United States
Environmental Protection
Agency

Office of Research and
Development
Washington, DC 20460

EPA 600 3-90 051
June 1990

EPA

The Across North America Tracer Experiment (ANATEX) Model Evaluation Study

EPA/600/3-90/051
June 1990

THE ACROSS NORTH AMERICA TRACER EXPERIMENT (ANATEX) MODEL EVALUATION STUDY

by

Terry L. Clark*
Atmospheric Sciences Modeling Division
National Oceanic and Atmospheric Administration
U.S. Department of Commerce
Research Triangle Park, NC 27711

and

Richard D. Cohn
Analytical Sciences, Incorporated
100 Capitola Drive, Suite 106
Durham, NC 27713

U.S. Environ
Dearborn S
Chicago, IL 6060

*On assignment to the Atmospheric Research and Exposure Assessment Laboratory,
U.S. Environmental Protection Agency

ATMOSPHERIC RESEARCH AND EXPOSURE ASSESSMENT LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NC 27711

NOTICE

The information in this document has been funded wholly or in part by the U.S. Environmental Protection Agency, the U.S. Air Force, and the National Oceanic and Atmospheric Administration of the U.S. Department of Commerce under U.S. Department of Commerce Contract No. 50-EANR-8-00025 to Analytical Sciences, Incorporated. It has been subjected to the EPA peer and administrative review, and it has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

FOREWORD

The U.S. Environmental Protection Agency, the National Oceanic and Atmospheric Administration, and the U.S. Air Force have completed an evaluation of 11 operational models to assess the performances of state-of-the-science, long-range transport and diffusion models. The model calculations were compared to observations of surface concentration data compiled during the Across North America Tracer Experiment (ANATEX).

Before the distribution of the ANATEX data, modelers applied their models in a "blind" applications mode using required meteorological input data and the actual periodic 3-hour (h) ANATEX tracer emission rates according to the prescribed schedule during the first 3 months of 1987. Some of these models were very similar to others in this study; the only differences were variations of the modeling assumptions or the selection of modeling options.

Model performance measures were developed on the basis of the features of the surface sampling network and the sampling protocol. These measures were quantified using either ensemble concentration means or relative distances of the centroids of tracer "footprints"--composite tracer plumes defined by the 24-h-mean measurements.

Several aspects of the evaluation study are discussed in this report. First, the performances of the three genres of models--(1) single-layer Lagrangian, (2) multiple-layer Lagrangian, and (3) multiple-layer Eulerian--are compared to each other to relate model performance to model approach. Secondly, the performances of the various model versions are related to the differences in the modeling codes to relate model performance to model assumptions/options. Thirdly, model performance is related to three meteorological scenarios to relate model performance to the degree of complexity of the air flow.

ABSTRACT

During the first three months of 1987, three perfluorocarbon tracer gases were released at 2.5-day or 5.0-day intervals from two sites in central North America (Glasgow, Montana and St. Cloud, Minnesota) and sampled for 24-h periods at 77 surface sites. The source-receptor distances ranged from less than 30 km to 3,000 km. These Across North America Tracer Experiment (ANATEX) data serve as a unique evaluation data set with which to evaluate the long-range transport and diffusion simulations of acid deposition models and to establish a range of uncertainty for various model genres.

The performances of three single-layer Lagrangian, six multiple-layer Lagrangian, and two multiple-layer Eulerian models are assessed using quantifiable measures based on comparisons of ensemble mean concentrations and plume widths as well as trajectory errors expressed as a function of transport time. In general, the multiple-layer Lagrangian models performed best in simulating the transport of the tracers, while the Eulerian models performed best in simulating the ensemble concentration frequency distributions.

After 0.5 day of transport, trajectory errors ranged from 100 km to 400 km; after 2.5 days, the errors ranged from 300 km to 800 km. Beyond 2.5 days, errors from four Lagrangian models plateaued, while errors for the other models continued to increase, peaking at nearly 1,100 km after 3.5 days. Of the three single-layer models, one performed as well as any other model in simulating the transport; the other two single-layer models performed worst. Thus, a slight difference in the method of calculating single-layer transport vectors can yield a significant difference in model performance. For six of the eleven models, the greatest errors in transport speed and location tended to occur when the tracer was intercepted by cyclones and/or fronts.

The ensemble mean concentrations along three bands of sites 1,000 to 2,300 km downwind from the release sites were nearly always overpredicted by the single-layer Lagrangian models, in some cases by as much as a factor of 5; overpredictions by a factor of 3 were common. The two Eulerian models tended to underpredict these ensemble means by about 40%, especially along the nearest band. Finally, the multiple-layer Lagrangian models tended to underpredict the ensemble means for the tracer released from Montana, yet tended to overpredict the ensemble means for the tracer released from Minnesota.

Horizontal spreading of the tracer plumes beyond 1,000 km transport distance occurred at a rate of 30% to 60% per 600 km. Generally the models replicated this spreading rate, although there was a wide range of plume widths. The plume widths of the Eulerian models tended to be the greatest, which partially explained their tendency to underpredict mean concentrations. In virtually all cases, the plume widths of the Lagrangian models were equivalent to or less than the actual plume widths; for some models, plume widths were much less, by factors of 2 and 3.

CONTENTS

Foreword iii
Abstract iv
Figures vii
Tables xiii
Abbreviations and Symbols xiv
Acknowledgments xvii

1. Introduction 1

2. Previous Model Evaluation Studies 8

3. Model Evaluation Protocol 10

3.1 The Evaluation Approach 10
3.2 Model Performance Measures 16

3.2.1 Ensemble Concentration Means and Distributions 16
3.2.2 Ensemble Mean Horizontal Diffusion 18
3.2.3 Transport Speed of Discrete Tracer Footprints 19
3.2.4 Tracer Footprint Location Error 22

4. Assessment of the Model Performances Based on Ensemble Data 23

4.1 Comparison of Concentration Magnitudes 23

4.1.1 Ensemble Concentration Distributions 23
4.1.2 Mean Concentrations Along Bands 29

4.2 Comparison of Horizontal Diffusion of Tracer Puffs 35

5. Assessment of the Model Performances Based on Comparisons of Individual Tracer Releases 43

5.1 Performance as Related to Meteorological Scenario 43
5.2 Mean Separation Error 49
5.3 Footprint Transport Speeds and Centroid Locations 52

5.3.1 Summary of Model Performances 54
5.3.2 Single-Layer Lagrangian Models 58
5.3.3 Multiple-Layer Lagrangian Models 66
5.3.4 Multiple-Layer Eulerian Models 82

CONTENTS (CONCLUDED)

6. Summary and Conclusions 86

6.1 Single-Layer Lagrangian (SLL) Models 88

6.1.1 Summary 88
6.1.2 Conclusions 89

6.2 Multiple-Layer Lagrangian (MLE) Models 90

6.2.1 Summary 90
6.2.2 Conclusions 92

6.3 Multiple-Layer Eulerian (MLE) Models 93

6.3.1 Summary 93
6.3.2 Conclusions 94

7. Recommendations 95

8. References 96

Appendix A. Model Descriptions A-1

Appendix B. Time Series Plots of 24-h Mean Concentrations of PTCH and PDCH Tracers Along Bands of Sites B-1

FIGURES

Number		Page
1	The 77-site ANATEX surface network for 24-h-average tracer concentration measurements and the location of the tracer release sites at Glasgow, Montana (GGW) and St. Cloud, Minnesota (STC)	4
2	Illustration of the 6 January 1987 "footprint", or a 24-h composite of instantaneous tracer puffs near the surface, and the centroid C_m, for the measured PDCH concentrations from release 1	12
3	The 24-h locations of ARL (---) and actual (——) footprint centroids for the first release of PTCH tracer	14
4	The bands of sites along which ensemble-mean calculated and measured horizontal diffusion were estimated for (a) PTCH and (b) PDCH tracers	17
5	The model performance measures characterizing (a) the agreement between calculated and measured transport speed and (b) footprint location errors	20
6	Box plots of cumulative frequency distribution of uncertainty (U) versus excess concentration (χ) for PMCH, PDCH, and PTCH	21
7	Box plots comparing quartiles and means of calculated and measured PTCH tracer concentrations at the 77 surface sites during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987	25
8	Frequency distributions of calculated and measured PTCH tracer concentrations at the 77 surface sites during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987	26
9	Box plots comparing quartiles and means of calculated and measured PDCH tracer concentrations at the 77 surface sites during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987	28
10	Frequency distributions of calculated and measured PDCH tracer concentrations at the 77 surface sites during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987	30

FIGURES (CONTINUED)

Number		Page
11	Mean calculated and measured PTCH concentrations along three bands of sites downwind of GGW during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987	32
12	Mean calculated and measured PDCH concentrations along three bands of sites downwind of STC during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987	36
13	Mean horizontal diffusion of calculated and measured tracer puffs as indicated by the average distance of nonzero PTCH concentrations along three bands of sites downwind of GGW during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987	38
14	Mean horizontal diffusion of calculated and measured tracer puffs as indicated by the average distance of nonzero PDCH concentrations along three bands of sites downwind of STC during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987	41
15	Comparison of the mean separation errors of PTCH and PDCH footprint centroids, (D_e) as a function of transport time	51
16	The comparison of trajectory error ranges determined from ANATEX and CAPTEX model evaluations	53
17	Percentages of footprint-days when the models deviate by 20% and 50% from the actual (a) transport speed, (b) centroid location within 24-h, (c) between 24 h and 48 h, and (d) after 48 h of transport	56
18	Errors in SRL transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH for the designated tracer releases	59
19	Errors in TCAL transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints for the designated tracer releases	61

FIGURES (CONTINUED)

Number		Page
20	Errors in VCAL transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints for the designated tracer releases	63
21	Comparison of TCAL and VCAL (V) errors in transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints common to both sets for the designated tracer releases	65
22	Errors in ARL transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints for the designated tracer releases	67
23	Errors in BAT transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints for the designated tracer releases	69
24	Errors in GAMUT transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints for the designated tracer releases	71
25	Errors in HY-SPLIT transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints for the designated tracer releases	73
26	Comparison of ARL and HY-SPLIT (H) errors in transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints common to both sets for the designated tracer releases	75
27	Errors in MLAM-FINE transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints for the designated tracer releases	77
28	Errors in MLAM-COARSE transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints for the designated tracer releases	79
29	Comparison of MLAM-COARSE and MLAM-FINE (F) errors in transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints common to both sets for the designated tracer releases	81

FIGURES (CONTINUED)

Number		Page
30	Errors in ADOM transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints for the designated tracer releases	83
31	Errors in ADPIC transport speeds for (a) PDCH and (b) PTCH footprints; and centroid locations for (c) PDCH and (d) PTCH footprints for the designated tracer releases	85
B-1	Time series plots of 24-h mean-measured and calculated PTCH concentrations along the 1,000-km band for January 1987	B-2
B-2	Time series plots of 24-h mean-measured and calculated PTCH concentrations along the 1,000-km band for February 1987	B-3
B-3	Time series plots of 24-h mean-measured and calculated PTCH concentrations along the 1,000-km band for March 1987	B-4
B-4	Time series plots of 24-h mean-measured and calculated PTCH concentrations along the 1,600-km band for January 1987	B-5
B-5	Time series plots of 24-h mean-measured and calculated PTCH concentrations along the 1,600-km band for February 1987	B-6
B-6	Time series plots of 24-h mean-measured and calculated PTCH concentrations along the 1,600-km band for March 1987	B-7
B-7	Time series plots of 24-h mean-measured and calculated PTCH concentrations along the 2,300-km band for January 1987	B-8
B-8	Time series plots of 24-h mean-measured and calculated PTCH concentrations along the 2,300-km band for February 1987	B-9
B-9	Time series plots of 24-h mean-measured and calculated PTCH concentrations along the 2,300-km band for March 1987	B-10
B-10	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 300-km band for January 1987	B-11
B-11	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 300-km band for February 1987	B-12

FIGURES (CONTINUED)

Number		Page
B-12	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 300-km band for March 1987	B-13
B-13	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 700W-km band for January 1987	B-14
B-14	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 700W-km band for February 1987	B-15
B-15	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 700W-km band for March 1987	B-16
B-16	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 700E-km band for January 1987	B-17
B-17	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 700E-km band for February 1987	B-18
B-18	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 700E-km band for March 1987	B-19
B-19	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,000W-km band for January 1987	B-20
B-20	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,000W-km band for February 1987	B-21
B-21	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,000W-km band for March 1987	B-22
B-22	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,000E-km band for January 1987	B-23
B-23	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,000E-km band for Febuary 1987	B-24
B-24	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,000E-km band for March 1987	B-25
B-25	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,400-km band for January 1987	B-26

FIGURES (CONCLUDED)

Number		Page
B-26	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,400-km band for February 1987	B-27
B-27	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,400-km band for March 1987	B-28
B-28	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,800-km band for January 1987	B-29
B-29	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,800-km band for February 1987	B-30
B-30	Time series plots of 24-h mean-measured and calculated PDCH concentrations along the 1,800-km band for March 1987	B-31

TABLES

Number		Page
1	The ANATEX schedule and amounts of GGW releases of PTCH and STC releases of PDCH and PMCH tracers	3
2	Number of "site-days" in the ANATEX database for each tracer	5
3	The 11 long-range transport and diffusion models evaluated in the ANATEX Model Evaluation Study (AMES)	6
4	Model performance measures	15
5	The mean calculated and measured tracer concentrations along bands of samplers for the two halves of the ANATEX period	33
6	The mean distance of nonzero calculated and measured concentrations and mean spread rates along bands of samplers	39
7	Number of calculated and measured footprint pairs and the type of meteorological scenario for each tracer release	44
8	The three types of meteorological scenarios for which model performance was assessed	46
9	Number of footprint-days intercepted (I) and not intercepted (NI) by cyclones and/or fronts after 10-h of transport	46
10	Relationship between model performance and meteorological scenario	48
11	Percent of Model A footprint-days common to both Models A and B	50
12	Number of calculated and measured footprint pairs for each model as a function of transport time	50
13	A summary of model performances in calculating transport speeds and footprint centroid locations	55
14	Overall biases in the model results as indicated by the model performance measures	87
A-1	SRL attributes	A-2
A-2	TCAL and VCAL attributes	A-3
A-3	ARL attributes	A-4
A-4	BAT attributes	A-5
A-5	GAMUT attributes	A-6
A-6	HY-SPLIT attributes	A-7
A-7	MLAM (COARSE and FINE) attributes	A-8
A-8	ADOM attributes	A-9
A-9	ADPIC attributes	A-10
B-1	List of 24-h-mean concentrations exceeding 50 dfL/L	B-32

ABBREVIATIONS AND SYMBOLS

ABBREVIATIONS

ADOM	Acid Deposition and Oxidant Model
ADPIC	Atmospheric Diffusion Particle-In-Cell model
AFGWC	Air Force Global Weather Center
AGL	above ground level
AMES	ANATEX Model Evaluation Study
ANATEX	Across North America Tracer Experiment
ARL	Air Resources Laboratory model
BAT	Branching Atmospheric Trajectory model
CAPTEX	Cross-Appalachian Tracer Experiment
dfL/L	decifemtoliters per liter (10^{-16} L/L)
GAMUT	Global Atmospheric Multilayer Transport model
GGW	Glasgow, Montana
GMT	Greenwich Mean Time
h	hour
HIRAS	High Resolution Analysis System
HY-SPLIT	HYBRID Single-Particle Lagrangian Integrated Trajectories model
LST	local standard time
MESOPUFF	Mesoscale Puff model
METREX	Metropolitan Tracer Experiment
MLAM	Multilayer Air Mass Model
MLAM-COARSE	a coarse-resolution version of MLAM
MLAM-FINE	a fine-resolution version of MLAM
MLE	Multiple-Layer Eulerian model type
MLL	Multiple-Layer Lagrangian model type
NGM	NOAA's Nested Grid Model
NOAA	National Oceanic and Atmospheric Administration
PDCH	perfluoro-ortho-dimethylcyclohexane
PMCH	perfluoro-methylcyclohexane
PTCH	perfluoro-trimethylcyclohexane
rms	root-mean-square
SLL	Single-Layer Lagrangian model type
SRL	Savannah River Laboratory Adjusted Geostrophic Model
STC	St. Cloud, Minnesota
TCAL	Trajectory Calculation model
VCAL	Variable Layer Trajectory Calculation model

SYMBOLS

C_c	centroid of the calculated footprint
C_m	centroid of the measured footprint
D_c	cumulative distance between C_c and the tracer source
D_c/D_m	transport speed ratio
D_e	distance between C_c and C_m
D_e/D_m	normalized separation distance between C_c and C_m
D_m	cumulative distance between C_m and the tracer source
d_i	distance from site i to trajectory starting point
K_h	horizontal turbulent eddy diffusivity coefficient
K_z	vertical turbulent eddy diffusivity coefficient
L	Monin-Obukhov length
N	number of footprint days
p	atmospheric pressure
R_i	gradient Richardson number
T	temperature
T_v	virtual temperature
t	time
Δt	time step
U_A	advection velocity
U_D	diffusivity velocity
u	eastward wind speed component
u_*	friction velocity
v	northward wind speed component
w	vertical wind speed component
w_i	vertically-averaged wind at site i or weight for site i used in calculating foortprint centroids
z	height
Δz	thickness of model layer
z_o	surface roughness height
∇	gradient operator ($\partial/\partial x + \partial/\partial y + \partial/\partial z$)
θ	potential temperature of atmospheric layer
θ_B	potential temperature of bottom layer
θ_T	potential temperature of top layer

SYMBOLS

θ_w	wet bulb potential temperature
θ_*	scaled potential temperature
$\Delta\theta$	potential temperature difference across a layer
σ_h	horizontal dispersion parameter
σ_y	crosswind distance from plume centerline
σ_z	vertical dispersion parameter
σ_θ	standard deviation of wind direction
φ_m	nondimensional wind shear
χ	concentration magnitude

ACKNOWLEDGMENTS

The authors acknowledge the valuable assistance received during this project from Roland R. Draxler and Jerome L. Heffter of the National Oceanic and Atmospheric Administration, U.S. Department of Commerce. Also acknowledged are the other modelers who participated in this model evaluation study, without whom the study would not have been possible:

William E. Davis and Anthony R. Olsen, Pacific Northwest Laboratory
Stephen E. Masters, ENSCO, Inc.
Marvin P. Olson and Keith J. Puckett, Environment Canada
Malcolm Pendergast, Savannah River National Laboratory
Daniel J. Rodriguez, Lawrence Livermore National Laboratory

The technical assistance of Steven K. Seilkop, Analytical Sciences, Inc.; the report-preparation assistance of Christine Maxwell, Computer Sciences Corporation; and the financial support of the U.S. Air Force Technical Applications Center were also greatly appreciated.

SECTION 1

INTRODUCTION

Computer models have been developed to simulate the fate of gases and particles in the atmosphere over spatial scales ranging from the urban scale to the global scale. These models are applied to study such phenomena as acidic deposition, Arctic haze, photochemical oxidants, accidental releases of radioactive material, etc. Each of these models is comprised of a series of modules, each simulating a physical process (e.g., vertical diffusion, horizontal diffusion, transport, and for some, wet and dry deposition) relating to the atmospheric concentrations of gases and particles. In addition, some of these models simulate chemical transformations while still others simulate radioactive decay.

It is essential to assess the limitations and reliability of these models, as well as the uncertainty limits of their results, in both an operational mode (i.e., performance assessments of the final model results), as well as a diagnostic model (i.e., performance assessments of the results of each module of the model). However, the scarcity of adequate measurements (due to the enormous costs of obtaining them) prohibits a comprehensive diagnostic model evaluation and limits the simpler operational evaluation.

The concentration data from two field studies--the 1983 Cross-Appalachian Tracer Experiment (CAPTEX) [Ferber et al., 1986] and the 1987 Across North America Tracer Experiment (ANATEX) [Draxler and Heffter, 1989]--provide virtually the sole means to evaluate the transport and diffusion modules of long-range models. Additional tracer data are required to minimize the uncertainties of the evaluation results. The other modules cannot be evaluated until appropriate measurements are available.

The CAPTEX study measured 3-h and 6-h mean concentrations of an inert perfluorocarbon tracer gas released on seven occasions during September and October from either Dayton, Ohio or Sudbury, Ontario. The sampling network consisted of 86 sites located within 300 km to 1,100 km of Dayton. The results of model evaluations using this data set are summarized in Section 2.

The more recent Across North America Tracer Experiment (ANATEX) provided additional tracer concentration data for many more days and over a much larger region (Draxler and Heffter, 1989). During the 12-week period between 5 January and 29 March 1987, three different inert perfluorocarbon tracers

were released periodically during thirty-three 3-hour (h) periods from near-surface sources near Glasgow, Montana (GGW) and St. Cloud, Minnesota (STC). Two of these tracers, PTCH [1] and PDCH [2], were uniformly released for 3 h every 2.5 days from GGW and STC, respectively (Table 1). The third tracer--PMCH [3]--was released from STC simultaneously with every other PDCH release. Figure 1 shows the 77-site surface network locations where 24-h-average concentrations of each tracer were measured for the period 1400-1400 GMT. The sites are oriented along eight bands, each of which is to the east of Glasgow by the approximate designated distances (km). Table 2 indicates the number of "site-days" when concentrations were available and passed the screening criteria. [4]

The ANATEX Model Evaluation Study (AMES) relies on these 24-h surface concentrations to assess the performance of the 11 long-range transport and diffusion models listed in Table 3 and described in Appendix A. Because of resource limitations, MLAM-FINE and ADOM were applied only in the first-half period (i.e., 5 January - 16 February 1987). All but two models use routinely available wind data; HY-SPLIT and ADOM use higher resolution data generated by diagnostic meteorological models. The primary objective of each model is to simulate the fate of atmospheric pollutants emitted from a single source in the form of either a single puff or a continuous plume. In the AMES, these models were applied in a prognostic mode by the respective modelers independent of each other and without benefit of the evaluation data set--that is, the models were also applied in a "blind-test" mode. Trajectories and concentrations computed by several of these models, as well as others, have already been compared to the CAPTEX data to assess model performance. However, with the ANATEX measurements, much more can be learned about the performance of these models, since:

(1) the ANATEX sampling domain was much larger than the CAPTEX network (enabling us to assess model performance over longer spatial and temporal scales);

(2) the ANATEX tracers were periodically released, as opposed to the CAPTEX tracer that was released only for two simple meteorological scenarios (enabling us to evaluate the models over a much larger range of meteorological scenarios); and

(3) there were 66 discrete tracer releases in ANATEX, as opposed to only 7 CAPTEX tracer releases (enabling us, for the first time, to base the model performance assessment on a statistically meaningful data set).

