

An Intercomparison of Airborne Nitrogen Dioxide Instruments

G. L. GREGORY,¹ J. M. HOELL, JR.,¹ M. A. CARROLL,² B. A. RIDLEY,³ D. D. DAVIS,⁴
J. BRADSHAW,⁴ M. O. RODGERS,⁴ S. T. SANDHOLM,⁴ H. I. SCHIFF,⁵
D. R. HASTIE,⁵ D. R. KARECKI,⁶ G. I. MACKAY,⁶ G. W. HARRIS,⁶
A. L. TORRES,⁷ AND A. FRIED^{8,9}

Results from an airborne intercomparison of techniques to measure tropospheric levels of nitrogen dioxide (NO₂) are discussed. The intercomparison was part of the National Aeronautics and Space Administration's Global Tropospheric Experiment and was conducted during the summer of 1986. Instruments intercompared included a two-photon nitric oxide (NO) laser-induced fluorescence system with laser photolysis of NO₂ to NO, an NO/O₃ chemiluminescence detector using FeSO₄ for conversion of NO₂ to NO, an NO/O₃ chemiluminescence detector with arc lamp photolysis of NO₂ to NO, and a tunable diode laser multipath absorption system. All intercomparisons were for NO₂ mixing ratios of <200 pptv with most at mixing ratios of <100 pptv. The FeSO₄ converter was found to convert peroxyacetyl nitrate (PAN) to NO, resulting in NO₂ values a factor of 2–3 higher than reported by the other techniques. Thus the FeSO₄ converter data are excluded from the analyses. Intercomparison data were analyzed in various mixing ratio ranges. Good correlation was observed between the remaining three instruments for those data sets which included mixing ratios to 100 or 200 pptv, showing on the average a 30–40% level of agreement among the techniques. However, when the data were restricted to mixing ratios of <50 pptv, little correlation among the measurements was observed. Even though correlations were poor at mixing ratios of <50 pptv, the tunable diode laser system tended to be high compared to data reported by the two-photon laser and arc lamp chemiluminescence systems, and agreement between the latter two instruments was generally better than 20 pptv with an equal tendency for one to be high relative to the other.

INTRODUCTION

As part of the National Aeronautics and Space Administration (NASA) Tropospheric Chemistry Program, a series of field intercomparisons have been initiated to evaluate state-of-the-art capability for measuring key tropospheric species [McNeal *et al.*, 1983; Beck *et al.*, 1987]. These intercomparisons, designated as Chemical Instrumentation Test and Evaluation (CITE), are conducted as part of NASA's Global Tropospheric Experiment (GTE). This paper reports the results from CITE 2, when airborne measurements from four NO₂ instruments were intercompared. The objectives of CITE 2 were (1) to evaluate instrumentation for measuring NO₂, HNO₃, and PAN in remote environments and (2) to determine the relative abundances of these major odd-nitrogen species for various tropospheric environments. CITE 1, an intercomparison of carbon monoxide, nitric oxide, and hydroxyl radical instrumentation, demonstrated the existing capability for aircraft measurements of NO in remote environments [Hoell *et al.*, 1987]. Results from CITE

2 evaluate the measurement capability for other nitrogen species believed to play a major role in daytime tropospheric chemistry. Results of the HNO₃ and PAN intercomparisons are addressed in companion papers.

INSTRUMENTATION/INTERCOMPARISON PROTOCOL

NO₂ Instrumentation

CITE 2 instrumentation for the nitrogen dioxide intercomparisons included four techniques, each fundamentally different in detection principle: (1) a two-photon NO laser-induced fluorescence laser system with laser photolysis of NO₂ to NO operated by Georgia Institute of Technology (GIT/LIF), (2) an NO/O₃ chemiluminescence detector using FeSO₄ for conversion of NO₂ to NO operated by NASA/Wallops Flight Facility (Wallops/Chem), (3) an NO/O₃ chemiluminescence detector with arc lamp photolysis of NO₂ to NO operated jointly by NOAA and NCAR (NOCAR/Chem), and (4) a tunable diode laser multipath absorption system operated by York University (York/TDLAS). It was acknowledged that the York/TDLAS instrument with a detection limit of about 25 pptv would not be able to make measurements in some tropospheric environments. It was desirable to include the tunable laser instrument in CITE 2, since it spectroscopically measures NO₂ directly.

The laser-induced fluorescence system (GIT/LIF) uses a photofragmentation/two-photon detection scheme [Bradshaw and Davis, 1982; Bradshaw *et al.*, 1985; Davis *et al.*, 1987] to detect, simultaneously, NO and NO₂. Air samples from independent aircraft inlets are provided to each of two detection cells. In the first cell the ambient NO concentration is measured, while in the second, ambient NO plus NO produced from photolysis of NO₂ is detected. In the latter case the NO₂ is photolyzed by the passage of a 353-nm XeF excimer laser beam through the sampled air stream. Depending upon atmospheric conditions, the degree of photodecom-

¹Atmospheric Sciences Division, NASA Langley Research Center, Hampton, Virginia.

²National Oceanic and Atmospheric Administration, Aeronomy Laboratory, Boulder, Colorado.

³National Center for Atmospheric Research, Boulder, Colorado.

⁴School of Geophysical Sciences, Georgia Institute of Technology, Atlanta.

⁵Chemistry Department, York University, North York, Ontario, Canada.

⁶Unisearch Associates, Concord, Ontario, Canada.

⁷NASA Wallops Flight Facility, Wallops Island, Virginia.

⁸Center for Analytical Chemistry, National Institute of Standards and Technology, Gaithersburg, Maryland.

⁹Now at National Center for Atmospheric Research, Boulder, Colorado.

Copyright 1990 by the American Geophysical Union.

Paper number 89JD02943.
0148-0227/90/89JD-02943\$05.00

position of NO_2 is between 50 and 90%. Ambient NO_2 levels are determined by subtracting the simultaneous ambient NO signal from the total NO signal from the NO_2 cell (after factoring in the NO_2 photodecomposition efficiency). In both cells the NO species is detected using two-photon laser-induced fluorescence. The resulting blue-shifted UV fluorescence from excited NO molecules is monitored using a solar blind photomultiplier tube (PMT) with a photon-counting detection system. The signal levels from the ambient NO cell and the second cell (NO + NO from photolyzed NO_2) are related to the ambient atmospheric mixing ratios of each species by standard addition calibration with NO and NO_2 , respectively. As operated during CITE 2, sample integration time was either 1 or 2 min with the latter being more prevalent at the low ambient NO_x concentrations. Data submitted for intercomparison were for a 6-min integration time in which three of the above 2-min data were averaged. For some flights, data were submitted in triplicate and included 2-, 6-, and 10-min integration periods; however, only the 6-min data are intercompared in these discussions. Instrument accuracy (1σ) was stated at about $\pm 16\%$ (8% transfer measurement precision plus 8% estimated systematic error). The precision obtainable with the GIT/LIF approach is dependent upon the integration time as well as the ambient levels of NO and NO_2 species. Under clean air conditions, and ambient mixing ratios of 15 pptv of NO and 50 pptv of NO_2 , the measurement precision for NO_2 (6-min integration) was about 17% (1σ). For the same NO and NO_2 levels of 15 and 50 pptv (i.e., NO_2 to NO ratio of 3.33), estimated total uncertainty (2σ) was of the order of ± 20 pptv (67%), ± 23 pptv (46%), and ± 31 pptv (31%) for NO_2 measurements at 30, 50, and 100 pptv, respectively. Sampling interferences due to thermal decomposition of pernitric acid [Ridley *et al.*, 1988a] or other NO_y species are minimized by temperature control of the sampling lines and by short (<1.5 s) residence times within the sampling system. Photolytic interferences are estimated to contribute less than 1 pptv to the measurements. The reader is referred to Sandholm *et al.* [this issue] for a detailed discussion of the technique and the stated accuracy, precision, and total uncertainty values.

The chemiluminescence FeSO_4 system (Wallops/Chem) uses the conventional NO/O_3 chemiluminescent reaction [Clough and Thrush, 1967; Fontijn *et al.*, 1970] to detect ambient NO and NO resulting from conversion of NO_2 . A single-sample inlet is used to supply ambient air to the reaction cell in which alternate detection of ambient NO and ambient NO plus NO from converted NO_2 are performed. Output from the chemiluminescent reaction of $\text{O}_3 + \text{NO}$ is detected by a PMT using photon-counting techniques [Torres, 1985]. For the NO_2 measurement the air sample is passed through a humidified, crystalline, FeSO_4 converter in which NO_2 is converted to NO. The converter is housed in a 100-cm³ Teflon-lined chamber. Based on gas-phase titration experiments during CITE 2, the NO_2 conversion efficiency is about 96%. The ambient NO_2 concentration is determined from the difference of the ambient NO measurement and the measurement of ambient NO plus NO from converted NO_2 . In this case the two signals are not simultaneously obtained. The instrument sample cycle as used during CITE 2 was of the order of 7 min, consisting of 1 min of NO detection, 2 min of NO_2 (NO plus converted NO_2), and the remaining 4 min in some combination of determining

baseline (zero) and/or switching between the measurement and zero modes. Sample integration time for photon counting was 1 s. For intercomparison purposes the data were averaged into 1-min periods. For levels of NO_2 below 100 pptv, measurement precision (1σ) for a 1-min average is about 6 pptv. For NO_2 greater than 100 pptv the precision is of the order of 10%. At the higher mixing ratios the total uncertainty of the technique for NO_2 , in the absence of interferences or artifacts from the converter, is about 20% (2σ) plus an additional 4 pptv due to uncertainties in the instrument offset. Thus at 30, 50, and 100 pptv, total uncertainty is 53, 28, and 24%, respectively. Calibration is by dynamic gas dilution of NO_2 from a permeation system.

The arc lamp/broadband photolytic NO_2 converter system (NOCAR/Chem) also uses the NO/O_3 chemiluminescent reaction to detect ambient NO and the NO originating from converted NO_2 [Kley and McFarland, 1980]. Two separate chemiluminescent detectors are used for the measurement—one dedicated to ambient NO detection and the second to detection of ambient NO plus NO from photolyzed NO_2 . Sample flows (3 L/min) are humidified prior to entering the reaction chambers to eliminate background variability and suppress ozone-related background signals. The basic integration period for each detector is 10 s, and for intercomparison purposes, six 10-s values were accumulated and reported as 1-min averages. Both chemiluminescent detectors are operated in a cyclic fashion consisting of modes for zero, measure, calibrate, and artifact check [Ridley *et al.*, 1987]. Instrument cycle time for CITE 2 was of the order of 30 min, during which five to six contiguous 1-min averages were reported. For NO_2 detection the air sample is irradiated with a 300-W xenon arc lamp to photodissociate NO_2 to $\text{NO} + \text{O}(^3\text{P})$ [Kley and McFarland, 1980]. The wavelength employed for photodissociation of NO_2 is ≥ 320 nm, and thus it is possible that NO_2 arising from homogeneous or heterogeneous reactions involving HO_2NO_2 (daytime), NO_3 (nighttime), and N_2O_5 (nighttime) may be present. Therefore the photolytic converter is normally operated at 10°C to minimize the thermal decomposition of these species. Calibration is by standard addition of NO_2 derived from a back titration of a NO standard. For the 1-min data reported, instrument precision (2σ) was estimated to be of the order of ± 10 pptv, systematic error at $\pm 6\%$, and offset uncertainty at $\pm 50\%$ of the observed instrument artifact. For CITE 2 the observed NO_2 instrument artifact was about 4.4 pptv. Thus for a 1-min average at NO_2 mixing ratios of 30, 50, and 100 pptv the overall uncertainty (2σ) is estimated to be ± 14 pptv (47%), ± 15 pptv (30%), and ± 18 pptv (18%), respectively. A detailed description of the operation and performance of the instrument is given by Ridley *et al.* [1988a].

The tunable diode laser system (York/TDLAS) utilizes infrared absorption to detect NO_2 [Hastie *et al.*, 1983; Schiff *et al.*, 1987]. The radiation source is a Pb salt semiconductor laser operated at cryogenic temperatures. The TDLAS uses a multipass absorption cell through which the air sample is drawn. The inlet system was constructed from or lined with Teflon. The NO_2 absorption feature employed during CITE 2 was centered at approximately 1600 cm^{-1} . The change in radiation transmitted through the cell is proportional to the concentration of NO_2 in the air sample. Data were reported as 2- to 3-min averages. Minimum detection limit for a 150-s integration period is estimated to be about 25 pptv and is based on laboratory tests. Detection limits vary slightly

depending on the sampling environment. The total uncertainty (1σ) for mixing ratios well above its 25-pptv detection limit (e.g., 100 pptv) was estimated at 25%. At 50 pptv the total uncertainty is of the order of 45%. Calibration was by gas dilution from a NO_2 permeation source. The same instrument was also used for detection of HNO_3 as part of those intercomparisons. For a given flight the instrument was dedicated to the detection of either NO_2 or HNO_3 . Details of the instrument and its operations are given by Schiff *et al.* [this issue].

The experimental layout of the NO_2 instrumentation as well as other supporting instrumentation aboard the NASA Wallops Electra aircraft is discussed by Hoell *et al.* [this issue].

Data Protocol

The CITE 2 activities included intercomparison of measurements of project-supplied calibration standards (ground-based but on board the aircraft) as well as in-flight intercomparisons of ambient measurements. All intercomparisons were blind, conducted with no exchange of information between the investigator teams before or after submittal of their results. Results from the tests were submitted to the GTE project office during the field operations and were analyzed by project personnel to monitor progress of the tests and to provide inputs into subsequent tests. The ground-based standards and flight tests were, in effect, conducted as separate intercomparisons with the various standards tests generally taking place on nonflight days. For the standards intercomparisons a reference calibration system from the National Bureau of Standards (NBS), now the National Institute of Standards and Technology, provided the NO_2 test gas mixtures. The gas mixtures were provided to individual instruments or subsets of instruments on different days depending upon project schedules and/or instrument readiness.

Only a qualitative assessment of the progress and results of the flight intercomparisons were provided to the investigators in the field. However, the data protocol adopted by the CITE 2 science team did provide for release of more specific details of the results if discrepancies were observed, which, in the opinion of the project staff, required disclosure for continuation of meaningful tests. The release of specific information was subject to (1) documentation of the accuracy/precision of the previously submitted data and (2) agreement that any changes between the preliminary and final data as a result of the additional information would be documented and reported. As will be discussed later, this protocol was activated twice, once during the standards tests (GIT/LIF-NBS initial intercomparison) and once during the flight tests (Wallops/Chem). The data protocol also provided for a full disclosure of standards intercomparison results while in the field, but after completion of the tests and submittal of final standards data. The data protocol did not allow for disclosure of flight results at the conclusion of the field work. Those results remained blind until final data had been submitted and a preliminary intercomparison analysis was performed by the project (approximately 6 months after the field work). It is recognized (and within the protocol guidelines) that the final and some of the in-the-field flight data were submitted knowing the results from the standards tests. However, and as agreed to by the investigators, any

changes between preliminary data submitted in the field and final data required an explanation and are documented. The flight data were not normalized using results from the ground-based standards intercomparisons.

GROUND-BASED STANDARDS INTERCOMPARISONS

Description of Reference System/Operational Procedures

The ground-based standards intercomparison was carried out by having each instrument sample from the output of an NBS reference NO_2 calibration system aboard the aircraft. The NO_2 reference source was designed, constructed, and evaluated at NBS specifically for the CITE 2 tests [Fried *et al.*, this issue]. The methodology of the tests was based upon providing to each instrument, as installed on the aircraft, known mixing ratios of NO_2 . This methodology inherently couples uncertainties associated with the measurement technique as well as those associated with its calibration procedure. Accordingly, where differences occur, it is not possible to partition them between instrumental or calibration sources. However, in the absence of differences the methodology provides a strong indication of agreement between calibration sources.

