

UC Irvine

Faculty Publications

Title

Results of the Gas-Phase Sulfur Intercomparison Experiment (GASIE): Overview of experimental setup, results and general conclusions

Permalink

<https://escholarship.org/uc/item/3zb95613>

Journal

Journal of Geophysical Research, 102(D13)

ISSN

0148-0227

Authors

Stecher, Hilmar A
Luther, George W
MacTaggart, Douglas L
[et al.](#)

Publication Date

1997-07-01

DOI

10.1029/97JD01362

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Results of the Gas-Phase Sulfur Intercomparison Experiment (GASIE): Overview of experimental setup, results and general conclusions

Hilmar A. Stecher III,¹ George W. Luther III,¹ Douglas L. MacTaggart,² Sherry O. Farwell,² David R. Crosley,³ William D. Dorko,⁴ Paul D. Goldan,⁵ Norbert Beltz,⁶ Udo Krischke,⁶ Winston T. Luke,⁷ Donald C. Thornton,⁸ Robert W. Talbot,⁹ Barry L. Lefer,⁹ Eric M. Scheuer,⁹ Richard L. Benner,¹⁰ Jainguo Wu,¹⁰ Eric S. Saltzman,¹¹ Michael S. Gallagher,¹¹ and Ronald J. Ferek¹²

Abstract. Seven techniques for the field measurement of trace atmospheric SO₂ were compared simultaneously over 1 month in 1994 using samples produced in situ by dynamic dilution. Samples included SO₂ in dry air, in humid air, and in air with potentially interfering gases added. In addition, 2 days of comparison using diluted ambient air were conducted. Six of the seven techniques compared well, with good linear response and no serious interferences but with a range of calibration differences of about 50%.

Introduction

Measurement of SO₂ in the remote atmosphere often requires quantitation over the range of several tens to several hundreds of parts per trillion. Such sensitivity is complicated by the reactivity of this species with surfaces, both with the walls of instrumentation and with atmospheric particles, particularly in the presence of water vapor. Results of the third Chemical Instrumentation and Testing Experiment (CITE 3), an intercomparison involving a number of sulfur gases, bear out the difficulty of measuring SO₂ and, as a logical consequence, of producing accurate calibration standards in the field [Gregory *et al.*, 1993].

This paper describes the experimental details, results, and general conclusions of the Gas-Phase Sulfur Intercomparison Experiment (GASIE) that was conducted in the fall of 1994 in Lewes, Delaware. In this experiment, samples of SO₂ in air ranging in nominal concentration from 20 to 500 parts per

trillion by volume (pptv) were produced in real time by dynamic dilution. During part of the experiment, gaseous species that might interfere with the participating measurement methods were introduced in various concentrations. An automated sulfur gas dilution system (ASGDS) was designed and built at the University of Idaho (UI) and transported to the University of Delaware (UD) in a mobile laboratory. (Note that the UI group has since moved to the South Dakota School of Mines.) The measurement groups (principal investigators (PIs)) were housed in two other mobile labs at UD. The UI group provided fixed, blind concentrations of SO₂ and interference gases over 90-min measurement periods (runs) 5 times each day for 21 days. During two additional days of the experiment, ambient air was diluted with “zero” air (1:5 or 1:10) and measured using the same schedule of five 90-min runs/d. For the historical background and a general description of the experiment and the participants, see *Luther and Stecher* [this issue, preceding paper].

Specific technical attributes of the methods compared in GASIE are given in Table 1. Note that the parameters given in this table were taken from the literature or were obtained by personal communication from the individual PIs. They are included here strictly for the purposes of summary comparisons. For more detailed information the reader is urged to see the individual papers by the PIs which appear in this special section. In like manner, the dilution apparatus used in GASIE is described below in the most general terms (see Experiment); for a thorough account of its design and performance, please see the paper after this one [MacTaggart *et al.*, this issue].

Experiment

Site Setup

The experiment was conducted in three trailers on the Hugh R. Sharp Campus of the University of Delaware, Lewes, Delaware on a site prepared specifically for this experiment. The PI equipment was housed in two 8 × 40 foot (2.44 × 12.19 m) trailers (UD trailers) that were installed parallel to each other.

The 8 × 20 foot (2.44 × 6.10 m) trailer housing the dilution system was installed perpendicular to the UD trailers on

¹College of Marine Studies, University of Delaware, Lewes.

²Graduate Office, South Dakota School of Mines and Technology, Rapid City.

³SRI International, Molecular Physics Laboratory, Menlo Park, California.

⁴National Institute of Standards and Technology, Gaithersburg, Maryland.

⁵NOAA Aeronomy Laboratory, Boulder, Colorado.

⁶Zentrum für Umweltforschung der J. W. Goethe-Universität, Frankfurt am Maim, Germany.

⁷NOAA Air Resources Laboratory, Silver Spring, Maryland.

⁸Department of Chemistry, Drexel University, Philadelphia, Pennsylvania.

⁹Institute for the Study of Earth, Oceans and Space, University of New Hampshire, Morse Hall, Durham.

¹⁰Geophysical Institute and Department of Chemistry, University of Alaska, Fairbanks.

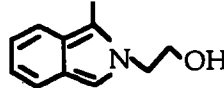
¹¹Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida.

¹²Atmospheric Sciences Department, University of Washington, Seattle.

Copyright 1997 by the American Geophysical Union.

Paper number 97JD01362.
0148-0227/97/97JD-01362\$09.00

Table 1. Measurement Processes and Analytical Details of GASIE Measurement Techniques

Method	Principal Investigator	Prefilter	Physical Interface With SO ₂				
Aqueous chemiluminescence (Aq. CL)	Beltz	none	passed through filter (Delbag Microsorban 98) impregnated with HgCl ₄ ²⁻ solution ^a				
TECO model 43S pulsed-fluorescence detector (mod. TECO)	Luke	2 μ Teflon® or none	pumped through permeation drier into fluor cell (pump downstream of optical cell)				
Isotope dilution-gas chromatograph/mass spectrometer (ID-GC/MS) ^b	Bandy/Thornton	none	~1 ppb SO ₂ (³⁴ S enriched) added/passed through Nafion® drier/cryogenically preconcentrated at -186°C measured by GC/MS				
Mist chamber/ion chromatograph (mist-IC)	Talbot	2 μ Teflon® filter	5 cm 3/8 inch OD glass tube into porcelain-coated steel manifold to glass mist chambers; SO ₂ dissolves in water mist ^c				
Diffusion denuder sulfur chemiluminescence detector (DD-SCD)	Benner	none	sample is alternately scrubbed (K ₂ CO ₃ -coated denuder) and unscrubbed of SO ₂ , then total S measured; SO ₂ calculated by difference				
HPLC/fluorescence detector (fluor-HPLC)	Saltzman	Teflon® filter	liquid/gas exchange coil with 10 μM H ₂ CO and 0.84 M Na ₂ EDTA (pH = 4.6)				
Carbonate filter/ion chromatograph (filter-IC)	Ferek	1 μ Teflon® filter	passed through 90 mm cellulose filter impregnated with 1% K ₂ CO ₃ /10% glycerol ^d				
Method	Chemistry/Fate of SO ₂		Final Analyte				
Aq. CL	SO ₂ + HgCl ₄ ²⁻ + 2H ₂ O → Hg(SO ₃) ₂ ²⁻ + 4HCl Hg(SO ₃) ₂ ²⁻ + 2H ⁺ → Hg ²⁺ + 2HSO ₃ ⁻ Ce ⁴⁺ + HSO ₃ ⁻ → Ce ³⁺ + HSO ₃ // 2HSO ₃ → S ₂ O ₆ ²⁻ + 2H ⁺ S ₂ O ₆ ²⁻ → *SO ₂ + SO ₄ ²⁻ // *SO ₂ → SO ₂ + hν		SO ₂ * hν (chemiluminescence) 450–500 nm				
Mod. TECO	none		SO ₂ hν (fluorescence)				
ID-GC/MS	no chemistry; see column 4 above		³² SO ₂ and ³⁴ SO ₂				
Mist-IC	SO ₂ + H ₂ O → HSO ₃ ⁻ ; add NaCO ₃ to 1 mM (prev. rxn w/H ₂ CO) add H ₂ O ₂ to 0.88 mM; wait 16–24 hours; analyze SO ₄ ²⁻ w/IC		SO ₄ ²⁻				
DD-SCD	SX _n + O ₂ + H ₂ (flame) → SO + "products" ^{ee} SO + O ₃ → SO ₂ * + O ₂ // SO ₂ * → SO ₂ + hν		SO ₂ * hν (chemiluminescence)				
Fluor./HPLC	SO ₂ + H ₂ O → HSO ₃ ⁻ + H ⁺ (at pH 4.6) HSO ₃ ⁻ + borate buffer (pH 9) → SO ₃ ²⁻ SO ₃ ²⁻ + H ₂ NCH ₂ CH ₂ OH + OPA → N-(2-hydroxyethyl)-isoindole-2-sulfonate bring to pH 5.7 with acetate buffer HPLC on C ₁₈ w/fluorescence detector		N-(2-hydroxyethyl)-isoindole-2-sulfonate SO ₃ ²⁻				
Filter-IC	extraction with 0.01% H ₂ O ₂ several hours to oxidize SO ₂ → SO ₄ ²⁻ removal of CO ₃ ²⁻ w/H-form ion chromatography cartridge measurement of SO ₄ ²⁻ using ion chromatography		SO ₄ ²⁻				
							
Method	Estimated LOD	Estimated Uncertainty From All Sources	Distance From Manifold	Flow Rate GASIE/Norm	GASIE Sample Time	Samp. Vol. in GASIE	Standards
Aq. CL	3 ± 1 pptv for 500 L ^f	7–10% of measured SO ₂ value	3.5 m	8/10 slpm	10–35 min ^g	80–320 L	Na ₂ S ₂ O ₅ in 0.01 M HgCl ₄ ²⁻ ; Scott-Marin SO ₂ , 0.183 ppm
Mod. TECO	30 pptv ^h	20 pptv ± 12% measured value	1 m	1/1 slpm	typically 25 min	25 L ⁱ	Scott-Marin SO ₂ , 1.88 ppm
ID-GC/MS	1 pptv for 1.7 L sample ^j	±1.5 pptv at 30 pptv and below ^k	3 m ^l	15/35 slpm	3–8 min (91% were 4 min)	60–120 L ^m	³⁴ SO ₂ calibrated against a gravimetric permeation tube
Mist-IC	5–47 pptv 14 pptv avg ⁿ	10%	2.5 m	28/35–70 slpm ^o	5–30 min ^p	140–700 L avg ~400 L	NIST SRM 3181 SO ₄ ²⁻
DD-SCD	10 pptv ^q	not available	1.5 m	1/1 slpm	3 min	3 L	GC Industries perm. tube cross calibrated with Scott-Marin OCS and SF ₆
Fluor.-HPLC	4 pptv for 6 L ^r	±73 pptv@610 pptv ±5 pptv@12 pptv ^s	1 m	1.5/1–2 slpm	4 min	6 L	VICI Metr. wafer perm tube (11 and 70 ng/min) referenced to Scott SO ₂ 10 ppm ± 2%
Filter-IC	3–12 pptv for 6000 L ^t	~10% of measured SO ₂ conc. ^u	3.5 m	25/200 slpm ^v	90 min	2250 L	filters, SO ₄ ²⁻ s.d. to blank filter; TECO, Scott-Marin 0.205 ppm SO ₂

Further information about how these techniques performed in GASIE and improvements made since GASIE may be found in papers in this issue.

^aCollection efficiency was 98.7% ± 6.4% [Jaeschke et al., this issue].

^bNo paper was submitted in this special section for the ID-GC/MS method.

^cCollection efficiency for mist chambers using water was measured to be 100 ± 2% [Talbot et al., this issue].

^dCollection efficiency is variable with SO₂ concentration and carbonate loading; extensive breakthrough tests before and during the CITE 3 study indicated that below 1 ppb SO₂ the collection efficiency is "essentially 100%" [Ferek et al., 1991]. It appears that the filters can be tuned to the field situation.

^eActual flame chemistry is very complex. For a more complete explanation, see Benner et al. [this issue].

^fThis represents 3 s.d. of the blank value of the HgCl₄²⁻ solution; it does not take into account any other measurement uncertainty at this low level [Jaeschke et al., this issue].