---

[1] Perfluoro-trimethylcyclohexane.

[2] Perfluoro-ortho-dimethylcyclohexane (oPDCH), hereafter referred to as "PDCH".

[3] Perfluoro-methylcyclohexane.

[4] Concentrations were excluded when the QA/QC flags indicated that the sample was suspect (Draxler and Heffter, 1989).

TABLE 1. THE ANATEX SCHEDULE AND AMOUNTS OF GGW RELEASES OF PTCH AND STC RELEASES OF PDCH AND PMCH TRACERS

Release number[a]	Date[b]	Time period, GMT	Release amount, kg PTCH Glasgow	PDCH St. Cloud	PMCH St. Cloud
1	Jan 5	1700-2000	76.6	50.7	49.5
2	8	0500-0800	88.8	48.8	0.0
3	10	1700-2000	83.5	50.8	49.7
4	13	0500-0800	83.9	50.6	0.0
5	15	1700-2000	83.4	51.1	50.2
6	18	0500-0800	83.9	51.5	0.0
7	20	1700-2000	83.5	50.0	48.6
8	23	0500-0800	83.7	51.1	0.0
9	25	1700-2000	83.6	50.5	49.3
10	28	0500-2000	83.6	51.0	0.0
11	30	1700-2000	83.7	50.8	49.6
12	Feb 2	0500-0800	83.6	51.0	0.0
13	4	1700-2000	83.7	50.7	49.6
14	7	0500-0800	83.7	47.2	2.9
15	9	1700-2000	83.7	46.8	52.4
16	12	0500-0800	83.7	47.0	2.9
17	14	1700-2000	83.7	46.8	52.5
18	17	0500-0800	83.7	47.0	2.9
19	19	1700-2000	84.1	47.1	53.0
20	22	0500-0800	83.7	47.4	2.9
21	24	1700-2000	83.7	47.1	53.0
22	27	0500-0800	83.7	47.6	3.0
23	Mar 1	1700-2000	83.7	46.9	52.6
24	4	0500-0800	83.7	47.2	2.9
25	6	1700-2000	83.7	47.1	53.0
26	9	0500-0800	83.6	51.3	0.0
27	11	1700-2000	83.8	47.1	53.0
28	14	0500-0800	83.6	51.4	0.0
29	16	1700-2000	83.6	47.0	52.7
30	19	0500-0800	83.7	49.9	0.1
31	21	1700-2000	83.7	46.8	52.4
32	24	0500-0800	83.6	47.2	2.9
33	26	1700-2000	75.1	44.6	49.0
Average			83.4	48.7	51.2[c]

a. Odd-numbered releases are daytime releases, while even-numbered releases are nighttime releases.

b. Year = 1987.

c. Not including releases of 2.9 kg.

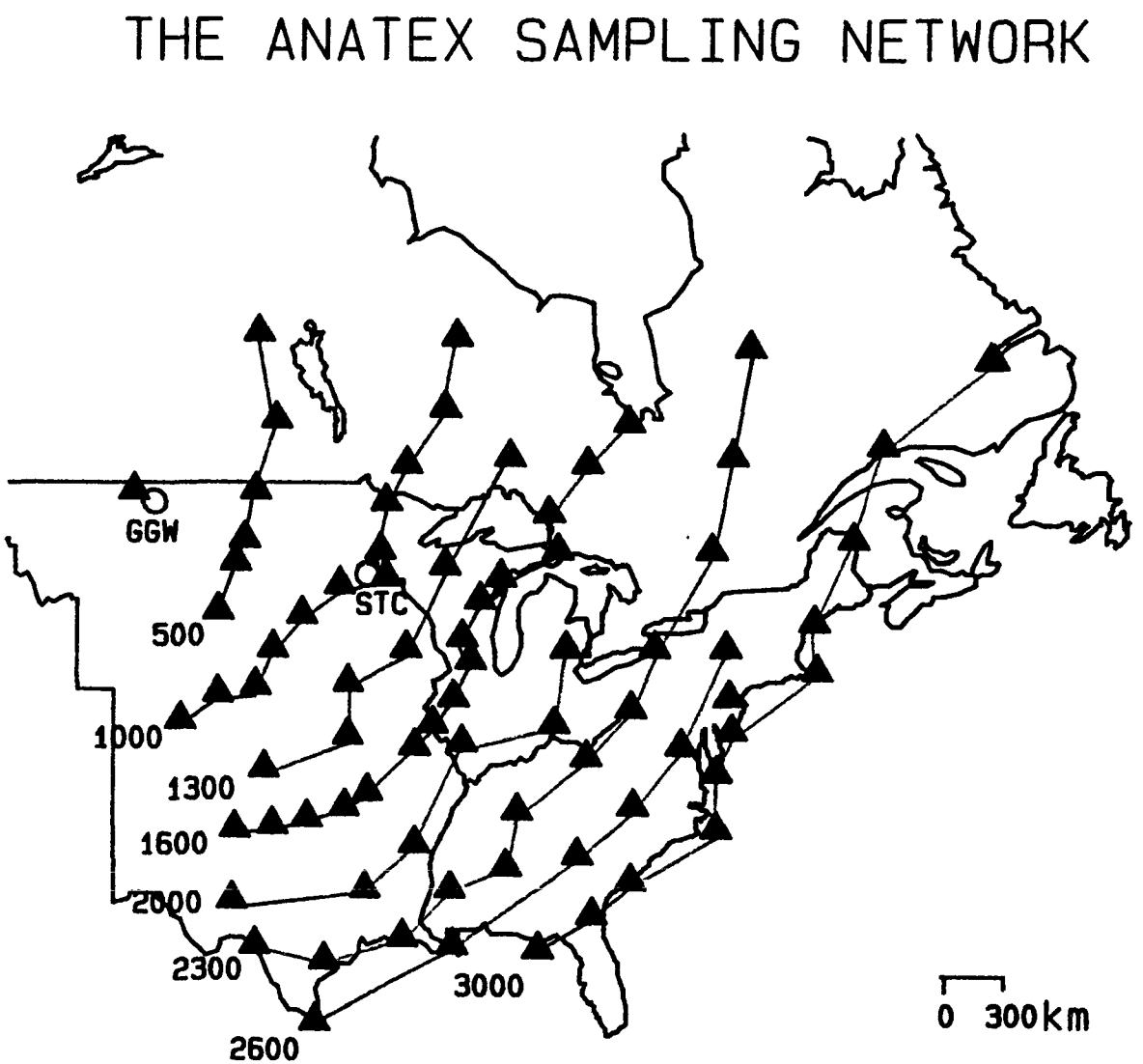


Figure 1. The 77-site ANATEX surface network for 24-h-average tracer concentration measurements and the location of the tracer release sites at Glasgow, Montana (GGW) and St. Cloud, Minnesota (STC)

TABLE 2. NUMBER OF "SITE-DAYS" IN THE ANATEX DATABASE FOR EACH TRACER

	ANATEX tracer		
Site-days	PMCH	PDCH	PTCH
Total possible	6,468	6,468	6,468
Missing sample	-892	-944	-830
Failing screening criteria	-558	-348	-342
Total in database	5,018	5,176	5,296

TABLE 3. THE 11 LONG-RANGE TRANSPORT AND DIFFUSION MODELS EVALUATED IN THE ANATEX MODEL EVALUATION STUDY (AMES)

Model acronym	Model name	Model developer
	SINGLE-LAYER LAGRANGIAN (SLL) MODELS	
SRL	Adjusted Geostrophic Model	Savannah River National Laboratory
TCAL	Trajectory Calculation Model	ENSCO, Inc.
VCAL[a]	Variable Layer Trajectory Calculation Model	ENSCO, Inc.
	MULTIPLE-LAYER LAGRANGIAN (MLL) MODELS	
ARL	Air Resources Laboratory Model	National Oceanic and Atmospheric Administration
BAT	Branching Atmospheric Trajectory Model	National Oceanic and Atmospheric Administration
GAMUT	Global Atmospheric Multilayer Transport Model	ENSCO, Inc.
HY-SPLIT[b]	Hybrid Single-Particle Lagrangian Integrated Trajectories Model	National Oceanic and Atmospheric Administration
MLAM-FINE and MLAM-COARSE	Multilayer Air Mass Model	Pacific Northwest Laboratory
	MULTIPLE-LAYER EULERIAN (MLE) MODELS	
ADOM	Acid Deposition and Oxidant Model	Environment Canada--Atmospheric Environment Service
ADPIC	Atmospheric Diffusion Particle-In-Cell Model	Lawrence Livermore National Laboratory

a. A model very similar to the TCAL model.

b. A model very similar to the ARL model.

As indicated by the objectives listed below, this study focuses on the performance of the models during the first three months of 1987. Explanations of model behavior are limited in this study in that additional analyses are required to substantiate the observations. The modelers are encouraged to conduct these analyses using the results of this study as a starting point.

The objectives of AMES are fourfold:

(1) to assess the overall performance, as well as the model errors on temporal scales of 24 h, of prognostic long-range atmospheric models with respect to transport and diffusion as a function of transport time and distance,

(2) to intercompare the model performances and relate performance to fundamental modeling approaches,

(3) to identify the periods and associated meteorological conditions when each model performed best and worst, and

(4) to compare and contrast the AMES conclusions with those of similar studies using CAPTEX data.

Section 2 summarizes the conclusions of previous model evaluation studies. Section 3 discusses the model performance measures and analyses that were developed specifically for the ANATEX database. The model evaluation results based on ensemble statistics are presented and interpreted in Section 4, while the model evaluation results based on the 24-h concentration patterns of individual tracer releases are presented and interpreted in Section 5. Section 6 presents the conclusions of this study and compares the final results to those of similar studies.

SECTION 2

PREVIOUS MODEL EVALUATION STUDIES

Past studies have assessed the reliability and uncertainties of long-range transport and diffusion models by comparing model-calculated concentrations to measured concentrations of inert tracer gases. One of the early studies compared calculations of a Lagrangian puff model to twice-daily averages of Krypton-85 concentrations 1,500 km downwind from the point source (Draxler, 1982). This study concluded that, unless a model simulates wind shear effects, surface concentrations will be overestimated.

More recent studies used the 3- and 6-h surface and 6- to 10-minute aircraft measurements of an inert tracer gas during the 1983 CAPTEX (Ferber et al., 1986) and the concentrations of radioactive material from the 25 April 1986 Chernobyl reactor accident (Persson et al., 1986). Haagenson et al. (1987) concluded that the root-mean-square (rms) separation distances between all CAPTEX tracer and transport model trajectories for all time periods ranged from 220 km to 260 km. They also concluded that absolute trajectory errors were 120 km and 200 km after 12 h and 24 h of transport, respectively. These estimates were within 20% of those of Kahl and Samson (1988). For the MESOPUFF-II model, Godowitch (1989) calculated a mean error of approximately 150 km after 24 h to 50 h of transport.

Draxler (1987) assessed the mean transport errors of a simple puff trajectory model (Draxler, 1982) using routine meteorological data and the tracer data from all CAPTEX time periods. Mean trajectory errors for this model were slightly lower than those of others--approximately 80 km and 150 km after 12 h and 24 h of transport, respectively. Between 30 h and 42 h of transport, the error was virtually constant at approximately 180 km. The mean ratios of predicted to measured concentrations also were a function of time after release; for the first 24 h, the model tended to underpredict by 50%, but after 30 h, the underprediction decreased to 10% or less.

Using data from each CAPTEX tracer release, Lee (1987) assessed mean transport errors of a three-dimensional particle-in-cell model. The author concluded that the position of the simulated plumes in general corresponded well to those of the measured plumes. The frequency histogram of residuals computed from spatially- and temporally-paired concentrations illustrated a modal value near zero--72% of the pairs of concentrations fell between ±400 decifemtoliters per liter (10^{-16} L/L or dfL/L), and 59% of the

time, the model overpredicted the concentrations. Because the frequency of concentrations in excess of 50 dfL/L was greater for the predictions, the author concluded that the simulated horizontal diffusion was overpredicted.

For two CAPTEX tracer releases, Kao and Yamada (1988) computed plume trajectories and surface tracer concentrations using a random-particle statistical method and four-dimensional data assimilated wind fields. For these two cases--(1) a light-wind, fair-weather situation with a wide surface plume and (2) a cold frontal situation with a rather narrow plume--the simulated concentration patterns corresponded well with those of the measurements in that the center lines of the composite simulated and measured plumes coincided after 54 h.

These studies indicated that after 12 h and 24 h of transport, absolute model trajectory errors ranged from 80 km to 150 km and 150 km to 220 km, respectively. One study showed that trajectory errors approximated 180 km after 30 h to 42 h of transport. After 72 h of transport, another study indicated that, 50% of the time, the error would exceed 350 km. In Section 5.2, these errors are compared to those computed in the ANATEX Model Evaluation Study. A range of concentration prediction errors could not be established on the basis of the published results.

SECTION 3

MODEL EVALUATION PROTOCOL

3.1 THE EVALUATION APPROACH

The AMES objectives are to:

(1) assess the overall performance (for the first three months of 1987), as well as the model errors on temporal scales of 24 h, of prognostic long-range atmospheric models with respect to transport and diffusion as a function of transport time and distance,

(2) intercompare the model performances and relate performance to fundamental modeling approaches,

(3) identify the periods and associated meteorological conditions when each model performed best and worst, and

(4) compare and contrast the AMES conclusions with those of similar studies using CAPTEX data.

These objectives are accomplished by first developing a set of model performance measures relating the basic attributes of the model-calculated transport and diffusion to those of the measurements. The values of these performance measures are determined from comparisons between model calculations and ANATEX measurements of the 24-h surface tracer concentrations at each of the 77 sites in the network. From the values of these measures, the performance of each model is compared to that of a perfect model as well as to each other.

Basic attributes of the transport and diffusion characterized from the ANATEX surface data are:

(1) magnitude of tracer concentrations,

(2) extent of horizontal diffusion,

(3) transport speed, and

(4) transport direction.

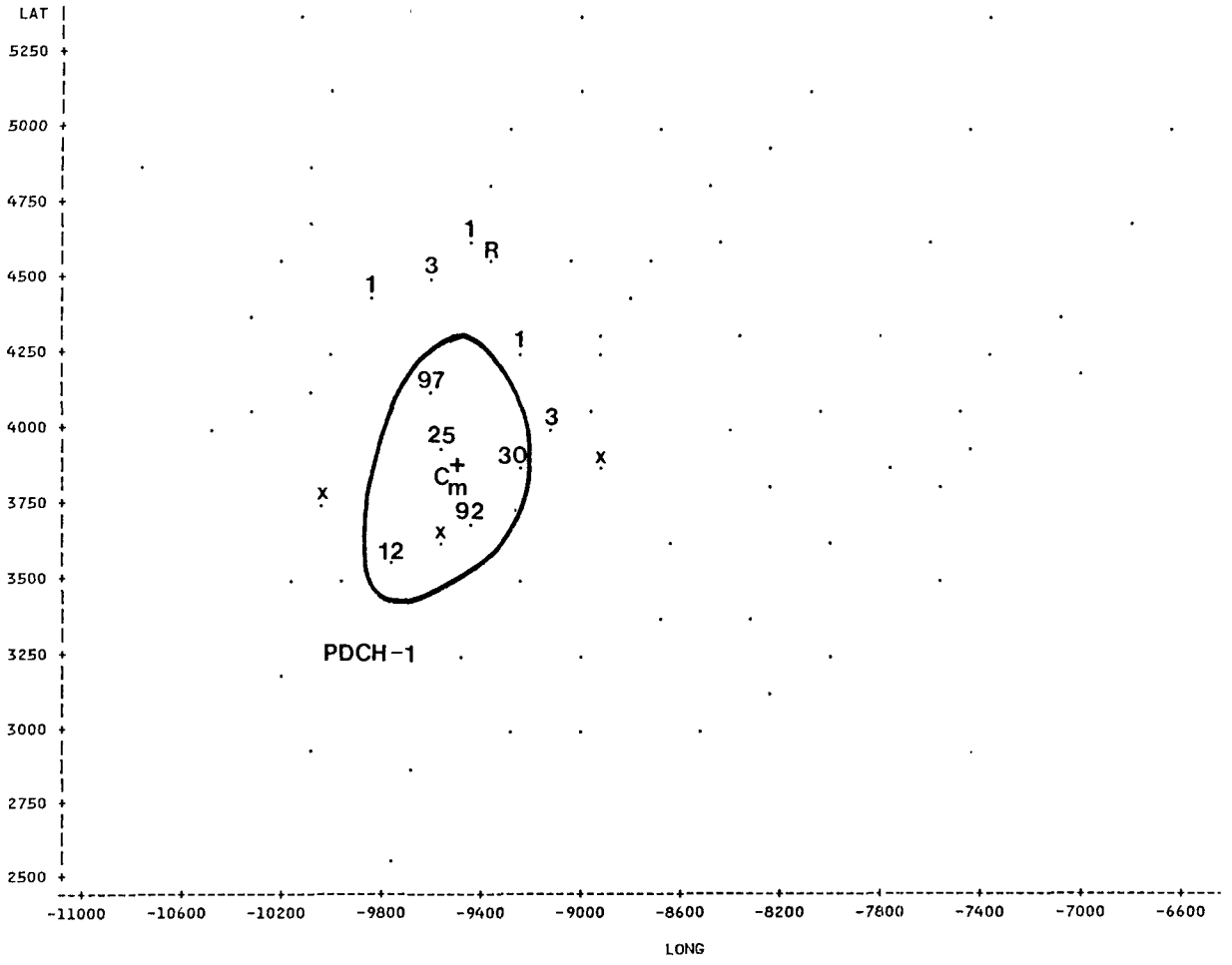
The most accurate evaluation of a predictive model is carried out on an individual event, rather than average behavior over an ensemble of events. The reason for this is that a model can err on individual events, but if those errors are unbiased, it may still replicate an ensemble of measurements on the average. Thus, it may appear to be accurate, but in fact it may not be accurate for the right reasons. If, on the other hand, a model successfully replicates each individual event, it is likely that it has done so through accurate representation of the relevant meteorological factors.

Therefore, one would ideally evaluate each of the characteristics listed above through the thorough analysis of each individual release as it is carried through time and space. The 24-h-mean surface concentrations from a single tracer release define a geometric entity of concentrations exceeding threshold values (see page 19), heretofore referred to as a "footprint." This entity differs from an instantaneous tracer puff in that the footprint is a composite of the surface manifestations of all instantaneous puffs within the 24-h sampling period (Figure 2).

In the ANATEX database, it is not always possible to associate a concentration with a particular tracer release. Footprints from separate releases have a tendency to be merged together in both the measurements and model calculations. While many individual footprints remain identifiable for the measurements and for each model, it is extremely rare that one particular release can be followed for the measurements as well as for <u>all</u> of the models. Thus, the intercomparison of models using the ANATEX database is somewhat difficult.

An additional complication with regard to tracer concentration levels and horizontal diffusion is the spatial and temporal resolution of the sampling network. That is, the determination of horizontal diffusion requires knowledge of the temporal changes in the width of the footprints, yet the separation of monitors along a band is generally greater than 100 km. Thus, the potential error in estimating the width of a single footprint can be as great as 200 km. Similarly, the uncertainty of the estimated mean tracer concentration of a footprint can also be unacceptably large for one 24-h period, since the higher concentrations can often elude the monitoring sites.

For these reasons, measures of the magnitude and diffusion of concentrations are assessed using average behavior over time, despite the acknowledged weakness therein. Specifically, they are characterized for two ensembles (both halves of the ANATEX period: 5 January - 16 February 1987 and 17 February - 29 March 1987), and for neither individual 24-h periods nor individual footprints. The splitting of the ANATEX period was necessitated by the fact that the sets of meteorological scenarios of the two halves were generally quite different. Due to the higher frequency of days with frontal systems stretching across



x = missing data

R = STC release site

Figure 2. Illustration of the 6 January 1987 "footprint", or a 24-h composite of instantaneous tracer puffs near the surface, and the centroid, C_m, for the measured PDCH concentrations from release 1.

the northcentral U.S. after 16 February 1987 (24 days versus 14 days in the first half period [Draxler, 1988a]), the transport and horizontal spreading of the tracer tended to be different than that in the first half. Consequently, the model performances would be expected to differ for each half.

In interpreting these measures of magnitude and diffusion, it is important to realize that any conclusions must be somewhat one-sided. That is, if a model performs poorly on average, it can be inferred that it must perform poorly for individual events as well. However, if a model performs well on average, that does not necessarily mean that it successfully replicates individual events.

With regard to transport speed and direction, one must base any meaningful analysis on concentrations associated with a specific tracer release. Such an approach is used in this evaluation, acknowledging the weakness inherent in utilizing different subsets of the data for each model as well as the measurements. In order to minimize this weakness, certain analyses of transport characteristics are performed on the ensemble database. While not as powerful as the analyses of individual releases, the ensemble results can be used to lend support to those more specific analyses.

Describing attributes of transport first requires a tracer footprint reference point so that a distance can be calculated as a function of elapsed time. Options include the farthest edge from the source, the maximum concentration, the concentration-weighted center, and the geometric center.

Haagenson et al. (1987) considered the first three options, or variations thereof, in their CAPTEX evaluation of isentropic, isobaric, and sigma models. They concluded that the third option--the concentration-weighted center--was best, since it consistently yielded the smallest rms separation for all models. Furthermore, this option is more suitable than the geometric center option, since it is more closely related to the core of the footprint (i.e., the highest concentrations) and is less affected by one-sided shearing of the footprint. Also, the concentration-weighted center is minimally affected by uncertainties in defining the footprint edges. Therefore, the concentration-weighted center or centroid was selected to represent both calculated and measured ANATEX footprint locations.

For each of the calculated and measured footprints, the coordinates (X, Y) of the centroid were computed from the set of n nonzero 24-h concentrations (χ_i) and the coordinates of the sampling sites (x_i, y_i) as:

$$X = \sum w_i x_i \qquad i = 1, \ldots, n$$
$$Y = \sum w_i y_i \qquad i = 1, \ldots, n$$

where the weights, w_i, were defined as--

$$w_i = \chi_i \,/\, [\sum \chi_i] \qquad i = 1, \ldots, n$$

As an example, Figure 3 illustrates the location of the centroids for ARL and actual footprints of PTCH tracer at 24-h intervals up to five days after release number 1.