The reference source utilized an NO_2 permeation wafer contained in a temperature-controlled oven with a carrier gas continuously flowing through the oven followed by two stages of dilution. Ultrahigh-purity zero cylinder air was employed for the carrier and dilution flows. The output from the final dilution stage was directed through a 3-m length of PFA Teflon tubing and into a 2.5-cm (ID) by 30-cm-long pyrex manifold containing four 0.6 cm (ID) sampling ports along its length. Each instrument obtained a sample from the reference source at different times using a Teflon tube attached to one of the ports. In each case the Teflon sample line was provided by the respective investigators. Depending upon the required flow rate, one or more of the unused ports were vented to maintain the sampling manifold at atmospheric pressure. The system was designed to deliver a maximum flow rate of 12 standard liters per minute (slpm) with mixing ratios that could be varied over a nominal range from 0.4 to 70 ppbv with an accuracy of $\pm 10\%$. With this output the sample flow rates used in-flight by most of the instruments could be accommodated. The reference calibration system was tested at NBS to characterize the output and to define a set of operational procedures that could be used on the aircraft. A detailed description of the reference system, characterization tests, and intercomparison procedures are provided in a companion paper [Fried *et al.*, this issue].

The ground-based standards intercomparison tests were conducted after integration and checkout of each instrument aboard the aircraft. All intercomparisons were conducted on the aircraft. The procedures adopted for the intercomparisons involved moving the reference system aboard the aircraft and as close as possible to the instrument to be tested. The NOCAR/Chem and York/TDLAS tests were always performed on the same day. For these tests the reference system was placed equidistant between the two instruments to allow for switching test instruments without disturbing the operation of the calibration system. This necessitated the previously mentioned 3-m length of PFA

TABLE 1. Nitrogen Dioxide Standard Results

Date	NBS	GIT/LIF	WFF/Chem	NOCAR/Chem	York/TDLAS	Percent Difference, ^a %
Aug. 4	1079 ± 98 1045 ± 111 ^b 1050 ± 95		1042	1056	1035 ^c	-3.4 +1.1 -1.4
Aug. 14	669 ± 66 390 ± 40 676 ± 70 396 ± 37			686 ^d 650 ^e 384 ^d 348 ^c	642 ^{c,d} 614 ^{c,f} 357 ^{c,d} 329 ^{c,f}	+2.5 -2.8 -1.5 -10.8 -5.1 -9.2 -9.8 -16.8
Aug. 18	690 ± 68 397 ± 43		722 409			+4.6 +3.0
Aug. 22	732 429	1720 ^g 928 ^g				+135 +116
Aug. 24	672 ± 64 672 ± 64	814 ^h	737			+21.1 +9.7

Mixing ratios are in parts per trillion by volume.

^aDifference = [Principal Investigator - NBS] * 100/NBS.

^bReflects a +2.9% postmission change to reflect the average of 2 flows (before and after the intercomparison) rather than the final flow as originally used.

^cReflects a +10.7% postmission change in calibration as the result of a recalibration of the NO₂ permeation source.

^dMixing ratio with no calibration system "zero" subtraction.

^eMixing ratio with a 36-pptv calibration system "zero" subtraction.

^fMixing ratio with a 28-pptv calibration system "zero" subtraction.

^gBefore correcting GIT/LIF procedural error.

^hAfter correcting GIT/LIF procedural error.

tubing which was employed in all tests. After each move the reference system was checked for leaks, flow rates were allowed to stabilize, and the 3-m transfer line/sampling manifold was equilibrated for 30 min to 1 hour before intercomparisons were carried out. Flow calibrations were conducted before and after each intercomparison and for each set of new flow operating conditions. In addition and between calibration runs at different NO₂ mixing ratios for the same instrument the system, transfer line, and manifold were allowed to equilibrate for 10–30 min. In general and as necessitated by the various equilibration times, the reference calibration system output was sampled for several hours by the NO₂ instrument being tested. Upon completion of an intercomparison series for each instrument the reference system "zero" was sampled by switching the NO₂ permeant flow out of the sample stream. After completion of tests with a given NO₂ instrument the reference system was either removed from the aircraft or moved to a second instrument location for additional intercomparisons.

In reporting the test data to the project, each investigator was encouraged to report the average value of the NO₂ sampled, the reference system NO₂ zero, and any available information as to observed variations of the NO₂ mixing ratio with time. As discussed by *Fried et al.* [this issue], the NBS reference NO₂ mixing ratios were based upon the permeation source emission rate (determined by three independent techniques) and the various system flow rates. The calculated reference mixing ratios and those reported by the investigators established the level of agreement between the instrument and the reference system.

Results From the Standards Intercomparison

Table 1 summarizes the results from the standards intercomparison tests. Nominal mixing ratios (NBS system) ranged from about 400 to 1000 pptv. The first intercomparisons took place on August 4 at Wallops Flight Facility (WFF), prior to any flight intercomparisons, and included tests of Wallops/Chem, NOCAR/Chem, and York/TDLAS. The remaining tests took place at Ames Research Center (ARC) on August 14, 18, 22, and 24. The uncertainties given with each NBS mixing ratio represent the total calibration system uncertainty estimated by quadrature addition of the various systematic and random errors as discussed by *Fried et al.* [this issue]. Investigators' uncertainties are not given in the table but are of the order of 10–15%. As the result of the relatively long equilibration and stabilization times associated with reference system transport and mixing ratio changes, there was a limited time available during the ARC intercomparisons for sampling of the NO₂ background and zero air from the calibration system. This was particularly true for the tests of August 14, where both York/TDLAS and NOCAR/Chem were intercompared at two different mixing ratios. For these tests, both investigator groups felt some reservations in quoting NO₂ values which had been corrected for system zero (i.e., zero air) and thus reported total NO₂ and a measured reference system zero of 28 and 36 pptv, respectively. If the above two zero values truly represent the background NO₂ from the system, including zero air, then they should be subtracted from the individual results for intercomparison with the NBS sample. If, on the

other hand, the above values are predominantly a result of NO_2 outgassing from the system, sample tube, and sample manifold associated with inadequate time to sample system zero, then the total reported NO_2 mixing ratios should be directly compared with the NBS reference without subtraction. Thus at the 400- and 700-pptv levels sampled, this potentially causes an additional 5–9% uncertainty when comparing these results with the NBS reference. However, and since the above two system zero values are in reasonable agreement with one another and are of the approximate magnitude measured during some of the other tests (18, 24, and 26 pptv), the former scenario is the more probable, suggesting that the true results may be closer to the lower values (i.e., subtracting the zero value). These values are footnoted *e* or *f* in the table. Our conclusions are based on this assumption. For completeness, Table 1 does show the results using both considerations. For the other intercomparisons the investigator groups subtracted a system zero determination in presenting their results.

As noted in Table 1, the values represent the original data submitted by each group in a blind fashion with two exceptions. The NBS reference mixing ratios during the NOCAR/Chem intercomparison of August 4 was changed by +2.9% to reflect what NBS felt, after the fact, was a more representative second-stage permeation flow during this test. This flow dropped by 5.8% during the NOCAR/Chem intercomparison, and the final lower value was initially used. The NBS result of 1045 pptv reported in Table 1 is based upon the average of the initial and final flow determinations during this test. York/TDLAS investigators, postmission and with the knowledge of the results from the standards, performed a reanalysis of their permeation emission rate and found that the source emission rate of their perm tube was 10.7% higher than the premission value. Subsequently corrected results, which are reported in Table 1, were submitted. The corrected data result in an improved agreement with the reference standards. The final flight data also reflected this change but were submitted not knowing the effect on the in-flight agreement, since flight intercomparisons remained blind until after final submittal of all investigators' data. It should be mentioned, however, that a correction of only half this amount (5.4% and obtained by averaging the premission and postmission permeation rates) quite possibly would have been made if there were no prior knowledge of the magnitude and direction of the bias.

As noted by the last column of Table 1, with the exception of the GIT/LIF results, the agreement between the various instruments and the reference standard was generally well within 10% (one York/TDLAS value of –17%). In the case of GIT/LIF the August 22 results showed disagreement of greater than a factor of 2. After discussions with the GIT/LIF investigator team concerning test procedures used, and certification that the field-submitted GIT/LIF values were final, the results from the tests were revealed to the GIT/LIF investigators. This was done while the NBS reference system was still connected to the instrument with hopes that the disagreement could be quickly resolved. However, the factor of 2 disagreement was not resolved, and no additional tests were performed on August 22. A second GIT/LIF standards test was scheduled for the afternoon and evening of August 24 following a scheduled night flight for August 23/24. The GIT/LIF investigator team was requested to review their standards and procedures used in hopes of

identifying the cause for the discrepancy. The project also reviewed its procedures and decided to request Wallops/Chem to participate in the August 24 tests to add a third “independent” measurement of the NO_2 mixing ratios provided by the reference system. The source of the discrepancy was found by the GIT/LIF investigator team, and agreement between the reference standard, GIT/LIF, and Wallops/Chem during this second test was within 21%. The August 24 tests were conducted blind with respect to both GIT/LIF and Wallops/Chem. An explanation of the source of the discrepancy as provided postmission by the GIT/LIF investigator team is paraphrased below:

On August 24, shortly before the second standards test was implemented, it was discovered that the gas flow from the GIT/LIF NO_2 standard tank (used in the August 22 tests) was substantially below the minimum flow rate previously established as necessary for reliable calibration results. The latter flow criteria had been determined while the instrument was being tested in the laboratory prior to CITE 2 field activities. The August 24 tests were conducted with the NO_2 standard tank flow increased by approximately a factor of 10. The GIT/LIF investigators attribute the poor agreement observed during the August 22 intercomparison to the flow setting of their calibration system. They state that this conclusion is reinforced by airborne calibration tests performed by GIT/LIF during the intercomparison flight of August 23/24. The analyses of these data, completed during the day of August 24, revealed a discrepancy between the NO_2 calibration derived from NO_2 standard addition and known photolysis efficiency of NO_2 to NO and that derived from NO calibration in the NO_2 cell. The discrepancy was found to be of similar magnitude to that of the August 22 standards results. On all flights after correcting the tank flow rate, these two analyses produced NO and NO_2 calibration results which agreed to about 10%. As a result of the calibration problem, all flight data prior to August 25 are based on calibration factors derived from NO calibration and NO_2 photolytic efficiency, whereas data after August 25 are slightly more accurate (16% compared to 26%) and based on calibration factors derived from standard addition of NO_2 .

Standards Intercomparison Conclusions

From the ground-based standards intercomparisons one concludes that with the exception of two intercomparisons (York/TDLAS nominal 400-pptv test of August 14 and GIT/LIF 700-pptv test of August 24) the reported values from the instruments are within the uncertainty stated for the NBS reference mixing ratios. While the York/TDLAS 400 pptv can be brought into agreement with the NBS reference by not applying the system zero subtraction (earlier discussion), the indicated –16.8% is still within the stated York/TDLAS accuracy, since the instrument's stated uncertainty $\pm 15\%$ overlaps the NBS reference uncertainty. The same is true for the GIT/LIF 700-pptv intercomparison because their $\pm 16\%$ uncertainty overlaps the indicated NBS uncertainty. Thus one concludes that the investigators' (after correcting the GIT/LIF error) measurements do agree with the NBS value to within their stated uncertainties. If one is willing to normalize results based on the NBS values, the results suggest that the GIT/LIF internal standard may be high relative to the others, the York/TDLAS standard may be low compared to the others, and disagreements as large as about

TABLE 2. CITE 2 Flight Summary

Date	Flight Number	Type of Flight	Takeoff, GMT	Landing, GMT	Altitudes, km	NO ₂ Range, pptv
Aug. 11*	4	ferry, day	1257	1714	4.8	10–240
Aug. 11*	5	ferry, day	1811	2301	4.8	20–140
Aug. 15*	6	oceanic, day	1824	2335	4.8	5–480
Aug. 19	7	oceanic, day	1718	2238	4.8, 0.15	0–180
Aug. 21	8	oceanic, day	1721	2214	4.8	0–50
Aug. 24	9	continental, night	0349	0847	4.8	20–200
Aug. 26†	10	oceanic, day	1725	2230	4.8, 0.9, 0.15	0–40
Aug. 28†	11	continental, day	1728	2315	4.7, 3.8, 2.6	0–100
Aug. 30	12	oceanic, day	1723	2227	1.0–5.0	0–70
Aug. 31	13	continental, day	1744	2253	4.8	5–70
Sept. 2	14	continental, day	1700	2224	6.1, 4.6, 2.3	0–240
Sept. 5†	15	ferry, day	1431	1911	5.2	5–50
Sept. 5†	16	ferry, day	2040	0035	5.2	0–530

*Denotes that LIF was not operational for nitrogen dioxide.

†Denotes that TDLAS was configured for nitric acid.

20% among the instruments during flight may be the result of internal standard differences. Finally, the fact that an NO₂ calibration by GIT/LIF was initially in error by about a factor of 2, but that it was not identified by the investigator team as having a potential error even when it was made known that agreement with the NBS reference was not good, cannot go unnoted. This experience suggests the importance of continual attention to instrumentation calibration for field work, particularly when comparison with independent standards is not a part of the data protocol.

IN-FLIGHT INTERCOMPARISONS

Sixteen flights were conducted as part of the CITE 2 program. The first three were shakedown flights based from WFF, Virginia, and no intercomparison data were requested from these flights. The remaining 13 flights were intercomparison flights, including the four transit flights (two each way) between WFF and ARC, California. The York/TDLAS instrument was configured for NO₂ in nine of the 13 intercomparison flights. The GIT/LIF instrument was not operational for NO₂ during the transit flights to and the first flight from ARC. Measurements were reported by Wallops/Chem and NOCAR/Chem for each of the 13 intercomparison flights.

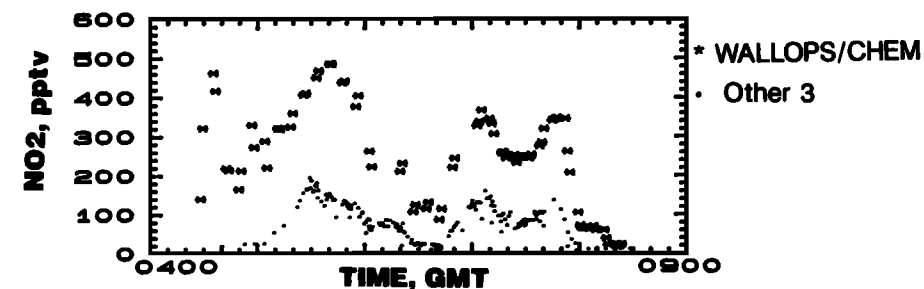
The nine flights based from ARC sampled a variety of air masses including tropical and nontropical, maritime, and continental. Flight altitudes ranged from 150 to 5000 m above sea level. All but one flight was made during daylight hours. Table 2 summarizes the flights, the study areas, flight times and altitudes, and the nominal range of NO₂ values reported. An asterisk with the mission date indicates those flights where the GIT/LIF instrument was not operational; a dagger indicates flights where the York/TDLAS instrument was configured for nitric acid. Details of the flights, types of air masses, and meteorological analyses are discussed by Shipham *et al.* [this issue].

