

# Trend analysis of greenhouse gases over Europe measured by a network of ground-based remote FTIR instruments

T. Gardiner<sup>1</sup>, A. Forbes<sup>1</sup>, M. de Mazière<sup>2</sup>, C. Vigouroux<sup>2</sup>, E. Mahieu<sup>3</sup>, P. Demoulin<sup>3</sup>, V. Velasco<sup>4</sup>, J. Notholt<sup>4</sup>, T. Blumenstock<sup>5</sup>, F. Hase<sup>5</sup>, I. Kramer<sup>5</sup>, R. Sussmann<sup>6</sup>, W. Stremme<sup>6</sup>, J. Mellqvist<sup>7</sup>, A. Strandberg<sup>7</sup>, K. Ellingsen<sup>8</sup>, and M. Gauss<sup>8</sup>

<sup>1</sup>National Physical Laboratory, Teddington, UK

<sup>2</sup>Belgisch Instituut voor Ruimte-Aëronomie, Brussels, Belgium

<sup>3</sup>University of Liège, Institute of Astrophysics and Geophysics, Liège, Belgium

<sup>4</sup>University of Bremen, Department of Physics, Institute of Environmental Physics, Bremen, Germany

<sup>5</sup>Forschungszentrum Karlsruhe, Institute of Meteorology and Climate Research, Atmospheric Trace Constituents and Remote Sensing, Karlsruhe, Germany

<sup>6</sup>Forschungszentrum Karlsruhe, Institute of Meteorology and Climate Research, Atmospheric Environmental Research, Garmish-Partenkirchen, Germany

<sup>7</sup>Chalmers University of Technology, Göteborg, Sweden

<sup>8</sup>University of Oslo, Oslo, Norway

Received: 4 October 2007 – Published in Atmos. Chem. Phys. Discuss.: 12 November 2007

Revised: 13 February 2008 – Accepted: 27 September 2008 – Published: 24 November 2008

**Abstract.** This paper describes the statistical analysis of annual trends in long term datasets of greenhouse gas measurements taken over ten or more years. The analysis technique employs a bootstrap resampling method to determine both the long-term and intra-annual variability of the datasets, together with the uncertainties on the trend values. The method has been applied to data from a European network of ground-based solar FTIR instruments to determine the trends in the tropospheric, stratospheric and total columns of ozone, nitrous oxide, carbon monoxide, methane, ethane and HCFC-22. The suitability of the method has been demonstrated through statistical validation of the technique, and comparison with ground-based in-situ measurements and 3-D atmospheric models.

## 1 Introduction

Global climate change is one of the most important environmental issues facing the world today. A key element of this issue is understanding the atmospheric behaviour of radiatively active gases (direct greenhouse gases), and also gases

involved in the chemical production of greenhouse gases (indirect greenhouse gases). Long-term measurements of such gases provide the experimental data to study the evolution of these gases and the changing sources and sinks. These data are often expressed in terms of an annual trend in the amount of a particular gas. In order for these trend results to be used appropriately it is vital that the uncertainty associated with the trend value is properly quantified. An accurate determination of the trend value is challenging due to influence of large seasonal variations and other effects reflected in the data (Oltmans et al., 1998).

This paper describes the development and implementation of a trend analysis method to determine the annual trend and associated uncertainties, based on a statistical model that makes minimal assumptions about uncertainty distributions associated with the raw data. The method has been applied to measurements of direct and indirect greenhouse gases measured by a network of six ground-based solar Fourier Transform Infrared (FTIR) sites across Europe. The outputs from the analysis are the annual trends in the total, tropospheric and stratospheric amount of each gas at each of the sites and their associated uncertainties.

Section 2, below, gives a short description of the measurement network and the derivation of tropospheric and stratospheric columns from the data. The trend analysis method is described in Sect. 3, while Sect. 4 covers the validation of the



Correspondence to: T. Gardiner  
(tom.gardiner@npl.co.uk)

**Table 1.** The UFTIR Network of ground-based solar FTIR sites.

Site	Latitude	Longitude	Altitude (km)	Data Period
Ny Ålesund	79° N	12° E	0.02	1994–2004
Kiruna	68° N	20° E	0.4	1996–2004
Harestua	60° N	11° E	0.6	1995–2004
Zugspitze	47° N	11° E	2.96	1995–2004
Jungfraujoch	46.5° N	8° E	3.58	1995–2004
Izana	28° N	16.5° W	2.36	1999–2004

method. Section 5 gives the main results of the trend analysis, including comparison with in-situ trend measurements and atmospheric model results. The conclusions are given in Sect. 6.

## 2 The UFTIR remote sensing network

The work described in this paper was carried out as part of an EC Project on “Time Series of Upper Free Troposphere Observations from a European Ground-based FTIR Network” – UFTIR ([www.nilu.no/uftir](http://www.nilu.no/uftir)) (De Mazière et al., 2005). The UFTIR remote sensing network comprises six sites across Europe making solar absorption measurements using high-resolution FTIR spectrometers. Table 1 gives the location and altitude of these sites, which range in latitude from 28° N (Izana, Tenerife) to 79° N (Ny Ålesund, Spitzbergen), together with the general period over which data were available for the trend analysis. These sites have been making total column measurements of a range of atmospheric gases for many years, and the results from these measurements are held on the database of the Network for the Detection of Atmospheric Composition Change ([www.ndacc.org](http://www.ndacc.org), formerly the Network for the Detection of Stratospheric Change, NDSC). The work within the UFTIR project has focussed on the derivation of vertical profiles of a number of key tropospheric gases – ozone, nitrous oxide, carbon monoxide, methane, ethane and HCFC-22. The data have been produced from a combination of re-analysis of previous data, and new measurements made during the UFTIR project, giving a series of datasets that typically cover the period from the beginning of 1995 to the end of 2004. A significant amount of effort was made during the course of the project to harmonise the data analysis procedures used by each group (De Mazière et al., 2005).

