Comparison of the regional CO\textsubscript{2} mole fraction filtering approaches at a WMO/GAW regional station in China

S. X. Fang\textsuperscript{1}, P. P. Tans\textsuperscript{2}, M. Steinbacher\textsuperscript{3}, L. X. Zhou\textsuperscript{4}, and T. Luan\textsuperscript{4}

\textsuperscript{1}Meteorological Observation Centre (MOC), China Meteorological Administration (CMA), Beijing, China
\textsuperscript{2}Earth System Research Laboratory (ESRL), National Oceanic and Atmospheric Administration (NOAA), Boulder, CO, USA
\textsuperscript{3}Empa, Swiss Federal Laboratories for Materials Science and Technology, Laboratory for Air Pollution/Environmental Technology, Duebendorf, Switzerland
\textsuperscript{4}Chinese Academy of Meteorological Sciences (CAMS), China Meteorological Administration (CMA), Beijing, China

Correspondence to: S. X. Fang (fangsx@cams.cma.gov.cn)

Received: 28 May 2015 – Published in Atmos. Meas. Tech. Discuss.: 10 July 2015
Revised: 26 November 2015 – Accepted: 5 December 2015 – Published: 18 December 2015

Abstract. The identification of atmospheric CO\textsubscript{2} observation data which are minimally influenced by very local emissions/removals is essential for trend analysis, for the estimation of regional sources and sinks, and for the modeling of long-range transport of CO\textsubscript{2}. In this study, four approaches are used to filter the atmospheric CO\textsubscript{2} observation records from 2009 to 2011 at one World Meteorological Organization/Global Atmosphere Watch (WMO/GAW) regional station (Lin’an, LAN) in China. The methods are based on the concentration of atmospheric black carbon (BC), on a statistical approach (robust extraction of baseline signal, REBS), on CH\textsubscript{4} as an auxiliary tracer (AUX), and on meteorological parameters (MET). All approaches do suitably well to capture the seasonal CO\textsubscript{2} cycle at LAN. Differences are observed in the average regional mole fractions with annual values in the REBS method at least 1.7 ± 0.2 ppm higher than the other methods. The BC method may underestimate the regional CO\textsubscript{2} mole fractions during the winter–spring period and should be treated with caution. The REBS method is a purely statistical method and it may also introduce errors on the regional CO\textsubscript{2} mole fraction evaluations, as the filtered trend may be influenced by the “noisy” raw data series. Although there are correlations between CH\textsubscript{4} and CO\textsubscript{2} mole fractions at LAN, the different source/sink regimes may introduce bias on the regional CO\textsubscript{2} estimation in the AUX method, typically in summer. Overall, the MET method seems to be the most favorable because it mainly focuses on the influence of potential local sources and sinks, and considers diurnal variations and meteorological conditions. Using the MET method, the annual growth rate of regional CO\textsubscript{2} at LAN is determined to be 3.1 ± 0.01 ppm yr\textsuperscript{-1} (standard error) from 2009 to 2011.

1 Introduction

Carbon dioxide (CO\textsubscript{2}) is the most important greenhouse gas in the atmosphere. It contributes more than 60 % of total radiative forcing (RF) of the long-lived greenhouse gases (AGGI, 2014). The large increase of atmospheric CO\textsubscript{2} of nearly 120 ppm above preindustrial levels has been unequivocally attributed to human-caused emissions (Keeling, 1993; WMO, 2015). Using atmospheric CO\textsubscript{2} observations, the source/sink estimations can be constrained through inverse models, which is an important way to understand the carbon cycle in the land biosphere (Chevallier et al., 2011; Thompson et al., 2009). For this purpose, lots of ground-based stations have been set up to monitor the CO\textsubscript{2} mole fractions around the world. So far, there are more than 150 sites worldwide where greenhouse gas mole fractions are measured (Artuso et al., 2009; Dlugokencky et al., 1995; Necki et al., 2003; Sirignano et al., 2010; Tans et al., 1990; WMO, 2015). Due to technical and logistical constraints like access to the measurement site, power supply or internet connection, very few monitoring stations are sufficiently remote to be permanently exposed to pristine air masses, while many Global Atmosphere Watch (GAW) stations are occasionally too frequently affected by local sources or sinks (Tsutsumi...
Several methods have been applied in the past for the extraction of background mole fractions (or regionally representative) from ground-based measurements: (1) filters based on specific trace gases or ratios of trace gas:trace gas (Brunke et al., 2004; Tsutsumi et al., 2006; Zanis et al., 2007). For example, Tsutsumi et al. (2006) used carbon monoxide (CO) as an indicator to filter the observed CO$_2$ mole fractions at the station Yonagunijima located in east Asia. Zanis et al. (2007) used the total reactive nitrogen (NO$_x$) to CO ratio to distinguish different regimes at the high-altitude station Jungfraujoch in central Europe. Brunke et al. (2004) used radon and CO to classify their observations at Cape Point, South Africa. (2) Meteorological filters are the most commonly used method (Artuso et al., 2009; Chmura et al., 2008; Collaud Coen et al., 2011; Zellweger et al., 2003). This method considers various factors such as the local wind speed, wind direction, boundary layer height, information on the atmospheric stability, solar input, or general weather conditions and others. Sometimes it further considers the diurnal CO$_2$ variation as it can be closely linked to the above parameters (Zhou et al., 2005). (3) Statistical methods is an approach that generally uses the variations (e.g., a low standard deviation) of observed data in certain time windows as a threshold to select the regional values (Cunnold et al., 2002; Morimoto et al., 2003; Pickers and Manning, 2015; Zhang et al., 2007). (4) Numerical transport methods use atmospheric dispersion modeling (e.g., air mass back trajectories) to study the advection regimes with subsequent distinction in periods with potential influence of local or regional source/sink and uninfluenced conditions (Cape et al., 2000; Manning et al., 2011; Ryall et al., 2001). Some of the studies combined two or more methods above to select the best well-mixed CO$_2$ mole fractions (e.g., Thoning et al., 1989). However, due to the different characteristics of each station such as location (e.g., continental sites, coastal sites), proximity to biosphere and diffuse or point sources etc., the best data filtering approach has to be carefully selected for each station (Ruckstuhl et al., 2012). One method can be also more useful than another at the same station depending on the parameter of interest. In brief, there is not a standard method for selecting the background mole fractions from a continuous data series.