Typical Flight Data

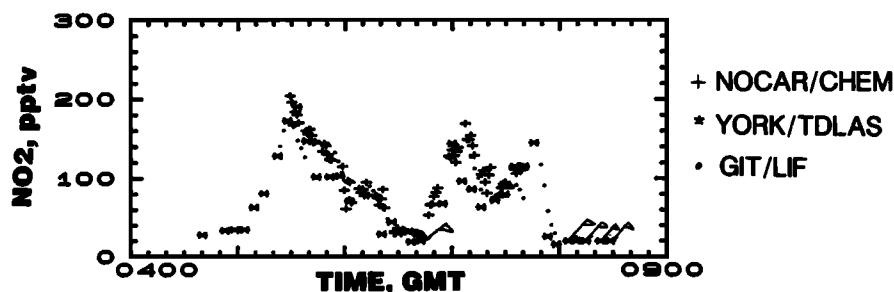
Figures 1 and 2 illustrate typical nitrogen dioxide flight data as reported by the investigators. Figure 1 shows measurements from the night flight, while Figure 2 shows measurements from one of the oceanic flights. For the Ames-

based flights, GMT time is local plus 7 hours. The upper panels of each figure show data from all four instruments, while the lower panels show the same results but with the Wallops/Chem data omitted. In the upper panels, only the Wallops/Chem data are plotted with a different symbol. The lower panels have an expanded ordinate scale and separate symbols for the various measurements.

The results shown on these two figures illustrate several general characteristics of the flight data. First, a comparison of the data in the upper panels suggest that the Wallops/Chem data are consistently higher (factor of 2 or 3) than those reported by the other investigators. This was the case for about 90% of the Wallops/Chem data. This observation, combined with subsequent studies by the Wallops/Chem investigator, indicated that the FeSO₄ converter is not specific to NO₂. This is verified by recently published results [Fehsenfeld *et al.*, 1987; Ridley *et al.*, 1988b] which indicate that the FeSO₄ converter can convert nitrogen species other than NO₂ to NO and hence introduce significant artifacts to the measurement. Accordingly, the Wallops/Chem data are not included in the intercomparison analyses. These data are briefly discussed in Appendix A, in which attention is focused on identifying correlations between the data and other CITE 2 measurements in an effort to identify interfering species for the FeSO₄ converter used by instrument. Second, GIT/LIF, NOCAR/Chem, and York/TDLAS tended to track changes in NO₂ as illustrated in Figure 1b at about 0500 GMT. Wallops/Chem also tracked the same trends, even though the reported values for NO₂ were much higher than the other instruments. Finally, the GIT/LIF and York/TDLAS investigators reported some data as upper limits (referred to here as less-than values). These data are flagged in the lower panels of Figures 1 and 2 and are plotted at half value (i.e., less than 25 pptv is interpreted as 12.5 ± 12.5 pptv). Wallops/Chem and NOCAR/Chem did not report data as less-than; however, NOCAR/Chem did report a 1σ with each NO₂ value. At mixing ratios below about 10 pptv this 1σ was always equal to or greater than the reported NO₂ value. NOCAR/Chem also reported a few negative NO₂ values (−1 to −5 pptv), indicating that at times the NO₂ was below detection to the point that corrections for offsets and artifacts exceeded the NO₂ related signal. Negative values are also flagged in the figures.



a. All 4 instruments

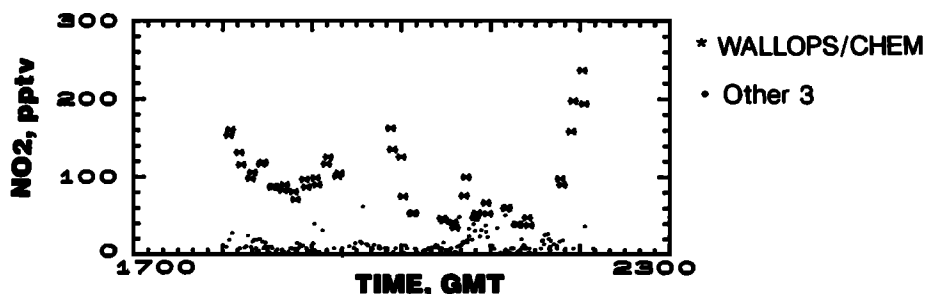


b. WALLOPS/CHEM data omitted

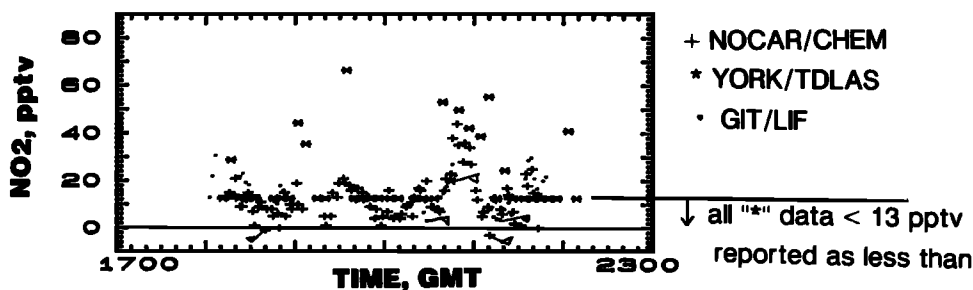
Fig. 1. Flight data for the night flight of August 24, 1986.

One additional group of NO_2 data has been omitted from the intercomparisons. During the CITE 2 flights, various investigator groups conducted inlet tests to verify the potential effects of HO_2NO_2 on the measurement of NO_2 [Ridley

et al., 1989]. These tests generally consisted of varying inlet temperature and/or flow rate and observation of any noted effects. Data obtained during these tests or during preparation for these tests (e.g., inlet heating or cooling) are omitted



a. All instruments



b. WALLOPS/CHEM data omitted

Fig. 2. Flight data for the flight of August 30, 1986.

TABLE 3. Summary of Overlapped Data Bases: Nitrogen Dioxide

Date	Flight Number	Number of Overlapped Data Periods			
		Data Base 1	Data Base 2	Data Base 3	Data Base 4
Aug. 11	4	NA	NA	NA	8
Aug. 11	5	NA	NA	NA	13
Aug. 15	6	NA	NA	NA	9
Aug. 19	7	0	2	1	5
Aug. 21	8	3 (1)*	4 (1)	3 (1)	3
Aug. 24	9	10	14	12	14 (1)
Aug. 26	10	NA	18 (3)	NA	NA
Aug. 28	11	NA	14 (1)	NA	NA
Aug. 30	12	14 (14)	14 (2)	15 (15)	31 (23)
Aug. 31	13	4 (1)	7	4 (1)	6 (1)
Sept. 2	14	13 (3)	13 (1)	15 (3)	19 (3)
Sept. 5	15	NA	7	NA	NA
Sept. 5	16	NA	4 (2)	NA	NA
Total CITE 2 overlaps		44	97	50	108
Overlaps excluding less-than data		25	87	30	79

Data base 1 was constructed by determining overlaps among GIT/LIF, NOCAR/Chem, and York/TDLAS instruments. Data base 2 was constructed by determining overlaps between GIT/LIF and NOCAR/Chem instruments. Data base 3 was constructed by determining overlaps between GIT/LIF and York/TDLAS instruments. Data base 4 was constructed by determining overlaps between NOCAR/Chem and York/TDLAS instruments. NA means not appropriate because one or more of the instruments for the data base was not operational.

*Number in parentheses indicates the number of overlap periods excluded from analyses as the result of one of more of the instruments reporting NO₂ as an upper limit or negative value.

from the intercomparison analyses. Appendix B briefly discusses some of these test results and observations.

Intercomparison Data

Table 3 summarizes the intercomparison data bases that were constructed from the measurements reported. Data base 1 contains measurements reported by GIT/LIF, NOCAR/Chem, and York/TDLAS; data base 2 from GIT/LIF and NOCAR/Chem; data base 3 from GIT/LIF and York/TDLAS; and data base 4 from NOCAR/Chem and York/TDLAS. In each case the data bases were obtained by defining a "simultaneous" measurement as one having an overlap between any two portions of the sample period reported by the respective investigators. For each data base the instrument with the longest integration time was used to initiate the overlapping data period, and as such only a single measurement from that instrument is used for the overlap period. Where more than one value of NO₂ was reported by any one of the remaining instruments during the overlap, the average of those measurements was used as the intercomparison value. Table 3 data entries are on a flight-by-flight basis with the first numerical entry representing the number of total overlaps obtained for a given flight. The number in parentheses indicates how many of the total overlaps have been excluded from the intercomparison as the result of including one or more values reported as less-than or negative. For example, for data base 1 and flight 12, all of the 14 overlap periods included some form of less-than or negative data. Thus all are excluded in the following quantitative analyses.

Analysis Procedures

Three methods of analysis were employed to evaluate the level of agreement among the measurements contained in the

data bases. The first method, referred to as "delta," is based upon the numerical difference between the investigators' values for each overlap period. The delta approach is useful for identifying constant biases among instruments. A disadvantage of the delta approach is that it does not normalize instrument agreement with respect to the NO₂ mixing ratio at which the observation was made. Accordingly, a delta of 20 pptv at 100 pptv must be interpreted differently from a delta of 20 pptv at 600 pptv. Accordingly, a second analysis approach which evaluates the percentage difference between overlapping measurements is also included. The percent difference is evaluated as

$$\% \text{ diff} = 100 * (X_1 - X_2) / \text{ave} \quad (1)$$

where X_i is the NO₂ mixing ratio reported by any one of the investigators during the overlap period and ave is the numerical average of the NO₂ mixing ratios reported during the overlap period. The % diff parameter normalizes the level of agreement to the average NO₂ mixing ratio at which the measurements were made and more readily is compared to an instrument uncertainty stated in percent. A disadvantage associated with the % diff parameter is it tends to lose significance as the detection limit of an instrument is approached. It is noted that for data base 1, ave is calculated from all three instrument measurements, while for data bases 2, 3, and 4, ave is calculated from the two respective instrument data.

The third analysis approach is based upon the correlation existing between the measurements reported by pairs of instruments. The correlation is quantified via parameters obtained from an unweighted linear regression of the results from one instrument against the results from a second instrument. The parameters characterizing the linear regression line include the slope, the intercept, the correlation coefficient (r), the number of samples contained in the data

set (N), and the standard deviation on the slope (σ_3) and intercept (σ_1). The correlation coefficient provides a measure of the likelihood that the instruments have observed the same parameter. The intercept provides a measure of any constant bias existing between the respective data sets, and the slope provides a measure of any existing proportional bias. An advantage of the regression analysis is its statistical approach in combining the results into a single analysis which provides a level of agreement which includes both a constant and proportional bias estimate. If a regression slope plus or minus the 2σ on the slope does not include the slope of one, then it can be concluded that at a 2σ level the indicated slope bias is statistically significant. Similarly for the intercept, if the regression intercept plus or minus its 2σ does not include the intercept zero, the indicated intercept bias is statistically significant. A disadvantage of the regression analysis is its sensitivity to a few data points (e.g., outliers) at the higher mixing ratios. As such, one must ensure that the regression analyses are not biased by a few data points which do not reflect the majority of the data range. For comparison of the results from the % diff analyses and regression analyses, it is noted that a factor of 2 disagreement (slope of 2) between two instruments translates into a % diff of 67% (i.e., 100 and 50 pptv investigator measurements gives $(100 - 50)/75$ or 67%).

Using the above methods, the data bases of Table 3, excluding the less-than and the negative data, were examined for the presence of measurements that might tend to bias the overall results, to identify data categories (i.e., subsets) under which intercomparison results should be stated independently, and to identify outlying events for which results are not representative of the data base. In particular, the data bases were examined with the view to consider the influence of (1) the degree of temporal overlap (i.e., the ratio of common sample time of any one instrument measurement to the total overlap period), (2) systematic day-to-day variability, (3) ambient variability during the overlap period, (4) the type of air mass sampled, and (5) the distribution of NO_2 mixing ratios. Data that are clearly "outliers" are omitted from further analysis or separated for independent analysis. In general, while all the analyses discussed above were performed on each data set, only those results most pertinent to describing the level of agreement among the instruments are included in the discussions.

GIT-NOCAR-York Intercomparison

Data base 1 is discussed first, since it provides a direct comparison of the results from the GIT/LIF, NOCAR/Chem, and York/TDLAS instruments. Nineteen of the 44 overlap periods in data base 1 contain less-than or negative values reported by the investigators (Table 3) and are not included in the quantitative analyses discussed below.

Figures 3 and 4 illustrate the general characteristics of the 25 samples in data set 1. Figure 3 shows the frequency distribution of the data as a function of the average (ave) of the mixing ratios reported for each overlap period. For this data base the overlap period was set by the GIT/LIF integration period and therefore contains only one GIT/LIF measurement. The NOCAR/Chem values for each period were typically the average of four to five 1-min samples, while the York/TDLAS values were either a single measurement (2- or 3-min sample) or the average of two such

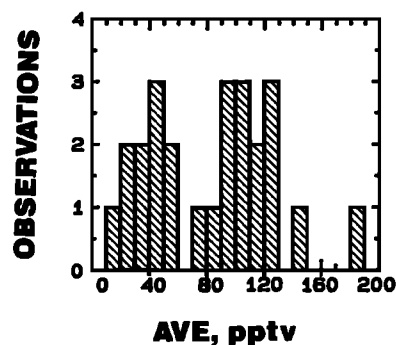


Fig. 3. Average NO_2 mixing ratio histogram: data base 1.

measurements. Note that all of the overlap periods occurred at ave mixing ratios of <200 pptv, and approximately 70% are for mixing ratios of <100 pptv. Thus the intercomparison results should be particularly meaningful for evaluating the ability for NO_2 measurements in clean, remote environments. Figure 4 shows the distribution of the temporal overlap for data base 1. The abscissa (time ratio) is the ratio of the common overlap time for the respective instrument to the total overlap period encompassed by the three techniques. Time ratios are typically greater than 0.3, suggesting that discrepancies associated with variability in ambient NO_2 should be minimal. The time ratio for GIT/LIF is expected to be greater than the other two, since its integration period controlled the selection of the overlap periods.

The effect of the degree of temporal overlap was evaluated by analyzing subsets of the base set—those for measurements having a time ratio of ≥ 0.3 and those (only one overlap) for measurements having a time ratio of <0.3 . No significant difference existed between the analysis parameters obtained for the two subsets. Changes in instrument performance associated with environmental factors or calibration can often result in systematic day-to-day offsets between data sets. To examine the data for such effects, data sets were intercompared on a flight-by-flight basis. Again no significant trends were noted. The effects of air mass type were evaluated by considering subsets as a function of altitude (i.e., less than 1 km and greater than 4 km) and as a function of air mass source (i.e., maritime polar, maritime tropical, and continental tropical). No significant trends or differences were revealed. The NOCAR/Chem data were used to identify overlap periods that might be susceptible to ambient variability in NO_2 . The criterion selected was the ratio of NOCAR's 1σ on the average to that average value reported by NOCAR/Chem. Overlap periods for which this ratio was greater than 1 were considered to most likely have some effects associated with NO_2 variability and the different sample times used by the three NO_2 instruments. The rationale for using the NOCAR/Chem data was based upon the assumption that ambient variability would best be observed by the instrument having the highest sampling rate. Using this criteria, only two overlap periods were found to have a ratio greater than 1. Since this criterion only identifies the potential for an overlap to be affected by ambient variability, a scatter diagram of the overlap data was examined to verify which of the identified overlaps indeed did appear to be an "outlier." Figure 5 shows these scatter diagrams in which the two overlaps are noted (circled). Only

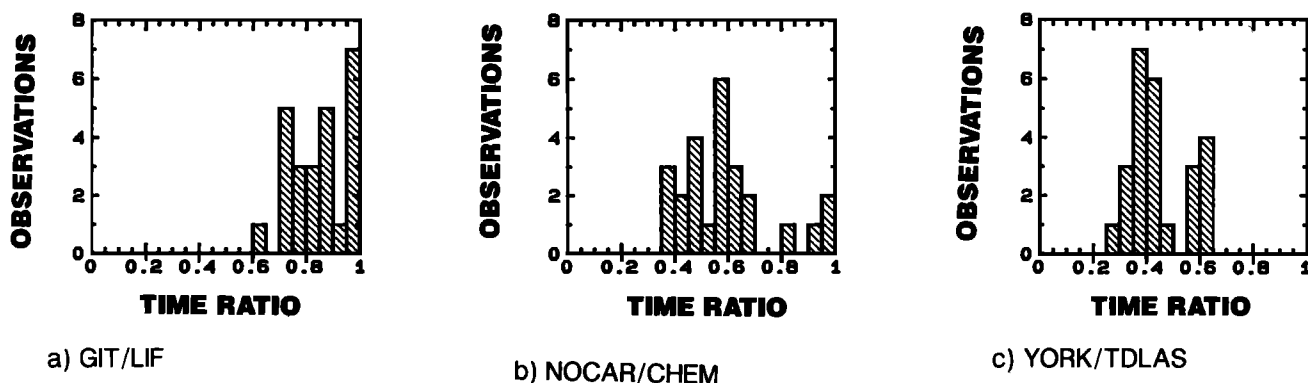


Fig. 4. Sample overlap histograms: data base 1.

overlap number 1 appears to be a potential outlier and thus has been omitted from data base 1.