For each of the features of transport and diffusion listed in Section 3.1, one or more model performance measures were adopted. These measures, identified in Table 4 and discussed in the following sections, could be either quantitative or qualitative, where the latter would take the form of maps or plots. Note that, due to the spatial and temporal dependence of the measurements, formal statistical inference (i.e., testing the null hypothesis that calculated and measured parameters are statistically similar) is generally not possible.

Figure 3. The 24-h locations of ARL (---) and actual (——) footprint centroids for the first release of PTCH tracer.

TABLE 4. MODEL PERFORMANCE MEASURES[a]

Ensemble statistics		24-h concentration patterns of individual tracer releases	
Concentration magnitude (χ)	Horizontal diffusion	Transport speed	Footprint location
Box plots of nonzero χ's [Section 4.1.1][b]	Mean distance of nonzero χ's along bands [Section 4.2]	Transport speed ratio:[c] D_c / D_m [Section 5.3]	Normalized location error:[d] D_e / D_m [Section 5.3]
Distributional bar charts [Section 4.1.1]			Mean separation error:[e] D_e [Section 5.2]
Mean χ along bands [Section 4.1.2]			

a. Used to describe the agreement between characteristics of calculated and measured tracer transport and diffusion.

b. Results are discussed in the section number given in brackets.

c. The ratio of the cumulative distance (D) between the centroid of calculated footprints (C_c) and the tracer source over the cumulative distance between the centroid of measured footprints (C_m) and the tracer source.

d. The ratio of the cumulative distance separating the centroids of calculated and measured footprints over the cumulative distance between the centroid of measured footprints (C_m) and the tracer source.

e. The mean distance between the centroids of the calculated and measured footprints.

No calculated concentration was considered for those cases when the measured concentration was either missing or ignored on the basis of the screening criteria, thus, avoiding an inconsistency that has the potential of dominating the centroid locations. For all but one model, there were 5,296 and 5,176 pairs (or 82% and 80% of the maximum possible pairs) for the PTCH and PDCH tracers, respectively. Since HY-SPLIT did not calculate concentrations at three sites along the border of the network, the number of pairs decreased by approximately 230.

3.2 MODEL PERFORMANCE MEASURES

3.2.1 Ensemble Concentration Means and Distributions

One commonly-used approach of assessing model performance regarding transport and diffusion is to compare calculated concentrations to measured concentrations paired in space and time. This approach is not suitable for an evaluation using the ANATEX data, since the temporal and spatial resolutions of the ANATEX network is too coarse to resolve individual plumes. Furthermore, this approach has a serious disadvantage in that the simulated mass transport could differ only slightly from the actual transport (i.e., be well-within the expectations of the model), yet a statistical comparison of concentrations paired both in space and time would infer that the model performed dismally. For example, suppose after 48 h of transport a model calculates a concentration of 20 dfL/L at site A and 0 dfL/L at site B, 30 km from site A, while the measurements indicate that the opposite is true. A transport error of only 30 km after 48 h is hardly significant and one could conclude that the model performed very well. However, a statistical measure based on concentrations paired in space and time (e.g., the mean square error) would falsely imply total failure and would not by itself inform the evaluator of the proximity of the calculated plume to the measured plume.

As an alternative to this approach, a model performance measure was adopted to assess model performance on the basis of the differences between calculated and measured ensemble mean concentrations along six bands of sites (Figures 4a and 4b) for each half of the ANATEX period. The sites within each band were approximately equidistant to the tracer release site; 1,000-, 1600-, and 2,300-km for the PTCH tracer released from GGW and 300-, 700-, and 1,400-km for the PDCH tracer released from STC. Supplementing these analyses are time-series plots of 24-h mean calculated and measured concentrations (Appendix B). Only the three bands were selected for PTCH because the sampler spacing--and thus, our uncertainty of the true mean concentration--was considerably larger for the 1,300- and 2,000-km bands. Some sites located near the band end points were not considered when calculating this measure to preserve the angle from source to band end points and to exclude some of the sites that often reported no data.

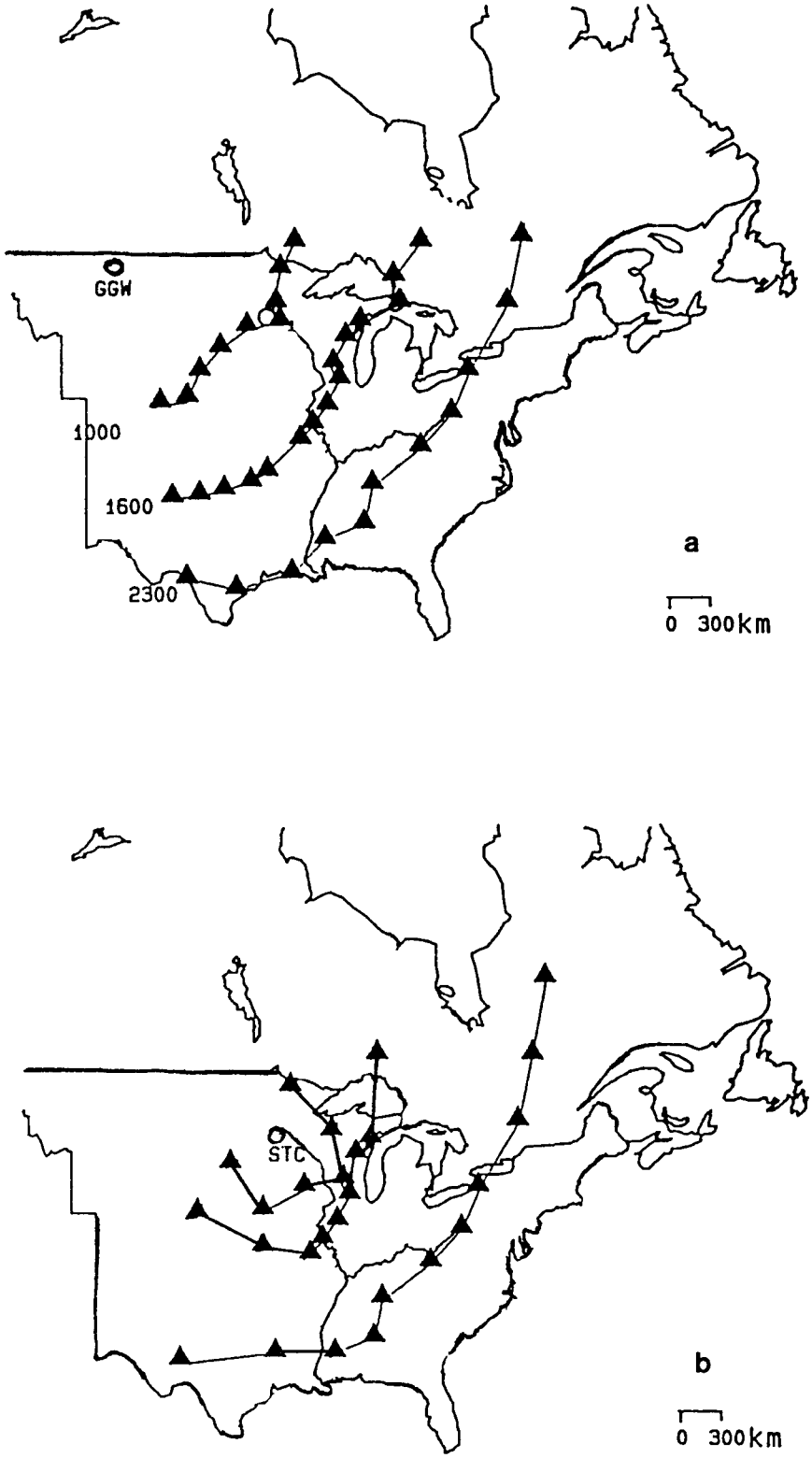


Figure 4. The bands of sites along which ensemble-mean calculated and measured horizontal diffusion were estimated for (a) PTCH and (b) PDCH tracers.

With this approach, differences in mean concentrations (i.e., model biases) can be related to transport distances. Furthermore, the ensemble-mean-concentration approach avoids a comparison of 24-h concentrations when, for example, the centerline of the measured plume avoids the sampling sites, while that of the calculated plume passes over a sampling site. By comparing ensemble mean concentrations, the influence of such cases on the model performance measure can be minimized.

The ensemble-mean concentrations ($\overline{\chi}$) were calculated as follows:

$$\overline{\chi} = 1/n \sum \chi_i \qquad i = 1, \ldots, n$$

where: n is the total number of concentrations along the band for every day in the ensemble period (i.e., the number of *site-days*), and

χ_i is the concentration at site-day i.

The mean concentrations were based on only those site-days for which both a calculated and a measured concentrations were simultaneously available. For this measure, both zero and nonzero concentrations were used.

In addition to the ensemble mean concentration comparisons, the ensemble distributions of measured and calculated concentrations for each half period are compared via box plots and distributional bar charts. These graphic tools quickly display any biases that may exist. The distributional box plots consider only nonzero concentrations to focus on the concentrations in the tracer plumes. Meanwhile, the distributional bar charts were restricted to those calculated concentrations corresponding to available measured concentrations passing the screening criteria. This was done to present a better comparison of the distributions of calculated and measured concentrations.

3.2.2 Ensemble Mean Horizontal Diffusion

As mentioned earlier, it is difficult to adequately calculate the horizontal diffusion of the tracer footprints, especially for the first 24 h of transport, due to the spatial resolution of the monitoring network. Rather than attempting to estimate the horizontal spread for each 24-h period, we used a measure based on the average distance along the 1,000-, 1,600-, and 2,300-km bands for PTCH tracer and the 300-, 700-, and 1,400-km bands for PDCH tracer (Figures 4a and 4b) of nonzero concentrations. Specifically, the mean horizontal spread is approximated by the ensemble product of the mean daily number of samplers along a band with nonzero concentrations and the average distance between samplers for that band.

Only the three bands were selected for PTCH because the sampler spacing--and thus, our uncertainty of the true footprint width--was considerably larger for the 1,300- and 2,000-km bands. Some sites located near the band end points were not considered when calculating this measure to preserve the angle from source to band end points and to exclude some of the sites that often reported no data.

The model performance measure is the ratio of the average distance for the model calculations over the average distance for the measurements. The averaging periods are the first and second halves of the ANATEX sampling period. For a model that, on the average, underpredicts the horizontal diffusion, the ratio will be less than 1.

3.2.3 Transport Speed of Discrete Tracer Footprints

The agreement between calculated and measured transport speed is quantified by the transport speed ratio defined as: the cumulative distance between the centroid of calculated footprints and tracer source, D_c , divided by the cumulative distance between the centroid of measured footprints and tracer source, D_m , for the same period (Figure 5). The cumulative distance is measured along the line segments connecting the centroids of each footprint series (from one 24-h centroid to the next), as opposed to a straight-line distance.

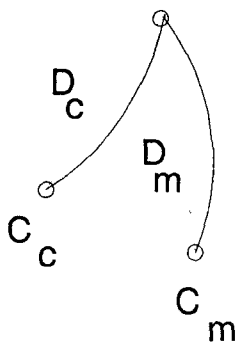
In calculating the centroids for the measured and calculated footprints, only concentrations at or above a specified threshold were considered. The threshold concentrations--defined as 6 dfL/L for PTCH and 9 dfL/L for PDCH--were necessary since: (1) many of the sites with concentrations below the threshold values were isolated and located where actual tracer was not expected, and (2) uncertainties of concentrations [5] below the threshold values tended to be at least ±25% (Figure 6). These threshold concentrations were defined so that:

(1) nearly all concentrations greater than the threshold are statistically significant at the $\rho = 0.1$ significance level, and

(2) as many discrete footprints as possible could be identified with minimal disturbance of the size and shape of the footprints.

A measure based on cumulative distances, as opposed to one based on discrete segment distances for single 24-h periods, was better suited for model evaluation, as illustrated by the following example. Suppose that a model underpredicts the transport speed for one 24-h period and overpredicts the transport speed by the same magnitude for the next 24-h period. The transport speed ratio for the first period would

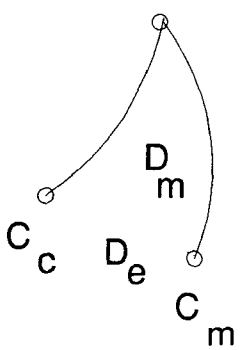
[5] Uncertainties were quantified by Draxler and Heffter (1989) for two terms: (1) the uncertainty in the gas chromatograph background, or noise level, and (2) the uncertainty in the concentration from the analytical precision of measuring reference standard tracer volumes for each of the three analytical laboratories.



a

Transport Speed Comparison

$$\frac{D_c}{D_m}$$



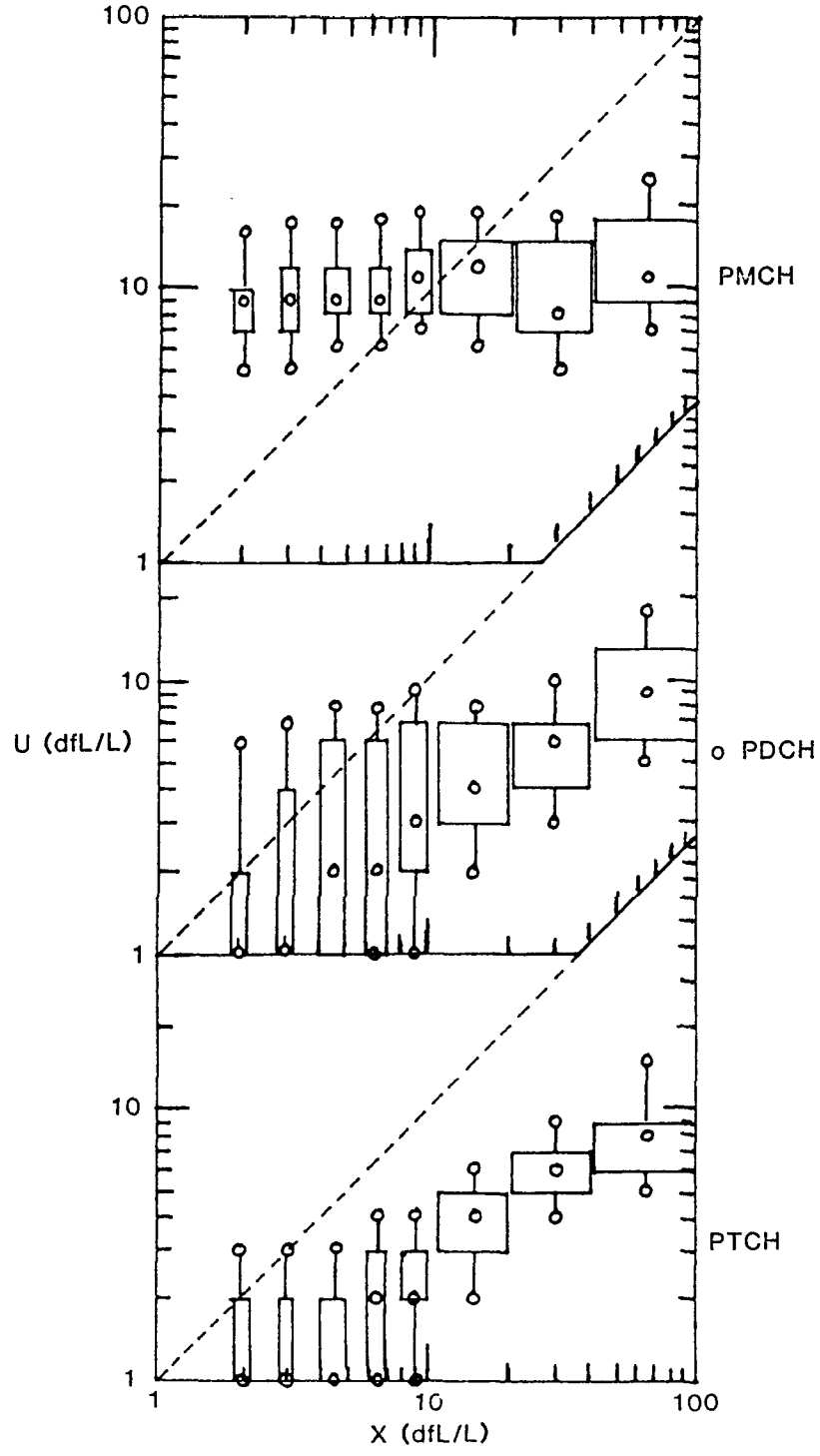
b

Centroid Location Error

$$\frac{D_e}{D_m}$$

D_c = distance between calculated centroid and tracer source
D_m = distance between measured centroid and tracer source
C_c = centroid of the calculated footprint
C_m = centroid of the measured footprint
D_e = distance between C_c and C_m

Figure 5. The model performance measures characterizing (a) the agreement between calculated and measured transport speed and (b) footprint location errors.



Note: boxes indicate the 25th and 75th percentile values, while circles show the 10th, 50th, and 80th percentile values.

Figure 6. Box plots of cumulative frequency distribution of uncertainty (U) versus excess concentration (χ) for PMCH, PDCH, and PTCH (from Draxler and Heffter, 1989).

be less than 1, while the ratio for the second period would be 1--a perfect score. Since for the 48-h period the calculated transport speed was identical to that measured, the model receives credit yet the model error for the first 24-h period is considered. If discrete distances were used, the model would be penalized for both 24-h periods.

For a perfect model, the ratio would equal 1 for every 24-h measurement period. If the value usually exceeded 1, this measure would indicate that the model tended to overpredict the transport speed. On the other hand, if the value usually is less than 1, the measure would indicate that the model tended to underpredict the transport speed.

3.2.4 Tracer Footprint Location Error

Two performance measures describe the agreement between calculated and measured transport direction. The primary measure is the normalized separation distance, D_e / D_m, between the centroids of the calculated and measured footprints (Figure 5). The normalization factor is the cumulative distance between the centroid of the measured footprint and tracer source. By normalizing the separation distance, less emphasis is placed on smaller model errors after long transport distances. This measure was selected over one based on angle differences, since the separation-distance measure is commonly used to determine calculated transport errors (Haagenson et al., 1987; Draxler, 1987; Kahl and Samson, 1988; Godowitch, 1989).

Unlike the transport speed ratio, this measure has a value of zero for a perfect model. When the calculated centroid is to the right (left) of the measured centroid--relative to the source location--the value of this measure is defined to be positive (negative). As was the case of the transport speed ratio, a model can be properly rewarded for calculating the correct location of the centroid after 48 h of transport even if it incorrectly calculated the centroid after 24 h of transport.

The second measure is the mean separation error expressed as a function of cumulative transport time. In addition to providing a useful description of model performance, this measure is often calculated by other model evaluators. It, therefore, affords us the means to compare the results of this model evaluation with those of others.

SECTION 4

ASSESSMENT OF THE MODEL PERFORMANCES BASED ON ENSEMBLE DATA

4.1 COMPARISON OF CONCENTRATION MAGNITUDES

4.1.1 Ensemble Concentration Distributions

Graphical displays of the concentration distributions in the form of box plots and distributional bar charts were constructed for the calculated and measured concentrations for two periods--the first half and the second half of the ANATEX period--and for both tracers. For the box plots, the range of concentrations, excluding those exceeding the 99th-percentile, is indicated by the two short horizontal bars. The first, second, and third quartiles are indicated by the longer horizontal bars, where the second quartile is the median. Means are represented by the asterisks. Since only nonzero concentrations are considered in the box plots, a somewhat different sample is considered for each model.

When interpreting the box plots of measured concentrations, one should understand that the distribution could slightly overstate the frequencies of very low concentrations. This is a consequence of those cases when some low measured concentrations were actually zero. Although the data uncertainties were assessed and a screening procedure was executed, some false occurrences of nonzero concentrations probably remained.

The distributional bar charts were constructed from every pair of calculated and measured concentrations. Each of these charts considers four classes:

for PTCH-- 0 dfL/L[6] , 1-5 dfL/L, 6-99 dfL/L, and >99 dfL/L, and

for PDCH-- 0 dfL/L, 1-8 dfL/L, 9-99 dfL/L, and >99 dfL/L.

The ranges differed slightly for the two tracers to distinguish the concentrations below and above the thresholds. Natural background concentrations are <0.3 dfL/L for PTCH and <1.0 dfL/L for PDCH, both of which are below current analytical sensitivity.

---

[6] 1 dfL/L = 10^{-16} L/L

PTCH (GGW) Concentration Distributions--

The PTCH box plots (Figures 7a and 7b) indicate that the median measured concentrations were 5 dfL/L and 6 dfL/L for the first and second halves, respectively. The third quartiles approximated the measured means (denoted by the asterisks in Figures 7a and 7b) and ranged from 10 to 15 dfL/L.

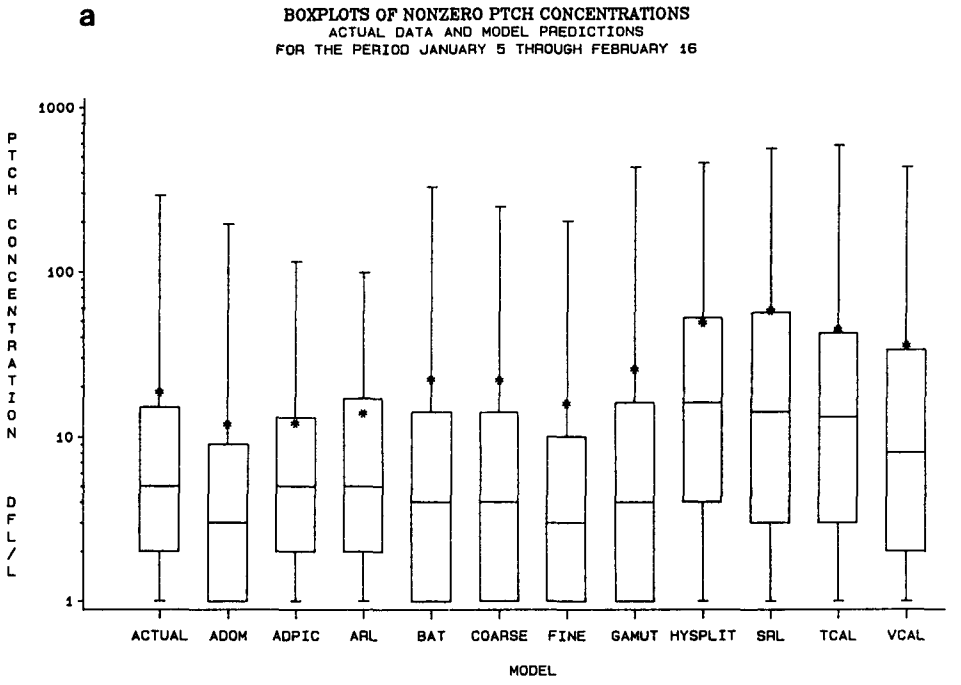
The PTCH distributional bar chart (Figures 8a and 8b) shows that approximately 60% of the time, the measured PTCH concentration was zero. Concentrations exceeded 5 dfL/L 20% of the time and exceeded 99 dfL/L less than 2% of the time.

