Atmos. Meas. Tech. Discuss., 5, 6949–6989, 2012
www.atmos-meas-tech-discuss.net/5/6949/2012/
doi:10.5194/amtd-5-6949-2012

Atmospheric Measurement Techniques Discussions

A statistical approach to quantify uncertainty in carbon monoxide measurements at the Izaña global GAW station: 2008–2011

A. J. Gomez-Pelaez[1], R. Ramos[1], V. Gomez-Trueba[1,2], P. C. Novelli[3], and R. Campo-Hernandez[1]

[1]Izaña Atmospheric Research Center, Meteorological State Agency of Spain (AEMET), Spain
[2]Air Liquide Canarias, Spain
[3]NOAA-ESRL-GMD, USA

Received: 17 August 2012 – Accepted: 30 August 2012 – Published: 21 September 2012

Correspondence to: A. J. Gomez-Pelaez (agomezp@aemet.es)

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

Abstract

Atmospheric CO in-situ measurements are carried out at the Izaña (Tenerife) global GAW mountain station using a ~~RGA (~~Reduction Gas Analyser). In-situ measurements at Izaña are representative of the subtropical North-East Atlantic free troposphere, specially during ~~the night period~~. We present the measurement system configuration, the response function, the calibration scheme, the data processing, the Izaña~~'s~~ 2008–2011 CO nocturnal time series, and the mean diurnal cycle by months.

We have developed a rigorous uncertainty analysis for carbon monoxide measurements carried out at the Izaña station which could be applied to other GAW stations. We determine the combined standard uncertainty from four components of the measurement: uncertainty of the WMO standard gases interpolated over the range of measurement, the uncertainty that takes into account the agreement between the standard gases and the response function used, the uncertainty due to the repeatability of the injections, and the propagated uncertainty related to the response function parameters ~~uncertainties~~ (which also takes into account the covariance between the parameters). The mean value of the combined standard uncertainty decreased significantly after March 2009, from 2.37 nmol mol^{-1} to 1.66 nmol mol^{-1}, due to improvements in the measurement system. A fifth type of uncertainty we call representation uncertainty is considered when some of the data necessary to compute ~~exactly~~ the mean are absent. Any computed mean has also a propagated uncertainty arising from the uncertainties of the data used to compute the mean. The law of propagation depends on the type of uncertainty component (random or systematic).

In-situ hourly means are compared with simultaneous and collocated NOAA flask samples. The uncertainty in the differences is ~~determined~~ and whether ~~these~~ are significant. For 2009–2011, only 24.5 % of the differences are significant, and 68 % of the differences are between −2.39 and 2.5 nmol mol^{-1}. Total and annual mean differences are computed using conventional expressions but also expressions with weights based

on the minimum variance method. The annual mean differences for 2009–2011 are well within the ±2 nmol mol^{-1} compatibility goal of GAW.

1 Introduction

Carbon monoxide affects the oxidizing capacity of the troposphere, and, in particular, plays an important role ~~on~~ the cycles of hydroxyl radical (OH), hydroperoxy (HO_2), and ozone (O_3); e.g. see Logan et al. (1981). Carbon monoxide atmospheric lifetime ranges from 10 days in summer over continental regions to more than a year over polar regions in winter (Novelli et al., 1992). Its relatively short lifetime and its uneven distribution of sources produce large temporal and spatial CO variations. The major sources of carbon monoxide are the combustion of fossil fuels, biomass burning, the oxidation of methane, and the oxidation of non-methane hydrocarbons. The major sink of CO is the reaction with OH, whereas surface deposition is a small sink (Ehhalt et al., 2001).

Comparisons of CO measurements among laboratories have shown differences larger than the quality objectives stated by the World Meteorological Organization (WMO) in its Global Atmosphere Watch Programme (GAW), WMO (2010). The Izaña station (28.309° N, 16.499° W, 2373 m a.s.l.) is located on the top of a mountain in the island of Tenerife (Canary Islands, Spain), well above a strong subtropical temperature inversion layer. In-situ measurements at Izaña are representative of the subtropical North-East Atlantic free troposphere, specially during the night period 20:00–08:00 UTC (e.g. Schmitt et al., 1988; Navascues and Rus, 1991; Armerding et al., 1997; Fischer et al., 1998; Rodríguez et al., 2009): air from below the inversion layer ~~can not~~ pass above it, and there is a regime of downslope wind produced by radiative cooling of the ground. The station is located on the top of a crest, where there is horizontal divergence of the downslope wind and subsidence of the air from above the station. During daytime an upslope wind produced by radiative heating of the ground transports to Izaña a small amount of contaminated air coming from ~~bellow~~ the subtropical temperature inversion layer (Fischer et al., 1998; Rodríguez et al., 2009),

producing a diurnal increase in carbon monoxide (Sect. 6). In this paper, we present the measurement system configuration, the response function, the calibration scheme, the data processing, the Izaña~~'s~~ 2008–2011 carbon monoxide nocturnal time series, and the mean diurnal cycle by months (Sects. 2, 3, and 6).

Reporting uncertainties associated with measurement results is strongly recommended by the WMO greenhouse gases measurement community (WMO, 2010, 2011). However, carrying out a rigorous uncertainty analysis taking into account uncertainty propagation and covariances between uncertainty components (JCGM, 2008) is a challenging task. In this paper, we present a rigorous uncertainty analysis for the carbon monoxide measurements carried out at the Izaña station (Sect. 4). The concepts presented here ~~my~~ be applied to other GAW stations.

The comparison between continuous (or quasi-continuous) measurements obtained by in-situ instruments and discrete measurements from collocated weekly flask samples analysed by another laboratory~~,~~ is an independent way of assessing the quality of the continuous in-situ measurements (WMO, 2011). ~~In this paper, as~~ part of our quality assurance procedure~~s~~, we compare the Izaña~~'s~~ in-situ quasi-continuous measurements with NOAA collocated flasks (Sect. 5). The difference between the measurements ~~is~~ evaluated in terms of their comparison uncertainty. Temporally averaged differences (e.g. annual means) also take into account the uncertaint~~ies of the differences~~.

2 Measurement system configuration

The ~~general~~ ambient air inlet line of the station is an 8 cm ID (inner diameter) stainless steel pipe that crosses the station building from the roof till the ground floor, with the entrance located 30 m ~~a.g.l.~~ A pump located on the ground floor produces a high flow rate (cubic meters per minute) of ambient air. On the third floor, there is a dedicated 4 mm ID PFA line that takes air from the general inlet using a KNF diaphragm pump ~~to the analytical system~~. Water vapour is removed by flowing the air through a 300 ml

glass flask immersed in a −67 °C alcohol bath. A multi-position selection valve (MPV) delivers ambient air or standard gas to the instrument.

The measurement system is based on a modified Trace Analytical gas chromatograph with mercuric oxide reduction detection (RGA). The RGA uses two chromatographic columns maintained at 105 °C: Unibeads 1S 60/80 mesh as pre-column, and Molecular Sieve 5A 60/80 mesh as main column. The pre-column separates CO and H_2 from other trace gases in an air sample. The main column separates H_2 and CO before entering a bed (265 °C) containing solid mercuric oxide. Reduced gases entering the bed are oxidized and HgO reduced to Hg, which is then measured by UV radiation absorption. High purity synthetic air is used as carrier gas. We use stainless ~~steal~~ sample loop volume of 1 ml. ~~The two columns and the sample loop are connected to a ten-port two-position injection valve. In one of the positions of such valve, there is back-flush in the pre-column. The back-flush removes water vapour, carbon dioxide and non-methane hydrocarbons.~~ Figure 1 shows a typical chromatogram, ~~appearing~~ the H_2 peak ~~firstly~~, ~~and then~~ the CO peak. ~~The system is controlled by a computer.~~ Working standard gas and ambient air are injected alternatively. There is an injection every ten minutes.