The outputs of the FTIR measurements are time series of vertical profiles for each species with a profile for each day on which a measurement was made. Where more than one measurement was made on a particular day, the daily mean profile has been taken. The vertical profiles of the gases are determined from an optimal estimation method which generates a model atmosphere that best reproduces the observed solar absorption. The model atmosphere consists of a se-

ries of layers typically 1 to 2 km thick. The temperature and pressure in each layer is estimated and an a priori “partial column” of target gas in the layer (in units of molecules per m<sup>2</sup>, i.e. the number of molecules in a one square metre vertical column through the layer) assumed. The a priori gas distribution is adjusted to give the best match between measured and modelled absorption given the covariance characteristics of the various parameters. The result is an optimal estimate of the vertical profile of the gas made up of the partial column amounts in each layer. The total column amount (also in units of molecules per m<sup>2</sup>) is the summation of the partial columns. The exact details of the layer heights vary between species and sites. In this work the total column is taken as the summation of the partial columns up to 50 km (or nearest available layer boundary) in order to give consistency between all sites and species. More details of the FTIR analysis methods are given by De Mazière et al. (2005) and the references therein.

The work described in this paper addresses the determination of the trends and the associated uncertainties for the UFTIR datasets, with a focus on calculating separate tropospheric and stratospheric trends for each species.

### 2.1 Tropospheric column determination

Since one of the primary objectives for the analysis was to determine separate tropospheric and stratospheric trends for the FTIR datasets, a key issue was how to quantify the tropospheric content of the atmospheric profile results. It was decided that the best option was to use tropopause altitude information from the NCEP meteorological database to determine appropriate tropopause heights and variabilities for each site. The average tropopause height ranged from 10.14 km at Ny Ålesund to 14.85 km at Izana. The ( $1\sigma$ ) variability of the tropopause was between 1.10 km (at the Jungfraujoch) and 1.55 km (at Izana).

The tropopause information was then used to produce a weighting function to apply to the partial column profile data. The tropospheric weighting function,  $W$ , is a sigmoid function of altitude of the form:

$$W(z) = 1 - 1/[1 + \exp\{-a(z - z_T)\}] \quad (1)$$

where  $z$  = mean layer altitude,  $z_T$  = mean tropopause altitude, and  $a = e/(\text{standard deviation of tropopause altitudes})$ .

The weighting function acts to give values of approximately 1 well below the tropopause and approximately 0 well above, with a transition from 1 to 0 around the tropopause with a vertical extent governed by the variability of the tropopause height at that site. The tropospheric partial column was then determined by integrating the weighted profile. The stratospheric partial column was then calculated as the difference between the total column and the tropospheric partial column. Separate trends were then calculated for the total, tropospheric and stratospheric columns.

### 3 Trend analysis method

#### 3.1 Trend analysis requirements

The objective of the trend analysis method is to assess whether there are statistically significant long-term trends in the various datasets. The most straightforward approach to determining a trend from data is to fit a straight line to the data, using a least squares criterion for example. The gradient of the fitted line can then be used to indicate the long term trend. In order to associate an uncertainty or confidence limits with the gradient it is necessary to estimate the contribution of random effects in the data to the likely variation in the slope that would be computed if the data were gathered a number of times over identical conditions. If the random effects can be assumed to be independently and identically normally distributed, then it is straightforward to show that the gradient parameter is also associated with a normal (Gaussian) distribution, allowing confidence limits to be calculated easily. However, the FTIR measurements show significant intra-annual effects so that the likely departure of the data points from a straight line fit has a significant time-dependent correlation, and hence is not independent. Secondly, even for measurements at the same site at the same period, the observed distribution of measurements can have significantly non-normal features. In order to determine valid estimates of the trends, it is necessary to take into account both the intra-annual variability and the potential non-normality of the distributions associated with the measurement data.

The approach described in this paper augments the basic linear trend model with an intra-annual function in order to represent the intra-annual effects, and uses least-squares regression in conjunction with a bootstrap resampling method in order to determine confidence limits associated with the trend estimates. The advantage of the approach is that it uses well-known least squares techniques without requiring an assumption of normality at the same time as accounting for the significant intra-annual effects present in the data.

#### 3.2 Intra-annual model

Since the intra-annual (seasonal) variability is of a periodic nature, it is appropriate to model these effects in terms of a Fourier series,  $V$ :

$$V(t, \mathbf{b}) = b_0 + b_1 \cos 2\pi t + b_2 \sin 2\pi t + b_3 \cos 4\pi t + b_4 \sin 4\pi t + \dots \quad (2)$$

where  $t$  is measured in years, and  $b_0$  to  $b_n$  are the parameters of the Fourier series contained in the vector  $\mathbf{b}$ . The total variation in measurements due to the trend and the intra-annual effects is then modelled by a function,  $F$ :

$$F(t, a, \mathbf{b}) = at + V(t, \mathbf{b}) \quad (3)$$

where  $a$  is the annual trend in the data. This model captures the underlying periodicity of the data and reduces the impact of sparse data. It also enables regular gaps in the data series,

such as those during the winter months in high latitude sites, to be accommodated without causing discontinuities in the intra annual function. See Sect. 6.2 for examples of the fitted intra-annual models.