With no independent absolute measure of ambient NO_2 to serve as a reference the numerical average of the NO_2 mixing ratios obtained for each overlapping period was initially used as the comparison reference. Figures 6, 7, and 8 show the agreement among the instruments using ave as the comparison reference. Figure 6 shows the results of the linear regression analysis (solid curve) for each instrument relative to average (ave) for data base 1. The regression equation associated 1σ uncertainty on the slope (σ_y) and intercept (σ_i), the correlation coefficient (r), and the number of samples (N) are given on each panel. For comparison, the 1:1 correlation is shown as a dashed line in each panel. The delta between the respective pairs of measurements relative to the ave is shown in Figure 7 with the corresponding % diff (calculated by equation (1)) shown in Figure 8. From Figure 6 it is noted that all correlation coefficients are ≥ 0.93 , and the instruments agree with the ave to within 19% (based on slopes) and among themselves to within 30% (NOCAR/Chem slope of 1.19 and York/TDLAS slope of 0.89). The data of Figures 7 and 8 suggest that for comparisons which include the York/TDLAS instrument a distinct change in instrument agreement occurs at mixing ratios below 50 or 60 pptv. This is most apparent in Figure 8b (comparison with GIT/LIF) and Figure 8c (comparison with NOCAR/Chem), where % diff drops precipitately for NO_2 mixing ratios below 50 pptv, while the % diff between GIT/LIF and NOCAR/Chem (Figure 8a) remains uniform. From these observations

it appears that the comparison reference (ave) may be biased by the York/TDLAS measurements for NO_2 mixing ratios less than about 50 pptv and possibly for mixing ratios up to 100 pptv.

To evaluate the influence of any possible bias by including the York/TDLAS mixing ratios, an instrument-to-instrument comparison for data base 1 was performed. Figures 9–11 show the results of a linear regression for instrument versus instrument for mixing ratio ranges of <200 , <100 , and <50 pptv, respectively. From Figure 9 (<200 pptv) the correlation coefficients for the regressions range from 0.85 to 0.91, and the level of agreement among the instruments is of the order of 38% (based on the NOCAR/Chem versus York/TDLAS slope of 1.38 and the fact that $\text{LIF/GIT} < \text{NOCAR/Chem}$ and $\text{LIF/GIT} > \text{York}$). While none of the slopes or intercepts are statistically significant at a 2σ level, it is interesting that the 38% agreement compares well with the 30% level of agreement determined from Figure 6. Note that the regression parameters associated with the GIT/LIF versus NOCAR/Chem regression for each mixing ratio range (i.e., Figures 9a, 10a, and 11a) indicate similar correlation ($r = 0.93\text{--}0.90$), while the regression coefficients for each instrument when paired with York/TDLAS clearly indicate a poorer correlation at the lower mixing ratios, with no correlation for the data set of mixing ratios of <50 pptv. (Thus no regression line is shown in Figures 11b and 11c.) This agrees with the earlier observations which suggested that the ave may be biased by the York/TDLAS measurements at the lower mixing ratios. The 25-pptv detection limit of the

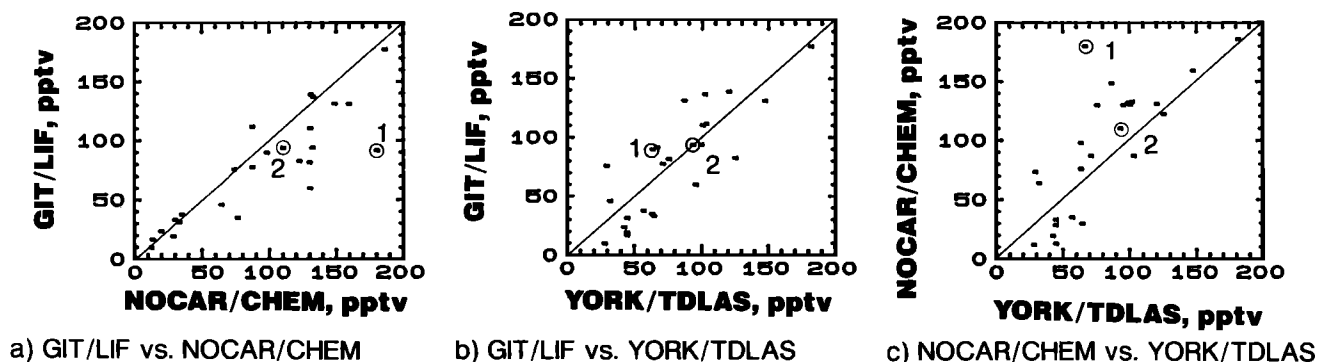


Fig. 5. Scatter plots: data base 1.

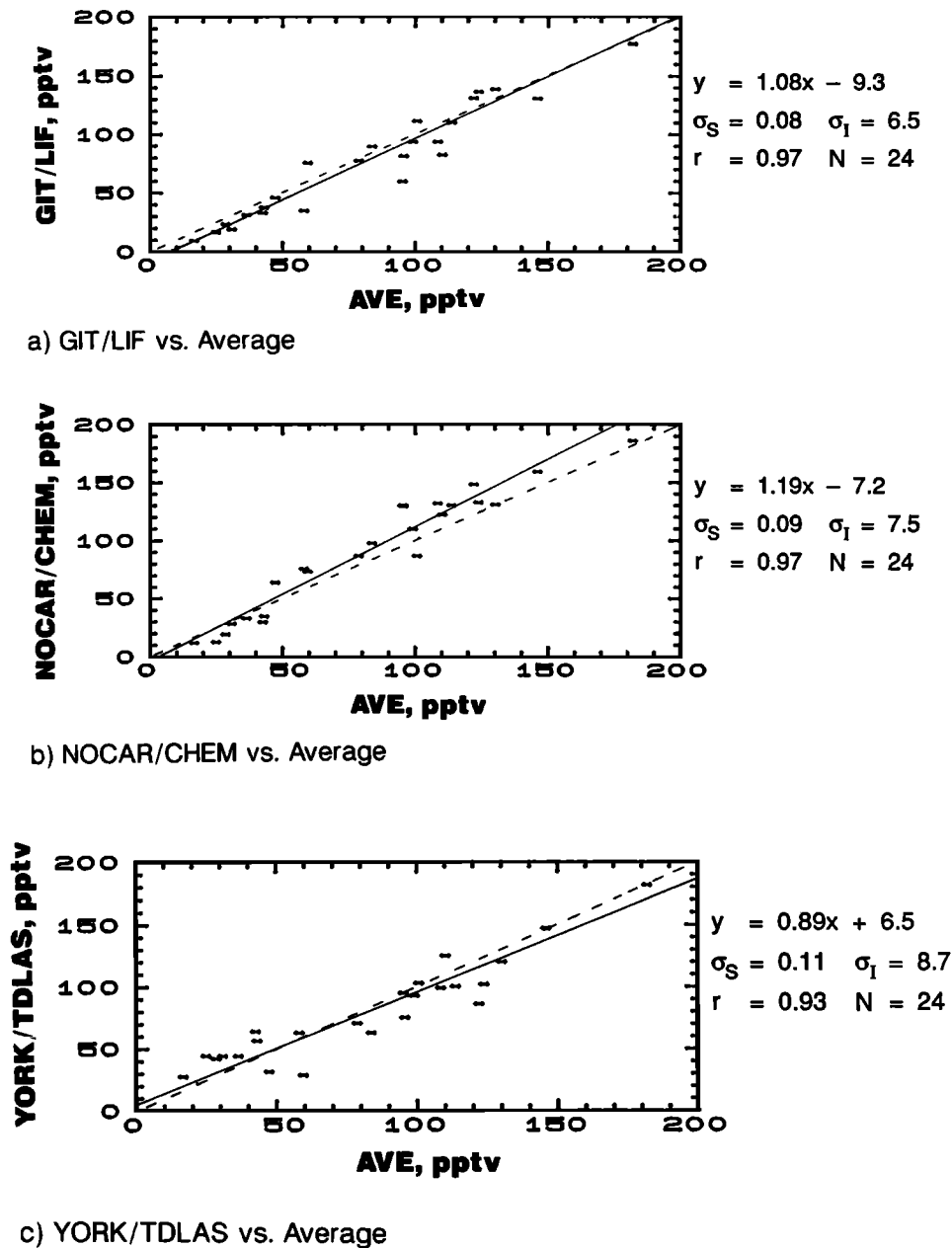


Fig. 6. Linear regression with average mixing ratio: data base 1.

TDLAS is shown as a shaded area in Figures 11b and 11c. Of the indicated slopes and intercepts on Figures 9–11, only the GIT/LIF versus NOCAR/Chem slope of Figure 10a (mixing ratios of <100 pptv) is statistically significant at a 2σ level (i.e., this slope plus or minus its 2σ does not include the slope of one). While not statistically significant at 2σ , it is interesting that all the regressions (Figures 9, 10, and 11) involving York/TDLAS exhibit a sizeable negative intercept bias (–20 to –71).

Each investigator provided an estimate of instrument total uncertainty (see instrument section). If these estimates (appropriate 1σ for the measured mixing ratio) are applied to the investigators' measurement during each overlap period to arrive at an estimate of the "expected" uncertainty between two measurements, then one can obtain a feel for how often the actual difference between individual pairs of measure-

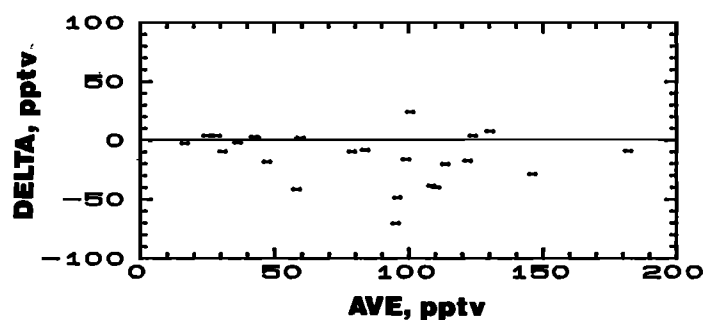
ments was greater than an expected difference. Such an estimate can be obtained by considering equation (2),

$$E_{12} = \sqrt{\text{uncert}_1^2 + \text{uncert}_2^2} \quad (2)$$

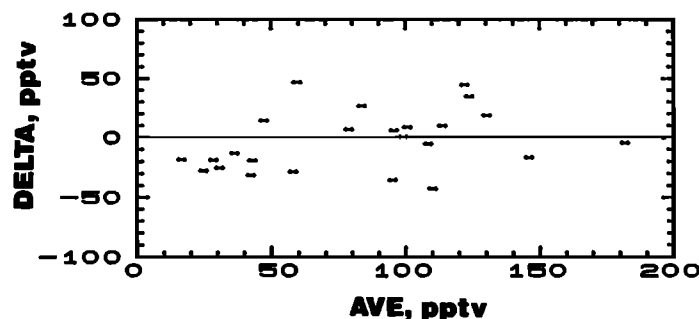
where

- E_{12} expected uncertainty between instruments 1 and 2, in parts per trillion by volume (pptv);
- uncert_1 uncertainty associated with technique 1 measurement during the overlap period, pptv;
- uncert_2 uncertainty associated with technique 2 measurement during the overlap period, pptv.

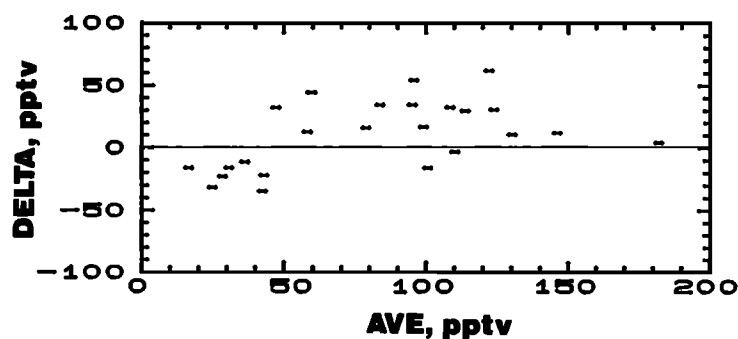
Of the 24 overlaps and for those cases involving York/TDLAS, delta exceeded the estimated uncertainty (E) for eight (with GIT/LIF) and nine (with NOCAR/Chem) cases.



a) GIT/LIF - NOCAR/CHEM



b) GIT/LIF - YORK/TDLAS



c) NOCAR/CHEM - YORK/TDLAS

Fig. 7. Delta difference between techniques: data base 1.

On the other hand, GIT/LIF-NOCAR/Chem discrepancies were generally within their estimated uncertainties, producing only five overlaps in which delta exceeded E . The fact that comparisons with York/TDLAS resulted in a higher frequency of delta exceeding E is attributed in part to the fact that many of the measurements were at or near the 25-pptv detection limit of the TDLAS technique. If the above analysis is performed for mixing ratios only above 100 pptv (nine overlaps), then delta (York/TDLAS) is exceeded only once each with GIT/LIF and NOCAR/Chem. It is noted that application of equation (2) at a 2σ uncertainty level results (as was the case for all data sets) in almost all the delta values being within expected limits.

As noted initially for data base 1, 19 of the original 44 overlaps contained less-than or negative values. Of these 19, 16 included both GIT/LIF and York/TDLAS reported less-than data. In each instance, NOCAR/Chem results were

lower than the other two reported values. The other three of the 19 overlaps were the result of various combinations of less-than or negative values, and in each case these data were consistent among the techniques. In general for data base 1, York/TDLAS less-than data were at the level of <25 pptv; GIT/LIF data were in the range of <10 to <25 pptv.

