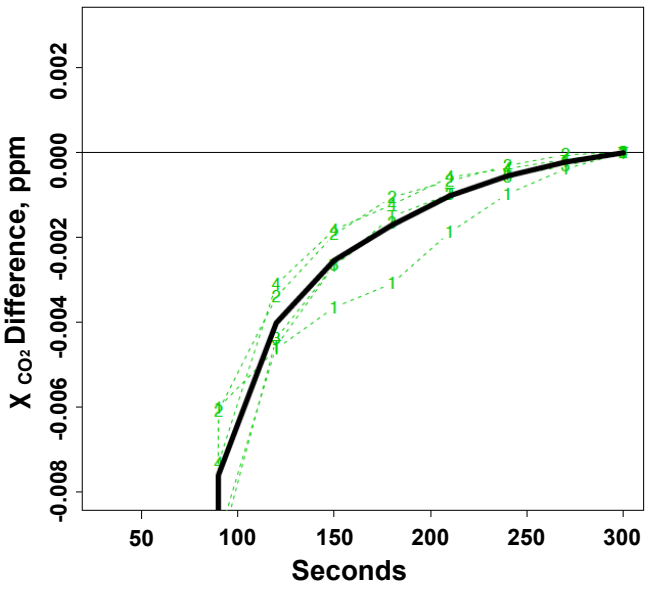

Fig. 9. Approach to equilibration for CO_2 calibration standards for WBI 5–7 August 2009 normalized by the $\Delta\chi_{CO_2}$ difference from the final χ_{CO_2} value from the previous sampling interval. Values represent the χ_{CO_2} difference from the 300-s value per ppm difference from previous sampling interval. The heavy solid line corresponds to the mean response computed over all the calibration and target modes. Green dashed lines with alphanumeric symbols correspond to individual standard gases (1 = CO2C1, 2 = CO2C2, 3 = CO2C3, 4 = CO2C4, T = CO2TGT).

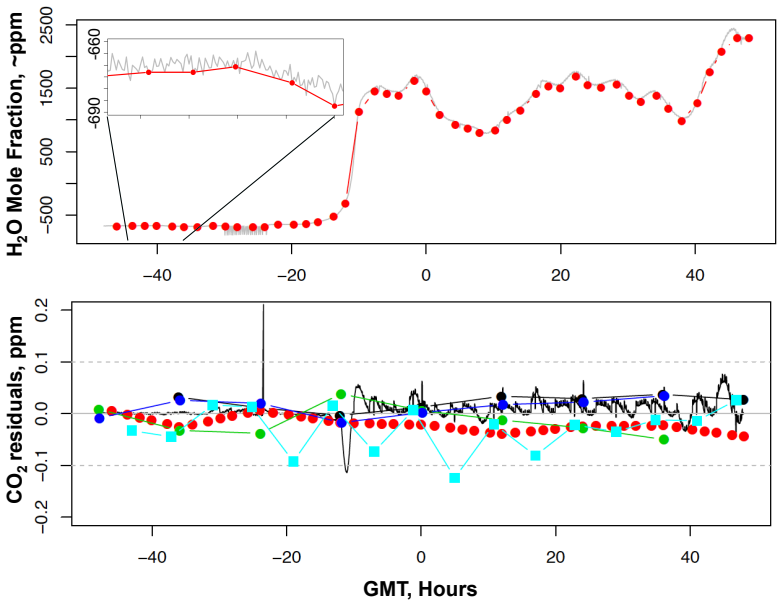

Fig. 10. (a) χ_{H_2O} reported by the Li-7000 at the SNP site for 13–15 February 2010 when the Nafion drier counter-flow was lost. Red symbols correspond to baseline calibration measurements. **(b)** The corresponding values of u_{wv} (black curve) along with the CO2 calibration and target residuals (CO2C1 = black circles, CO2C2 = red circles, CO2C3 = green circles, CO2C4 = blue circles, CO2TGT = cyan squares). Negative χ_{H_2O} values result from inaccurate (manufacturer-specified) zero-offset values for the Li-7000, but relative changes can be interpreted with some confidence.