With the rapid development of its economy, China became the largest fossil fuel CO$_2$ emitter in 2006 and emitted 1.8 Pg carbon in 2011 (LeQuéré et al., 2013; Marland, 2012). The Yangtze Delta area is one of the most developed regions in China and is one of the largest global CO$_2$ emission regions (Gregg et al., 2008). The total population in this area was ~159 million in 2010 (National Bureau of Statistics, 2011). Moreover, this area is a highly productive region for paddy rice and winter wheat in China, with a potentially large influence on the atmospheric CO$_2$ concentrations in this region. For example, winter wheat and rice production in this region represent 20 % of the total Chinese wheat harvest and 5 % of the entire Chinese grain production (Colby et al., 1992; Yan et al., 2003). To understand the character and the abundance of greenhouse gases in this region, the Chinese Meteorological Administration (CMA) built the Lin’an (LAN) station in the center of the Yangtze Delta area in 1983. The station has been included in the World Meteorological Organization/Global Atmosphere Watch (WMO/GAW) as a regional station and is named after the small town of Lin’an, which is approximately 6 km southwest of the station. There was no in situ CO$_2$ measuring system until January 2009, when a cavity ring-down spectrometer (G1301, Picarro Inc.) was installed to continuously monitor atmospheric CO$_2$ and CH$_4$ mole fractions. Since then we have been acquiring the first-hand greenhouse data at this station (Fang et al., 2013; Pu et al., 2014).

We have found that the CO$_2$ mole fractions at LAN were the highest out of the four WMO/GAW stations in China (Liu et al., 2009; Fang et al., 2011). Based on the meteorological methods, we filtered the CO$_2$ records from 2009 to 2011 and estimated that 16.6 % of the data are likely to be regionally representative (Fang et al., 2014). However, in a different study, Pu et al. (2014) used black carbon (BC) as a chemical tracer to identify the influence of anthropogenic emissions and found that 27.3 % of the data were regionally representative. In a previous study (Fang et al., 2011), we also applied a purely statistical method to filter the data which resulted in regionally representative conditions during 63.5 % of the time in 2009. The different data filtering approaches have produced different results in terms of the regional CO$_2$ mole fractions. Our previous study (Fang et al., 2014) obtained an average regional CO$_2$ mole fraction of 404.2 ± 3.9 ppm in 2011 at LAN by using a meteorological filter, while Pu et al. (2014) estimated a corresponding average value of 407 ± 5.3 ppm during the same period by using a different filtering approach. The difference (~3 ppm) between these two methods can induce biases in the estimation of CO$_2$ abundances at the regional scale as well as in the calculation of sources/sinks by inverse models. In this paper, we applied four approaches to filter the observed data from 2009 to 2011 at LAN station, and studied the applicability of them. The four methods use black carbon as a tracer, a statistical method, methane (CH$_4$) as a tracer, and meteorological parameters.
2 Experiment

2.1 Measurement system

The LAN station (119° 44′ E, 30° 18′ N; 138.6 m a.s.l.) is about 50 km from Hangzhou (capital of the province of Zhejiang) and 150 km from Shanghai (the largest economic center and the second largest population city in China) (Fig. 1). North of the station (1.4 km away) is a small factory where charcoal is manufactured from bamboo wood. The town of Lin’an (with a population of ~ 100,000) is approximately 6 km southwest of the station. The observatory is built on the top of a small hill (160 m a.s.l.) and is surrounded with hilly lands and farming areas, with dense vegetation coverage. The site is located in a humid subtropical monsoon climate zone, with a mean annual precipitation of 1480 mm and a mean temperature of 15.3°C.