^gSample integration time depended on SO₂ concentration; higher concentrations = shorter times = more replicates per run.

- ^bDetermined as 3 times the standard deviation of four 25-min baseline determinations.
- ^cInstantaneous method; 1 slpm × 25 min averaging time = 25 L [see *Luke*, this issue].
- ^dThis was determined in the laboratory. However, more air than necessary is sampled in real situations.
- ^eThis represents the instrumental variability from all sources. Flow and SO₂ standard (permeation tube) biases are very small compared to the precision at low values, but the standard bias is significant at the higher range of concentrations in GASIE.
- ^fThe ³⁴SO₂ added about 5 cm downstream from the main manifold.
- ^gThis represents the total taken; the amount into GCMS = 0.54 slpm × 4 min. = 2.2 L.
- ^hThis range reflects differences in sample size; LOD is roughly inversely proportional to sample size.
- ⁱRecent conversations with R. Talbot indicated that mist chambers operating at 120 slpm are feasible, representing a fourfold increase in sample volume per unit time over that taken in the GASIE experiment.
- ^jSample integration time not dependent on SO₂ concentration; no auxilliary measurement technique used to estimate concentration and no real-time results.
- ^kDefined as the lowest level that can be identified above noise 50% of the time.
- ^lLOD = 3 × s.d. of the daily blank [*Gallagher et al.*, this issue].
- ^mThese values include the estimated error of the permeation rate of the tube used for standardization [*Gallagher et al.*, this issue].
- ⁿThis represents 3 s.d. of the average blank value of 5 or more unexposed carbonate-impregnated filters; it does not take into account any other measurement uncertainty. Range reflects batch variability in filters.
- ^oAn "implied 2 s.d. measurement uncertainty" of ±2–8 pptv for 6000 L is based on the blank value of the carbonate-impregnated filter. Additional uncertainty in IC analysis and volume measurement are estimated to total 10% of measured value.
- ^pThe 200 slpm is maximum stated in GASIE planning questionnaire; actual flow used in the field depends on sample and SO₂ concentration measured by TECO.

August 21, 1994, and UI began setting up for the intercomparison (see Figure 1). A conduit of 4 inch PVC pipe was installed among the three trailers in which the Teflon® (PFA) manifold and wiring for the associated sensors were housed. The ASGDS and all the monitoring and quality assurance (QA) equipment were tested to confirm the performance specifications of the system at the GASIE field site in Lewes. The ambient temperature in Lewes dropped as the experiment progressed, and it became necessary to insulate the conduit and install a resistive heater inside the manifold to prevent condensation during the phases where water vapor was added.

Details of the installation and performance of the ASGDS are presented by *MacTaggart et al.* [this issue]. Briefly, the UI equipment consisted of a commercially available proportioning-valve dilution apparatus (PVDA) that was extensively modified to allow finer proportion control. Diluent was pro-

vided by a commercial "zero" air generator and was not further purified. Sulfur dioxide and interference gases (except H₂O and O₃) were added from aluminum cylinders of compressed certified mixtures of these gases. Water vapor was added by passing the zero air output through a commercial humidity and temperature controller and adding it to the manifold downstream of the PVDA. Ozone was added from a commercial O₃ generator downstream of the PVDA. In addition, analytical systems monitored the concentrations of SO₂ and all the added matrix gases, as well as cloud condensation nuclei, temperature, and pressure in the manifold. This information was logged in real time and became part of the UI data record.

The PIs, with the exception of the ID-GC/MS group from Drexel, arrived the week of September 5, 1994, and installed their equipment in the UD trailers as shown in Figure 1. The Drexel group was involved in another field program and expe-

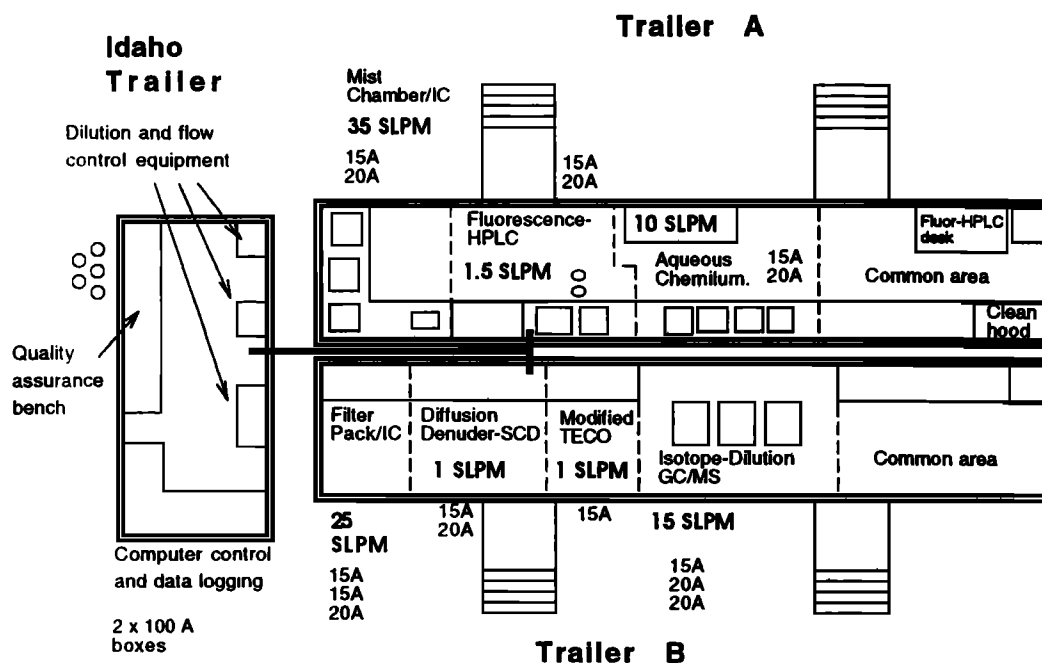


Figure 1. Diagram of the trailer layout in GASIE showing the orientation of principal investigators (PIs) and the University of Idaho (UI) group, their power and circuit requirements, their flow allotment, and their approximate footprint in the trailers.

Table 2. GASIE Chronology

Phase	Date	Run	Sample Type
1	Sept. 11–15	1–25	SO ₂ in dry, interference-free air
2a	Sept. 17–21	26–50	SO ₂ in humid air
3a	Sept. 24–27	51–70	SO ₂ in dry air with O ₃ and NO _x added
3b	Sept. 29 to Oct. 3	71–90	SO ₂ in dry air with CO, CO ₂ , CH ₄ , and DMS added
2b	Oct. 5–7	91–105	SO ₂ in humid air
4	Oct. 8–9	106–115	diluted ambient air

rienced unforeseen delays. Their equipment arrived in Lewes on September 13 and was ready to collect data on September 20.

Schedule

The intercomparison measurements began as scheduled on September 11, 1994, and proceeded according to the schedule shown in Table 2. Five 90-min runs, serially numbered throughout the experiment, were conducted each day starting at 0900, 1100, 1330, 1530, and 1730. A small number of predetermined nominal concentrations of SO₂ (set points) were used so that statistical techniques requiring replication could be employed. These set points were chosen to allow adequate assessment of techniques with a wide range of detection limits. The detailed schedule of set points, along with the nominal concentrations of interferences, is given in Table 3. The statistical design rationale behind these set points is given by *MacTaggart et al.* [this issue].

At the beginning of each day, the UI group started delivering the concentration for the first run of the day and then synchronized their computer clock to digital clocks with one second resolution in each of the UD trailers. This fixed concentration was delivered throughout the run and at least 2 min into the changeover period between runs, at which time the concentration for the next run was set. It was determined both in the field and in the UI laboratory that within the precision of the QA instrumentation the SO₂ concentration required ≤ 20 min to stabilize in the manifold. This was borne out by examination of the time-resolved PI data. The PIs were asked to make as many replicate measurements within the 90-min run as possible, but the timing of these was at their discretion. They were also asked to calculate their best estimate of the SO₂ concentration for each replicate and to report it by the following day to the UD group (see Data Handling below). After the last run of the day, the manifold was flushed with zero air until the next morning.

On October 7, 1994, after the conclusion of phase 2b the delivery manifold was modified for the diluted ambient air experiment, as shown in Figure 2. Ambient air was introduced into the sampling manifold using two metal bellows pumps run

in parallel. Each pump had a nominal, no-load flow rating of 12 lpm.

The dilution factor was estimated by measuring the flow between the metal bellows pump and the connection with the zero air line with a mass flowmeter (MFM), while the zero air was adjusted to deliver a total flow of 100 standard liters per minute (slpm). The MFM was removed before sampling by the PIs began. This system lowered the ambient SO₂ concentration to levels more relevant to this intercomparison while retaining all of the matrix components in ambient air, albeit at reduced levels because of the dilution. This is particularly significant for CO₂ and H₂O and may be for other interferences as well. The quality assurance instrumentation in the UI trailer used in the first three phases (continuous total sulfur monitor, GC/SCD sulfur species detector, cloud condensation nuclei (CCN) counter, and O₃ and NO_x monitors [see *MacTaggart et al.*, this issue]) was employed to characterize the diluted ambient air and to track the temporal variability of the total sulfur concentration.

Data Handling Procedure

The following procedure was used to strike a balance between maintaining blindness and monitoring quality during the data collection. Each day, UI would report to one member of the UD group (Luther) the nominal set points from the previous day. Another member of the UD group (Stecher) was designated to receive the PI data (start/stop times for each replicate, replicate SO₂ concentrations, and any comments) the day after they were generated, to code them and then to pass them along to Luther. Because the methods had widely different integration times and inherent precision levels, the within-run variability and number of replicates were removed by reporting only the run mean for each investigator. Luther compared this coded, run-averaged data to the theoretical set points to monitor the progress of the experiment. Thus Stecher monitored the within-run homogeneity of the samples delivered and Luther the correspondence of the PI data with the ASGDS set points. These coded, run-averaged data were shared with the Advisory Committee on a regular basis.

In the case of one PI, the filter pack-IC system, it was necessary to modify this procedure. Under normal field conditions, this PI collects filters in the field and sends them back to the home laboratory for analysis. Because of problems at the home lab, which did not become apparent until after the experiment started, this PI was not able to report data within the time frame necessary to assess the progress of the experiment. However, as a matter of course, this PI continuously monitors the incoming sample with a TECO pulsed-fluorescence instrument to determine the necessary sample time. Starting on September 20 the results from this TECO became the preliminary data set for this method because of the good correlation between the TECO and filter data. In the final analysis of

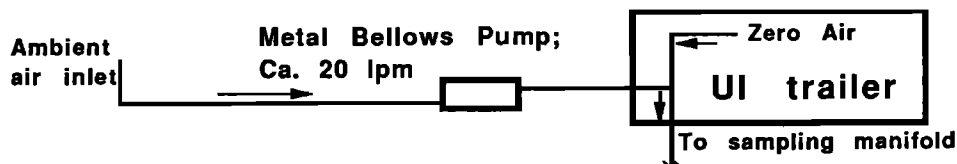


Figure 2. For the diluted-ambient portion of the experiment, air was drawn through a metal bellows pump, diluted with zero air from UI system and delivered to the PIs through the manifold used throughout the experiment.

