Inter-comparison of four different carbon monoxide measurement techniques and evaluation of the long-term carbon monoxide time series of Jungfraujoch

C. Zellweger, C. Hüglin, J. Klausen, M. Steinbacher, M. Vollmer, and B. Buchmann
Empa, Swiss Federal Laboratories for Materials Testing and Research, Laboratory for Air Pollution/Environmental Technology, 8600 Dübendorf, Switzerland

Received: 19 November 2008 – Published in Atmos. Chem. Phys. Discuss.: 26 January 2009
Revised: 18 May 2009 – Accepted: 18 May 2008 – Published: 2 June 2009

Abstract. Despite the importance of carbon monoxide (CO) for the overall oxidative capacity of the atmosphere, there is still considerable uncertainty in ambient measurements of CO. To address this issue, an inter-comparison between four different measurement techniques was made over a period of two months at the high-alpine site Jungfraujoch (JFJ), Switzerland. The measurement techniques were Non-dispersive Infrared Absorption (NDIR), Vacuum UV Resonance Fluorescence (VURF), gas chromatographic separation with a mercuric oxide reduction detector (GC/HgO), and gas chromatographic separation followed by reduction on a nickel catalyst and analysis by a flame ionization detector (GC/FID). The agreement among all techniques was better than 2% for one-hourly averages, which confirmed the suitability of the NDIR method for CO measurements even at remote sites. The inter-comparison added to the validation of the 12-year record (1996–2007) of continuous CO measurements at JFJ. To date this is one of the longest time series of continuous CO measurements in the free troposphere over Central Europe. This data record was further investigated with a focus on trend analysis. A significant negative trend was observed at JFJ showing a decrease of 21.4±0.3% over the investigated period, or an average annual decrease of 1.78%/yr (2.65±0.04 ppb/yr). These results were compared with emission inventory data reported to the Long-range Transboundary Air Pollution (LRTAP) Convention. It could be shown that long range transport significantly influences the CO levels observed at JFJ, with air masses of non-European origin contributing at least one third of the observed mole fractions.

1 Introduction
Carbon monoxide (CO) plays an important role in atmospheric chemistry. Reactions involving CO provide the dominant sink for the hydroxyl radical (Logan et al., 1981), and together with nitrogen oxides, the level of CO largely controls the overall oxidative capacity of the atmosphere. As a consequence, changes in CO emissions have an influence on climate by affecting methane and other greenhouse gases that are oxidized by the OH radical (Daniel and Solomon, 1998; Wild and Prather, 2000). Furthermore, CO plays an important role as a precursor of tropospheric ozone (Levy et al., 1997). CO has a relatively long atmospheric lifetime, ranging from 10 days in summer over continental regions to more than a year over polar regions in winter (Holloway et al., 2000b). This lifetime is long enough to make use of CO as a sound tracer for anthropogenic pollution.

In-situ measurements at remote sites are often made using gas chromatographic technique combined with a mercuric oxide detector (HgO) (Novelli, 1999). This technique has a low detection limit and good precision; however, non-linearity issues require careful calibration, and drift of standards with ambient CO mole fractions required for the calibration of this technique may further affect the accuracy of such measurements (Novelli et al., 2003). In addition to this method, several other techniques for the detection of CO with different temporal resolutions and detection limits have become available. The most common techniques currently applied comprise gas chromatographic (GC) techniques in combination with a flame ionization (FID) detector, and photometric methods such as non-dispersive infrared absorption (NDIR), vacuum ultra-violet resonance fluorescence (VURF) and tunable diode lasers spectroscopy (TDLS).

Despite its importance and the relatively large numbers of different measurement techniques employed, there is still considerable uncertainty in ambient measurements of CO.
To date, no comprehensive CO instrument inter-comparisons have been published, although a few older studies compare TDLs instruments with GC/HgO (Hoell et al., 1987), NDIR (Fried et al., 1991), and a VURF instrument (Holloway et al., 2000a). A short inter-comparison campaign with an NDIR and a VURF instrument at Jungfraujoch (JFJ) showed good correlation between the two techniques, but with absolute differences of 20–30 ppb (Whalley et al., 2004). These differences were attributed to the use of different calibration gases or interfering species in one of the techniques. A short inter-comparison between a GC/HgO and an NDIR instrument showed good overall agreement (coefficient of determination $r^2=0.88$), but slightly larger deviations at mole fractions below 100 ppb (Tsutsumi and Matsueda, 2000). A more recent study (Tanimoto et al., 2007) investigated differences between an NDIR and a GC/HgO instrument, along with the use of different NDIR monitors. The agreement between the NDIR and the GC/HgO instrument was within 10 ppb for 60% of the one-hourly averages, but much larger deviations were found in the comparison of different NDIR monitors.

