2003

Intercomparison of NDSC ground-based solar FTIR measurements of atmospheric gases at Lauder, New Zealand

D W. T Griffith
University of Wollongong, griffith@uow.edu.au

Nicholas Jones
National Institute of Water and Atmospheric Research, njones@uow.edu.au

B McNamara
National Institute of Water and Atmospheric Research

Clare Paton-Walsh
National Physical Laboratory, clarem@uow.edu.au

W. R. Bell
National Physical Laboratory

See next page for additional authors

Intercomparison of NDSC ground-based solar FTIR measurements of atmospheric gases at Lauder, New Zealand

Abstract

A formal intercomparison of atmospheric total column measurements of N_2O, N_2, CH_4, O_3, HCl, HNO_3, and HF by two ground-based solar Fourier transform infrared (FTIR) spectrometers conducted as part of the Network for the Detection of Stratospheric Change (NDSC) instrument certification procedure at Lauder, New Zealand, is presented. The two instruments were nominally very similar, collocated, and collected data at the same times. Collected spectra were analyzed independently by the individual operators in a blind-phase intercomparison, then reanalyzed by a single operator using identical analysis methods to eliminate any potential bias from the spectral analysis. From the consistent reanalysis, gases with predominantly tropospheric distributions and pressure-broadened spectral lines, such as N_2O and CH_4, showed differences between retrieved columns of typically less than 1%. For predominantly stratospheric gases, such as HCl and O_3, differences were less than 3%. In most cases, the differences were greater than the scatter in the individual measurements and were significant at the 95% confidence level. The worst case observed was for HF, which showed a 7% systematic bias between instruments. The differences are consistent in magnitude with those expected for known types of imperfection in spectrometer alignment and operation, but attempts to quantify these effects through instrument line shape analysis, phase error, zero offsets, and channel spectra did not remove the apparent differences.

Keywords

lauder, zealand, measurements, ftir, solar, ground, ndsc, gases, atmospheric, intercomparison, GeoQuest

Disciplines

Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Griffith, D. W. T., Jones, N. B., McNamara, B., Paton-Walsh, C., Bell, W. & Bernado, C. (2003). Intercomparison of NDSC ground-based solar FTIR measurements of atmospheric gases at Lauder, New Zealand. Journal of Atmospheric and Oceanic Technology, 20 (8), 1138-1153.

Authors

D W. T Griffith, Nicholas Jones, B McNamara, Clare Paton-Walsh, W. R. Bell, and Cirilo Bernado

Intercomparison of NDSC Ground-Based Solar FTIR Measurements of Atmospheric Gases at Lauder, New Zealand

D. W. T. GRIFFITH

University of Wollongong, Wollongong, Australia

N. B. JONES* AND B. MCNAMARA+

National Institute of Water and Atmospheric Research, Lauder, New Zealand

C. PATON WALSH* AND W. BELL

National Physical Laboratory, Teddington, United Kingdom

C. BERNARDO

University of Wollongong, Wollongong, Australia

(Manuscript received 17 June 2002, in final form 3 December 2002)

ABSTRACT

A formal intercomparison of atmospheric total column measurements of N_2O, N_2, CH_4, O_3, HCl, HNO_3, and HF by two ground-based solar Fourier transform infrared (FTIR) spectrometers conducted as part of the Network for the Detection of Stratospheric Change (NDSC) instrument certification procedure at Lauder, New Zealand, is presented. The two instruments were nominally very similar, collocated, and collected data at the same times. Collected spectra were analyzed independently by the individual operators in a blind-phase intercomparison, then reanalyzed by a single operator using identical analysis methods to eliminate any potential bias from the spectral analysis. From the consistent reanalysis, gases with predominantly tropospheric distributions and pressure-broadened spectral lines, such as N_2O and CH_4, showed differences between retrieved columns of typically less than 1%. For predominantly stratospheric gases, such as HCl and O_3, differences were less than 3%. In most cases, the differences were greater than the scatter in the individual measurements and were significant at the 95% confidence level. The worst case observed was for HF, which showed a 7% systematic bias between instruments. The differences are consistent in magnitude with those expected for known types of imperfection in spectrometer alignment and operation, but attempts to quantify these effects through instrument line shape analysis, phase error, zero offsets, and channel spectra did not remove the apparent differences.

1. Introduction

The Network for the Detection of Stratospheric Change (NDSC; Kurylo 1991) is dedicated to the detection and characterization of long-term changes in stratospheric chemical composition, especially as it affects stratospheric ozone. To this end, the NDSC coordinates the operation of an increasing number of ground-based remote sensing stations (currently over 50) from high northern to high southern latitudes, employing spectroscopic techniques including lidar, microwave, infrared, and UV/visible spectroscopy. The NDSC stations also provide measurements of predominantly tropospheric trace gases as well as a coordinated resource for comparison and calibration of satellite-borne instruments for atmospheric composition measurements. Further details can be obtained from the NDSC Web site (information online at http://www.ndsc.ncep.noaa.gov).

From the measurement perspective, the trends of interest may be quite small, of the order of a few percent per decade (Zander et al. 1998), and thus require a high degree of stability in the calibration of the instruments, both with time for each station and between stations. NDSC instruments must therefore satisfy rigorous certification and validation criteria, with the results open to scrutiny by the scientific peer establishment. The

* Current affiliation: University of Wollongong, Wollongong, Australia.
+ Deceased.

Corresponding author address: Prof. David Griffith, Dept. of Chemistry, University of Wollongong, Wollongong, NSW 2522, Australia.
E-mail: David_Griffith@uow.edu.au

present paper concerns the intercomparison of high-resolution Fourier transform infrared (FTIR) spectrometers operated as part of the NDSC for solar absorption spectrometry. As part of the procedure for instrument certification, each NDSC spectrometer should make and intercompare measurements in parallel with another spectrometer, collocated and at the same time, so as to probe the same air mass. For this purpose a spectrometer operated by the U.K. National Physical Laboratory (NPL) is designated as the ''traveling spectrometer'' for intercomparisons with other NDSC FTIR instruments at their normal sites of operation. Previous intercomparisons in this series have been described for instruments in Ny-Ålesund, Spitzbergen (79°N), and Harestua, Norway (60°N), by Paton Walsh et al. (1997) and Table Mountain, California (34°N), by Goldman et al. (1999). This paper describes the intercomparison of the NPL spectrometer with that operated by the National Institute for Water and Air Research (NIWA) at Lauder, New Zealand (45°S), in February 1997. The next intercomparisons in the series took place at Kiruna, Sweden, in 1998 and Eureka, Nunaret, Canada, in 1999 and will be reported on separately.

Analysis of ground-based FTIR absorption spectra of the sun provides the total column density of trace gases above the measurement site, with limited vertical profile information possible in some cases. The process can be summarized in three distinct steps: 1) recording of the solar intensity as a raw interferogram by the FTIR spectrometer, 2) phase correction and Fourier transformation of the interferogram to provide the solar spectrum as modified by the spectrometer, and 3) analysis of the solar spectrum to provide total vertical column amounts of trace gases. With appropriate software, vertical profile information may also be retrieved in this step.

It is most important to define the scope of the intercomparison at the outset. It is the aim of NDSC instrument certification to answer the question, To what extent do spectra measured by two spectrometers, collocated and operating in parallel, provide the same vertical column amounts of key trace gases when subjected to identical analysis? We thus wish to compare performance in steps 1 and 2, spectrum collection, without bias from step 3, the spectrum analysis. In this context, it is difficult to compare the spectra themselves (step 2) in any quantitative sense. Fortunately, the software algorithms commonly used in the NDSC for spectrum analysis (step 3) have been extensively intercompared in previous rigorous exercises and have been shown to have only small differences (Goldman 1996; Zander 1995; Zander et al. 1993).

In the present study the two spectrometers were of the same make and model, providing the most favorable circumstances for an intercomparison. Total column amounts of gases determined from spectra collected simultaneously were always compared using the same analysis procedure, thus avoiding bias from the analysis method.

2. Description of the measurements

The intercomparison took place at the NIWA Lauder laboratory (45.0°S, 169.7°E, 370 m ASL) in February 1997. Both groups used Bruker IFS 120M Fourier transform spectrometers. The NIWA spectrometer was coupled to the suntracker routinely used at Lauder for NDSC measurements. This tracker was built at the University of Denver and used a quadrant detector and feedback loop to actively keep the solar image on the entrance aperture of the spectrometer. The NIWA instrument was fitted with a KBr beamsplitter for all measurements, a mercury–cadmium–telluride (MCT) detector for spectra below 1800 cm^{-1}, and an indium–antimonide (InSb) detector for spectra above 1800 cm^{-1}. The exit beam from the interferometer was split to the two detectors via a dichroic beamsplitter. The NPL spectrometer was coupled to a second suntracker and located in the same laboratory as the NIWA instrument. The NPL tracker was an altazimuth-mounted passive tracker controlled by a small computer and supplied by NIWA. Both trackers were designed to image approximately the central third of the solar image on the entrance aperture. The NPL spectrometer was also equipped with MCT and InSb detectors and a dichroic beamsplitter but used a CaF_2 beamsplitter for all the InSb spectra and a KBr beamsplitter for the MCT spectra.

The procedures for recording and analyzing spectra generally followed the format of the Table Mountain, California, intercomparison (Goldman et al. 1999) except where described explicitly below. Spectra were collected in three separate spectral regions using three different bandpass interference filters as specified in Table 1. The filters for the InSb region were nominally identical filters from the NDSC-recommended set (filter 1, 4000–4300 cm^{-1}, for HF; filter 3, 2400–3200 cm^{-1}, for N_2, N_2O, CH_4, HCl, and O_3). In the MCT spectra NIWA used the standard NDSC filter 6 (longpass filter, cutoff ca. 1350 cm^{-1}), while NPL used a 1000 cm^{-1} cutoff filter. In the regions below 4000 cm^{-1}, spectra were normally taken at 180-cm maximum optical path difference (OPD), and in the region above 4000 cm^{-1} at 150-cm OPD, corresponding to resolutions (1/OPD) of 0.0056 and 0.0067 cm^{-1}, respectively. Some spectra were also collected at full resolution (257-cm OPD, 0.0039 cm^{-1}) in order to test the ability of the cell measurements to characterize the optical alignment under the most stringent conditions. Time constraints meant that these full resolution tests were limited to the NPL instrument.