Table 3. Nominal Set Points for SO₂ and Interference Gases by Run

Phase 1: SO ₂ in Dry Zero Air With No Interferences			Phase 2a: SO ₂ With Water Vapor				Phase 3a: SO ₂ With O ₃ and NO _x				
Date	Run	SO ₂ pptv	Date	Run	SO ₂ pptv	RH %	Date	Run	SO ₂ pptv	O ₃ ppbv	NO _x ppbv
Sept. 11	1	41	Sept. 17	26	18	35	Sept. 24	51	0	100	10
	2	260		27	127	35		52	41	100	10
	3	18		28	501	35		53	127	100	10
	4	501		29	41	70		54	501	100	10
	5	64		30	127	0		55	127	0	0
Sept. 12	6	260	Sept. 18	31	41	35	Sept. 25	56	501	100	10
	7	18		32	501	70		57	127	100	10
	8	0		33	127	0		58	260	100	10
	9	64		34	0	50		59	0	100	10
Sept. 13	10	501	Sept. 19	35	501	50	Sept. 26	60	41	100	10
	11	501		36	260	0		61	127	100	10
	12	41		37	0	0		62	18	100	10
	13	64		38	18	0		63	41	100	10
	14	18		39	501	0		64	260	100	10
Sept. 14	15	0	Sept. 20	40	260	0	Sept. 27	65	501	100	10
	16	64		41	0	35		66	41	0	0
	17	501		42	41	35		67	260	0	0
	18	41		43	18	35		68	501	0	0
	19	0		44	501	35		69	127	0	0
Sept. 15	20	260	Sept. 21	45	127	35	70	18	100	10	
	21	0		46	18	35					
	22	64		47	0	35					
	23	260		48	41	35					
	24	127		49	127	35					
	25	501		50	501	35					

Phase 3b: SO ₂ With CO, CO ₂ , CH ₄ , DMS							Phase 2b: SO ₂ With Water Vapor				Phase 4: Diluted Ambient Air		
Date	Run	SO ₂ pptv	DMS pptv	CO ₂ ppmv	CH ₄ ppmv	CO ppbv	Date	Run	SO ₂ pptv	RH %	Date	Run	SO ₂ pptv
Sept. 29	71	0	100	350	2	120	Oct. 5	91	0	50	Oct. 8	106	amb/4.98
	72	41	100	350	2	120		92	18	50		107	amb/5.15
	73	501	100	350	2	120		93	41	50		108	amb/5.15
	74	127	0	0	0	0		94	127	50		109	amb/5.15
	75	18	100	350	2	120		95	501	50		110	amb/5.15
Sept. 30	76	18	100	350	2	120	Oct. 6	96	18	50	Oct. 9	111	amb/5.00
	77	501	100	350	2	120		97	41	50		112	amb/10.09
	78	18	0	0	0	0		98	18	50		113	amb/5.15
	79	260	100	350	2	120		99	501	50		114	amb/5.15
Oct. 1	80	501	0	0	0	0	Oct. 7	100	127	50	115	amb/5.15	
	81	127	100	350	2	120		101	0	50			
	82	41	100	350	2	120		102	127	50			
	83	260	100	350	2	120		103	501	50			
	84	0	100	350	2	120		104	127	0			
Oct. 3	85	18	0	0	0	0	105	41	50				
	86	41	100	350	2	120							
	87	260	0	0	0	0							
	88	290	100	350	2	120							
	89	501	100	350	2	120							
	90	127	0	0	0	0							

coded results, the filter pack data were substituted for the TECO data.

As mentioned above, one method (the ID-GC/MS) was absent for a substantial amount of the experiment. In order to protect this investigator's anonymity, "dummy" numbers were manufactured by Stecher, passed along to Luther, and became part of the "preliminary data set" used for the blind data evaluation. It was determined after the code was broken that there were no faulty conclusions reached as a result of the dummy data.

Amendments to the Procedure

As with any field experiment, unforeseen circumstances occur which require changes in planned procedure. Significant changes are explained below.

1. On September 12 during run 10 the SO₂ concentration dropped dramatically and then recovered. This drop was observed with the UI continuous SCD and by several of the investigators who had real-time monitoring capability. This happened again on September 13 during run 11; once again,

Table 4. PI Results: Run Means of SO₂ (pptv), Within-Run Standard Deviation, and Number of Replicates

Run	Set	Phase	UI/QA	Aq. CL		Mod TECO		ID-GC/MS		Mist-IC		DD-SCD		Fluor.-HPLC		Filter-IC							
51	0	3a		12	6	2	21	6	2	39	3	11	7 ^a	7 ^a	6	170	10	7	3 ^a	3 ^a	2	13	1
52	41	3a	61	41	3	3	67		1	83	17	11	51	6	6	167	12	7	61	2	3	60	1
53	127	3a	155	130	18	3	173	3	2	222	6	11	153	7	6	236	11	7	195	4	4	156	1
54	501	3a	530	494	38	3	617	12	2	729	6	11	533	15	6	515	11	7	584	3	4	567	1
55	127	3a-z	129	129	2	2	162	7	2	185	3	14	146	7	6	168	11	7	196	4	3	155	1
56	501	3a	545	608	22	3	601	18	2	726	6	15	531	26	6	600	19	8	520	6	4	537	1
57	127	3a	143	128	23	3	161	8	2	199	4	13	159	10	5	222	9	8	151	5	5	140	1
58	260	3a	293	272	7	3	328		1	389	10	14	305	16	6	302	11	8	305	2	4	266	1
59	0	3a		2	2	2	10 ^a	10 ^a	2	10	1	5	19	3	5	90	13	8	2 ^a	2 ^a	3	22	1
60	41	3a	71	25	6	2	63		1	73	3	9	68	8	5	120	6	8	56	1	4	73	1
61	127	3a	159	133	2	3	168		1	184	3	13	143	8	5	224	21	7	151	4	4	134	1
62	18	3a		19	8	2	23		1	33	1	12	39	7	5	96	9	7	36	1	4	34	1
63	41	3a	60	33	2	3	54		1	73	1	10	64	7	5	103	10	7	70	3	4	63	1
64	260	3a	293	256	12	3	317		1	385	8	10	306	7	4	257	9	7	315	2	4	290	1
65	501	3a	558	469	18	3	609		1	729	5	14	538	22	4	454	9	7	557	3	4	582	1
66	41	3a-z	55	79	8	3	73	5	2	76	1	11	71	7	4	84	8	8			0	68	1
67	260	3a-z	314	350	37	3	336	0	2	405	5	8	311	15	4	296	24	8	344	3	3	282	1
68	501	3a-z	568	677	48	3	592	9	2	727	4	13	553	21	5	466	19	8	535	4	5	570	1
69	127	3a-z	170	192	25	3	166	23	2	198	4	8	173	18	5	166	25	8	171	3	4	179	1
70	18	3a		9	6	2	39		1	41	1	8	37	4	6	110	18	8	21	2	3	86	1
71	0	3b		4 ^a	4 ^a	3	10 ^a	10 ^a	2	2		0	26	4	3	-66	30	8	2 ^a	2 ^a	4	15	1
72	41	3b	60	40	29	3	59	2	2		0	85	11	6	-26	20	8	54	2	4	110	1	
73	501	3b	526	650	52	3	620	3	2		0	550	14	6	403	14	8	533	7	3	552	1	
74	127	3b-z	139	169	5	3	172	12	2		0	175	11	6	146	13	8	153	1	3	172	1	
75	18	3b	27	23	4	3	32	10	2		0	71	8	5	-42	10	8	25	0	3	25	1	
76	18	3b	26	34	2	3	32	15	2	30	1	9	41	8	5	-72	43	8	25	1	4	88	1
77	501	3b	588	580	23	3	648	6	2	696	6	12	553	29	6	438	24	8	552	11	4	604	1
78	18	3b-z	16	33	10	2	22 ^b	17 ^b	2	30	1	10	46	14	6	52	10	8	29	2	3	34	1
79	260	3b	283	322	8	3	317	4	2	354	6	5	292	12	6	200	10	8	307	5	3	292	1
80	501	3b-z	555	530	10	3	648	10	2	675	59	9	546	20	6	558	22	7	571	10	5	577	1
81	127	3b	163	148	10	3	181	16	2	193	8	10	148	9	6	90	15	7	144	3	3	118	1
82	41	3b	62	45	4	2	68	9	2	68	3	11	53	4	6	-41	9	7	55	1	7	43	1
83	260	3b	292	283	15	3	330	6	2	375	8	12	282	8	6	209	17	7	295	11	3	263	1
84	0	3b		2 ^a	2 ^a	2	10 ^a	10 ^a	2	1 ^a	1 ^a	10	4 ^a	4 ^a	5	-95	25	7	2 ^a	2 ^a	3	0	1
85	18	3b-z	37	23	3	2	16	12	2	31	1	10	23	4	5	51	27	7	29	1	3	31	1
86	41	3b	58	42	6	2	62	6	2	62	1	9	50	6	6		0	53	1	2	86	1	
87	260	3b-z	301	303	18	2	341	17	2	373	3	12	291	12	6		0	293	2	4	285	1	
88	290	3b	296	334	37	3	375	14	2	404	10	12	305	14	6		0	331	6	3	345	1	
89	501	3b	578	579	49	3	627	17	2	692	15	13	516	24	6	315	22	7	565	18	4	590	1
90	127	3b-z	164	154	6	3	148	17	2	185	8	10	146	3	6	126	38	7	176	1	3	168	1
91	0	2b		10	1	2	10 ^a	10 ^a	2	9	4	9	5 ^a	5 ^a	5	532	109	8	1 ^a	1 ^a	3	6	1
92	18	2b	29	19	3	2	51	7	2	33	4	11	19	1	6	311	70	8	27	1	3	14	1
93	41	2b	61	42	2	2	65	0	2	62	1	11	39	3	5	361	101	8	55	3	3	43	1
94	127	2b	164	125	6	3	181	3	2	174	4	11	118	8	6	237	22	8	151	3	4	94	1
95	501	2b	524	415	12	3	608	0	2	652	8	9	461	30	6	522	105	8	553	13	3	455	1
96	18	2b	25	30	1	10 ^a	10 ^a	2	30	3	9	20	3	6	327	71	8	24	1	5	12	1	
97	41	2b	55	57	0	2	74	8	2	65	1	11	43	2	5	247	36	8	52	3	3	38	1
98	18	2b	27	31	5	2	23	5	2	31	1	11	20	2	5	352	115	8	27	2	4	13	1
99	501	2b	529	620	17	3	622	7	2	688	14	13	509	17	5	421	60	7	539	8	4	447	1
100	127	2b	168	179	12	3	176	7	2	185	5	11	135	5	6	433	82	7	165	2	3	132	1
101	0	2b		3 ^a	3 ^a	2	11	8	2	1 ^a	1 ^a	7	21	2	5	785	163	7	1 ^a	1 ^a	4	37	1
102	127	2b	141	133	5	3	161	5	2	172	5	12	147	11	6	496	53	7	139	5	5	163	1
103	501	2b	544	545	34	3	635	30	2	655	6	11	491	23	6	693	137	7	498	13	3	467	1
104	127	2b-z	191	188	26	3	191	12	2	213	2	10	182	13	6	160	23	7	185	5	4	237	1
105	41	2b	52	55	4	2	65	12	2	60	2	12	74	8	6	571	71	7	63	3	3	67	1

1. UI quality assurance data (continuous SCD for set points above 100 pptv and GC-SCD for the 18 and 41 pptv set points) are included to provide an independent measurement against which to plot the PI data. The nearly unanimous feeling among the PIs and the Planning Committee was that the UI-QC data provided a better *x* axis for regression than the set points (based on the readability of the plots), although there is no valid statistical basis for this.

2. Any value reported as "below the limit of detection" was recoded as $0.5 \times (\text{limit of detection}) \pm 0.5 \times (\text{limit of detection})$ as stated by the PI. Some PIs calculated an average LOD for the whole experiment, some for each run and some for each replicate. The most detailed system provided by the investigator was used in calculating the run statistics. Superscript-a values indicate runs in which all replicates were reported below the detection limit. Superscript-b values indicate that at least one replicate of the run (but not all) was reported as below detection limit and that $0.5 \times (\text{det. lim.})$ was used to calculate the average and the standard deviation. This procedure, while statistically unsatisfactory, was undertaken to provide a rough numerical estimate of measurements near an instrument's limit of detection for the purposes of plotting. This is particularly useful in the figures demonstrating low-level performance (Figure 4).

3. Limits of detection for the fluorescence-HPLC system depend on daily calibration and thus can change from day to day. In some cases, drift in the daily blanks required the calibration to be corrected for drift. For these days the maximum daily LOD for the experiment (excluding phase 2a) was used (5.6 pptv), and the value 2.8 pptv was taken as the estimate of "below LOD" for run averages.

4. Some of the investigators (aq. chemilum., mod. TECO, ID-GC/MS, mist-IC) reported replicate values for at least one run that represented different integration times. Strictly speaking, the mean of these replicates should be a time-weighted mean. In order to avoid the problems associated with calculation of time-weighted standard deviations, the means reported herein are not time-weighted.