CO trends in the troposphere are important for the oxidizing capacity of the atmosphere and have been studied using data from observation networks. Early measurements of total column carbon monoxide from JFJ (Zander et al., 1989) showed an increase in CO of 1–2 ppb/yr between 1950 and 1980. A decrease of total column CO was reported for the 1980s and 1990s (annual change for selected periods of $-0.63\%$ and $-0.27\%$, respectively), and more stable mixing ratios were observed between 2004 and 2005 (Zander et al., 2008). A positive global trend was reported between 1980 and 1988 (Khalil and Rasmussen, 1988), but a negative trend was observed after 1988 (Khalil and Rasmussen, 1994). Decreasing CO mole fractions have also been reported between 1991 and 1993 (Novelli et al., 1994), and a more recent analysis showed an ongoing significant but less pronounced negative trend for global CO flask observations after 1995 (Novelli et al., 2003; Meszaros et al., 2005). Chevallier et al. (2008) analyzed long-term trends of CO over Western Europe. They estimated a negative trend of $-0.84\pm0.95$ ppb/yr for the Zugspitze (ZUG) site between 1991 and 2004. Most of the overall negative trend was attributed to the trend in spring (January–April: $-1.49\pm1.50$ ppb/yr), whereas no trend was evident for July–September ($-0.28\pm1.36$ ppb/yr).

CO measurements from JFJ have been used for the assessment of meteorological influences on trace gas levels (Forrer et al., 2000), and the validation of chemical transport models (Holloway et al., 2000b) and Lagrangian models (Folini et al., 2008). Independent emission control is becoming increasingly important for verification of international treaties such as the Montreal and Kyoto protocols. CO measurements are often used as a proxy for such estimations because CO emission inventories are relatively well known. For example, CO inventories from the European Monitoring and Evaluation Programme (EMEP) (Vestreng et al., 2005) in combination with in-situ CO and halocarbon measurements from JFJ have successfully been used to estimate halocarbon emissions in Europe (Reimann et al., 2005). Especially applications combining data of emission inventories with in-situ CO measurements for source apportionment require CO data of known high quality.

This study presents results from an inter-comparison of several currently used in-situ techniques (NDIR, VURF, GC/HgO, GC/FID) for the measurement of atmospheric CO, which are normally used in international programs and networks (such as GAW, EEA, EMEP). The measurements were carried out at the high alpine research station Jungfraujoch (JFJ), Switzerland. The aim of the study was to evaluate differences between various techniques and to estimate their uncertainties with respect to different temporal resolutions. In addition, the study added to the validation of an ongoing long NDIR CO time series of the JFJ site because it could be demonstrated that accurate and sufficiently precise CO measurements are possible with the NDIR technique for the use of source apportionment and trend analysis. The 12-year CO data record of JFJ is further presented with a focus on climatology and trends of CO in the remote continental troposphere.

2 Experimental

2.1 Measurement site

The high alpine research station Jungfraujoch (JFJ) (46°33′N, 7°59′E, 3580 m a.s.l.) is located on the main crest of the Bernese Alps, Switzerland. Details of the location and the measurements program can be found in the GAW Station Information System (GAWSIS, 2008). Further details of the station including the inlet system have been described elsewhere (Zellweger et al., 2000, 2003). JFJ is an excellent platform for long-term observations of the free troposphere due to its high elevation and year-round accessibility. It is part of the Swiss National Monitoring Network (NABEL) and one of the global stations of the Global Atmosphere Watch (GAW) programme.

2.2 Instruments and calibration procedures

2.2.1 NDIR: Horiba APMA-360CE

CO has been continuously monitored since 1996 using a commercially available NDIR monitor (APMA-360, Horiba) as part of the Swiss National Air Pollution Monitoring Network (NABEL). Modification of the instrument included drying of the air by a Nation dryer in split flow mode (Permapure PD-50T-24”). The instrument was calibrated approximately in monthly intervals using a commercial CO calibration gas referenced against NIST (National Institute of Standards and Technology) SRM (Standard Reference Material) standards. Automatic instrument zero checks were performed every 49 h using zero air (heated CO/CO$_2$ converter,
Molecular Sieve 3 Å, Sofnocat 423). The detection limit for individual 1-min samples is 20 ppb, and the overall measurement uncertainty is estimated to be ±5% (1σ) (Zellweger et al., 2000), which includes the uncertainty of the calibration standard, the H₂O interference, and the instrument precision.

In contrast to other commercially available NDIR CO monitors the Horiba APMA-360 uses “cross-flow modulation” to compensate for matrix effects in the NDIR absorption measurements. The air passes over a heated oxidation catalyst to selectively remove CO from the sample air with a frequency of 1 Hz. Other commercial instruments use a gas filter correlation technique; these instruments have shown reduced performance concerning zero drift in the past. Results of this study can therefore not easily be transferred to gas filter correlation NDIR CO monitors. This limitation is also confirmed by a recent inter-comparison by Tanimoto et al. (2007). They compared a Horiba APMA-360 instrument with two gas filter correlation monitors and found significant deviations, which were attributed to both the analytical performance of the NDIR instruments and the reference gases.

### 2.2.2 VURF: VUV-fluorescence: Aerolaser AL5001

VUV fluorescence measurements were made using a commercially available instrument (Aerolaser AL5001). The instrument was calibrated every 60 min using a natural air (i.e. partly spiked/purified ambient air) working standard. The instrument was operated using CO₂ (99.995%) in Ar (99.9999%) and N₂ (99.9999%) with an additional purifier (Aeronex Gate Keeper SS-400KGC-I-4S) as auxiliary gases. The sensitivity of the instrument decreased from initially 40 counts per second (cps) per ppb to 10 cps per ppb at the end of the campaign, which is still above the specified limits for operation. The operating principle is described elsewhere (Gerbig et al., 1999).

