



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

A. J. Gomez-Pelaez¹, R. Ramos¹, V. Gomez-Trueba^{1,2}, P. C. Novelli³, and R. Campo-Hernandez¹

¹Izaña Atmospheric Research Center (IARC), Meteorological State Agency of Spain (AEMET), Izaña, 38311, Spain

²Air Liquide España, Delegación Canarias, Candelaria, 38509, Spain

³National Oceanic and Atmospheric Administration, Earth System Research Laboratory, Global Monitoring Division (NOAA-ESRL-GMD), Boulder, CO 80305, USA

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

Received: 17 August 2012 – Published in Atmos. Meas. Tech. Discuss.: 21 September 2012

Revised: 20 February 2013 – Accepted: 21 February 2013 – Published: 20 March 2013

Abstract. Atmospheric CO in situ measurements are carried out at the Izaña (Tenerife) global GAW (Global Atmosphere Watch Programme of the World Meteorological Organization – WMO) mountain station using a Reduction Gas Analyser (RGA). In situ measurements at Izaña are representative of the subtropical Northeast Atlantic free troposphere, especially during nighttime. We present the measurement system configuration, the response function, the calibration scheme, the data processing, the Izaña 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 measurement uncertainty taking into consideration four contributing components: 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 temporal consistency of the response function parameters (which also takes into account the covariance between the parameters). The mean value of the combined standard uncertainty decreased significantly after March 2009, from $2.37 \text{ nmol mol}^{-1}$ to $1.66 \text{ nmol mol}^{-1}$, 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 the temporal 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 of the differences is computed and used to determine whether the differences are significant. For 2009–2011, only 24.5 % of the differences are significant, and 68 % of the differences are between -2.39 and $2.5 \text{ nmol mol}^{-1}$. 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 $\pm 2 \text{ nmol mol}^{-1}$ compatibility goal of GAW.

1 Introduction

Carbon monoxide affects the oxidizing capacity of the troposphere, and, in particular, plays an important role in the cycles of hydroxyl radical (OH), hydroperoxyl radical (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 (as compared with long-life greenhouse gases) and uneven distribution of its sources leads to 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 data 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 on the island of Tenerife (Canary Islands, Spain), well above a strong subtropical temperature inversion layer. Mean solar time is UTC−1. In situ measurements at Izaña are representative of the subtropical Northeast Atlantic free troposphere, especially 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 cannot pass above it, and there is a regime of downslope wind caused by radiative cooling of the ground. The station is located on the top of a crest, where horizontal divergence of the downslope wind and subsidence of the air from above the station occurs. During daytime an upslope wind caused by radiative heating of the ground transports to Izaña a small amount of contaminated air coming from below 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 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 may 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). As part of our quality assurance procedure, we compare the Izaña in situ quasi-continuous measurements with NOAA collocated flasks (Sect. 5). The differences between the measurements are evaluated in terms of their comparison uncertainty. Temporally averaged differences (e.g. annual means) also take into account the comparison uncertainty.

2 Measurement system configuration

The 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 above ground level. 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 to the analytical system using a KNF diaphragm pump. Water vapour is removed by flowing the air through a 300-mL glass flask immersed in a −67 °C alcohol bath. The residual level of water vapour downstream this trap is 5.3 ppm. 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. For both columns, the outer diameter is 3.2 mm and the length is 76.8 cm. The pre-column separates CO and H₂ from other trace gases in an air sample. The main column separates H₂ 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 used a stainless steel sample loop volume of 1 mL. Figure 1 shows a typical chromatogram, where the H₂ peak appears first, followed by the CO peak. Working standard gas (also called reference gas) and ambient air are injected alternatively 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), which is hosted by NOAA-ESRL-GMD. 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 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. Stability of the Izaña laboratory standards was evaluated in two ways, indicating there was no statistically significant drift in the mole fraction of these gases. First, in 2009, the WMO World Calibration Centre (WCC) for CO, which is hosted by EMPA, carried out an audit at the Izaña station (Zellweger et al., 2009) which included a blind analysis of five WCC travelling CO-in-air mixtures with mole fraction ranging from 88 to 201 nmol mol^{−1}. In