3 Standard gases, calibrations, response function, and processing

Instrument calibrations are performed every two weeks using 3–5 WMO CO standard gases. These CO-in-air mixtures were purchased from the WMO CO CCL (Central Calibration Laboratory). They range from 62.6 to 221.2 nmol mol^{-1} and are referenced to the WMO-2004 scale. These five high pressure cylinders serve as our ~~primary~~ laboratory standards. Table 1 shows their mole fractions with the 1-sigma uncertainty assigned in 2006 by the CCL. Before March 2009 we used 3 standard gases to define instrument characteristics, then five standards were used. ~~In order to minimize potential artefacts due to changes in instrument response with time, the response curves are~~

~~built using the ratio of the standard response to the response of a working gas (also called reference gas).~~

~~The standard and working high pressure tanks used are made of aluminium. Ceodeux brass valves with connection GCA-590 and Scott 14C two-stage high-purity regulators~~ are used, following the procedure described by Lang (1998) ~~to conditioning them~~. Working gas tanks ~~are~~ filled with natural air at the Izaña station using a filling system similar to that described by Kitzis (2009). The lifetime of a working gas high pressure tank is between 3 and 5 months (tanks are used till they reach 25 bar).

We determine the response function of the instrument based on the standard/reference peak height ratios:

$$r = r_{\mathrm{wg}} \left(\frac{h}{h_{\mathrm{wg}}} \right)^{\beta}, \tag{1}$$

where, r is CO mole fraction of the sample, h is peak height, and h_{wg} is the mean peak height of the bracketing working standard. In each calibration, the coefficients of the response function, r_{wg} and β, are obtained fitting (through least-squares) the mole fractions of the standards and the mean relative heights to the logarithm of the response function. From these definitions, it follows that r_{wg} is the working gas CO mole fraction. In this paper, carbon monoxide is measured by mole fraction (nmol mol^{-1}) on the WMO-2004 scale (WMO, 2011). To quantify the goodness of the fit, we use the RMS (Root Mean Square) residual,

$$u_{\mathrm{fit}} = \sqrt{\frac{\sum_{i=1}^{n} \left[r_i - R\left(h_i/h_{\mathrm{wg}}\right) \right]^2}{n-2}}, \tag{2}$$

where n is the number of standards, $n-2$ represents the number of degrees of freedom (JCGM, 2008) of the residuals (since the n standards have been used to compute two regression parameters) and $R(h_i/h_{\mathrm{wg}})$ is the fitted response function. Figure 2 shows

the least-squares fitting of a typical calibration, ~~whereas Table 2 shows~~ the residuals with respect to ~~the least-squares~~ fit ~~for such calibration~~. Figure 3 shows the working gas mole fractions and the response function exponents obtained from calibrations conducted during 2008 to 2011.

The time dependent response function for the working gas in use is computed using the response functions determined in its calibrations: β is computed as the mean of the calibration values, whereas a linear drift in time is allowed for r_{wg}. The CO mole fractions contained in high pressure cylinders are known to drift with time (e.g. Novelli et al., 2003). We evaluate potential drift in working standards using a Snedecor F statistical test (e.g. Martin, 1971, chapter 8) with the null hypothesis being "mole fraction is constant", and with its alternative being "linear drift in time". We require a 95 % confidence level to reject the null hypothesis. Constant mole fraction and the linear drift rate are computed using a least-squares fit with weights. The test takes into account the relative reduction of the chi-square computed with the residuals when using the linear drift instead of the constant mole fraction. To carry out the weighted least-squares fitting, a 1-sigma uncertainty for each value of r_{wg} has to be provided. The main advantage of using a Snedecor F test instead of a Chi-square test is that the 1-sigma uncertainties can be multiplied by a common factor without affecting the result of the test. Therefore, the test is not sensitive to the exact values of the uncertainties, only to their relative values. We have used u_{fit} as the 1-sigma uncertainties necessary. Six of the sixteen working gases used (see the upper graph of Fig. 3) show significant drift: five with rates ranging from −0.58 to −1.63 nmol ~~(mol month)$^{-1}$~~ and one with a positive drift of 2.75 nmol ~~(mol month)$^{-1}$~~. These changes ~~may~~ result from the interaction of CO with the internal surface of the cylinder, and the ~~large~~ decrease in ~~the~~ internal pressure (from 125 to 25 bar) ~~of the cylinder along~~ the few months they ~~last~~.

~~Once the response functions are known (correcting for drift in the working gases), mole fractions can be assigned to the Izaña air samples~~. Identification and discarding of outlayers uses an iterative process of three filtering steps. We begin considering the time series of working gas injections, in detail, the h_{wg}/r_{wg} time series. The first step

uses a running mean of 7 days and the RMS departure (σ_{run}) of the residuals is computed. Data with a departure from the running mean larger than $5\sigma_{run}$ are discarded. Note that the running mean is carried out only for evaluating data departures (i.e. it is not used for smoothing actual data). This procedure is run again with a 2 day running mean and a $4\sigma_{run}$ threshold for discarding. Lastly, a 0.19 day running mean and a $3.5\sigma_{run}$ threshold for discarding are used. Summarizing, 0.40 %, 0.64 %, and 0.61 % of the working gas injections were discarded in the first, second, and third step, respectively. The quality of measured air mole fractions is also considered. First, mole fractions are calculated only if both, the previous and the posterior working gas injections are present (3.11 % of the ambient air injections were discarded by this reason). As for the working gas injections time series, an iterative process of three filtering steps is ~~done~~ using running means of 30, 3, and 0.26 days, and thresholds $4.5\sigma_{run}$, $4\sigma_{run}$, and $3.5\sigma_{run}$ for the first, second, and third step, respectively. Summarizing, 0.11 %, 0.30 %, and 1.08 % of the ambient air samples were discarded in the first, second, and third step, respectively. Figure 4 shows daily night means (20:00–08:00 UTC) for the carbon monoxide mole fraction measured at Izaña Observatory. As indicated in Sect. 1, the air sampled at the station at night is representative of the free troposphere. Processed data are submitted to the WMO World Data Centre for Greenhouse Gases.

4 Uncertainty analysis

We compute the combined standard uncertainty for hourly means, as quadratic combination of four uncertainty components: the uncertainty of the WMO standard gases interpolated over the range of measurement (u_{st}), the uncertainty that takes into account the agreement between the standard gases and the response function used (u_{fit}), the uncertainty due to the repeatability of the injections (u_{rep}), and the propagated uncertainty ~~due~~ to the response function parameters ~~uncertainties~~ (u_{par}), which also takes into account the covariance between the parameters. ~~So, we use the following~~

equation:

$$u_{\mathrm{tot}} = \sqrt{u_{\mathrm{st}}^2 + u_{\mathrm{fit}}^2 + u_{\mathrm{rep}}^2 + u_{\mathrm{par}}^2}, \tag{3}$$

where

$$u_{\mathrm{st}} = 7.40 \times 10^{-5} r^2 - 1.80 \times 10^{-2} r + 1.92 \tag{4}$$

$$u_{\mathrm{rep}} = \frac{\beta r h_{\mathrm{wg}} \sigma_{h/h_{\mathrm{wg}}}}{\sqrt{3} h}, \tag{5}$$

$$u_{\mathrm{par}}^2 = u_{\mathrm{pr}}^2 + u_{\mathrm{p}\beta}^2 + c, \tag{6}$$

$$u_{\mathrm{pr}} = \frac{r}{r_{\mathrm{wg}}} \sigma_{r_{\mathrm{wg}}}, \tag{7}$$

$$u_{\mathrm{p}\beta} = r \sigma_\beta \log \frac{h}{h_{\mathrm{wg}}} \tag{8}$$

$$c = 2 \frac{r^2}{r_{\mathrm{wg}}} \mathrm{covar}\left(r_{\mathrm{wg}}, \beta\right) \log \frac{h}{h_{\mathrm{wg}}}, \tag{9}$$

u_{tot} is the combined standard uncertainty, the unit of u_{st} in Eq. (4) is nmol mol^{-1}, u_{fit} is defined in Eq. (2), $\sigma_{h/h_{\mathrm{wg}}}$ is the repeatability (standard deviation) of the relative height, which has been divided by $\sqrt{3}$ in Eq. (5) to take into account the improvement in repeatability due to using hourly means, $\sigma_{r_{\mathrm{wg}}}$ quantifies the consistence of the working gas mole fraction along its lifetime (RMS departure from linear drift or from constancy), σ_β is the standard deviation of the exponent, and $\mathrm{covar}\left(r_{\mathrm{wg}}, \beta\right)$ is the covariance between r_{wg} and β.