It is recognised that, with the relatively simple intra-annual model used here, the residuals are unlikely to be fully independent. The trend data presented here result from a common algorithm with a compact set of coefficients across the full set of UFTIR species and sites to give an overview of the basic trend behaviours. With some minor modification, the approach described could be extended to accommodate more complex models that, for example, capture atmospheric processes such as the QBO and solar cycles, or include non-linear trend behaviours, and thereby reducing any dependence within the residuals.

#### 3.3 Bootstrap resampling

The technique of bootstrap resampling enables non-normally distributed data to be treated robustly (Gatz and Smith, Part II, 1995). It is based on the idea that the distribution associated with the random effects reflected in the data is best represented by the residual deviations of a model fit to the data. The appropriateness of the bootstrap resampling technique to measuring trends in air quality data sets has already been demonstrated (Cox et al., 2002). In this technique, the model function  $F(t, a, \mathbf{b})$  is fitted to the data  $(t_i, M_i)$  to determine estimates  $a_0$  and  $\mathbf{b}_0$  that minimise

$$\sum_{i=1}^m (M_i - F(t_i, a, \mathbf{b}))^2 \quad (4)$$

with respect to  $a$  and  $\mathbf{b}$ . Since the function  $F$  is a linear function of the parameters  $a$  and  $\mathbf{b}$ , these estimates can be found using standard linear least squares methods (Lawson and Hansen, 1974).

Once the initial fit has been determined, the residual deviations

$$R_{i,0} = M_i - F(t_i, a_0, \mathbf{b}_0) \quad (5)$$

are then regarded as a discrete representation of the distribution associated with the random effects reflected in the data. Given  $R_{i,q}$  sampled at random from the set  $\{R_{i,0}\}$  (with replacement),  $i=1, \dots, m$ , a new data set  $\{(t_i, M_{i,q})\}$  is generated with:

$$M_{i,q} = F(t_i, a_0, \mathbf{b}_0) + R_{i,q} \quad (6)$$

The model is refitted to this data set to give parameter estimates  $a_q$  and  $\mathbf{b}_q$ .

This procedure is repeated a large number of times,  $q=1, \dots, N$ , to generate the 1 by  $N$  vector  $\mathbf{A}$  containing the set of trend results  $\{a_q\}$  and the  $n$  by  $N$  matrix  $\mathbf{B}$  with the set of intra-annual variability parameters  $\{\mathbf{b}_q\}$ . Each row of these matrices contains a sample from the distribution for the corresponding parameter, and therefore provides a (discrete)

**Table 2.** Precision of the confidence limits for the trends (as % of 2000 value) in the Jungfraujoch tropospheric column data.

Species	Lower limit of estimated trend		Upper limit of estimated trend	
	Value	precision range	Value	precision range
CH <sub>4</sub>	0.076	0.075 to 0.076	0.147	0.146 to 0.148
C <sub>2</sub> H <sub>6</sub>	-1.73	-1.74 to -1.72	-0.944	-0.958 to -0.927
CO	-0.545	-0.554 to -0.537	0.031	0.042 to 0.025
HCFC-22	2.29	2.28 to 2.30	2.66	2.66 to 2.67
N <sub>2</sub> O	0.203	0.202 to 0.204	0.272	0.271 to 0.273
O <sub>3</sub>	-0.288	-0.295 to -0.281	0.825	0.783 to 0.878

approximation to this distribution. Since the elements in **A** form a sample from the distribution for the trend parameters, the 2.5 and 97.5 percentiles of this empirical distribution specify a 95% confidence interval associated with the value of the trend. Using standard matrix factorisation techniques, it is possible to organise the computation so that determination of the addition parameter fits can be done efficiently.

This method allows the uncertainty associated with any of the model parameters to be evaluated without making any assumptions about the statistical distribution of the residuals. It can therefore be applied generally to results for different species and sites. It can also be extended to combinations of parameters, and an example of this is described in Sect. 5.3, where methods for aggregating the results obtained from many sites are discussed.

#### 4 Validation of analysis method

As in many areas of data analysis, tests must be carried out in order to demonstrate that the results obtained are valid and that the model chosen provides a satisfactory explanation of the data. In this application, it is necessary to choose an appropriate number of terms in the intra-annual model (i.e. the order of the Fourier series). The method used to evaluate the confidence intervals associated with the estimates relies on bootstrap resampling from the distribution of residual errors. This approach requires some further demonstration that there is no significant bias introduced and that the confidence intervals are reliable.

In addition to the statistical validation of the model, the results of the trend analysis of the data from the UFTIR sites can be compared to the results of ground-based in-situ monitoring networks, and the output from atmospheric models, to give confidence that similar long-term behaviour is seen in these different datasets.

##### 4.1 Number of factors in the intra-annual model

The choice of the number of terms in the Fourier series has to balance the need to determine a faithful representation of the underlying periodic behaviour with that of avoiding over-fitting the data. An investigation was carried out to as-

sess the appropriate order for the Fourier series by looking at the root-mean-square (RMS) residuals for different orders for each of the UFTIR species. As the order is increased the RMS is calculated. The point at which the RMS value shows no significant reduction usually represents a good balance between faithfulness and economy of representation. This study showed that a 3rd order series with a total of 7 coefficients (a constant and 3 sine and 3 cosine components) provided the best overall representation of the typical intra-annual variability without over-fitting the data.

##### 4.2 Bias

It is possible for a bootstrap resampling process to introduce biased confidence intervals. Efron and Tibshirani (1993) describe how any bias in the analysis may be quantified and, if it is large, how bias-corrected intervals may be estimated. As shown by Efron and Tibshirani (1993), the bias correction,  $z_0$ , for any statistic is given by  $z_0=C^{-1}(r)$ , where  $C^{-1}$  indicates the inverse function of the standard normal cumulative distribution function and  $r$  is the proportion of bootstrap resample values less than the original estimate. A value of  $r=1/2$ , giving  $z_0=0$ , implies that there is no bias.