5. There were 38 measurements reported in which all or part of the integration time fell outside the run window due to accidental mistiming, typographical or transcribing error, etc. The following replicates were discarded before the run statistics were calculated: any replicate for which more than half of the integration time fell outside the run window; any replicate for which the start time was reported to be more than 5 min before the start of a run or the stop time was reported to be more than 2 min after a run was discarded before the average was computed. The following replicates were included in calculation of the run statistics: any runs in which the start time was reported to be less than 5 min before the start of a run and the stop time was reported to be less than 2 min after a run, and in which more of the integration time fell within the run window than outside of it.

6. Because none of the techniques (as used during GASIE) were capable of taking more than one sample simultaneously, any replicate datum (measurement) which overlapped in time with any other was discarded. This resulted in the deletion of 13 data points in four runs for the ID-GC/MS.

several of the PIs noticed it. The unintentional "drop-outs" in SO₂ concentration during runs 10 and 11 were due to electrical noise being introduced to the ASGDS connection to its on-line PC triggering a shut-down command to the PVDA. When this was discovered, the PVDA was operated manually. Results for these runs are provided but were not used in any data analysis. It must be noted that all PIs with the temporal resolution sufficient to track these drop-outs did so.

2. On September 15 (the last day of phase 1), with data from the first seven runs, it became clear that the set points were significantly higher than the PI data. (The delay in obtaining the data was due to the late arrival of filter pack results from the home lab; see above.) Luther indicated to the UI group that the PIs were reporting concentrations significantly different from the set points and suggested that they perform checks on the ASGDS system to verify its performance. The UI personnel searched for leaks but found none at that time (see item 4 below).

3. On September 18 during runs 32 and 34, large droplets of water were observed in the delivery manifolds in both UD trailers. It was discovered that the ambient temperature in trailer B was 4°C colder than in either of the other trailers (see Figure 1). In addition, the manifold plumbing in trailer A had been incorrectly plumbed such that a substantial local pressure drop was possible. These problems were corrected; in addition, the conduit, which housed the manifold, was insulated, and a resistive heater was installed. A number of the PIs expressed concern about the control of water vapor in the manifold, and they requested that RH, temperature, and pressure sensors be placed in the manifold with displays for PI use. The Planning Committee agreed with this and modified the schedule to allow the appropriate equipment to be ordered and installed. This required splitting phase 2 into two parts, with the second part, phase 2b, near the end of the experiment. Information from these additional sensors was also monitored and recorded by UI.

4. A storm on September 22 (rest day between phases 2a and 3a) caused the electric power in Lewes to fail at ~1600. The power was restored at ~1730 but then failed again shortly thereafter. The second failure only affected the trailers, but this was not discovered and corrected until 0830 on September 23. Several of the investigators and the Idaho group needed time on September 23 to warm up and test their instrumentation. The schedule was revised to make September 23 an off day and to reduce the number of days in phase 4 to two. The ASGDS purged the manifold with dry zero air for part of the day and then delivered a constant, high concentration of SO₂. During the day off after the power outage the UI team conducted a systematic search for leaks in the ASGDS. Two gas leaks were discovered, one in the mixing chamber and another in the fittings within the proportioning valve dilution apparatus. These leaks were repaired; it is not known why they went undiscovered in the previous examinations.

5. During the first run of October 9 (run 111, the second

day of phase 4) the total sulfur was extremely high and variable as judged by the UI continuous SCD. The dilution factor was doubled (from 1:5 to 1:10) in order to bring the SO₂ down to a more relevant level. The next run (run 112) showed a much lower SO₂ concentration, and the dilution factor was returned to its original value for the remainder of the experiment.

Data Evaluation and Dissemination

The data turned in during the experiment constituted the preliminary data. The PIs were given 2 months (until December 15, 1994) to recalculate these data based on instrument and standard recalibrations and submit it to Stecher of the UD group. This period was considered to be consistent with the processing of typical field data. The final run averages were calculated, the dummy numbers were revised slightly to protect their anonymity, and the data set was annotated to explain the processing methods used. The same code was maintained, and the final coded, run-averaged data were disseminated to the rest of the Planning Committee and to the PIs on January 3, 1995. Note that this was the first time the PIs had any knowledge of the SO₂ and interference set points.

The blind data analysis carried out by the Planning Committee culminated in a meeting (March 30 to April 1, 1995) at which the conclusions reached by each member during the previous three months were discussed and refined. The conclusions reached at this meeting were integrated into a preliminary report by UD which was distributed to the PIs and to National Science Foundation (NSF) on May 26, 1995 [Stecher *et al.*, 1995]. The last page of that report contained the code key and the key to the dummy values; this was the first time that anyone except Stecher had access to this information.

After the preliminary report was issued, the final replicate data were compiled into a single electronic file and distributed to the Planning Committee and to the PIs. This full data set is available in electronic form from George Luther at the University of Delaware, Lewes. The Planning Committee and the PIs met together in January 1996 to discuss the final conclusions from the project and the presentation and publication of the work. Conclusions from that meeting form the basis of this special section.

Results and Discussion

Because of the enormity of the raw data set, only the most relevant subset of the data will be considered here. Because of problems with the dilution system and the accompanying set point inconsistencies outlined in the Experiment section, only results for phases 2b, 3, and 4 will be discussed in detail; discussion of phases 1 and 2a are given by Stecher *et al.*, [1995]. The run-averaged data, within-run standard deviations, and the number of replicates for phases 2b and 3 are given in Table 4.

Results for the interference phases (phases 3a, 3b, and 2b) are shown in Plate 1 as plots of average PI response for a given

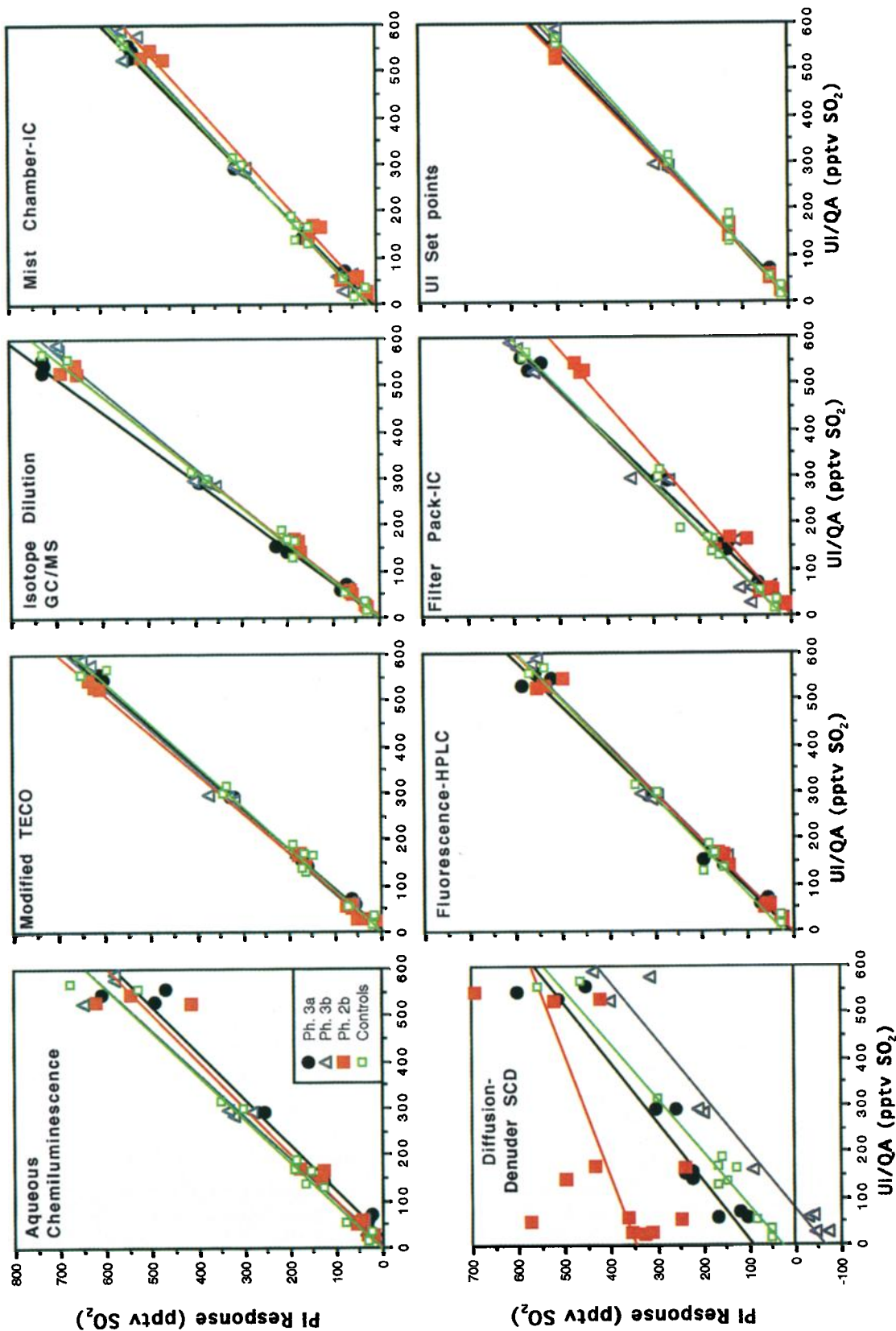


Plate 1. Plots of SO₂ measured by each PI versus UI/quality assurance (QA) data separated by phase (phase 3a, 100 ppb O₃/10 ppb NO_x; phase 3b, 350 ppm CO₂/2 ppm CH₄/120 ppb CO/100 ppt DMS; phase 2b, 50% RH). Control samples were scattered within all three phases and contained only SO₂ in dry “zero” air. Parameters for the one-way least squares linear regression lines shown are given in Table 5a. The plot of set point versus UI/QA data is provided for comparison; the negative intercept suggests that a small amount of SO₂ was present in the zero air. This is supported by the regression statistics for plots of PI data versus the set points (Table 5b).

run versus the UI quality assurance (UI/QA) data for that run. It must be noted that this unusual format was chosen during discussions with the PIs during a meeting in January 1996 in preference to plotting against the set point or one of the PI data sets for three reasons. First, while the delivered concentration of SO₂ within each run was quite consistent, the concentration did vary from one run at a given set point to another (see Precision section). Although the UI/QA system is not a "proven" technique, it would be expected to (and indeed did) respond to these variations in actual SO₂ concentration. Secondly, the UI/QA data set provided a neutral variable against which to plot the PI data. It was judged unfair to use the data of any one PI as an *x* axis for comparison because it would suggest a superiority or a validation of that technique. Finally, using the set points leads to very cluttered plots that are quite difficult to interpret because there is no variation in the *x* direction within a group of data from a single set point. Using the UI/QA data as an *x* axis maintains the relative ordering of responses for the investigators but separates the points for each set point horizontally so that the PI data may be viewed more easily.

One-way regression parameters for the plots of PI response versus UI/QA data shown in Plate 1 are given in Table 5a. The one-way regression parameters for plots of PI data versus the set points yield valuable information about the dilution system and are presented in Table 5b. Finally, the one-way regression parameters for plots of data from each versus each other PI are presented in Table 5c in order to make direct comparisons.

The data for the diluted ambient air (phase 4) are presented graphically in Plates 2a and 2b. In this case, time-resolved plots of the individual replicates are plotted so that both the instrumental response and the ability of the PIs to track changes in concentration can be compared. Given the wide variation in integration time and number of replicates per run, the methods cannot easily be compared quantitatively, as expected from the CITE 3 results [Gregory *et al.*, 1993]. However, viewing the time-resolved results this way allows a direct qualitative comparison.

General PI Response

Six of the seven techniques compared well with one another with no evidence of interference from O₃, NO_x, DMS, CO, CO₂, or CH₄ and only minor interference from water vapor (see Interferences section). In addition, these techniques tracked one another quite well during the diluted ambient air portion of the experiment (see Results with Diluted Ambient Air). The other method exhibited mild to severe difficulties during the interference phases, but these problems appeared reversible (see Interferences section). When the data from the investigators were plotted against the UI quality control data (Table 5a), against the set points (Table 5b) or against each other (Table 5c), good linear relationships were observed with a range of slopes of about 50%. The broad range of slopes is attributed to calibration differences among the laboratories and is discussed under Accuracy.