The term u_{st} (Eq. 4) was obtained through a least-squares fit of the standard uncertainties for the WMO standard gases of the Izaña station provided by Table 1. So, it represents the mole fraction dependent uncertainty due to the WMO standard

gases. The term u_{fit} takes into account the disagreement among the response function and the WMO standard gases. Note that the residuals of the standards in the calibrations can have an important systematic component that remains constant for the same standard gas between successive calibrations. Therefore, an hypothetical decrease of u_{fit} when combining the information of successive calibrations can not be considered. A mean value of u_{fit} is computed for each working gas used. As indicated in Sect. 3, 3 WMO standard gases were used before March 2009, with CO mole fractions, 83.9 nmol mol^{-1}, 151.6 nmol mol^{-1}, and 165.7 nmol mol^{-1}. In this case, the computed u_{fit} is abnormally small because the fitting is abnormally good due to the fact that the mole fractions of two of the standards are near. To avoid such underestimation, before March 2009, the u_{fit} used is forced to be at least equal to the mean value of u_{fit} after March 2009.

The terms $u_{\mathrm{rep}}^2 + u_{\mathrm{par}}^2$ in Eq. (3) come from the propagation of the response function uncertainty (JCGM, 2008). Taking differentials in Eq. (1),

$$dr = \frac{r}{r_{\mathrm{wg}}} dr_{\mathrm{wg}} + \beta r \frac{h_{\mathrm{wg}}}{h} d\left(\frac{h}{h_{\mathrm{wg}}}\right) + r \log \frac{h}{h_{\mathrm{wg}}} d\beta, \tag{10}$$

which relates errors (differentials). Obtaining the square of Eq. (10) and averaging over an appropriate ensemble, the terms $u_{\mathrm{rep}}^2 + u_{\mathrm{par}}^2$ are obtained. The only non-null covariance is that between the two parameters of the response function. The variables $\sigma_{r_{\mathrm{wg}}}$, σ_β, and $\mathrm{covar}\left(r_{\mathrm{wg}}, \beta\right)$ are computed using the residuals of these parameters respect to the considered linear drift in time or constancy in time. A single value for each variable per working gas in use is obtained. The typical value of $\sigma_{r_{\mathrm{wg}}}$ is 1.09 nmol mol^{-1} before March 2009, and 0.40 nmol mol^{-1} after March 2009; whereas the typical value of σ_β is 0.030 before March 2009, and 0.0044 after March 2009. The correlation coefficient between r_{wg} and β reaches significant values as high as 0.73, and as low as −0.91, depending its sign on the mole fraction of the working gas. So, the associated covariance has to be considered in the uncertainty computation.

The repeatability (standard deviation) of the relative height, $\sigma_{h/h_{wg}}$, is determined from the repeated injections for each standard made during instrument calibrations. Moreover, it is necessary to know the dependence of $\sigma_{h/h_{wg}}$ on relative height, h/h_{wg},

$$\sigma_{h/h_{wg}} = k\sqrt{1+\left(\frac{h}{h_{wg}}\right)^2}, \tag{11}$$

where k is a parameter equal to $(\sigma_h)/h_{wg}$, which depends on the mole fraction of the working gas and possibly on time. For the computation of the uncertainty component given by Eq. (5), Eq. (11) is used to provide $\sigma_{h/h_{wg}}$ using a single (mean) value of k for each working gas used. Equation (11) has been obtained taking into account that the statistical properties of the height error do not depend on mole fraction (the error in the placement of the peak baseline does not depend on peak height, but on baseline noise).

Figure 5 shows the uncertainty components for the period 2008–2011. Table 3 summarizes the mean values of each uncertainty component before and after March 2009. The mean combined standard uncertainty decreased significantly after March 2009, from 2.37 nmol mol^{-1} to 1.66 nmol mol^{-1}. After March 2009 the components u_{pr}, $u_{p\beta}$, and u_{par} are significantly smaller than before reflecting an improvement in the determination and consistency of the response function parameters. Those values are particularly high during the first half of 2008. After March 2009, the single largest uncertainty component was u_{fit}, whereas before March 2009 it was u_{par}.

4.1 The representation uncertainty and the propagated uncertainty of the mean for quasi-continuous and flask measurements

There is a fifth type of uncertainty we call representation uncertainty, u_{rs}. This is present when computing means from a number of available data (n) that is smaller than the theoretical number of data necessary to compute exactly the mean (N). The computed

mean will be different from the actual mean (unknown) and the representation uncertainty quantifies this difference. In time series analysis a hierarchy of data assemblages are possible (e.g. hourly mean, daily mean, monthly mean, annual mean), each being computed from the means of the previous level. An additional representation uncertainty is associated with each assemblage. For example, an additional representation uncertainty will appear when computing a daily mean from only 22 available hourly means (in this case, $N = 24$, and $n = 22$). The value N is known without doubts for each level except for the first one. That is, for computing the hourly means, the theoretical maximum number of independent measurements within an hour necessary to compute exactly the mean is not easy to assess. The additional representation uncertainty is given by the equation

$$u_{rs}^2 = \frac{\sigma_{sam}^2}{n}\left(\frac{N-n}{N-1}\right), \tag{12}$$

where

$$\sigma_{sam} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(r_i - \langle r\rangle\right)^2} \tag{13}$$

is the standard deviation of the sample of data, $\langle r\rangle$ is the mean, and r_i is the data number i used to compute the mean. Indeed, the standard deviation of the sample of data includes the dispersion due to measurement repeatability. So, before using Eq. (12), the uncertainty due to the repeatability should be subtracted quadratically from σ_{sam}, and if the result is negative convert it to zero. Equation (12) is a general statistical result that holds for the variance of the mean of a sample without replacement of size n from a finite population of size N (e.g. Martin, 1971, chapter 5). Note that when $N \gg n$, the term between parenthesis in Eq. (12) becomes equal to 1, so in such case, it does not matter the exact value of N.