A cavity ring-down spectrometer (CRDS; Picarro Inc., model G1301) is used for continuous measurements of atmospheric CO₂ and CH₄. This type of instrument has been proven suitable for making precise measurements of CO₂ and CH₄ mole fractions, since its response is both highly linear and very stable (Chen et al., 2010; Crosson, 2008). The factory reported that the precision of the instrument is 50 ppb for CO₂ and 0.7 ppb for CH₄ (1 standard deviation, 1σ) in 5 min. Sample air is drawn regularly from about 10 m a.g.l. (above ground level). At the end of 2010, a new sampling tower (50 m a.g.l.) was built and another sampling port was installed at 50 m. The Picarro system then switched the air sample stream between the 10 m and the 50 m intake every 5 min. The sample air is filtered and temperature stabilized and dried to meet the high-quality target of the WMO/GAW network. Details of the system are described in Fang et al. (2013). Two standard gases are used to calibrate the measurements and a target gas is used to check the precision of the system routinely. All of the standards are linked to the WMO X2007 scale (Zhao and Tans, 2006). The CRDS system responds quickly to the sample and reports data with a frequency of 0.3 Hz. For the long-term time series, ambient air data are recorded as 5 min averages. Excluding the periods of system maintenance and calibration, more than 97% of the total 5 min average data points were retained. After computing the CO₂ mole fractions, the data were manually inspected to flag any analytical or sampling problems. More than 97% of the 5 min data remained after this filtering step. Then the data were aggregated to hourly averages for further study. Except when noted differently, the averaged values in this study are reported with 95% confidence intervals (CI). The CO₂ concentrations are atmospheric CO₂ dry air mole fractions. The black carbon particles in this study were measured by an aethalometer (AE-31, American Magee Scientific, USA). Seven wavelengths are used for the observations, with values of 370, 470, 520, 660, 880, and 950 nm.

2.2 Data filtering approach

There are significant CO₂ diurnal variations at LAN in all seasons, indicating the strong influence of biological activity near the site including absorption by photosynthesis and emissions by plants and soil, and the variation of boundary layer height. The CO₂ differences between the parallel measurements (10 and 50 m a.g.l.) at LAN also showed distinct diurnal variations with the most stable and lowest differences (less than 0.2 ± 0.2 ppm) occurring from 10:00 to 16:00 (local time, LT). At remote sites where the local sources and sinks are negligible, the CO₂ diurnal variations are generally weak (e.g., Keeling et al., 1976; Zhou et al., 2005). Thus, in this study, the CO₂ data from 10:00 to 16:00 LT were first selected to represent better mixed conditions. Then the selected data were used for the study of different filtering approaches. We adopted four data selection methods to filter the selected CO₂ mole fractions. They are described as follows.

- Black carbon (BC) tracer: we adopt a similar routine as used by Pu et al. (2014). The observed CO₂ was filtered based on the observation of both black carbon concentration and meteorological parameters. Pu et al. (2014) found a correlation coefficient of 0.53 (R) between black carbon concentrations and CO₂ mole fractions and concluded that both CO₂ and black carbon have some common sources such as fossil fuel combustion and biomass burning within this area. Thus, first
we excluded the episodes when the black carbon concentration exceeded 5000 ng m\(^{-3}\). During the wet season precipitations when the black carbon concentration was very low, we used air mass back trajectory analysis to further flag the data which were likely influenced by anthropogenic emissions from cities nearby. Finally, we studied the average standard deviations (\(\sigma\)) of hourly CO\(_2\) mole fractions as a function of wind speed using all data from 10:00 to 16:00 LT from 2009 to 2011. As shown in Fig. 2, the average \(\sigma\) decreased sharply when local surface wind speed was faster than 1.5 m s\(^{-1}\). Clearly, higher local surface wind caused better mixed conditions and consequently more stable CO\(_2\) mole fractions. Thus the remaining data were further flagged when surface wind was below 1.5 m s\(^{-1}\) to minimize the influence of very local sources and sinks.