In the first half, the median and mean PTCH concentrations for HY-SPLIT, SRL, TCAL, and VCAL were considerably higher than those of the measurements (Figure 7a). The medians of the first three models exceeded 10 dfL/L and their first quartile concentrations approximated the median of the measurements, indicating a tendency for these models to overpredict. The quartiles of the other models were very similar to that of the measurements.

Compared to the distribution of measurements (Figure 8a), HY-SPLIT, SRL, TCAL, and VCAL also calculated fewer occurrences of nonzero concentrations, but more occurrences of concentrations exceeding 99 dfL/L in the first half. Thus, providing that the calculated transport speeds are accurate, the horizontal and/or vertical diffusion calculated by these models might be underestimated. The ARL, BAT, GAMUT, and MLAM-COARSE also calculated fewer occurrences of nonzero concentrations, but unlike the four models mentioned above, they did not calculate more occurrences of concentrations in excess of 99 dfL/L. While the ADOM and MLAM-FINE distribution closely resembles the distribution of the measurements, the ADPIC distribution shows considerably more occurrences of nonzero concentrations, suggesting the possibility that this model's horizontal diffusion rate is too high and/or the calculated transport speed is too slow.

In the second half, the SRL median and mean concentrations were an order of magnitude greater than those of the measurements (Figure 7b). Its first quartile exceeded 10 dfL/L and was higher than the third quartile measured concentration. Figure 8b illustrates a serious discrepancy between the SRL distribution and the distribution of the measurements: 98% of the time, SRL calculates a zero concentration. Apparently, the diffusion rates in this half were much too low and differed from those of the first half.

Unlike the SRL median, the other medians were within ±3 dfL/L of the median of measurements (Figure 7b). Again HY-SPLIT and VCAL concentrations were higher than the measurements, but less so than in the first half. The second quartiles for ARL and BAT were considerably lower than those for the measurements. The distributions for the remaining models were very similar to the distribution of the measurements. Note that ADOM and MLAM-FINE results were unavailable for the second half.



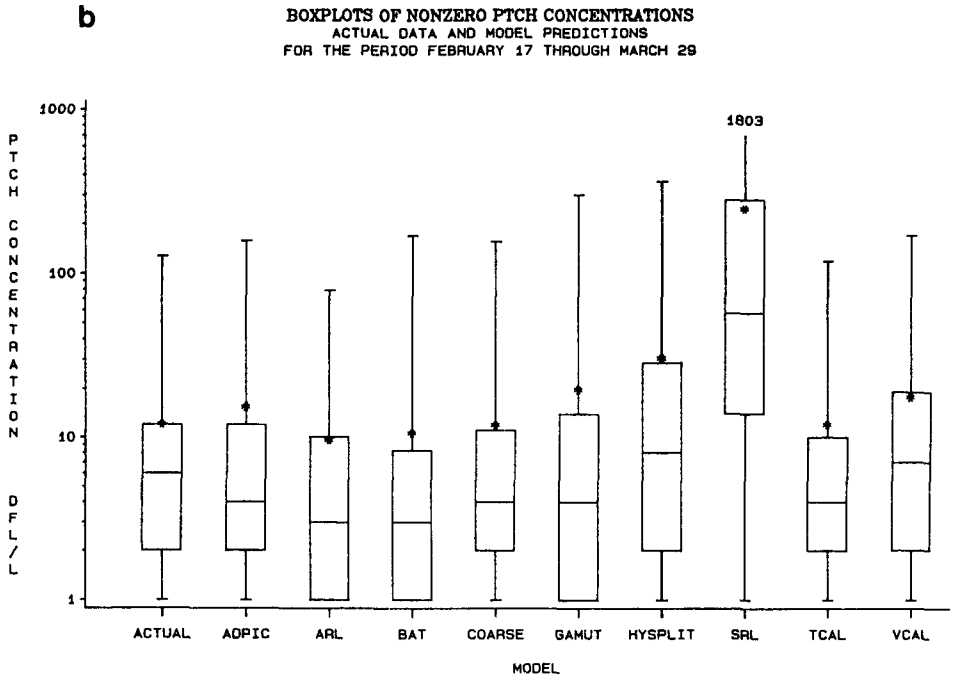
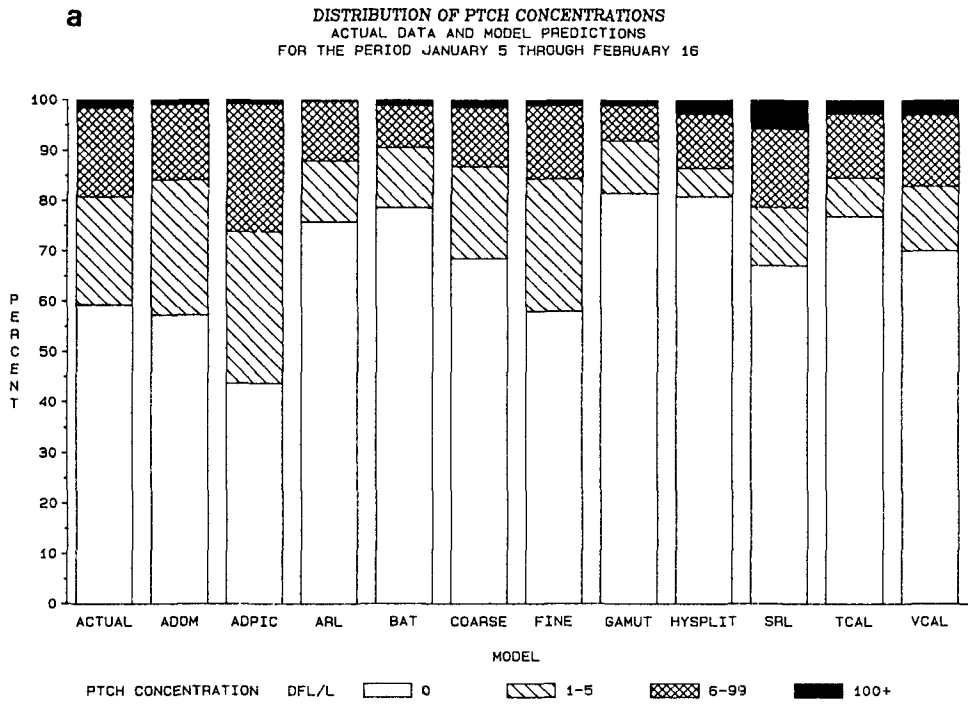


Figure 7. Box plots comparing quartiles and means (*) of calculated and measured PTCH tracer concentrations at the 77 surface sites during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987.



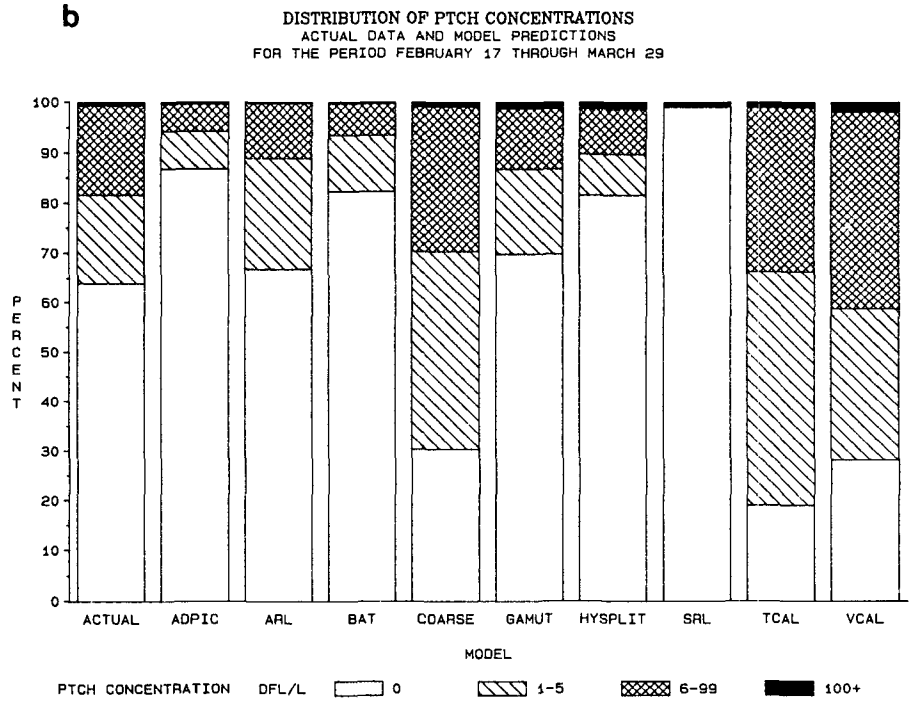


Figure 8. Frequency distributions of calculated and measured PTCH tracer concentrations at the 77 surface sites during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987.

Nonzero concentration occurred much more often for MLAM-COARSE, TCAL, and VCAL than for the measurements during the second half (Figure 8b). This suggests that the horizontal diffusion rates calculated by these models may be too high for this half, which has a higher frequency of cyclonic and/or frontal passages and accompanying wind shears.

It is noteworthy to contrast the poor agreement of TCAL and VCAL (both SLL models) with the good agreement of GAMUT. For the complicated wind patterns common to the second half, the multilayer model produced a distribution much closer to that of the measurements than did the single-layer models. In addition to the GAMUT distribution, the ARL distribution closely resembled that of the measurements. The remaining distributions (i.e., ADPIC, BAT, and HY-SPLIT) showed a higher frequency of zero concentrations. For both BAT and HY-SPLIT, this was consistent with the first half; however, for ADPIC, the first-half distribution showed a much lower frequency of zero concentrations.

It is also noteworthy to mention that HY-SPLIT--a version of ARL that used finer resolved meteorological input data--predicted PTCH concentrations in excess of 99 dfL/L more often than ARL, suggesting that the HY-SPLIT diffusion rate was slower than that of ARL. This difference was also observed in the first half.

PDCH (STC) Concentration Distributions--

Distributional differences were detected between PDCH and PTCH. As the second pair of box plots (Figures 9a and 9b) illustrate, low concentrations were measured more often for PDCH than for PTCH. The median concentrations for both halves were 3 dfL/L, about a factor of 2 lower than the PTCH medians. Also lower were the third quartiles--9 dfL/L and 8 dfL/L for the first half and the second half, respectively. The higher occurrence of low PDCH concentrations can be explained by the lower PDCH emission rates--on the average, 41.6% lower than those for PTCH.

For both halves, all mean calculated concentrations exceeded the mean measurements and, in some cases, exceeded them by an order of magnitude. (This was not observed for the PTCH concentrations since the mean concentration, which is greatly influenced by the highest values, was based on concentrations at sites at least 500 km downwind of the tracer source--much farther than the closest sites to the PDCH source.) In addition, nearly all the models calculated maximum concentrations considerably higher than those measured; some exceeded 500 dfL/L--more than double those measured. As was the case with the first-half PTCH concentrations, first-half PDCH quartiles were highest for GAMUT, HY-SPLIT, SRL, TCAL, and VCAL. However, all model medians were higher than the median of the measurements. In addition, the third quartiles were at least twice those of the measurements for all models but ADOM and ADPIC. The ADOM and ADPIC quartiles closely resembled those of the measurements.

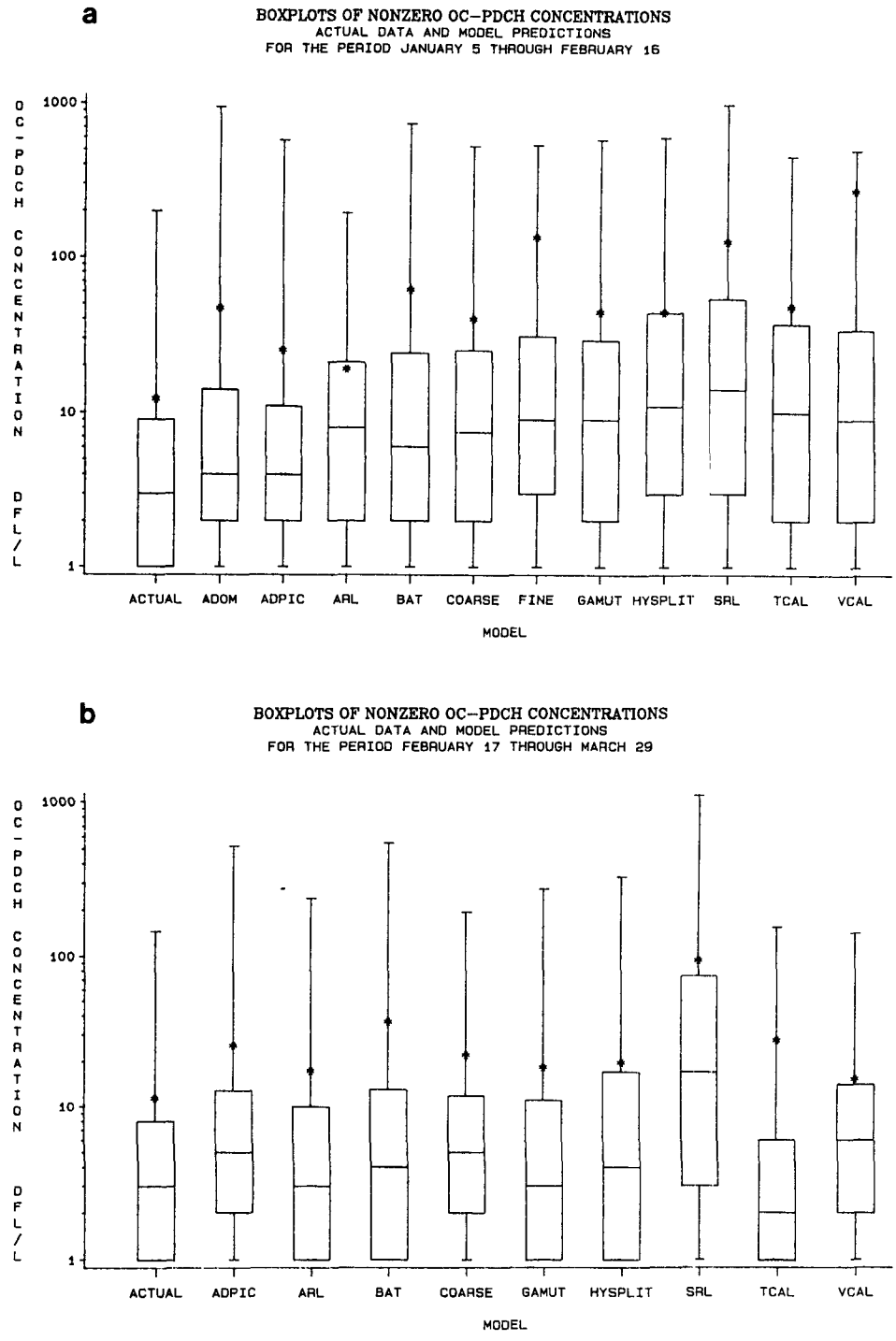


Figure 9. Box plots comparing quartiles and means (*) of calculated and measured PDCH tracer concentrations at the 77 surface sites during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987.

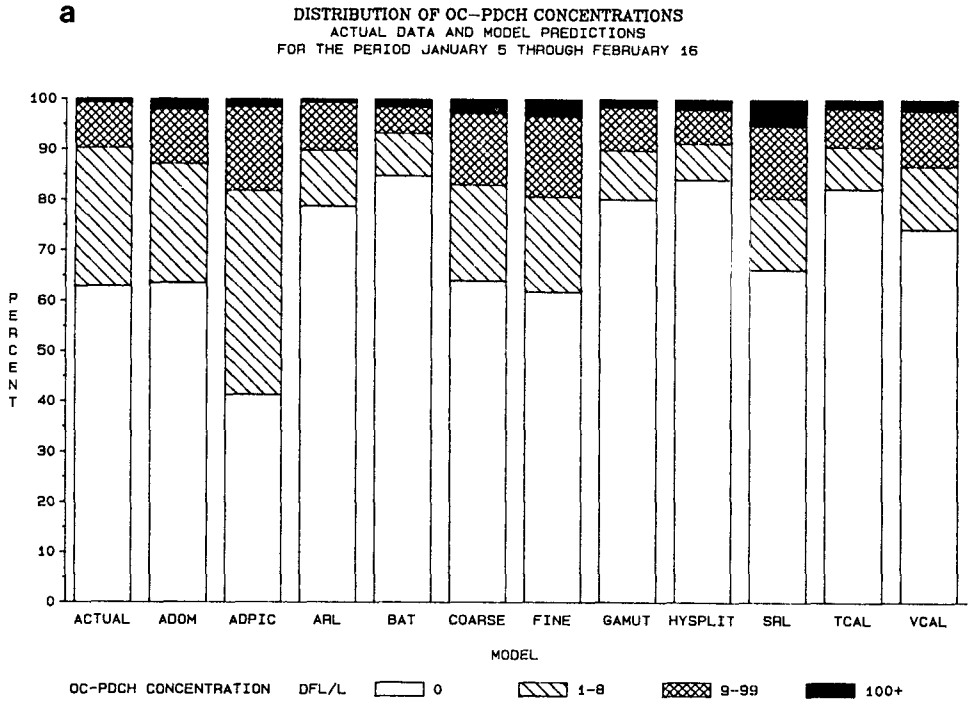
The differences between the model and measured PDCH distributional bar charts (Figures 10a and 11b) for the first half resembled those in the first-half PTCH charts. For example, SRL calculated concentrations in excess of 99 dfL/L most often, while ADPIC calculated nonzero concentrations most often. In addition, ARL, BAT, GAMUT, HY-SPLIT, TCAL, and VCAL nonzero concentrations occurred less often than the nonzero measurements--15% to 25% versus nearly 40%, respectively. More importantly, however, the frequency of PDCH concentrations exceeding the threshold for these six models was about the same as that for the measurements. The ADOM distribution closely resembled the distribution of measurements. The distributions of the final two models--MLAM-COARSE and MLAM-FINE--were similar to the distribution of measurements in that the frequencies of nonzero concentrations were virtually the same, but their frequencies of concentrations above the threshold were nearly double that for the measurements.

In the second half, SRL quartiles were once again the highest (Figure 10b). Also, as the case for PTCH, the SRL first quartile approximated the median measured concentration. The medians of the other models were within ±2 dfL/L of the median of the measurements. With the exception of TCAL, the model third quartiles were higher, but not nearly as much as in the first half.

As was the case for its PTCH concentrations, SRL PDCH concentrations were zero much more often and MLAM-COARSE, TCAL, and VCAL concentrations of PDCH were nonzero more often than those measured indicating a problem with diffusion. To a lesser degree, BAT and HY-SPLIT concentrations also were zero more often, as was the case in the first half; this suggests that these two models slightly underestimate horizontal diffusion. In contrast, ADPIC, ARL, and GAMUT concentration distributions resembled those of the measurements.

4.1.2 Mean Concentrations Along Bands

One test assessing the combined effects of model transport speeds and diffusion rates is the comparison between model and measured mean concentrations at specific distances downwind of GGW and STC. If, for example, the model mean concentrations are appreciably greater than the measured mean concentrations, one or more of the following is true about the model: (1) the transport speeds are too high, (2) the horizontal diffusion rates are too low, or (3) the vertical diffusion rates are too low. Unfortunately, the evaluation data set cannot distinguish between these error sources. Across three bands of samplers--where samplers along a band are quasi-equidistant from the tracer sources--the ensemble mean concentrations of the calculations and measurements are determined and compared for each tracer for each half of the ANATEX period.



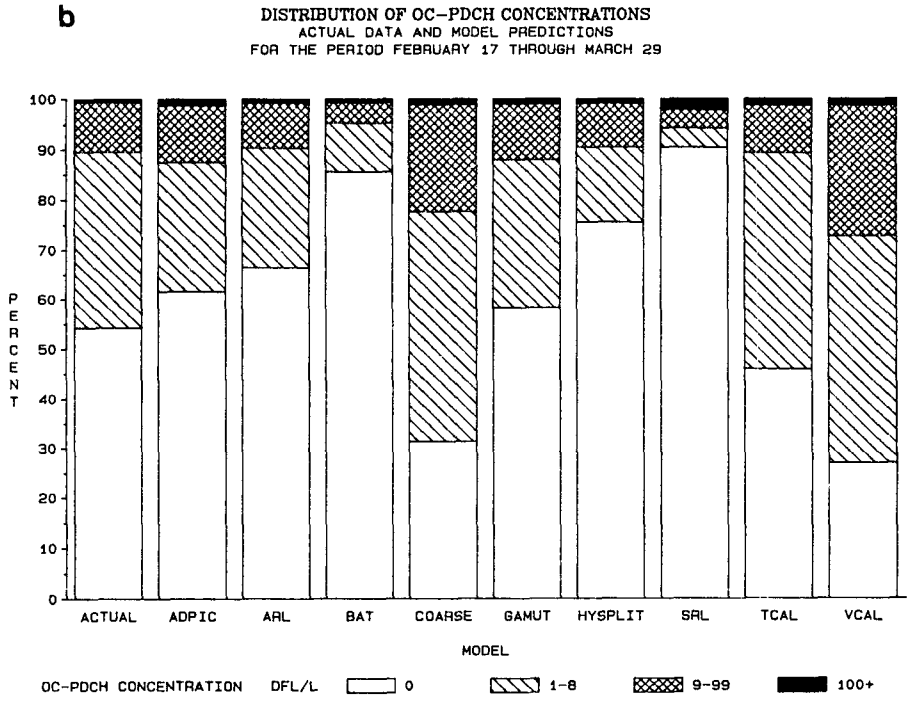


Figure 10. Frequency distributions of calculated and measured PDCH tracer concentrations at the 77 surface sites during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987.

Mean PTCH (GGW) Concentrations--

The mean PTCH concentration plot for the first half (Figure 11a) shows a wide range of model mean concentrations as the tracer plumes were transported from their sources to the 1,000-, 1,600- and 2,300-km bands. The mean measured concentration decreased sharply from nearly 18 dfL/L at 1,000 km to 5 dfL/L at 1,600 km and then decreased at a much slower rate to more than 3 dfL/L at 2,300 km (Table 5). The models tended to smooth through the steepest segment of the gradient; that is, the model mean concentrations generally were lower closest to the source, while the model mean concentrations were more comparable to the mean measured concentrations at the mid-range. Most models were within ±1 dfL/L of the mean measured concentration at the 2,300-km band.