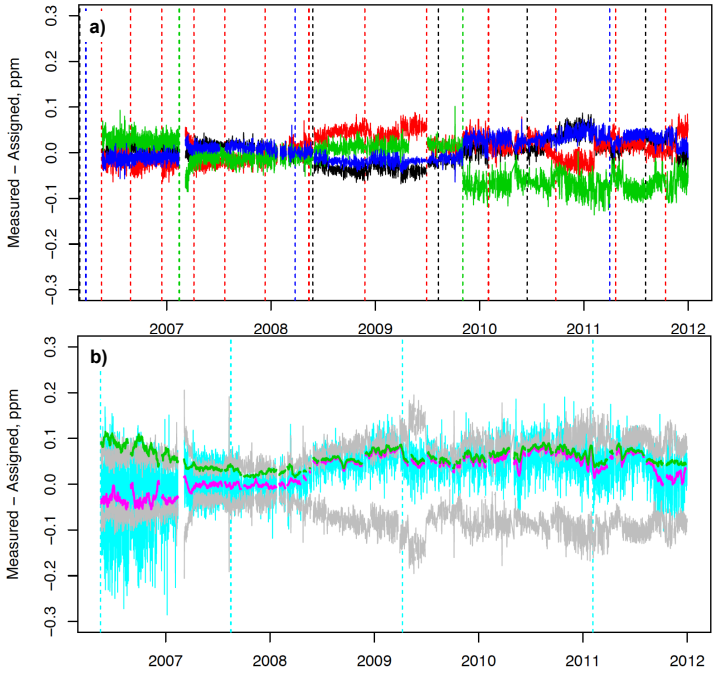

Fig. 11. Standard gas residuals (i.e. assigned minus measured χ_{CO_2} values) for **(a)** CO2C1 (black), CO2C2 (red), CO2C3 (green), CO2C4 (blue) calibration and **(b)** CO2TGT (cyan) standards at WKT for the period 2006–2011. Dashed vertical lines in both panels correspond to dates that standards were replaced. The grey curves in **(b)** correspond to the estimated analytical uncertainty. The 10-day running mean (magenta) and root-mean-square (green) residuals for CO2TGT are also shown.

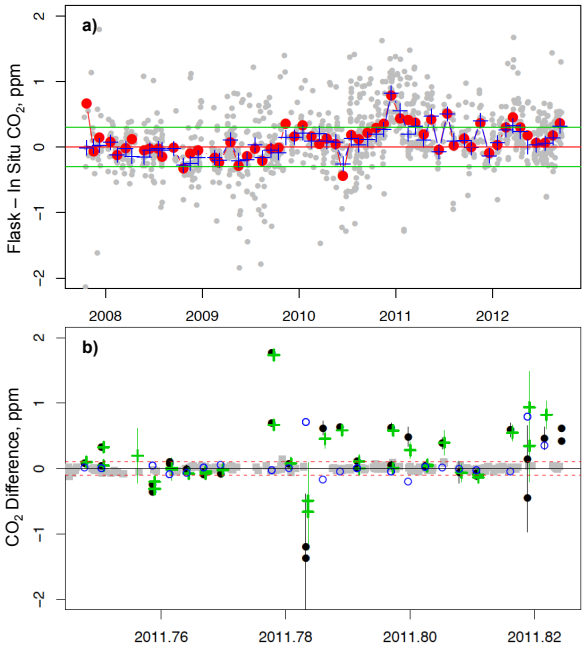

Fig. 12. (a) Comparison of individual (grey filled circles), monthly mean (red filled circles) and monthly median (blue crosses) PFP flask and in situ CO_2 measurements from the BAO tower for samples collected when the standard deviation of the in situ data within a 1.25 h window < 0.5 ppm. True pair samples were collected starting in January 2011. Horizontal lines correspond to ±0.3 ppm. **(b)** PFP minus in situ Licor (black filled circles), PFP minus in situ Picarro (green crosses), Licor minus Picarro corresponding to the PFP sample times (blue open circles) and Licor minus Picarro hourly averages for hours with standard deviations < 0.3 ppm (grey squares, N = 193).

