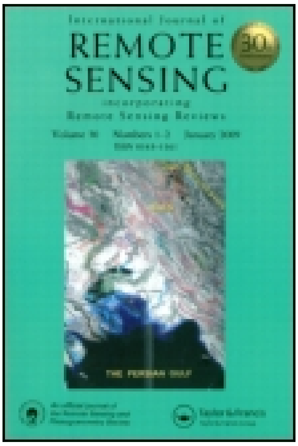
This article was downloaded by: [East China Normal University]
On: 30 October 2013, At: 22:25
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tres20

A correlation analysis of monthly mean CO_2 retrieved from the Atmospheric Infrared Sounder with surface station measurements

Cong Zhou[ab], Runhe Shi[ab], Chaoshun Liu[ab] & Wei Gao[abc]

[a] Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, Shanghai, China

[b] Joint Laboratory for Environmental Remote Sensing and Data Assimilation, ECNU and CEODE, Shanghai, China

[c] Department of Ecosystem Science and Sustainability, Colorado State University, Fort Collins, CO 80523, USA

To cite this article: Cong Zhou, Runhe Shi, Chaoshun Liu & Wei Gao (2013) A correlation analysis of monthly mean CO_2 retrieved from the Atmospheric Infrared Sounder with surface station measurements, International Journal of Remote Sensing, 34:24, 8710-8723

To link to this article: http://dx.doi.org/10.1080/01431161.2013.847295

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms &

International Journal of Remote Sensing, 2013
Vol. 34, No. 24, 8710–8723, http://dx.doi.org/10.1080/01431161.2013.847295

A correlation analysis of monthly mean CO_2 retrieved from the Atmospheric Infrared Sounder with surface station measurements

Cong Zhou[a,b], Runhe Shi[a,b]*, Chaoshun Liu[a,b], and Wei Gao[a,b,c]

[a]*Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, Shanghai, China;* [b]*Joint Laboratory for Environmental Remote Sensing and Data Assimilation, ECNU and CEODE, Shanghai, China;* [c]*Department of Ecosystem Science and Sustainability, Colorado State University, Fort Collins, CO 80523, USA*

(*Received 12 April 2013; accepted 11 September 2013*)

As one of the major greenhouse gases, atmospheric carbon dioxide (CO_2) concentrations have been monitored by both top-down satellite observations and air sampling systems on surface stations. The Atmospheric Infrared Sounder (AIRS) on board NASA's Aqua low Earth orbit (LEO) satellite is a high-resolution infrared sounder that has been in operation for more than 10 years. The World Data Centre for Greenhouse Gases (WDCGG) archives and provides data on CO_2 and other greenhouse gases measured mainly from surface stations. In this article, we focus on the correlation between the two different sources of CO_2 data and the influencing factors. In general, we find that a linear positive correlation occurs at most stations. However, the variation in the correlation coefficient is large, especially for stations in the Northern Hemisphere. The station's location, including its latitude, longitude, and altitude, is an important influencing factor because it determines how much its CO_2 measurements are influenced by human activities. We also use root mean square difference (RMSD) and bias as evaluation indicators and find that they have similar trends like correlation coefficients.

Introduction

Atmospheric carbon dioxide (CO_2) is a prominent greenhouse gas and a marker of climate change and the carbon cycle. CO_2 absorbs radiation in the infrared and near-infrared range and plays a significant role in the greenhouse effect. As a trace gas, the current concentration of CO_2 was only approximately 390 ppm (parts per million) in 2011, with an annual mean growth rate of 1.68 ppm year^{-1} (Dlugokencky and Tans 2013). However, the CO_2 concentration has been growing roughly exponentially with respect to its pre-industrial concentration of 280 ppm. The increased concentration of CO_2 in the atmosphere influences Earth's radiation balance and has long-term effects on climate change. Determining the sources and sinks of CO_2 quantitatively to understand its distribution and dynamic change is an urgent task for scientists in the face of severe climate consequences and risks. To support related scientific studies and provide reliable data to policy-makers to enable them to respond to climate change, it is of prime importance to gather and analyse CO_2 measurements and data from different sources.

Currently, there are two major means to obtain atmospheric CO_2 concentrations: satellites and surface stations. Satellites such as the Atmospheric Infrared Sounder (AIRS), Infrared Atmospheric Sounding Interferometer (IASI), Scanning Imaging

---

*Corresponding author. Email: rhshi@geo.ecnu.edu.cn

Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY), and Greenhouse Gases Observing Satellite (GOSAT) have played an important role in CO_2 remote sensing. These satellites can obtain the temporal and spatial variation of CO_2 in long time series and their trends at both global and regional scales (Crevoisier et al. 2009; Hammerling et al. 2012; Wang et al. 2011). The World Data Centre for Greenhouse Gases (WDCGG) archives and provides data on CO_2 and other greenhouse gases measured mainly from surface stations. These stations are under the direction of Global Atmosphere Watch (GAW) and other programmes and have been built throughout the world to acquire information about the variation in CO_2 (Conway et al. 1994; WMO GAW Report No. 161. 2005). Studies have been conducted to define the annual and seasonal changes of CO_2 using CO_2 measurements from stations (Worthy, Higuchi, and Chan 2003; Rutgersson, Norman, and Astrom 2009). One important result is that the fluctuation in CO_2 concentration exhibits long-range power-law correlations using monthly mean values of CO_2 concentration measured at Mauna Loa station over the period 1958–2004 (Varotsos, Assimakopoulos, and Efstathiou 2007). As a result, a correctly rescaled subset of the original time series of the CO_2 concentrations resembles the original time series.