### 2.2.3 GC/FID: Agilent 6890N GC

An Agilent 6890-N gas chromatograph equipped with a flame ionization detector (FID) was used for the detection of CO and CH₄. CO was analyzed as CH₄ after passing a hydrogen-flushed nickel catalyst heated to 375°C. Chromatographic separation was achieved isothermally at 60°C by means of a Unibeads 1S and a Molecular Sieve 5Å column. Nitrogen (99.9999%), further purified by passing through a nitrogen purifier (ALL-Pure Nitrogen Purifier, Alltech), was used as the carrier gas. The sample loop size was 10 ml. The air was dried prior to injection using a Nafion dryer (Permapure MD-110-72SS). Air samples were measured every 30 min and were bracketed by working standard measurements. This system was found to be linear for CO based on a dilution experiment as described below.

### 2.2.4 GC/HgO-reduction detector: trace analytical RGA-3

The RGA-3 gas chromatographic analysis is based on mercuric oxide reduction and ultraviolet light detection. Details of the modifications to the original setup design can be found in Vollmer et al. (2007). Chromatographic separation was achieved isothermally at 105°C by means of Unibeads 1S (1/8” OD, 80 cm) and Molecular Sieve 13-X (1/8” OD, 130 cm) columns. Synthetic air, further purified with Sofnocat 514, was used as the carrier gas. The sample air and standard gases were passed through a Nafion dryer prior to injection. Instantaneous air samples were measured every 30 min and were bracketed by working standard measurements to determine and correct for short-term instrumental drift. The RGA-3 data were corrected for nonlinear instrument response which was characterized by dynamic dilution of a reference gas (1.3 ppm CO in synthetic air) with CO free synthetic air using two mass flow controllers and calibrated flow meters. These dilution ratios were independently checked by simultaneous analysis of CH₄ on the above mentioned GC-FID, for which linearity was assumed.

### 2.2.5 Calibration standards

All measurements were traced back to a common reference standard (CA02854, 295.5±3.0(2σ)) ppb CO in natural (ambient) air, certified NOAA/ESRL (National Oceanic and Atmospheric Administration / Earth System Research Laboratory) standard, WMO-2000 CO calibration scale). The traceability of the measurements to this standard is illustrated in Fig. 1. Table 1 gives an overview of working standards used including traceability and uncertainties. Traceability of the long-term NDIR time series to this common reference was assured by propagation of NIST SRM traceable working standards. All NIST SRMs were cross-checked against secondary standards to assure the internal consistency of the working standards.

The common WMO-2000 reference standard was regularly checked for stability against primary reference standards from NIST, NMI (Nederlands Meetinstituut) and NPL (National Physical Laboratory), along with secondary laboratory standards starting in 2000; no observable drift was found since then. A comparison of the NIST SRM 2612a 23-F-06 standard against the NOAA/ESRL WMO-2000 scale through dynamic dilution showed that the NIST SRM was higher by 0.28±0.16(2σ)% compared to the WMO-2000 scale. This was not corrected in the data evaluation of the NDIR instrument because the difference was smaller than the certified uncertainty of the NIST standard gas; however, it has to be considered in the calculation of the uncertainty of the working standard used for the NDIR calibration. In addition, a number of comparisons of our NOAA/ESRL WMO-2000 standard with other NIST standard gases (SRM 1677c, SRM 2612a 23-F-32) always showed an agreement better than ±5% (1σ) (Zellweger et al., 2000).
Table 1. Overview of standards used for the calibration of the CO instruments. The uncertainty for the individual standards was estimated including the traceability to the common reference (WMO-2000); the expanded uncertainty includes the uncertainty of the NOAA WMO-2000 standard. All uncertainties are given for the 95% confidence level (2σ).

<table>
<thead>
<tr>
<th>Instrument/WS</th>
<th>Mole fraction (ppb)</th>
<th>Traceability chain</th>
<th>Uncert. (%)</th>
<th>Expanded uncert. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VURF / CA06439</td>
<td>438.8</td>
<td>WMO-2000</td>
<td>0.3</td>
<td>1.0</td>
</tr>
<tr>
<td>NDIR / SL68874</td>
<td>2020.0</td>
<td>NIST – WMO-2000</td>
<td>2.1</td>
<td>2.3</td>
</tr>
<tr>
<td>GC/FID / CC106830</td>
<td>177.9–180.0(^a)</td>
<td>WMO-2000</td>
<td>0.9</td>
<td>1.4</td>
</tr>
<tr>
<td>GC/HgO / E-033</td>
<td>241.1</td>
<td>CC106830 – CA06439 – WMO-2000</td>
<td>1.7</td>
<td>2.0</td>
</tr>
</tbody>
</table>

\(^a\) A time dependent drift correction was applied

Fig. 1. Traceability of the four CO instruments to a common reference standard. The black arrows indicate the calibrations made during the inter-comparison campaign.

than 0.5% (average 0.15±0.09(2σ)%), which is well within the individual stated uncertainties of the NIST SRM standards (0.5–1.2%,2σ). Therefore the NOAA/ESRL WMO-2000 scale based on this particular cylinder is not considered to be significantly different from the NIST CO scale.