Any computed mean has also a propagated uncertainty arising from the uncertainties of the data used to compute ~~the~~ mean. Therefore, a mean will have an additional representation uncertainty and a propagated uncertainty (both to be summed quadratically) that includes, among others, the propagated representation uncertainty arising from the previous levels of means. The uncertainty components are of two types: systematic, u_{syst}, and random, u_{rand}, and are combined quadratically. The law of propagation depends on the type of uncertainty. Therefore, we can write

$$u_{\langle r\rangle} = \sqrt{u_{\mathrm{rs}}^2 + u_{\langle r\rangle,\mathrm{rand}}^2 + u_{\langle r\rangle,\mathrm{syst}}^2}, \tag{14}$$

where $u_{\langle r\rangle}$ indicates the uncertainty of the mean, $u_{\langle r\rangle,\mathrm{rand}}$ indicates the random component of the propagated uncertainty, and $u_{\langle r\rangle,\mathrm{syst}}$ indicates the systematic component of the propagated uncertainty. For the propagation of random uncertainty, the equation

$$u_{\langle r\rangle,\mathrm{rand}}^2 = \frac{1}{n^2}\sum_{i=1}^{n} u_{\mathrm{rand}_i}^2 \tag{15}$$

holds; whereas for the propagation of systematic uncertainty, the equation

$$u_{\langle r\rangle,\mathrm{syst}}^2 = \frac{1}{n}\sum_{i=1}^{n} u_{\mathrm{syst}_i}^2 \tag{16}$$

holds, where the subindex i indicates the uncertainty of the data number i used to compute the mean. Note that in Eq. (15) there is partial cancellation of random errors, whereas in Eq. (16) there is ~~not any~~ cancellation ~~of systematic errors~~ because the systematic error is the same (or nearly the same) for all the data used in the computation ~~of the mean~~. The random uncertainty can be expressed as

$$u_{\mathrm{rand}_i} = \sqrt{u_{\mathrm{rep}_i}^2 + u_{\mathrm{rs}_i}^2}, \tag{17}$$

~~whereas~~ for the systematic uncertainty

$$u_{\mathrm{syst}_i} = \sqrt{u_{\mathrm{st}_i}^2 + u_{\mathrm{fit}_i}^2 + u_{\mathrm{par}_i}^2}. \tag{18}$$

Note that u_{par} behaves as systematic for computing hourly, daily, and monthly means, but behaves as random for computing annual means. The component u_{fit} has systematic and random contributions, but we consider it as systematic for the uncertainty propagation (~~so, overestimating a bit~~ the propagated uncertainty).

Table 4 shows mean values of the uncertainty components for hourly, daily night, monthly, and annual means during the period 25 March 2009–31 December 2011. The hourly means correspond to the night period (20:00–08:00 UTC). The mean representation uncertainty in the hourly means is 0.63 nmol mol^{-1} for the night period, and 0.83 nmol mol^{-1} for the daytime period (08:00–20:00 UTC). The larger value during daytime is due to the CO diurnal cycle (Sect. 6), which makes σ_{sam} larger during daytime. Since the time coverage of the continuous in-situ measurements is very high, no additional representation uncertainty components appear when computing the successive means (daily night, monthly, and annual) but only the propagated representation uncertainty. Therefore, the uncertainties associated to random errors (repeatability and representation) are smaller for longer periods of averaging, while the uncertainties associated to systematic errors (u_{st} and u_{fit}) are the same for all the periods of averaging. The uncertainty u_{par} has a mixed behaviour due to the fact that its character (random or systematic) depends on the period of averaging.

~~Since 1991, ambient air samples were collected weekly at Izaña Observatory~~ for analysis at NOAA-ESRL-GMD Carbon Cycle Greenhouse Group (CCGG) as part of Cooperative Air Sampling Network (Komhyr et al., 1985; Conway et al., 1988; Thoning et al., 1995). In every sampling event, two flasks are collected nearly simultaneously ~~and the difference between the two is used to identify artifacts in sampling and analysis. A difference greater than 3 nmol mol^{-1} is considered indicative of a problem and both flasks are flagged accordingly~~. Monthly and annual means computed with such sparse flask data (4 independent values per month) are subject to a large representation

uncertainty. Table 5 shows mean values of the representation uncertainty in the different ~~types of~~ means ~~for the NOAA flasks~~. For the hourly and daily night means based on a single pair of flasks, the associated σ_{sam} were computed using the quasi-continuous in-situ measurements. The bias ~~associated to sampling out of the period of free troposphere background conditions will be considered in Sect. 6 where the CO diurnal cycle is considered~~. The standard deviation σ_{sam} of ~~flasks within a month has been computed for every month of 2009 and 2010 and then averaged to yield~~ 9.80 nmol mol^{-1}. Following the discussion above, it is not surprising that the representation uncertainties in the in situ means (Table 4) are much smaller than those from flask sampling.

5 Flasks-continuous comparison, comparison uncertainty, and means

~~The comparison between continuous (or quasi-continuous) measurements obtained using in-situ instruments and collocated flask samples is~~ an independent way of assessing the quality of the continuous in-situ measurements (WMO, 2011). A significant difference ~~between~~ a flask sample and a simultaneous in-situ hourly mean ~~can~~ be obtained due to two causes: (1) the measurements have different, potentially large, errors (note that the concept of error includes the bias in the measurements of any of the laboratories), and/or (2) the air sampled by the two methods is different (i.e. both measurements have different "true values"). The second potential cause for differences between measurements will be quantified through what we call the comparison uncertainty. The statistical significance of each difference (i.e. if there are significantly different errors in both measurements) will be evaluated comparing it with its comparison uncertainty. Note that the error (unknown) is the difference between the true value and the value provided by the measurement system (JCGM, 2008). To compare in-situ hourly means with simultaneous NOAA flask samples we proceed as follows.

First. Flasks results are used only if are defined by NOAA as representative of background conditions, their sampling and analysis ~~are all right~~ and results from both members of the pair are available. Each pair of mole fractions, r_{f1} and r_{f2}, is substituted by

its mean, $\langle r_f \rangle$, and its standard deviation,

$$\sigma_f = \frac{|r_{f2} - r_{f1}|}{\sqrt{2}}. \quad (19)$$

~~Such~~ standard deviation is indicative of the internal consistency of the pair.

Second. ~~Each pair is compared with the hourly mean simultaneous in time (the time interval of the hourly mean must contain the time of the pair sampling).~~ We denote the hourly mean as r_c, and the standard deviation of the sample of data within the hour as σ_c, which quantifies the departures of the instantaneous measurements from the mean. So, we compute the difference,

$$\text{dif} = \langle r_f \rangle - r_c, \quad (20)$$

and its comparison uncertainty,

$$\sigma_{dif} = \sqrt{\sigma_f^2 + \sigma_c^2}. \quad (21)$$

~~Note that, in the comparison, we are using the in-situ hourly mean instead of the in-situ instantaneous data (not measured) simultaneous to the flask sampling. This is why~~ σ_c must be used in Eq. (21) instead of the standard deviation of the hourly mean. The comparison uncertainty asseses if the difference is significant. If $|\text{dif}| \leq 2\sigma_{dif}$, ~~this means that~~ the difference is not significant, whereas if $|\text{dif}| > 2\sigma_{dif}$, ~~this means that~~ the difference is significant.

Figure 6 shows the time series of differences between NOAA flask samples and simultaneous in-situ hourly means. Error bars indicate comparison uncertainty. Dots in red do not have associated error bar due to the absence of σ_c (corresponding to hours with only one valid in-situ ~~injection~~). For 2008, 47.4 % of the differences are significant, whereas for 2009–2011, only 24.5 % of the differences are significant. Computing percentiles for the CO differences, we conclude that for 2009–2011, 68 % of the differences

are between −2.39 and 2.5 nmol mol^{-1} (a large fraction of this dispersion is caused by the comparison uncertainty, since the 68 percentile of σ_{dif} is equal to 2.28 nmol mol^{-1}), whereas for 2008, 68 % of the differences are between −1.26 and 6.58 nmol mol^{-1}.

5.1 Annual and multi-annual means

Here, total (full period of data) and annual mean differences are computed using conventional expressions and expressions weighted by the comparison uncertainty. Note that the mean difference is equal to the difference of systematic errors of both laboratories (averaged in the mole fractions of the samples) plus a possible systematic difference between the flask and the continuous data that would appear if there was a systematic trend within the hourly means and the minute of flask sampling was not random.