AIRS is sensitive to CO_2 in the middle to upper troposphere (Yoshida et al. 2011). The CO_2 values measured by stations represent its concentrations at different altitudes. Data on the spatial distribution and temporal evolution of CO_2 are limited due to the rather sparse network of surface stations. Retrievals from the satellites have the potential to overcome these limitations (Buchwitz et al. 2005). Comparisons of the CO_2 concentrations retrieved by AIRS and those measured by stations have been carried out in many studies. The results of comparisons using aircraft measurements demonstrate the remarkable ability of AIRS to track seasonal and latitudinal variations of CO_2 in the middle to upper troposphere with an accuracy better than 2 ppmv (Olsen 2009). AIRS CO_2 products have also shown good agreement with five ground-based station observations (Bai, Zhang, and Zhang 2010). AIRS estimates and National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory/Global Monitoring Division (ESRL/GMD) aircraft measurements obtained during 2005 correlate very well (76% correlation), with an root mean square difference (RMSD) of 2.05 ppmv and a bias (AIRS–ESRL) of –1.03 ppmv (Maddy et al. 2008). On the one hand, however, different station measurements have diverse correlations with AIRS CO_2 retrievals, but there are no comprehensive comparisons between AIRS and surface stations at the global scale. On the other hand, the influencing factors also need to be analysed.

This article studies the correlation of CO_2 data from AIRS with those from surface stations. The influencing factors are also analysed to make better use of these two data sources.

CO_2 data sources

Satellite-derived data

The satellite data used in this article were retrieved from the AIRS instrument and downloaded via NASA's official AIRS CO_2 product site. AIRS has orbited the Earth on NASA's Aqua satellite in a Sun-synchronous near-polar orbit since 2002. The satellite is equipped with a cross-track scanning grating spectrometer covering a spectral range of 3.74 μm to 15.4 μm with 2378 channels. AIRS, for the first time, has allowed the retrieval of daily CO_2 concentrations globally, including over land, oceans, and polar regions during daytime and night-time and in the presence of clouds (Chahine et al. 2008). Level2 (L2) CO_2 products are retrieved by the Vanishing Partial Derivative (VPD) algorithm (Chahine et al. 2005). These products have a spatial resolution of 90 km × 90 km at nadir. Level3 (L3) CO_2 products are created by binning on 2.0° latitude by 2.5° longitude grids from L2 CO_2 standard products. The global spatial coverage of AIRS CO_2 data ranges from 90° N to 60° S. In this article, L3 calendar monthly data from September 2002 to January 2012 are used.

Station-measured data

The station measurements are from the World Meteorological Organization (WMO) WDCGG website. The WDCGG, first established in 1990, has been operating for more than 20 years and is one of the World Data Centres (WDCs) under the WMO GAW programme. The WDCGG gathers, archives, and provides data on greenhouse gases (e.g. CO_2, CH_4, CFCs, N_2O and surface ozone) and related gases (e.g. CO, NO_x, SO_2 and VOC) in the atmosphere and oceans, as observed under GAW and other programmes. The data vary due to different observation categories and sampling and data types. Monthly mean data measured by air sampling observations at stationary platforms from September 2002 to January 2012 are used in this article (see Figure 1 for station locations). There are a total of 132 CO_2 measurements at 114 stations. The sampling types of these data are ‘cn’ and ‘fl’, that is, continuous or quasi-continuous *in situ* measurements and analysis of air samples in flasks, respectively. More information is described in the WDCGG guide.

Methodology

Association of surface stations with AIRS CO_2 grid points

The spatial resolution of AIRS CO_2 retrievals is 2.0° latitude by 2.5° longitude, so the global continents are divided into 12,960 grids of the same size, and a station is associated with a CO_2 grid point. In a certain grid in which the station is located, the CO_2 data of this grid from AIRS and from the station were matched for analysis. A total of 106 grids were matched between AIRS and the stations. The sample numbers, which varied from 4 to 110,

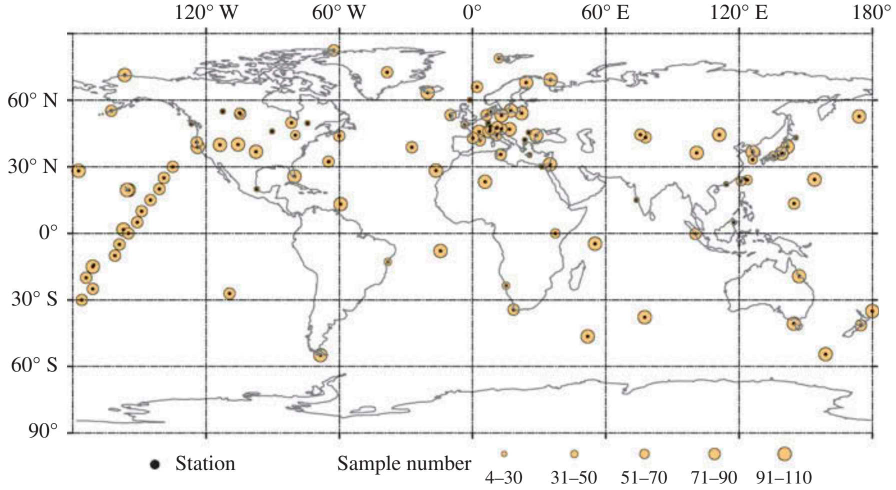


Figure 1. Geographical distribution of surface stations and sample numbers in associated grid cells.

differed for each grid cell due to variation in periods of CO_2 data measured by the stations (Figure 1). Of these grids, 83% had more than 50 sample numbers.

Evaluation indicators

The evaluation indicators used in this article are the correlation coefficient (r), the RMSD, and the mean difference (bias). The correlation coefficient is one of the most common and useful statistical methods for studying the relationship between two variables, and its results range from –1 to +1. Zero indicates that there is no relationship between the variables, while a negative correlation indicates that as one variable goes up, the other goes down. A positive correlation indicates that both variables move together in the same direction. To eliminate the influence of sample number, the significance of the correlation coefficient was tested by employing the t distribution (Zimmerman 1986). The correlation analysis was performed to define the relationship between the CO_2 data from AIRS and those from the surface stations. RMSD is an indicator that is frequently used to quantify the differences between values predicted by a model or estimator and the values actually measured. In this article, RMSD was used as an indicator of the difference in CO_2 concentrations between AIRS observations and station measurements. The bias defined in the study measured the inclination of the AIRS CO_2 value to be above or below a station's CO_2 value. Thus, if it has a positive bias, on average the AIRS observation exceeds the station-measured value. The formulas are described in Table 1.