When the bias check is applied to the bootstrap resample values of the trend estimated for each of the species and each of the sites, the values of  $r$  are all close to one half, ranging from 0.43 to 0.54 scattered (apparently) randomly about one half, with a mean bias of 0.49. It can therefore be concluded that the use of the bootstrap resampling method is not introducing significant bias, and a bias correction is not necessary.

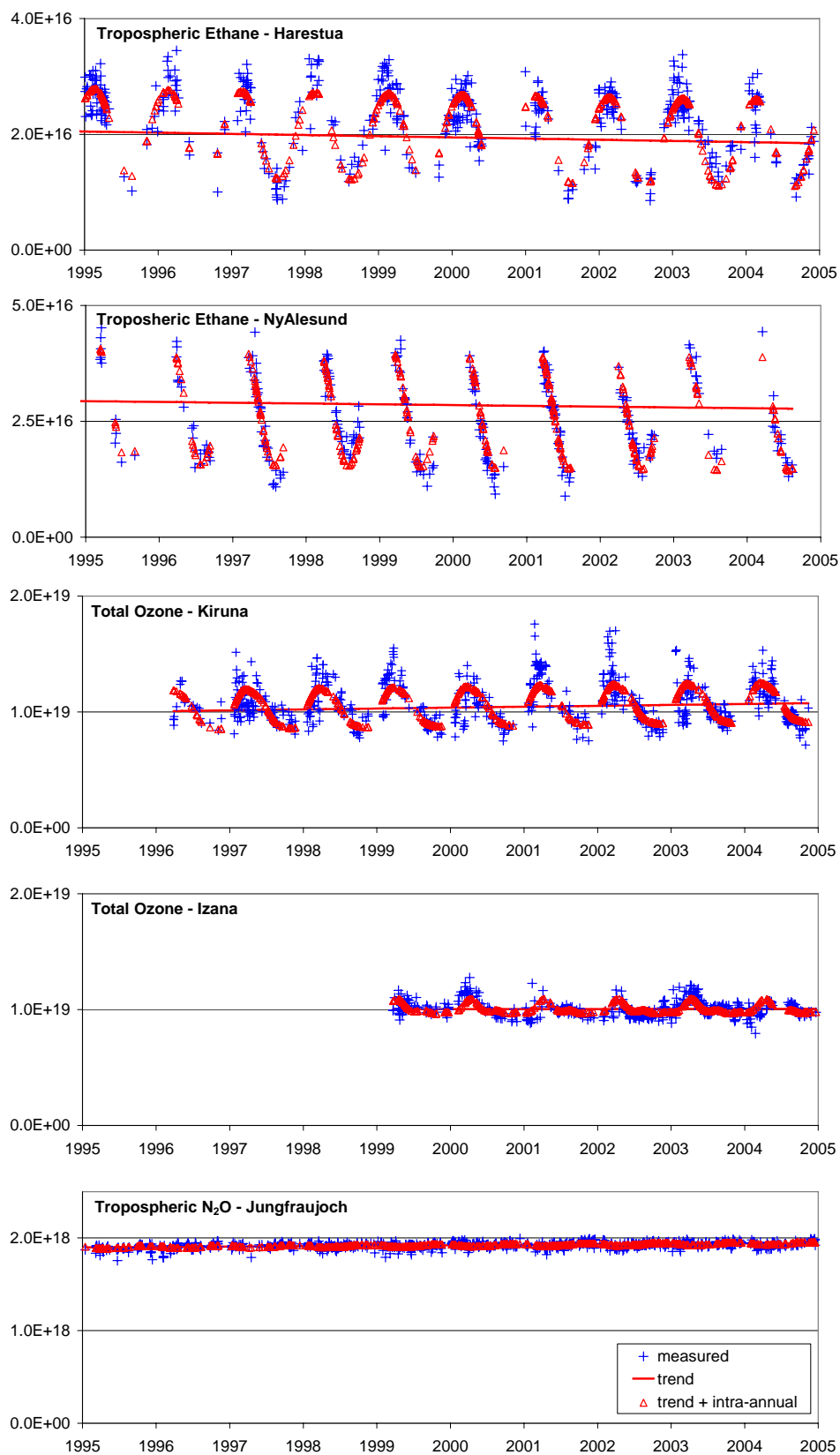
##### 4.3 Reliability of confidence limits

The bootstrap resampling method has been used to estimate the non-parametric 95% confidence intervals associated with the underlying trends. However, it is necessary to assess the level of confidence that can reasonably be placed on these intervals for the finite number  $N$  of resamplings ( $N=5000$  in our case) used. Berthouex and Brown (1994, p. 68) state that the precision of the quantiles estimated in the manner used here decreases rapidly as the estimates move towards the extreme tails of the distribution. They provide formulae for quantifying this precision, as follows. Let  $p$  denote the fractional quantile of interest (0.025 and 0.975 in our case). Then the 95% confidence intervals associated with the  $p$ th quantile are (for a large sample, as in our case):

$$\left[ p(N+1)-1.96Np(1-p)^{1/2}, p(N+1)+1.96Np(1-p)^{1/2} \right] \quad (7)$$

The corresponding values of the statistic of interest (the trend in our case) are then obtained immediately from the corresponding values in the empirical error distribution.