- Statistical method (REBS): here we applied the robust extraction of baseline signal (REBS) to extract the regional CO\(_2\) mole fractions, which was similar to those used in the Global Atmospheric Gases Experiment/Advanced Global Atmospheric Gases Experiment (GAGE/AGAGE) network (Ruckstuhl et al., 2001) to filter halocarbons and other non-CO\(_2\) gases. Ruckstuhl et al. (2012) suggested that a meteorological filtering of the data should be applied prior to the application of the REBS method, as the polluted conditions might induce a bias on the background classification. In this study, this was taken into consideration by using data from 10:00 to 16:00 LT. The REBS method is a purely non-parametric technique and assumes that the background signal varies very slowly relative to contributions of the regional signal. The observed concentrations \(Y(t_i)\) are defined by a regional concentration \(g(t_i)\), plus a polluted concentration \(m(t_i)\), plus the measurement errors \(E_i\). The measurement errors \(E_i\) are assumed to be independent and Gaussian-distributed with mean 0 and variance \(\sigma^2\). If the regional signal \(m(t_i)\) is zero in a time period around \(t_0\), the baseline signal \(g(t_0)\) can be estimated, even when the form of the curve \(g\) is unknown. Hence the curve \(g(t_i)\) is approximated as being linear in a sufficiently small neighborhood around any given time point \(t_0\). Details of the method are described by Ruckstuhl et al. (2012). A bandwidth of 60 days was used in this study, while other bandwidths of 90, 120, and 180 days were also tested. The bandwidth choice did not considerably influence the retrieved averages, and trends of the regionally representative CO\(_2\) mole fractions were similar. In comparison with other methods, this approach did not have to be considerably adapted to the conditions at the individual measurement site.

- Auxiliary (AUX) tracer: this tracer uses CH\(_4\) as an auxiliary indicator to filter the CO\(_2\) time series. Many previous studies found positive correlations (mostly in winter) between the atmospheric CH\(_4\) and CO\(_2\) mole fractions (Conway et al., 1989; Tohjima et al., 2014; Wong et al., 2015; Worthy et al., 2009) as well as the respective fluxes from ecosystems (Jamali et al., 2013; Repo et al., 2007). For the data series of CO\(_2\) and CH\(_4\) at LAN, we also observed an apparent correlation between them during the observing period (Fig. 3). The correlation coefficient \((R)\) is higher than 0.5 for all seasons, which indicates that there are similar patterns of CO\(_2\) and CH\(_4\) sources. This phenomenon is more distinct in spring \((R = 0.7)\) and winter \((R = 0.8)\) when photosynthetic activity of the vegetation, i.e., the CO\(_2\) uptake, is
weak. In summer and autumn, the active absorption of \( \text{CO}_2 \) by terrestrial ecosystems may partly alter the \( \text{CO}_2-\text{CH}_4 \) correlation. Indeed, the positive coefficients still suggest that the anthropogenic emissions dominate the carbon cycle at the LAN station. In remote areas, an uncorrelated or negatively correlated relationship is generally observed (e.g., Necki et al., 2003). As described above, we also used the robust extraction of baseline signal (REBS) method to filter the \( \text{CH}_4 \) data because it has proven to be suitable for extracting the background mole fractions of \( \text{CH}_4 \) at remote sites (Cunnold et al., 2002). By doing so, we flagged the hourly \( \text{CO}_2 \) records which correspond to the periods of locally influenced or regionally representative events of \( \text{CH}_4 \) filtered by the REBS method. Moreover, although there were correlations between the atmospheric \( \text{CH}_4 \) and \( \text{CO}_2 \) in all seasons at LAN, they were not perfectly correlated, meaning that some \( \text{CO}_2 \) events could not be determined by \( \text{CH}_4 \) mole fractions, especially for data points which were far away from the linear fit in Fig. 3. To reduce this influence, we calculated the standard deviations (1\( \sigma \)) of the differences between measured \( \text{CO}_2 \) mole fraction and linear fit in each season. If the absolute difference was larger than 1\( \sigma \), the data point was considered as locally influenced and was flagged as an outlier. This additional filter excluded most events with poor \( \text{CH}_4-\text{CO}_2 \) correlation (Fig. 3).

Meteorological (MET) method: as used in previous studies, the diurnal variation of \( \text{CO}_2 \) mole fractions, local surface wind direction, local surface wind speed, and information on nearby sources were all considered for the \( \text{CO}_2 \) data filtering (Fang et al., 2014; Zhou et al., 2004, 2005). According to the nearby potential contamination sources (nearby villages, industry etc.), the data when local surface winds were from SSW–SW and N sectors were excluded. Then the data were further flagged by discarding the events when the local surface wind speed was lower than 1.5 m s\(^{-1}\) to minimize the influence of very local sources or sinks as discussed above.

After the multiple steps filter, there were still a few discrete data points remaining with high/low \( \text{CO}_2 \) mole fractions in the BC, AUX, and MET methods. These odd outliers unlikely represented regional \( \text{CO}_2 \) conditions as they should not peak within a few hours. Thus we used a mathematical method to further flag the remainder of data in the BC, AUX, and MET methods. The standard deviation (\( \sigma \)) of hourly \( \text{CO}_2 \) data in a 60-days bandwidth (similar to the REBS method) was calculated. The differences between every data point and the 60-day average were calculated. Data were flagged and excluded if the difference exceeded 3\( \sigma \). After that, the remaining data in the BC, AUX, and MET were considered as the least influenced by local sources or sinks.