Reprocessing

There were 119 CO_2 measurements at 104 stations that passed the significance test (with a significance level of 0.05) for the relationship between CO_2 concentrations from AIRS and those from the stations. Twelve stations had different data sources, so the similarities and differences were analysed further. Moreover, two stations had a slightly negative correlation with AIRS, so the time series of CO_2 at the two stations were compared with the AIRS CO_2 concentration.

Table 1. Formulas of the three indicators.

Indicator	Formula	Range
Correlation coefficient (r)	$r_{xy}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$	[–1, +1]
Root mean square difference (RMSD)	$\mathrm{RMSD}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-x_i)^2}$	[0, +∞)
Bias	$\mathrm{bias}(x)=\frac{1}{N}\sum_{i=1}^{N}(y_i-x_i)$	[0, +∞)

Notes: n is the sample number; x_i represents the value of x for the sample i, that is, the CO_2 data measured at surface stations; y_i represents the value of y for the sample i, that is, CO_2 retrieved from AIRS; $\bar{x}$ is the mean value for all x_i; and $\bar{y}$ is the mean value for all y_i.

Results and analysis

AIRS global CO_2 concentration

The distribution of global averaged CO_2 concentration, based on the AIRS 2003–2011 Level3 CO_2 monthly product, is shown in Figure 2. The white area of the map indicates that no data were retrieved by AIRS. It will be clearly observed that the CO_2 concentration is higher in the Northern Hemisphere than in the Southern Hemisphere. The highest CO_2 concentration belt is within the region between 30° N and 60° N, which may be strongly influenced by surface sources and the large-scale circulations of mid-latitude Northern Hemisphere pollution belts (Zhang et al. 2006). The lowest CO_2 concentration region is the 0–30° S belt of the Atlantic Ocean. The latitudinal gradient in the CO_2 concentration is a result of the larger amount of land in the Northern Hemisphere than in the Southern Hemisphere (Engelen and McNally 2005).

Evaluation indicator results

Figure 3 depicts the geographical distribution of the three evaluation indicators between AIRS and the stations' monthly mean CO_2 data. For most stations located in the region between 60° S and 30° N, the correlation coefficients are higher than 0.8 between their CO_2 measurements and those observations from AIRS. The bias and RMSD are smaller as well, with averages of 0.16 and 2.49 ppm, respectively. The strongest correlation is found at the Cape Point (CPT) station, where the coefficient is 0.988 with a bias of 0.7 ppm and an RMSD of 2.7 ppm. The CPT station is administered by the South Africa Weather Service (SWAS). Continuous measurements of CO_2 at CPT started in 1993 using the non-dispersive infrared (NDIR) absorption technique (Zellweger et al. 2011). The CPT station

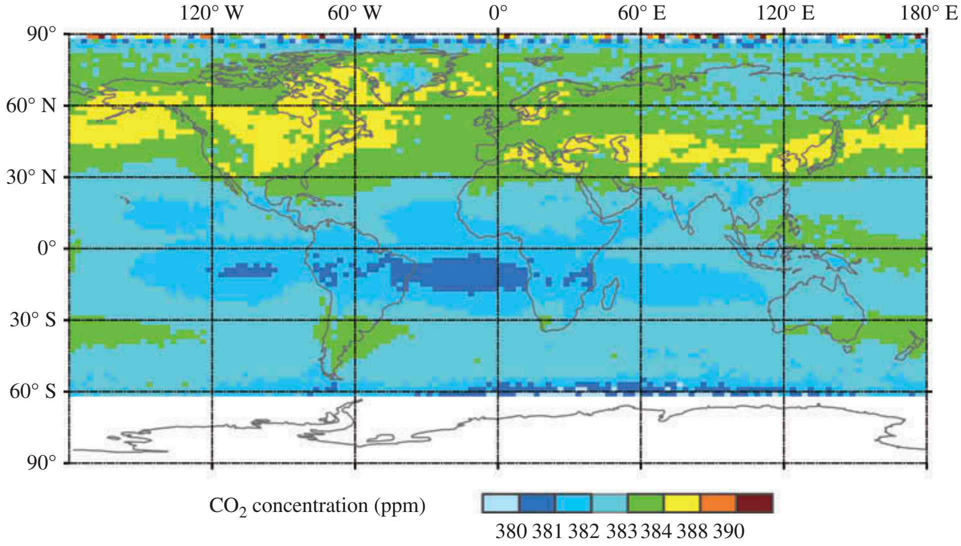


Figure 2. Global distribution of AIRS-averaged CO_2 concentrations from 2003 to 2011.

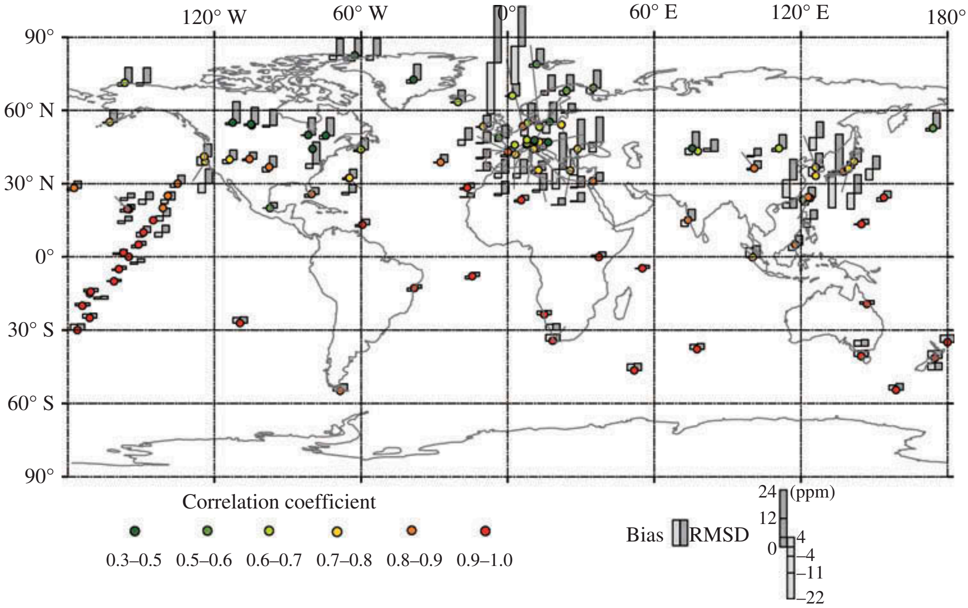


Figure 3. Evaluation indicators between AIRS and stations' monthly mean CO_2.

is located in a nature reserve at the southern end of the Cape Peninsula, which is approximately 60 km south of the city of Cape Town. The station is exposed to the sea on the top of a cliff that is 230 m high and is subjected to maritime air from the South Atlantic most of the time. The surrounding environment of the station is seldom influenced by human activities and land-use changes. Thus, CO_2 is conserved after its long-range transport from the near surface to the mid-troposphere because it is chemically stable, which may explain why the highest correlations of CO_2 observations between AIRS and the surface stations are found in maritime areas.