The NPL instrument was damaged in transit to Lauder and required major realignment after arrival. In particular both of the beamsplitters were cracked, but could not be replaced for the intercomparison measurements and were used in that state. The resulting extra scattering caused a significant reduction in the signal-to-noise ratio (SNR), but otherwise seemed not to have affected the

TABLE 1. Details of all spectral measurements.

	Set 1	Set 2	Set 3	Set 4	Set 5
(a) NDSC filter 3 (2400–3130 cm^{-1}) N_2, N_2O, CH_4, HCl, and O_3					
Date (1997)	15 Feb	15 Feb	16 Feb	18 Feb	19 Feb
Time (NZ std)	1056–1127	1334–1358	0958–1057	1046–1120	0815–1005
SZA range (°)	41.4–37.8	33.2–34.7	50.2–41.5	43.6–39.5	68.5–53.4
No. of spectral[a]	11/12	12/12	8/14	12/12	22/24
Max OPD NIWA	180	180	180	180	180[b]
(cm) NPL	180	180	180	180	257
Aperture[c] NIWA	0.50	0.50	0.50[d]	0.50	0.50
(mm) NPL	0.65	0.65	0.65	0.65	0.65
(b) NDSC filter 1 (4030–4280 cm^{-1}) HF					
Date (1997)	15 Feb	15 Feb	16 Feb	18 Feb	19 Feb
Time (NZ std)	1137–1157	1406–1425	1106–1130	1046–1120	1012–1040
SZA range (°)	36.5–43.9	35.5–37.2	40.4–37.5	36.8–33.8	48.7–44.6
No. of spectra	12	12	12	12	15
Max OPD (cm)	150	150	150	150	150
Aperture[c] NIWA	0.65	0.65	0.65	0.65	0.65
(mm) NPL	0.65	0.65	0.65	0.65	0.65
(c) NDSC filter 6 [$<$1350 cm^{-1} (NIWA), $<$1000 cm^{-1} (NPL)] HNO_3					
Date (1997)	15 Feb	15 Feb	18 Feb	19 Feb	19 Feb
Time (NZ std)	0900–1023	1434–1527	1315–1418	0728–0805	1051–1203
SZA range (°)	60.0–46.6	38.1–44.8	33.5–37.1	77.0–70.3	43.5–36.0
No. of spectra[a]	10	12	14	12	14
Max OPD (cm)	180	180	180	180	180
Aperture[c] NIWA	1.1	1.1	1.1	1.1	1.1
(mm) NPL	1.4	1.4	1.1	1.4	1.4

[a] (No. of spectra included in analysis)/(no. of spectra collected in set). Spectra were excluded if the retrieved columns were $>$2.5 std dev from the mean for the set; in most cases these outliers corresponded to spectra with identifiably perturbed baselines or line shapes due to cloud or transient instrumental effects.

[b] Last two NIWA spectra were collected at 257 cm (for HBr cell measurements).

[c] In both spectrometers the FOV was limited by the input aperture from the interferometer. The output aperture was set to be one stop larger than the input aperture.

[d] First six NIWA spectra were collected with 0.65-mm aperture; all showed -6% zero offsets and have been excluded.

[e] All filter 6 measurements are 3 scans except for NPL set 4 where the last 8 measurements of the set are single scans.

alignment of the instrument. The reduced SNR reduced the precision of the measurements, but it should not have introduced any bias into the results. Once the NPL beamsplitters were replaced back in the United Kindgom, the SNR improved to a level comparable to that of the NIWA instrument. Both spectrometers were aligned to be within Bruker specifications for resolution based on the apparent widths of absorption lines of low pressure N_2O in a small absorption cell.

The intercomparison consisted of an initial open phase, where discussion and intercomparison of spectra and analyses was permitted, followed by a blind phase where no results could be exchanged. In the open phase, spectra were collected simultaneously by both spectrometers in sets of at least 10 single-scan spectra per filter. MCT and InSb spectra were run alternately, not simultaneously. Data were collected from both forward and reverse scans of the interferometer mirror, phase corrected, and Fourier transformed into spectra using boxcar apodization and standard Bruker OPUS software. Two full sets of open-phase spectra were collected for each filter on 13 and 14 February. Preliminary analysis of these spectra using the analysis protocols described below showed the two instruments to retrieve total columns of all gases to within 2% except HNO_3 and HF (5% and 7%, respectively, with the NIWA instrument higher in both cases). The reasons for these differences were not understood and no further adjustment was made to either spectrometer. The instruments were declared ready for the blind intercomparison.

In the blind phase of the intercomparison, five separate sets of spectra were collected in each filter region over four days, 16–19 February. Sky conditions were predominantly clear on all days. Full details of these measurements are given in Table 1, including dates, times, solar zenith angles, and spectrometer parameters. In each set, both spectrometers were started within 10 s of each other for at least 10 separate scans, except set 5, filter 3, when NPL spectra were recorded at 257 cm OPD. Single scans were saved, transformed, and analyzed individually for all InSb measurements, while three scans were coadded for each MCT measurement (filter 6). Spectra of low pressure (0.2 hPa) N_2O and HBr in 10-cm pathlength cells were recorded by both spectrometers before and after the intercomparison as a diagnostic for the instrument line shape (ILS). Spectra were measured with both a blackbody (globar) source and in the solar beam. For HBr in the fundamental band

TABLE 2. Details of spectrum analyses.

Setting/parameter	Blind analysis	Reanalysis
Fit software	SFIT 1.09d	SFIT 1.09e
Microwindows	Table 3	Table 3
Line parameters	HITRAN92[a]	HITRAN96
Fitted species	Table 3	Table 3
PT profile	5-yr mean for respective day	Single profile used for all days (5-yr mean for all Feb)
VMR profiles	Provided by NIWA[b]	Provided by NIWA
Layering (FSCATM)	36 (2 km to 50 km, 5 km to 100 km)	Same as above
Layering (SFIT)	29[c]	29
SZAs	Forward scans okay; reverse scan time same as forward scan[d]	Use correct SZAs for both forward and reverse scans
Effective apodization	Not fitted	Not fitted
100% line fit	Straight with slope	Straight with slope
Zero level	Not fitted	Not fitted
Monochromatic point spacing	Table 3	fix at 0.0004 cm^{-1}
Max iterations	30	20
Cm^{-1} shift	Single shift for whole spectrum	Single shift for whole spectrum
Pressure shifts	No	Yes
Channel spectra	No	No
Line widths	Air	Air

[a] Both groups agreed to use HIRAN92, except for HNO_3 (HIRAN96 without hot bands, as used in Table Mountain intercomparison). HITRAN96 used consistently in reanalysis.

[b] NIWA provided a consistent set of VMR profiles for southern midlatitudes in file REFNIWA.DAT. After distribution but before the blind analysis, the water vapor profile was adjusted in one set but not the other.

[c] In the blind intercomparison, each group used a different 29-level layering in SFIT: NIWA used 2-km layers to 50, 10–80, 20–100 km (29 layers); NPL used 1-km layers to 10, 2–30, 5–50, 10–100 km.

[d] The NIWA OPUS software version did not distinguish times for forward and reverse scans, and the same solar zenith angle was used for both forward and reverse scans in the blind analysis. The forward scan ZPD times and thus SZAs should thus be correct: the time error of ca. 200 s in the calculated reverse scan ZPD times and solar zenith angles corresponds to an error in air mass of 0.5%–1%, but both groups used the same incorrect SZAs in the blind analysis. Correct ZPD times and solar zenith angles were used in the reanalysis.

centered at 2558 cm^{-1} and for the N_2O υ_1 band centered at 1286 cm^{-1} the Doppler widths are sufficiently small that the true line widths at low pressure are less than the maximum spectral resolution (0.004 cm^{-1}) and can be used to deconvolve or otherwise determine the instrument line shapes. During the intercomparison period the cell spectra were used as a qualitative check on alignment by measuring the widths of N_2O absorption lines. Subsequently, spectra of HBr measured on the last day of the intercomparison period, 19 February, were analyzed to deconvolve the true ILS for each instrument.

TABLE 3. Details of microwindows and fitting parameters.

Target gas	Region no.	Window (cm^{-1})	Interfering gases fitted
HNO_3	1	868.8–870	H_2O, OCS*
HNO_3	2	872.35–875.0	H_2O, OCS*
N_2	1	2403.26–2403.86	
N_2	2	2410.86–2411.64	H_2O
N_2	3	2418.4–2418.9	
N_2O	1	2441.8–2444.4	
N_2O	2	2481.2–2482.5	
CH_4	1	2903.5–2904.2	HCl, H_2O, HDO
HCl	1	2925.8–2926	CH_4
O_3	1	3027.42–3027.6	
O_3	2	3045.08–3045.38	
HF	1	4038.78–4039.1	H_2O

* The 867.0–869.2 cm^{-1} window used in the Table Mountain intercomparison was reduced to avoid an electronic noise spike traced to the power supply and to avoid possible interference from NH_3 absorption, which can be strong in summer in Lauder (Murcray et al. 1989). The 871.8–875.8 cm^{-1} window was initially fitted in the blind analysis, but several features due to the much higher water vapor column were poorly fitted by both groups and the results are not included. This window was reduced in the reanalysis and OCS was explicitly fitted as well as water vapor.