3 Results

3.1 Filtered regional \( \text{CO}_2 \) mole fractions

Figure 4 illustrates the filtered \( \text{CO}_2 \) results in the four approaches. From top to bottom are the results of the BC, REBS, AUX, and MET methods, respectively. The filtered regional mole fractions account for \( \sim 12.2 \% \) in BC, \( 15 \% \) in REBS, \( 12.8 \% \) in AUX, and \( 16.5 \% \) in MET of the total valid hourly data. The low proportions of regional \( \text{CO}_2 \) in the four methods reflect the strong influences of local sources and sinks. The overall seasonal patterns of regional \( \text{CO}_2 \) retrieved in the four approaches are similar with peaks in winter and troughs in summer, which is consistent with other observations in the Northern Hemisphere (Nevison et al., 2008).

Table 1 compares the annual regional \( \text{CO}_2 \) mole fractions in the four methods. The annual mole fractions in the REBS method are apparently higher than the others with the smallest discrepancy of \( 1.7 \pm 0.2 \text{ ppm} \) in 2010. As the REBS
method uses a purely mathematical method, the high CO₂ mole fractions induced by local sources may enhance the σ values and enlarge the regional CO₂ band, and subsequently introduce higher regional CO₂ values. The annual mean CO₂ mole fractions display increasing trends. Although the annual CO₂ determined by the BC method increases from 402.0 ± 0.1 in 2010 to 402.7 ± 0.2 ppm in 2011, the absolute increase (∼0.7 ppm) is too small compared to increases from the other methods, as well as the global means. For example, the global average CO₂ increases were 2.3 ppm in 2009–2010 and 2.0 ppm in 2010–2011 based on WMO/GAW’s statistics (WMO, 2011; 2012) and were 2.39 ppm in 2009–2010 and 1.71 ppm in 2010–2011 based on observations from NOAA’s network (Dlugokencky and Tans, 2015). Considering the global CO₂ growth rate and the increasing emissions of CO₂ in China (CDIAC, 2015; Tohjima et al., 2014), it is unlikely that the regional CO₂ mole fraction at LAN almost remained constant from 2010 to 2011. In fact, the long-term trend in the BC method (polynomial part of the curve-fitting function) decreases after February in 2011 until end of the year (data not shown), which is opposite to the variations of total CO₂ emissions (or black carbon) expected from fossil fuel emissions in China with increasing value from 0.15 Pg C in 2009–2010 to 0.21 Pg C in 2010–2011 (CDIAC, 2015). As the BC method uses a fixed black carbon concentration (5000 ng m⁻³ in this study) as a threshold to filter the CO₂ record, a large proportion of high regional CO₂ mole fractions in 2011 may be flagged as local events, and consequently, a decreasing long-term trend was acquired. The absolute CO₂ increases from the other three methods indicate smaller increases from the 2009 to 2010 period and larger increases from 2010 to 2011, which better corresponds with the trend in total CO₂ emissions in China.

3.2 Mean seasonal cycles and trends

We used the curve-fitting method described by Thoning et al. (1989) to derive the seasonal CO₂ trends based on the filtered regional mole fractions from 2009 to 2011 (Figs. 4 and 5). The data were fitted into a function with three polynomial terms for the long-term trend and four annual harmonic terms. To minimize the influence of inconsistent records in the four methods, the data were fitted and interpolated with the same time interval (2 h). For comparison, surface CO₂ mole fractions at similar latitudes to the marine boundary layer (MBL) reference computed by NOAA/GMD

<table>
<thead>
<tr>
<th>Year</th>
<th>BC (ppm)</th>
<th>REBS (ppm)</th>
<th>AUX (ppm)</th>
<th>MET (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>398.9 ± 0.1</td>
<td>401.5 ± 0.2</td>
<td>398.6 ± 0.2</td>
<td>398.6 ± 0.2</td>
</tr>
<tr>
<td>2010</td>
<td>402.0 ± 0.1</td>
<td>403.7 ± 0.2</td>
<td>400.8 ± 0.2</td>
<td>401.7 ± 0.2</td>
</tr>
<tr>
<td>2011</td>
<td>402.7 ± 0.2</td>
<td>407.4 ± 0.2</td>
<td>405.1 ± 0.2</td>
<td>404.9 ± 0.2</td>
</tr>
</tbody>
</table>

Figure 5. (a) Variations of monthly CO₂ mole fractions in the four methods, also compared to the surface values at similar latitudes (30° N) from the MBL reference (MBL, 2015). The data in this figure are values smoothed by the curve-fitting method of Thoning et al. (1989). (b) The detrended seasonal CO₂ cycles in the four methods. This is the smoothed curve minus the trend. Error bars indicate confidence intervals of 95 %.
concentrations. Emissions of black carbon and CO$_2$ from fossil fuel and biomass burning occur at both the local and regional scale (Baumgardner et al., 2002), and the BC method robustly flags the CO$_2$ mole fractions when black carbon concentrations exceed the threshold value (5000 ng m$^{-3}$). However, it is difficult to distinguish the local emissions of black carbon from the regional contents. Especially during the winter–spring seasons, the regional black carbon and CO$_2$ concentrations are both high due to the increase of fossil fuel consumption and cement burning (Feng et al., 2014). These high concentrations should still represent the volumes at regional scale and should not be flagged. Thus in the BC method, the flagging of higher CO$_2$ mole fractions as local representatives is probably the reason for the lower regional CO$_2$ values during the winter–spring period. It should also be mentioned that the use of a constant threshold in the BC method may lead to slight errors in the estimation of regional CO$_2$ mole fractions. Secondly, the BC method is mainly geared towards polluted air masses altered by anthropogenic sources, whereas the influence of the land biosphere during daytime remains unconsidered. Moreover, the large difference of the lifetime, which is 4–12 days for black carbon (Cape et al., 2012) and more than decades for CO$_2$ (Moore and Braswell, 1994), may also contribute to the bias of this method.