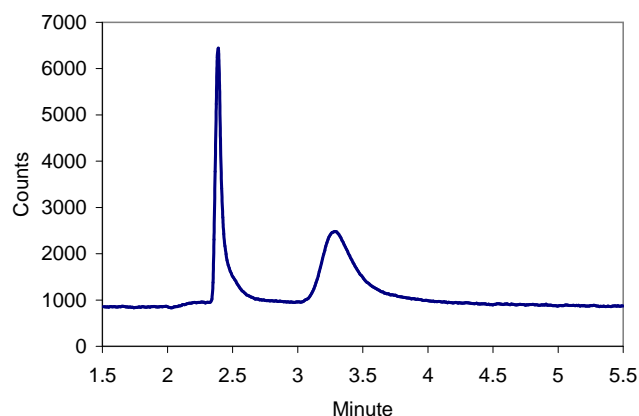


Fig. 1. Typical RGA chromatogram. The sample was injected at the 2 minute mark. The first eluted peak corresponds to H_2 , whereas the second one corresponds to CO.

each analysis, repeated injections of travelling cylinder gas alternate with working gas injections. The WCC assignments initially used were on an earlier version of the WMO scale, WMO-2000 (Zellweger et al., 2009). When the WCC travelling standards were revised to the WMO-2004 scale used at Izaña, the differences in the mole fractions assigned by Izaña and the WCC ranged from -1.69 to $2.63 \text{ nmol mol}^{-1}$ (C. Zellweger, personal communication, 2010). If we consider only the three travelling cylinders within the ambient range at Izaña ($\sim 60 - 150 \text{ nmol mol}^{-1}$), the differences range from -1.69 to $0.45 \text{ nmol mol}^{-1}$. The later values are compatible with the uncertainty in the Izaña RGA measurements (Sect. 4). Second, the stability of the laboratory standards was also evaluated by comparing Izaña in situ measurement results with results from air samples collected weekly in flasks at Izaña and analysed by NOAA-ESRL-GMD. The annual mean differences between CO results by the two laboratories are not significant for the years 2009, 2010, and 2011, and show no significant trend over this period (Sect. 5), indicating there was no significant change in the Izaña laboratory standards relative to their NOAA assignments.

The laboratory and working standards are contained in aluminium high-pressure cylinders fitted with Ceodeux brass valves (connection GCA-590). The 29-L cylinders containing the laboratory standards were obtained from Scott-Marrin, Inc, whereas the 20-L cylinders containing the working standards were obtained from Air Liquide Spain. They may differ in the type of aluminium alloy used and their internal conditioning. Two-stage high-purity regulators from Scott Specialty Gases (model 14C) are used, following the procedure for conditioning described by Lang (1998). Working gas tanks were filled to 125 bar 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).

Table 1. WMO CO standard gases of the Izaña station: CO mole fraction and standard uncertainty referenced to the WMO-2004 CO scale. The mole fractions were assigned by the WMO CO CCL in 2006.

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

We determine the response function of the instrument based on the standard/reference peak height ratios (in order to minimize potential artefacts due to changes in instrument response with time):

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

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 by 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 expressed as 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,

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

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 $R(h_i/h_{\text{wg}})$ is the fitted response function. Figure 2 shows the least-squares fitting of a typical calibration and the residuals with respect to this fit. 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”,

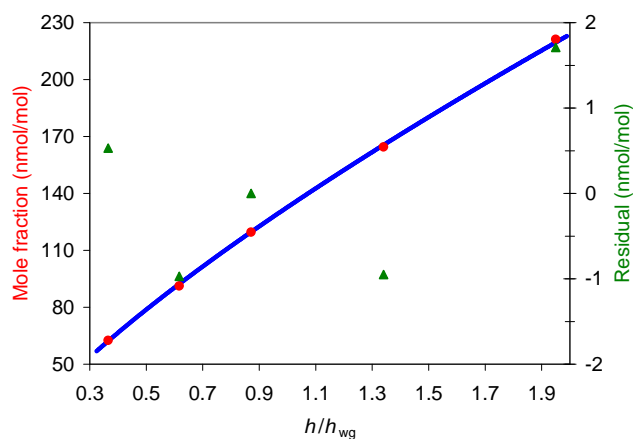


Fig. 2. Least-squares fitting of a typical calibration. The fitting curve is plotted in blue (left y-axis), the measured means are plotted in red (left y-axis), whereas the residuals with respect to the fitting curve are plotted in green (right y-axis).

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 for carrying out the weighted least-squares fitting. 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 $-1.63 \text{ nmol mol}^{-1} \text{ month}^{-1}$, and one with a positive drift of $2.75 \text{ nmol mol}^{-1} \text{ month}^{-1}$. These rapid changes likely result from the interaction of CO with the internal surface of the cylinders and the rapid decrease in their internal pressure (from 125 to 25 bar) during the few months they are in use. According to the experience of other laboratories, CO-in-air mixtures stored in aluminium tanks are usually prone to positive drift rates, whereas drift in our working gases is primarily negative. This may be due to the type of tanks used to store the working gases or to issues related to the Izaña station filling system. We consider the drifts are accurately determined and accounted for in the data processing.