Following the blind phase measurements, each group analyzed their own spectra without further exchanges with the other. Each group employed the same version (1.09d) of SFIT software (Rinsland et al. 1982, 1998) for the spectral analysis, nominally using an agreed set of input settings and parameters detailed in Table 2. NIWA provided the climatological pressure, temperature, and composition vertical profiles to be used in the analysis, and spectral microwindows and line parameters were agreed upon. Details of the microwindows and fitted species used are given in Table 3. In general the parameters were chosen to be the same as those used for the previous intercomparison at Table Mountain, California (Goldman et al. 1999). When the blind phase analyses were complete, both groups submitted their results to an independent referee (Griffith), who had taken no part in the measurements. The blind phase results were analyzed by the referee and the results of each group then exchanged.

Systematic differences between results from the two instruments from 0.5% (N_2O) to 7% (HF) were apparent, and are described below (section 3). These may

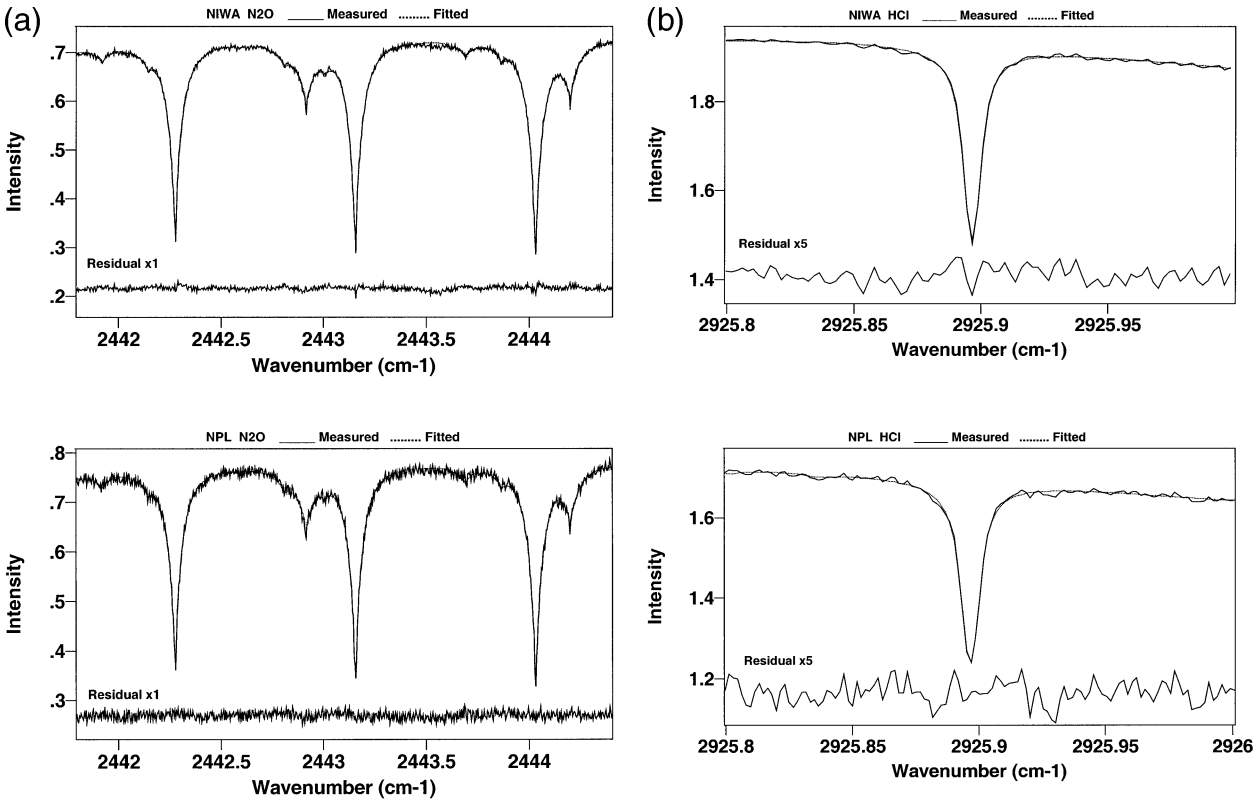

FIG. 1. Examples of measured (solid line), fitted (dotted line), and residual (lower solid line) spectra for each species. (top) NIWA and (bottom) NPL: (a) N_2O, (b) HCl, (c) O_3, (d) HF, and (e) HNO_3.

have been due to differences in the spectra as collected by the instruments, or to inadvertent differences in the analysis procedures. To identify and eliminate the latter, the two groups exchanged spectra and each analyzed a subset of the other's spectra using their own analysis procedure. Despite the fact that both groups had used the same analysis software (SFIT1.09d), several differences emerged from this exercise due to different choices of input options, including slight differences in the spectral regions chosen for the fits, differences in the way that interfering species were handled, and allowance for zero offsets. A full reanalysis of all spectra from both NIWA and NPL by one group (Jones, NIWA) was therefore carried out to eliminate any bias between groups due to the analysis. The results of this reanalysis are also presented below; any remaining differences should be due only to instrumental differences, the objective of the intercomparison.

3. Results

Figure 1 illustrates a subset of the spectral microwindows and typical fits for each species. In most cases the spectral residuals are close to the noise levels in the spectra. Some systematic residuals near the centers of sharp lines are indicative of errors in the assumed concentration profile and/or incorrect characterization of the ILS.

a. Blind intercomparison

The full results of the blind intercomparison are given in the appendix (Table A1) and the differences between the two instruments are summarized in Table 4. The following general points may be noted.

- For the predominantly tropospheric gases with pressure-broadened absorption lines (N_2O, N_2, CH_4, all retrieved from filter 3 spectra), the NIWA-retrieved columns are within 1% of NPL columns but systematically higher, except in set 5 where the NIWA columns were typically 2%–3% higher.
- For the predominantly stratospheric gases HCl, O_3, and HNO_3, NIWA columns were typically 2%–4% higher than NPL values. For the gases retrieved from filter 3 spectra (HCl, O_3), in set 5 NIWA columns exceeded NPL columns by 4%–6%.
- For HF, NIWA columns were 6%–9% higher than NPL columns in all sets.
- The NPL spectra consistently showed SNRs to be about half of those of the NIWA spectra, which we assume is traceable to the damage suffered by the NPL spectrometer during shipping. The higher noise levels

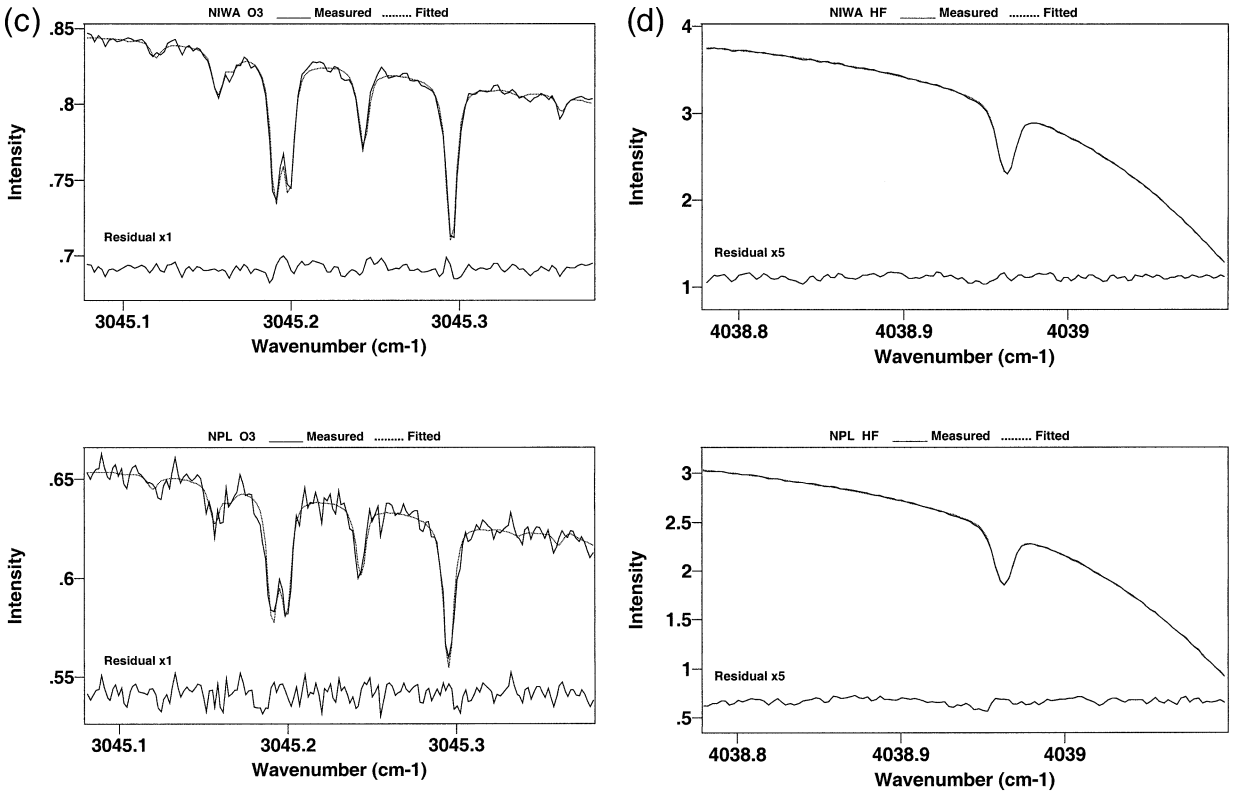

FIG. 1. (*Continued*)

are reflected in typically higher standard deviations about the mean columns within each set.

- In general the differences between the mean columns derived from each instrument are greater than the standard deviations of these means, and hence the differences between instruments are statistically significant. More detailed statistical significance testing is reported below for the consistent reanalysis of the spectra.
- The retrieved N_2 columns are about 4% higher than predicted by the airmass calculation in the forward models for an atmosphere with a uniform N_2 mixing ratio of 0.781. A similar discrepancy is also present in the data from the Table Mountain intercomparison

TABLE 4. Summary of reanalyses of all datasets.