GIT-NOCAR Intercomparison Data Set

Data base 2 (see Table 3) initially contained 97 overlap periods. Ten have been excluded from the analyses due to the reporting of either less-than or negative NO_2 values. For all 10 cases the reported data were internally consistent between the two instruments. Eight additional overlaps were deleted from data base 2, identified as outliers due to suspected NO_2 ambient variability. For each of the eight the NOCAR/Chem σ to NOCAR/Chem average ratio for the

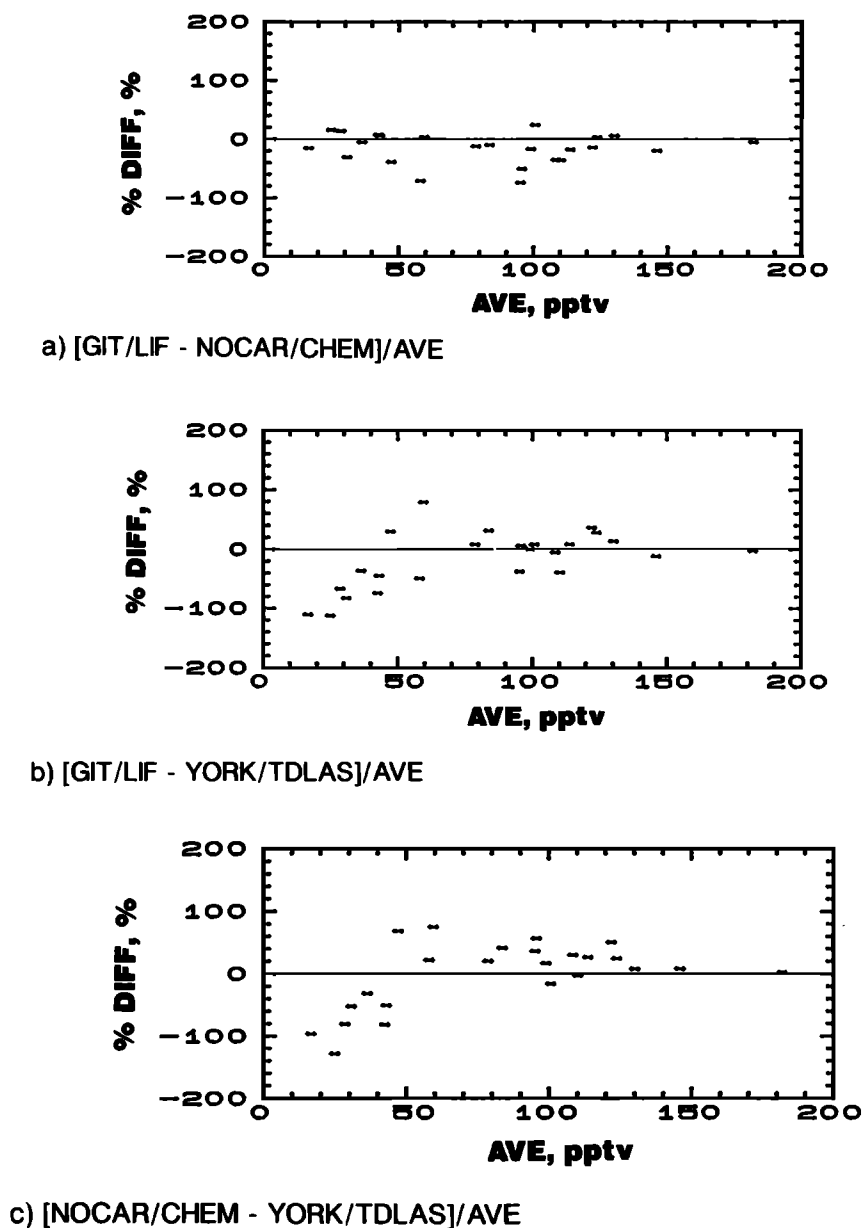
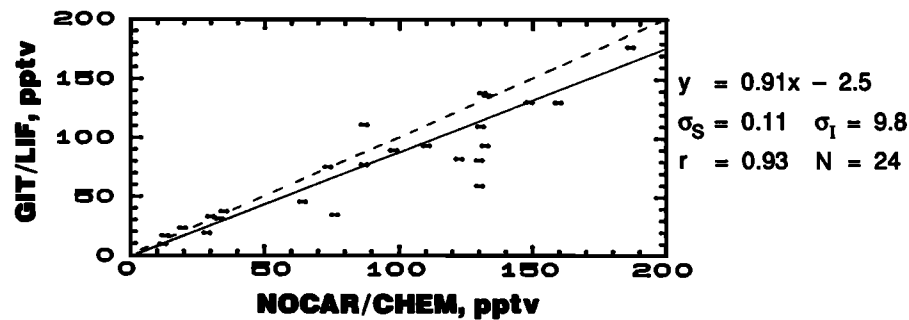


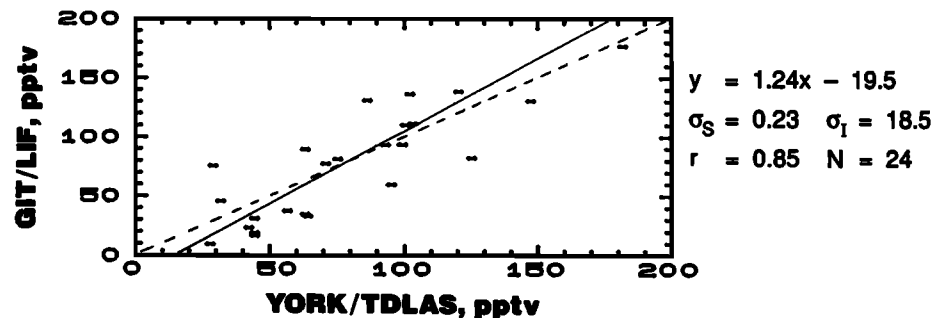
Fig. 8. Percent difference between techniques: data base 1.

overlap period was greater than 2. As was the case for data base 1, the ave NO_2 mixing ratios for all overlap periods (now 79 samples) were <200 pptv. For over 70% of the data the average NO_2 mixing ratio was <50 pptv. Temporal overlap of the two instruments was reasonable with 75% of the data having time ratios between 0.3 and 0.6 and only four having a time ratio of <0.3 . While complete analyses of the data base were performed (i.e., identical to the analysis discussed earlier for data base 1), only the overall results are presented. Subset analyses showed no evidence of biases as a result of test conditions (except the eight outliers already noted). Figure 12 shows the results of the linear regression analyses for the data sets of mixing ratios less than 200, 100, 50, and 25 pptv. The correlation coefficients for the <200 - and <100 -pptv data sets are 0.95 and 0.87, but the coefficient deteriorates to 0.37 and 0.05 for the lower mixing ratio data sets. Results from the two higher mixing ratio data sets show a similar level of agreement between the instruments with

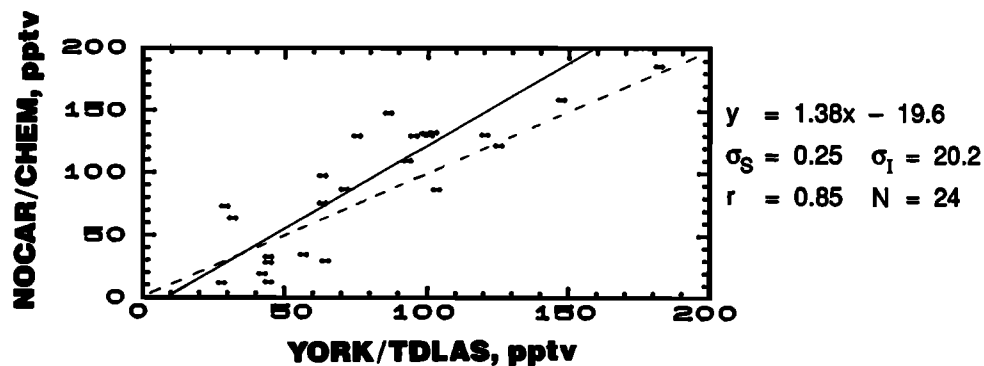
GIT/LIF being about 20% lower than NOCAR/Chem. Both regression slopes are statistically significant at a 2σ level, while the approximate 5-pptv intercept biases are not significant at 2σ . It is noted that in data base 1 (discussed earlier) the only statistically significant slope bias was for the regression of GIT/LIF versus NOCAR/Chem and for the data set of mixing ratios of <100 pptv. The results from data base 1, while indicating GIT/LIF to be low relative to NOCAR/Chem, provided a rather large range (as the result of a large slope σ) for the magnitude of the bias, namely, a slope of 0.71 ± 0.28 (2σ). Data base 2 narrows this range to about 0.81 ± 0.16 . This level of agreement with GIT/LIF being about 20% low compared to NOCAR/Chem is in the opposite direction of the results from the standards intercomparisons that were discussed earlier. (If the NBS reference system calibration data are used as a transfer standard between the two instruments, then GIT/LIF is about 20% higher than NOCAR/Chem.)



a) GIT/LIF vs. NOCAR/CHEM



b) GIT/LIF vs. YORK/TDLAS



c) NOCAR/CHEM vs. YORK/TDLAS

Fig. 9. Linear regression results for data base 1: ave mixing ratios of <200 pptv.

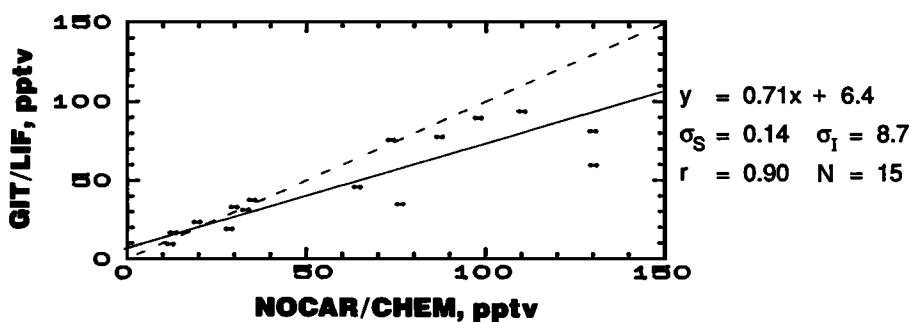
The lack of correlation between the two instruments at lower mixing ratios (Figures 12c and 12d) has to be viewed with other information. First, while from Figure 13a the % diff below about 50 pptv can be large (ranges from +100 to -100%) and highly variable, Figure 13b (with the exception of one data point) shows that the delta (below 50 pptv) between the two instruments is always within 20 pptv with approximately equal tendency for any one technique to be high or low compared to the other. Second, for the entire data set (mixing ratios of <200 pptv), in 23 of 79 overlaps (29%) delta exceeded the expected uncertainty, E , as calculated from equation (2). This percentage did not change when considering mixing ratios of <25 pptv (45 data points) and was 28.8%.

In summarizing data base 2 one concludes that there are no inconsistencies with results stated earlier from data base 1. As a result of the larger number of samples in data base 2

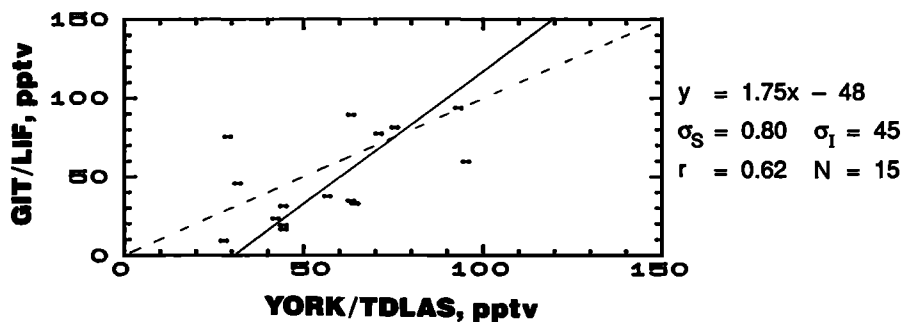
a statistically significant proportional bias between the two instruments (at mixing ratios of >100 pptv) is identified and more accurately quantified. As calculated from equation (2) the bias was generally (two of three times) within the stated uncertainties for the two instruments.

GIT-York Intercomparison Data Set

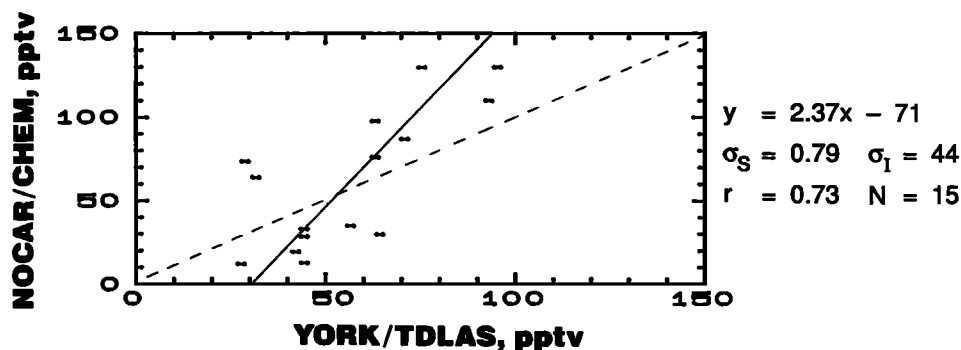
Data base 3 (see Table 3) consists of 30 overlaps after removing those 20 overlaps which included less-than data. Of these 20, 17 were due to York/TDLAS reported upper limits, one to GIT/LIF reporting upper limits, and two to both. In all cases the reported upper limits were internally consistent. Subset analyses suggested no biases in the data set as the result of test conditions. Again, all overlap periods occurred at NO_2 mixing ratios below 200 pptv. The maximum time ratio for any overlap was 0.66 with about 80% of



a) GIT/LIF vs. NOCAR/CHEM



b) GIT/LIF vs. YORK/TDLAS



c) NOCAR/CHEM vs. YORK/TDLAS

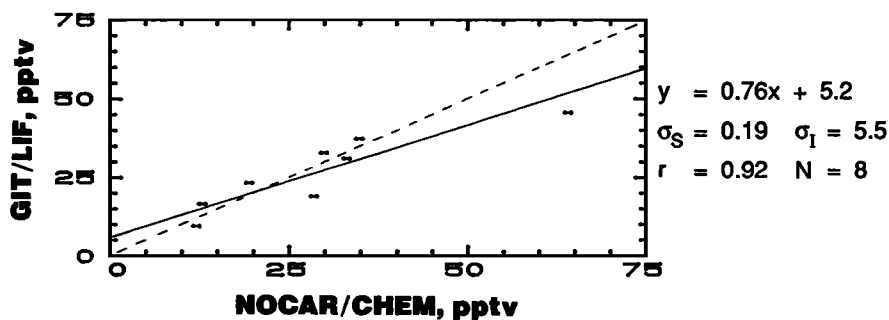
Fig. 10. Linear regression results for data base 1: ave mixing ratios of <100 pptv.

the overlaps having a time ratio between 0.3 and 0.5. The results from these data are in general agreement with the earlier results from data base 1 (Figures 9b, 10b, and 11b); thus no data are shown. This is not surprising, since 24 of the 30 overlap periods are coincident with those analyzed in data base 1. The major observation from data base 3 is its supporting evidence that comparisons with York/TDLAS deteriorate at mixing ratios below 50 pptv and that at these mixing ratios York/TDLAS has a strong tendency to be high. For mixing ratios above 50 pptv and based on equation (2), 25% (five of 20) of the overlaps have a delta greater than expected. For mixing ratios greater than 100 pptv, two of 10 overlaps have a delta greater than expected.