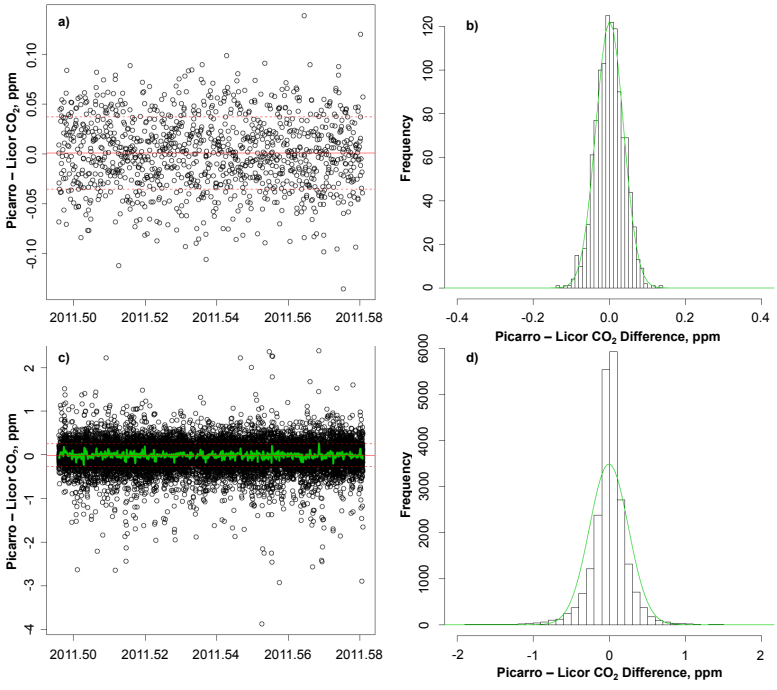

Fig. 13. Comparison of Licor and Picarro CO_2 analyzers at WGC. **(a)** Timeseries and **(b)** histogram of measurements of standard gases for the period 1–31 July 2011 with a mean difference of 0.00 ± 0.04 ppm (1σ). **(c)** Timeseries and **(d)** histogram for ambient air samples with 30-s standard deviations < 0.3 ppm. The mean difference is −0.01 ± 0.26 ppm (1σ).

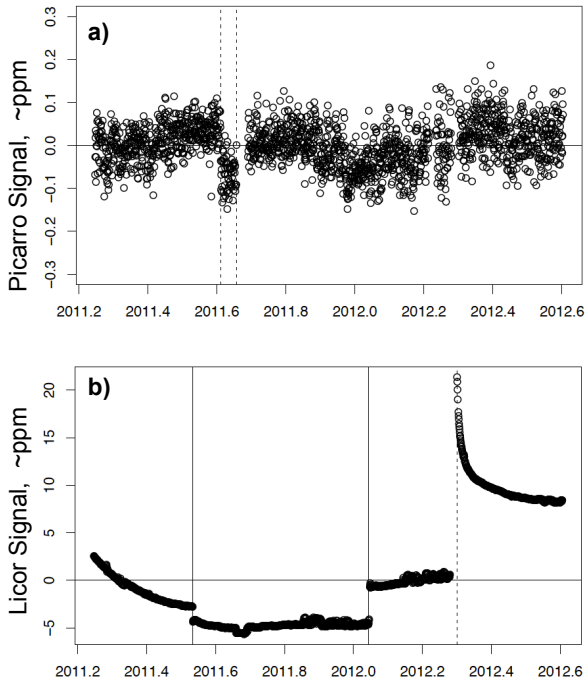

Fig. 14. Uncorrected signal minus the mean value for all CO2C3 measurements (assigned χ_{CO_2} = 407.77 ppm) or the period 1 April 2011 to 9 August 2012 for the WGC **(a)** Licor and **(b)** Picarro analyzers. Dashed vertical lines in **(a)** correspond to 12 and 29 August 2011, a period when the CO2C2 reference tank was offline, which caused flow and pressure disruptions in the Picarro sample cell. Solid lines in **(b)** correspond to dates when the CO_2 reference gas was changed. The dashed vertical line in **(b)** corresponds to 20 April 2012, when the analyzer was restarted after a power supply failure.