	Window (cm^{-1})	NPL mean column[a]	Std error obs[b] (%)	Std error MALT[c] (%)	NIWA mean column[a]	Std error obs[b] (%)	Std error MALT[c] (%)	Diff[d] reanalysis (%)	Diff[d] blind (%)
N_2O-1	2441.8–2444.4	6.235E+18	0.14	0.26	6.265E+18	0.08	0.17	0.49	0.94
N_2O-2	2481.2–2482.5	6.106E+18	0.12		6.127E+18	0.09		0.34	0.89
N_2-3	2418.4–2418.9	1.714E+25	0.68		1.716E+25	0.47		0.58	1.33
CH_4	2903.5–2904.2	3.378E+19	0.21		3.415E+19	0.12		1.11	1.20
O_3-1	3027.42–3027.6	6.374E+18	0.98		6.529E+18	0.37		2.43	2.85
O_3-2	3045.08–3045.38	6.397E+18	0.79		6.560E+18	0.39		2.57	2.68
HCl	2925.8–2926	3.613E+15	0.40	1.12	3.718E+15	0.19	0.79	2.90	3.25
HNO_3	868.8–870	9.151E+15	0.59		9.411E+15	0.37		4.00	2.14
HF	4038.78–4039.1	1.009E+15	0.61	0.51	1.077E+15	0.24	0.50	6.78	7.81

[a] Average over all sets of the retrieved column (molec cm^{-2}).

[b] Average over all sets of the std error of the mean of the retrieval for each set derived from observed std dev given in Table A2 [std error = (std dev/sqrt (no. of measurements)].

[c] Average over all sets of std error of the mean derived from the MALT least squares analyses. This statistic is derived from the covariance matrix and spectral residual of the fit.

[d] Average percentage absolute difference over all sets between NPL and NIWA mean columns (NIWA/NPL-1)100 for the blind intercomparison and SFIT reanalysis of all spectra, from Tables A1 and A2.

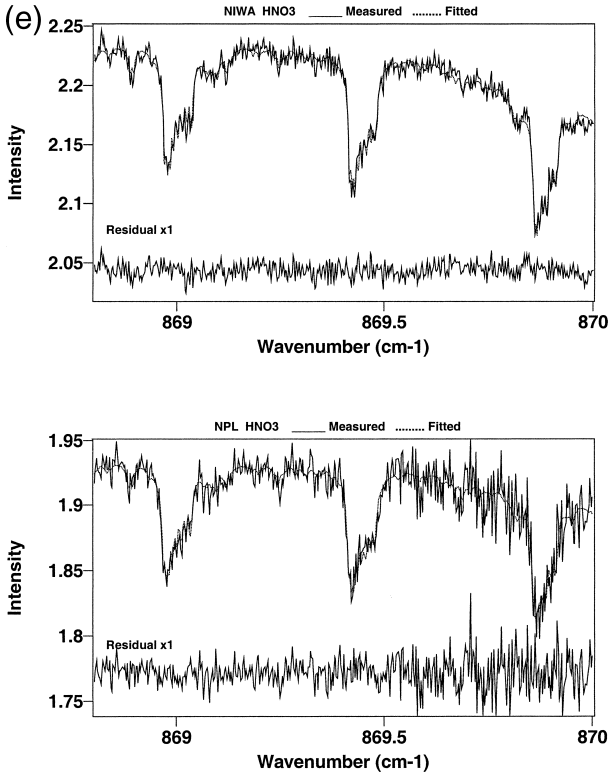

FIG. 1. (*Continued*)

(Goldman et al. 1999). The difference is currently being investigated in an intercomparison of N_2 measurements from all NDSC sites.

Differences between the two instruments' results may be due to either instrumental differences or differences in the spectral analysis procedures. The former include instrument line shape (due to different alignment), zero levels (due to detector saturation and/or phase errors near strongly absorbed bands), channel (etalon) spectra, SNR, and digital filtering of the spectra. The latter include several of the parameters used for the SFIT analysis as detailed in Table 2, such as the exact time (and hence solar zenith angle) of analysis; atmospheric layering; spectral microwindows; assumed vertical profiles of pressure, temperature, and composition; and fitting of zero levels and channel spectra. To remove all uncertainties due to the analysis procedure, all spectra from both instruments were reanalyzed by one person (Jones) using identical SFIT analysis parameters. In addition, a subset of the spectra was analysed using an entirely independent forward model and analysis procedure (MALT-CLS; Griffith 1996) and the results compared.

b. SFIT reanalysis

The full reanalysis with SFIT (version 1.09e) used the parameters defined in Table 2. Within each set there was no significant variation in total columns with time or zenith angle, and we therefore report here only the mean, standard deviation, and number of spectra for each set. The results are tabulated in the appendix (Table A2) and summarized in Table 4. Figure 2 shows the mean total column amounts retrieved for selected molecules from each set for each instrument. Between 15 and 17 February there was a significant change in atmospheric structure, with Lauder station radiosonde records showing that the tropopause height dropped from approximately 17 to 15 km. This is evident in the retrieved columns of all species, with predominantly tropospheric species (N_2O, CH_4) decreasing and stratospheric species (HCl, HF, HNO_3, O_3) increasing.

It is clear from Fig. 2 and Table 4 and the appendix (Table A2) that significant differences remain between the two instruments. Specific points to note are as follows.

- Random error may be estimated from the relative standard deviation about the mean within each set (i.e., the error bars in Fig. 2). For the NPL instrument this ranged from 0.30% (N_2O) to 5.4% (O_3), and for the NIWA instrument from 0.14% (N_2O) to 2.2% (N_2). The higher random errors for the NPL data are consistent with the observed noise levels in the spectra.
- Systematic differences between instruments remain in the reanalysis, varying from 0.12% (N_2O) to 6.8% (HF). Except for the noisiest microwindows, the mean columns retrieved from each spectrometer in each set are significantly different at the 95% confidence level; that is, there is a less than 5% chance that the observed differences were due to random error and the two instruments were in fact operating identically. Details of the significance tests are given in the appendix (Table A2).
- Definition of the exact time of the zero path difference (ZPD) crossing and hence the effective solar zenith angle for each spectrum caused considerable confusion due to different reporting of this time in different versions of the OPUS spectrometer software and individual operators' software. Correct assignment of all ZPD times removed some scatter in the results of sets 1–4, and removed the consistently larger bias found for filter 3 spectra in set 5 in the blind analysis. This correction is the major factor in reducing the differences between instruments from the blind to the reanalyses of the spectra.
- In set 5, NPL spectra were collected at 257-cm OPD (0.0039 cm^{-1} resolution) while NIWA spectra were run at the normal 180-cm OPD. The NPL spectra were retransformed to 180-cm OPD and reanalyzed; the mean difference in retrieved columns between the two resolutions was less than 0.3% in all cases.
- Correction of zero offsets in MCT spectra below 1000 cm^{-1}, typically about 3% in both instruments, increased in all columns by a similar factor as expected for weak absorption lines (Abrams et al. 1994). How-

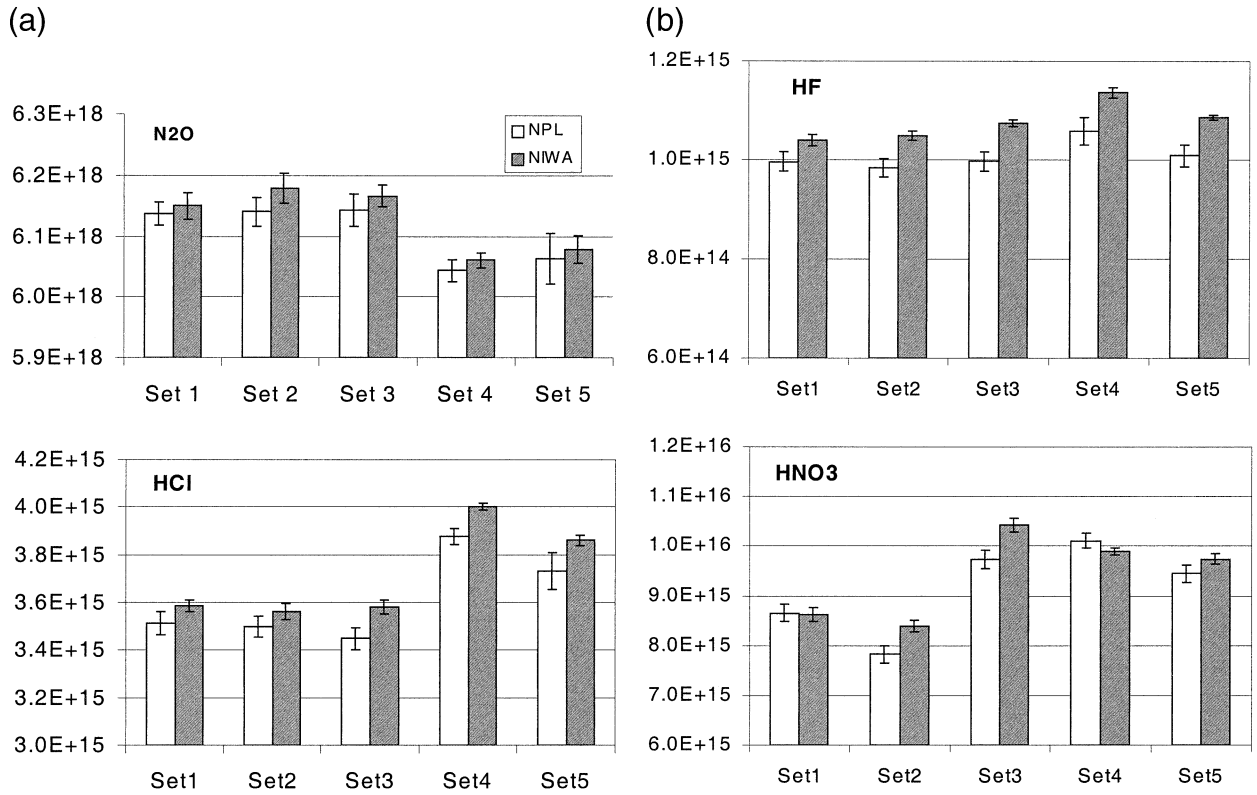

FIG. 2. Mean total column amounts determined from each instrument and set. Error bars are the std dev for the measurements in each set. A change in air mass is evident between set 3 and set 4 in all measured species.

ever, this correction did not remove the bias between instruments.

c. *MALT analyses*

Subsets of spectra from each instrument were analyzed for N_2O and HCl (50 spectra) and HF (59 spectra) by the referee using an independent forward model and linear least squares software (Griffith 1996). The same line parameters, microwindows, and vertical profiles were used in all cases. Table 5 summarizes the differences between the SFIT and MALT analyses. The absolute differences between the two algorithms were always less than 0.7%, consistent with previous rigorous intercomparisons (Zander et al. 1993). More importantly, the MALT analyses changed the differences between instruments by less than 0.2% for N_2O and HCl and by 0.7% for HF compared to SFIT. Thus the independent spectral analysis confirms the differences between instruments.