Working standards were calibrated using the NOAA/ESRL and NIST laboratory standards for field calibrations at the JFJ, as illustrated in Fig. 1. The following working standards were used for the calibration of the instruments during the campaign:

NDIR – Horiba APMA 360: A 10 liter aluminum Luxfer cylinder (Messer Schweiz GmbH) containing CO in nitrogen was used as a calibration gas. This cylinder was assigned 2.02 ppm CO based on initial calibration against a NIST SRM 5-i-04 (9.66 ppm CO in N\(_2\)) in April 2005. This value was confirmed after use of the cylinder in December 2006 with NIST SRM 23-F-06 (9.75 ppm CO in air). The uncertainty of the NDIR working standard due to calibration was estimated to be ±2.3%(2σ) from the inter-comparison between the NOAA/ESRL and NIST reference standards (0.56%, 2σ), the contribution of an imperfect calibration on the laboratory NDIR system due to instrument noise of 12 ppb at zero and 16 ppb at span (2.0%, 2σ), and the uncertainty of the NOAA/ESRL standard (1%, 2σ).

VURF – Aerolaser AL5001: A 301 Scott Marrin aluminum cylinder (Luxfer) containing pressurized ambient air (RIX SA-3 oil free compressor) was used as a calibration gas. This standard was calibrated several times against the NOAA/ESRL certified standard (CA02854, 295.5±3.0(2σ) ppb CO in natural (ambient) air, WMO-2000 scale) before and after the campaign, and was assigned a CO mole fraction of 438.8 ppb. No significant drift was observed in this cylinder over its lifetime between July 2005 and September 2006. The uncertainty of the VURF working standard was estimated to be 1.0% (2σ) based on multiple calibrations against the NOAA/ESRL standard. Note that the VURF instrument requires a standard in natural air because the UV fluorescence reaction is quenched by oxygen. No matrix effects for standards balanced with air or nitrogen are known for the other analytical techniques.

GC/FID: A working standard with the same cylinder type and material as the one for the Aerolaser instrument was used for automatic calibrations; however, CO in this cylinder was less stable, and a continuous and constant upward drift of CO was observed over time. The cylinder was calibrated against the NOAA/ESRL certified standard described above several times before and after the campaign, and against the working standard of the VURF instrument during the campaign. Based on these measurements, a linear correction was applied, with a CO increase rate of 0.033 ppb per day and an initial mole fraction of 172.0 ppb on 15 July 2005. This allowed the calculation of the working standard mole fractions at all times during the campaign. The estimated uncertainty of this standard is 1.4% (2σ) from a linear interpolation of multiple calibrations against the NOAA/ESRL standard.

GC/HgO: An electro-polished stainless steel tank (Essex Cryogenics) filled with natural ambient air was used as a working standard. This standard was stable at 241.1 ppb CO
for the period of the campaign. The standard was referenced against the standard of the GC/FID instrument. The uncertainty of the standard was estimated to be 2.0% (2σ).

3 Results and discussion

3.1 Field inter-comparison

Measurements with four CO instruments employing four different analytical techniques were performed over a period of approximately two months between 11 January and 15 March 2006. Data availability based on one-hourly averages was 96.7% (NDIR), 86.4% (VURF), 86.7% (GC/FID), and 98.9% (GC/HgO). For the continuous techniques, one-hourly averages were only calculated when at least four 10-min averages were available. Hourly averages of the GC observations typically represent the average of two single injections. Figure 2 shows the available time series for all four techniques and the difference between the NDIR / GC techniques and the VURF instrument. The overall variability was well captured by all techniques. The CO mole fractions during the campaign ranged from approximately 100 to 260 ppb, which is consistent with other studies at the JFJ site (Forrer et al., 2000; Zellweger et al., 2003). Figure 2 shows that the NDIR results were slightly higher throughout the entire campaign, with an almost constant bias irrespective of the CO level. However, the deviations of the NDIR from the VURF results seem to decrease slightly towards the end of the inter-comparison. The reason for this may be changing NDIR zero readings during the campaign. Figure 3 shows the individual zero readings of the NDIR instrument (1-min averages) made automatically every 49 h, as well as the corresponding 30 min averages. These zero readings averaged −0.3±8.4 (2σ) ppb and were not significantly different from zero over the entire period. Consequently, due to the relatively high uncertainties of the individual zero readings, no further correction was applied to the data. However, neither a potential drift nor an offset can be excluded based on these data, which potentially explains the small observed difference.

Table 2 shows the parameters obtained from orthogonal regression analysis (York, 1966) between different techniques based on one-hourly averages. A generally good agreement was found among all techniques, with the highest correlation between the two continuous methods (VURF and NDIR, \(r^2=0.992\)). A slightly lower but still excellent correlation was found between the continuous and the GC methods (\(r^2\) between 0.962 and 0.981), while the lowest correlation was observed between the two GC methods (GC/FID and GC/HgO, \(r^2=0.935\)). Due to the large number of measurement points, the estimated uncertainties of slope and intercept were relatively small, and significant differences were found between all time series. A pair-wise Wilcoxon-Mann-Whitney test based on one-hourly averages also confirmed significant differences between all possible combinations (p-value<0.01 with Bonferroni correction for multiple testing), with the exception of the VURF-NDIR instrument pair (p-value=0.42). These results suggest that the differences observed by Whalley et al. (2004) between the JFJ NDIR system and their VURF instrument of 20–30 ppb are likely due to the use of calibration standards that were not traceable to NIST or WMO-2000 CO scales, or from instrumental faults, such as leaks.