After correcting for drift, the response curves were constructed and mole fractions are determined from the air measurements. Identification and discarding of outliers uses an iterative process of three filtering steps. We begin by 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 applied to the ambient air mole fraction time series 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-time 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 a 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 related to the temporal consistency of the response function parameters (u_{par}), which also takes into account the covariance between the parameters. The combined standard uncertainty (u_{tot}) is therefore given by

$$u_{tot} = \sqrt{u_{st}^2 + u_{fit}^2 + u_{rep}^2 + u_{par}^2}, \quad (3)$$

where

$$u_{st} = 7.40 \times 10^{-5} r^2 - 1.80 \times 10^{-2} r + 1.92, \quad (4)$$

u_{fit} is defined by Eq. (2),

$$u_{rep} = \frac{\beta r h_{wg} \sigma_{h/h_{wg}}}{\sqrt{3}h}, \text{ and} \quad (5)$$

$$u_{par}^2 = u_{pr}^2 + u_{p\beta}^2 + c. \quad (6)$$

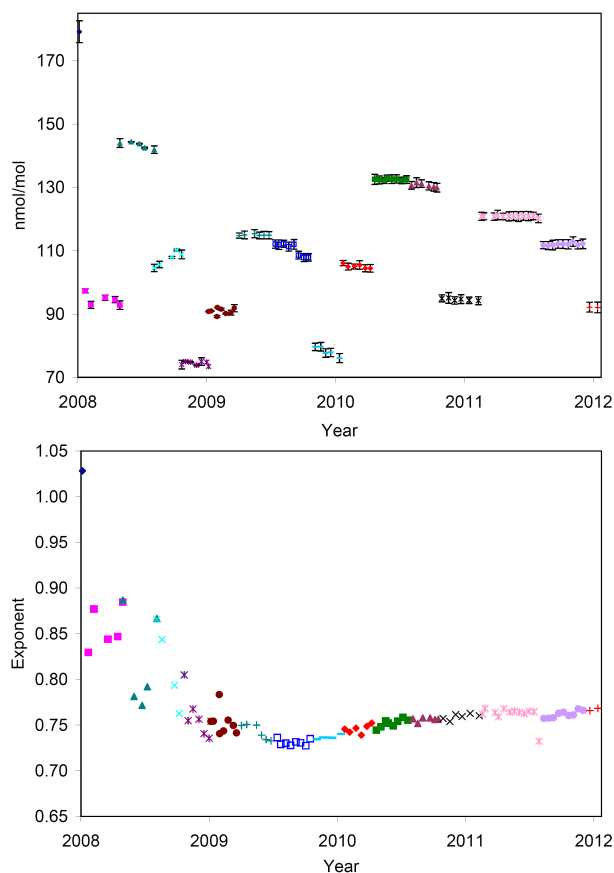


Fig. 3. Upper graph: working gas mole fractions obtained from calibrations conducted during 2008 to 2011. Error bars represent the RMS residual of each calibration, i.e. $\pm u_{\text{fit}}$. Lower graph: response function exponents obtained in the calibrations. Different colours and symbols are used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

$$u_{\text{pr}} = \frac{r}{r_{\text{wg}}} \sigma_{r_{\text{wg}}}, \quad (7)$$

$$u_{\text{p}\beta} = r \sigma_{\beta} \left| \log \frac{h}{h_{\text{wg}}} \right|, \text{ and} \quad (8)$$

$$c = 2 \frac{r^2}{r_{\text{wg}}} \text{covar}(r_{\text{wg}}, \beta) \log \frac{h}{h_{\text{wg}}}. \quad (9)$$

The units of r and u_{st} in Eq. (4) are nmol mol^{-1} ; $\sigma_{h/h_{\text{wg}}}$ 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; $\sigma_{r_{\text{wg}}}$ 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 $\text{covar}(r_{\text{wg}}, \beta)$ is the covariance between r_{wg} and β .

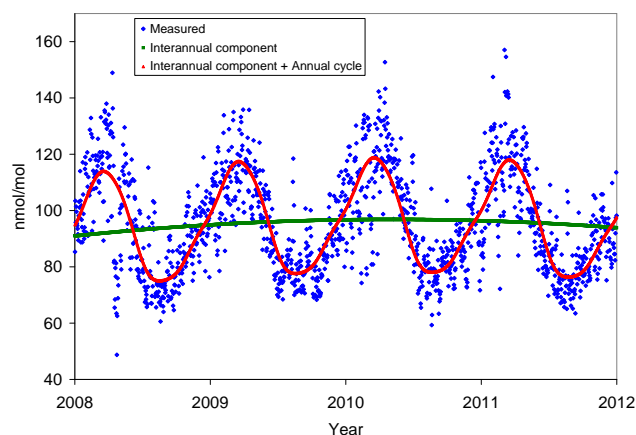


Fig. 4. Daily nighttime (20:00–08:00 UTC) mean CO mole fractions measured at Izaña Observatory (blue squares). The interannual component and the annual cycle from Eq. (26) are shown as the green and red lines, respectively.