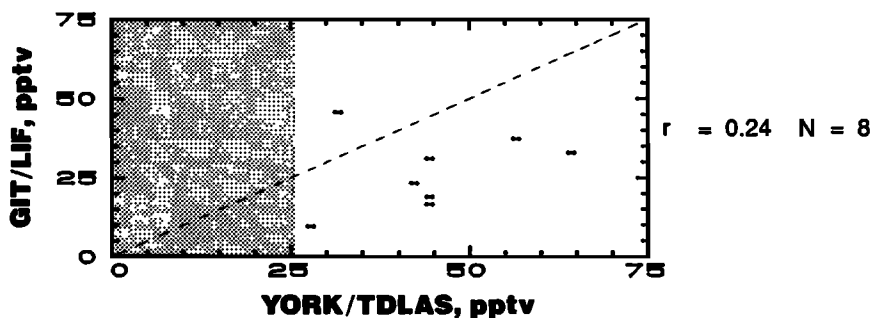
NOCAR-York Intercomparison Data Set

Data base 4 initially contained 108 overlaps. Of these, 29 are excluded because they contain York/TDLAS less than

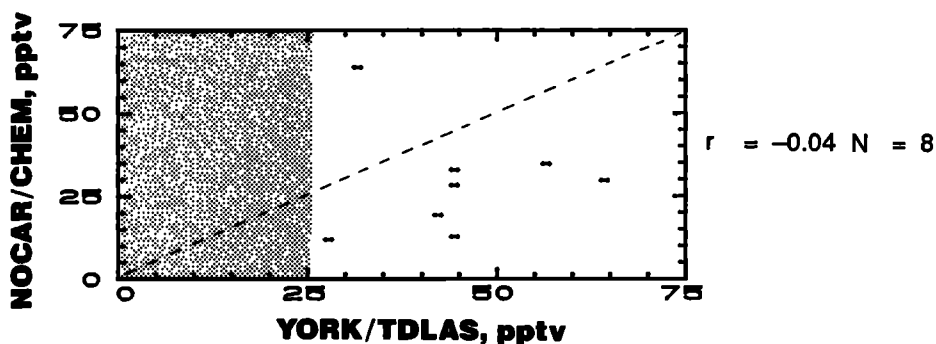
data (27 overlaps) or NOCAR/Chem negative data (two). For all these 29 overlaps the reported data are internally consistent between the two instruments. Six additional overlaps are excluded from analysis due to suspected ambient NO₂ variation identified by the earlier discussed NOCAR sigma to average ratio test. Subset analyses were performed on data base 4 (now 73 samples) with the results that instrument agreement from flight 4 appeared to be different than observed from the other flights. For the five overlap periods from flight 4, York/TDLAS reported abnormally high (factor of 8) NO₂ compared to NOCAR/Chem. While (as will be shown) there is a tendency in data base 4 for York/TDLAS to be high compared to NOCAR/Chem, the flight 4 results are unique in that York/TDLAS is consistently about 150 pptv higher than NOCAR/Chem NO₂ values (all in the range of 50–70 pptv). As a result these five overlaps are considered as outliers and are excluded from the analyses presented



a) GIT/LIF vs. NOCAR/CHEM



b) GIT/LIF vs. YORK/TDLAS



c) NOCAR/CHEM vs. YORK/TDLAS

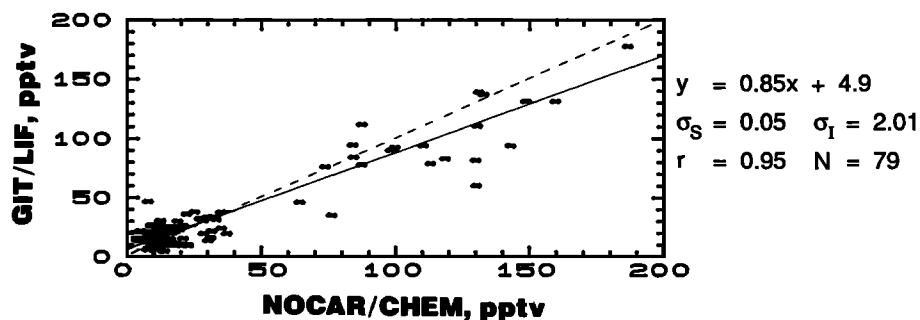
Fig. 11. Linear regression results for data base 1: ave mixing ratios of <50 pptv.

below. The suspected cause of the flight 4 anomaly is a leak in the York/TDLAS inlet. While in the field, York reported a potential leak in its inlet system, and discussions were held with the project concerning the submittal of the data. The York/TDLAS data of flight 4 were resubmitted several times while in the field.

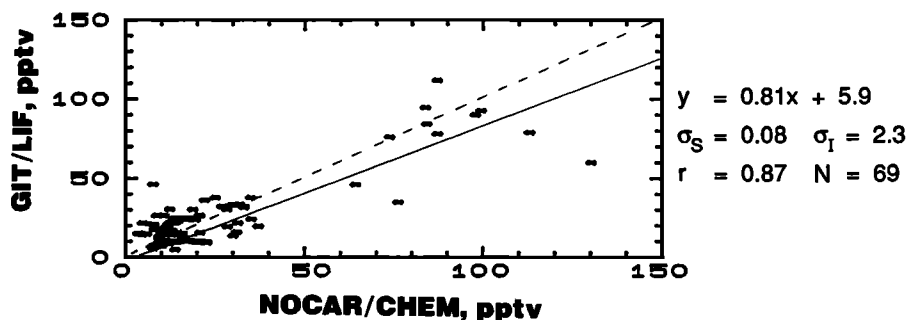
Figures 14 and 15 show the results from regression, delta, and % diff analyses of data base 4. (It is noted that York/TDLAS did report a few data points, all in data base 4, that were below the previously stated 25-pptv detection limit.) Two data points (circled and labeled as 1 and 2 in Figures 14 and 15) have been omitted from the linear regression analyses of Figures 14a and 14b. Data point 1 is from flight 6, and data point 2 from flight 7. Both data points are excluded due to suspected ambient NO₂ variability (confirmed by various tests and data). Exclusion of the data points does not change the regression equations to any degree but does improve the

correlation coefficient from about 0.7 to 0.84. The results from analyses of data base 4 agree with those presented earlier in that (1) correlation with York/TDLAS deteriorates at the lower mixing ratios (no correlation at <50 pptv) and (2) below about 50 pptv there is a strong tendency for York/TDLAS to be high. Concerning these two observations the York/TDLAS investigators, after examining the CITE 2 results and performing additional laboratory tests, have provided an explanation for some of the disagreements at the lower concentrations. Their explanation is paraphrased below. A more detailed discussion is given by Schiff *et al.* [this issue].

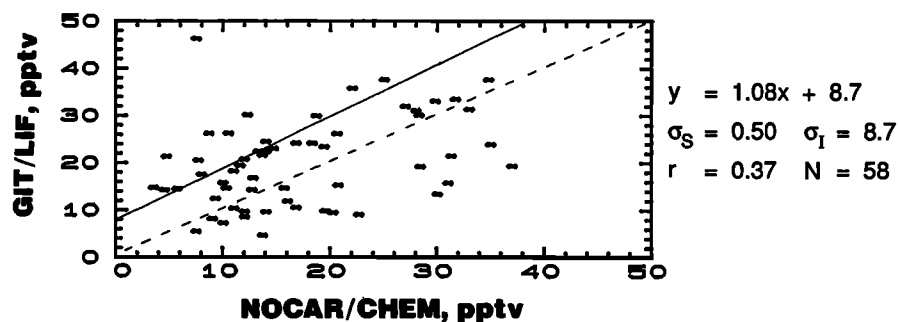
The deviations in the York/TDLAS values below an ave of 50 pptv were all due to higher values than reported by GIT/LIF or NOCAR/Chem. After examining the comparison data set, the York investigators observed that there was a bias in the data analysis procedure which preferentially



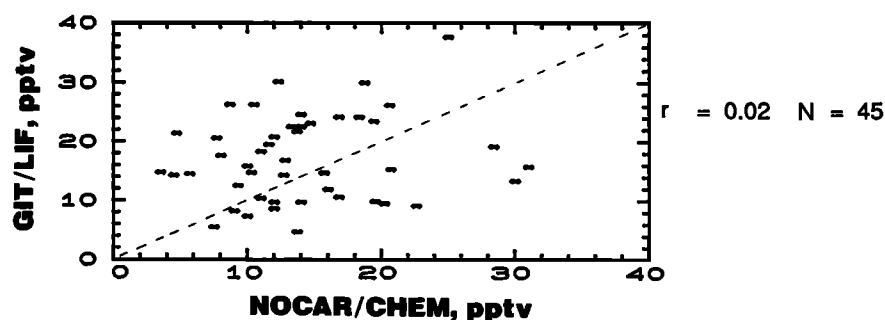
a) Mixing ratios < 200 pptv



b) Mixing ratios < 100 pptv



c) Mixing ratios < 50 pptv

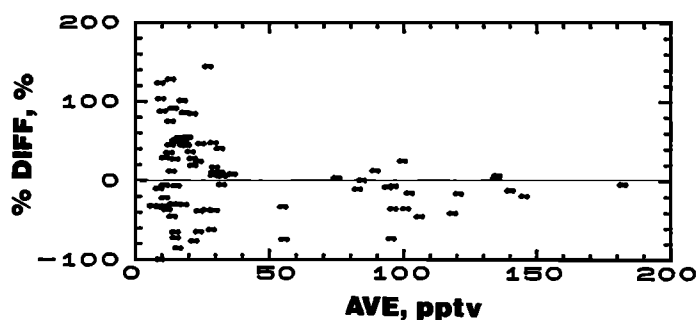


d) Mixing ratios < 25 pptv

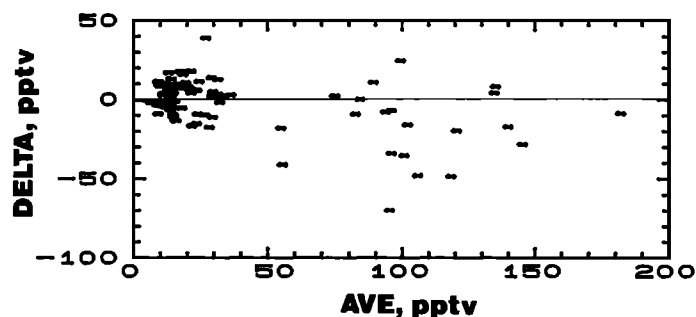
Fig. 12. Scatter plots and linear regression results: data base 2. Mixing ratio ranges shown in the panels are based on ave mixing ratio; i.e., Figure 12b shows data in which ave < 100 pptv.

selected higher values for their measurements at mixing ratios approaching their detection limit of 25 pptv. The procedure involves fitting the shape of the absorption feature of the ambient air sample to that of the calibration sample

and determining a correlation coefficient which is taken as a measure of the “goodness” of the fit. This correlation factor typically ranged from values greater than 0.95 at mixing ratios above 100 pptv to values less than 0.5 near the



a) % DIFF vs. AVE: $\% \text{ DIFF} = [\text{GIT/LIF} - \text{NOCAR/CHEM}] / \text{AVE}$



b) DELTA vs. AVE: $\text{DELTA} = [\text{GIT/LIF} - \text{NOCAR/CHEM}]$

Fig. 13. Instrument agreement as function of ave mixing ratio: data base 2.

detection limit as the measurement signal-to-noise approached unity. Values having correlation factors less than 0.5 were subjectively rejected or recorded as “less than 25 pptv” rather than, for example, as 11 ± 6 pptv with a correlation factor of 0.38. If the measurements are normally distributed about the correct value, this procedure biases the averages of the measurement toward the higher values regardless of how many measurements are combined in the average (2- to 3-min data submitted). In retrospect, the investigators state that the subjective decision to reject data on the basis of the correlation coefficient should not have been used in the analysis algorithm.

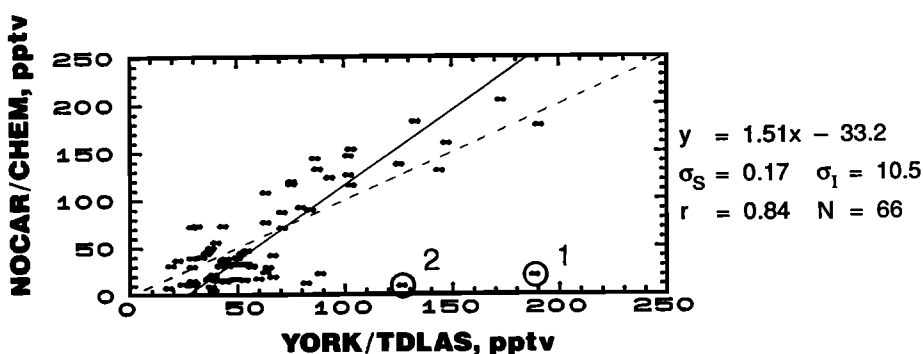
Other important points from the analyses are that for the subset of mixing ratios of <200 pptv, both the indicated slope and intercept biases are statistically significant at 2σ , verifying the earlier noted tendency (other data bases) for York/TDLAS to be high compared to NOCAR/Chem and to have a sizeable negative intercept bias. For data base 4 and above 50 pptv, 55% (15 of 27) of the overlaps have delta values greater than the expected. Above 100 pptv, 46% (six of 13) of the overlaps have larger than expected delta values.

SUMMARY

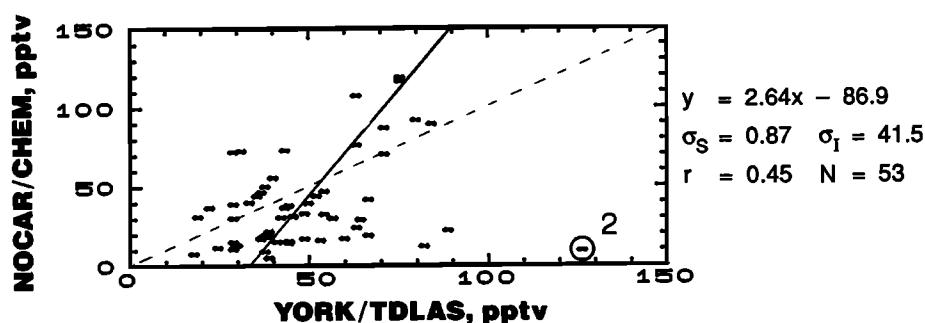
The CITE 2 activities included ground-based NO_2 intercomparisons of project-supplied standards as well as airborne measurements from four different measurement techniques. The four techniques included a two-photon NO laser-induced fluorescence laser system with laser photolysis of NO_2 to NO (GIT/LIF), and NO chemiluminescent detector using iron sulfate (FeSO_4) conversion of NO_2 to NO (Wallops/Chem), an NO chemiluminescence detector with arc lamp photolysis of NO_2 to NO (NOCAR/Chem), and a tunable diode laser multipath absorption system (York/TDLAS). The standards intercomparison was performed by

having each instrument sample from the output of a portable NO_2 reference source provided by the National Bureau of Standards. During these tests a significant calibration error associated with a flow-dependent loss in the GIT/LIF calibration system was identified. A second GIT/LIF standards test was performed after correcting the calibration error. The standards tests also revealed a 10.7% calibration error in the York/TDLAS system which was corrected postmission after recalibration of the NO_2 permeation source. After correcting these calibration errors the agreement between the NO_2 instruments and the NBS reference source ranged from -9% to +21% at the 700-pptv level. At the 400-pptv level the agreement ranged between 3% and -17%. These levels of agreement were basically within the relative uncertainties stated for the instruments and NBS standards. If one accepts the NBS data as a type of “transfer” standard among the instruments, then the results would tend to suggest that as a result of calibration standards/procedures, GIT/LIF is about 20% high and York/TDLAS 10% low relative to the chemiluminescence instruments.

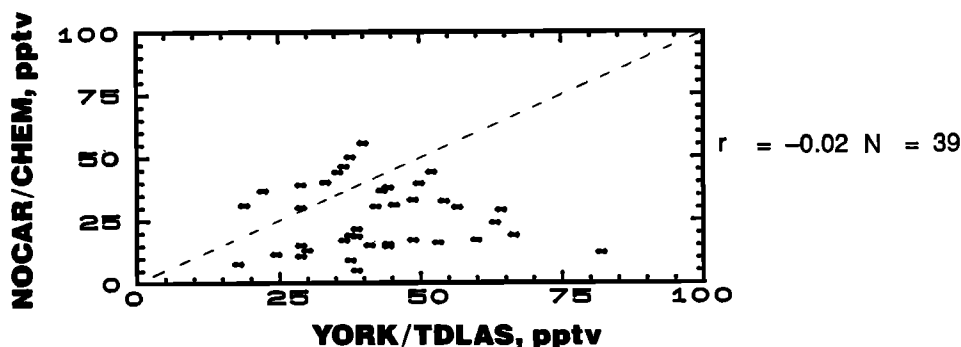
During the CITE 2 mission there were approximately 65 flight hours for intercomparison of NO_2 measurements. The York/TDLAS also participated in the CITE 2 intercomparison of HNO_3 instruments. Accordingly, the data base for York/TDLAS NO_2 intercomparison was about 45 hours of flight. The GIT/LIF participated in about 50 hours of flight intercomparison. Wallops/Chem measurements were consistently a factor of 2–3 higher than measurements from the other three instruments. The FeSO_4 converter used by Wallops/Chem was not specific to NO_2 , and other nitrogen species (e.g., PAN) were converted to NO and introduced a significant interferent to NO_2 . As such the Wallops/Chem NO_2 measurements were not included in the intercomparison analyses.



a) Mixing ratios <200 pptv



b) Mixing ratios <100 pptv



c) Mixing ratios <50 pptv

Fig. 14. Scatter plots and linear regression results: data base 4. Mixing ratio ranges shown in the panels are based on ave mixing ratio; i.e., Figure 14b shows data in which ave <100 pptv.