The monthly CO$_2$ mole fractions in the AUX method are lower than in the other methods in summer. During this time of the year, a large amount of CH$_4$ is emitted from wetlands (e.g., from rice paddy fields) on the eastern China Plain (Lu et al., 2000; Zhang et al., 2010). Thus, these high values should represent regional conditions rather than local events (Fang et al., 2013). On the other hand, the active processes of photosynthesis by local and regional vegetation in summer reduce the observed CO$_2$ mole fractions, especially by local vegetation, which has a strong negative influence on the CO$_2$ mole fractions during afternoon. This influence can also be seen from the frequently lower CO$_2$ mole fractions at 10 m a.g.l. than at 50 m a.g.l. in the daytime. These different source/sink regimes may cause some CO$_2$ mole fractions being flagged as regional even though they are actually influenced by the absorption of local vegetation, subsequently leading to lower regional values in summer.

Compared with the BC, REBS, and AUX methods, there is no apparent disadvantage of the MET method, which attempts to eliminate the influence of local sources and sinks using meteorological information. Figure 5b illustrates the detrended seasonal cycles of CO$_2$ in the four approaches. The peak to trough amplitudes of regional CO$_2$ are 14.4 ± 0.1, 18.6 ± 0.1, 22.7 ± 0.1, and 20.4 ± 0.1 ppm for the BC, REBS, AUX, and MET methods, respectively. The amplitude for the BC method is the lowest, which is ascribed to the lower CO$_2$ mole fractions during the winter–spring period, and the higher values than AUX and MET method in summer. The higher CO$_2$ mole fractions in the BC method in summer may be due to the lower black carbon concentrations (Feng et al., 2014). The highest CO$_2$ amplitude is observed in the AUX method, which is ascribed to the lowest CO$_2$ mole fractions in summer.

In addition to providing seasonal cycles, the method of Thoning et al. (1989) also provides an estimate of the trend over the full measurement period 2009–2011. The regional CO$_2$ mole fractions all show positive trends with annual growth rates of 1.8 ± 0.01 for BC, 2.8 ± 0.01 for REBS, 3.2 ± 0.01 for AUX, and 3.1 ± 0.01 ppm yr$^{-1}$ (standard error) for MET. According to the statistics from the WMO Greenhouse Gas Bulletins (2011, 2012), the global CO$_2$ growth rate exceeded 2 ppm yr$^{-1}$ from 2009 to 2012. The growth rate for the BC method is lower than the global average, which may partly be caused by the different CO$_2$ to black carbon ratios in the considered years. As the regional CO$_2$ value in the BC method is based upon the black carbon concentrations, the increasing fossil fuel standards (upgraded from Chinese national stage 3 to 4 since 2010) and exhaust efficiency may induce different ratios between CO$_2$ and black carbon concentrations and hence extract a smaller CO$_2$ growth rate. Similar to the annual CO$_2$ mole fractions (Table 1), the annual growth rates of the AUX and MET method are close. It should be mentioned that only 3 years of data are used to evaluate the annual CO$_2$ growth rate. The relatively short time series here may inevitably induce bias on the growth rate estimation, which needs to be treated with caution.

### 3.3 Comparison of local CO$_2$ events

The benefit of a successful extraction of regional values is twofold. The identified regional values can be used e.g., for the determination of regionally representative trends. In addition, the data considered to be locally influenced can be used to learn more about the sources and sinks in the vicinity of the station. Fig. 6a displays the seasonal variations of local CO$_2$ mole fractions in the four approaches. Due to the considerable proportion of local events below the regional band in REBS (blue dots Fig. 3), the local CO$_2$ mole fractions were separated into “REBS-P” events (in or above the regional band) and “REBS-D” events (below the regional band). The data were also fitted and smoothed by the method of Thoning et al. (1989). The local CO$_2$ events all reveal a broad spring maximum peaking in May, and a distinct winter maximum with the highest value in December. Minimum values are all observed in August. The peak in December and trough in August agrees with the seasonal pattern of the regional data. However, there is another distinct peak in May in all the approaches. Feng et al. (2014) investigated the black carbon measurements in Shanghai, China (150 km from LAN), and found three peaks in January–February, April–June, and November–December from 2010 to 2011. As the anthropogenic emissions of black carbon and CO$_2$ in the Yangtze area have a similar spatial distribution (Qin and Xie, 2012), the peaks in May and December are probably due
to enhanced anthropogenic emissions in these months. However, the peak in May is blurred due to the dampening effect caused by the CO$_2$ uptake with the onset of the growing season.