The MALT analysis also provides a statistical estimate of the standard error of each individual retrieval based on the covariance matrix of the absorption coefficients and the spectral fitting residual (Box et al. 1978). These standard errors are reported in Table 4 and can be compared with the standard errors determined from the measured standard deviations and number of measurements in each set.[1] The relative values are con-

TABLE 5. SFIT MALT intercomparison. Relative differences between mean columns for spectra analyzed by SFIT and MALT algorithms reported as (MALT/SFIT-1)100.

	NPL (%)	NIWA (%)
N_2O	+0.29	+0.34
HCl	+0.50	+0.66
HF	−0.04	+0.69

[1] Standard error = standard deviation/$\sqrt{\text{no. of measurements}}$. Note that Table A2 quotes the standard deviations and Table 4 gives the standard errors.

TABLE 6. Sensitivities of retrieved total columns to nonideal ILS parameters. Values reported as percent error relative to retrieval with the ideal ILS.

Molecule	Modulation (20%)	Misalignment (0.2 mm)	Phase error (+2°)
N_2O	−0.5	−0.1	+0.1
HCl	−3.1	−1.7	−1.2
HF	−2.5	−1.8	−1.5
O_3	−3.2	−2.0	−0.8
HNO_3-1	−0.8	−0.1	−0.1
NO_3-2	−1.2	−0.1	−1.1

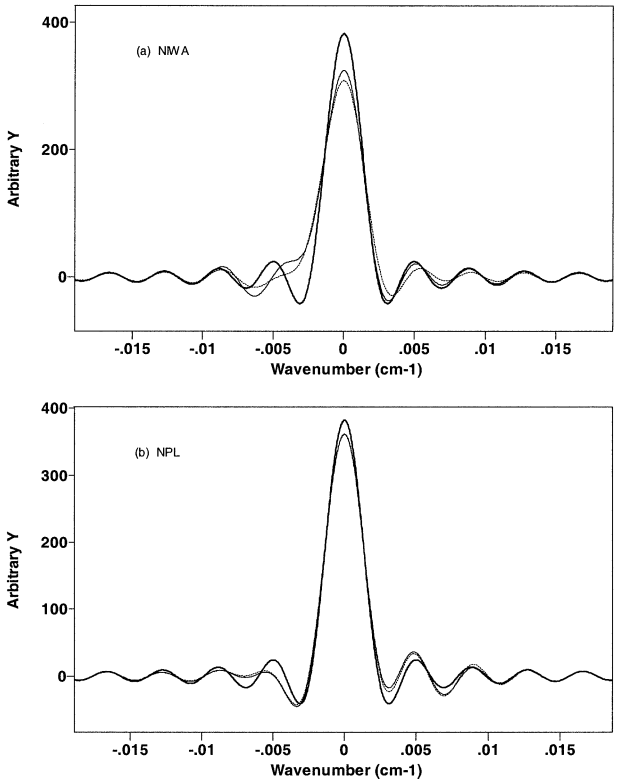

FIG. 3. ILSs for both instruments retrieved by Fourier deconvolution from narrow lines of HBr in a low pressure cell. Heavy solid line is the ideal ILS; solid line is the forward scans; dashed line is the backward scans. (a) NIWA and (b) NPL.

sistent, except for the NPL HF case, where the observed variation is dominated by a systematic bias between forward and backward interferometer scans, discussed further below. The MALT standard errors are systematically larger than the observed values because they include both a random and a systematic lack of fit of the spectra; the latter does not contribute to the observed standard deviations, but may contribute to the measured values of the mean total column.

4. Discussion

After elimination of potential biases due to spectral analysis procedures, statistically significant differences between the two instruments remain for all gases. We have made several attempts to model these instrumental differences and explain the differences in retrieved columns, with varying degrees of success.

1) Simulations of spectra with nonideal contributions to the ILS due to loss of modulation efficiency, phase errors, and optical misalignment were used to assess the sensitivity of the retrieved columns to reasonable values of these parameters.
2) The ''true'' ILS for each instrument was estimated by Fourier deconvolution from measured spectra of very narrow lines of HBr in a low pressure cell. This ILS was then used in the calculated spectra to fit the measured solar spectra in place of an assumed ideal ILS.
3) Measured spectra were empirically corrected for phase errors leading to offset and slope in observed zero levels in the centers of saturated absorption lines and refitted with ideal calculated spectra.

a. Retrieval sensitivity to instrument line shape (ILS)

The MALT forward model (Griffith 1996) has been extended to include the effects of loss of modulation efficiency (''effective apodization''), phase error, and off-axis misalignment of the interferometer collimator aperture (Kauppinen and Saarinen 1992). We have simulated the sensitivity of total column retrievals to each of these nonideal ILS parameters by calculating test spectra of the target atmospheric species with a nonideal ILS and fitting them with calculated spectra that assume an ideal ILS. The sensitivities to nonideal parameters are given in Table 6 for an indicative subset of the molecules and microwindows.

Here, N_2O is typical of tropospheric species with pressure-broadened absorption lines (N_2O, N_2, CH_4); the retrieved total columns are relatively insensitive ($<1\%$) to realistic distortions of the ILS from ideal. For predominantly stratospheric species with narrow absorption lines (HF, HCl, O_3, HNO_3) realistic distortion of the ILS may modify retrieved total columns by amounts of up to 3%, and provides a possible explanation for the differences between instruments for all gases in this study except perhaps HF. Phase error affects the retrieved total columns in two separate ways; first through the change in line shape of the actual absorption line, and second through a change in the effective continuum level at the center of the absorption line if there is a strong absorption line nearby whose wings overlap the narrow line of interest. In the latter case the phase error distorts the strong band and affects the local zero and continuum levels in the observed spectrum. This is the case for HCl (strong adjacent CH_4 band) and HF (H_2O band); the apparent sensitivity of the total column retrieval to phase error is dominated by the continuum offset effect, and may lead to either a lower or a higher estimate of retrieved total column, depending on the sign of the phase error. This is illustrated later (Fig. 4) and its correction is discussed further below.

The spectral residuals that appear when fitting with nonideal ILSs such as those above, with the marginal exception of those with phase error, have amplitudes similar to the noise level at the typical signal-to-noise ratios obtained in solar spectra and will therefore be difficult to observe. Thus the spectra themselves do not readily provide a sensitive indicator of the actual ILS, and an independent method is required. Measurements of very narrow absorption lines (Hase et al. 1999) of low pressure gases or emission lines of a suitable laser

TABLE 7. Total column retrievals with nonideal ILS. Average column retrievals of 50 spectra assuming ideal and retrieved ILS.

	Ideal (molec cm^{-2})	Retrieved (molec cm^{-2})	Error (%)
	(a) NPL		
N_cO	6.143 E18	6.178 E18	+0.6
HCl	3.572 E15	3.621 E15	+1.4
	(b) NIWA		
N_2O	6.172 E18	6.212 E18	+0.6
HCl	3.674 E15	3.725 E15	+1.5

(Vance 1998) provide two such methods; here we have applied the former and used Fourier deconvolution to retrieve the actual ILS for each instrument.

b. ILS retrieval by Fourier deconvolution

The ''true'' ILS was determined from the line shapes of several absorption lines of HBr from 2400 to 2700 cm^{-1} at about 0.3-hPa pressure in a 10-cm cell measured on the last day of the intercomparison by both spectrometers. Full details of the Fourier deconvolution procedure are reported elsewhere (Bernardo 2001): in summary, spectra of the P5–P8 (2410–2470 cm^{-1}) and R4–R7 (2648–2688 cm^{-1}) lines of both $H^{79}Br$ and $H^{81}Br$ were calculated using the MALT forward model and recent line parameters for HBr, which incorporate hyperfine structure (Coffey et al. 1998), without instrumental contribution to the line shape. These ''monochromatic'' spectra of the individual lines were then deconvolved from the measured spectra to retrieve 16 estimates of the ILS, which were averaged in two sets (P and R branch lines). Figure 3 shows the true ILS retrieved from the deconvolution procedure for each instrument as well as the theoretical, ideal line shape. The ILSs determined by Fourier deconvolution are in close agreement with those determined by the LINEFIT routine of Hase et al. (1999). These averaged ILSs were used in the forward calculation of the solar spectra in a MALT-based solar spectrum analysis and the results compared to those obtained during the blind intercomparison and reanalyses stages, which assumed an ideal ILS.

The theoretical field-of-view (FOV) contribution to the ILS is dependent on both wavenumber and collimator aperture size, and the retrieved ILS must therefore be adjusted to the spectral region and FOV of the solar measurements. This was approximated by deconvolving the theoretical FOV contribution at the measured wavenumber (a sinc function in interferogram space with width proportional to frequency and FOV^2) and reconvolving it with the theoretical FOV contribution appropriate to the spectral region of interest. The adjustment was made based on the ILS determined from R branch lines of HBr for HCl spectra near 2930 cm^{-1} and from P branch lines of HBr for N_2O spectra near 2440 cm^{-1}. In addition, since the HBr spectra were inadvertently recorded on the NIWA instrument with a collimator aperture of 0.65 mm but solar spectra were recorded with a 0.5-mm aperture, the ILS determined for a 0.65-mm aperture was adjusted to 0.5 mm in a similar manner.