Precision

Before any statement of precision of an analytical technique can be made, the variability of the samples being measured must be assessed; this is particularly important when the samples are being prepared in situ. Upper limit estimates of the variability in SO₂ concentration delivered by the ASGDS were based on the precision of the PI data because, collectively, the

GASIE PIs represent the current state of the art in SO₂ measurement. The ASGDS variability was assumed to be at least as precise as the most reproducible of the PI data. The within-run precision for the seven techniques was estimated by plotting the average of the within-run standard deviations for each of the six set points from phases 3a, 3b, and 2b (nominally 0, 18, 41, 127, 260, and 501 pptv) against the average response for those set points (see Plate 3a). The intercepts of the resulting regressions estimate the constant variability of an instrumental technique regardless of concentration and measures the error in the blank over the timescale of interest. The slopes estimate how the variability of the instrumental response depends on (or scales with) concentration. The run-to-run precision was treated similarly and is shown in Plate 3b.

Using the parameters obtained from Plate 3, the upper limit estimate of the average ASGDS variability within a run is $\pm(1.5 \text{ pptv} + 1.5\% \text{ of the set point})$, based on the data for the ID-GC/MS and fluorescence-HPLC methods. The differences in PI calibration mentioned above will have an effect on the concentration-dependent term, but this amounts at most to a few tenths of a percent. (For a more thorough discussion of the PI calibrations, see the Accuracy section.) Similarly, the run-to-run precision of the ASGDS (i.e., the ability of the dilution system to return to a previously delivered concentration after being set at a different value) was estimated to be at or below $\pm(7 \text{ pptv} + 2\% \text{ of the set point})$. This estimate is a composite of the data from the modified TECO, the ID-GC/MS, the mist chamber-IC, and the fluorescence-HPLC methods.

With this in mind, it is possible to assess precision for the methods themselves, keeping in mind that at least some of the ASGDS variability must be subtracted from the values obtained for the PIs. The regression parameters from Plates 3a and 3b are summarized in Table 6 and must be treated as upper limits of PI variability. Note that the filter pack method generated a single measurement for each run, so no within-run analysis was possible. Also note that for the diffusion-denuder SCD, only the data from control runs were used.

Three items are noted. First, the aqueous chemiluminescence system is set up to trap SO₂ samples in the field as the bis-sulfitomercurate complex and to analyze the resulting stable solutions in the laboratory after the field work. This procedure allows more time to construct the calibration curve and allows all the samples to be run on the same calibration curve, leading to much better precision. During the GASIE campaign, separate calibration curves were constructed daily by this group, as mandated by the experimental protocol. Thus it may be expected that this method would exhibit more precise data under normal operating conditions. Secondly, both the mist chamber-IC and the filter pack-IC methods were constrained to sample sizes lower than they could have taken (by factors of about 2 and 3, respectively). Since a large portion of the error in both methods comes from random errors in the blank, a larger sample should significantly increase the signal/noise ratio. Thus a more accurate estimate of the precision requires that the intercepts of these regression data be adjusted down accordingly. Finally, the modified TECO system showed by far the least dependence of variability on the measured concentration, both within run and run to run, suggesting a single, constant source of random error.

Accuracy

Determination of accuracy for specific techniques is problematic because there is no validated standard for SO₂ at these

Table 5a. Summary Regression Statistics of PI Response Versus UI/QA Data for Late Phases

Method	Parameter	Phase 3a	Phase 3b	Phase 2b	Late, No Controls	Late, Controls	All Late Runs
Aq. CL	slope	1.103	1.088	0.996	1.032	1.068	1.038
	S.E. slope	0.06	0.053	0.067	0.035	0.056	0.031
	intercept	-28.8	-11.63	-4.86	-13.84	0.59	-9.5
	S.E. Int	19.2	17.02	18.77	10.91	15.72	9.21
Modified TECO	R^2	0.9697	0.977	0.9565	0.9627	0.9734	0.9626
	slope	1.135	1.123	1.17	1.14	1.107	1.134
	S.E. slope	0.018	0.026	0.016	0.012	0.04	0.013
	intercept	-8.75	2.28	-3.58	-2.96	-4.6	-3.79
ID-GC/MS	S.E. int	5.96	8.16	4.55	3.84	11.32	4.01
	R^2	0.9976	0.995	0.9981	0.9961	0.987	0.9938
	slope	1.35	1.207	1.261	1.279	1.252	1.273
	S.E. slope	0.025	0.035	0.022	0.02	0.03	0.017
Mist-IC	intercept	-7.28	4.62	-9.8	-5.15	-2.86	-4.77
	S.E. int	7.94	11.32	6.24	6.31	8.81	5.16
	R^2	0.997	0.9943	0.9969	0.9924	0.9948	0.9928
	slope	0.983	0.926	0.92	0.947	0.949	0.947
DD-SCD	S.E. slope	0.02	0.035	0.027	0.019	0.027	0.016
	intercept	2.77	17.18	-4.74	4.43	13.2	6.82
	S.E. int	6.49	11.13	7.63	5.71	7.71	4.71
	R^2	0.9962	0.9863	0.9913	0.9875	0.9918	0.9879
Fluor. HPLC	slope	0.778	0.821	0.381	0.567	0.856	0.635
	S.E. slope	0.075	0.062	0.166	0.132	0.053	0.105
	intercept	85.51	-66.51	340.75	158.47	24.7	124.81
	S.E. int	24.3	21.26	46.42	41.18	14.76	32.15
Filter IC	R^2	0.924	0.9557	0.3443	0.374	0.967	0.4635
	slope	1.011	0.988	0.994	0.999	0.959	0.991
	S.E. slope	0.039	0.027	0.026	0.017	0.044	0.016
	intercept	5.94	1.34	-0.94	1.72	18.61	5.29
Filter IC	S.E. int	1.25	8.6	7.34	5.19	12.8	4.9
	R^2	0.987	0.9928	0.9931	0.9907	0.9817	0.9886
	slope	1.04	0.995	0.865	0.975	0.979	0.975
	S.E. slope	0.028	0.048	0.032	0.029	0.036	0.024
Filter IC	intercept	-10.46	14.24	-5.09	-0.42	14.64	3.69
	S.E. int	8.98	15.42	8.91	9.04	10.25	7.32
	R^2	0.9936	0.9775	0.9866	0.9711	0.9864	0.9729

low levels. Standards in this range must be produced dynamically (diluted from higher concentrations for immediate use), and validation involves certifying both the standard and the dilution system (including the diluent). However, because this experiment probably represents the best technology available for the measurement of SO_2 in this concentration range, the "calibration differences" discussed above can be used to estimate an experiment-wide (and, it might be argued, community wide) accuracy. These differences (Table 7) are taken from the regression statistics given in Table 5a. The absolute numbers have no absolute relevance because they are based on the UI/QA data, but they may be used to compare the calibrations of the techniques. This is displayed graphically in Figure 3 and indicates that the best accuracy currently attainable for SO_2 below a half a part per billion within the atmospheric community is about $\pm 25\%$. The implication is that a report of 500 pptv

from any given investigator constrains the value of SO_2 at best between about 375 and 625 pptv.

The fact that the precision of most of the methods is much better than the range of calibrations indicates that the major source of error in absolute measurement of SO_2 below 500 pptv lies with the standardization and not the measurement itself. This suggests that there is a substantial discrepancy in either the commercial standards as obtained by the PIs and UI or in the way these standards are diluted during calibration (or in the case of UI, during delivery). A summary of the calibration standards appears in Table 1 and more detailed procedures in the accompanying papers by the PIs.

Interferences

As discussed in the Experiment section, the interferences examined were grouped into three phases: water vapor (phases

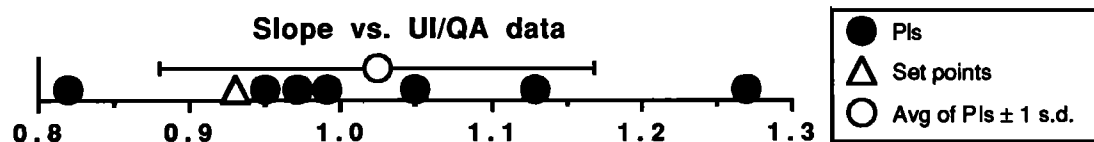


Figure 3. Graphical representation of the linear least squares best fit slopes of PI response to the response of the UI/QA system for phases 3a, 3b, and 2b. The UI set points are included for comparison. The average and standard deviation is included to give a measure of the breadth of the calibrations.

Table 5b. Summary Regression Statistics of PI Response Versus Set Points

Method	Parameter	Phase			Late, No Controls	Late, Controls	All Late Runs
		3a	3b	2b			
Aq. CL	slope	1.049	1.198	1.035	1.097	1.177	1.113
	S.E. slope	0.043	0.028	0.063	0.030	0.070	0.029
	intercept	-4.0	-2.0	9.5	1.4	16.7	5.2
	S.E. int	10.7	7.4	15.0	7.5	17.2	7.2
Modified TECO	R^2	0.9791	0.9928	0.9580	0.9711	0.9656	0.9662
	slope	1.190	1.236	1.217	1.215	1.228	1.217
	S.E. slope	0.010	0.013	0.017	0.008	0.033	0.009
	intercept	13.2	11.3	13.2	12.5	10.5	12.2
ID-GC/MS	S.E. int	2.4	3.3	4.1	2.1	8.1	2.2
	R^2	0.9991	0.9936	0.9976	0.9982	0.9928	0.9972
	slope	1.416	1.372	1.313	1.368	1.379	1.371
	S.E. slope	0.015	0.013	0.013	0.015	0.034	0.014
Mist-IC	intercept	19.0	7.6	8.0	12.5	17.6	13.6
	S.E. int	3.7	3.6	3.1	3.8	8.6	3.5
	R^2	0.9986	0.9988	0.9989	0.9956	0.9945	0.9953
	slope	1.038	1.026	0.954	1.011	1.051	1.019
DD-SCD	S.E. slope	0.013	0.023	0.021	0.015	0.026	0.014
	intercept	19.5	21.1	9.8	16.9	16.5	19.3
	S.E. int	3.2	6.1	5.1	3.8	6.4	3.5
	R^2	0.9980	0.9941	0.9941	0.9910	0.9939	0.9900
Fluor.-HPLC	slope	0.816	0.926	0.215	0.580	0.953	0.638
	S.E. slope	0.060	0.053	0.228	0.154	0.046	0.128
	intercept	105.0	-65.8	417.5	177.3	35.9	151.6
	S.E. int	15.1	14.6	54.9	39.2	11.2	32.3
Filter-IC	R^2	0.9340	0.9685	0.0680	0.2667	0.9793	0.3315
	slope	1.086	1.092	1.041	1.074	1.059	1.074
	S.E. slope	0.027	0.012	0.021	0.013	0.043	0.014
	intercept	16.3	7.4	10.9	11.8	32.9	15.4
Filter-IC	S.E. int	6.9	3.3	5.0	3.3	11.1	3.5
	R^2	0.9918	0.9985	0.9953	0.9442	0.9852	0.9915
	slope	1.066	1.104	0.892	1.029	1.078	1.039
	S.E. slope	0.029	0.043	0.027	0.023	0.049	0.025
Filter-IC	intercept	19.4	18.6	10.7	16.5	29.3	19.6
	S.E. int	7.2	11.5	6.5	7.0	12.1	6.3
	R^2	0.9905	0.9819	0.9890	0.9709	0.9794	0.9701

2a and 2b), O₃ and NO_x (phase 3a), and DMS, CO₂, CO, and CH₄ (phase 3b). Because of the operating difficulties experienced in phase 2a, the humidity test was repeated in phase 2b, and the following discussion is limited to phases 3a, 3b, and 2b. It must be noted in any discussion of phase 3 that the experiment was set up to overtest the PIs. While SO₂ levels were chosen to span a range of values commonly observed from remote marine environments to coastal areas, the interference gases were delivered at or above concentrations associated with continental air.

Except for the diffusion-denuder/SCD, which is discussed in detail below, none of the methods suffered from serious interference due to any of the added matrix gases, nor due to species present in the diluted ambient air. However, two minor interferences were noted.

First, the filter pack/IC method gave a significantly lower slope (10–15%) for measurements in the presence of water vapor than for dry air, while the intercept appeared to be unaffected. This phenomenon was observed in both the late phases (3a/3b versus 2b) and by comparing phases 1 and 2a control runs with phase 2a wet runs.