Table 2 shows an example of the results of the confidence limit precision. In this case they are calculated for the tropospheric partial column trends from the Jungfraujoch dataset as a percentage of the average value in 2000. The table gives



**Fig. 1.** Examples of measured (blue crosses) and fitted (red triangles) time series of vertical column amount (in molecules m<sup>-2</sup>) including the underlying trend (red line).

the lower (2.5% quantile) and upper (97.5% quantile) confidence limits for each species, together with the associated precision range calculated using Eq. (7).

The results shown in Table 2 indicate that the overall confidence intervals are not affected significantly when this “additional” uncertainty is included. We can therefore taken the original 95% confidence interval as a reasonable estimate of the uncertainty in the trend value.

## 5 Results of trend analysis

### 5.1 Time series and intra-annual variability

The output of the bootstrap resampling analysis produces an estimate of the average trend and intra-annual variability parameters for a given dataset. The first step in the analysis of the results was to see how well these parameters captured the variability and trends in the measurements. The five panels of Fig. 1 show a series of examples of the measured time series, the results of the model function determination, and the underlying trend. The top two panels are the time series for tropospheric ethane for Harestua and Ny Ålesund, showing that the Fourier series provides a good fit for species with a large variability even when there are regular gaps in the data, as for Ny Ålesund where measurements are not possible during the Arctic winter. The third and fourth panel show the total column ozone time series for Kiruna and Izana, showing that the general structure is captured well even if the intra-annual behaviour is very different from site to site. The final panel shows that the method is equally appropriate in the cases where there is little intra-annual variability as in this example of tropospheric nitrous oxide measured at the Jungfraujoch.

In summary, the results shown in Fig. 1 demonstrate the suitability of the 3rd order Fourier series, discussed in Sect. 4.1, in capturing the range of intra-annual variability in the various datasets.

### 5.2 Trend results from individual sites

The null hypothesis to be tested for each analysis is that “there is no underlying straight-line trend over the time span of the data”, i.e. the gradient of the underlying long-term trend in the regression model is zero. The sampling distribution of the gradient of the underlying straight-line trend term is determined empirically using bootstrap resampling. If the 95% confidence interval associated with the gradient, computed from this empirical distribution, does not contain zero then, in a formal statistical sense, there is reason to doubt the null hypothesis.

Table 3 shows the annual trends in the total, tropospheric, and stratospheric columns for each species and site, together with the associated uncertainties based on the 95% confidence limits of the bootstrap resampling. The bootstrap

method provides separate estimates of the positive and negative uncertainties, however, in these analyses the difference between the magnitude of these uncertainties were very small. Therefore, for reasons of clarity, the uncertainty values given in Table 3 are the mean magnitude of the positive and negative uncertainties. Also shown are the site latitudes and altitudes. All trends are reported as a percentage of the average value in the year 2000 for that particular parameter. Those annual trends shown in bold identify those cases where the confidence interval does not contain zero, and the null hypothesis is not met, i.e. it indicates those cases where there is a statistically significant (positive or negative) trend.

For comparative purposes the 95% confidence intervals associated with the estimates in Table 3 were also calculated under the (untested) assumption that Gaussian statistics apply. In most cases these were comparable to the bootstrap resampled results. A valid bootstrap approach can be expected to provide reliable confidence intervals that may be smaller than or greater than those obtained under the assumption that Gaussian statistics apply. Whether the intervals are smaller or greater depends on the nature of the sampling distribution, which the bootstrap resampling method estimates in an unbiased way. One particular point is that the use of the bootstrap technique is designed to accommodate outliers but, at the same time, provide estimates comparable with least squares estimates when a Gaussian model is appropriate. In this way, the bootstrap method gives a suitable generalisation of a standard approach that provides an important advantage in data of this type where outliers may be an issue. We concluded that our approach is valid because of the careful validation carried out on the results.

### 5.3 Estimating combined trends from all sites

In this section we aim to consider how the results obtained for each site can be aggregated to evaluate the long-term trend over the whole network. There are many ways of combining the results from all six sites to obtain representative values for the whole network. In this paper, the selected approach is to take the mean of the individual site values. Computing this statistic is straightforward, but standard approaches to evaluating the associated uncertainties can be misleading because only a small number of data points are available – six in this case. We show here how bootstrap resampling can be used to overcome this problem (Efron, 1982; Efron and Tibshirani, 1993).

As a result of the calculations described earlier, for each of the  $p$  sites, a row vector of length  $N$ , that describes the sampling distribution of the gradient in ozone concentration was calculated. These vectors can be arranged into a single  $p$  by  $N$  matrix,  $\mathbf{G}$ , that contains all  $N$  ( $=5000$ ) bootstrap estimates from all six sites.

This array of bootstrap estimates can form the basis for estimating the confidence interval associated with any estimator formed by taking a function of the trends – in this case

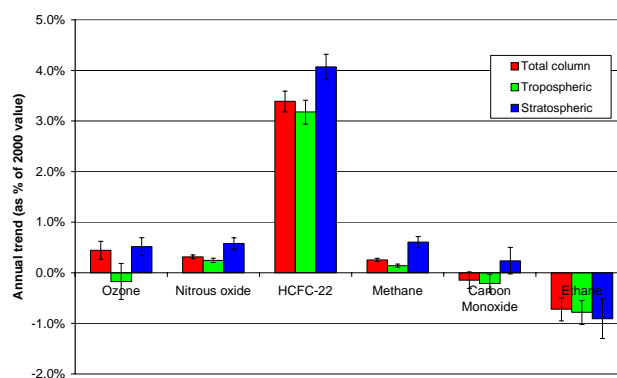
**Table 3.** Annual trend results for UFTIR measurements.