The agreement between the various time series is further illustrated in relative difference histograms (Fig. 4) compared to the VURF as the reference instrument for averages of 1-, 10-, and 60 min respectively. Single injections of the GC techniques were compared to 1 min and 10 min data, and the average of (usually) 2 GC injections was used to compare with for hourly averages. The relative differences \((x−x_{\text{ref}})/x_{\text{ref}}\) are shown where \(x_{\text{ref}}\) is the CO mole fraction measured by the reference instrument (VURF). In addition, the individual uncertainties of the calibration standards (cf.
Table 2. Results of the orthogonal regression analysis between the different measurement techniques, where x and y are the corresponding instruments, and a and b are the intercept and slope of the regression line with 95% confidence intervals. $r^2$ is the coefficient of determination, and N is the number of data points. In addition, the p-value of the Wilcoxon-Mann-Whitney test is shown. Comparisons are based on one-hourly averages.

<table>
<thead>
<tr>
<th>y</th>
<th>x</th>
<th>a [ppb]</th>
<th>b</th>
<th>$r^2$</th>
<th>N</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VURF</td>
<td>NDIR</td>
<td>-1.9±0.8</td>
<td>0.997±0.005</td>
<td>0.992</td>
<td>1273</td>
<td>0.424</td>
</tr>
<tr>
<td>VURF</td>
<td>GC/FID</td>
<td>6.1±1.2</td>
<td>0.967±0.008</td>
<td>0.981</td>
<td>1165</td>
<td>3.95e-12</td>
</tr>
<tr>
<td>VURF</td>
<td>GC/HgO</td>
<td>-4.4±1.5</td>
<td>1.033±0.010</td>
<td>0.970</td>
<td>1292</td>
<td>4.03e-04</td>
</tr>
<tr>
<td>NDIR</td>
<td>GC/FID</td>
<td>9.8±1.3</td>
<td>0.960±0.009</td>
<td>0.975</td>
<td>1271</td>
<td>3.19e-10</td>
</tr>
<tr>
<td>NDIR</td>
<td>GC/HgO</td>
<td>-2.0±1.6</td>
<td>1.033±0.010</td>
<td>0.962</td>
<td>1448</td>
<td>1.50e-05</td>
</tr>
<tr>
<td>GC/FID</td>
<td>GC/HgO</td>
<td>-1.2±2.2</td>
<td>1.070±0.016</td>
<td>0.935</td>
<td>1299</td>
<td>2.41e-03</td>
</tr>
</tbody>
</table>

Table 1) and the additional uncertainty due to imperfect zero compensation of the NDIR instrument are shown. The maximum of the distribution of the differences was in all cases within the uncertainty limits of the calibration standards. It can therefore be concluded that the mean differences of the various time series are due to differences in the calibration standards. It can further be seen that the averaging time has a significant influence on the width of the relative difference distribution for the NDIR technique. The standard deviation of the relative difference distribution is comparable for all techniques for one-hourly averages, with the lowest value for the NDIR technique. This implies that the performance of the NDIR technique for one-hourly averages is equal or even slightly better compared to the GC methods. At the 10 min level the noise of the NDIR technique was significant, but the performance was still comparable to the GC techniques. The GC techniques showed a considerable number of values with large deviation compared to the VURF method, resulting in long tails of the distribution of the relative differences at the 1- and 10 min levels; this can be explained with the different temporal coverage was different (single injections vs. integration over ten min). The number of these outliers was relatively insensitive to the level of aggregation, but the overall width of the frequency distribution increased slightly due to the fact that instrument noise was becoming more of an issue. This was also the case for the VURF technique. Instrument noise was clearly the dominating factor for the NDIR technique on the 1 min level. The width of the distribution of the 1 min relative differences of the NDIR technique was comparable to data obtained from our laboratory experiments with a Horiba APMA-360 NDIR CO monitor. This was demonstrated through an additional experiment, during which the 1 min average noise over 24 h was determined at a constant mole fraction of 152 ppb, which is similar to the average mole fraction during the JFJ campaign. Based on this experiment a standard deviation of the relative difference for a mole fraction of 152 ppb of 0.0750 was calculated during the laboratory experiment compared to 0.0768 during the JFJ campaign. For the 10 min and hourly averages, the corresponding numbers were 0.0336 (0.0310 at JFJ), and 0.0208 (0.0171 at JFJ). This clearly demonstrates that instrument noise is a limiting factor for the determination of CO levels with the NDIR technique if high temporal resolution is required. It can be seen from Fig. 4 that the averaging interval has a significant influence on the uncertainty of the NDIR measurements. One-hourly averages of the NDIR instrument achieve a data quality that is comparable to the GC instruments, with even slightly lower relative differences because of fewer outliers. This is confirmed by the Wilcoxon-Mann-Whitney test, which also does not yield significant differences between the NDIR and the GC techniques when one-hourly averages are compared.

The mean value of the relative difference is a measure of the difference in location of the data points compared to the VURF technique. The largest mean deviation (1.58%) was observed between the VURF and NDIR technique. However, the calibration standard of the NDIR instrument also has a relatively high uncertainty, and at least part of the bias can be explained by differences in the calibration. In addition, imperfect compensation of the zero offset (cf. Fig. 3) may also contribute to the observed bias. Automatic zero checks were made every 49 h, which is potentially insufficient for an accurate compensation of the zero offset. The standard deviation of the zero readings obtained during the campaign was 4.2 ppb (31 observations); this results in an additional uncertainty of the NDIR measurements of 1.0%.