For the majority of the stations located within the range 30° N to 90° N, the coefficients vary from 0.2 to 0.8. The average bias is –2.65 ppm and the average RMSD is 6.57 ppm. On the one hand, there is a large amount of land in this region, and the terrestrial biosphere is one of the most important components of the Earth system that influences atmospheric CO_2 concentrations (Erickson et al. 1996). On the other hand, the increase in CO_2 is also partly affected by human activities. There are numerous stations in the regions of 10° W to 30° E and 30° N to 60° N and in Korea, as well as in Japan, regions where the population is relatively high (as of 5 July 2013, the website 'http://www.populationlabs.com/World_Population.asp'). With a growing population, more energy is being used for socioeconomic development and CO_2 emissions have been increasing as a result (Figure 4). These stations are likely to be influenced by human activity. Therefore, CO_2 measurements at these stations have distinct fluctuations compared with the mid-tropospheric CO_2 retrieved from AIRS.

The influencing factors

The first influencing factor is the latitude of the stations. The variations of the three indicators in five latitudinal gradients are displayed in Figure 5. The correlation

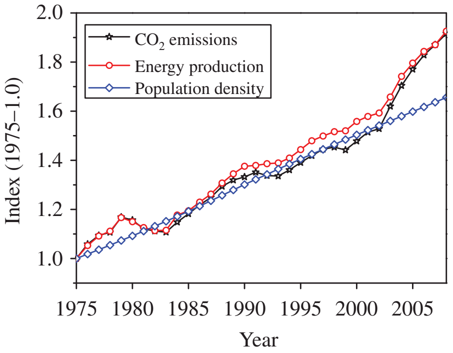


Figure 4. Growth of energy use, CO_2 emissions, and global population from 1975 to 2008; the index of the *y*-axis denotes the growth extent compared to 1975 (Data obtained from World Bank Development Indicators).

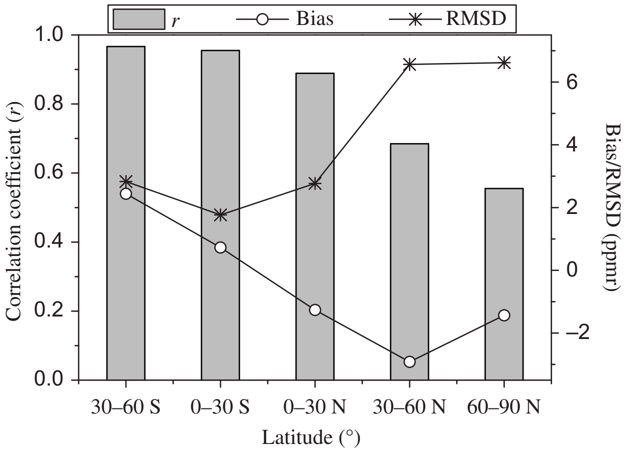


Figure 5. Variation in the three indicators in five latitudinal gradients.

coefficients continue decreasing with latitude from south to north. The average RMSD is lower in the Southern Hemisphere than in the Northern Hemisphere. It can be shown that the biases of the Northern Hemisphere are negative, which illustrates that CO_2 concentrations retrieved from AIRS are lower than the stations' CO_2 measurements. In the Southern Hemisphere, however, the phenomenon is precisely the opposite.

The second influencing factor is the altitude of the stations, which also influences the relationship of CO_2 concentrations between AIRS and the surface stations (Figures 6(*a*)–(*c*)). The correlation coefficients increase with an increase in station altitude, with decreasing RMSD and an absolute bias value. First, all stations of altitude higher than 2 km have high correlation coefficients and smaller biases and RMSDs. Then, among the 20 stations of altitude higher than 2 km, all of the high stations located between 30° N and 60° N have a bias of less than 1.3 ppm and an RMSD of less than 5 ppm. Stations far from the sea are also included (i.e. Niwot Ridge in the USA, Pic du Midi in France, and Mt Waliguan in China). Finally, stations of altitude near 0 km but with low bias and RMSD are all located near the sea.