Most obvious is the fact that SRL mean concentrations in the first half were the highest for each band and are 2 to 3 times higher than the measured mean concentrations. In contrast, SRL mean concentrations in the second half (Figure 11b) were the lowest and approach zero at the 1,600- and 2,300-km bands. Thus, as seen before, this model's diffusion rates and/or transport speeds appeared to be inconsistent with reality.

ADPIC, HY-SPLIT, TCAL, and VCAL replicated the measured mean concentration gradients rather well for the first half, but of these models only ADPIC and HY-SPLIT replicated the gradients for the second half. ADOM also replicated the gradient in the first half, while that of MLAM-FINE was much flatter. The mean concentrations of both ADOM and MLAM-FINE tended to be lower than the mean measured concentrations, especially at the 1,000-km band, where calculated concentrations were approximately 50% lower. In the second half, both TCAL and VCAL mean concentrations were much higher than the mean measured concentrations. This concurred with the distributional bar charts discussed earlier (Figure 8b). Unlike the mean concentrations of the single-layer TCAL and VCAL, GAMUT mean concentrations closely resembled the mean measured concentrations with the exception of the first-half mean at the 1,000-km band, where its concentration was a factor of 3 too low.

ARL, BAT, and MLAM-COARSE oversmoothed the gradient as well as underestimated the mean concentrations for both halves. This suggests that these models did one or both of the following:

(1) transported the tracer too slowly to allow more dilution before the tracer plume reaches each band, and/or

(2) diffused the tracer too quickly in the horizontal and/or vertical to overstate the dilution of the tracer plume at the surface.

The analyses in subsequent sections will shed more light on the causes of the apparent underestimates of mean concentrations.

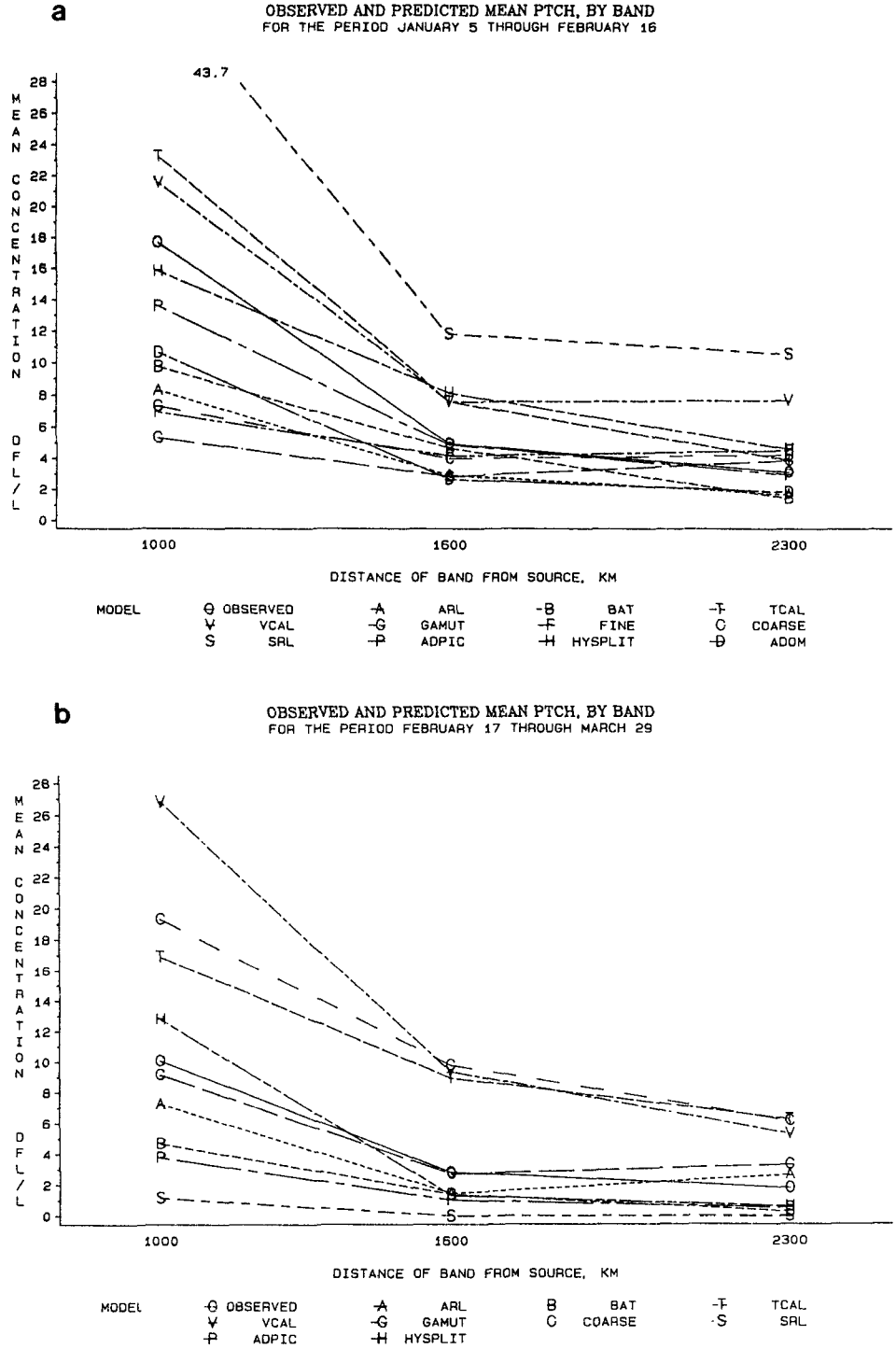


Figure 11. Mean calculated and measured PTCH concentrations along three bands of sites downwind of GGW during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987.

TABLE 5. THE MEAN CALCULATED AND MEASURED TRACER CONCENTRATIONS ALONG BANDS OF SAMPLERS FOR THE TWO HALVES OF THE ANATEX PERIOD

PTCH

Model		Tracer concentration, dfL/L (Ratio of calculated to measured) 1,000 km	1,600 km	2,300 km	Concentration ratios between bands 1,000-1,600 km	1,600-2,300 km
First half: 1/5/87 - 2/16/87						
Measured		17.7	4.9	3.1	3.6	1.6
SLL	SRL	43.7 (2.5)	11.8 (2.4)	10.6 (3.4)	3.7	1.1
	TCAL	23.3 (1.3)	7.6 (1.6)	3.9 (1.3)	3.1	1.9
	VCAL	21.6 (1.2)	7.6 (1.6)	7.7 (2.5)	2.8	1.0
MLL	ARL	8.3 (0.5)	2.9 (0.6)	1.7 (0.5)	2.9	1.7
	BAT	9.8 (0.6)	4.7 (1.0)	1.5 (0.5)	2.1	3.1
	GAMUT	5.3 (0.3)	2.8 (0.6)	3.9 (1.3)	1.9	0.7
	HY-SPLIT	15.8 (0.9)	8.1 (1.7)	4.6 (1.5)	2.0	1.8
	MLAM-FINE	6.9 (0.4)	4.2 (0.9)	4.5 (1.5)	1.6	1.9
	MLAM-COARSE	7.3 (0.4)	4.0 (0.8)	4.3 (1.4)	1.8	0.9
MLE	ADOM	10.7 (0.6)	2.7 (0.6)	1.9 (0.6)	4.0	1.4
	ADPIC	13.6 (0.8)	4.9 (1.0)	3.0 (1.0)	2.8	1.6
Second half: 2/17/87 - 3/29/87						
Measured		10.1	2.8	1.8	3.6	1.6
SLL	SRL	1.2 (0.1)	0.0 (0.0)	0.0 (0.0)	--	--
	TCAL	16.9 (1.7)	8.9 (3.2)	6.2 (3.4)	1.9	1.4
	VCAL	26.8 (2.7)	9.3 (3.3)	5.3 (2.9)	2.9	1.8
MLL	ARL	7.3 (0.7)	1.5 (0.5)	2.6 (1.4)	4.9	0.6
	BAT	4.7 (0.5)	1.4 (0.5)	0.3 (0.2)	3.4	4.7
	GAMUT	9.2 (0.9)	2.7 (1.0)	3.3 (1.8)	3.4	0.8
	HY-SPLIT	12.8 (1.3)	1.3 (0.5)	0.6 (0.3)	9.8	2.2
	MLAM-FINE	--	--	--	--	--
	MLAM-COARSE	19.3 (1.9)	9.8 (3.5)	6.1 (3.4)	2.0	1.6
MLE	ADOM	--	--	--	--	--
	ADPIC	3.8 (0.4)	1.0 (0.4)	0.5 (0.3)	3.8	2.0

(continued)

TABLE 5. (CONCLUDED)

PDCH

Model		Tracer concentration, dfL/L (Ratio of calculated to measured) 300 km		700 km		1,400 km		Concentration ratios between bands 300 - 700 km	700 - 1,400 km
First half: 1/5/87 - 2/16/87									
Measured		8.7		4.2		3.9		2.1	1.1
SLL	SRL	46.1	(5.3)	20.8	(5.0)	17.5	(4.5)	2.2	1.2
	TCAL	13.4	(1.5)	13.7	(3.3)	4.7	(1.2)	1.0	2.9
	VCAL	32.2	(3.7)	15.8	(3.8)	5.9	(1.5)	2.0	2.7
MLL	ARL	9.4	(1.1)	5.2	(1.2)	3.6	(0.9)	1.8	1.4
	BAT	21.7	(2.5)	10.5	(2.5)	2.7	(0.7)	2.1	3.9
	GAMUT	19.1	(2.2)	12.8	(3.0)	5.1	(1.3)	1.5	2.5
	HY-SPLIT	16.1	(1.9)	7.1	(1.7)	4.1	(1.1)	2.3	1.7
	MLAM-FINE	46.9	(5.4)	15.3	(3.6)	13.4	(3.4)	3.1	1.1
	MLAM-COARSE	44.2	(5.1)	16.3	(3.9)	9.8	(2.5)	2.7	1.7
MLE	ADOM	17.0	(2.0)	13.2	(3.1)	1.8	(0.5)	1.3	7.3
	ADPIC	18.2	(2.1)	7.9	(1.9)	4.5	(1.2)	2.3	1.8
Second half: 2/17/87 - 3/29/87									
Measured		15.1		5.0		2.0		3.0	2.5
SLL	SRL	10.6	(0.7)	5.4	(1.1)	8.4	(4.2)	2.0	0.6
	TCAL	12.4	(0.8)	10.3	(2.1)	3.8	(1.9)	1.2	2.7
	VCAL	18.5	(1.2)	14.2	(2.8)	6.4	(3.2)	1.3	2.2
MLL	ARL	8.0	(0.5)	4.7	(0.9)	2.4	(1.2)	1.7	2.0
	BAT	3.6	(0.2)	4.5	(0.9)	1.1	(0.6)	0.8	4.1
	GAMUT	12.2	(0.8)	8.2	(1.6)	3.5	(1.8)	1.5	2.3
	HY-SPLIT	6.5	(0.4)	3.8	(0.8)	2.8	(1.4)	1.7	1.4
	MLAM-FINE	--		--		--		--	--
	MLAM-COARSE	14.7	(1.0)	9.0	(1.8)	5.3	(2.7)	1.6	1.7
MLE	ADOM	--		--		--		--	--
	ADPIC	7.6	(0.5)	3.5	(0.7)	1.9	(1.0)	2.2	1.8

Mean PDCH (STC) Concentrations--

For the first half, the mean measured concentrations (Figure 12a) peaked to nearly 9 dfL/L at 300 km and decreased at a much slower rate than those for PTCH to approximately 4 dfL/L at both 700 km and 1,400 km. In contrast, the second-half mean concentrations (Figure 12b) peaked at 15 dfL/L at 300 km, decreased sharply to 5 dfL/L at 700 km, and then decreased at a slower rate to 2 dfL/L at 1,400 km.

The ARL model replicated these characteristics to the greatest extent, especially during the first half. However, for the second half, like the other models, it oversmoothed the gradient between 300 km and 700 km. In addition, its 300-km mean concentration was about half that of the mean measured concentration.

Decreases of ADPIC, BAT, GAMUT, and HY-SPLIT mean concentrations with distance more or less paralleled those of the mean measurements in the first half, but the model means were considerably higher. All four models oversmoothed the gradient in the second half. However, for both halves, the model mean concentrations at 1,400 km resembled those of the measurements. At the closer bands, the ADOM mean concentrations exceeded the mean measured concentrations by factors of 2 and 3, but at 1,400 km, the ADOM mean resembled that of the measurements.

The MLAM-FINE, MLAM-COARSE, SRL, and VCAL means were much higher than measured means and showed much steeper gradients between 300 km and 700 km for the first half. The relatively high concentrations near STC were also illustrated by the distributional bar charts (Figure 10a). Also during this half, the TCAL mean at 700 km was comparable to its 300-km mean. In the second half, based on the few nonzero concentrations, the SRL mean at 1,400 km was greater than its mean at 700 km. The VCAL and MLAM-COARSE means tended to be higher than the mean measured concentrations, especially at 1,400 km, where calculated concentrations were 2-to-3 times greater.

4.2 COMPARISON OF HORIZONTAL DIFFUSION OF TRACER PUFFS

Given the distribution of samplers in the ANATEX network, horizontal diffusion of a tracer could not be determined precisely. As a means of assessing the differences in ensemble-average diffusion rates of the models and the actual tracer, the average distances of nonzero concentrations calculated by each model along three bands for each half of the ANATEX period were compared to those of the measurements. If a model's average horizontal diffusion rate differed from the actual average rate, the differences should become evident from these comparisons.

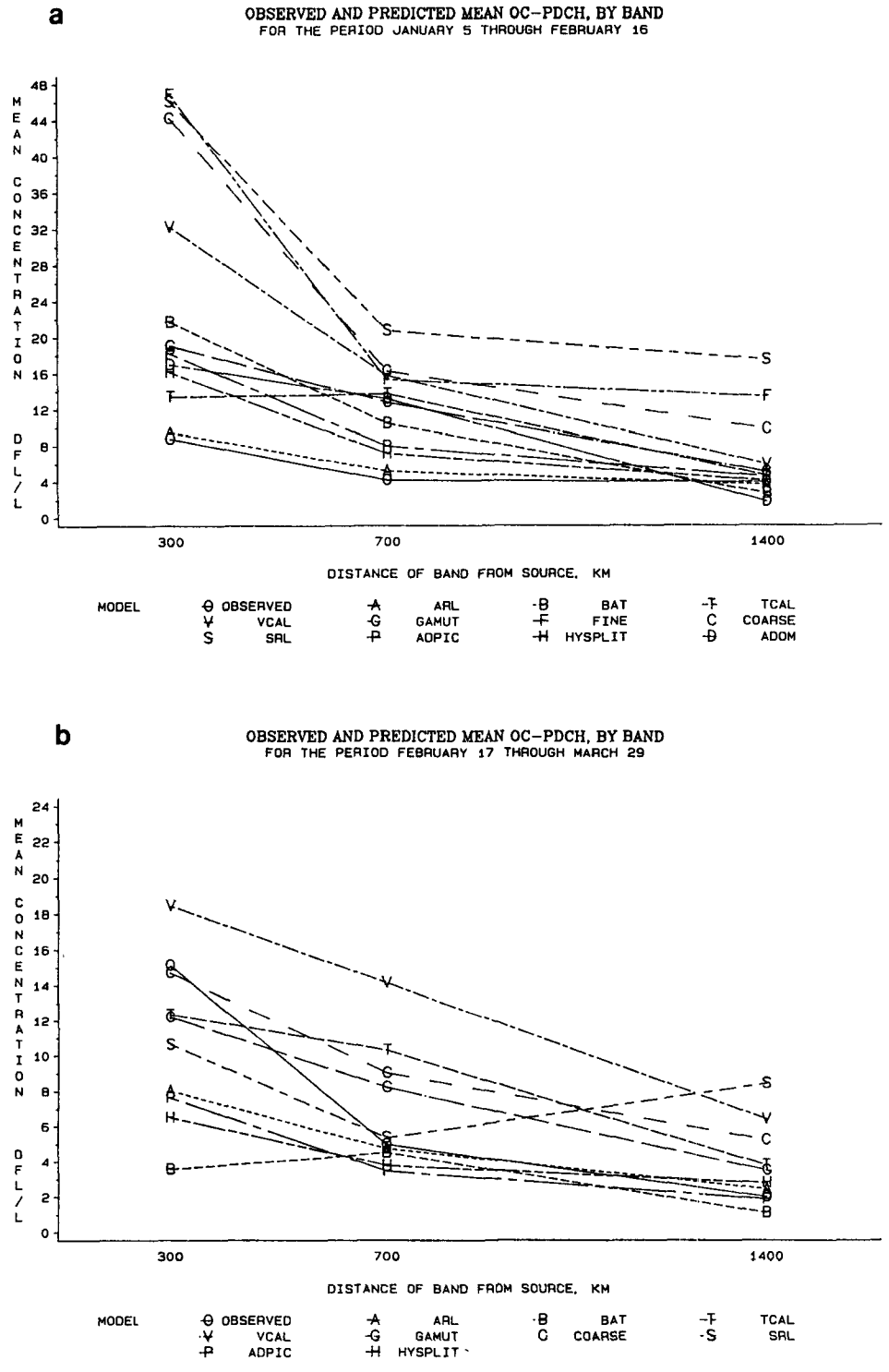


Figure 12. Mean calculated and measured PDCH concentrations along three bands of sites downwind of STC during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987.

PTCH (GGW) Concentrations--

As Figures 13a and 13b illustrate, the mean distance of nonzero measured PTCH concentrations along three bands of sites--approximately 1,000 km to 2,300 km downwind of GGW--increased from 590 km to 1,441 km during the first half and from 546 km to 1,091 km during the second half (Table 6). The mean rate of change relative to the transport distance (i.e., the "mean spread rate") approximated 0.42 km/km for both half periods, with the exception of the 0.83 km/km first-half-period spread rate between 1,600 km and 2,300 km.

Compared to the mean distances for the measured concentrations, the first-period means for the calculated concentrations tended to be lower by 10% to 70%. ADPIC was a noticeable exception; its mean distances were 20% to 60% greater. Mean distances for ADOM, MLAM-FINE, and SRL most closely resembled those for the measurements; all were within 30% for the three bands.

First-period spread rates also tended to be lower than those for the measurements. MLAM-FINE, the one exception, had spread rates of nearly 1.0 km/km. The spread rates for ARL and its sibling, HY-SPLIT, were much lower, approximately 0.15 to 0.25 km/km.

During the second period, the mean distances for TCAL, VCAL, and MLAM-COARSE concentrations tended to be factors of 2.0 to 2.5 greater than those for the measurements. In contrast to the first period, ADPIC mean distances were 50% to 70% lower than the means for the measurements. The mean distances for ARL and GAMUT were within 10% and 30% for those of the measurements, respectively. The extremely low SRL mean distances indicated a serious modeling problem during the second period.

The range of spread rates during the second period varied much more than that of the first period. In fact, some models, particularly BAT and HY-SPLIT, indicated a decrease in the mean distances with increasing transport distances, while ADPIC mean distances changed little. ARL most closely resembled the spread rates of the measurements.

PDCH (STC) Concentrations--

The mean distances of nonzero measured PDCH concentrations (Figures 14a and 14b) ranged from approximately 650 km at 300 km downwind to 1,350 km at 1,400 km downwind of STC during each period (Table 6). Meanwhile, the spread rates ranged from 0.37 km/km to 0.92 km/km, similar to the range for PTCH.

During the first period, the mean distances for the calculated concentrations again tended to be lower than those for the measured concentrations. In addition, the ratios of calculated-to-measured mean distances for PDCH were very similar to those for PTCH. As before, ADOM, MLAM-FINE, and SRL (in

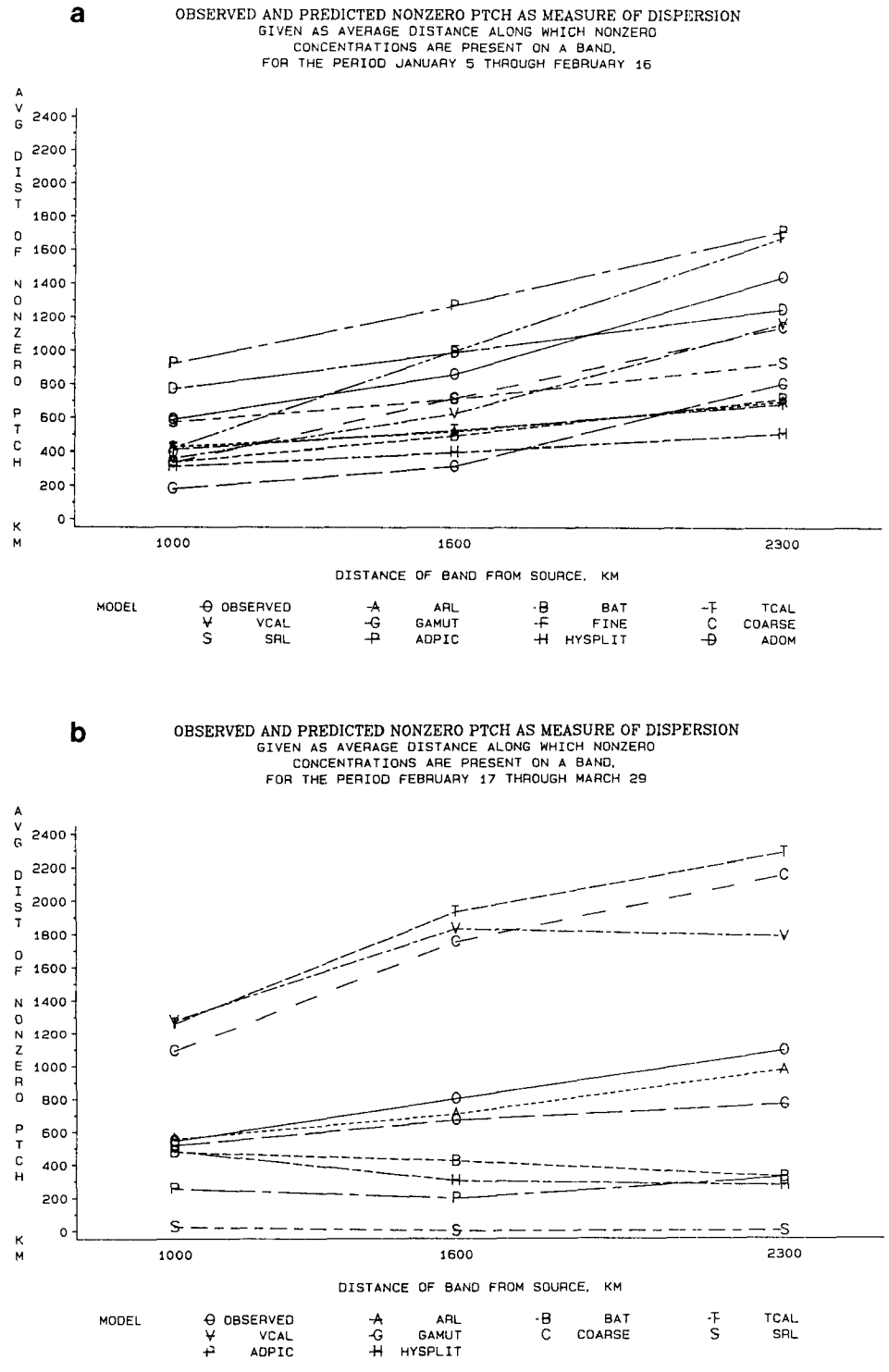


Figure 13. Mean horizontal diffusion of calculated and measured tracer puffs as indicated by the average distance of nonzero PTCH concentrations along three bands of sites downwind of GGW during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987.