Site	Latitude	Altitude (km)	Annual Trend in Total Column (as % of 2000 value)					
			CH <sub>4</sub>	C <sub>2</sub> H <sub>6</sub>	CO	HCFC-22	N <sub>2</sub> O	O <sub>3</sub>
Ny Ålesund	79° N	0.02	<b>0.14 (±0.08)</b>	<b>-0.76 (±0.54)</b>	-0.58 (±0.69)	<b>2.89 (±0.84)</b>	<b>0.33 (±0.16)</b>	<b>0.93 (±0.65)</b>
Kiruna	68° N	0.4	<b>0.35 (±0.08)</b>	<b>-0.63 (±0.37)</b>	-0.12 (±0.38)	<b>4.71 (±0.25)</b>	<b>0.43 (±0.10)</b>	<b>0.76 (±0.39)</b>
Harestua	60° N	0.6	<b>0.40 (±0.06)</b>	<b>-0.65 (±0.32)</b>	-0.14 (±0.25)	<b>4.23 (±0.15)</b>	<b>0.46 (±0.08)</b>	0.04 (±0.42)
Zugspitze	47° N	2.96	<b>0.12 (±0.05)</b>	<b>-1.14 (±0.60)</b>	-0.10 (±0.46)	<b>2.57 (±0.79)</b>	<b>0.12 (±0.06)</b>	0.22 (±0.30)
Jungfraujoch	47° N	3.58	<b>0.17 (±0.03)</b>	<b>-1.05 (±0.35)</b>	<b>-0.31 (±0.25)</b>	<b>2.68 (±0.16)</b>	<b>0.22 (±0.04)</b>	<b>0.41 (±0.21)</b>
Izana	28° N	2.36	<b>0.14 (±0.09)</b>	0.00 (±0.92)	<b>1.70 (±0.41)</b>	<b>3.64 (±0.24)</b>	<b>0.34 (±0.09)</b>	0.05 (±0.31)
Site	Latitude	Altitude (km)	Annual Trend in Tropospheric Column (as % of 2000 value)					
			CH <sub>4</sub>	C <sub>2</sub> H <sub>6</sub>	CO	HCFC-22	N <sub>2</sub> O	O <sub>3</sub>
Ny Ålesund	79° N	0.02	0.08 (±0.08)	<b>-0.74 (±0.54)</b>	-0.62 (±0.72)	<b>2.68 (±1.10)</b>	<b>0.35 (±0.22)</b>	0.26 (±0.99)
Kiruna	68° N	0.4	<b>0.11 (±0.06)</b>	<b>-0.75 (±0.38)</b>	-0.21 (±0.37)	<b>4.67 (±0.24)</b>	<b>0.26 (±0.07)</b>	0.22 (±0.58)
Harestua	60° N	0.6	<b>0.33 (±0.05)</b>	<b>-0.72 (±0.36)</b>	-0.13 (±0.28)	<b>4.82 (±0.18)</b>	<b>0.37 (±0.07)</b>	<b>-1.05 (±0.65)</b>
Zugspitze	47° N	2.96	<b>0.17 (±0.05)</b>	<b>-1.09 (±0.62)</b>	-0.12 (±0.53)	<b>1.76 (±0.81)</b>	<b>0.18 (±0.06)</b>	-0.14 (±0.76)
Jungfraujoch	47° N	3.58	<b>0.11 (±0.04)</b>	<b>-1.34 (±0.39)</b>	-0.25 (±0.29)	<b>2.48 (±0.19)</b>	<b>0.24 (±0.03)</b>	0.26 (±0.56)
Izana	28° N	2.36	<b>-0.26 (±0.09)</b>	0.09 (±0.95)	<b>1.60 (±0.43)</b>	<b>3.62 (±0.24)</b>	0.07 (±0.08)	-0.21 (±1.10)
Site	Latitude	Altitude (km)	Annual Trend in Stratospheric Column (as % of 2000 value)					
			CH <sub>4</sub>	C <sub>2</sub> H <sub>6</sub>	CO	HCFC-22	N <sub>2</sub> O	O <sub>3</sub>
Ny Ålesund	79° N	0.02	<b>0.14 (±0.08)</b>	<b>-3.31 (±1.31)</b>	-0.22 (±0.97)	<b>4.10 (±1.16)</b>	0.26 (±0.35)	<b>1.03 (±0.66)</b>
Kiruna	68° N	0.4	<b>0.35 (±0.08)</b>	0.56 (±1.05)	0.47 (±0.85)	<b>4.82 (±0.28)</b>	<b>0.90 (±0.30)</b>	<b>0.81 (±0.40)</b>
Harestua	60° N	0.6	<b>0.71 (±0.24)</b>	0.43 (±0.58)	-0.13 (±0.50)	<b>1.97 (±0.23)</b>	<b>0.83 (±0.33)</b>	0.19 (±0.42)
Zugspitze	47° N	2.96	<b>0.12 (±0.05)</b>	<b>-1.71 (±0.58)</b>	0.04 (±0.30)	<b>5.15 (±0.83)</b>	-0.05 (±0.20)	0.27 (±0.30)
Jungfraujoch	47° N	3.58	<b>0.17 (±0.03)</b>	<b>-0.70 (±0.33)</b>	<b>-0.37 (±0.28)</b>	<b>2.88 (±0.14)</b>	<b>0.20 (±0.05)</b>	<b>0.42 (±0.20)</b>
Izana	28° N	2.36	<b>0.14 (±0.09)</b>	-1.11 (±1.40)	<b>1.41 (±0.33)</b>	<b>3.73 (±0.26)</b>	<b>1.42 (±0.33)</b>	0.07 (±0.31)

the arithmetic mean. For a specific estimator, we form its value for each column of the matrix  $\mathbf{G}$ , (i.e. for each bootstrap sample). The result is a 1 by  $N$  matrix whose elements estimate the sampling distribution for that estimator. The 2.5 and 97.5 percentiles of this empirical distribution specify a 95% confidence interval associated with the value of the estimator. Figure 2 shows the combined trend values for the total, tropospheric and stratospheric columns for each of the UFTIR species. The error bars on each trend show the associated 95% confidence intervals obtained from the estimated sampling distribution.