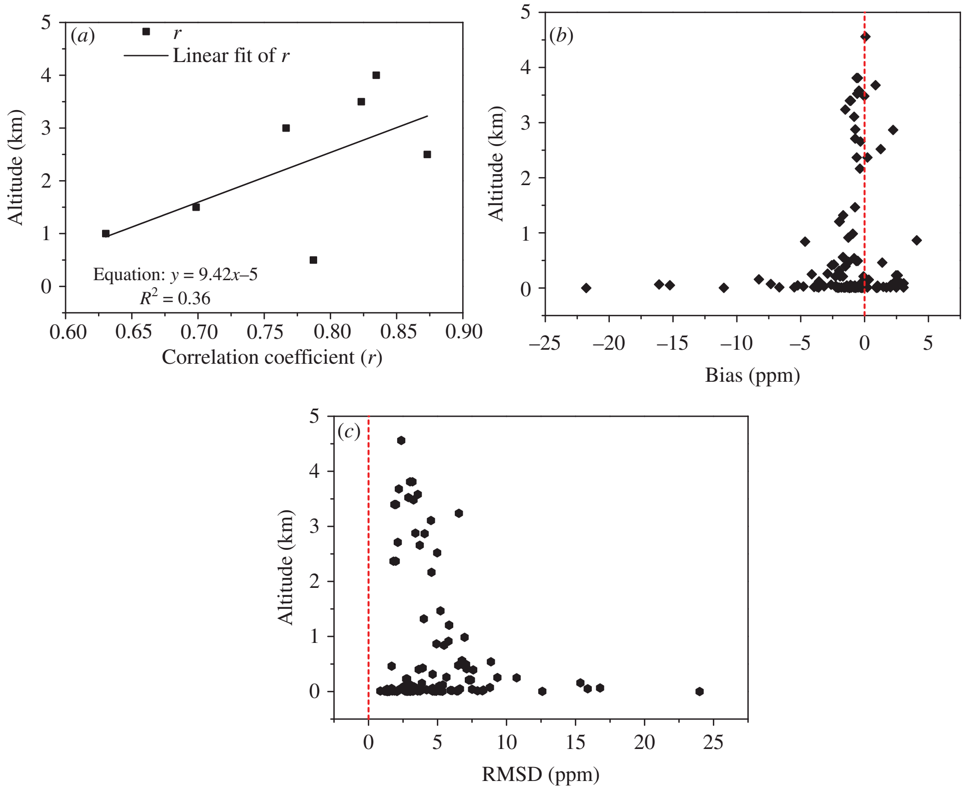


Figure 6. Variations in (*a*) correlation coefficient, (*b*) bias, and (*c*) RMSD with altitude.

Different sampling types or contributors

Some stations have different data sources (i.e. measurement techniques). Data sampling types are mainly divided into flask and continuous, and the instruments contributing the data are non-identical. The continuous method often samples data at an hourly frequency, and flask sampling measurements are taken at a frequency of approximately once per week. High-frequency data sampling can record the changes in CO_2 concentrations effectively, and continuous measurements are usually considered better for use under adequate background conditions (Tsutsumi et al. 2009). Figure 7 presents the correlation coefficient variability of these 12 stations with different sources. The CO_2 data taken at Lampedusa station, shown in red in the figure, exhibit the greatest difference.

Lampedusa station is located on a small island in the Mediterranean Sea, 120 km or more from larger islands or continents. The island is rocky and relatively flat and has very poor vegetation. The Italian National Agency for New Technology, Energy and the Environment (ENEA) and the NOAA/ESRL have presented CO_2 data to the WDCGG. The measurement method utilized by ENEA is NDIR, and the sampling type is flask. Air samples are collected every Friday. NOAA/ESRL uses the NDIR instrument to analyse data at a weekly sampling frequency.

The correlation coefficient between CO_2 data contributed by NOAA/ESRL at Lampedusa station and AIRS is 0.63, with a bias of 0.03 ppm and RMSD of 2.96 ppm. However, the correlation coefficient between CO_2 data contributed by ENEA at Lampedusa station and AIRS is 0.78, with a bias of –1.39 ppm and RMSD of 3.77 ppm.

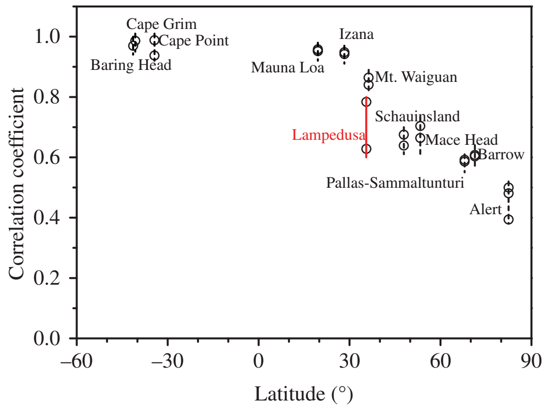


Figure 7. Correlation coefficients of 12 stations with different data sources.

The seasonal cycle is similar for CO_2 measured at Lampedusa station and that observed by AIRS satellite (Figure 8), that is, CO_2 has its maximum value in spring and minimum in autumn. The AIRS CO_2 data behave similarly to the CO_2 values from Lampedusa station, but are much more uniform and with only slight fluctuations. Various contributors use different sampling frequencies to measure the CO_2 data and are able to acquire coincident values. The most probable reason for the different correlations is the number of observation matches between AIRS and Lampedusa station (*N*). The correlation coefficient tells us about the strength of the linear relationship between two variables. However, the reliability of the linear model also depends on how many observed data points are in the sample (Bloom 2010). Therefore, when there are more data, the results should be reanalysed. Further surveys should be carried out to investigate this relationship of diverse sampling types or instrument contributors and their correspondence with AIRS CO_2 observations.

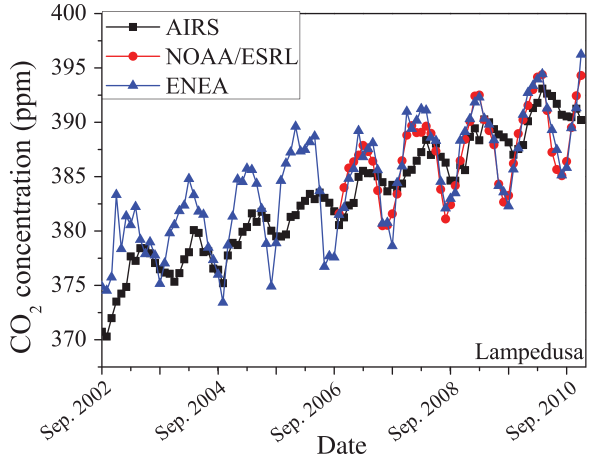


Figure 8. CO_2 concentration measurements at Lampedusa station with different contributors and CO_2 retrievals from AIRS.

Table 2. Station information, including two stations slightly negatively correlated with AIRS and the stations used for validation.

Station name	Country	Lat. (°)	Long. (°)	Alt (m)	Type	*r*	RMSD (ppm)	Bias (ppm)	*N*
BEO Moussala	Bulgaria	42.1792	23.5865	2925	Continuous	−0.14	56.28	47.03	30
Black Sea*	Romania	44.17	28.67	3	Flask	0.75	12.6	−11	100
Finokalia*	Greece	35.3378	25.6694	150	Flask	0.73	3.8	0.33	23
Hok Tsui	Hong Kong	22.2095	114.2579	60	Continuous	−0.06	9.8	−0.35	12
Yonagunijima*	Japan	24.47	123.02	30	Continuous	0.89	4.2	−3.17	112

Note: The symbol '*' indicates the station used for validation.