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 (Table 1). Therefore, u_{st} represents the mole fraction dependent uncertainty due to the WMO standard gases, assuming they have been stable over time. As stated in Sect. 3, within the uncertainty of the measurements we do not observe significant drift in our laboratory standards. In a more general case of Eq. (4), u_{st} would account for (1) laboratory standard gas uncertainties increasing linearly in time due to undetermined potential drifts in the laboratory standards, and (2) the uncertainty in laboratory standard drift rates in case significant drifts had been determined. The term u_{fit} takes into account the disagreement between 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, a hypothetical decrease of u_{fit} when combining the information of successive calibrations cannot be expected. A mean value of u_{fit} is computed for each working gas used. As indicated in Sect. 3, the five laboratory standards detailed in Table 1 were used to determine the response function after March 2009. Before this date, a different set of three WMO standard gases was used, with CO mole fractions, $83.9 \text{ nmol mol}^{-1}$, $151.6 \text{ nmol mol}^{-1}$, and $165.7 \text{ nmol mol}^{-1}$. In this period, u_{fit} is unrealistically small due to the fact that the mole fractions of two of the three standards are near. To provide a better estimate of u_{fit} for curves determined before March 2009, this uncertainty component was forced to be at least equal to the mean value of u_{fit} after March 2009.

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

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

which relates errors (differentials). Obtaining the square of Eq. (10) and averaging over an appropriate ensemble, the terms $u_{\text{rep}}^2 + u_{\text{par}}^2$ are obtained. The only non-null covariance is that between the two parameters of the response function. The variables $\sigma_{r_{\text{wg}}}$, σ_{β} , and $\text{covar}(r_{\text{wg}}, \beta)$ are computed using the residuals of these parameters with respect to the considered linear drift in time or constancy in time. A single value for each variable per working gas is obtained. The typical value of $\sigma_{r_{\text{wg}}}$ is 1.09 nmol mol⁻¹ before March 2009, and 0.40 nmol mol⁻¹ after March 2009. 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, with its sign dependant on the mole fraction of the working gas. Therefore, the associated covariance has to be considered in the uncertainty computation. Note that the term u_{par} has been obtained, propagating only the parameter repeatabilities. u_{par} does not include other components of the parameter uncertainties. For example, it does not include the parameter uncertainties that could be estimated for each calibration following Sect. 8.1.2 of Martin (1971). Also, it does not include the whole uncertainty of r_{wg} (the mole fraction of the working gas). Therefore, what u_{par} takes into account is the temporal consistency of the response function. The term $\sqrt{u_{\text{st}}^2 + u_{\text{fit}}^2}$ provides the uncertainty in the response function for each calibration event (every two weeks). However, for the rest of time instants, the response function is used without performing any calibration, and therefore the term $\sqrt{u_{\text{st}}^2 + u_{\text{fit}}^2 + u_{\text{par}}^2}$ provides the uncertainty in the response function.

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

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

where k is a parameter equal to $(\sigma_h)/h_{\text{wg}}$, 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 $\sigma_{h/h_{\text{wg}}}$ 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 base does not depend on peak height, but on baseline noise).

Figure 5 shows the uncertainty components for the period 2008–2011. Table 2 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} , $u_{\text{p}\beta}$, and u_{par} are

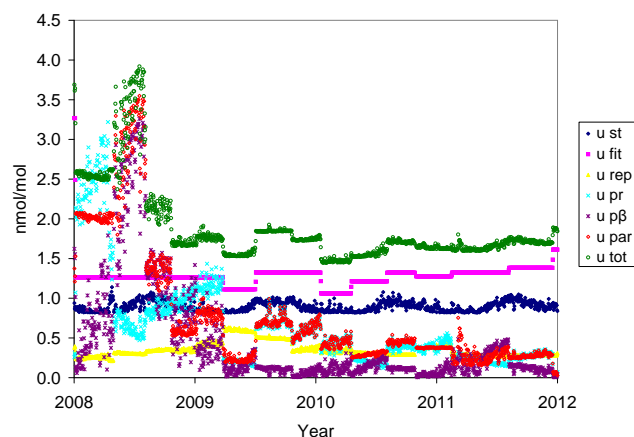


Fig. 5. Uncertainty components (daily means) of the measured CO mole fraction.