The results shown in Fig. 2 can be taken as estimates of the trends for each species over a large spatial scale. These results are in good agreement with the behaviour predicted by a model of the atmosphere – see Sect. 6.5. The results of the combined trend behaviour are also in good agreement with the trends determined by long term ground level monitoring. For example:

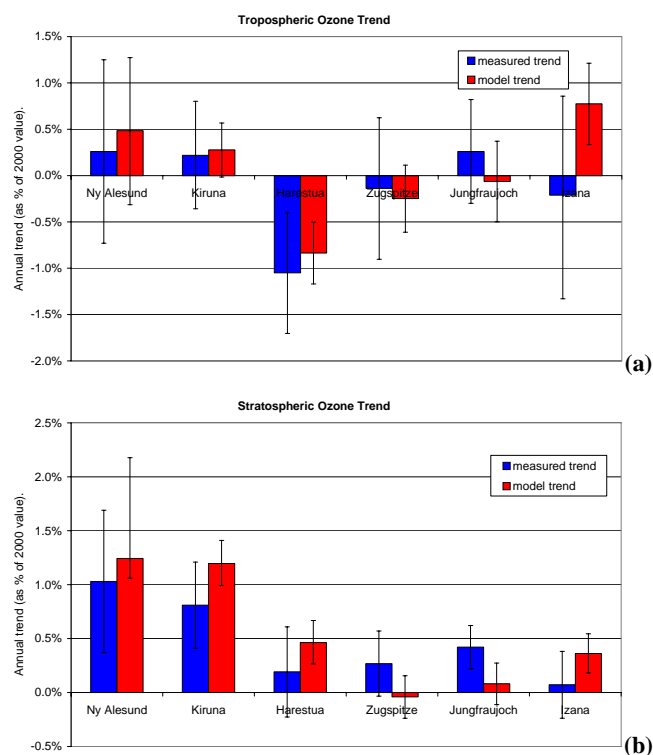
- the measured value for the tropospheric N<sub>2</sub>O trend of 0.245 (±0.044)%/yr compares to a global mean rate of 0.25%/yr determined by the AGAGE and NOAA/CMDL network results (WMO, 2003).
- the measured tropospheric HCFC-22 trend of 3.18(±0.24)%/yr can be compared to the GC/MS

**Fig. 2.** Total column, tropospheric and stratospheric trends for each species combined over all sites.

measurements from the AGAGE site at Mace Head in Ireland (53° N) which give a growth rate of 3.02%/yr for the period 1999–2003.

#### 5.4 Comparison with the CTM model

An alternative approach to looking at the combination of results from the different sites is to compare the measured trends to those predicted by 3-D atmospheric models. The



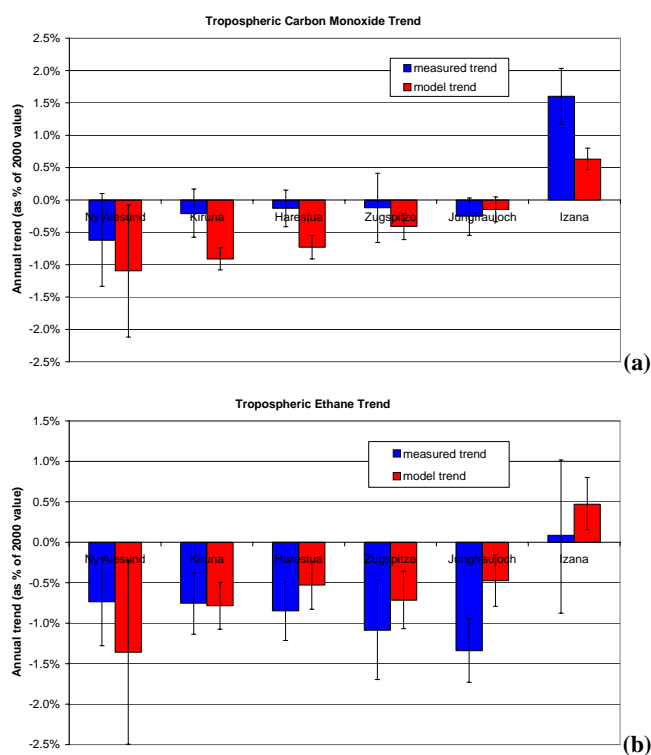
**Fig. 3.** Comparison between measured and modelled trends in (a) tropospheric and (b) stratospheric ozone for each UFTIR site.

advantage of this method is that it enables systematic differences in the behaviour at different sites (and latitudes) to be taken into account, and it also provides a useful validation tool for the long-term behaviour of the models themselves.

The model used within the UFTIR project was the Oslo CTM2 model developed within the Department of Geosciences at the University of Oslo (Isaksen et al., 2005; Gauss et al., 2006). This model is a global 3-D chemical transport model (CTM) driven by real meteorological data from the European Centre for Medium Range Weather Forecasting (ECMWF). The model was run with a  $2.8^\circ$  by  $2.8^\circ$  horizontal resolution and 40 vertical layers from the surface up to 10 hPa. The chemical scheme includes comprehensive stratospheric and tropospheric chemistry. The spatial and temporal variation in emissions has been included based on the EDGAR3.2 inventories for anthropogenic emissions (Olivier et al., 1999) and the GEIA inventories ([www.geiacenter.org/](http://www.geiacenter.org/)) for natural emissions.