Two exceptional stations

There are two stations that have a slightly negative correlation with AIRS, as shown in Table 2. The Black Sea and Finokalia stations are used to verify the Basic Environmental Observatory (BEO) Moussala station, and the Yonagunijima station is used to verify the Hok Tsui station. The principle for selecting validation stations is the distance between two stations, with the nearest stations chosen because of their similar background environments.

The BEO Moussala (42° 10′ 45″ N; 23° 35′ 07″ E) is located at the top of the highest mountain on the Balkan Peninsula, Moussala (Angelov et al. 2011), which is 2925 m above sea level. This site is one of the best for environmental monitoring in the Balkan region due to its low anthropogenic influence. The station was accepted into the GAW regional station group in 2010 for two reasons: the acknowledged significance of its location in Eastern Europe and its high-quality measurements. The main feature of the BEO Moussala station is its complexity (Stamenov et al. 2007). The time series of CO_2 at BEO Moussala, the Black Sea, the Finokalia stations, and from AIRS are shown in Figure 9(*a*). The Finokalia CO_2 data are highly coincident with those from AIRS. The Black Sea CO_2 observations are mostly higher than those from AIRS, and its RMSD and bias are relatively

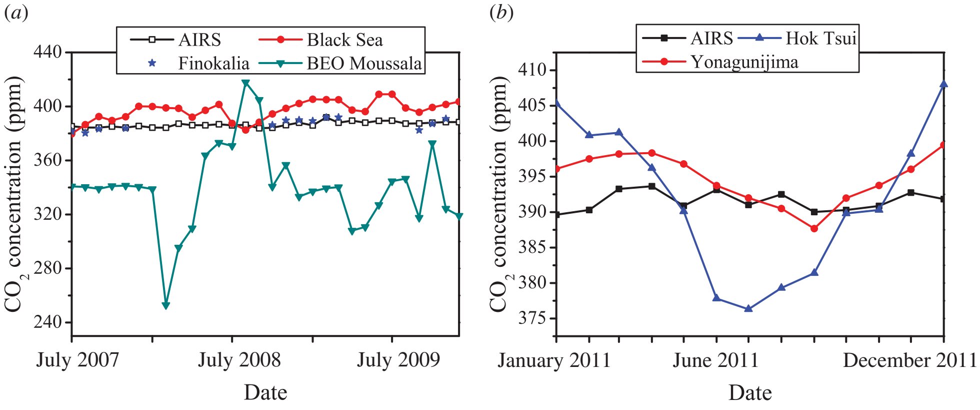


Figure 9. (*a*) Time series of CO_2 concentration measurements at three stations (BEO Moussala, Black Sea, and Finokalia) and corresponding observations from AIRS; (*b*) time series of CO_2 concentrations at two stations (Hok Tsui and Yonagunijima) and corresponding observations from AIRS.

high (Table 2). However, the CO_2 measurements at BEO Moussala fluctuated dramatically from July 2007 to December 2009, ranging from 252.98 to 417.89 ppm with an average standard deviation of 39.06 ppm. Thus, the CO_2 values from BEO Moussala should be used prudently.

Hok Tsui, with an elevation of 60 m above sea level, is a relatively remote coastal site located at the southeastern tip of Hong Kong Island. The site is situated on a cliff in a relatively clean area of Hong Kong. The urban areas of Hong Kong are approximately 10 km from the site and are normally downwind under the prevailing east–northeast flow in spring. The atmospheric background environment of the site could be affected by emissions from populated areas such as the Pearl River Delta due to its close proximity to urban centres. Emissions from ships in and around Hong Kong can also be an important source of CO_2 (Wang et al. 2003). The time series of CO_2 from the Hok Tsui and Yonagunijima stations and from AIRS are displayed in Figure 9(b). AIRS CO_2 concentration has a relatively minor fluctuation, from 390 to 395 ppm. The CO_2 measurements at the Yonagunijima station fluctuate slightly compared with AIRS. However, CO_2 values measured at the Hok Tsui station range from 375 to 410 ppm, with a mean standard deviation of 6.2 ppm. These values may be affected by emissions from the polluted areas mentioned above.

Station classification

Through the above analysis, the stations are classified into three grades: 'Class One', 'Class Two', and 'Class Three', under the following conditions. First, because the difference between the column-averaged CO_2 mixing ratio and the surface value varies from 2 to 10 ppm, depending on location and time of year (Olsen and Randerson 2004), the four stations with biases greater than 10 ppm in comparison with AIRS are rejected. The remaining stations are then classified according to the criterion shown in Table 3. Stations are classified as 'Class One' if the coefficients are equal to or greater than the average of the coefficients of all the stations (i.e. 0.78), and if the bias and RMSD are equal to or less than the average. 'Class One' comprises 29 measurements accounting for 29% of the total. Other stations whose coefficients are equal to or greater than 0.78, but with a larger bias and RMSD, are considered as 'Class Two', with 29 stations (29% of the total). The remaining 42 stations are classified as 'Class Three', these accounting for 42% of the total. The classification results are displayed on a map (Figure 10) according to their locations.

When performing research on CO_2 retrieval from the AIRS instrument, it is crucial to validate the retrieval results, and the 'Class One' stations can be used for validation and comparison. However, for CO_2 retrieval in a region without any 'Class One' station, CO_2 results from model predictions or from other satellites should be adopted to perform the test. Also, the reason for the large difference between surface and mid-tropospheric CO_2 should be researched further.

Table 3. The criterion for station classification.