TABLE 6. THE MEAN DISTANCE OF NONZERO CALCULATED AND MEASURED CONCENTRATIONS AND MEAN SPREAD RATES ALONG BANDS OF SAMPLERS

PTCH

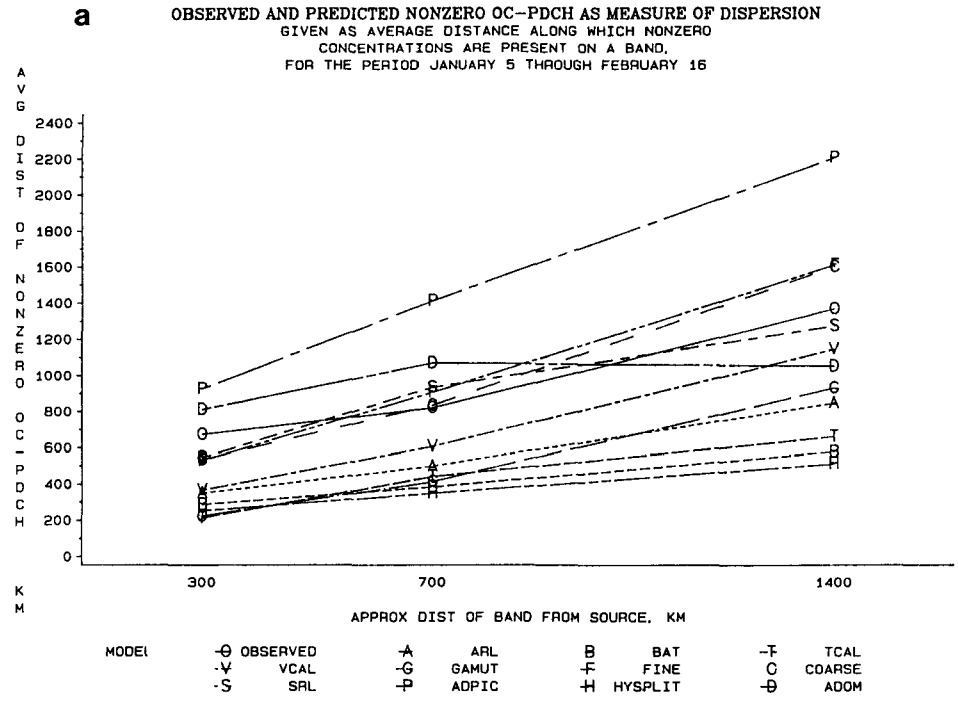
Model		Mean distance of nonzero concentrations, km (Ratio of calculated to measured) 1,000 km		1,600 km		2,300 km		Mean spread rates between bands, km/km 1,000-1,600 km	1,600-2,300 km
First half: 1/5/87 - 2/16/87									
Measured		590		861		1,441		0.45	0.83
SLL	SRL	575	(1.0)	714	(0.8)	930	(0.7)	0.23	0.30
	TCAL	412	(0.7)	527	(0.6)	686	(0.5)	0.19	0.23
	VCAL	363	(0.6)	627	(0.7)	1,167	(0.8)	0.44	0.77
MLL	ARL	427	(0.7)	520	(0.6)	702	(0.5)	0.16	0.26
	BAT	340	(0.6)	496	(0.6)	717	(0.5)	0.26	0.32
	GAMUT	182	(0.3)	317	(0.4)	808	(0.6)	0.23	0.70
	HY-SPLIT	314	(0.5)	401	(0.5)	511	(0.4)	0.15	0.16
	MLAM-FINE	420	(0.7)	999	(1.2)	1,678	(1.2)	0.97	0.97
	MLAM-COARSE	337	(0.6)	718	(0.8)	1,144	(0.8)	0.64	0.61
MLE	ADOM	760	(1.3)	960	(1.1)	1,190	(0.8)	0.33	0.33
	ADPIC	923	(1.6)	1,269	(1.5)	1,716	(1.2)	0.58	0.64
Second half: 2/17/87 - 3/29/87									
Measured		546		800		1,091		0.42	0.42
SLL	SRL	---		---		---		---	---
	TCAL	1,266	(2.3)	1,931	(2.4)	2,281	(2.1)	1.11	0.50
	VCAL	1,279	(2.3)	1,830	(2.3)	1,782	(1.6)	0.92	-0.07
MLL	ARL	559	(1.0)	708	(0.9)	974	(0.9)	0.25	0.38
	BAT	478	(0.9)	423	(0.5)	325	(0.3)	-0.09	-0.14
	GAMUT	520	(1.0)	671	(0.8)	766	(0.7)	0.25	0.14
	HY-SPLIT	482	(0.9)	303	(0.4)	275	(0.3)	-0.03	-0.04
	MLAM-FINE	---		---		---		---	---
	MLAM-COARSE	1,092	(2.0)	1,752	(2.2)	2,148	(2.0)	1.10	0.57
MLE	ADOM	---		---		---		---	---
	ADPIC	256	(0.5)	198	(0.3)	325	(0.3)	-0.10	0.18

(continued)

TABLE 6. (CONCLUDED)

PDCH

Model		Mean distance of nonzero concentrations, km (Ratio of calculated to measured) 300 km		700 km		1,400 km		Mean spread rates between bands, km/km 300 - 700 km	700 - 1,400 km
First half: 1/5/87 - 2/16/87									
Measured		674		820		1,368		0.37	0.78
SLL	SRL	547	(0.8)	934	(1.1)	1,275	(0.9)	1.00	0.49
	TCAL	213	(0.3)	442	(0.5)	661	(0.5)	0.57	0.31
	VCAL	369	(0.6)	606	(0.7)	1,144	(0.8)	0.59	0.77
MLL	ARL	351	(0.5)	499	(0.6)	847	(0.6)	0.37	0.50
	BAT	288	(0.4)	385	(0.5)	577	(0.4)	0.24	0.27
	GAMUT	225	(0.3)	414	(0.5)	930	(0.7)	0.47	0.74
	HY-SPLIT	253	(0.4)	349	(0.4)	512	(0.4)	0.24	0.23
	MLAM-FINE	530	(0.8)	906	(1.1)	1,610	(1.2)	0.94	1.01
	MLAM-COARSE	536	(0.8)	834	(1.0)	1,600	(1.2)	0.75	1.09
MLE	ADOM	805	(1.2)	1,050	(1.3)	1,015	(0.7)	0.61	-0.05
	ADPIC	927	(1.4)	1,412	(1.7)	2,205	(1.6)	1.21	1.13
Second half: 2/17/87 - 3/29/87									
Measured		658		1,026		1,349		0.92	0.46
SLL	SRL	139	(0.2)	230	(0.2)	297	(0.2)	0.23	0.01
	TCAL	700	(1.1)	1,093	(1.1)	1,816	(1.3)	0.98	1.03
	VCAL	1,086	(1.7)	1,668	(1.6)	2,262	(1.7)	1.46	0.85
MLL	ARL	531	(0.8)	815	(0.8)	1,147	(0.9)	0.71	0.47
	BAT	266	(0.4)	441	(0.4)	382	(0.3)	0.44	-0.08
	GAMUT	676	(1.0)	978	(1.0)	1,147	(0.9)	0.76	0.24
	HY-SPLIT	428	(0.7)	585	(0.6)	796	(0.6)	0.39	0.30
	MLAM-FINE	---		---		---		---	---
	MLAM-COARSE	1,261	(1.9)	1,841	(1.8)	1,805	(1.3)	1.45	-0.05
MLE	ADOM	---		---		---		---	--
	ADPIC	676	(1.0)	882	(0.9)	1,179	(0.9)	0.52	0.42



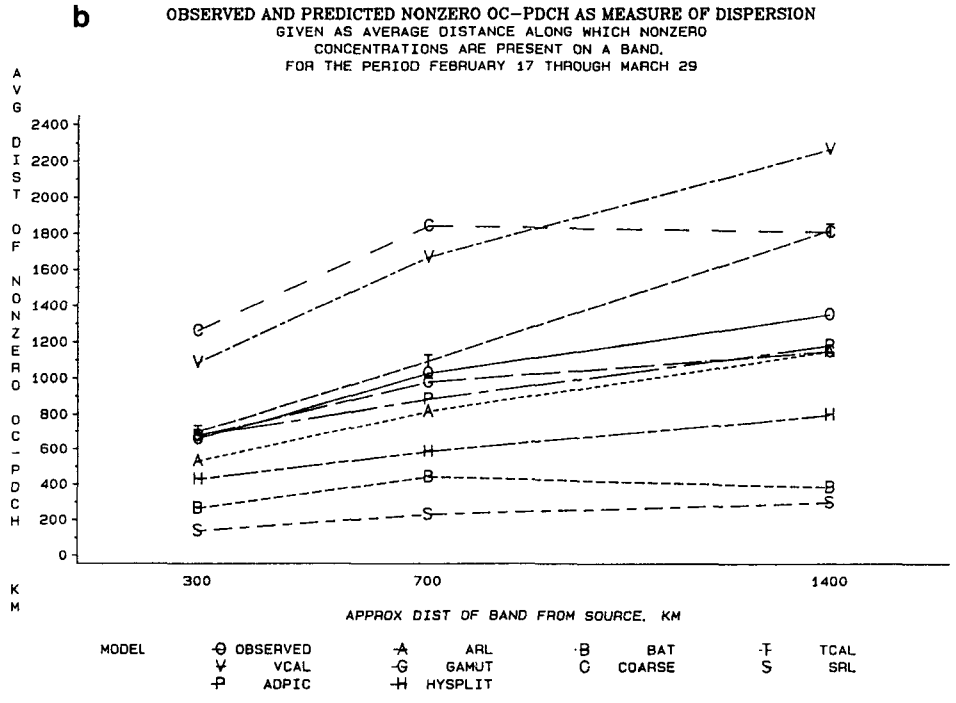


Figure 14. Mean horizontal diffusion of calculated and measured tracer puffs as indicated by the average distance of nonzero PDCH concentrations along three bands of sites downwind of STC during the periods: (a) 5 January - 16 February 1987 and (b) 17 February - 29 March 1987.

addition to MLAM-COARSE) mean distances were within ±30% of those for the measurements, while ADPIC mean distances were 40% to 70% greater. BAT, HY-SPLIT, and TCAL mean distances were consistently lower by 50% to 70%.

Although first-period mean distances for ARL, GAMUT, and VCAL were lower than those for the measurements, their spread rates (0.37 km/km to 0.77 km/km) most closely resembled those for the measurements. BAT and HY-SPLIT spread rates (approximately 0.25 km/km) were the lowest, while ADPIC, MLAM-FINE, and MLAM-COARSE spread rates were the greatest (approximately 0.75 km/km to 1.20 km/km). ADOM, SRL, and TCAL spread rates were lower between the 300 km and 700 km bands, but greater between the 700 km and 1,400 km bands.

Similar to the second-period mean distances for PTCH calculations, the mean distances for VCAL and MLAM-COARSE PDCH calculations were generally 60% to 90% greater than those for the measurements while SRL mean distances were much (80%) lower. ARL and GAMUT mean distances again best replicated those for the measurements (within ±20% at all three bands). For PDCH, TCAL mean distances were within ±30% of those for the measurements. As with the first period, second-period mean distances for BAT were low (by 60% to 70%) and for HY-SPLIT (by 30% to 40%). In contrast to those of the first period and the two PTCH periods, ADPIC mean distances for PDCH in the second period were within ±10% of those for the measurements.

As before, the mean spread rates for ARL and GAMUT (ranging from 0.24 km/km to 0.76 km/km) most closely resembled those for the measurements. ADPIC mean spread rates (approximately 0.50 km/km) also resembled those of the measurements. BAT and HY-SPLIT mean spread rates again were low (ranging from near 0.00 km/km to 0.44 km/km), as were those for SRL (<0.25 km/km). TCAL and VCAL mean spread rates were greatest (0.85 km/km to 1.46 km/km). MLAM-COARSE spread rates were much greater between the 300 km and 700 km bands, but much lower between the 700 km and 1,400 km bands.

SECTION 5

ASSESSMENT OF THE MODEL PERFORMANCES BASED ON COMPARISONS OF INDIVIDUAL TRACER RELEASES

5.1 PERFORMANCE AS RELATED TO METEOROLOGICAL SCENARIO

From a set of discrete 24-h tracer footprints (i.e., those that did not merge with the footprints of other releases of the same tracer), 24-h mean transport speeds and 24-h locations of the footprint centroids were determined. The speeds and locations of all footprints common to both the sets of measurements and the calculations for each model were compared to assess the accuracy of calculated transport speeds and directions. Table 7 identifies the number of 24-h footprints common to both model and measured sets for each tracer release. The relatively small number of ADPIC footprints results from the fact that this model often showed long residence times of tracer across the grid, making it difficult to distinguish discrete footprints. The time-series plots in Appendix B of the 24-h mean calculated and measured concentrations at various transport distances, often demonstrate this unique behavior.

Also, the relationship between model performance and meteorological scenario was investigated. This section discusses the results of these comparisons and the behavior of the models with respect to the meteorological scenarios. Section 5.2 discusses the mean model location errors for ANATEX and compares the range of errors with that for CAPTEX. Section 5.3 summarizes ANATEX performances by model genres and assesses the performances of individual models.

Three types of meteorological scenarios were considered. Each type was based simply on the timing of cyclonic and/or frontal interceptions of the measured footprints, if any, relative to the transport period (Table 8). The purpose was to distinguish the simpler flow patterns from the more complicated flow patterns. That is, cyclonic and frontal interceptions usually introduce complexities in the horizontal and vertical wind flow patterns that may accelerate vertical mixing and horizontal diffusion. Moreover, tracer looping in the horizontal plane can result from these phenomena.

The approach of relating model performances to meteorological scenarios was quite simple. Firstly, each footprint-day beyond 10 h of transport was labeled as either being or not being intercepted by a cyclone and/or front since the tracer release ("*I*" and "*NI*", respectively) [Table 9]. Model performances

TABLE 7. NUMBER OF CALCULATED AND MEASURED FOOTPRINT PAIRS AND THE TYPE OF METEOROLOGICAL SCENARIO FOR EACH TRACER RELEASE[a]

		PTCH										
	Type of	SLL[b]			MLL						MLE	
Release number	met. scenario	SRL	TCAL	VCAL	ARL	BAT	GAMUT	HY-SPLIT	MLAM FINE	MLAM-COARSE	ADOM	ADPIC
1	I		3	3	2	2	2	1	2	2	2	
2	I											
3	II	3	1	3	1	3	1		3	3		
4	III		5	2	3	5	5	5	3	3	2	3
5	II	2	2	2	3	2	2	2	2	2	3	2
6	III	1	1		1	1	1	1			1	1
7	III	5	2	2	2	2	1	5			1	
8	I	1	1		1	1			1			
9	II											
10	III		3	2	4	2		1		2		1
11	III		2	2	2		2		1		2	
12	III		1	1	1	1	1		1	1		
13	II		2	3	2	3	3	3	3	3		1
14	II											
15	II		1	1	1	1		1	1	1	1	
16	II											
17	I											
18	I											
19	III											
20	I											
21	I											
22	I				2		2					
23	II											
24	III											
25	III		2		2		1					
26	I		5		3							
27	III											
28	I											
29	II											
30	III				2	2	1	2		1		
31	I											
32	II											
33	III						1					
Total		12	31	21	32	25	23	21	17	18	12	8

(continued)

TABLE 7. (CONCLUDED)

PDCH												
	Type of	SLL[b]			MLL						MLE	
Release number	met. scenario	SRL	TCAL	VCAL	ARL	BAT	GAMUT	HY-SPLIT	MLAM FINE	MLAM-COARSE	ADOM	ADPIC
1	I	3	4	5	6	5	5	5	5	5	4	2
2	III	2	1	2	1	2	2	1	2	2	2	
3	I											
4	III	3	3	2	5	2	5	5	2	3	2	1
5	II		2		2		2	2			1	
6	III	2	3	4		4		2	2			2
7	II		2	2	2		2					
8	II	2	2	5	3	3	5	3	5	2	2	3
9	II		4	4	4	2	1	2	2	3		2
10	III	1	2	2	2	2	2	2	2	1	2	
11	II											
12	III		1		1				1			
13	I		2	3	2	2	2	2	2	2		
14	I		3	2		2	1	1	1			
15	III	2	3	2	3	3	3	1	3	2	2	1
16	III											
17	I			1	1	1	1	1		1	1	
18	I	1	2	2	2	2	1	2		1		2
19	III											
20	I	1	2	2	2	2	2	1		2		
21	II							1				
22	III	1										
23	I											
24	III	3		2	2	1	3					3
25	III	1										
26	I		1	2	2	2	1	3		2		
27	III	2		1		1		1				
28	I											
29	I											
30	II											2
31	III											1
32	III											
33	III											
Total		24	37	43	40	36	38	35	27	26	16	19

a. See Table 8 for descriptions of the types of meteorological scenarios.

b. SLL = Single-Layer Lagrangian; MLL = Multiple-Layer Lagrangian; MLE = Multiple-Layer Eulerian.

TABLE 8. THE THREE TYPES OF METEOROLOGICAL SCENARIOS FOR WHICH MODEL PERFORMANCE WAS ASSESSED

Type of met. scenario	Meteorological scenario with respect to actual footprints		PTCH release numbers	PDCH release numbers
I	Actual footprint is in or near surface pressure ridge and is not intercepted by a cyclone or front	First half:	1, 2, 8	1, 3, 13, 14
		Second half:	17, 18, 20, 21, 22, 26, 28, 31	17, 18, 20, 23, 26, 28, 29
II	Actual footprint is intercepted by a cyclone or front only after 48 h of transport	First half:	3, 5, 9, 13, 14, 15, 16	5, 7, 8, 9 11
		Second half:	23, 29, 32	21, 30
III	Actual footprint is intercepted by a cyclone or front within the first 48 h of transport	First half:	4, 6, 7, 10, 11, 12	2, 4, 6, 10, 12, 15, 16
		Second half:	19, 24, 25, 27, 30, 33	19, 22, 24, 25, 27, 31, 32, 33

TABLE 9. NUMBER OF FOOTPRINT-DAYS INTERCEPTED (*I*) AND NOT INTERCEPTED (*NI*) BY CYCLONES AND/OR FRONTS AFTER 10-h OF TRANSPORT

Model[a]		PDCH		PTCH		Total
		I	*NI*	*I*	*NI*	*I*, %
SLL:	SRL	11	6	8	2	70
	TCAL	13	16	17	11	53
	VCAL	16	18	13	7	54
MLL:	ARL	15	19	19	11	53
	BAT	13	16	15	7	55
	GAMUT	15	16	13	8	54
	HY-SPLIT	11	18	15	4	54
	MLAM-FINE	11	9	9	7	56
	MLAM-COARSE	7	14	12	5	50
MLE:	ADOM	5	7	6	4	50
	ADPIC	8	5	4	2	63

a. SLL = Single-Layer Lagrangian; MLL = Multiple-Layer Lagrangian; MLE = Multiple-Layer Eulerian.

were related to meteorological scenarios only for transport times beyond 10 h, since initial performances were often dissimilar to performances beyond this time. Daily surface weather maps and tracer concentration maps were superimposed to determine interception occurrences. Secondly, the footprint-days were identified for which transport speed errors and centroid location errors were least and greatest. Finally, the number of *I* and *NI* cases were compared for the least and greatest errors to establish the presence or absence of a relationship for all but two models (Table 10). SRL performances and meteorological scenarios were not related for SRL, since an inordinate percentage (70%) of its footprint-days were intercepted by cyclones and/or fronts, and ADPIC, since there were only 19 footprint-days while a high percentage of these (63%) were intercepted by cyclones and/or fronts. For the other nine models, the number of *I* cases was comparable to the number of *NI* cases (50% to 56% of the cases were *I* cases).

For four of the models (TCAL, ARL, BAT, and HY-SPLIT), the greatest errors in transport speed and centroid location tended to correspond to cyclonic and/or frontal interceptions. The same is true for VCAL and ADOM, but for these two models, the least errors also tended to correspond to cyclonic and/or frontal interceptions. Meanwhile, the least MLAM-FINE errors tended to correspond to cyclonic and/or frontal interceptions.

The model performance assessments discussed in this section rely primarily on three plots. The first plot compares mean separation errors (i.e., the mean distance between the centroids of the measured and calculated footprints) as a function of transport time. The second plot compares model and actual transport speeds by displaying for each case the ratio of calculated-to-measured transport speeds (D_c / D_m) as a function of transport time. Finally, the last plot displays the model error in footprint centroid locations via the ratio of separation distance over measured transport distance (D_e / D_m) also as a function of transport time.

A word of caution when interpreting the transport speed comparison plots: model errors are not necessarily greatest for the lowest ratios. It is possible for the calculated centroids to be very close to the measured centroids, yet at the same time, the transport speed ratios are <0.5 or >2.0. In some cases when both the model and actual wind speeds are low, the ratios can be very small. For these cases, one must also view the corresponding cases on the plots of footprint centroid separation.

Since each model has a unique set of discrete footprints, model performances can be intercompared only if the sample sizes are large enough to reveal a true assessment of model performance. That is, a comparison of the performances of two models may not always be appropriate, since the two models could perform equally well for the same set of footprints (if one were available).