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} . Note that u_{pr} was larger than u_{tot} during part of 2008. According to Eq. (3), u_{tot} is larger than any of its four components (u_{st} , u_{fit} , u_{rep} , and u_{par}). However, u_{pr} and $u_{\text{p}\beta}$ can be larger than u_{par} and even u_{tot} for a negative large enough covariance term c (see Eq. 6).

4.1 The representation uncertainty and the propagated uncertainty of the temporal 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 a temporal mean from a number of available data (n) that is smaller than the theoretical maximum number of independent data (N) within the time interval in which the temporal mean is defined. The temporal mean computed from the n available data may be different from the mean determined from the N data (unknown). The representation uncertainty quantifies this difference statistically. 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 ($N = 24$, and $n = 22$). The value N is known precisely for each level except for the first. For example, in the Izaña data, the lowest ensemble is the hourly mean, for which $n = 3$ and N is unknown but certainly greater than n . The additional representation uncertainty is given by the equation

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

Table 2. Mean values of the uncertainty components (in nmol mol⁻¹) before and after March 2009.

Period	u_{st}	u_{fit}	u_{rep}	u_{pr}	$u_{\text{p}\beta}$	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

where

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

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. 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. Note that when $N \gg n$, the term between parenthesis in Eq. (12) becomes equal to 1; in such cases the exact value of N does not matter. 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). It assumes that the missing values are randomly distributed with respect to the mean. If this is not the case, the actual representation error could be larger than what the computed representation uncertainty predicts. For example, if three consecutive hours of a day are missing and they are located at an extreme of a significant diurnal cycle, the representation uncertainty will underestimate the actual representation error.

Any computed mean has also a propagated uncertainty arising from the uncertainties of the data used to compute this mean. Therefore, a temporal 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, combined quadratically: systematic, u_{syst} ; and random, u_{rand} . The law of propagation depends on the type of uncertainty. Therefore, we can write

$$u_{\langle r \rangle} = \sqrt{u_{\text{rs}}^2 + u_{\langle r \rangle, \text{rand}}^2 + u_{\langle r \rangle, \text{syst}}^2}, \quad (14)$$

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

$$u_{\langle r \rangle, \text{rand}}^2 = \frac{1}{n^2} \sum_{i=1}^n u_{\text{rand}_i}^2 \quad (15)$$

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

$$u_{\langle r \rangle, \text{syst}}^2 = \frac{1}{n} \sum_{i=1}^n u_{\text{syst}_i}^2 \quad (16)$$

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 no cancellation because the systematic error is the same (or nearly the same) for all the data used in the computation. The random uncertainty can be expressed as

$$u_{\text{rand}_i} = \sqrt{u_{\text{rep}_i}^2 + u_{\text{rs}_i}^2}, \quad (17)$$

and for the systematic uncertainty

$$u_{\text{syst}_i} = \sqrt{u_{\text{st}_i}^2 + u_{\text{fit}_i}^2 + u_{\text{par}_i}^2}. \quad (18)$$

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 (thus the propagated uncertainty may be slightly overestimated).

Table 3 shows mean values of the uncertainty components for hourly, daily nighttime, monthly, and annual means during the period 25 March 2009–31 December 2011. The hourly means correspond to the nighttime period (20:00–08:00 UTC). The mean representation uncertainty in the hourly means is 0.63 nmol mol⁻¹ for the nighttime 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 nighttime, 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.

As an example of the large representation uncertainty introduced when using very sparse data, we consider the ambient air samples collected weekly at Izaña Observatory

Table 3. Mean values of the uncertainty components (in nmol mol^{-1}) for different averaging periods from 25 March 2009–31 December 2011. The hourly means correspond to the nighttime 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 nighttime	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

since 1991 for analysis at NOAA-ESRL-GMD Carbon Cycle Greenhouse Gases 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. Monthly and annual means computed with such sparse flask data (typically 3 or 4 independent values per month) are subject to a large representation uncertainty. Table 4 shows mean values of the representation uncertainty in the different means determined from the NOAA measurements of flask samples. For the hourly and daily nighttime means based on a single pair of flasks, the associated σ_{sam} were computed using the quasi-continuous in situ measurements. The bias between NOAA measurements of daytime upslope air samples and the nighttime free troposphere measurements from the in situ monitoring system is considered in Sect. 6. The average standard deviation σ_{sam} of mole fractions within a month determined from NOAA measurements between 2009 and 2010 was $9.80 \text{ nmol mol}^{-1}$. Following the discussion above, it is not surprising that the representation uncertainties in the in situ means (Table 3) are much smaller than those from flask sampling.