Grade	Criterion	
Class One	$r \geq \bar{r}$	$\text{bias} \leq \overline{\text{bias}}$, $\text{RMSD} \leq \overline{\text{RMSD}}$
Class Two		$\text{bias} > \overline{\text{bias}}$, $\text{RMSD} > \overline{\text{RMSD}}$
Class Three	$r < \bar{r}$	

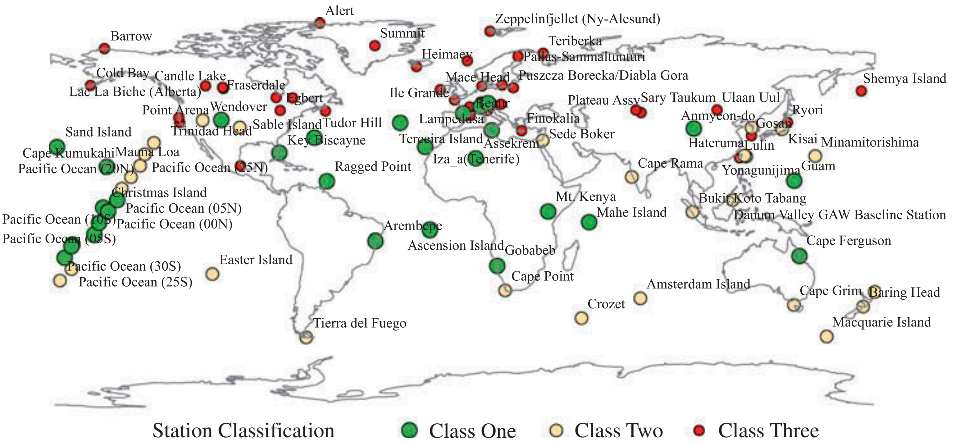


Figure 10. Classification results of stations according to their locations.

Conclusion

In this article, a correlation analysis from several aspects was performed on CO_2 values retrieved from AIRS and those from 132 surface station measurements. The CO_2 measurements from the stations located within the region between 60° S and 30° N were highly correlated with the AIRS CO_2 retrieval data, and the bias and RMSD between the two were small, at 0.16 and 2.49 ppm, respectively. The strongest correlation was found at the CPT station, whose surrounding environment is seldom influenced by human activities and land-use changes. The coefficient at the CPT station is 0.988, with a bias of 0.7 ppm and an RMSD of 2.7 ppm. However, for the stations located between 30° N and 90° N, the correlation coefficients dramatically change from 0.2 to 0.8. The average bias and RMSD are –2.65 and 6.57 ppm, respectively. The large mass of land within this latitude belt is one reason for this result. Second, the population of the region within 10°W to 30° E and 30° N to 60° N, as well as Korea and Japan, is relatively high, so these stations are influenced by human activities with more energy use emitting more CO_2.

Two influencing factors were analysed in this research. One is the latitude at which a station is located, with the coefficients decreasing from south to north. The average RMSD of sites in the Southern Hemisphere was less than that in the Northern Hemisphere. Another influencing factor is the altitude of the stations. With an increase in altitude, the correlation coefficients increase with decreasing RMSD and the absolute value of the bias. On the one hand, bias and RMSD between the CO_2 data from AIRS and sites higher than 2 km were lower, regardless of where the station was situated. On the other hand, stations of altitude near 0 km but with low bias and RMSD according to AIRS are all located near the sea.

Although the CO_2 data from the Lampedusa station were measured by two different contributors using different sampling strategies, concentrations were similar. The CO_2 measurements from the Lampedusa station display a similar seasonal cycle, reaching a maximum value in spring and minimum in autumn. The AIRS CO_2 retrieval is similar to that of the Lampedusa station, but it fluctuated slightly. When there are more data from different data sources, the results should be researched once more.

There were two exceptional stations with a slightly negative correlation with AIRS: the BEO Moussala station in Bulgaria and the Hok Tsui station in Hong Kong. The time series of CO_2 at the two stations fluctuate dramatically compared with AIRS and the surrounding stations, most likely due to the stations' own observation errors or the influence of polluted areas near the stations.

Finally, the stations were classified into three grades, 'Class One', 'Class Two', and 'Class Three', based on the three indicators. The proportions of each grade were 29%, 29%, and 42%, respectively. 'Class One' stations can be used for the validation of CO_2 retrieved by AIRS.

Acknowledgements

We would like to thank the WDCGG for the CO_2 data of the stations, the data contributors, and the GAW Station Information System (GAWSIS) website for providing information regarding the stations. We also thank NASA for providing the AIRS CO_2 product. Great thanks to Dr William Smith and Dr Hung-Lung Allen Huang for their suggestions about the modifications in this paper. Finally, special thanks to Professor Costas Varotsos and all the reviewers for their comments.

Funding

This project was supported by the National Basic Research Programme of China (No. 2010CB951603), the Shanghai Science and Technology Commission Programmes (No. 10DZ0581600 and 13231203804) and the National Science Foundation of China (No. 41201358). The computation was supported by the High Performance Computer Centre of East China Normal University.

References

Angelov, C., I. Angelov, T. Arsov, N. Archangelova, A. Boyukliiski, A. Damianova, M. Drenska, K. Georgiev, I. Kalapov, A. Nishev, N. Nikolova, I. Penev, I. Sivriev, J. Stamenov, A. Tchorbadjieff, S. Todorov, and B. Vachev. 2011. "BEO Moussala – A New Facility for Complex Environment Studies." In *Sustainable Development in Mountain Regions*, edited by G. Zhelezov, 123–139. Berlin: Springer.

Bai, W. G., X. Y. Zhang, and P. Zhang. 2010. "Temporal and Spatial Distribution of Tropospheric CO_2 over China Based on Satellite Observations." *Chinese Science Bulletin* 55: 3612–3618.

Bloom, R. 2010. *Linear Regression and Correlation: Testing the Significance of the Correlation Coefficient (Modified R. Bloom)*. Accessed July 5, 2013. http://cnx.org/content/m33270/1.2/content_info

Buchwitz, M., R. D. Beek, J. P. Burrows, H. Bovensmann, T. Warneke, J. Notholt, J. F. Meirink, A. P. H. Goede, P. Bergamaschi, S. KÖrner, M. Heimann, and A. Schulz. 2005. "Atmospheric Methane and Carbon Dioxide from SCIAMACHY Satellite Data: Initial Comparison with Chemistry and Transport Models." *Atmospheric Chemistry and Physics* 5: 941–962.

Chahine, M., C. Barnet, E. T. Olsen, L. Chen, and E. Maddy. 2005. "On the Determination of Atmospheric Minor Gases by the Method of Vanishing Partial Derivatives with Application to CO_2." *Geophysical Research Letters* 32: L22803.