To illustrate the performance as a function of the averaging time, a selected time period is shown in Fig. 5. During the period from 11 to 13 March, rapid changes in the ambient mole fractions occurred. Instrument noise of the NDIR technique was dominant at the one minute level, but good agreement was observed between all techniques for 10 min and one-hourly averages. All instruments were able to detect fast changes in the CO mole fractions that occurred in the second half of the selected period. More interesting is the first half of the period that was characterized by relatively small changes of the mole fractions. Part of this period was characterized by large short-term variations as apparent from...
Fig. 4. Relative difference histograms for the NDIR and GC instruments calculated relative to a common reference instrument (VURF). Each panel shows the frequency of data falling into 0.01 relative difference bins (normalized to the number of coincident data points). Relative differences for one-hourly, 10-min and 1-min averages are shown. 1- and 10-min averages of the GC techniques represent single injections. The red shaded areas represent the uncertainty of the calibration standards, and the blue shaded areas the uncertainty due to imperfect zero compensation (NDIR only). P(%) is the percentage of data falling within the uncertainty limits.

the one minute VURF data. These fast changes could only be detected with the VURF technique because the instrument noise of the NDIR monitor is too large to allow detection of mole fraction changes of a few ppb on a temporal scale ranging from seconds to a few minutes. The GC techniques were able to accurately reflect the CO mole fraction, but lack temporal resolution; consequently, differences of integrated values between continuous and GC methods may be significantly higher compared to periods with less pronounced short-term variation in the CO mole fractions. The period
with pronounced short-term variability is further highlighted in Fig. 5d. In the first six hours of the selected period significant short-term variability in the CO mole fractions was observed, which is visible in the VURF 1 min and 10 s averages. During this period, the agreement between the 10 min VURF average and the single GC injections was considerably lower compared to the following hours with more stable CO mole fractions. The lack of temporal coverage of the GC methods explains to a large extent the lower correlation between the quasi-continuous and continuous techniques. In conclusion, one-hourly NDIR CO data of the JFJ station can be considered to be fully comparable to data obtained with a VURF.

3.2 Carbon monoxide trend at Jungfraujoch between 1996 and 2007

The JFJ carbon monoxide time series is one of the longest continuous datasets of CO measurements in the remote continental troposphere in Europe. Figure 6a shows the 12-year CO time series from 1996 to 2007 of JFJ obtained with NDIR instruments as described in the previous section, and a summary of monthly and yearly mean mole fractions is shown in Table 3. To investigate trends and seasonal behavior, the one-hourly CO data were decomposed into a quadratic trend and average seasonal cycle as shown in Eq. (1) (Thoning et al., 1989). This function has been successfully used to determine the long-term trend of baseline data from the NOAA/ESRL flask sampling network (Novelli et al., 1998; Novelli et al., 2003).

$$f(t) = a_0 + a_2 t + a_3 t^2$$

$$+ \sum_{i=1}^{4} [a_{2i+2} \sin(2\pi i t) + a_{2i+3} \cos(2\pi i t)]$$

(1)

The complete fit including the seasonal variation (light blue line) and a linear trend (orange line, see below) are also plotted in Fig. 6a. In addition, individual data points were discriminated between baseline conditions (blue) and pollution (red) events. These events were defined by assuming normal distribution of baseline values around the fitted function. To
define conditions for pollution events, the negative residuals of the fit were first mirrored at zero. The standard deviation of the distribution of negative residuals and mirrored negative residuals was then used to calculate the condition for pollution events. Values higher by more than two standard deviations were considered as pollution events. The annual growth rate curve is plotted in Fig. 6b. The growth rate was calculated as the difference between two annual moving averages based on daily data. To avoid a bias in the growth rate due to missing values, a loess fit was applied to the data and gaps were filled with predicted data based on the fitting parameters. For better comparability of the growth rate with ambient data presented in Fig. 6a, the growth rate values are centered on the time axis such that, for example, the value of 1 July 1999 represents the growth rate for the period from 1 January 1999 thru 31 December 1999.

It can be seen that the CO mole fractions decreased significantly during the period between 1996 and 2007. The trend part of the fit is close to linear with a fitted value of $a_3=0.050 \pm 0.014$. Due to the small contribution of the non-linear term $a_3$ Eq. (1) was simplified by setting $a_3=0$ to calculate the linear baseline CO trend at JFJ (orange line in Fig. 6a). The result is an average change of $-2.65 \pm 0.04$ ppb/yr, which corresponds to a decrease in baseline CO of 21.4% over the period 1996 to 2007 at JFJ.