The model was used to predict the vertical profiles for each species above each site over the trend analysis period. The tropospheric/stratospheric partial column functions (see Sect. 2.1) were then applied to the profiles, and the bootstrap resampling algorithm applied as for the measured columns.

The detailed results of these analyses, and their scientific implications will be discussed in other papers, however a few examples are given here to demonstrate the comparability be-



**Fig. 4.** Comparison between measured and modelled trends in (a) tropospheric carbon monoxide and (b) tropospheric ethane for each UFTIR site.

tween the measured and modelled trends. Figure 3a and b shows the tropospheric and stratospheric ozone trends, with generally good agreement between the measured and modelled results (given the trend uncertainties) including the differences from site to site and between troposphere and stratosphere. Figure 4a and b shows the tropospheric trends for carbon monoxide and ethane, again with good agreement between the data sets.

## 6 Discussion and conclusion

The ability to reliably determine trends in atmospheric datasets is an important element in the study of the long term behaviour of the atmosphere and climate system. Conventional methods for estimating the uncertainties in the trends may give misleading results as they make unjustified assumptions about the statistical distribution of the data. We have established that the method described in this paper – bootstrap resampling with a low order Fourier series to capture the intra-annual variability – is a statistically robust method for determining the trends and uncertainties. A series of statistical and experimental validation tests have been carried out to demonstrate the suitability of the method. We have also shown how the method can be applied to the aggregated



data from different sites to give an assessment of the long-term trends across a network of measurement sites.

The trend analysis method has been applied to long-term datasets of direct and indirect greenhouse gases measured by the UFTIR network of six ground-based solar FTIR sites. The output from these analyses gives the trends and associated uncertainties for the total, tropospheric and stratospheric columns of ozone, nitrous oxide, carbon monoxide, methane, ethane and HCFC-22. These results provide a valuable data resource for the study and modelling of the changing sources, sinks and dynamics for each species. Further papers will address the scientific interpretation of the results for the various species.

*Acknowledgements.* The work on the UFTIR project has been funded by the respective national authorities and the European Commission (under contract EVK2-2002-00159).

Edited by: S. Wood

## References

- Berthouex, P. M. and Brown, L. C.: *Statistics for Environmental Engineers*, CRC Press, Inc., US, 1994.
- Cox, M. G., Harris, P. M., Milton, M. J. T., and Woods, P. T.: *Method for Evaluating Trends in Ozone Concentration Data and its Application to Data from the UK Rural Ozone Monitoring Network*, NPL Report CMSC 15/02, November, 2002.
- De Mazière, M., Vigouroux, C., Gardiner, T., Coleman, M., Woods, P., Ellingsen, K., Gauss, M., Isaksen, I., Blumenstock, T., Hase, F., Kramer, I., Camy-Peyret, C., Chelin, P., Mahieu, E., Demoulin, P., Duchatelet, P., Mellqvist, J., Strandberg, A., Velasco, V., Notholt, J., Sussmann, R., Stremme, W., and Rockmann A.: The exploitation of ground-based Fourier transform infrared observations for the evaluation of tropospheric trends of greenhouse gases over Europe, *Environ. Sci.*, 2, 283–293, 2005.
- Efron, B.: *The Jackknife, the Bootstrap, and Other Resampling Plans*. Society for Industrial and Applied Mathematics, Philadelphia, 1982.
- Efron, B. and Tibshirani, R.: *An Introduction to the Bootstrap*, *Monographs on Statistics and Applied Probability* 57, Chapman and Hall, 1993.
- Gatz, D. F. and Smith, L.: The standard error of a weighted mean concentration – II. Estimating confidence intervals, *Atmos. Environ.*, 29, 1195–1200, 1995.
- Gauss, M., Isaksen, I. S. A., Lee, D. S., and Søvde, O. A.: Impact of aircraft NO<sub>x</sub> emissions on the atmosphere – tradeoffs to reduce the impact, *Atmos. Chem. Phys.*, 6, 1529–1548, 2006, <http://www.atmos-chem-phys.net/6/1529/2006/>.
- Isaksen, I. S. A., Zerefos, C. S., Kourtidis, K., Meleti, C., Dalsoren, B. B., Sundet, J. K., Grini, A., Zanis, P., and Balis, D.: Tropospheric ozone changes at unpolluted and semipolluted regions induced by stratospheric ozone changes, *J. Geophys. Res.*, 110, D02302, doi:10.1029/2004JD004618, 2005.
- Lawson, C. L. and Hanson, R. J.: *Solving Least Squares Problems*, Prentice Hall, Englewood Cliffs, New Jersey, 1974.
- Olivier, J., Bouwman, A., Berdowski, J., Bloos, J., Visschedijk, A., van der Mass, C., and Zandveld, P.: Sectoral emission inventories of greenhouse gases for 1990 on a per country basis as well as on 1×1 degrees, *Environ. Sci. Policy*, 2, 241–263, 1999.
- Oltmans, S. J., Lefohn, A. S., Scheel, H. E., Harris, J. M., et al.: Trends in ozone in the troposphere, *Geophys. Res. Lett.*, 25, 139–142, 1998.
- WMO: *Scientific Assessment of Ozone Depletion: 2002*, Global Ozone Research and Monitoring Project-Report No. 47, 498 pp., Geneva, 2003.