A significant data base was obtained for intercomparison of the other three instruments and included over 40 overlap periods in which all three instruments were sampling ambient NO_2 . Separate data bases of overlapping measurements between pairs of instruments were also constructed and included from 50 to over 100 overlap periods. Analyses of these overlap data sets formed the basis for stating levels of instrument agreement. All the intercomparisons were at NO_2 mixing ratios of <200 pptv, with greater than 50% of the intercomparisons at mixing ratios of <100 pptv.

In-Flight Intercomparisons: Mixing Ratios Less Than 200 pptv

These data support a general level of agreement among the instruments of the order of 30–40% (regression slopes) with

GIT/LIF being low compared to NOCAR/Chem and high compared to York/TDLAS. The data are also consistent in showing NOCAR/Chem high relative to York/TDLAS. While correlation coefficients for linear regression analyses of the <200-pptv data sets ranged from 0.84 to 0.95, the indicated slopes and intercept biases were often not statistically significant at a 2σ level. Those regressions which were statistically significant showed GIT/LIF about 15% low compared to NOCAR/Chem and NOCAR/Chem high relative to York/TDLAS. The NOCAR/Chem versus York/TDLAS regression also exhibited a statistically significant negative intercept bias of -33 ± 21.0 pptv (2σ). All the York/TDLAS regression analyses with the other instruments showed sizeable negative intercept biases. It is noted that the GIT/LIF and NOCAR/Chem difference is in the opposite

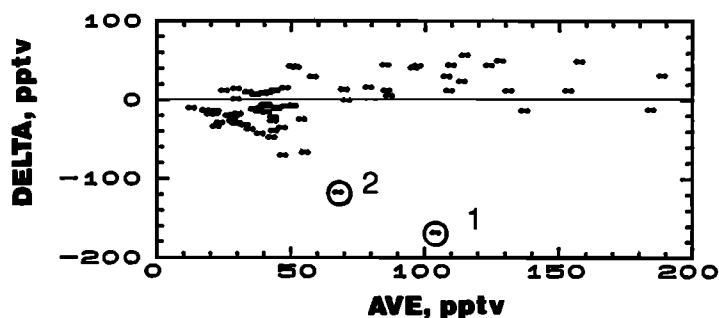
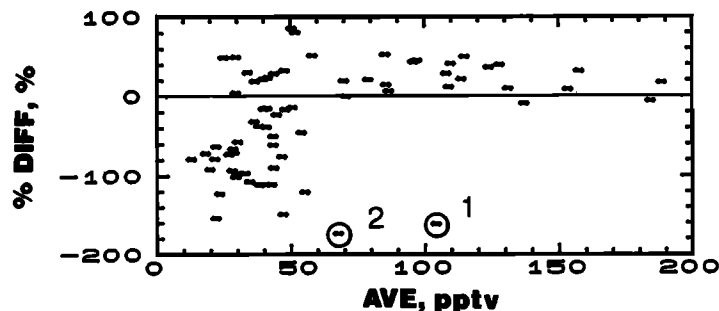
a) DELTA vs. AVE: $\text{DELTA} = [\text{NOCAR}/\text{CHEM} - \text{YORK}/\text{TDLAS}]$ b) % DIFF vs. AVE: $\% \text{ DIFF} = [\text{NOCAR}/\text{CHEM} - \text{YORK}/\text{TDLAS}] / \text{AVE}$

Fig. 15. Instrument agreement as function of ave mixing ratio: data base 4.

direction as observed during the standards (assuming NBS as a transfer standard) in which GIT was high compared to NOCAR by about 20%.

In-Flight Intercomparisons: Mixing Ratios Less Than 100 pptv

The major observation in this mixing ratio range is an apparent deterioration in those regression analyses which included the York/TDLAS instrument. Correlation coefficients for regressions involving York/TDLAS range from 0.45 to 0.73 compared to those which involved GIT/LIF versus NOCAR/Chem of 0.87 and 0.90. The correlation which existed between GIT/LIF and NOCAR/Chem is similar to that for the data set which included mixing ratios above 100 pptv. The data still support a level of instrument agreement of 30–40% among the instruments, but with the possibility that (because of the poorer correlations) York/TDLAS could be exhibiting a larger deviation. The GIT/LIF versus NOCAR/Chem regression gives a slope $\pm 2\sigma$ of 0.81 ± 0.16 ($N = 69$, $r = 0.87$) which is well within the level of agreement stated above for mixing ratios to 200 pptv. As noted from the earlier analyses, the York/TDLAS intercept bias is sizeable, as shown by the regression with both NOCAR/Chem and GIT/LIF.

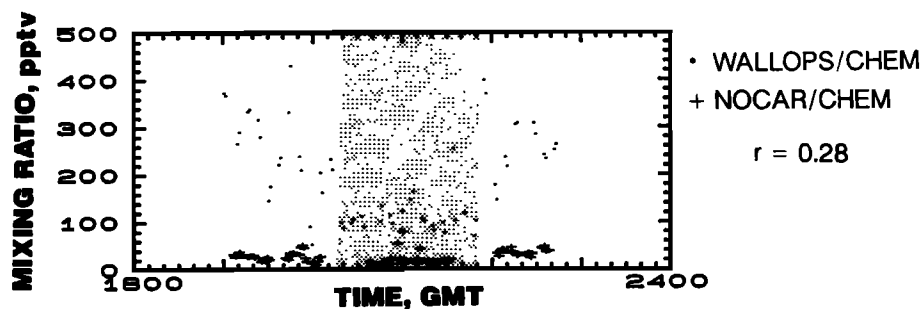
In-Flight Intercomparisons: Mixing Ratios Less Than 50 pptv

Generally, below 50 pptv the data show little correlation. GIT/LIF versus NOCAR/Chem show the best correlation. Regression of an eight-sample data set showed excellent correlation ($r = 0.92$). However, when the data set was expanded to 58 samples (data base 2), the correlation was

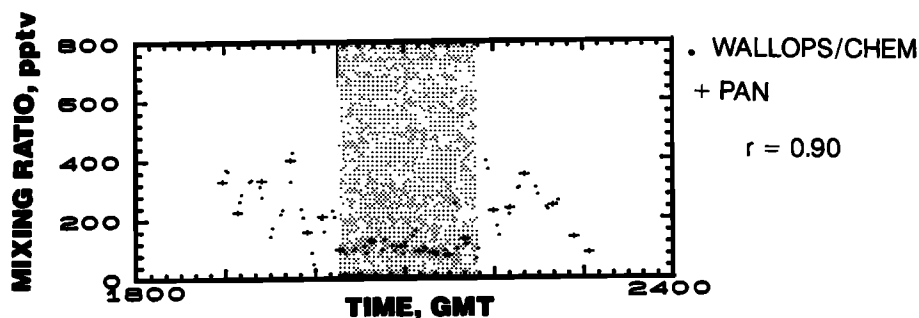
only 0.37. At the low mixing ratios there is a strong tendency for York/TDLAS to be high compared to the other instruments. The lack of correlation at the lower mixing ratios must be considered in view of additional information. First, the York/TDLAS detection limit was stated at 25 pptv. Thus measurements near this detection limit can be expected to be of poor quality for integration periods of 2–3 min used during CITE 2. Second, and based upon the results from the CITE 2 intercomparisons, the York/TDLAS investigators found an error in their analysis algorithm which has the effect of biasing results at the lower mixing ratios (e.g., below 50 pptv). The error involves the method in which the absorption features from calibration are applied to measurements of low mixing ratios. The net effect of the error is an overestimation of the NO_2 mixing ratio. Third, and in terms of GIT/LIF and NOCAR/Chem, while no correlation was noted between the two instruments at mixing ratios less than 50 pptv, their absolute agreement was within 20 pptv and with an equal tendency for one to be high or low compared to the other. This 20-pptv agreement is typically within stated uncertainties of the two instruments at these mixing ratios.

APPENDIX A: NITROGEN DIOXIDE MEASUREMENTS: THE FERROUS SULFATE CONVERTER

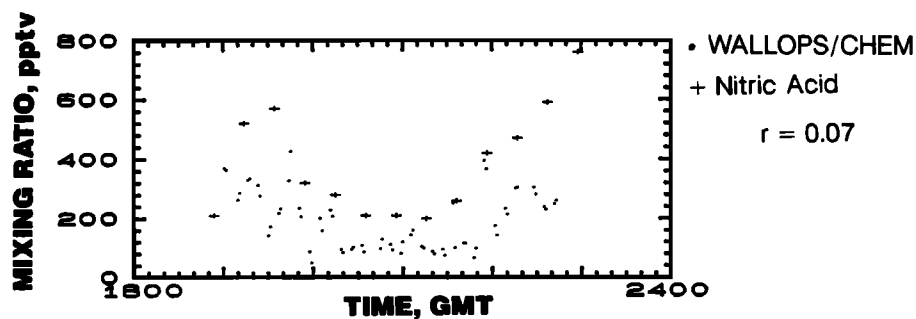
As discussed in the text, the Wallops/Chem nitrogen dioxide instrument, which used the ferrous sulfate converter (surface/chemical reduction of NO_2 to NO) as part of its detection principle, frequently measured higher values of NO_2 than the other techniques. Other work [e.g., Kelly *et al.*, 1980; Fehsenfeld *et al.*, 1987; Ridley *et al.*, 1989] has shown that the ferrous sulfate converter, while providing nearly 100% reduction of NO_2 to NO , also converts other



a) Comparison of nitrogen dioxide measurements: Chemiluminescent instruments



b) Comparison of WALLOPS/CHEM nitrogen dioxide measurement with PAN

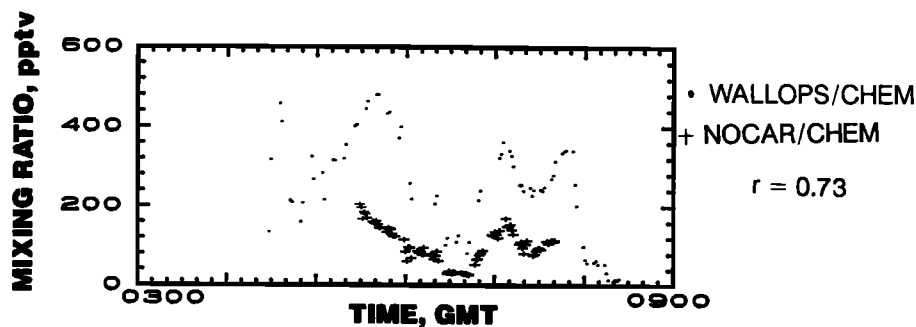


c) Comparison of WALLOPS/CHEM nitrogen dioxide measurement with nitric acid

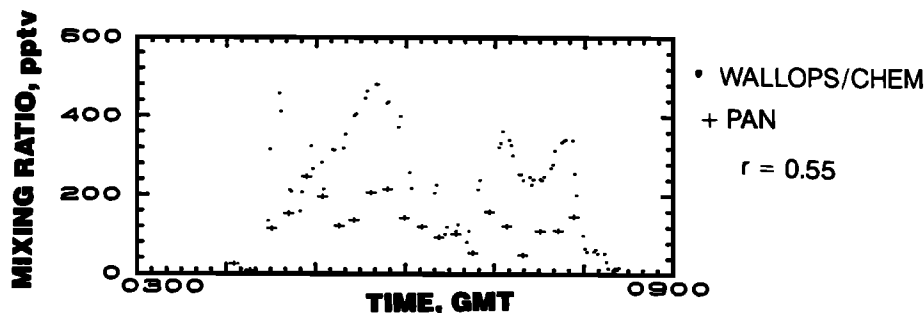
Fig. A1. Comparison of Wallops/Chem NO_2 measurements with other nitrogen species measurements: August 15, 1986.

nitrogen species (especially organic nitrogen compounds like PAN) to NO . The net effect of these additional conversions is an overestimation of the ambient NO_2 concentration. The CITE 2 data base (tropospheric aircraft data) supports these conclusions. The CITE 2 data base was analyzed to provide additional insight into the behavior of the ferrous sulfate converter. Wallops/Chem data were analyzed in conjunction with the other NO_2 data and the other nitrogen species data, namely, NO , PAN, and HNO_3 . Figures A1–A3 illustrate the results. Figures A1a, A2a, and A3a show the comparison of the Wallops/Chem and NOCAR/Chem NO_2 data. As discussed in the text, the major difference between the two techniques is the method of conversion of NO_2 to NO . NOCAR/Chem uses a photolytic conversion process. Figures A1b, A2b, and A3b compare the Wallops/Chem data with PAN measurements as provided by the National Center for Atmospheric Research (NCAR), and Figures A1c and A2c show the comparison with nitric acid (NASA/Langley DENUDER technique). The NCAR data were selected

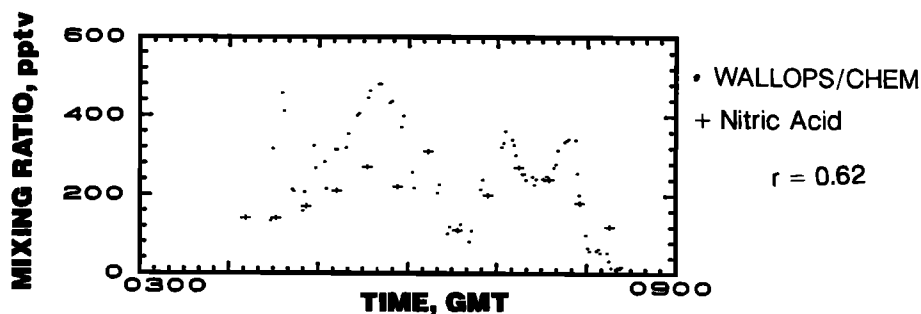
(Ames data also available) because their sample time more closely matched the Wallops/Chem sample periods. The nitric acid DENUDER data were selected because, of the two nitric acid techniques which provided data on all flights (DENUDER and FILTER), it had the highest time resolution. Figure A1 shows results from the August 15 flight (oceanic/daytime); Figure A2, the August 23 flight (continental/nighttime); and Figure A3, the September 5 flight (continental transit/California to Oklahoma). Also given in each figure is the correlation coefficient (r) calculated for each pair of measurements. The correlation coefficient was calculated for overlapping data periods for the indicated measurement pairs. In order to obtain a sufficient number of data points a 2-min overlap criterion was used (i.e., data were considered as overlapping if the two measurements were made within 2 min of each other), and overlap periods between the species were determined independently of each other (i.e., overlaps for the two NO_2 measurements were determined and a correlation performed, overlaps for Wallops/Chem and PAN



a) Comparison of nitrogen dioxide measurements: Chemiluminescent instruments



b) Comparison of WALLOPS/CHEM nitrogen dioxide measurement with PAN



c) Comparison of WALLOPS/CHEM nitrogen dioxide measurement with nitric acid

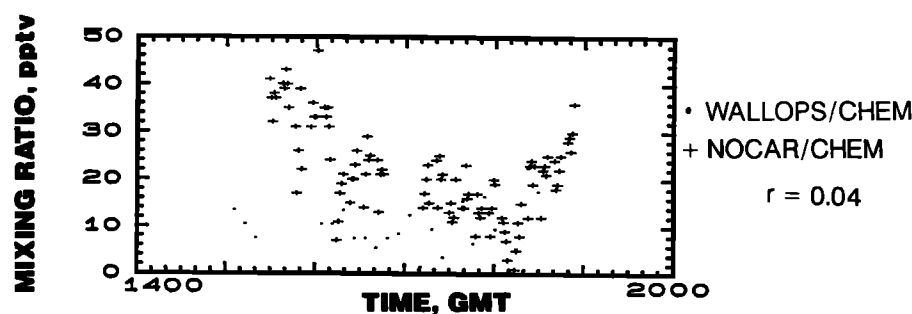
Fig. A2. Comparison of Wallops/Chem NO_2 measurements with other nitrogen species measurements: August 23, 1986.

measurements were determined and a correlation performed, and so on). Two perfectly correlated measurements would exhibit a coefficient of +1 (positive correlation) or -1 (negative correlation). Two random measurements would exhibit a coefficient of 0. A coefficient of 0.8 between Wallops/Chem and PAN would suggest a high degree of positive correlation between the measurements but is not to be interpreted as meaning that 80% of the Wallops/Chem measurements are the result of ambient PAN contributions.