5 Flasks-continuous comparison, comparison uncertainty, and means

Comparison of the results from in situ measurements and results from collocated flask air samples can be used as an independent way of assessing the quality of the continuous in situ measurements (WMO, 2011). A significant difference between a flask sample measurement result and a simultaneous in situ hourly mean is caused by two reasons: (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 collected at Izaña (see the last paragraph of Sect. 4.1), we proceed as follows.

1. Flasks results are used only if they are defined by NOAA as representative of background conditions, their sampling and analysis pass quality control checks (a pair of flasks is rejected if the difference between the two members of the pair is greater or equal to 3 nmol mol^{-1}), and the 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,

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

This standard deviation is indicative of the internal consistency of the pair.

2. The NOAA results are compared to hourly means determined in situ (the hour for which the mean is obtained must cover the time that the NOAA samples were collected). 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. Therefore, we compute the difference

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

and its comparison uncertainty

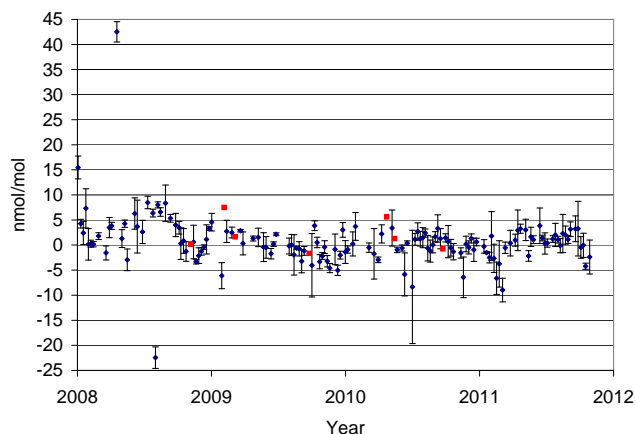
$$\sigma_{\text{dif}} = \sqrt{\sigma_f^2 + \sigma_c^2}. \quad (21)$$

Note that we use the in situ hourly means in the comparison since the grab samples are not collected simultaneously with measurements determined at the station. Due to the different temporal character of the compared measurements, σ_c must be used in Eq. (21) instead of the standard deviation of the hourly mean. The comparison uncertainty assesses if the difference is significant. If $|\text{dif}| \leq 2\sigma_{\text{dif}}$, the difference is not significant, whereas if $|\text{dif}| > 2\sigma_{\text{dif}}$, 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 measurement). 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 $2.5 \text{ nmol mol}^{-1}$ (a large fraction of this

Table 4. Mean values of the representation uncertainty (nmol mol^{-1}) for different averaging periods from the NOAA flask air samples.

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 nighttime	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

**Fig. 6.** Differences between NOAA flask samples and simultaneous in situ hourly means. Error bars indicate the 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.

dispersion is caused by the comparison uncertainty, since the 68 percentile of σ_{dif} is equal to $2.28 \text{ nmol mol}^{-1}$), whereas for 2008, 68 % of the differences are between -1.26 and $6.58 \text{ nmol mol}^{-1}$.

5.1 Annual and multi-annual means

We examine differences between the in situ and flask results over annual and longer periods of time. These differences are computed using conventional expressions and two expressions weighted by the comparison uncertainty. Note that the mean difference primarily results from potential systematic errors in the measurements from both laboratories.

The conventional mean is denoted as Mean,

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

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),

$$\langle \text{dif} \rangle_{\text{FW}} = \frac{1}{n} \sum_{i=1}^n \frac{\sigma_{\text{inv}}^2}{\sigma_{\text{dif}_i}^2} \text{dif}_i, \quad (23)$$

where

$$\frac{1}{\sigma_{\text{inv}}^2} = \frac{1}{n} \sum_{i=1}^n \frac{1}{\sigma_{\text{dif}_i}^2}. \quad (24)$$

This computation considers the quality of the measurements by applying weights to the differences. A difference with a larger uncertainty is considered to provide data of a lower quality and therefore the applied weight is smaller. WMean is an “intermediate” weighted mean for which Eq. (23) applies but $\sigma_{\text{dif}_i}^2$ is replaced by the median of σ_{dif}^2 for those $\sigma_{\text{dif}_i}^2$ smaller than the median of σ_{dif}^2 . This avoids an excessive weight being applied to those differences with a very small $\sigma_{\text{dif}_i}^2$. We believe that WMean is the most appropriate estimator. Differences without an associated uncertainty and three differences in 2008 exceeding (in absolute value) 10 nmol mol^{-1} are not included in the computation of the weighted means.