The conventional mean is denoted as Mean,

$$\langle \text{dif} \rangle = \frac{1}{n} \sum_{i=1}^{n} \text{dif}_i, \tag{22}$$

where n is the number of differences used to compute the mean. FWMean is a "full" weighted mean computed following the minimum variance method (equivalent to the maximum likelihood method for Gaussian distributions), e.g. Martin (1971, chapter 9),

$$\langle \text{dif} \rangle_{\text{FW}} = \frac{1}{n} \sum_{i=1}^{n} \frac{\sigma^2_{\text{inv}}}{\sigma^2_{\text{dif}\,i}} \text{dif}_i, \tag{23}$$

where

$$\frac{1}{\sigma^2_{\text{inv}}} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma^2_{\text{dif}\,i}}. \tag{24}$$

The basic idea: a difference with a larger uncertainty provides information of a lower quality to compute the mean, and therefore, the applied weight is smaller. WMean is an "intermediate" weighted mean for which Eq. (23) applies but $\sigma^2_{\text{dif}\,i}$ is replaced by the median of σ^2_{dif} for those $\sigma^2_{\text{dif}\,i}$ smaller than the median of σ^2_{dif}. The basic idea is to avoid an excessive weight of those dif with a very small $\sigma^2_{\text{dif}\,i}$. We believe that WMean is the most appropriate estimator. Differences without associated uncertainty and those with an absolute value larger than 10 nmol mol^{-1} (3 differences in 2008, as Fig. 6 shows) have not been used to compute the weighted means.

Table 6 provides the mean differences between flask and in situ measurements. Smaller differences are found in 2009–2011 than in 2008. The annual mean differences for 2009–2011 are well within the ±2 nmol mol^{-1} compatibility goal of GAW (WMO, 2011). Note that the various mean differences are not very different. The conventional annual mean differences are the closest to zero, except for 2008. However, for CO_2 and CH_4, we have observed many annual weighted mean differences closer to zero than the conventional mean difference.

Table 6 also shows the standard deviation of the mean, σ_{mean}. For the conventional mean, it can be computed in two different ways. On the one hand, it is equal to the standard deviation of the sample (SD) divided by the square root of the number of data, n, used to compute the mean (e.g. Martin, 1971, chapter 5). On the other hand, the relation

$$\sigma^2_{\text{mean}} = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2_{\text{dif}_i} \tag{25}$$

holds. For the weighted means, the relation $\sigma_{\text{mean}} = \sigma_{\text{inv}}/\sqrt{n}$ holds (e.g. Martin, 1971, chapter 9), where σ_{inv} is given by Eq. (24). Note that the σ_{mean} associated to FWMean is smaller than those associated to the other means since FWMean is obtained using the minimum variance method, and σ_{inv} is smaller for smaller values of σ_{dif}. For the conventional mean, Table 6 shows that the values of SD$/\sqrt{n}$ are larger than the values

of σ_{mean}, except for 2010 due to the presence of a difference with a very large value of σ_{dif} (11.3 nmol mol^{-1}) during this year. This means that the dispersion of the differences within one year is larger than it would be expected according to the values of σ_{dif_i}. Therefore, for computing weighted means, σ_{dif} does not include all the causes of variability within one year. In detail, σ_{dif_i} does not fully include the errors in the measurements that behave as random along one year. This means that, for computing weighted means, the smallest σ_{dif} are smaller than they should be and FWMean is not a very good estimator of the mean for this dataset.

Finally, we consider if the annual average flask versus in situ differences are significant. The mean difference, which is distributed normally according to the Central Limit theorem (e.g. see Martin, 1971, chapter 5), is significant for a 95 % confidence level if $|\langle \text{dif} \rangle| > 1.96\sigma_{mean}$, where $\langle$ dif $\rangle$ denotes annual mean difference. As Table 6 shows, the conventional mean and the "intermediate" weighted mean differences are not significant for 2009, 2010, and 2011, whereas they are significant for 2008. Mean is not significant over the full period 2008–2011, whereas WMean is significant. The "full" weighted mean difference is significant for all years except 2010, but we have stated previously that it is not a good estimator for this particular data set.

6 Time series analysis

To analyse the CO time series of daily night means, we carry out a least-squares fitting to a quadratic interannual component plus a constant annual cycle composed by 4 Fourier harmonics,

$$f(t) = a_0 + a_1 t + a_2 t^2 + \sum_{i=1}^{4} [b_i \cos(\omega_i t) + c_i \sin(\omega_i t)], \quad (26)$$

where t is time in days, being $t = 1$ for 1 January 2008, a_0, a_1, and a_2 are the parameters of the interannual component to be determined, b_i and c_i are the parameters

of the annual cycle to be determined, and $\omega_i = 2\pi i/T$ with $T = 365.25$ days. This fitting is the same as the one used by Novelli et al. (1998, 2003) and developed by Thoning et al. (1989).

Figure 4 shows the time series of daily night measured carbon monoxide, the fitted interannual component, and the fitted interannual component plus the fitted annual cycle. The RMS residual of the fitting is equal to 11.5 nmol mol^{-1}. The interannual component increases till the beginning of 2010, and then decreases. Table 7 shows nocturnal CO annual means. These were computed using measured data when available and values from the curve fitted data when measured data were not available. As Table 7 shows, the number of days with data not available are very small. From 2008 to 2010 the CO annual mean increased 4.0 nmol mol^{-1} (standard uncertainty: 2.3 nmol mol^{-1}), while a decrease of 2.5 nmol mol^{-1} (standard uncertainty: 2.2 nmol mol^{-1}) is found between 2010 and 2011.

Figure 7 shows the fitted annual cycle, which has an amplitude from the minimum to the maximum of 40.7 nmol mol^{-1}. The maximum occurs in late March, while the minimum is in middle August. This is the seasonal cycle common to the Northern Hemisphere which is primarily driven by reaction with OH and anthropogenic sources (e.g. Novelli et al., 1998). The annual cycle obtained here is similar to that obtained by Schmitt and Volz-Thomas (1997) using measurements carried out at Izaña from May 1993 to December 1995.

Now we consider the time series of the residuals from the curve fitting. A large change in the residuals indicates a change in air mass. The persistence of the residuals can be measured computing the autocorrelation (Fig. 8). For a time-lag of 1, 2, 3, and 7 days, the autocorrelation is 0.56, 0.30, 0.21, and 0.10, respectively. We conclude the residuals are not autocorrelated after a time-lag of 7 days. So, 7 days could be considered the typical period of persistence of an air mass for CO.

Figure 9 shows the carbon monoxide monthly mean diurnal cycle for every month relative to the nocturnal background, computed using hourly data for the period 2008–2011. In detail, this figure shows mean values of the differences: hourly CO minus the

nocturnal background level. ~~Such~~ reference background level was computed as follows. Firstly, the averages of the pre (00:00–07:00 UTC) and post (21:00–04:00 UTC) nights were computed. Then, the linear drift in time passing through both averages is used as the reference background level. Note for example that hour 1 means the hourly mean for the period 00:00–01:00 UTC. Carbon monoxide at Izaña is ~~tipically~~ stable during the night period 20:00–08:00 UTC, starts to increase around 09:00 UTC, reaches its maximum around 13:00–15:00 UTC, before returning to the nocturnal background (Fig. 9). The amplitude is ~~around~~ 5 ~~or~~ 6 nmol mol^{-1}, except in December, when the amplitude is around 4 nmol mol^{-1}. The mean time of flask sampling during 2008–2011 was 10:00 UTC. That is, sampling occurred during non background conditions. There is a mean bias of approximately +1.5 nmol mol^{-1} between the air sampled with flasks and the nocturnal background conditions. During 2002–2007 the mean time of flask sampling was 15:35 UTC. Given all other effects are similar to the more recent period, CO determined from flask air samples are approximately +4.5 nmol mol^{-1} higher than nocturnal background conditions. NOAA began air sampling at Izaña late 1991, until 2002 flasks were sampled during nighttime, but this was discontinued due to the absence of staff during nighttime.