Second, phase 3a data suggest that the aqueous chemiluminescence method has a negative slope interference due to O₃/NO_x. This is most apparent when only the data for phase 3a (interference and controls separated) are plotted [Stecher *et al.*, 1995]. However, all but one of the control runs for phase 3a were performed on the same day, and other data throughout

the experiment indicate that this method suffered from day-to-day calibration changes during GASIE [Stecher *et al.*, 1995]. As stated in the Precision section, the aqueous chemiluminescence system used separate calibration curves each day in contrast to the normal operation. Thus there is no conclusive evidence that can discern a real interference from a daily calibration variation.

The diffusion-denuder/SCD exhibited interferences to several of the added species, most notably water vapor. In the absence of interferences this instrument performs similarly to the other instruments in the experiment, showing a linear response and a calibration (slope) approximately 10% below the next lowest method. Water vapor caused a sudden and severe degradation in signal stability, observed in phases 2a and 2b. This effect seemed to be reversible (control runs throughout the experiment appear self-consistent), although there are insufficient data to conclude whether there were significant long-term effects on the instrument due to water vapor. This phenomenon has been studied by the investigator and has been corrected by using a Nafion® drier on the sample inlet [Benner *et al.*, this issue].

In phase 3b the DD/SCD exhibited a negative intercept bias of 80–100 pptv versus the control data that has been attributed to scavenging of DMS by the uncoated stainless steel denuder tube but not by the carbonate-coated denuder prior to analysis. This would lower the total sulfur signal but leave the signal for (total minus SO₂) unchanged, resulting in a negative bias pro-

Table 5c. Summary Regression Statistics of PI Response Versus Other PI Responses for Late Phases

Y Axis Method	Parameter	X-Axis Method					
		Aq CL	Modified TECO	ID-GC/MS	Mist-IC	Fluor.-HPLC	Filter-IC
Aq. CL	slope		1.057	1.213	0.892	0.930	0.908
	S.E. slope		0.028	0.032	0.021	0.028	0.029
	intercept		13.8	13.3	19.2	18.1	19.8
	S.E. int		8.0	9.0	6.0	8.1	8.2
	R^2		0.9640	0.9682	0.9717	0.9544	0.9491
Modified TECO	slope	0.912		1.123	0.835	0.881	0.852
	S.E. slope	0.024		0.015	0.013	0.012	0.021
	intercept	-5.5		0.6	9.5	5.3	9.4
	S.E. int	7.6		4.7	4.0	3.9	6.5
	R^2	0.9640		0.9916	0.9879	0.9901	0.9695
ID-GC/MS	slope	0.798	0.883		0.745	0.783	0.758
	S.E. slope	0.021	0.012		0.008	0.010	0.018
	intercept	-4.2	1.4		6.6	6.1	7.9
	S.E. int	7.4	4.2		2.8	3.7	6.4
	R^2	0.9682	0.9916		0.9945	0.9919	0.9738
Mist-IC	slope	1.090	1.183	1.336		1.047	1.021
	S.E. slope	0.026	0.018	0.014		0.016	0.019
	intercept	-15.4	-8.5	-7.4		-3.1	-3.7
	S.E. int	6.9	4.8	3.9		4.4	5.0
	R^2	0.9717	0.9879	0.9945		0.9877	0.9826
Fluor.-HPLC	slope	1.027	1.124	1.267	0.944		0.964
	S.E. slope	0.031	0.016	0.017	0.015		0.023
	intercept	-9.5	-3.8	-5.6	5.4		5.1
	S.E. int	8.8	4.4	4.8	4.1		6.5
	R^2	0.9544	0.9901	0.9919	0.9877		0.9715
Filter-IC	slope	1.046	1.138	1.284	0.962	1.008	
	S.E. slope	0.033	0.028	0.030	0.018	0.024	
	intercept	-10.7	-4.0	-3.4	3.7	0.6	
	S.E. int	9.2	7.7	8.4	4.9	6.7	
	R^2	0.9491	0.9695	0.9738	0.9826	0.9715	

portional to the DMS concentration. The nominal concentration of added DMS (100 pptv) supports this explanation. Finally, the presence of O_3/NO_x caused a 40–50 pptv intercept bias without affecting the slope. The mechanism of this bias has not yet been determined.

A repeat test of this instrument with the dilution system is being scheduled under the direction of the Advisory Committee in order to assess corrections that have been made since the original experiment. For a complete description of the instrumental performance and remedies to these problems, see *Benner et al.* [this issue].

Response Under 50 pptv

Approximately half the runs in the late phases were devoted to SO_2 concentrations below 50 pptv, with approximate nominal concentrations of 0, 20, and 40 pptv. These data were examined from the perspective of discrimination among these set points, and the data are presented in Figure 4 and summarized in Table 8. As discussed in the Precision section, a caveat must be placed on the conclusions reached from Figure 4a relating to the amount of sample allowed the filter-IC and mist-IC methods. Since a substantial amount of the random error resides in the blank, the precision of the low-level data for these methods will be particularly affected. This was remedied at least for the former method by constructing a similar plot and including the nominal 127 pptv set point (Figure 4b). An attempt was made to include the DD-SCD by using just the control data, but there were very few control runs below 50 pptv; this was remedied by including the 127 pptv set point as well (Figure 4c). The caveats mentioned above are reflected in

Table 8 and are discussed in detail in later papers [*Benner et al.*, this issue; *Ferek*, this issue; *Talbot et al.*, this issue].

Results With Diluted Ambient Air

Because the concentration of SO_2 was not constant during the sampling of dilute ambient air and because the methods used widely different integration times, the run means are expected to vary from one investigator to another according to when during a run each PI sampled. Thus the individual replicate data must be compared. Time-resolved plots of the data (Plates 2a and 2b) show that six of the seven methods tracked each other quite well despite the differences in integration time and that in general, the techniques agree with one another within the differences in calibration. As shown in Plate 4, the diffusion-denuder/SCD suffered from a very large and variable positive offset during this phase, although it tracked gross changes in concentration with the other methods. The source of this phenomenon has been studied by the investigator [*Benner et al.*, this issue].

As mentioned in the Experiment section, the ambient air was diluted by known factors of 5 or 10 during this phase. In addition to diluting all the potential interferences, the metal bellows pump/Teflon® manifold system would be expected to attenuate selected ones (reactive gases by stainless steel bellows and particles by the Teflon®). While this compromised the interference testing during this phase, it was judged necessary to allow comparison with the other phases of the experiment. Since the Lewes, Delaware, site is 22.5 km NNE of a coal-fired power plant and 6.5 km from the Atlantic Ocean, changes in wind direction effected large and rapid changes in

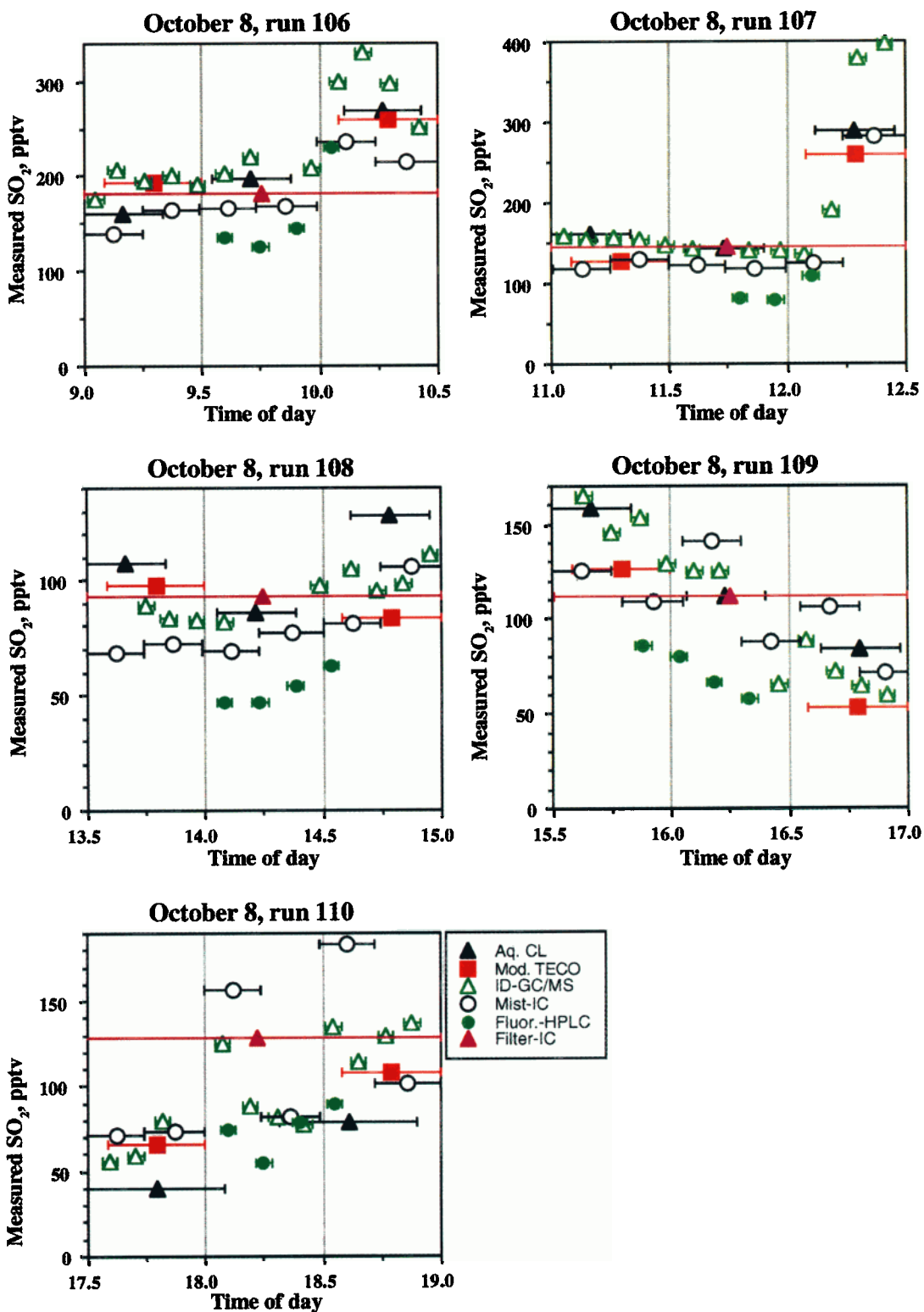


Plate 2a. Time-resolved response for all PIs except diffusion-denuder/SCD for all runs on October 8, 1994. The DD-SCD suffered from a very large and variable positive offset during this phase and was not included on these plots, although it tracked gross changes in concentration with the other methods (Plate 4). The horizontal bars represent integration times. Each run is scaled to maximize the vertical spread but starts at zero to show the relative differences among the methods.