Table 5 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 $\pm 2 \text{ nmol mol}^{-1}$ compatibility goal of GAW (WMO, 2011). The results determined by the three approaches are not very different. The conventional annual mean differences are the closest to zero, except for 2008. This is not a general property, since, for example, for CO_2 and CH_4 we have observed many annual weighted mean differences closer to zero than the conventional mean difference.

Table 5 also shows the standard deviation of the mean, σ_{mean} . The standard deviation of the conventional mean can be determined by two approaches. First, this parameter is simply the standard deviation of the sample (SD) divided by \sqrt{n} (e.g. Martin, 1971, chapter 5). Alternately, the relation

$$\sigma_{\text{mean}} = \sqrt{\frac{1}{n^2} \sum_{i=1}^n \sigma_{\text{dif}_i}^2} \quad (25)$$

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

Table 5. Mean differences between results from the NOAA flask samples and coincident in situ hourly means (NOAA minus in situ) and standard deviations of the means (in nmol mol⁻¹). *n* dif denotes the number of differences available.

Period	<i>n</i> 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

Eq. (24). Note that the σ_{mean} associated with FWMean is smaller than those associated with 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 5 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 would be expected according to the values of σ_{dif} . Therefore, when computing weighted means, σ_{dif} does not include all the causes of variability within one year, i.e. it does not fully include the error components that behave as random within one year. Thus, when computing the weighted means, the smallest σ_{dif} are smaller than they should be, which makes FWMean 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 at 95 % confidence level if $|\langle \text{dif} \rangle| > 1.96\sigma_{\text{mean}}$, where $\langle \text{dif} \rangle$ denotes annual mean difference. As Table 5 shows, the conventional mean and the “intermediate” weighted mean differences (Mean and WMean, respectively) 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 (FWMean) is significant for all years except 2010, but as stated previously, it is not considered a good estimator for this dataset.

6 Time series analysis

To analyse the CO time series of daily nighttime 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).

The daily nighttime means, the fitted interannual component, and the fitted interannual component plus the fitted annual cycle are presented in Fig. 4. The RMS residual of the fit is equal to 11.5 nmol mol⁻¹. The nocturnal annual means (Table 6) were determined using the measured data when available and values from the curve fitted data when measured data were not available. As Table 6 shows, the number of days with data not available is very small. From 2008 to 2010 the CO annual mean increased by 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.

The annual cycle, as defined by the curve (see Fig. 4), shows an amplitude from the minimum to the maximum of 40.7 nmol mol⁻¹. The annual maximum occurs in late March, while the minimum is in the middle of 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.

The residuals from the curve can provide information about the air parcels influencing the site. A large change in the residuals indicates a change in air mass. The persistence of the residuals can be measured computing the autocorrelation (Fig. 7). 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. Therefore, 7 days could be considered the typical period of persistence of an air mass for CO.

Figure 8 shows the carbon monoxide monthly mean diurnal cycle relative to the nocturnal background, computed using hourly data for the period 2008–2011. This reference background level was computed for each day as follows. Firstly, the averages of the pre- (00:00–07:00 UTC) and post-nights (21:00–04:00 UTC) were computed. Then, the linear drift in time passing through both averages was

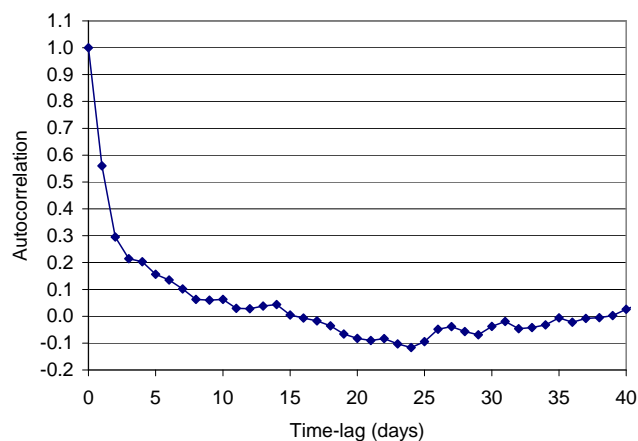


Fig. 7. Autocorrelation of the residuals from the fit determined from Eq. (26).

Table 6. Annual mean and the standard uncertainty of nighttime (20:00–08:00 UTC) CO mole fractions (in nmol mol^{-1}) measured between 2008 and 2011, and number of days with data available.

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

used as the reference background level for that day. Note, for example, that hour 1 means the hourly mean for the period 00:00–01:00 UTC. Carbon monoxide at Izaña is typically 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. 8). The amplitude of the diurnal cycle is 5–6 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 in late 1991, until 2002 flasks were sampled during nighttime, but this was discontinued due to the absence of staff during nighttime.