7 Summary and conclusions

A rigorous procedure to determine the uncertainties in the semi-continuous measurements of CO made at the Izaña global GAW station has been developed. This approach is applicable to other sites in the WMO GAW global network. The error in the measurements are reported as the combined standard uncertainty. This has four components: the uncertainty of the WMO standard gases interpolated over the range of measurement, the uncertainty that takes into account the agreement between the standard gases and the response function used, the uncertainty due to the repeatability of the injections, and the propagated uncertainty related to the response function parameters ~~uncertainties~~ (which also takes into account the covariance between the parameters).

The mean value of the combined standard uncertainty decreased significantly after March 2009, from 2.37 nmol mol^{-1} to 1.66 nmol mol^{-1}. The reason of this improvement is a very significant reduction in the response function parameter ~~uncertainties~~ (the dominant source of uncertainty before March 2009). There was an improvement in the determination and consistency of the response function parameters due to the following facts that apply after March 2009. The improvement reflects the use of a newer and larger set of WMO standard gases, more injections of the working gas in the calibration sequence, and use of an adjacent closed port in the multiposition selection valve to stop sample loop flushing for pressure equilibration. The dominant uncertainty component after March 2009 is the uncertainty that takes into account the agreement between the standard gases and the response function (1.27 nmol mol^{-1}). A fifth type of uncertainty we call representation uncertainty is considered when some of the data necessary to compute ~~exactly~~ the mean are absent. Any computed mean has also a propagated uncertainty arising from the uncertainties of the data used to compute the mean. The law of propagation depends on the type of uncertainty component (random or systematic), i.e. there is partial cancellation of random errors, ~~however~~ there is not cancellation of systematic errors. The representation uncertainties in the in-situ measurements are much smaller than those computed using flasks (e.g. the representation uncertainty in the monthly means is equal to 4.98 nmol mol^{-1} for the NOAA flasks, and 0.03 nmol mol^{-1} for the quasi-continuous measurements). The larger uncertainty in the flask air measurements reflects the relatively sparse sampling.

The 2008–2011 Izaña carbon monoxide nocturnal time series ~~is presented. The time series~~ is evaluated using a least squares fit to a quadratic interannual component plus a constant annual cycle composed by 4 Fourier harmonics. The interannual component increases till the beginning of 2010, and then decreases. ~~The fitted annual cycle has an~~ amplitude ~~of~~ 40.7 nmol mol^{-1}. ~~Its maximum occurs in late March, while its minimum occurs in middle August.~~ The autocorrelation of the residuals indicates that the typical period of persistence of an air mass for CO is 7 days. The monthly mean diurnal cycle ~~relative to the nocturnal background shows that during night the CO mole fraction is~~

~~stable representing the free troposphere. Upslope winds during daytime increase CO by 5 or 6~~ nmol mol^{-1}. The air samples collected during the daytime are biased high with respect to mid-tropospheric background conditions. The magnitude of the bias depends on the sampling hour.

We also examine differences between hourly means determined by the in-situ continuous measurements with flask air samples determined within the hour. The uncertainties in the mean results from each method allows determination ~~if the difference is significant~~. During 2008, 47.4 % of the differences were significant, with 68 % of these between −1.26 and 6.58 nmol mol^{-1}. During 2009–2011, only 24.5 % of the differences were significant and 68 % were between −2.39 and 2.5 nmol mol^{-1} (this range is largely due to the comparison uncertainty). Total and annual mean differences between the grab samples and in situ measurements were computed using conventional expressions but also expressions with weights based on the minimum variance method. During the period 2009–2011 the flask in situ differences are much closer to zero than during 2008, which likely results from the better performance of the Izaña measurement system during 2009–2011. The annual mean differences between NOAA and in-situ (AEMET) measurements for 2009–2011 are not significant and within the 2 nmol mol^{-1} inter-laboratory compatibility goal of GAW.

Acknowledgements. This study was developed within the Global Atmospheric Watch (GAW) Programme at the Izaña Atmospheric Research Center, financed by AEMET. We are grateful to Christoph Zellweger~~,~~ for ~~some~~ advice during the configuration of the RGA, and to the Izaña observatory staff. The collaborative air sampling network operated by the NOAA-ESRL-GMD carbon cycle group is acknowledged for its many years of measurements at the observatory.

References

Armerding, W., Comes, F. J., Crawack, H. J., Forberich, O., Gold, G., Ruger, C., Spickermann, M., Walter, J., Cuevas, E., Redondas, A., Schmitt, R., and Matuska, P.: Testing the

daytime oxidizing capacity of the troposphere: 1994 OH field campaign at the Izaña observatory, Tenerife, J. Geophys. Res., 102, 10603–10612, doi:10.1029/96JD03714, 1997. 6951

Conway, T. J., Waterman, L. S., Tans, P., Thoning, K. W., and Masarie, K. A.: Atmospheric carbon dioxide measurements in the remote global troposphere, 1981–1984, Tellus B, 40, 81–115, doi:10.1111/j.1600-0889.1988.tb00214.x, 1988. 6962

Ehhalt, D., Prather, M., Dentener, F., Derwent, R., Dlugokencky, E., Holland, E., Isaksen, I., Katima, J., Kirchhoff, V., Matson, P., Midgley, P., and Wang, M.: Atmospheric chemistry and greenhouse gases, chap. 4, in: Climate Change 2001: The Scientific Basis, edited by: Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., and Johnson, C. A., Cambridge University Press, Cambridge, UK, 239–287, 2001. 6951

Fischer, H., Nikitas, C., Parchatka, U., Zenker, T., Harris, G. W., Matuska, P., Schmitt, R., Mihelcic, D., Muesgen, P., Paetz, H.-W., Schultz, M., and Volz-Thomas, A.: Trace gas measurements during the oxidizing Capacity of the Tropospheric Atmosphere campaign 1993 at Izaña, J. Geophys. Res., 103, 13505–13518, doi:10.1029/97JD01497, 1998. 6951

JCGM: JCGM 100:2008, GUM 1995 with Minor Corrections, Evaluation of Measurement Data – guide to the Expression of Uncertainty in Measurement, Joint Committee for Guides in Metrology, Member organizations: BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, 2008. 6952, 6954, 6958, 6963

Kitzis, D.: Preparation and stability of standard reference air mixtures, Tech. rep., NOAA Earth System Research Laboratory, Boulder, USA, available at: http://www.esrl.noaa.gov/gmd/ccl/airstandard.html (last access: 14 September 2012), 2009. 6954

Komhyr, W. D., Gammon, R. H., Harris, T. B., Waterman, L. S., Conway, T. J., Taylor, W. R., and Thoning, K. W.: Global atmospheric CO_2 distribution and variations from 1968–1982 NOAA/GMCC CO_2 flask sample data, J. Geophys. Res., 90, 5567–5596, doi:10.1029/JD090iD03p05567, 1985. 6962

Lang, P. M.: Guidelines for standard gas cylinder and pressure regulator use, Tech. rep., NOAA Earth System Research Laboratory, Boulder, USA, available at: http://www.esrl.noaa.gov/gmd/ccl/reg.guide.html (last access: 14 September 2012), 1998. 6954

Logan, J. A., Prather, M. J., Wofsy, S. C., and McElroy, M. B.: Tropospheric chemistry – a global perspective, J. Geophys. Res., 86, 7210–7254, doi:10.1029/JC086iC08p07210, 1981. 6951

Martin, B. R.: Statistics for Physicists, Academic Press, London, UK, 1971. 6955, 6960, 6965, 6966, 6967

Navascues, B. and Rus, C.: Carbon dioxide observations at Izana baseline station, Tenerife (Canary Islands) – 1984–1988, Tellus B, 43, 118–125, doi:10.1034/j.1600-0889.1991.t01-1-00006.x, 1991. 6951

Novelli, P. C., Steele, P., and Tans, P. P.: Mixing ratios of carbon monoxide in the troposphere, J. Geophys. Res., 97, 20731–20750, doi:10.1029/92JD02010, 1992. 6951