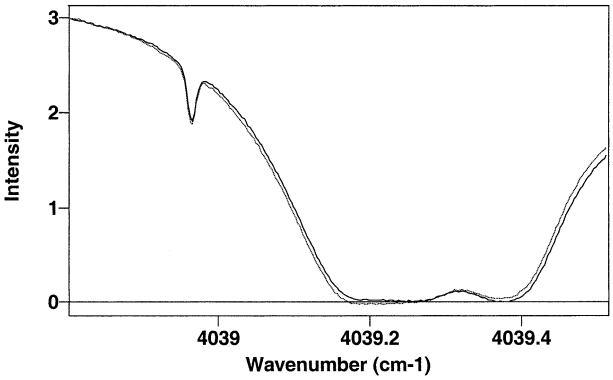

FIG. 4. Spectra in the HF analysis region from two consecutive scans of the NPL interferometer, illustrating the forward–backward bias. Solid line, forward mirror scan; dotted line, backward mirror scan.

Table 7 compares the mean columns determined in a reanalysis using the retrieved nonideal ILSs with those for an assumed ideal ILS. The test set consisted of the 50 spectra analyzed in the MALT–SFIT intercomparison. In summary, retrieved total columns increased by 0.6% for N_2O and by 1.4%–1.5% for HCl when the adjusted, retrieved ILS was used in the forward calculation. There was no significant difference in the spectral residuals between ILSs. The changes are consistent with the sensitivity studies above, but of very similar magnitude for both instruments; thus, the adoption of an ILS determined in this way is not able to account for the differences in total columns for the two instruments.

This analysis was not applied to HF: extrapolation of the deconvolved ILS to the 4000 cm^{-1} region resulted in worse fits to the spectra and was considered too unreliable. To address this problem we attempted to fit the retrieved ILSs by nonlinear least squares in terms of parameters for modulation efficiency (or ''effective apodization''), phase error, and off-axis misalignment of the collimator aperture. Using these fundamental parameters, the theoretical ILS could in principle be calculated at any aperture and wavenumber. However, these attempts were unsuccessful; fitting solar spectra with the parameterized ILS in the forward model resulted in poorer fits in all cases than those with the ideal or retrieved ILS; it appears that the parameters used to characterize the ILS are either not independent enough or do not accurately describe the true ILS, so that they do not provide a unique and accurate description of the true ILS. For example, if the sun tracker does not uniformly illuminate the entrance aperture of the spectrometer collimator, or there are significant ghosts in the measured

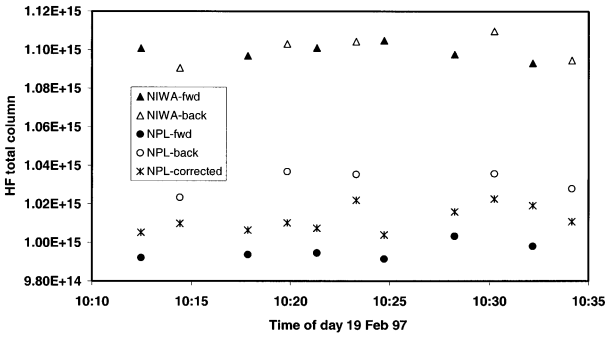

FIG. 5. Effect of the time delay correction to NPL spectra on HF retrievals for set 5. In the NPL uncorrected data, the higher values are all from backward scan spectra, and the lower values from forward scans.

spectra (e.g., Guelachvili 1981; Learner et al. 1996), this parameterization cannot be expected to be reliable.

c. Empirical correction of measured spectra

An ideal spectrum should show zero transmission in the center of a totally absorbing spectral line. As mentioned above, uncorrected phase error will cause the center of such a line to show nonzero transmission, as illustrated in Fig. 4, which shows the region around the HF absorption line at 4038.96 cm^{-1} used in the analysis. There is a strong water vapor absorption feature centered near 4039.22 cm^{-1}. Figure 4a shows typical spectra from consecutive forward and backward scans of the NPL instrument, where the center of the water feature has a positive or negative slope near zero. This sloping zero effect can be modeled quite accurately by a phase error of around $\pm 1.5°$ for the forward and backward scans, respectively. From Table 6, such phase errors will lead to retrieved errors in the HF total column of $\pm 1.2\%$, respectively, predominantly due to the $\pm 1.5\%$ change in the local continuum level around the HF line. The NIWA spectrometer showed insignificant differences between forward and backward scans of the interferometer. The HCl absorption at 2925.9 cm^{-1} also has a strong neighboring band, in this case due to CH_4. Spectra from both spectrometers show similar but much smaller effects to those in Fig. 4a on zero levels in this region.

The measured ILSs cannot explain this zero effect, which required a phase error that changes sign between forward and backward scans, such as an electronic time delay. The ILS for the NPL instrument cannot be dominated by this type of phase error, for this would cause the asymmetry in the ILS to reverse sign between forward and backward scans (positive and negative phase errors, respectively). The NIWA ILS (Fig. 3) is broadened to a lower wavenumber and is therefore consistent with a slight misalignment of the collimating mirror with respect to the optic axis (Kauppinen and Saarinen 1992). However, the NPL ILS is broadened to higher wavenumber, which cannot be explained in the same way. We have not been able explain the actual physical cause of the NPL line shape; it may also be due to the damage to beamsplitters suffered by the instrument during shipping.

The effects of zero offsets can be corrected in two ways; either the spectra can be adjusted to have correct zero levels then analyzed in the normal way or the phase error can be included in the calculated spectra used to fit the measured spectra. As might be expected, both approaches provide similar results, and we describe here an approach to the former. One plausible cause of the forward–backward asymmetry would be phase error introduced by the time delay between the time of laser fringe crossings used to trigger the sampling and the actual time of sampling of the detector signal. While the detector preamplifier and electronics should ideally compensate for such a delay, the preamplifier gain is increased by a factor of 8 after 1000 laser fringes. This gain switch will result in a different time delay, and the change is not normally compensated by the data system. Such a change would have an opposite effect on phase errors during forward and backward scans of the interferometer mirror. We have simulated this effect in the following way:

- recalculate the phase-corrected interferogram by Fourier transformation of the original spectrum,
- resample the interferogram with a selectable time delay after the gain switch position,
- combine the original interferogram points before the gain switch with the resampled interferogram points after the gain switch to obtain the time-delay-corrected interferogram, and
- retransform the time-delay-corrected interferogram to obtain the corrected spectrum.

Realistic time delay corrections of typically 0.2–0.5 μs are sufficient to correct all spectra to have acceptable zero levels.

As a test of the procedure, we corrected all NPL filter 1 spectra in set 5 by this means, reanalyzed the corrected spectra using MALT and an assumed ideal ILS, and compared the analyses with those of uncorrected spectra. Figure 5 shows the results from all set 5 spectra for NIWA together with those from corrected and uncorrected NPL spectra. The forward–backward bias in the NPL spectra is much reduced, but the mean NPL column for this set changes only from 1.014×10^{15} to 1.012×10^{15} molec cm^{-2}, compared to the mean NIWA column of 1.100×10^{15} molec cm^{-2}. Thus, while effectively removing the forward–backward bias, the time delay correction does not have any significant effect on the NIWA–NPL bias. The time delay correction would be less significant in the lower-wavenumber regions used for the analyses of other species because the zero effects are much less, and we conclude that this effect also cannot explain the NIWA–NPL bias.

5. Conclusions

After consistent reanalysis of all spectra, eliminating any bias due to the spectral analysis method, small differences remain in the total columns retrieved by the two instruments in this study. For predominantly tropospheric gases with pressure-broadened spectral lines, such as N_2O and CH_4, the differences between retrieved columns were typically less than 1%. For predominantly stratospheric gases, such as HCl and O_3, differences were less than 3%. In most cases, the differences were greater than the scatter in the individual measurements, and significant at the 95% confidence level. The worst case observed was for HF, which showed a 7% systematic bias between instruments. After elimination of ''avoidable'' errors due for example to logging of data collection times and solar zenith angles, we have shown that the analysis algorithm is only a minor source of error in the whole measurement procedure.

The magnitudes of these differences, except for the case of HF, are consistent with errors that might be expected due to small misalignments and other imperfections in the two nominally similar spectrometers. However, attempts to correct the analyses for such imperfections, based on empirical corrections for real instrument line shapes, zero offsets, channel spectra, and apparent uncorrected phase errors, tended to affect both instruments' results equally and did not remove the apparent biases. While the differences between results from two such collocated instruments must ultimately be due to the instruments themselves, it appears that the methods we have used to try to characterize these differences are currently inadequate. Instrumental imperfections should mostly be apparent in the observed instrument line shapes, and we find two independent methods to determine real ILS functions based on narrow lines from HBr cell spectra both provide very similar ILSs. We must ask if this method is not sufficiently realiable to provide the instrumental corrections required to correct for errors at this level. The extrapolation of the measured ILS in the region of the HBr absorption lines to other spectral regions may be one limitation. Unless the true contributions to the real, imperfect ILS are well understood, this extrapolation may be significantly in error.

Except for the case of HF, these biases are also consistent with those seen in earlier intercomparisons (Goldman et al. 1999; Paton Walsh et al. 1997), and we conclude that, for the time being, the existence of systematic biases of up to 3% must be considered when assessing the intercomparability of total column measurements from different instruments. This intercomparison probes only biases due directly to the instruments themselves; other sources of systematic error due to different locations and assumed atmospheric structure are additional and discussed by other authors (e.g., Paton Walsh et al. 1997). In the context of using the NDSC database of total column FTIR measurements in model studies and other applications, this level of error must be taken into account.

Acknowledgments. This work has been supported by the New Zealand Foundation for Science and Research through Contract C01523 and the U.K. Department of Environment, Transport and the Regions through Contract EPG 1/1/17.

REFERENCES

Abrams, M. C., G. C. Toon, and R. A. Schindler, 1994: Practical example of the correction of Fourier-transform spectra for detector nonlinearity. *Appl. Opt.,* **33,** 6307–6314.