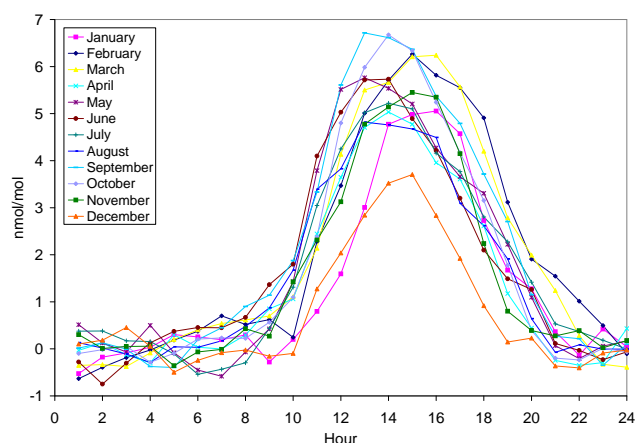


Fig. 8. Carbon monoxide mean diurnal cycle relative to the nocturnal background level.

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 temporal consistency of the response function parameters (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 for this improvement is a very significant reduction in the response function parameter standard deviations (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 the temporal 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 through averaging, whereas there is no cancellation of systematic errors. The representation uncertainties in the in situ measurements are much smaller than those computed using flask samples (e.g. the representation uncertainty in the monthly means is equal to $4.98 \text{ nmol mol}^{-1}$ for the NOAA flask measurements, and $0.03 \text{ nmol mol}^{-1}$ 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 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 amplitude of the seasonal cycle, determined from the curve, is $40.7 \text{ nmol mol}^{-1}$. The late winter maximum and mid-summer minimum observed at Izaña compare well with seasonal cycles reported from many other northern-hemisphere mid-latitude locations. 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 reaches a mid-day maximum of 4 to 6 nmol mol^{-1} above the stable nocturnal free troposphere background. This enhancement results from upslope winds which transport contaminated air coming from below the temperature inversion layer to the station. 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 of the statistical significance of the differences. During 2008, 47.4 % of the differences were significant, with 68 % of these between -1.26 and $6.58 \text{ nmol mol}^{-1}$. During 2009–2011, only 24.5 % of the differences were significant and 68 % were between -2.39 and $2.5 \text{ nmol mol}^{-1}$ (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 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. We thank the editor and two anonymous referees for their comments, which helped to improve the manuscript. AJGP thanks Celia Milford for English language comments.

Edited by: O. Tarasova

References

- Armerding, W., Comes, F. J., Crawack, H. J., Forberich, O., Gold, G., Ruger, C., Spiekermann, 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, 10603–10611, doi:10.1029/96JD03714, 1997.
- Conway, T. J., Tans, P., Waterman, L. S., Thoning, K. W., Masarie, K. A., and Gammon, R. H.: 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.
- 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, Chapter 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.
- 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.
- 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.
- 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.
- Komhyr, W. D., Gammon, R. H., Harris, T. B., Waterman, L. S., Conway, T. J., Taylor, W. R., and Thoning, K. W.: Global atmospheric CO₂ distribution and variations from 1968–1982 NOAA/GMCC CO₂ flask sample data, J. Geophys. Res., 90, 5567–5596, doi:10.1029/JD090iD03p05567, 1985.
- 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.

- 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.
- Martin, B. R.: *Statistics for Physicists*, Academic Press, London, UK, 1971.
- Navascues, B. and Rus, C.: Carbon dioxide observations at Izaña baseline station, Tenerife (Canary Islands) – 1984–1988, *Tellus B*, 43, 118–125, doi:10.1034/j.1600-0889.1991.t01-1-00006.x, 1991.
- 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.
- 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, 19015–19033, doi:10.1029/98JD01366, 1998.
- 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.
- 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.
- 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.
- 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.
- 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.
- Thoning, K. W., Conway, T. J., Zhang, N., and Kitzis, D.: Analysis system for measurement of CO₂ mixing ratios in flask air samples, *J. Atmos. Ocean. Tech.*, 12, 1349–1356, doi:10.1175/1520-0426(1995)012<1349:ASFMOC>2.0.CO;2, 1995.
- WMO: Guidelines for the Measurement of Atmospheric Carbon Monoxide, GAW Report No. 192, World Meteorological Organization, Geneva, Switzerland, 2010.
- 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.
- Zellweger, C., Klausen, J., Buchmann, B., and Scheel, H. E.: System and Performance Audit of Surface Ozone, Carbon Monoxide, Methane and Nitrous Oxide at the Global GAW Station Izaña, Spain, March 2009, WCC-Empa Report 09/1, 2009.