Novelli, P. C., Masarie, K. A., and Lang, P. M.: Distributions and recent changes of carbon monoxide in the lower troposphere, J. Geophys. Res., 103, 19015–19034, doi:10.1029/98JD01366, 1998. 6968

Novelli, P. C., Masarie, K. A., Lang, P. M., Hall, B. D., Myers, R. C., and Elkins, J. W.: Reanalysis of tropospheric CO trends: effects of the 1997–1998 wildfires, J. Geophys. Res.-Atmos., 108, 4464, doi:10.1029/2002JD003031, 2003. 6955, 6968

Rodríguez, S., González, Y., Cuevas, E., Ramos, R., Romero, P. M., Abreu-Afonso, J., and Redondas, A.: Atmospheric nanoparticle observations in the low free troposphere during upward orographic flows at Izaña Mountain Observatory, Atmos. Chem. Phys., 9, 6319–6335, doi:10.5194/acp-9-6319-2009, 2009. 6951

Schmitt, R. and Volz-Thomas, A.: Climatology of ozone, PAN, CO, and NMHC in the free troposphere over the Southern North Atlantic, J. Atmos. Chem., 28, 245–262, 1997. 6968

Schmitt, R., Schreiber, B., and Levin, I.: Effects of long-range transport on atmospheric trace constituents at the baseline station Tenerife (Canary Islands), J. Atmos. Chem., 7, 335–351, 1988. 6951

Thoning, K. W., Tans, P. P., and Komhyr, W. D.: Atmospheric carbon dioxide at Mauna Loa Observatory, II – Analysis of the NOAA GMCC data, 1974–1985, J. Geophys. Res., 94, 8549–8565, doi:10.1029/JD094iD06p08549, 1989. 6968

Thoning, K. W., Conway, T. J., Zhang, N., and Kitzis, D.: Analysis system for measurement of CO_2 mixing ratios in flask air samples, J. Atmos. Ocean. Tech., 12, 1349, doi:10.1175/1520-0426(1995)012<1349:ASFMOC>2.0.CO;2, 1995. 6962

WMO: Guidelines for the Measurement of Atmospheric Carbon Monoxide, GAW Report No. 192, World Meteorological Organization, Geneva, Switzerland, 2010. 6951, 6952

WMO: 15th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques, Jena, Germany, 7–10 September 2009, GAW Report No. 194, chap. Expert group recommendations, World Meteorological Organization, Geneva, Switzerland, 1–25, 2011. 6952, 6954, 6963, 6966

Table 1. WMO CO standard gases of the Izaña station: CO mole fraction and uncertainty ~~(1-sigma) as calibrated in 2006~~ by the WMO CO CCL in ~~the WMO-2004 CO scale~~.

Cylinder	CO (nmol mol^{-1})	1-sigma (nmol mol^{-1})
CA06768	62.6	1.2
CA06946	91.2	0.7
CA06988	119.6	0.8
CA06968	164.5	1.1
CA06978	221.2	1.5

Table 2. Residuals respect to the least-squares fitting for the typical calibration plotted in Fig. 2. CO mole fraction of the standard gases, 1-sigma uncertainty and residuals are given in nmol mol^{-1}.

CO	1-sigma	Residual
62.6	1.2	0.53
91.2	0.7	−0.97
119.6	0.8	0.00
164.5	1.1	−0.95
221.2	1.5	1.71

Table 3. Mean values of the uncertainty components in nmol mol^{-1} before and after March 2009.

Period	u_{st}	u_{fit}	u_{rep}	u_{pr}	$u_{p\beta}$	u_{par}	u_{tot}
1 Jan 2008–24 Mar 2009	0.89	1.28	0.33	1.27	1.13	1.64	2.37
25 Mar 2009–31 Dec 2011	0.90	1.27	0.36	0.36	0.13	0.39	1.66

Table 4. Mean values of the uncertainty components (in nmol mol^{-1}) for ~~the different types of means during the period~~ 25 March 2009–31 December 2011. The hourly means ~~considered~~ correspond to the night period (20:00–08:00 UTC).

Type of mean	u_{st}	u_{fit}	u_{par}	u_{rep}	u_{rs}
Hourly	0.90	1.27	0.39	0.36	0.63
Daily night	0.90	1.27	0.39	0.10	0.18
Monthly	0.90	1.27	0.39	0.02	0.03
Annual	0.90	1.27	0.11	0.01	0.01

Table 5. Mean values of the representation uncertainty (nmol mol^{-1}) ~~in the different types of means for~~ the NOAA flasks.

Type of mean	Additional u_{rs}	Propagated u_{rs}	Total u_{rs}	n	N
Hourly	1.09	0.00	1.09	1	$\gg 1$
Daily night	3.44	1.09	3.61	1	12
Monthly	4.64	1.81	4.98	4	30
Annual	0.00	1.44	1.44	12	12

Table 6. Mean differences ~~(flasks~~ minus in-situ) and standard deviations in nmol mol^{-1}. n dif denotes the number of differences available.

Period	n dif	Mean	σ_{mean}	SD/$\sqrt{n}$	WMean	σ_{mean}	FWMean	σ_{mean}
2008–2011	147	0.79	0.20	0.42	0.61	0.16	0.59	0.07
2008	39	3.23	0.35	1.36	2.48	0.31	2.16	0.15
2009	35	−0.43	0.36	0.47	−0.47	0.31	0.81	0.12
2010	38	0.02	0.48	0.44	0.25	0.31	−0.11	0.18
2011	35	0.12	0.42	0.49	0.15	0.33	−0.68	0.14

Table 7. Annual mean ~~carbon monoxide mole fraction for the daily night period~~ 20:00–08:00 UTC~~, standard uncertainty~~ in nmol mol^{-1} and number of days with data ~~availability~~.

Year	CO (nmol mol^{-1})	Standard uncertainty	Available days
2008	93.63	1.63	355
2009	94.73	1.56	355
2010	97.64	1.56	351
2011	95.16	1.56	356

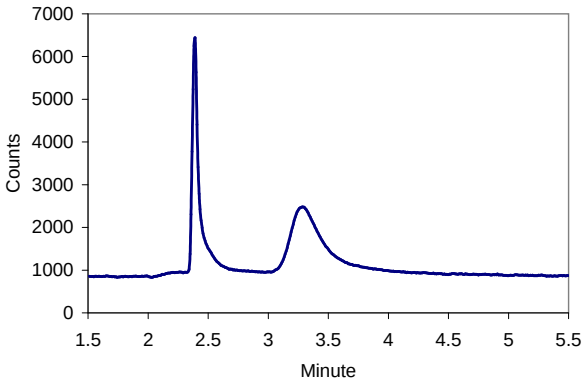

Fig. 1. Typical RGA chromatogram. The first eluted peak corresponds to H_2, whereas the second one corresponds to CO.

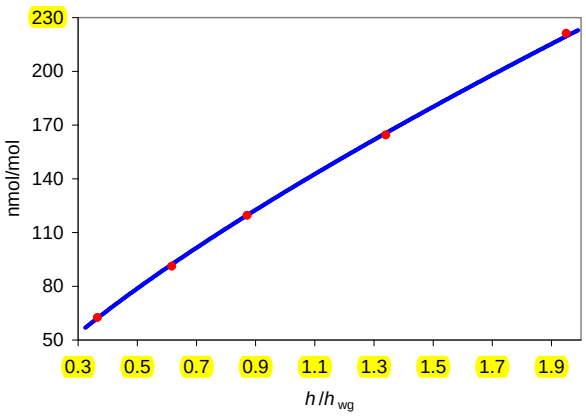

Fig. 2. Least-squares fitting of a typical calibration. The fitting is plotted in blue, ~~whereas~~ the measured means are plotted in red.