The average and seasonal diurnal cycles of mean CO mole fractions are shown in Fig. 7a. A significant diurnal cycle with lowest values in the early morning and maximum values in late afternoon local time could only be observed during the warmer seasons (spring and summer). This is in line with observations from previous studies that documented the influence of more polluted atmospheric boundary layer air lifted by thermally induced flow systems (Forrer et al., 2000; Zellweger et al., 2003; Henne et al., 2004). It can also be seen from the yearly seasonal cycles (Fig. 7b) that the CO mole fraction decreased over the observation period, and reached its lowest values during summer 2007. Part of this inter-annual variability can be explained by global biomass burning. Elevated CO mole fractions were observed during periods with increased biomass burning, e.g. 1996 and 1998 (Yurganov et al., 2004; Wotawa et al., 2001), and also for the years 2002 and 2003 (Yurganov et al., 2005). These events are well captured by the growth rate as illustrated in Fig. 6b. However, elevated mole fractions during the summer of 2003 may also be explained by increased thermally induced vertical upward transport due to the extremely high Central European summer temperatures in 2003 and the forest fires in Portugal (Luterbacher et al., 2004; Tressol et al., 2008). Despite significant year-to-year variability, the trend over the observation period was rather constant and no significant seasonality was observed in the decrease (not shown). An analysis of the monthly JFJ data showed that the decrease was significantly lower for February and March at approx. $-1.2$ ppb/yr for both months. This observation is in line with a model study by Pfister et al. (2004) showing the largest contributions of Asian and North American anthropogenic CO in Europe between January and May. However, it is in contrast to observations at Zugspitze (ZUG) (Chevalier et al., 2008), where the overall annual downward trend was mainly attributed to a decrease in the winter / spring period (January–April). Furthermore, the magnitude of the CO decrease is significantly higher at JFJ ($-2.65 \pm 0.04$ ppb/yr, 1996–2007) when compared with ZUG ($-0.84$ ppb/yr, 1991–2004) (Chevalier et al., 2008). It should also be noted that the
Table 3. Monthly and yearly mean CO mole fractions (ppb) at Jungfraujoch for the years 1996 to 2007. Monthly mean values were only calculated if at least 75% of the data where available. Yearly averages were only calculated if the data coverage was larger than 90%.

<table>
<thead>
<tr>
<th>Month</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>NA</td>
<td>NA</td>
<td>203.7</td>
<td>NA</td>
<td>162.2</td>
<td>126.4</td>
<td>126.2</td>
<td>126.4</td>
<td>115.4</td>
<td>140.2</td>
<td>134.5</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>148.5</td>
<td>145.2</td>
<td>145.3</td>
<td>159.5</td>
<td>150.0</td>
<td>133.3</td>
<td>115.1</td>
<td>116.4</td>
<td>126.4</td>
<td>130.8</td>
<td>164.4</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>197.1</td>
<td>182.6</td>
<td>182.5</td>
<td>208.6</td>
<td>179.5</td>
<td>132.5</td>
<td>139.3</td>
<td>149.4</td>
<td>162.7</td>
<td>148.9</td>
<td>174.3</td>
<td>181.0</td>
<td>168.9</td>
</tr>
<tr>
<td>1999</td>
<td>173.8</td>
<td>172.1</td>
<td>179.2</td>
<td>180.6</td>
<td>148.5</td>
<td>138.2</td>
<td>130.5</td>
<td>NA</td>
<td>135.3</td>
<td>129.9</td>
<td>166.1</td>
<td>139.1</td>
<td>153.2</td>
</tr>
<tr>
<td>2000</td>
<td>NA</td>
<td>134.4</td>
<td>158.1</td>
<td>191.8</td>
<td>171.8</td>
<td>154.1</td>
<td>102.0</td>
<td>116.4</td>
<td>126.4</td>
<td>130.8</td>
<td>164.4</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>150.8</td>
<td>160.6</td>
<td>149.8</td>
<td>158.6</td>
<td>136.1</td>
<td>109.9</td>
<td>108.2</td>
<td>107.4</td>
<td>99.3</td>
<td>98.0</td>
<td>124.4</td>
<td>127.5</td>
<td>127.3</td>
</tr>
<tr>
<td>2002</td>
<td>NA</td>
<td>156.9</td>
<td>143.6</td>
<td>159.5</td>
<td>146.6</td>
<td>130.0</td>
<td>130.1</td>
<td>116.4</td>
<td>130.4</td>
<td>119.1</td>
<td>147.5</td>
<td>141.5</td>
<td>137.8</td>
</tr>
<tr>
<td>2003</td>
<td>150.8</td>
<td>156.1</td>
<td>145.2</td>
<td>156.6</td>
<td>133.3</td>
<td>117.5</td>
<td>120.7</td>
<td>123.0</td>
<td>124.8</td>
<td>121.4</td>
<td>124.3</td>
<td>130.8</td>
<td>133.7</td>
</tr>
<tr>
<td>2004</td>
<td>135.7</td>
<td>151.2</td>
<td>155.1</td>
<td>163.2</td>
<td>132.3</td>
<td>101.8</td>
<td>112.4</td>
<td>106.2</td>
<td>94.8</td>
<td>102.4</td>
<td>111.0</td>
<td>122.8</td>
<td>124.2</td>
</tr>
<tr>
<td>2005</td>
<td>134.3</td>
<td>166.6</td>
<td>144.8</td>
<td>NA</td>
<td>127.6</td>
<td>103.2</td>
<td>101.8</td>
<td>109.4</td>
<td>110.8</td>
<td>103.6</td>
<td>126.7</td>
<td>136.8</td>
<td>126.0</td>
</tr>
<tr>
<td>2006</td>
<td>148.3</td>
<td>151.8</td>
<td>153.0</td>
<td>148.7</td>
<td>132.7</td>
<td>119.0</td>
<td>115.4</td>
<td>105.6</td>
<td>108.6</td>
<td>98.5</td>
<td>108.7</td>
<td>120.5</td>
<td>125.7</td>
</tr>
<tr>
<td>2007</td>
<td>129.5</td>
<td>140.4</td>
<td>164.6</td>
<td>153.0</td>
<td>132.5</td>
<td>106.7</td>
<td>84.5</td>
<td>86.8</td>
<td>90.5</td>
<td>98.4</td>
<td>132.8</td>
<td>145.2</td>
<td>121.9</td>
</tr>
<tr>
<td>AVG</td>
<td>148.4</td>
<td>156.1</td>
<td>156.5</td>
<td>169.8</td>
<td>146.9</td>
<td>124.2</td>
<td>115.5</td>
<td>116.5</td>
<td>120.1</td>
<td>117.0</td>
<td>138.8</td>
<td>139.7</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 8. Emissions reported to the Long-range Transboundary Air Pollution (LRTAP) Convention (EMEP) normalized to 1997 for Switzerland and neighboring countries (grey colors), and yearly mean CO mole fractions normalized to 1997 for JFJ and other sites in Switzerland (see text for details).