TABLE 10. RELATIONSHIP BETWEEN MODEL PERFORMANCE AND METEOROLOGICAL SCENARIO

		Number of footprints							
		Transport speed errors				Centroid location errors			
		Least		Greatest		Least		Greatest	
Model[a]		*I*[b]	*NI*[b]	*I*	*NI*	*I*	*NI*	*I*	*NI*
SLL:	SRL	--	--	--	--	--	--	--	--
	TCAL[c]	3	4	7	5	6	6	7	2
	VCAL	9	6	4	1	10	4	3	2
MLL:	ARL[c]	5	8	9	3	5	5	7	4
	BAT[c]	8	6	7	1	7	7	5	2
	GAMUT	5	5	5	4	7	5	2	2
	HY-SPLIT[c]	8	5	8	2	6	7	8	2
	MLAM-FINE[d]	7	3	2	2	10	4	0	0
	MLAM-COARSE	9	5	2	1	9	7	2	0
MLE:	ADOM	3	2	6	2	6	3	4	1
	ADPIC	--	--	--	--	--	--	--	--

a. SLL = Single-Layer Lagrangian; MLL = Multiple-Layer Lagrangian; MLE = Multiple-Layer Eulerian.

b. I ≡ footprint intercepted by a cyclone and/or front since tracer release;
NI ≡ footprint not intercepted by a cyclone and/or front since tracer release.

c. Greatest errors appeared to be related to the occurrences of cyclonic and/or frontal interceptions.

d. Least errors appeared to be related to the occurrences of cyclonic and/or frontal interceptions.

5.2 MEAN SEPARATION ERROR

For each of the 11 models, the mean distance separating C_c and C_m for the combined set of PTCH and PDCH footprints was calculated as a function of transport time. Ideally, for a model intercomparison, the set of footprints for each model should be identical so that the performance of each model is assessed with a standard data set. However, due to the inconsistencies of the sets of footprints between models (Table 11), a standard set would not be useful: the number of footprints common to each model would be too small to draw statistically meaningful conclusions about model performance intercomparisons.

There are enough footprints within the first 61.5 h of transport (32-60 for the Lagrangian models and 27 for the Eulerian models) to characterize the mean model error (Figure 15), while minimizing the possibility that differences between model performances were due to the differences in footprint sets. The number of cases for each model expressed as a function of transport time (Table 12), can be used as a qualitative confidence factor when interpreting Figure 15; that is, the greater the number of cases, the greater the probability that the results presented in Figure 15 are representative. A precautionary note: the ADOM mean error decreases with increasing transport time from 37.5 h to 61.5 h, probably reflects the small sample size at 61.5 h (i.e., 6) and not the true behavior of the model.

There appears to be two clusters of model behavior in Figure 15. The first cluster of models--BAT, MLAM-COARSE, MLAM-FINE, and VCAL--show relatively low mean errors that gradually increase with increasing transport time within the first 85.5 h of transport. These mean errors range from 100 km to 400 km. The second cluster of models--ADOM, ADPIC, ARL, GAMUT, HY-SPLIT, SRL, and TCAL--show higher mean errors than the other four models. These mean errors also increase at a higher rate as a function of transport time, from 180 km to 380 km after 13.5 h of transport to 750 km to 1,080 km after 85.5 h of transport.

TABLE 11. PERCENT OF MODEL A FOOTPRINT-DAYS COMMON TO BOTH MODELS A AND B[a]

Model B	Model A: SLL[b] SRL	SLL TCAL	SLL VCAL	MLL ARL	MLL BAT	MLL GAMUT	MLL HY-SPLIT	MLL MLAM-FINE	MLL MLAM-COARSE	MLE ADOM	MLE ADPIC
SRL	--	64	67	64	75	64	67	56	56	44	42
TCAL	34	--	72	82	71	66	60	53	53	37	29
VCAL	41	77	--	77	83	70	63	63	64	39	33
ARL	32	79	68	--	67	72	61	49	56	38	33
BAT	44	79	87	79	--	72	74	66	69	39	36
GAMUT	38	74	74	85	72	--	67	59	59	43	31
HY-S	45	73	68	79	82	73	--	55	61	39	38
MLAM-F	45	82	91	80	91	82	70	--	75	55	39
MLAM-C	45	82	93	91	95	82	72	75	--	48	36
ADOM	57	89	89	96	86	93	79	86	75	--	39
ADPIC	56	74	78	78	81	70	78	63	59	41	--

a. Range = 29% - 96%; mean = 64%.

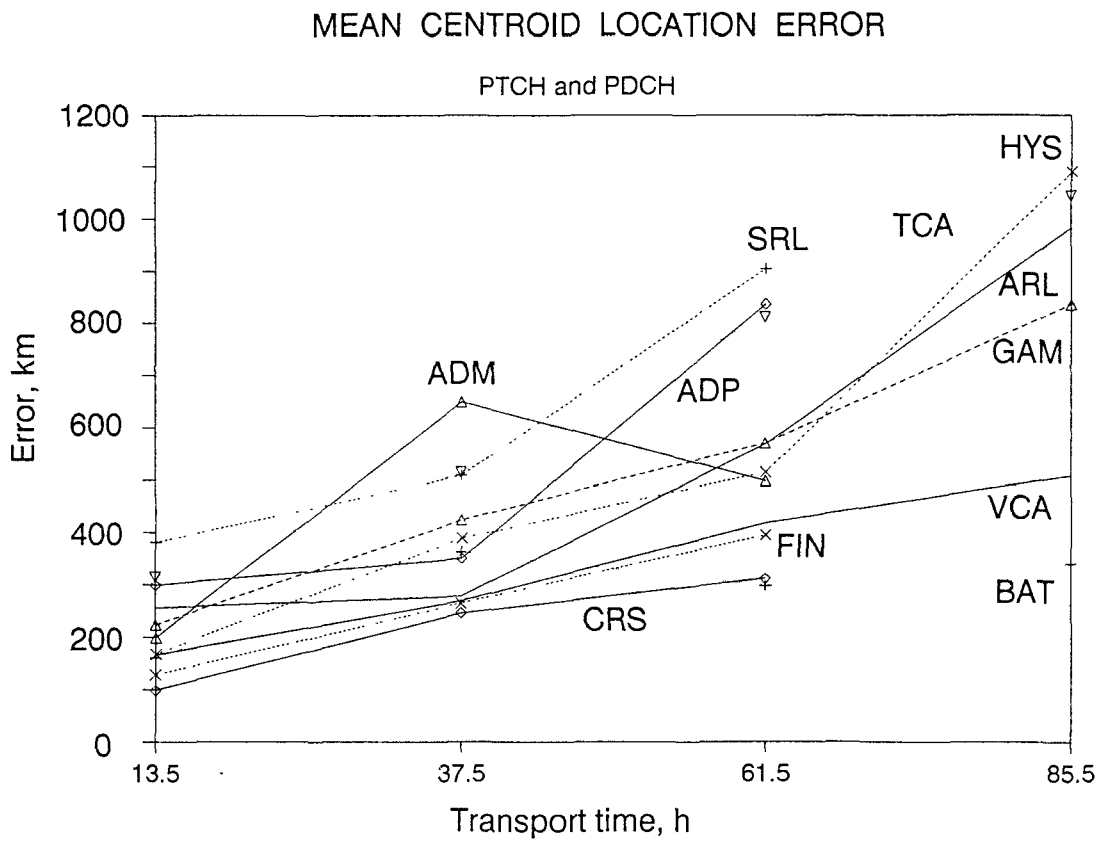
b. SLL = Single-Layer Lagrangian; MLL = Multiple-Layer Lagrangian; MLE = Multiple-Layer Eulerian.

TABLE 12. NUMBER OF CALCULATED AND MEASURED FOOTPRINT PAIRS FOR EACH MODEL AS A FUNCTION OF TRANSPORT TIME[a]

Model		Total number	Transport time, h: 13.5	37.5	61.5	85.5	109.5	133.5
SLL:[b]	SRL	36	12	14	6	2	1	1
	TCAL	68	14	23	20	6	3	2
	VCAL	64	14	26	14	6	4	
MLL:	ARL	72	14	26	20	6	5	1
	BAT	61	12	25	16	6	2	
	GAMUT	61	13	24	15	5	4	
	HY-SPLIT	56	13	19	14	5	4	1
	MLAM-FINE	44	10	16	12	4	2	
	MLAM-COARSE	44	8	17	15	3	1	
MLE:	ADOM	28	9	12	6	1		
	ADPIC	27	11	9	7			

a. Number of pairs for 6 h preceding and following designated transport times.

b. SLL = Single-Layer Lagrangian; MLL = Multiple-Layer Lagrangian; MLE = Multiple-Layer Eulerian.



Note: The number of footprint pairs decreases sharply beyond 61.5 h (see Table 12); plot includes those errors with at least 5 cases.

Figure 15. Comparison of the mean separation errors of PTCH and PDCH footprint centroids ($\overline{D}_e$), as a function of transport time.

Two of the three single-layer models--SRL and TCAL--showed the greatest mean errors. However, the mean errors of the third single-layer model--VCAL--were considerably less. For example, the VCAL mean error after 85.5 h of transport approximated the mean errors of SRL and TCAL after only 13.5 h of transport. Since the TCAL and VCAL codes are very similar to each other, the magnitude of these differences was unexpected. These differences can be attributed to two possible causes. First, VCAL uses a varying mixing depth based on potential temperature profiles, while TCAL uses a fixed mixing depth of 1,500 m. The VCAL approach seems more reasonable, especially since the TCAL mixing depth of 1,500 m is very high for winter applications. Holzworth (1972) empirically determined that typical midday winter mixing depths across the central United States range from about 500 m in January to 900 m in March. The VCAL approach would more likely reflect these representative values. Secondly, the differences in the results may be attributed, in part, to the dissimilarity of the two samples of footprints. However, as shown in Table 7, the two sets are quite similar with respect to the release numbers.

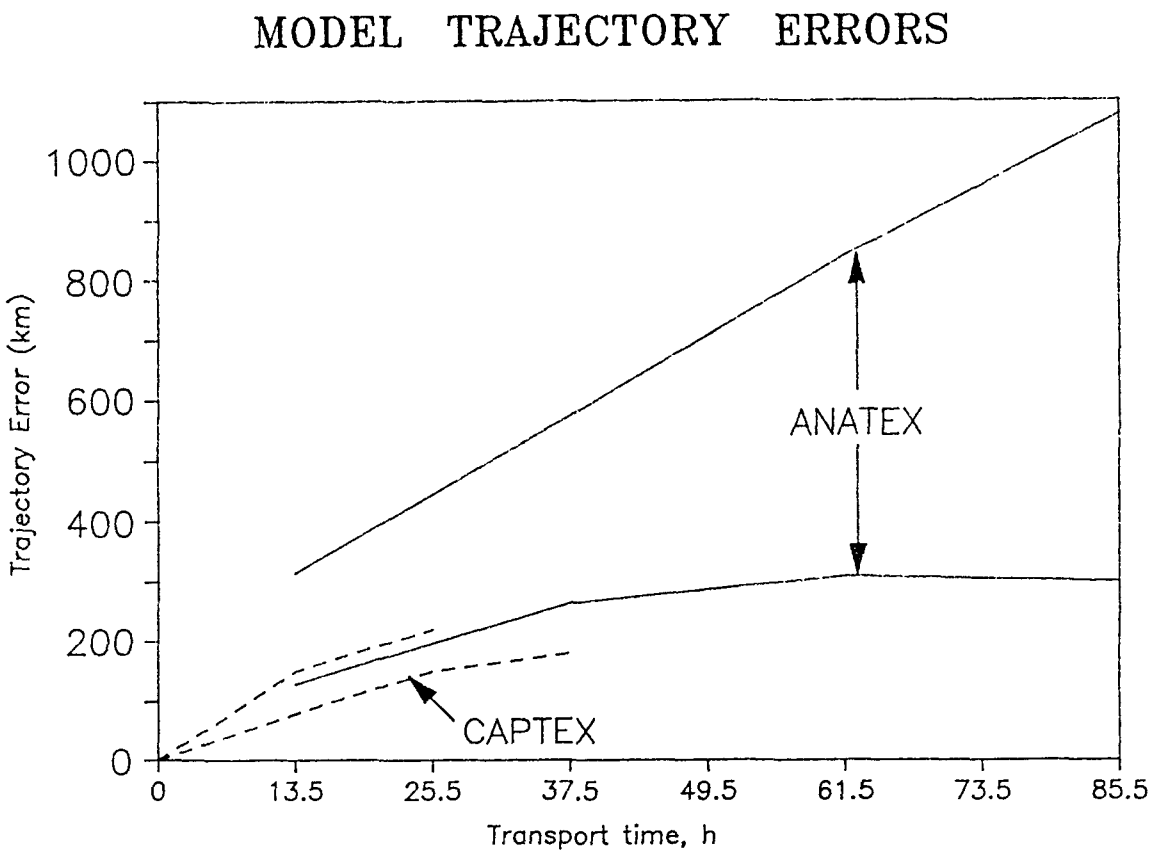
The ARL and HY-SPLIT sibling models produce similar results, despite the HY-SPLIT model's use of the higher resolution Nested Grid Model (NGM) meteorological data. One possible explanation for this is the rather turbulent K_z profiles used by HY-SPLIT, which could exaggerate vertical diffusion (Draxler, personal communication, 1989). Likewise, the MLAM-COARSE and MLAM-FINE family produced similar results, indicating that the higher frequency of puff mass redistribution in the vertical (and hence, additional computational time) did not improve the model performance.

The average trajectory errors of the two MLE models--ADOM and ADPIC-- resembled those of the MLL models at 13.5 h, but deviated significantly (i.e., increased sharply) from them either at 37.5 h (ADOM) or 61.5 h (ADPIC). Since the number of footprint pairs exceeded six for both models at these times, the sudden deviation did not appear to be caused by a very small sample size. However, more data are required to substantiate this. Based on the results illustrated in Figure 15 and discussed in Sections 4 and 5.3, the MLE models appeared to excel in replicating the frequency distribution of the ensemble concentrations, but did not perform as well as the MLL in trajectory accuracy. This seems to indicate that, in general, the MLE simulated the diffusion process better than the transport process.

Compared to the range of model trajectory errors during CAPTEX, as reported in the literature (Kahl and Samson 1988; Haagenson et al. 1987; Draxler 1987; and Godowitch 1989), the range of model trajectory errors during ANATEX was not only greater, but the magnitude of the errors tended to be greater also (Figure 16). The simpler flow patterns of CAPTEX were a definite factor.

5.3 FOOTPRINT TRANSPORT SPEEDS AND CENTROID LOCATIONS

The model performances based on footprint transport speeds and locations are discussed in two formats. First, in Section 5.3.1 the overall performances are summarized for the combined sets of PTCH and PDCH tracer footprints. Secondly, the performances for individual footprints are discussed separately for each model in Section 5.3.2 through Section 5.3.4. These sections also compare the performances of the three model genres: the single-layer Lagrangian (SLL), the multiple-layer Lagrangian (MLL), and the multiple-layer Eulerian (MLE) models, as well as compare the performances of sibling models.



Note: The single least and greatest mean ANATEX errors are ignored for each transport time interval.

Figure 16. The comparison of trajectory error ranges determined from ANATEX and CAPTEX model evaluations.

5.3.1 Summary of Model Performances

Regarding transport speeds, the MLL models as a group performed slightly better than the SLL and the MLE models (Table 13 and Figure 16a-d). With one exception (MLAM-FINE), the transport speeds of approximately half of the MLL footprints were within ±20% of the speeds of the measured footprints. In contrast, the transport speeds of only one-third of the footprints for all but one of the SLL models (VCAL) and both MLE models were within ±20% of the measured speeds. Likewise, the MLL models tended more often to calculate transport speeds within ±50% of the measured speeds (64% to 89% for the MLL models versus 61% to 78% and 68% to 70% for the SLL and MLE models, respectively).

Although the range in the MLL model performances in calculating footprint centroid locations increased over that for transport speeds, the MLL models again tended to perform slightly better than the others. With increasing transport time, no consistent trend in the performances appeared. That is, performances did not systematically improve or worsen as a function of transport time. However, the calculated centroids generally were within ±20% x D_m more frequently for $T \leq 24$ h than for $T > 48$ h.

The following comments highlight the results and conclusions of Section 5.3:

(1) The model performances in calculating the location of the centroids paralleled those for the transport speeds (Table 13 and Figures 17a-d).

(2) Compared to the range of model trajectory errors during CAPTEX, as reported in the literature (Kahl and Samson 1988; Haagenson et al. 1987; Draxler 1987; and Godowitch 1989), the range of model trajectory errors during ANATEX was not only greater, but the magnitude of the errors tended to be greater also (Figure 16). The simpler flow patterns of CAPTEX were a definite factor.

(3) The MLL models as a group tended to perform as well or better than the SLL and MLE models (Figure 15).

(4) The performances were worst for two of the three SLL models--SRL and TCAL; both use simple approximations to the transport wind (SRL uses surface pressure fields to calculate geostrophic wind vectors, while the TCAL model calculates mean-wind vectors for a constant mixing depth of 1,500 m).

(5) The VCAL performance paralleled those of many of the MLL models, while the MLAM-FINE performance paralleled those of SRL and TCAL--SLL models.

TABLE 13. A SUMMARY OF MODEL PERFORMANCES IN CALCULATING TRANSPORT SPEEDS AND FOOTPRINT CENTROID LOCATIONS

A. Percentage of *N* footprint-days when calculated transport speed was within ±20% and ±50% of the measured transport speed

	SLL[a]			MLL						MLE	
	SRL	TCAL	VCAL	ARL	BAT	GAMUT	HY-SPLIT	MLAM-FINE	MLAM-COARSE	ADOM	ADPIC
N=	36	68	64	72	61	61	56	44[b]	44	28[b]	27
±20%:	36	35	55	47	57	46	46	35	59	32	30
±50%:	61	62	78	75	80	72	68	82	89	68	70

B. Percentage of *N* footprint-days when calculated centroids were within ±20% x D_m and ±50% x D_m of measured centroids as a function of transport time (T)

	SLL[a]			MLL						MLE	
	SRL	TCAL	VCAL	ARL	BAT	GAMUT	HY-SPLIT	MLAM-FINE	MLAM-COARSE	ADOM	ADPIC
For $T \leq 24$ h,											
N=	12	14	14	14	12	13	13	10[b]	8	9[b]	11
±20%:	25	14	50	7	50	54	54	30	50	44	27
±50%:	42	21	57	64	67	69	77	50	88	56	64
For 24 h $< T \leq$ 48 h,											
N=	14	23	26	26	25	24	19	16[b]	17	12[b]	9
±20%:	21	17	35	31	32	17	32	19	41	0	22
±50%:	43	43	58	77	52	54	42	81	88	42	44
For $T >$ 48 h,											
N=	10	31	24	32	24	24	24	18[b]	19	7[b]	7
±20%:	0	13	21	25	42	21	13	39	42	14	0
±50%:	30	52	79	53	88	58	58	89	95	71	71

a. SLL = Single-Layer Lagrangian; MLL = Multiple-Layer Lagrangian; MLE = Multiple-Layer Eulerian.

b. Through February 16, 1987, only.

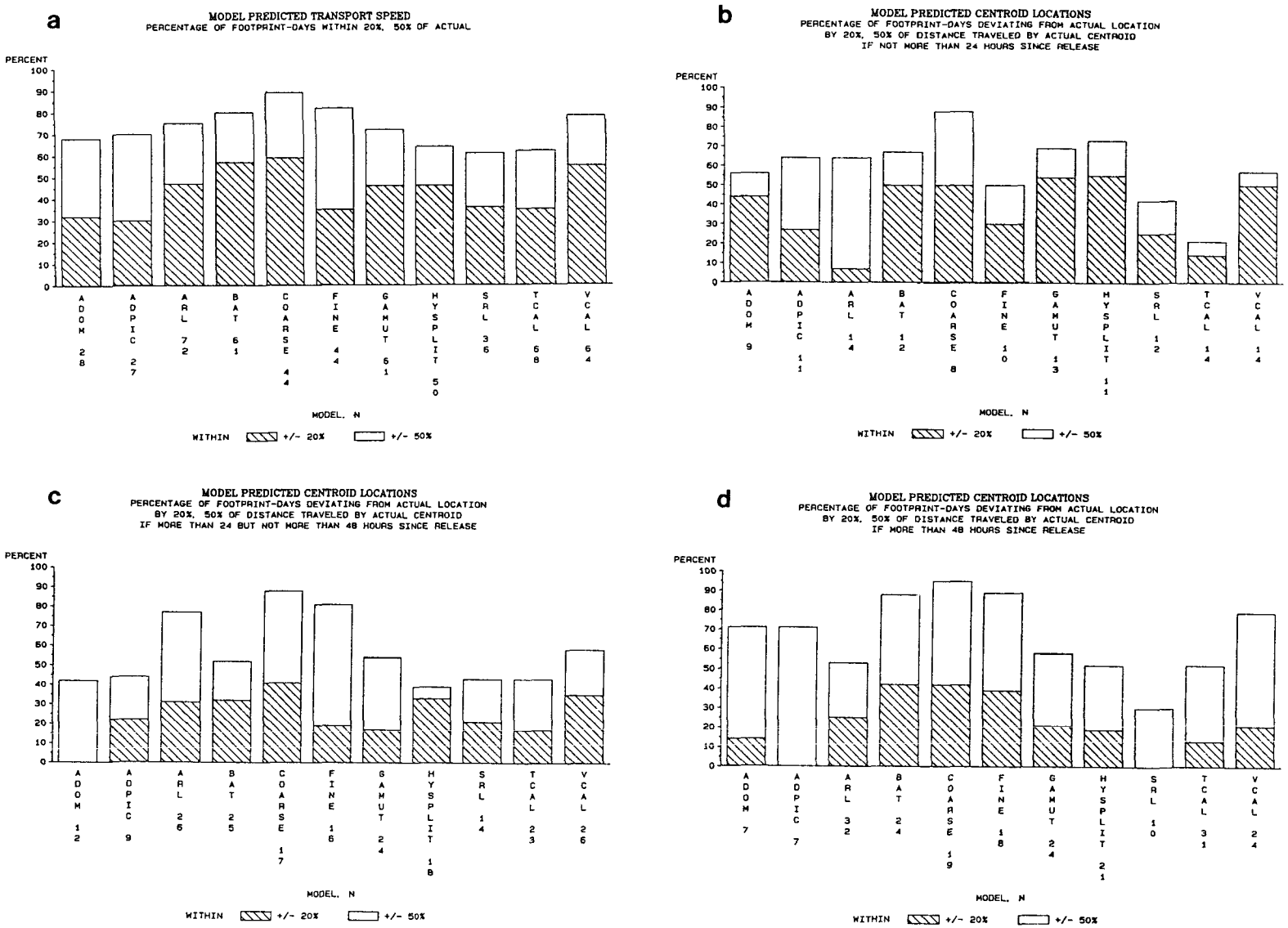


Figure 17. Percentages of footprint-days when the models deviate by 20% and 50% from the actual (a) transport speed, (b) centroid location within 24 h, (c) between 24 h and 48 h, and (d) after 48 h of transport.