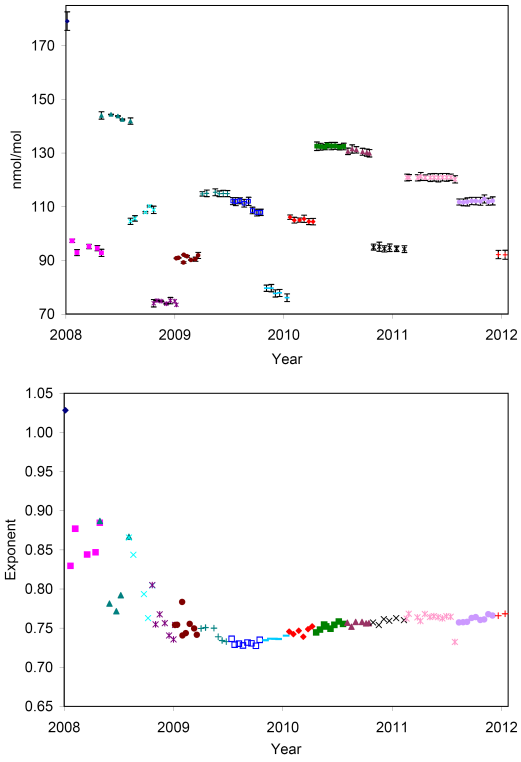

Fig. 3. Upper graph: working gas mole fractions obtained from calibrations conducted during 2008 to 2011. Error bars represent the RMS residual of each calibration, i.e. $\pm u_{\text{fit}}$. Lower graph: response function exponents obtained in the calibrations. Different colours and symbols are used for the different working gases ~~in use~~.

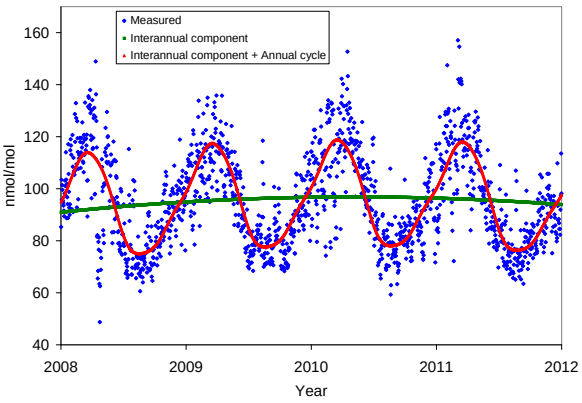

Fig. 4. Daily night ~~means~~ (20:00–08:00 UTC) ~~for the~~ carbon monoxide mole fraction measured at Izaña Observatory (blue squares). ~~Fitted~~ interannual component ~~(green curve) and fitted~~ annual cycle ~~superposed (red curve) that are explained in Sect. 6~~.

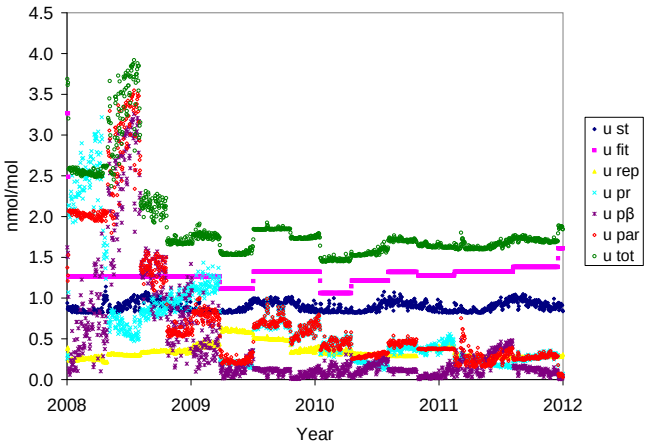

Fig. 5. Uncertainty components (daily means) of the measured CO mole fraction.

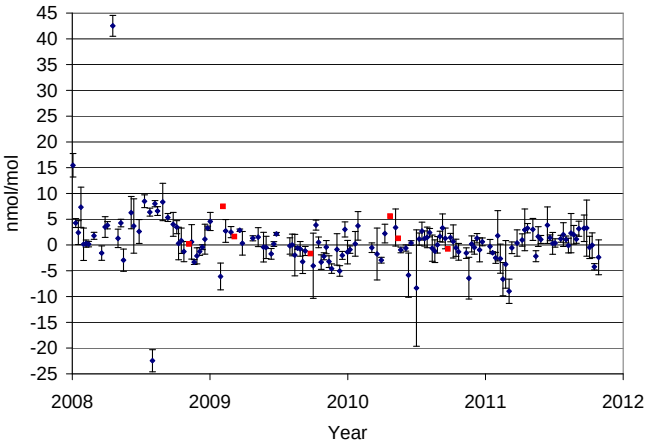

Fig. 6. Differences between NOAA flask samples and simultaneous in-situ hourly means. Error bars indicate comparison uncertainty. Differences plotted in red do not have associated uncertainty due to the presence of only one ambient air injection within the associated hour.

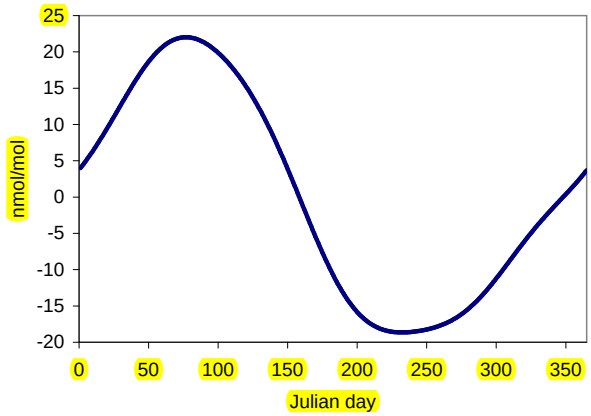

Fig. 7. Carbon monoxide fitted annual cycle.

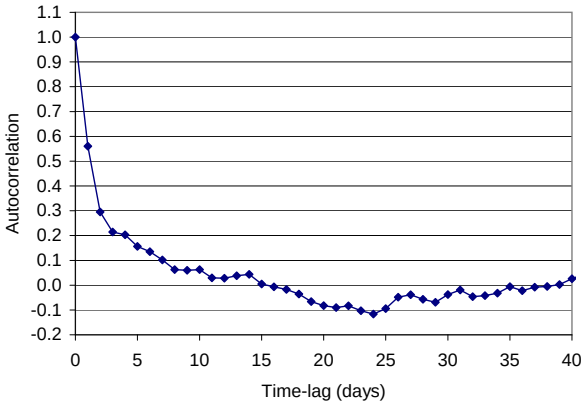

Fig. 8. Autocorrelation of the residuals from the fitting ~~given by~~ Eq. (26).

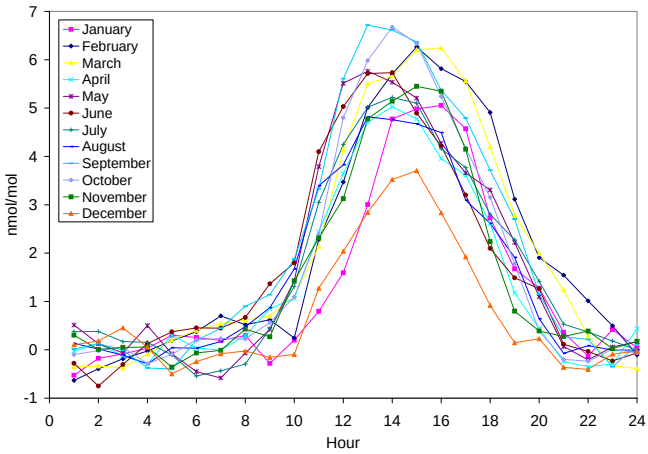

Fig. 9. Carbon monoxide mean diurnal cycle relative to the nocturnal background level.