Bernardo, C., 2001: Measurement of instrument line shape functions of high-resolution FTIR spectrometers and their application to the analysis of spectra. Ph.D. thesis, University of Wollongong.

Box, G. E. P., W. G. Hunter, and J. S. Hunter, 1978: *Statistics for Experimenters: An Introduction to Design, Data Analysis and Model Building.* John Wiley and Sons, 653 pp.

Coffey, M. T., A. Goldman, J. W. Hannigan, W. G. Mankin, W. G. Schoenfeld, C. P. Rinsland, C. Bernardo, and D. W. T. Griffith, 1998: Improved vibration–rotation (0–1) HBr line parameters for validating high resolution infrared atmospheric spectra measurements. *J. Quant. Spectros. Radiat. Transfer,* **60,** 863–867.

Goldman, A., 1996: Intercomparison of analysis of ground-based IR synthetic spectra. *Proc. Network for the Detection of Stratospheric Change Infrared Working Group Meeting,* Garmisch-Partenkirschen, Germany.

——, and Coauthors, 1999: Network for Detection of Stratospheric Change Fourier transform infrared intercomparison at Table Mountain Facility, November 1996. *J. Geophys. Res.,* **104,** 30 481–30 503.

Griffith, D. W. T., 1996: Synthetic calibration and quantitative analysis of gas phase infrared spectra. *Appl. Spectros.,* **50,** 59–70.

Guelachvili, G., 1981: Distortions in Fourier spectra and diagnosis. *Spectrometric Techniques,* G. A. Vanasse, Ed., Academic Press, 1–53.

Hase, F., T. Blumenstock, and C. Paton-Walsh, 1999: Analysis of the intrumental line shape of high resolution FTIR spectrometers with gas cell measurements and new retrieval software. *Appl. Opt.,* **38,** 3417–3422.

Kauppinen, J., and P. Saarinen, 1992: Line shape distortions in misaligned cube corner interferometers. *Appl. Opt.,* **31,** 69–74.

Kurylo, M. J., 1991: Network for the Detection of Stratospheric Change. *Proc. Soc. Photo-opt. Instrum. Eng.,* **1491,** 169–174.

Learner, R. C. M., A. P. Thorne, and J. W. Brault, 1996: Ghosts and artifacts in Fourier transform spectroscopy. *Appl. Opt.,* **35,** 2947–2954.

Murcray, F. J., A. Matthews, A. Goldman, P. Johnston, and C. Rinsland, 1989: NH_3 column abundances over Lauder, New Zealand. *J. Geophys. Res.,* **94,** 2235–2238.

Paton Walsh, C., and Coauthors, 1997: An uncertainty budget for ground-based Fourier transform infrared column measurements of HCl, HF, N_2O and HNO_3 deduced from results of side-by-side instrument intercomparisons. *J. Geophys. Res.,* **102,** 8863–8873.

Rinsland, C. P., A. Goldman, F. J. Murcray, D. G. Murcray, M. A. H. Smith, J. R. K. Seals, J. C. Larsen, and P. L. Rinsland, 1982: Stratospheric N_2O mixing ratio profile from high resolution balloon-borne solar absorption spectra and laboratory spectra near 1880 cm^{-1}. *Appl. Opt.,* **21,** 4351–4355.

——, and Coauthors, 1998: Northern and Southern Hemisphere ground-based infrared spectroscopic measurements of tropospheric carbon monoxide and ethane. *J. Geophys. Res.,* **103,** 28 197–28 217.

Vance, A., 1998: Fourier transform instrument line shape determi-

nation using an infrared helium neon laser and a low pressure gas cell. Polar stratospheric ozone: 1997 European Commission Air Pollution Res. Rep. 66, 735–738.

Zander, R., 1995: IR retrieval algorithms intercomparison for the NDSC. *Extended Abstracts, Fourier Transform Spectroscopy: New Methods and Applications*, Santa Fe, NM, 104–106.

——, P. Demoulin, E. Mathieu, G. P. Adrian, C. P. Rinsland, and A. Goldman, 1993: ESMOS II/NDSC spectral fitting algorithms intercomparison exercise. *Proc. Atmospheric Spectroscopy Applications Workshop*, Riems, France, 7–12.

——, and Coauthors, 1998: An overview of NDSC-related activities at the Jungfraujoch through high resolution infrared solar observations. *Proc. 18th Quadrennial Ozone Symp.*, L'Aquila, Italy.

APPENDIX

Results of Blind Intercomparisons

TABLE A1. Total columns are means of each set, rsd is the relative standard deviation within set (%), no. is the number of spectra in each set, Diff is the percent difference between instruments (NIWA/NPL-1)100 for each window. Mean differences (boldface) over sets are means of the *absolute* differences.

		NPL mean (molec cm^{-2})	Rsd (%)	No.	NIWA mean (molec cm^{-2})	Rsd (%)	No.	Diff (%)
N_2O-1	Set 1	6.09E+18	0.49	11	6.11E+18	0.32	11	0.29
	Set 2	6.15E+18	0.49	12	6.20E+18	0.45	12	0.76
	Set 3	6.08E+18	0.49	13	6.11E+18	0.40	14	0.47
	Set 4	5.97E+18	0.34	7	6.01E+18	0.21	12	0.71
	Set 5	5.81E+18	1.89	21	5.95E+18	0.80	24	2.48
	Mean	**6.02E+18**	**0.74**		**6.08E+18**	**0.43**		**0.94**
N_2O-2	Set 1	6.23E+18	0.48	11	6.25E+18	0.36	11	0.38
	Set 2	6.29E+18	0.32	12	6.32E+18	0.37	12	0.42
	Set 3	6.21E+18	0.48	13	6.23E+18	0.76	14	0.31
	Set 4	6.11E+18	0.98	7	6.15E+18	0.19	12	0.73
	Set 5	5.94E+18	2.02	21	6.10E+18	0.78	24	2.63
	Mean	**6.16E+18**	**0.86**		**6.21E+18**	**0.49**		**0.89**
N_2-1	Set 1	1.67E+25	4.19	11	1.66E+25	1.82	11	−0.89
	Set 2	1.72E+25	4.07	12	1.65E+25	3.12	12	−3.88
	Set 3	1.69E+25	4.14	13	1.80E+25	10.32	14	6.25
	Set 4	1.60E+25	3.75	2	1.63E+25	2.92	12	1.74
	Set 5				1.61E+25	4.03	24	
	Mean	**1.67E+25**	**4.04**		**1.67E+25**	**4.44**		**3.19**
N_2-2	Set 1	1.65E+25	10.30	11	1.64E+25	5.81	11	−0.62
	Set 2	1.67E+25	10.78	12	1.70E+25	3.93	12	1.76
	Set 3	1.65E+25	6.06	13	1.73E+25	4.00	14	5.03
	Set 4	1.62E+25	4.94	2	1.73E+25	2.36	12	7.06
	Set 5	1.64E+25	6.10	21	1.65E+25	3.55	24	0.64
	Mean	**1.65E+25**	**7.64**		**1.69E+25**	**3.93**		**3.02**
N_2-3	Set 1	1.68E+25	3.57	11	1.71E+25	0.94	11	1.74
	Set 2	1.74E+25	3.45	12	1.73E+25	2.28	12	−0.60
	Set 3	1.69E+25	2.96	13	1.71E+25	2.15	14	1.31
	Set 4	1.70E+25	1.76	7	1.71E+25	0.89	12	0.53
	Set 5	1.63E+25	3.07	21	1.67E+25	1.43	24	2.47
	Mean	**1.69E+25**	**2.96**		**1.71E+25**	**1.54**		**1.33**
CH_4	Set 1	3.36E+19	0.60	11	3.38E+19	0.38	11	0.55
	Set 2	3.40E+19	0.88	12	3.45E+19	0.61	12	1.48
	Set 3	3.35E+19	0.60	13	3.36E+19	1.68	14	0.33
	Set 4	3.28E+19	2.74	8	3.30E+19	0.37	12	0.70
	Set 5	3.19E+19	1.88	21	3.28E+19	1.07	24	2.93
	Mean	**3.32E+19**	**1.34**		**3.36E+19**	**0.82**		**1.20**

TABLE A1. (*Continued*)