The local CO$_2$ mole fractions in REBS-P are always higher than for the other methods, as most of the events identified by REBS-P are observed during nighttime (00:00–08:00 LT) when local emissions are strong and the boundary layer becomes lower. Contrarily, the CO$_2$ mole fractions in REBS-D are mostly observed at midday (12:00–16:00 LT) and are apparently lower than the other methods, reflecting the strong absorption by local vegetation. Thus, it can be seen that at LAN station, the REBS method tends to define a median band of the CO$_2$ record as regional representative. Except from December to January, the local CO$_2$ mole fractions in the BC method are higher than those in the AUX and MET methods. As discussed above, the tendency of flagging higher CO$_2$ mole fractions as locally representative is probably the main reason for the higher CO$_2$ values. This result also indicates that the BC method induces bias on the local CO$_2$ estimations.

Meteorological data (such as surface wind direction and speed) could help to understand the greenhouse gas emissions and transport (Dlugokencky et al., 1995; Massen and Beck, 2011). Figure 6b shows the wind rose distribution patterns of local CO$_2$ mole fractions in the four methods. The distributions are similar with the highest CO$_2$ values in the SW–SSW sectors, including in REBS-P and REBS-D. This is due to the anthropogenic emissions from the town of Lin’an, located at approximately 6 km southwest of LAN (Fig. 1). The local CO$_2$ mole fractions in the WSW to SSW sectors in the BC method are apparently higher than in the other methods. It is also probably due to the tendency of flagging higher CO$_2$ mole fractions emitted from the town.

Figure 7. The filtered regional and local CO$_2$ mole fractions from 28 to 31 December 2010 (LT). The black dots represent the regional events and the gray dots denote the local events. The phase 1 and phase 2 represent periods from 06:00 to 16:00 LT in 29 and 30 December 2010, respectively.

3.4 Case analysis

To further investigate the difference of the four data filtering approaches, we used a period in winter as a case study. In winter, CO$_2$ mole fractions at LAN are pretty high due to strong emissions and weak absorption by the regional terrestrial ecosystems. Here we selected two time periods which were from 06:00 to 16:00 LT on 29 December 2010 (period 1) and 30 December 2010 (period 2), respectively (see Fig. 7). Period 1 features elevated CO$_2$ mole fractions, while period 2 reveals “normal” CO$_2$ values. We compute the 3-day back trajectories with 500 m a.g.l. for the period with elevated CO$_2$ (28 December 2010 19:00 to 29 December 2010 06:00) using the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) dispersion model (Draxler and Rolph, 2003). The model is based on NCEP/NCAR Re-
result, some low regional CO\textsubscript{2} mole fractions may frequently deviate the regional events from the real trend. As a consequence, it may also induce errors when evaluating the regional CO\textsubscript{2} mole fractions. In a previous study (Fang et al., 2011), we estimated an annual mole fraction growth rate and the local CO\textsubscript{2} values in the region, the filtered CO\textsubscript{2} mole fractions at LAN. Moreover, it mainly gears to the regional contents. Especially during rainy periods. In this study, the annual CO\textsubscript{2} mole fractions, the annual growth rate, and the local CO\textsubscript{2} values in the BC method are different from the other three methods. It should be mentioned that the BC method in Pu et al. (2014) is different from the one in the present study. Besides using different wind speeds as the filter criterion (2 m s\textsuperscript{-1}) and excluding outliers, they used all data including both from daytime and nighttime. The emissions from local vegetation and accumulation in the shallow boundary layer in the night definitely enhanced the filtered CO\textsubscript{2} mole fractions and induced higher annual values than those in our study.

The REBS method is based on a purely statistical method. This method is appealing as it requires no additional information (site specific criteria, additional observations). However, it may also induce errors when evaluating the regional CO\textsubscript{2} mole fractions, e.g., overestimating the regional values. For example, the different lifetime between atmospheric CO\textsubscript{2} and black carbon may also introduce errors on the estimation, especially during rainy periods. In this study, the annual CO\textsubscript{2} mole fraction, the annual growth rate, and the local CO\textsubscript{2} values in the BC method are different from the other three methods. It should be mentioned that the BC method in Pu et al. (2014) is different from the one in the present study. Besides using different wind speeds as the filter criterion (2 m s\textsuperscript{-1}) and excluding outliers, they used all data including both from daytime and nighttime. The emissions from local vegetation and accumulation in the shallow boundary layer in the night definitely enhanced the filtered CO\textsubscript{2} mole fractions and induced higher annual values than those in our study.