Chahine, M. T., L. Chen, P. Dimotakis, X. Jiang, Q. B. Li, E. T. Olsen, T. Pagano, J. Randerson, and Y. L. Yung. 2008. "Satellite Remote Sounding of Mid-Tropospheric CO_2." *Geophysical Research Letters* 35: L17807.

Conway, T. J., P. P. Tans, L. S. Waterman, and K. W. Thoning. 1994. "Evidence for Inter-Annual Variability of the Carbon Cycle from the NOAA/CMDL Global Air Sampling Network." *Journal of Geophysical Research* 99: 22831–22855.

Crevoisier, C., A. Chédin, H. Matsueda, T. Machida, R. Armante, and N. A. Scott. 2009. "First Year of Upper Tropospheric Integrated Content of CO_2 from IASI Hyperspectral Infrared Observations." *Atmospheric Chemistry and Physics* 9: 4797–4810.

Dlugokencky, E., and P. Tans. 2013. *Ed Dlugokencky and Pieter Tans*, *NOAA/ESRL*. Accessed July 5, 2013. www.esrl.noaa.gov/gmd/ccgg/trends/

Engelen, R. J., and A. P. McNally. 2005. "Estimating Atmospheric CO_2 from Advanced Infrared Satellite Radiances within an Operational Four-Dimensional Variational (4D-Var) Data Assimilation System: Results and Validation." *Journal of Geophysical Research* 110: D18305.

Erickson III, D. J., P. J. Rasch, P. P. Tans, P. Friedlingstein, P. Ciais, E. Maier-Reimer, K. Six, C. A. Fischer, and S. Walters. 1996. "The Seasonal Cycle of Atmospheric CO_2: A Study Based on the NCAR Community Climate Model (CCM2)." *Journal of Geophysical Research* 101: 15,079–15,097.

Hammerling, D. M., A. M. Michalak, C. O'Dell, and S. R. Kawa. 2012. "Global CO_2 Distributions over Land from the Greenhouse Gases Observing Satellite (GOSAT)." *Geophysical Research Letters* 39: L08804.

Maddy, E. S., C. D. Barnet, M. Coldberg, C. Sweeney, and X. Liu. 2008. "CO_2 Retrievals from the Atmospheric Infrared Sounder: Methodology and Validation." *Journal of Geophysical Research* 113: D11301.

Olsen, E. T. 2009. *AIRS Version 5 Release Tropospheric CO_2 Products*. Pasadena, CA: Jet Propulsion Laboratory.

Olsen, S. C., and J. T. Randerson. 2004. "Differences Between Surface and Column Atmospheric CO_2 and Implications for Carbon Cycle Research." *Journal of Geophysical Research* 109: D02301.

Rutgersson, A., M. Norman, and G. Astrom. 2009. "Atmospheric CO_2 Variation over the Baltic Sea and the Impact on Air–Sea Exchange." *Boreal Environment Research* 14: 238–249.

Stamenov, J., C. Angelov, I. Angelov, A. Antonov, L. Branekov, D. Belokapov, A. Damianova, K. Georgiev, M. Gelev, E. Hristova, P. Ivanov, I. Kalapov, I. Kirov, A. Mishev, A. Nishev, N. Nikolova, I. Sivriev, S. Todorov, and B. Vachev. 2007. "BEO Moussala – A Site for Integrated Environmental Studies-First Results." *Observatoire De Montagne De Moussala OM2* 12: 9–23.

Tsutsumi, Y., K. Mori, T. Hirahara, M. Ikegami, and T. J. Conway. 2009. *Technical Report of Global Analysis Method for Major Greenhouse Gases by the World Data Center for Greenhouse Gases*. GAW Report No. 184. Geneva: World Meteorological Organization.

Varotsos, C., M. N. Assimakopoulos, and M. Efstathiou. 2007. "Technical Note: Long-Term Memory Effect in the Atmospheric CO_2 Concentration at Mauna Loa." *Atmospheric Chemistry and Physics* 7: 629–634.

Wang, K., H. Jiang, X. Y. Zhang, and G. M. Zhou. 2011. "Analysis of Spatial and Temporal Variations of Carbon Dioxide over China Using SCIAMACHY Satellite Observations during 2003–2005." *International Journal of Remote Sensing* 32: 815–832.

Wang, T., A. J. Ding, D. R. Blake, W. Zahorowski, C. N. Poon, and Y. S. Li. 2003. "Chemical Characterization of the Boundary Layer Outflow of Air Pollution to Hong Kong during February–April 2001." *Journal of Geophysical Research* 108 (D20): 8787.

WMO GAW Report No. 161. 2005. *12th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracers Measurement Techniques*. Geneva: World Meteorological Organization.

Worthy, D. E. J., K. Higuchi, and D. Chan. 2003. "North American Influence on Atmospheric Carbon Dioxide Data Collected at Sable Island, Canada." *Tellus* 55B: 105–114.

Yoshida, Y., Y. Ota, N. Eguchi, N. Kikuchi, K. Nobuta, H. Tran, I. Morino, and T. Yokota. 2011. "Retrieval Algorithm for CO_2 and CH_4 Column Abundances from Short-Wavelength Infrared Spectral Observations by the Greenhouse Gases Observing Satellite." *Atmospheric Measurement Techniques* 4: 717–734.

Zellweger, C., B. Schwarzenbach, M. Steinbacher, and B. Buchmann. 2011. *System and Performance Audit of Source Ozone, Methane, Carbon Dioxide and Carbon Monoxide at the Global GAW Station Cape Point South Africa*. WCC-Empa Report 11/1. Geneva: World Meteorological Organization.

Zhang, L., D. J. Jacob, K. W. Bowman, J. A. Logan, S. Turquety, R. C. Hudman, Q. Li, R. Beer, H. M. Worden, J. R. Worden, C. P. Rinsland, S. S. Kulawik, M. C. Lampel, M. W. Shephard, B. M. Fisher, A. Eldering, and M. A. Avery. 2006. "Ozone-CO Correlations Determined by the TES Satellite Instrument in Continental Outflow Regions." *Geophysical Research Letters* 33: L18804.

Zimmerman, D. W. 1986. "Tests of Significance of Correlation Coefficients in the Absence of Bivariate Normal Populations." *Journal of Experimental Education* 54: 223–227.