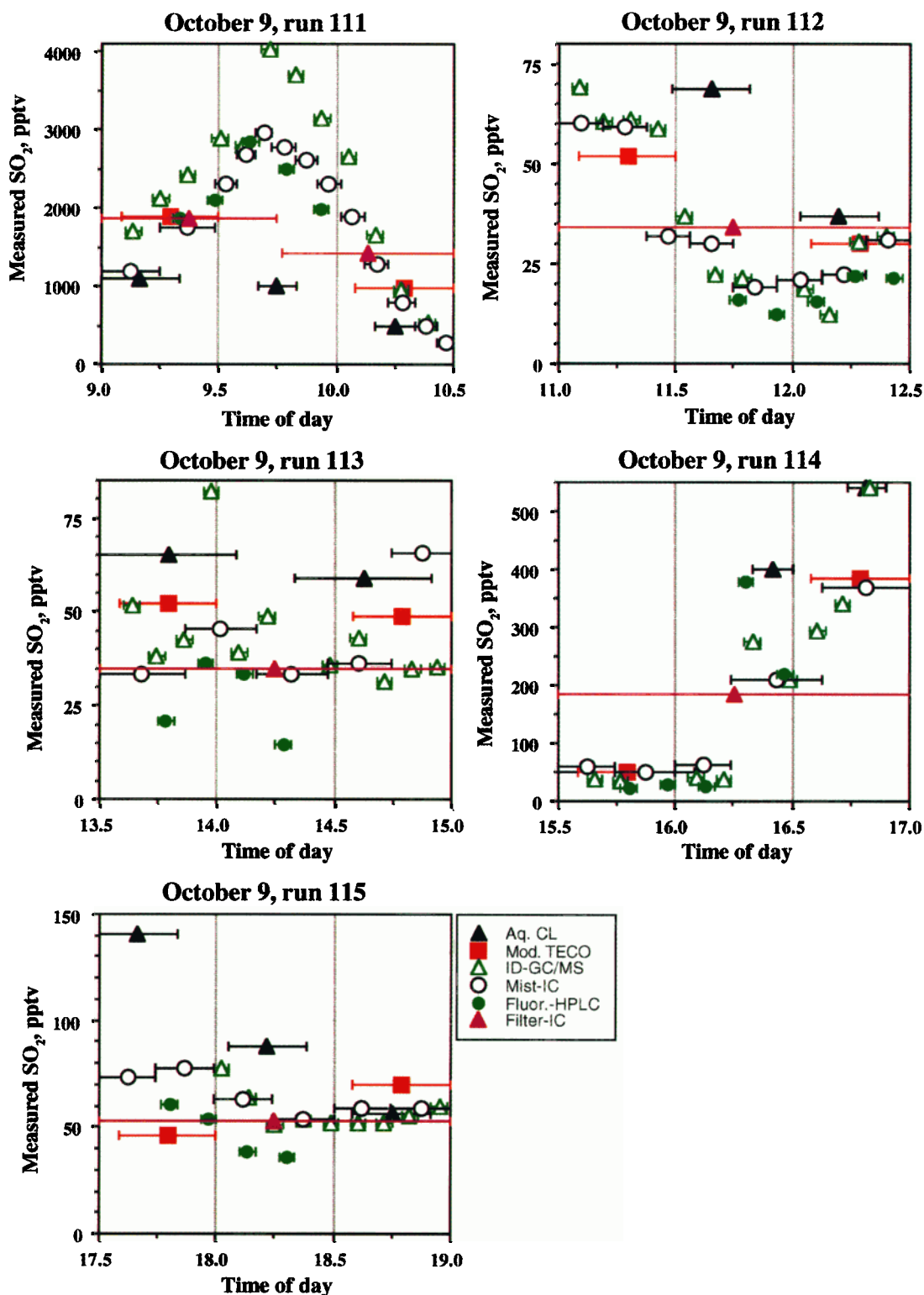


Plate 2b. Time-resolved response for all PIs except diffusion-denuder/SCD for all runs on October 9, 1994. The DD-SCD suffered from a very large and variable positive offset during this phase and was not included on these plots, although it tracked gross changes in concentration with the other methods (Plate 4). The horizontal bars represent integration times. Each run is scaled to maximize the vertical spread but starts at zero to show the relative differences among the methods. The passage of an SO_2 plume (run 111) was detected by all investigators except the aqueous chemiluminescence.

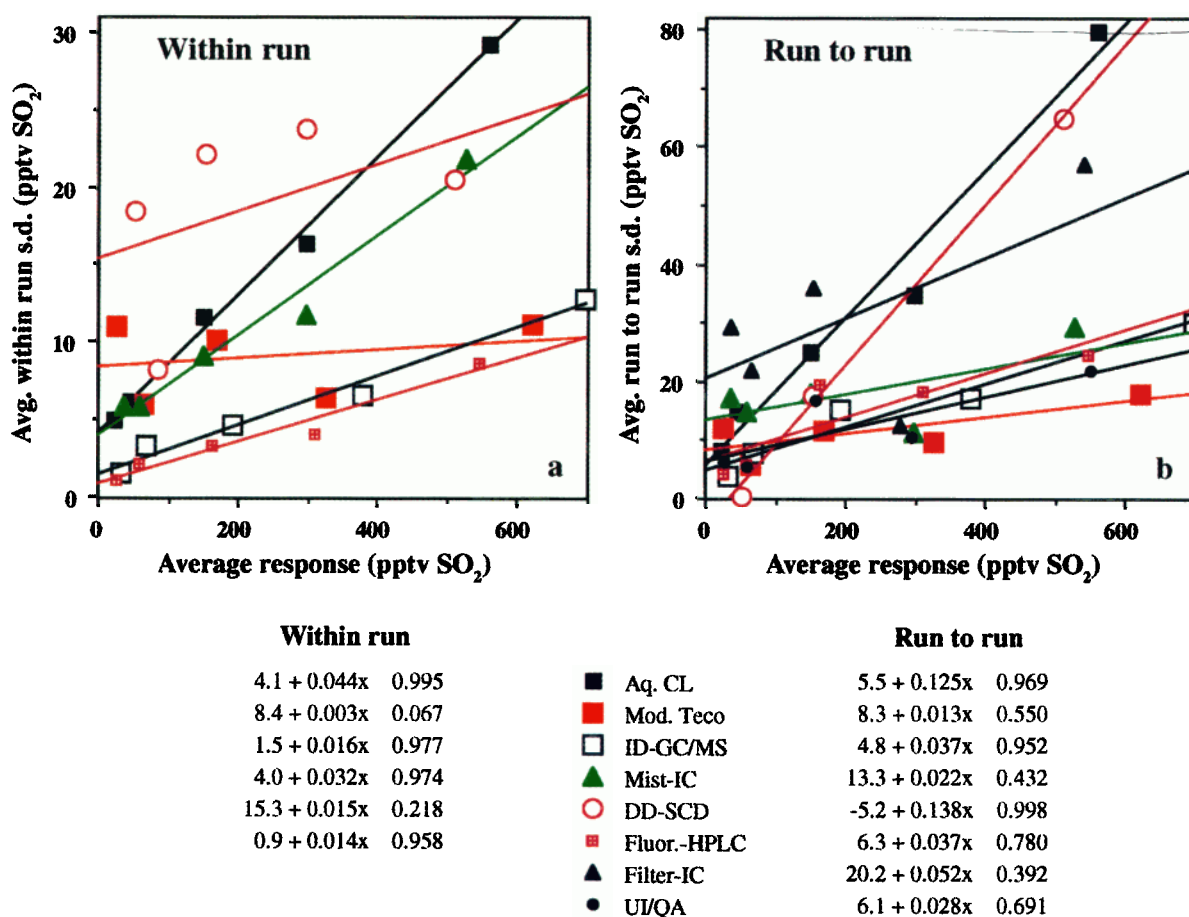


Plate 3. Mean value of the standard deviations (a) within run and (b) run to run for all runs of a given set point plotted against the mean response for those set points. The numbers in the accompanying table represent the equation of the least squares best fit line through the data followed by the R^2 value. These data were used to assess not only the variability of the measurement techniques but also the ISGDS delivery system. The filter pack/IC only gave one replicate per run, so no within-run s.d. could be calculated. Note that only the control data for the DD-SCD method are used because of the interferences noted with this method (see text). Note also that the UI/QA run-to-run results are plotted for comparison.

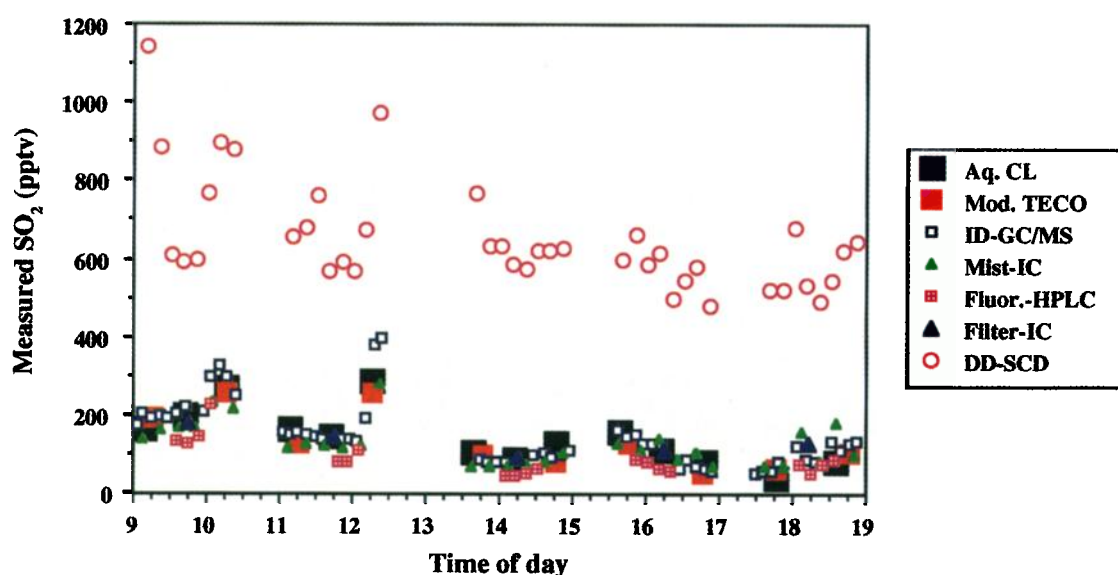


Plate 4. Replicate data for all investigators for the first day of phase 4. Note that the DD-SCD results show a substantial and variable positive bias, although there is ample evidence that large concentration changes are being followed.

Table 6. Summary of Within-Run and Run-to-Run Precision for PIs

Method	Within Run	Run to Run
A, aq. chemilum.	3.7 pptv + 4.5%	4.9 pptv + 13%
B, modified TECO	8.6 pptv + 0.2%	7.0 pptv + 1.6%
C, isotope-dil. GC/MS	1.6 pptv + 1.5%	8.0 pptv + 3.0%
D, mist chamber/IC	4.0 pptv + 3.2%	12 pptv + 2.6%
E, dif. den./SCD (control only)	15 pptv + 1.5%	-5.3 pptv + 14%
F, fluorescence HPLC	1.2 pptv + 1.3%	4.4 pptv + 4.2%
G, carbonate filter pack/IC	not determined	18 pptv + 5.8%
UI/quality assurance data	not determined	6.1 pptv + 2.8%

SO₂ mixing ratio, usually well above the target range for this experiment. Specific observations are discussed below.

Since each of the instruments integrated replicate measurements over different times, any pair-wise comparison of replicate data will be complicated by contributions to each of the values of sample that was not measured by the other instrument. Thus there is no rigorous parametric statistical method for comparing the results. In the CITE 3 data analysis, only pair-wise comparisons were used; this analysis included an extensive series of qualitative judgments aimed at finding pairs of replicate data with acceptable temporal overlap [Gregory *et al.*, 1993]. While it is tempting to call upon the vast array of numerical methods to add rigor to analysis of such a data set, the GASIE Planning Committee found this approach to be unacceptable. No matter how much effort is expended in developing criteria for, finding, and analyzing data with a specified degree of temporal overlap, the result is a numerical assessment based on a changing sample measured by one but not the other technique and represents the use of statistics for statistics sake.

Eschewing statistical analysis still allows several qualitative observations. Examination of Plates 2a and 2b indicates that except for the DD-SCD, the instruments tracked rapid changes of SO₂ over 2 orders of magnitude and appeared to agree within the calibration differences discussed above in the Accuracy section. The ability of a method to detect these changes depends on the integration time of a single replicate measurement, the precision of the replicate, and the amount of time between replicates. For some of the methods (mist-IC, filter-IC, aqueous chemiluminescence, in particular) the integration time and the precision are almost inversely proportional, and this combination can be tailored to the field situation and the research question. Because the former two methods sampled smaller volumes than normal, they were forced to compromise either precision or time resolution (or both). Thus

Table 7. Summary of Calibration Differences Among PIs and UI Dilution System

Method	Slope Versus QA
Isotope-dilution GC/MS	1.27
Modified TECO	1.13
Aqueous chemiluminescence	1.05
Fluorescence HPLC	0.99
Carbonate filter pack/IC	0.97
Mist chamber/IC	0.95
Set points (for comparison)	0.93
Diffusion denuder/SCD	0.82

PI average, 1.026; 1 s.d., 0.144.

while we can observe the general agreement of the methods in this experiment (both in the measurement and in the response to changes), we cannot assess how these methods would compare if they had chosen different integration times.

One exception to this general agreement is the apparent lower response of the fluorescence-HPLC method compared to the other phases of GASIE. This method exhibited a calibration very near the mean of the other techniques during the first three phases (see Table 7 and Figure 3 in the Accuracy section). While there is no statistical test that can be used to state unambiguously that a given fluor-HPLC replicate measurement in phase 4 was lower than those of the other investigators, inspection of Plates 2a and 2b suggest that data from this method were uniformly at or below the minimum value reported by the other PIs. This is addressed by Gallagher *et al.* [this issue].