As noted in Figure A1a, Wallops/Chem NO_2 is significantly higher than NOCAR/Chem throughout the flight. While NOCAR/Chem shows NO_2 consistently at values below about 50 pptv, Wallops/Chem shows values as high as 400 pptv and a significant time variation in NO_2 . From Figure A1b it is observed that the Wallops/Chem NO_2 mixing ratios compare well with the PAN data, in both absolute value and time history. From Figure A1c it is observed that the nitric acid and Wallops/Chem data also

show a similar time history. The indicated correlation coefficients suggest that the Wallops/Chem results are highly correlated with PAN, exhibit some degree of correlation with NO_2 , and are not correlated with nitric acid. From Figure A1 one concludes that the ferrous sulfate converter is sensitive to PAN and the conversion efficiency is relatively high. A comparison of the Wallops/Chem NO_2 , NOCAR/Chem NO_2 , and PAN absolute values would suggest a PAN conversion efficiency upward of 80%. As expected, the Wallops/Chem and NOCAR/Chem data show better agreement at the lower values of PAN (see shaded areas of Figures A1a and A1b).

Figure A2 illustrates results when ambient PAN and NO_2 (NOCAR/Chem) values were at about the same levels. Observations from Figure A2 are similar to those above with two noted differences. First, the correlation between the two NO_2 instruments is much better than shown in Figure A1 (0.73 versus 0.28), indicating that a significant quantity of the



a) Comparison of nitrogen dioxide measurements: Chemiluminescent instruments



b) Comparison of WALLOPS/CHEM nitrogen dioxide measurement with PAN

Fig. A3. Comparison of Wallops/Chem NO_2 measurements with other nitrogen species measurements: September 5, 1986.

NO from the ferrous sulfate converter is due to NO_2 reduction. Comparison of the Wallops/Chem, NOCAR/Chem, and PAN data suggests that the higher Wallops/Chem NO_2 results can, allowing for some temporal overlap deficiencies, be accounted for by adding the NOCAR/Chem NO_2 mixing ratios to the PAN mixing ratios after having assumed a reasonably high PAN conversion efficiency of 80% or higher. The second noted difference concerning the data of Figure A2 is the higher correlation between the Wallops/Chem and nitric acid data. This is not the result of conversion of nitric acid to NO by the ferrous sulfate converter but is due to a natural correlation between NO_2 and nitric acid in the air being sampled as the correlation between NOCAR/Chem NO_2 and nitric acid is significant at 0.62. It is noted that the August 23 data of Figure A2 are for a night flight.

The results from Figures A1 and A2 were typical of about 80% of the CITE 2 results in that a relatively high conversion efficiency for PAN to NO by the ferrous sulfate converter was noted. However, for the other data a much lower PAN conversion efficiency was suggested. Figure A3 illustrates these data. It is observed that in Figure A3a the Wallops/Chem NO_2 values are generally lower than NOCAR/Chem, suggesting no converter problems. The agreement between the two instruments in Figure A3a is about the same as that observed between the other NO_2 instruments at mixing ratios of <50 pptv. As previously discussed, NOCAR/Chem and GIT/LIF values agree to within about 20 pptv at mixing ratios of <50 pptv but show poor correlation. This is about the level of agreement shown in Figure A3. The PAN measurements for this flight (Figure A3b) are sufficiently high that conversion efficiencies upward of 50% would be reflected in the Wallops/Chem data as compared to the

NOCAR/Chem results. Thus if these data are not an anomaly, then the CITE 2 results suggest a variable ferrous sulfate converter efficiency for PAN. The nitric acid results are not given in the figure, since values are low, typically below the 20-pptv detection limit of the DENUDER technique.

The data from CITE 2 clearly indicate that the ferrous sulfate converter is sensitive to PAN. While the majority of the data support a relatively high conversion efficiency of 80% or higher for PAN, some data indicate lower conversion efficiencies (less 50%). The Wallops/Chem NO_2 data were correlated with some of the other CITE 2 measurements, for example, NO_y , ozone, air temperature, dew point temperature, and altitude of the measurements. No cause for the apparent variable PAN conversion efficiency of the ferrous sulfate converter could be found.

APPENDIX B: NITROGEN DIOXIDE INSTRUMENTS: INLET TESTS

As part of the CITE 2 flight missions, the NO_2 investigator teams performed a series of inlet tests to define artifact problems which might be associated with chemical reactions occurring in sample inlets, e.g., pernitric acid and/or other NO_x reactions [Ridley *et al.*, 1989]. Test procedures and protocol were left to the discretion of the investigators with the general guideline that the project and other NO_2 investigators be advised (flight planning meetings) as to the test plan. While an inlet test was being performed, the other NO_2 investigator teams were requested to operate their instruments in their normal sampling mode. These investigators' data, along with the data of the investigator performing the inlet test, were analyzed to determine whether significant changes in reported ambient NO_2 values occurred during the

TABLE B1. Test Conditions for Nitrogen Dioxide Inlet Tests

Instrument	Date	Local Time	Type of Air	Altitude, km	Nominal Mixing Ratios,* pptv		Other NO ₂ Data
					NO ₂	NO _y	
GIT/LIF	Aug. 19	1200	mp, mt	5	53 ± 9	158 ± 47	NOCAR, York
	Sept. 5	1100	mp	5	27 ± 3	232 ± 39	NOCAR
	Sept. 5	1400	cp	5	47 ± 3	35 ± 7	NOCAR
	Sept. 5	1600	cp	5	86 ± NA	256 ± 6	NOCAR
NOCAR/Chem	Aug. 15	1500	mt	5	37 ± 5	1100 ± 90	York
	Aug. 21	1300	mt	5	8 ± 7	173 ± 28	York, GIT
	Aug. 31	1400	mp, mt	5	39 ± 13	605 ± 222	York, GIT
	Sept. 2	1500	mp	6	18 ± 16	223 ± 250	York, GIT
York/TDLAS	Aug. 23	0400	mt	5	83 ± 52	NA	NOCAR, GIT

The following abbreviations are used: mp, maritime polar air origin; mt, maritime tropical air origin; cp, continental air origin; and NA, not available or not appropriate.

*Mixing ratio $\pm 1\sigma$.

time period of the inlet tests, and if so, whether these changes were the result of the investigator's having changed the inlet and/or sampling conditions. Data obtained by the investigator performing the inlet test have been omitted from the formal intercomparisons (main text). The GIT/LIF inlet tests involved operating at both lower sample flow and elevated inlet temperature. The NOCAR/Chem tests consisted of heating the inlet from 10°C (normal) to 40°C. The rationale for these procedures was to create inlet conditions in which artifact reaction was favorable and to observe any resulting changes in the recorded NO₂ signal. The York/TDLAS tests used the opposite rationale in that their inlet was cooled for the tests. Inlet tests were performed at high altitudes (5–6 km) and during constant altitude flight. These conditions were thought to provide the best opportunities (among those possible in view of CITE 2 goals) for identifying inlet artifact effects and in particular any pernitr reactions. Wallops/Chem did not participate in the inlet tests. Analyses of the data revealed no conclusive evidence of inlet artifact reactions. While results showed no conclusive effects, test conditions were such that the results are not necessarily sufficient evidence for positively concluding that such problems may not exist for one or more of the instruments/inlets. As a result of ambient fluctuations in NO₂ and the noted "poor correlation" of the techniques at mixing ratios of <50 pptv (most inlet tests were performed at mixing ratios of <50 pptv), one must question whether a sufficient NO₂ mixing ratio comparison reference was established from which to make the judgment as to the presence or nonpresence of an inlet artifact. Table B1 summarizes the test environments in order to help establish those conditions under which the tests were performed and for which no artifact reactions were noted.

Acknowledgments. The authors would like to acknowledge all CITE 2 participants and express thanks for cooperation, support, and professionalism displayed throughout CITE 2. Special recognition is extended to the Wallops Electra flight and ground crews under the direction of Roger Navarro for the long hours spent in the field, preparing and flying the aircraft while at the Ames Research Center. To the hangar crew at Wallops, their support and helpful suggestions during installation of equipment and test flights at Wallops is appreciated. We also acknowledge the support and hospitality of the Ames aircraft/hangar personnel during the Ames-based portion of the field program.

REFERENCES

- Beck, S. M., et al., Operational overview of NASA GTE/CITE 1 airborne instrument intercomparisons: Carbon monoxide, nitric oxide, and hydroxyl instrumentation, *J. Geophys. Res.*, **92**, 1977–1985, 1987.
- Bradshaw, J., and D. D. Davis, Sequential two-photon laser-induced fluorescence: A new method for detecting atmospheric trace levels of NO, *Opt. Lett.*, **7**, 224–226, 1982.
- Bradshaw, J., M. O. Rodgers, S. T. Sandholm, S. KeSheng, and D. D. Davis, A two-photon laser-induced fluorescence field instrument for ground-based and airborne measurements of atmospheric NO, *J. Geophys. Res.*, **90**, 12,861–12,873, 1985.
- Clough, P. N., and B. A. Thrush, Mechanism of chemiluminescent reaction between nitric oxide and ozone, *Trans. Faraday Soc.*, **63**, 915–925, 1967.
- Davis, D. D., J. D. Bradshaw, M. O. Rodgers, S. T. Sandholm, and S. KeSheng, Free tropospheric and boundary layer measurements of NO over the central and eastern Pacific Ocean, *J. Geophys. Res.*, **92**, 2049–2070, 1987.
- Fehsenfeld, F. C., et al., A ground-based intercomparison of NO, NO_x, and NO_y measurement techniques, *J. Geophys. Res.*, **92**, 14,710–14,722, 1987.
- Fontijn, A., A. J. Sabadell, and R. J. Ronco, Homogeneous chemiluminescent measurement of nitric oxide with ozone, *Anal. Chem.*, **42**, 575–579, 1970.
- Fried, A., L. Nunnermacker, B. Cadoff, R. Sams, N. Yates, W. Dorko, R. Dickerson, and E. Winstead, Reference NO₂ calibration system for ground-based intercomparisons during NASA's GTE/CITE 2 mission, *J. Geophys. Res.*, this issue.
- Hastie, D. R., G. I. Mackay, T. Iguchi, B. A. Ridley, and H. I. Schiff, Tunable diode laser systems for measuring trace gases in tropospheric air, *Environ. Sci. Technol.*, **17**, 352a–364a, 1983.
- Hoell, J. M., G. L. Gregory, D. S. McDougal, A. L. Torres, D. D. Davis, J. Bradshaw, M. O. Rodgers, B. A. Ridley, and M. A. Carroll, Airborne intercomparison of nitric oxide measurement techniques, *J. Geophys. Res.*, **92**, 1995–2008, 1987.
- Hoell, J. M., Jr., D. L. Albritton, G. L. Gregory, R. J. McNeal, S. M. Beck, R. J. Bendura, and J. W. Drewry, Operational overview of NASA GTE/CITE 2 airborne instrument intercomparisons: Nitrogen dioxide, nitric acid, and PAN, *J. Geophys. Res.*, this issue.
- Kelly, T. J., D. H. Stedman, J. A. Ritter, and R. B. Harvey, Measurements of oxides of nitrogen and nitric acid in clean air, *J. Geophys. Res.*, **85**, 7417–7425, 1980.
- Kley, D., and M. McFarland, Chemiluminescence detector for NO and NO₂, *Atmos. Technol.*, **12**, 63–69, 1980.
- McNeal, R. J., J. P. Mugler, R. C. Harris, and J. M. Hoell, NASA Global Tropospheric Experiment, *Eos Trans. AGU*, **64**, 561–562, 1983.
- Ridley, B. A., M. A. Carroll, and G. L. Gregory, Measurements of nitric oxide in the boundary layer and free troposphere over the Pacific Ocean, *J. Geophys. Res.*, **92**, 2025–2048, 1987.

- Ridley, B. A., M. A. Carroll, G. L. Gregory, and G. W. Sachse, NO and NO₂ in the troposphere: Technique and measurements in regions of a folded tropopause, *J. Geophys. Res.*, **93**, 15,813–15,830, 1988a.
- Ridley, B. A., M. A. Carroll, A. L. Torres, E. P. Condon, G. W. Sachse, G. F. Hill, and G. L. Gregory, An intercomparison of results from ferrous sulfate and photolytic converter techniques for measurements of NO_x made during the NASA GTE/CITE 1 aircraft program, *J. Geophys. Res.*, **93**, 15,803–15,811, 1988b.
- Ridley, B. A., M. A. Carroll, D. D. Dunlap, M. Trainer, G. W. Sachse, G. L. Gregory, and E. P. Condon, Measurements of NO_x over the eastern Pacific Ocean and southwestern United States during the spring 1984 NASA GTE aircraft program, *J. Geophys. Res.*, **94**, 5043–5067, 1989.
- Sandholm, S. T., J. D. Bradshaw, K. S. Dorris, M. O. Rodgers, and D. D. Davis, An airborne compatible photofragmentation two-photon laser-induced fluorescence instrument for measuring atmospheric NO, NO_x, and NO₂, *J. Geophys. Res.*, this issue.
- Schiff, H. I., G. W. Harris, and G. I. Mackay, Measurement of atmospheric gases by laser absorption spectrometry, in *Chemistry of Acid Rain: Sources and Atmospheric Processes*, ACS Symp. Ser., vol. 349, edited by R. W. Johnson and G. E. Gordon, pp. 274–288, American Chemical Society, Washington, D. C., 1987.
- Schiff, H. I., D. R. Karecki, G. W. Harris, D. R. Hastie, and G. I. Mackay, A tunable diode laser system for aircraft measurements of trace gases, *J. Geophys. Res.*, this issue.
- Shipham, M. C., A. S. Bachmeier, and D. R. Cahoon, Meteorological conditions during the summer 1986 CITE 2 flight series, *J. Geophys. Res.*, this issue.
- Torres, A. L., Nitric oxide measurements at a nonurban eastern United States site: Wallops instrument results from July 1983 GTE/CITE mission, *J. Geophys. Res.*, **90**, 12,875–12,880, 1985.
- J. Bradshaw, D. D. Davis, M. O. Rodgers, and S. T. Sandholm, School of Geophysical Sciences, Georgia Institute of Technology, Atlanta, GA 30332.
- M. A. Carroll, Aeronomy Laboratory, National Oceanic and Atmospheric Administration, 325 Broadway, Boulder, CO 80303.
- A. Fried and B. A. Ridley, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307.
- G. L. Gregory and J. M. Hoell, Atmospheric Sciences Division, NASA Langley Research Center, Hampton, VA 23665.
- G. W. Harris, D. R. Karecki, and G. I. Mackay, Unisearch Associates, 222 Snidercroft Road, Concord, Ontario, Canada L4K 1B5.
- D. R. Hastie and H. I. Schiff, Chemistry Department, York University, 4700 Keele Street, North York, Ontario, Canada M3J 1P3.
- A. L. Torres, NASA Goddard Space Flight Center, Wallops Flight Facility, Wallops Island, VA 23337.

(Received January 9, 1989;
revised September 11, 1989;
accepted September 11, 1989.)