		NPL mean (molec cm^{-2})	Rsd (%)	No.	NIWA mean (molec cm^{-2})	Rsd (%)	No.	Diff (%)
O_3-1	Set 1	6.35E+18	2.05	11	6.47E+18	1.00	11	1.88
	Set 2	6.45E+18	1.86	12	6.63E+18	1.27	12	2.84
	Set 3	6.18E+18	1.62	13	6.32E+18	2.64	14	2.32
	Set 4	6.44E+18	2.95	7	6.65E+18	1.84	12	3.29
	Set 5	6.32E+18	6.65	21	6.57E+18	2.15	24	3.91
	Mean	**6.35E+18**	**3.02**		**6.53E+18**	**1.78**		**2.85**
O_3-2	Set 1	6.43E+18	4.04	11	6.48E+18	1.46	11	0.83
	Set 2	6.49E+18	4.47	12	6.63E+18	1.59	12	2.13
	Set 3	6.30E+18	3.02	13	6.42E+18	1.32	14	2.03
	Set 4	6.41E+18	4.06	7	6.60E+18	1.72	12	2.90
	Set 5	6.15E+18	3.09	21	6.49E+18	1.59	24	5.50
	Mean	**6.36E+18**	**3.73**		**6.52E+18**	**1.54**		**2.68**
HCl	Set 1	3.48E+15	1.44	11	3.56E+15	0.70	11	2.33
	Set 2	3.50E+15	1.43	12	3.57E+15	0.90	12	2.07
	Set 3	3.41E+15	1.47	13	3.48E+15	1.65	8	2.02
	Set 4	3.83E+15	0.78	8	3.98E+15	0.69	8	3.89
	Set 5	3.57E+15	3.36	21	3.78E+15	0.93	21	5.91
	Mean	**3.56E+15**	**1.70**		**3.67E+15**	**0.97**		**3.25**
HNO_3-1	Set 1	7.13E+15	6.17	10	7.39E+15	1.57	10	3.64
	Set 2	7.23E+15	7.47	12	7.42E+15	2.10	12	2.69
	Set 3	8.88E+15	6.19	12	9.07E+15	2.31	14	2.13
	Set 4	8.53E+15	3.63	12	8.38E+15	2.59	12	−1.80
	Set 5	8.05E+15	2.36	12	8.25E+15	1.4	14	2.50
	Mean	**7.96E+15**	**5.17**		**8.10E+15**	**2.02**		**2.14**
HNO_3-2	Set 1	7.05E+15	1.42	10	7.25E+15	2.21	10	2.85
	Set 2	6.98E+15	1.72	12	7.36E+15	2.12	12	5.42
	Set 3	8.00E+15	2.75	12	8.58E+15	1.81	14	7.30
	Set 4	7.73E+15	2.72	12	7.70E+15	3.40	12	−0.36
	Set 5	7.86E+15	0.89	12	8.00E+15	1.83	14	1.72
	Mean	**7.52E+15**	**1.90**		**7.78E+15**	**2.27**		**3.53**
HF	Set 1	9.89E+14	1.52	12	1.047E+15	1.06	12	5.92
	Set 2	9.87E+14	1.42	12	1.055E+15	0.81	12	6.86
	Set 3	9.86E+14	1.42	12	1.070E+15	0.62	12	8.57
	Set 4	1.052E+15	2.09	11	1.143E+15	0.81	12	8.67
	Set 5	9.99E+14	1.50	14	1.089E+15	0.89	15	9.02
	Mean	**1.003E+15**	**1.59**		**1.081E+15**	**0.84**		**7.81**

TABLE A2. Total columns are means of each set, rsd is the relative standard deviation within set (%), no. is the number of spectra in each set, Diff is the percent difference between instruments (NIWA/NPL-1)100. Blind is the percent difference from blind phase from Table A1. Mean differences (boldface) over sets are means of *absolute* differences.

		NPL mean[a] (molec cm^{-2})	Rsd (%)	No.	NIWA mean (molec cm^{-2})	Rsd (%)	No.	Diff[b] (%)	Blind[b] (%)	*T* test[c]		*F* test[d]	
N_2O_1	Set 1	6.270E+18	0.42	11	6.295E+18	0.33	11	0.40	0.29	0.02	x	0.47	
	Set 2	6.270E+18	0.30	12	6.296E+18	0.31	12	0.42	0.76	0.00	x	0.91	
	Set 3	6.260E+18	0.40	11	6.316E+18	0.31	7	0.89	0.47	0.00	x	0.63	
	Set 4	6.180E+18	0.92	6	6.198E+18	0.14	8	0.30	0.71	0.13		0.00	x
	Set 5	6.195E+18	0.42	21	6.222E+18	0.34	21	0.43	2.48	0.00	x	0.34	
	Mean	**6.235E+18**	**0.49**		**6.265E+18**	**0.28**		**0.49**	***0.94***				
N_2O_2	Set 1	6.138E+18	0.31	11	6.149E+18	0.35	11	0.19	0.38	0.19		0.71	
	Set 2	6.140E+18	0.38	12	6.179E+18	0.39	12	0.62	0.42	0.00	x	0.89	
	Set 3	6.143E+18	0.42	10	6.166E+18	0.29	7	0.38	0.31	0.04	x	0.35	
	Set 4	6.044E+18	0.30	7	6.060E+18	0.19	8	0.27	0.73	0.53		0.00	x
	Set 5	6.064E+18	0.68	21	6.079E+18	0.38	21	0.25	2.63	0.15		0.01	x
	Mean	**6.106E+18**	**0.42**		**6.127E+18**	**0.32**		**0.34**	***0.89***				
N_2_3	Set 1	1.698E+25	3.30	11	1.722E+25	0.98	11	1.40	1.74	0.24		0.19	
	Set 2	1.733E+25	3.21	12	1.724E+25	2.24	12	−0.50	−0.60	0.66		0.24	
	Set 3	1.704E+25	2.65	11	1.709E+25	1.89	8	0.33	1.31	0.77		0.39	
	Set 4	1.718E+25	1.88	6	1.726E+25	1.10	8	0.49	0.53	0.29		0.01	x
	Set 5	1.703E+25	1.79	21	1.705E+25	1.32	21	0.16	2.47	0.74		0.18	
	Mean	**1.714E+25**	**2.39**		**1.716E+25**	**1.64**		**0.58**	***1.33***				
CH_4	Set 1	3.411E+19	0.46	11	3.426E+19	0.34	11	0.44	0.55	0.02	x	0.34	
	Set 2	3.413E+19	0.73	12	3.465E+19	0.57	12	1.52	1.48	0.00	x	0.43	
	Set 3	3.399E+19	0.41	11	3.460E+19	0.42	8	1.79	0.33	0.00	x	0.85	
	Set 4	3.311E+19	0.90	6	3.348E+19	0.17	8	1.12	0.70	0.00	x	0.00	x
	Set 5	3.355E+19	1.22	21	3.377E+19	0.63	21	0.67	2.93	0.03	x	0.01	x
	Mean	**3.378E+19**	**0.74**		**3.415E+19**	**0.43**		**1.11**	***1.20***				
O_3_1	Set 1	6.404E+18	3.94	11	6.466E+18	1.44	11	0.98	1.88	0.45		0.00	x
	Set 2	6.412E+18	4.26	12	6.551E+18	1.56	12	2.17	2.84	0.11		0.00	x
	Set 3	6.274E+18	2.06	6	6.528E+18	1.13	7	4.04	2.32	0.00	x	0.22	
	Set 4	6.396E+18	4.21	6	6.544E+18	0.94	8	2.32	3.29	0.17		0.00	x
	Set 5	6.385E+18	2.66	21	6.554E+18	1.33	21	2.65	3.91	0.00	x	0.00	x
	Mean	**6.374E+18**	**3.43**		**6.529E+18**	**1.28**		**2.43**	***2.85***				
O_3_2	Set 1	6.372E+18	2.02	11	6.451E+18	0.98	11	1.25	0.83	0.08		0.03	x
	Set 2	6.394E+18	1.77	12	6.552E+18	1.31	12	2.46	2.13	0.00	x	0.37	
	Set 3	6.207E+18	1.62	11	6.496E+18	1.19	7	4.65	2.03	0.00	x	0.53	
	Set 4	6.464E+18	2.90	6	6.650E+18	1.42	8	2.87	2.90	0.55		0.00	x
	Set 5	6.547E+18	5.42	21	6.652E+18	1.93	21	1.60	5.50	0.21		0.00	x
	Mean	**6.397E+18**	**2.77**		**6.560E+18**	**1.37**		**2.57**	***2.68***				
HCl	Set 1	3.512E+15	1.35	11	3.586E+15	0.64	11	2.11	2.33	0.00	x	0.03	x
	Set 2	3.497E+15	1.28	12	3.562E+15	0.96	12	1.86	2.07	0.00	x	0.39	
	Set 3	3.448E+15	1.33	11	3.579E+15	0.84	8	3.80	2.02	0.00	x	0.28	
	Set 4	3.874E+15	0.87	6	4.002E+15	0.35	8	3.29	3.89	0.00	x	0.04	x
	Set 5	3.731E+15	2.07	21	3.860E+15	0.57	21	3.46	5.91	0.00	x	0.00	x
	Mean	**3.613E+15**	**1.38**		**3.718E+15**	**0.66**		**2.90**	***3.24***				

TABLE A2. (*Continued*)

		NPL mean[a] (molec cm^{-2})	Rsd (%)	No.	NIWA mean (molec cm^{-2})	Rsd (%)	No.	Diff[b] (%)	Blind[b] (%)	*T* test[c]		*F* test[d]	
HNO_3-1	Set 1	8.650E+15	2.00	10	8.610E+15	1.60	10	−0.47	3.64	0.57		0.50	
	Set 2	7.822E+15	2.25	12	8.384E+15	1.30	12	7.19	2.69	0.00	x	0.13	
	Set 3	9.727E+15	1.99	12	1.043E+16	1.30	12	7.21	2.13	0.00	x	0.25	
	Set 4	1.011E+16	1.47	8	9.897E+15	0.74	8	−2.09	−1.80	0.00	x	0.08	
	Set 5	9.448E+15	1.86	11	9.737E+15	1.02	11	3.06	2.50	0.00	x	0.08	
	Mean	**9.151E+15**	**1.89**		**9.411E+15**	**1.18**		**4.0**	***2.14***				
HF	Set 1	9.962E+14	2.03	12	1.040E+15	1.10	12	4.44	5.92	0.00	x	0.07	
	Set 2	9.829E+14	1.93	12	1.049E+15	0.96	12	6.67	6.86	0.00	x	0.05	x
	Set 3	9.969E+14	1.95	12	1.074E+15	0.64	12	7.75	8.57	0.00	x	0.00	x
	Set 4	1.059E+15	2.62	12	1.136E+15	0.97	12	7.29	8.67	0.00	x	0.00	x
	Set 5	1.009E+15	2.15	11	1.087E+15	0.47	11	7.74	9.02	0.00	x	0.00	x
	Mean	**1.009E+15**	**2.14**		**1.077E+15**	**0.83**		**6.78**	***7.81***				

[a] Mean column (molec cm^{-2}), relative standard deviation within each set, and number of spectra within each set, for NIWA and NPL spectra.
[b] Difference between NPL and NIWA mean columns (NIWA/NPL-1)100 for the blind intercomparison and SFIT reanalysis of all spectra.
[c] The *t* test is the probability that the NPL and NIWA mean measured columns are derived from distributions with the same mean. Rows marked with an ''x'' are significantly different at the 5% probability level.
[d] The *F* test is the probability that the NPL and NIWA measurements are derived from distributions with the same variance. Rows marked with an ''x'' are significantly different at the 5% probability level.