The BC method may be treated with caution, as it is difficult to distinguish the local emissions of black carbon from the regional contents. Especially during the winter–spring seasons, this method may underestimate the regional CO\textsubscript{2} mole fractions at LAN. Moreover, it mainly gears to the polluted air masses altered by anthropogenic sources, and does not consider the influence of the land biosphere. Additionally, the different lifetime between atmospheric CO\textsubscript{2} and black carbon may also introduce errors on the estimation, especially during rainy periods. In this study, the annual CO\textsubscript{2} mole fractions, the annual growth rate, and the local CO\textsubscript{2} values in the BC method are different from the other three methods. It should be mentioned that the BC method in Pu et al. (2014) is different from the one in the present study. Besides using different wind speeds as the filter criterion (2 m s\textsuperscript{-1}) and excluding outliers, they used all data including both from daytime and nighttime. The emissions from local vegetation and accumulation in the shallow boundary layer in the night definitely enhanced the filtered CO\textsubscript{2} mole fractions and induced higher annual values than those in our study.

The REBS method is based on a purely statistical method. This method is appealing as it requires no additional information (site specific criteria, additional observations). However, it may also induce errors when evaluating the regional CO\textsubscript{2} mole fractions, e.g., overestimating the regional values. In a previous study (Fang et al., 2011), we estimated an annual average of 405.3 ppm in 2009 by using REBS, which was apparently higher than the averages in this study (Table 1). Due to the “noisy” CO\textsubscript{2} mole fractions at the regional sites like LAN, the filtered regional trend may be drawn upward or pulled down from the real variation.

Although there are correlations between CH\textsubscript{4} and CO\textsubscript{2} at LAN, the different source/sink regimes may induce bias on the regional CO\textsubscript{2} estimation in the AUX method, typically in summer. Further, the atmospheric CH\textsubscript{4} and CO\textsubscript{2} at LAN

---

**Figure 8.** HYSPLIT 72 h back trajectories for every hour from 28 December 2010 19:00 (local time) to 30 December 2010 06:00.
are not perfectly correlated, meaning that some CO₂ events cannot be determined by the CH₄ mole fractions.

In comparison to these approaches, there are fewer disadvantages in the MET method for the data selection. As this method mainly focuses on the influence of potential local sources and sinks and considers diurnal variations and meteorological conditions, it is reasonable to identify the influence of local sources and sinks and it is suitable to be applied at other regional stations. We also studied the results at one other station in China (Longfengshan, 127°36′E, 44°44′N; 330.5 m a.s.l.) using different BC, AUX, and REBS, but a similar MET method. The results also indicated that the MET method was the most favorable (Fang et al., 2015). However, we have to mention that due to the intake height (only 10 m a.g.l.) in this study and the complex influence of the land biosphere, the data selected at LAN may not fully represent the volume at the regional scale. Although we selected the data from 10:00 to 16:00 LT when the boundary layer was the highest and the surface wind speed was faster than 1.5 m s⁻¹, the influence of local land biosphere could not be fully eliminated. This influence can be seen from the frequently lower CO₂ mole fraction at 10 m a.g.l. than 50 m a.g.l. during daytime in summer.

However, it cannot be concluded that the other three methods are not suitable for the CO₂ data selection. For example, the four methods were also applied to process the observed CO₂ record at Mt. Waliuguan station (100°54′E, 36°17′N; 3816 m a.s.l.) in China without excluding the nighttime data. As this station is a WMO/GAW global site and is located in a remote area, we found there were no distinct differences between the filtered CO₂ mole fractions, including the seasonal cycles, annual growth rates, and background mole fractions (data not shown); the results agreed very well with the background information in the Northern Hemisphere. In this study, we selected daytime data only for the four approaches to exclude the influences of very local sources/sinks (e.g., vegetation). However, for sites without strong local sources/sinks and indistinct diurnal CO₂ variations, the nighttime data may also represent the regional background and can be used. Atmospheric black carbon is mainly from fossil combustion and biomass burning. Thus the BC method may be applied at some remote sites to identify the anthropogenic influence on the observed CO₂ records. The theory of REBS assumes that the background signal varies very slowly relative to contributions of the regional signal. The results in this study prove that the REBS is not suitable for the CO₂ data filtering at regional stations like LAN. Instead, it may be applied at some remote stations like Mt. Waliuguan. The AUX can be applied at sites where the atmospheric CH₄ and CO₂ are subject to the same sources, and it can also be applied at some remote stations. Moreover, due to the different characteristics and source/sink regimes of various gas species, the suitability of a particular filtering method may even differ when looking at different trace gases at the same sampling site. This needs to be studied separately.

Acknowledgements. We express our great thanks to the staff at Lin’an station, who have contributed to the system installation and maintenance at the stations. This work is supported by the National Natural Science Foundation of China (no. 41375130, 41405129, 41175116). Martin Steinbacher acknowledges funding from QASAC Switzerland which is financially supported by MeteoSwiss and Empa.

Edited by: D. Brunner

References