The observed overall negative trend of CO at JFJ can mainly be explained by the reduction of European emission sources since the early 1990s. Fossil fuel emissions are the largest contributor to the CO burden in the northern extratropics (Duncan et al., 2007). European CO emissions of the EU-15 member states decreased from 39.1 Gt in 1996 to 24.2 Gt in 2005 (EEA, 2007), which corresponds to a decrease of 38.1%. During the same period the measured CO mole fraction at JFJ decreased by 17.8%, which is less than half of what would be expected if only European emissions were contributing to the CO levels observed at JFJ. A comparison of emission inventory data with CO measurements from sites situated at lower altitudes in Switzerland is shown in Fig. 8. Emission data were taken from the reports to the Long-range Transboundary Air Pollution (LRTAP) Convention (EEA, 2007). Yearly mean values of both station and emission inventory data were normalized to 1997. CO mole fractions decreased most at two curbside sites Bern (BER) and Lausanne (LAU). The agreement with the emission inventory data is very good if only road transportation emissions are considered (not shown). Data of the emission inventory also agree well with measured CO mole fractions at the two urban sites, Lugano (LUG) and Zürich (ZUE), as well as the two rural sites, Häringen (HAE) and Sion (SIO), which are both situated adjacent to a highway and thus highly influenced by traffic emissions. The observations at these stations seem to be representative for European CO emission trends and reliably represent the mixture of traffic and
The authors would like to thank the Interna-
tional Foundation Jungfraujoch and Gornergrat and the custodians of the JFJ Research Station for their support and hospitality. Parts of the measurements were made by the Swiss National Air Pollution Monitoring Network (NABEL) in collaboration with the Swiss Federal Office for the Environment (FOEN). The work was also supported by MeteoSwiss through engagement in the Global Atmosphere Watch programme.

4 Conclusions

An inter-comparison between four different measurements techniques (NDIR, VURF, GC/FID, and GC/HgO) for the measurement of atmospheric CO showed excellent agreement among all analytical techniques based on one-hourly averages. The observed differences could be explained with remaining biases of calibration standards. Thus, when other potential issues such as non-linearity are carefully considered in the measurement set-up, the limiting factor for accurate CO measurements is the uncertainty of the calibration standards. In addition, the NDIR technique requires careful zero compensation to achieve data of sufficiently high quality.

The inter-comparison experiment added to the validation of the 12 year long CO time series made by NDIR technique at the JFJ. Results show a clear decrease of the CO burden over Europe during the past decade, which is in agreement with decreasing CO emissions in Europe. Further examination of this time series showed that CO decreased by 21.4% in the period from 1996 to 2007 at JFJ. This trend fits well into the context of decreasing CO emissions in Europe. However, comparisons with emission inventory data showed that a significantly larger decrease would be expected if European emissions alone were driving CO mole fractions at JFJ. Therefore, long-range transport is considered to have a significant influence on the CO levels at JFJ. It was estimated that a least one third of the baseline CO measured at JFJ is of non-European origin after considering the fraction of CO produced by CH$_4$ and NMHC oxidation. This is in agreement with model studies that attribute a significant fraction of the European CO budget to non-European sources (Pfister et al., 2004).

Acknowledgements. The authors would like to thank the International Foundation Jungfraujoch and Gornergrat and the custodians of the JFJ Research Station for their support and hospitality. Parts of the measurements were made by the Swiss National Air Pollution Monitoring Network (NABEL) in collaboration with the Swiss Federal Office for the Environment (FOEN). The work was also supported by MeteoSwiss through engagement in the Global Atmosphere Watch programme.
References


http://www.atmos-chem-phys-discuss.net/8/3313/2008/


http://www.atmos-chem-phys.net/4/497/2004/


Tanimoto, H., Sawa, Y., Yonemura, S., Yumimoto, K., Matsueda, H., Uno, I., Hayasaka, T., Mukai, H., Tohjima, Y., Tsuboi, K., and Zhang, L.: Diagnosing recent CO emission sources and ozone...
C. Zellweger et al.: Carbon monoxide trend at Jungfraujoch

3503