One other noteworthy observation involves the response of the aqueous chemiluminescence method to rapid changes in mixing ratio. It appears that this technique exhibited substantial memory when the sample SO₂ rapidly decreased (see runs 112 and 113 after the decline from run 111 and the beginning of run 115 after the end of run 114 in Plate 2b). This phenomenon was recognized by the investigators and is discussed by Jaeschke *et al.* [this issue].

Despite the necessary qualifications, these results are illuminating given the recent SO₂ and DMS measurements reported by Bandy *et al.* [1996] in the equatorial Pacific atmosphere. Clear diurnal oscillations of SO₂ concentration were observed over the range of 10–80 pptv. We conclude from the GASIE data that above 50 pptv, six of the seven systems have the capability to respond to large SO₂ variations on this time-scale. For the lower SO₂ levels, measurement ability is dependent on the minimum detection limit and the integration time necessary to achieve it for each system.

General Conclusion

The general outcome of this intercomparison was that there are six viable methods available for measurement of SO₂ below 500 pptv with a seventh nearly ready. This array of methods offers a wide range of precision, time resolution, detection limit, cost, transportability, and labor intensity that must be taken into consideration for a given application. In addition, a mobile system is now available for production of diluted trace gases that in principal can be used for field standardization. It must be stressed that these results, while relevant to other sampling platforms, are only strictly applicable to ground-based measurements.

Table 8. Summary of PI Performance Below 50 pptv

Method	Comments
Aq. chemiluminescence	can distinguish 0/40 but has difficulty with 0/20, more so with 20/40
Modified TECO	can distinguish 40 from both 0 and 20, has trouble with 0/20
Isotope-dil. GC/MS	easily distinguishes 0, 20, and 40
Mist chamber/IC	should distinguish 0, 20, and 40 with higher-flow mist chamber
DD/SCD (controls)	cannot determine with just controls, appears to distinguish 20, 40, and 120
Fluorescence HPLC	easily distinguishes 0, 20, and 40
Filter pack/IC	should distinguish 0, 20, and 40 (with some overlap) with normal sample size

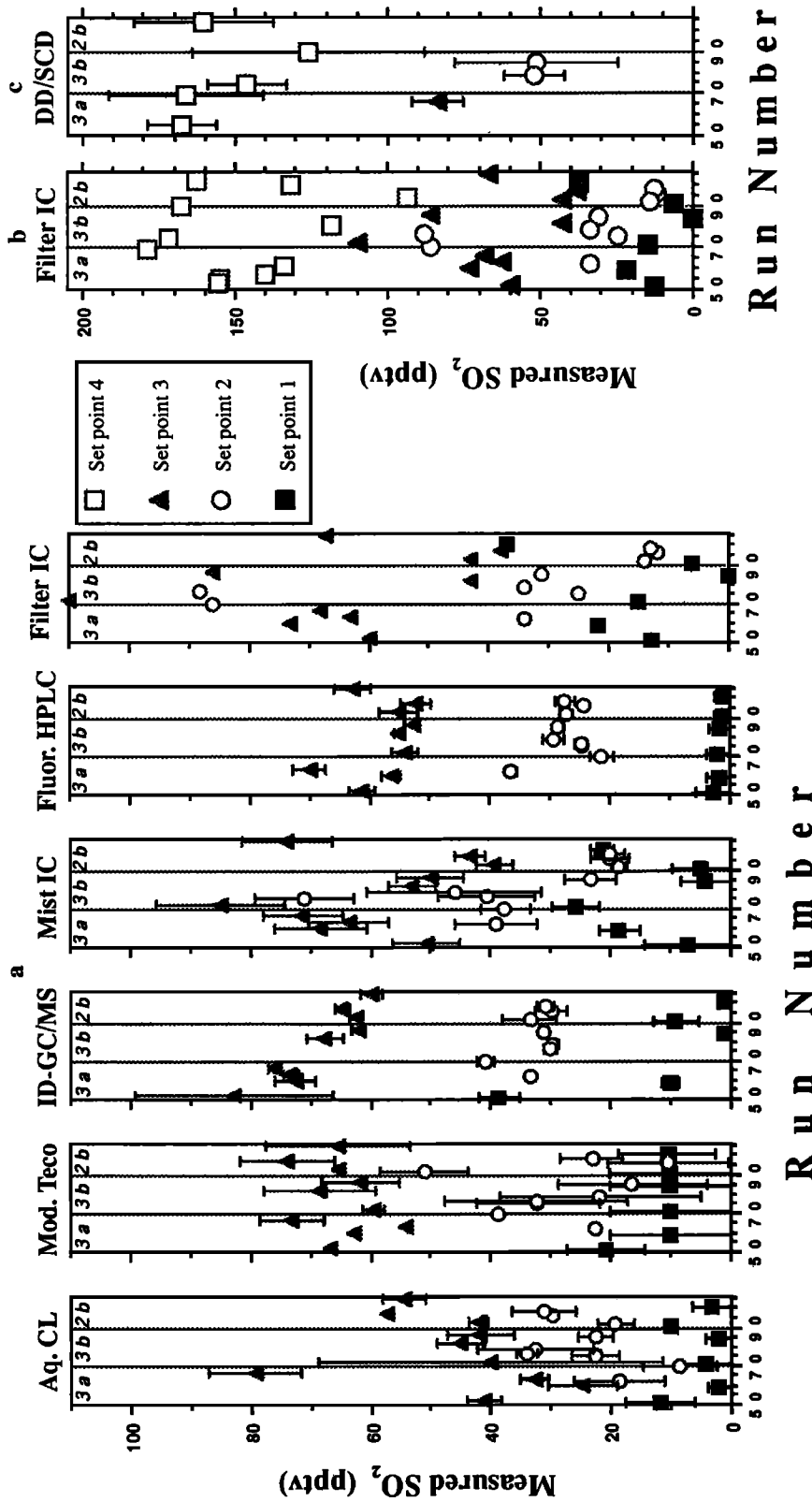


Figure 4. PI data for the three lowest set points (nominally 0, 18, and 41 pptv) in phases 3a, 3b, and 2b are shown in Figure 4a separated by run. Error bars = 1 s.d. (within run). When data were reported as “<det. lim.,” the values are given as $0.5 \times dl \pm 0.5 \times dl$ for ease of viewing but do not indicate a real distribution of values. Figure 4b shows data for the first four set points for the filter-IC method (set point 4 = 127 pptv). The 41 and 127 pptv levels represent a threefold increase in the SO_2 level over Figure 4a to compensate for the lower sample volume allowed this method during GASIE relative to its normal procedure. Data for the mist-IC in Figure 4a were compromised in similar fashion to a smaller degree (approximately twofold). Data for only the control runs of the diffusion denuder/SCD method are presented in Figure 4c. While it is not possible to form a solid conclusion regarding the low-level detection limit with such a small data set, it allows some assessment of the technique in the absence of interferences.

This program uncovered two general areas that require attention. The discrepancy among the calibrations of otherwise worthy methods is troubling because it suggests that either the commercially available "standards" used by the atmospheric community are not meeting the stated specifications or that the processing of the standards by the investigators is flawed. In addition, the complexity of the dilution system and the tasks asked of it require a much greater attention to detail in the field.

Acknowledgments. The authors acknowledge support from grant ATM-9002401 (GWL) from the Atmospheric Sciences Program of the National Science Foundation. In addition, they wish to express their sincerest gratitude to the support staff at the University of Delaware's College of Marine Studies. In particular, J. Tigue and P. Gerhart (purchasing), S. Dawson (shipping and receiving), L. Leidy (secretary), G. Sterling (carpenter), D. Lushis (HVAC), R. McGee (electrician), K. Lee (forklift artist), and D. Edwards (shop supervisor) were available on a moment's notice to help solve unforeseen problems during the 2 months of setup and running of this experiment.

References

- Bandy, A. R., D. C. Thornton, B. W. Blomquist, S. Chen, T. P. Wade, J. C. Ianni, G. M. Mitchell, and W. Nadler, Chemistry of dimethyl sulfide in the equatorial Pacific atmosphere, *Geophys. Res. Lett.*, **23**, 741–744, 1996.
- Benner, R. L., J. Wu, and D. K. Knicks Jr., Development and evaluation of the diffusion denuder-sulfur chemiluminescence detector for atmospheric SO₂ measurements, *J. Geophys. Res.*, this issue.
- Ferek, R. J., D. A. Hegg, J. A. Herring, and P. V. Hobbs, An improved filter pack technique for airborne measurement of low concentrations of SO₂, *J. Geophys. Res.*, **96**, 22,373–22,378, 1991.
- Ferek, R. J., P. A. Covert, and W. Luke, Intercomparison of measurements of sulfur dioxide in ambient air by carbonate-impregnated filters and TECO pulsed-fluorescence analyzers, *J. Geophys. Res.*, this issue.
- Gallagher, M. S., D. B. King, P.-Y. Whung, and E. S. Saltzman, Performance of the HPLC/fluorescence SO₂ detector during the GASIE instrument intercomparison experiment, *J. Geophys. Res.*, this issue.
- Gregory, G. L., D. D. Davis, N. Beltz, A. R. Bandy, R. J. Ferek, and D. C. Thornton, An intercomparison of aircraft instrumentation for tropospheric measurements of sulfur dioxide, *J. Geophys. Res.*, **98**, 23,325–23,352, 1993.
- Jaeschke, W., N. Beltz, W. Haunold, and U. Kruschke, Improvement of the Tetrachloromercurate-Absorption technique for measuring low atmospheric SO₂ mixing ratios, *J. Geophys. Res.*, this issue.
- Luke, W. T., Evaluation of a commercial pulsed-fluorescence detector for the measurement of low-level SO₂ concentrations during GASIE, *J. Geophys. Res.*, this issue.
- Luther, G. W., III, and H. A. Stecher III, Preface: Historical background, *J. Geophys. Res.*, this issue.
- MacTaggart, D. L., S. O. Farwell, T. J. Haakenson, W. L. Bamesberger, and W. D. Dorko, Generation and evaluation of test gas mixtures for the Gas-Phase Sulfur Intercomparison Experiment (GASIE), *J. Geophys. Res.*, this issue.
- Stecher, H. A., III, G. W. Luther III, D. L. MacTaggart, S. O. Farwell, P. D. Goldan, W. D. Dorko, and D. R. Crosley, Evaluation of GASIE data, 1, Coded, run-averaged data, *Natl. Sci. Found.*, in press, 1995.
- Talbot, R. W., E. M. Scheuer, B. L. Lefer, and W. T. Luke, Measurements of sulfur dioxide during GASIE with the mist chamber technique, *J. Geophys. Res.*, this issue.
- N. Beltz and U. Kruschke, Zentrum für Umweltforschung der J. W. Goethe-Universität, D-60325, Frankfurt am Main, Germany.
- R. L. Benner and J. Wu, Geophysical Institute and Department of Chemistry, University of Alaska, Fairbanks, Alaska 99701.
- D. R. Crosley, SRI International, Molecular Physics Laboratory, Menlo Park, CA 94025.
- W. D. Dorko, National Institute of Standards and Technology, B-324 Chemistry, Gaithersburg, MD 20889.
- S. O. Farwell and D. L. MacTaggart, Graduate Office, South Dakota School of Mines and Technology, Rapid City, SD 57701.
- R. J. Ferek, Atmospheric Sciences Department, University of Washington, Box 351640, Seattle, WA 98195.
- M. S. Gallagher and E. S. Saltzman, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149.
- P. D. Goldan, NOAA Aeronomy Laboratory, 325 Broadway, R/E/AL7, Boulder, CO 80303.
- B. L. Lefer, E. M. Scheuer, and R. W. Talbot, Institute for the Study of Earth, Oceans and Space, University of New Hampshire, Morse Hall, Durham, NH 03824.
- W. T. Luke, NOAA Air Resources Laboratory, SSMC3, Rm. 3439, 1315 East-West Highway, Silver Spring, MD 20910.
- G. W. Luther III, College of Marine Studies, 700 Pilottown Road, University of Delaware, Lewes, DE 19958-1298.
- H. A. Stecher III, 5236 40th Avenue, SW, Seattle, WA 98136.
- D. C. Thornton, Department of Chemistry, Drexel University, 32nd and Chestnut Streets, Philadelphia, PA 19104.

(Received July 10, 1996; revised January 10, 1997; accepted May 5, 1997.)