(6) Although the TCAL and VCAL model codes are nearly identical, the varying mixing depth in VCAL appeared to yield a better transport simulation. Since many of the TCAL and VCAL footprint-days correspond to the same days (Table 7), this conclusion is likely valid.

(7) Although MLAM-FINE and MLAM-COARSE are virtually identical models, the more-frequent mass redistribution in MLAM-FINE apparently afforded no benefit; in fact, consistently the performance was degraded. Although the MLAM-FINE footprint set was limited to the first half of the ANATEX period (unlike the MLAM-COARSE footprint set), few of the MLAM-COARSE footprints pertain to the second half--6 out of a total of 44 MLAM-COARSE footprints (Table 7); therefore, this conclusion also is likely valid.

(8) With few exceptions, the ARL and HY-SPLIT model performances were similar; use of the high-resolution NGM meteorological data in HY-SPLIT did not appear to substantially improve its performance.

The remainder of Section 5.3 discusses in more detail the performances of each of the 11 models. The discussions are ordered by model genre.

5.3.2 Single-Layer Lagrangian Models

SRL--

Approximately one-third (36%) of the 36 SRL footprint-days were within ±20% of the measured transport speed (Figures 18a and 18b). For seven footprint-days, the SRL transport speeds were more than double the measured speeds. After 48 h of transport, only 3 of 10 footprint-days were within ±50% D_m of the measured centroids (Figures 18c and 18d). SRL tended to overestimate transport speeds and place centroids to the right of the measured centroids.

There were relatively few SRL footprint-days when PTCH and PDCH footprints could be related to a single release. Because this model diluted the surface concentrations much more slowly than the measurements indicated, many of the SRL footprints blended with others. For example, after 150 h of transport, the SRL surface concentrations for PTCH-3 (the third release of PTCH tracer) exceeded 100 dfL/L (compared to measurements <10 dfL/L).

For two of the three PDCH footprint-days when SRL significantly underpredicted the transport speeds (PDCH-20, -27), low wind speeds and reversals in wind directions (complicated wind patterns) occurred. In contrast, winds for PDCH-8 had straight-line flow at a moderate speed of nearly 20 km/h. SRL overpredicted transport speeds by at least a factor of 2 for PDCH-6 and PDCH-24, both of which were intercepted by cyclones and/or fronts.

For the footprint-days with the largest separation distances beyond 40 h (i.e., PTCH-3 and PDCH-15, -24, -27), cyclonic and/or frontal interceptions occurred. The transport speed was overestimated for each of these footprint-days.

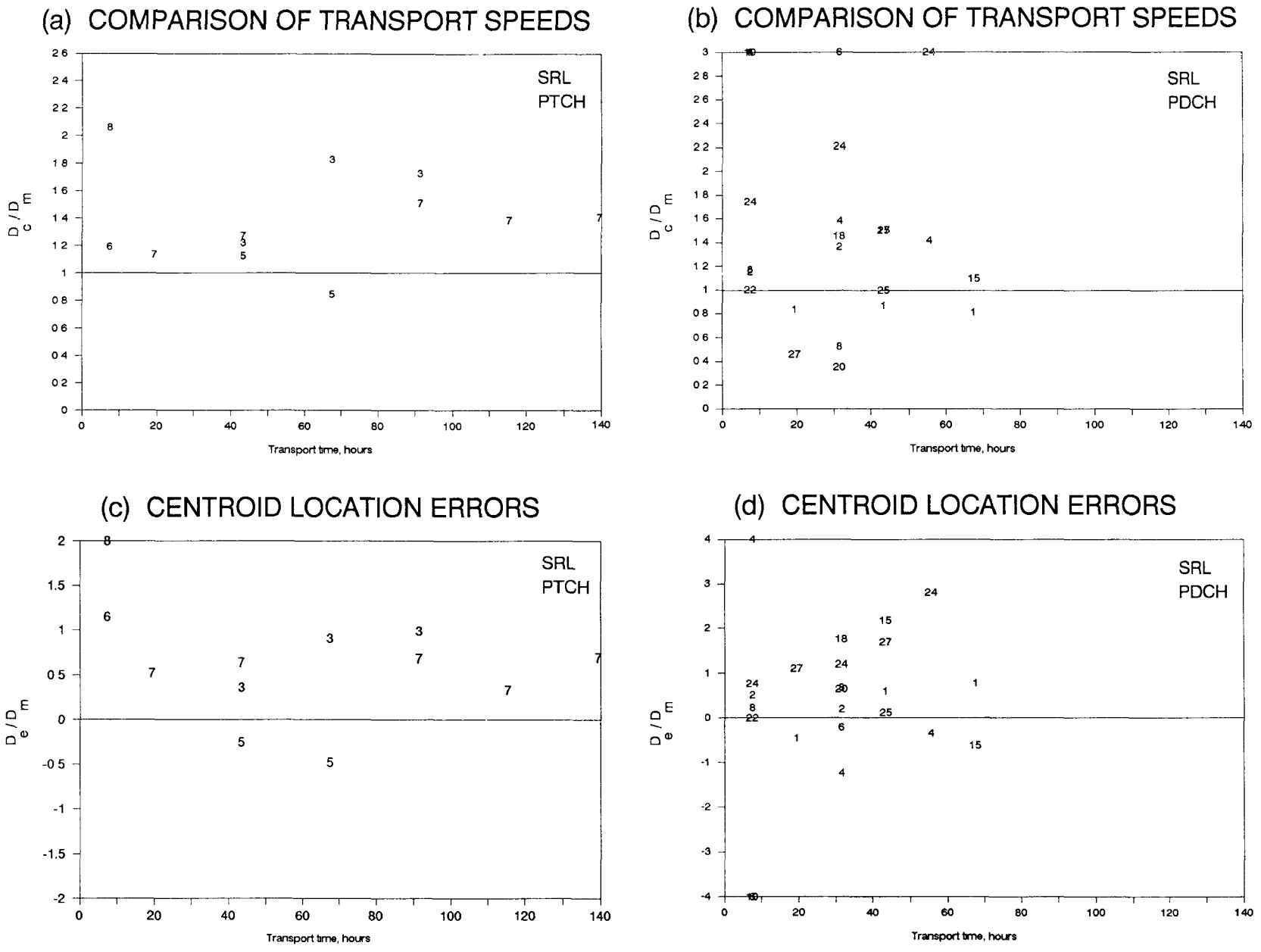


Figure 18. Errors in SRL transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

TCAL (Sibling of VCAL)--

Only 24 of the 68 TCAL footprint-days (35%) were within ±20% of the measured transport speed (Figures 19a and 19b). The transport speeds for the majority of footprint-days (62%) were overestimated; 40% of these were overestimated by more than a factor of 2. After 48 h of transport, 17 out of 30 footprint-days (57%) were within ±50% D_m of the measured centroids (Figures 19c and 19d). The greatest centroid location errors tended to occur for cyclonic and/or frontal interceptions (Table 10).

The model particularly showed problems with PDCH-4, -5, -6, -7, -14 and -15, where all but PDCH-5 and -14 were affected by cyclonic and/or frontal interceptions. In the cases of PDCH-5 and-14, the model substantially underestimated and overestimated, respectively, the transport speed during the first 24 h, when the actual boundary layer winds were very light.

For the footprint-days when the centroid separation distance was great, TCAL tended to locate its centroids to the right of the measured centroids. Because TCAL is a single-layer model that computes a layer-average transport vector, it may have overemphasized the higher-level winds for these footprint-days.

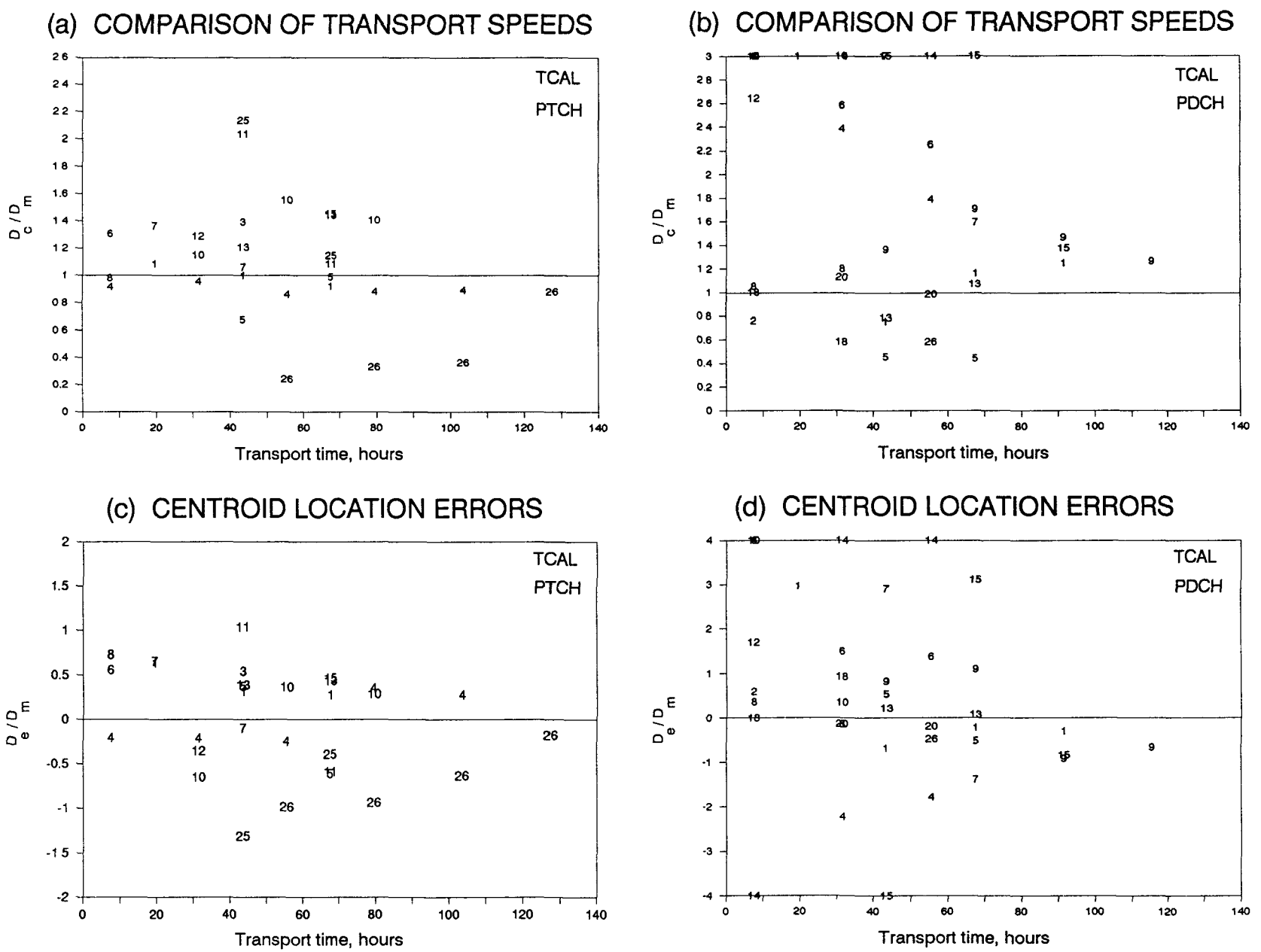


Figure 19. Errors in TCAL transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

VCAL (Sibling of TCAL)--

Slightly more than half (55%) of the 64 VCAL footprint-days were within ±20% of the measured transport speed (Figures 20a and 20b). Like TCAL, VCAL showed a tendency to overestimate transport speeds. However, there were only 6 footprint-days when the overestimates exceeded a factor of 2, compared to 24 for TCAL. After 48 h of transport, 19 out of 24 footprint-days (79%) were within ±50% D_m of the measured centroids (Figures 20c and 20d).

The model performed particularly well for PTCH-1, -3, -5, -11, -13, and -15 and for PDCH-1, -8, -9, and -13; these represent a mixture of meteorological scenarios. Beyond 80 h of transport, the VCAL centroids for these footprints were well within ±50% D_m of the measured centroids. VCAL tended to place its PTCH centroids to the right of the measured centroids. However, such a tendency was not apparent for PDCH.

Compared to the 47 TCAL footprint-days common to the set of VCAL footprint-days, the VCAL transport speeds and centroids in nearly all cases were as close or closer to the measured transport speeds and centroids than those of TCAL (Figures 21a - 21d). The VCAL centroids were slightly farther from the measured centroids for only three footprint days.

Like TCAL, the greatest VCAL errors tended to correspond to cyclonic and/or frontal interceptions. In contrast, however, the cases when VCAL errors were least tended also to correspond to cyclonic and/or frontal interceptions, suggesting that VCAL is better suited than TCAL to simulate transport during the more complicated flow patterns. In general, the single-layer wind vector computed by VCAL to represent the flow within its varying mixed depth appeared to be more representative of the actual flow than the single-layer wind vector computed by TCAL to represent the flow in the fixed 300-1,500 m layer.

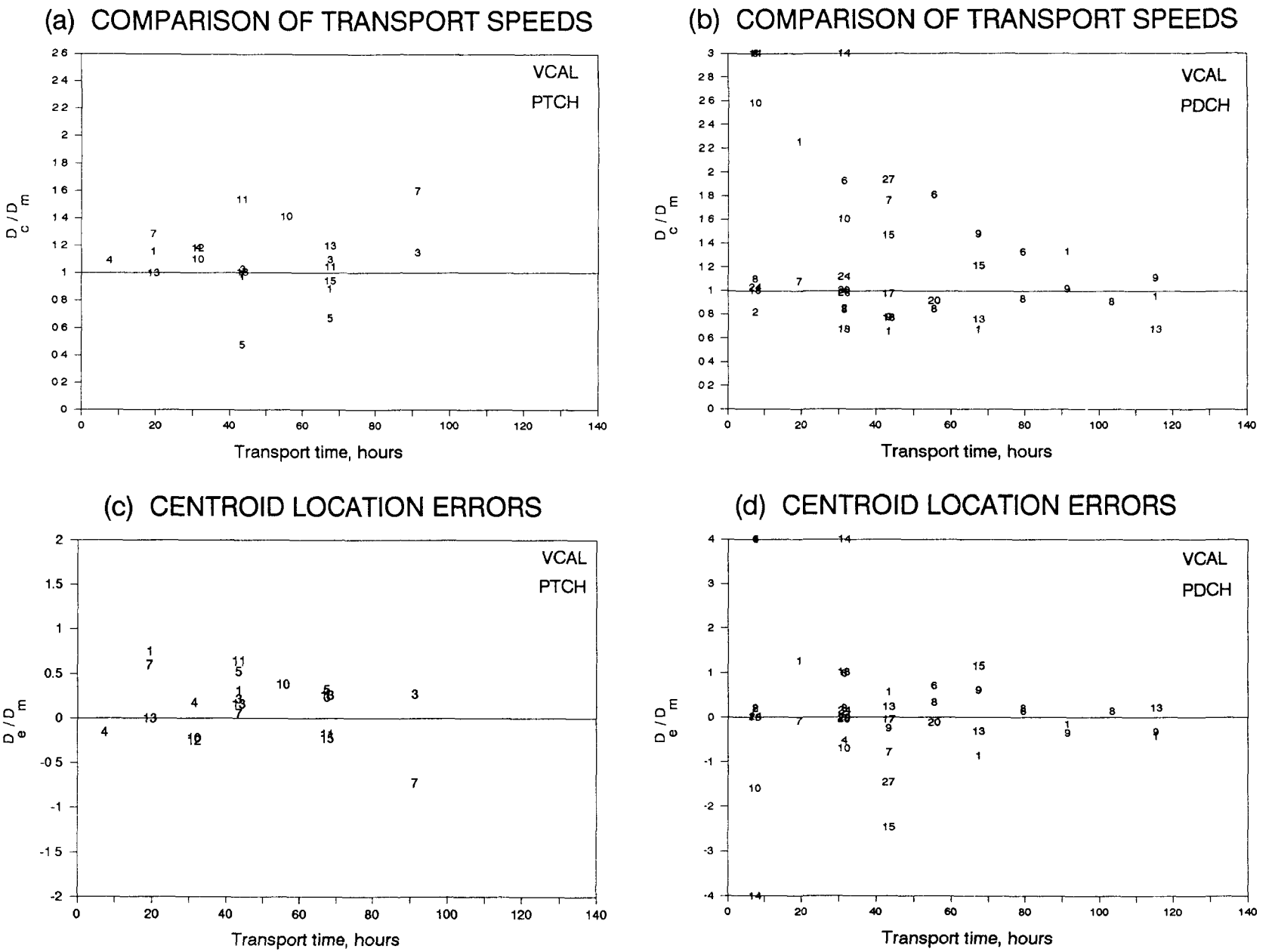


Figure 20. Errors in VCAL transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

(This page is intentionally left blank.)

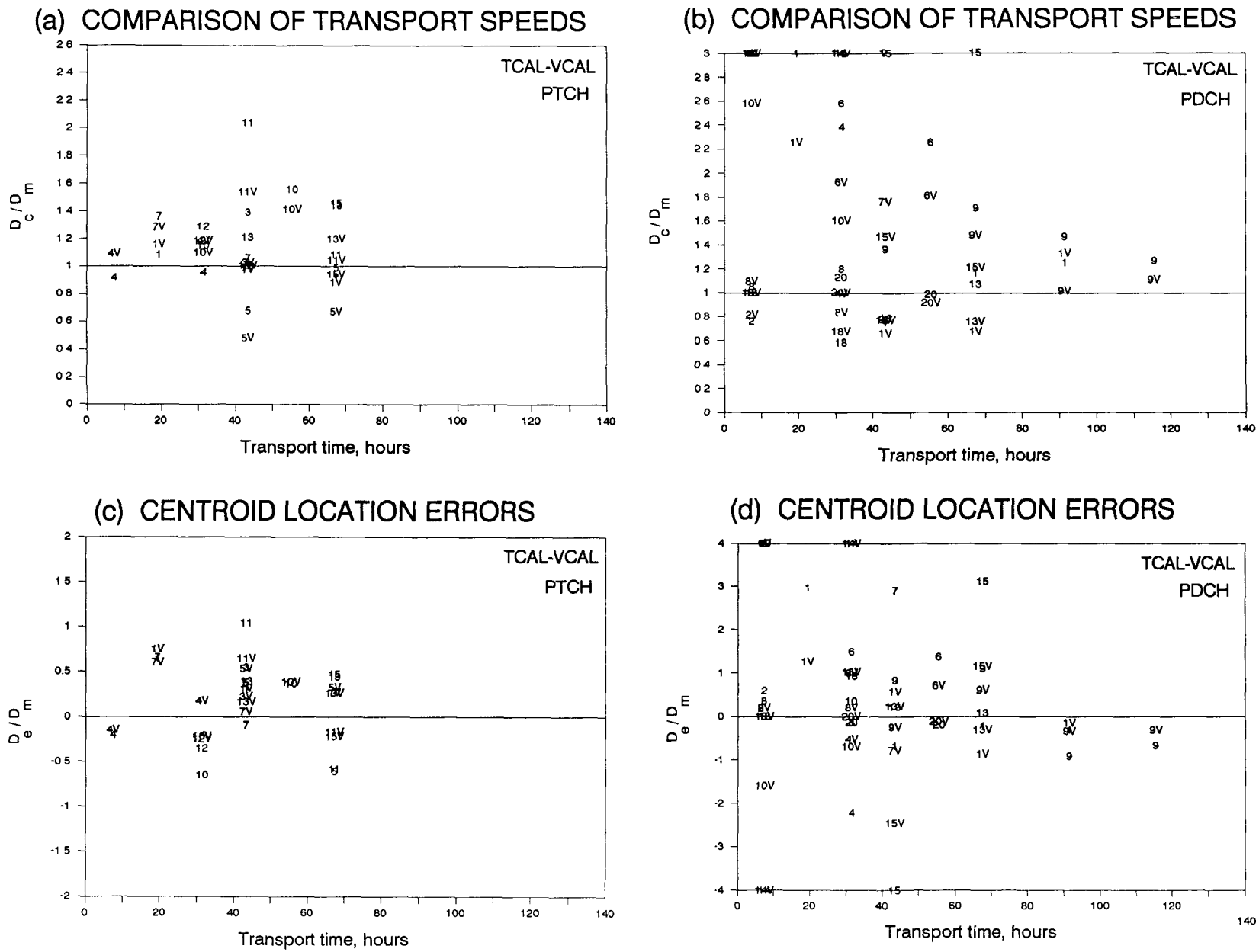


Figure 21. Comparison of TCAL and VCAL (V) errors in transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints common to both sets for the designated tracer releases.

5.3.3 Multiple-Layer Lagrangian Models

ARL--

Approximately half (47%) of the 72 ARL footprint-days were within ±20% of the measured transport speed (Figures 22a and 22b). When the error exceeded ±20%, the model tended to overpredict--24 overpredicted footprint-days versus 14 underpredicted footprint-days. After 48 h of transport, 44% of the footprint-days were within ±50% D_m of the measured centroids (Figures 22c and 22d).

ARL was particularly successful with the footprint-days involving PDCH-1 and -4; both sets were characterized by minimal wind direction changes with time. For each case, ARL's transport speeds and centroid locations were very similar to those of the measured footprints, even after 100 h of transport.

The greatest errors in transport speed and centroid location tended to correspond to footprint-days with cyclonic and/or frontal interceptions (Table 10), which introduce more complicated wind patterns and vertical shears. During these meteorological scenarios, it is possible that ARL either relied too heavily on the greater wind speeds at higher levels or did not replicate the looping of the actual trajectories (and thereby, transporting the tracer farther downwind).

For the PTCH footprint-days of overpredicted transport speeds, the ARL footprint centroids tended to be to the right of the measured centroids. Thus, higher-level wind speeds could have been overemphasized. However, for the PDCH footprint-days of overpredicted transport speeds, the opposite tended to be true. The cause of the difference warrants further investigation by the modeler.

The model underpredicted the transport speeds for PTCH-26 by 60% to 80% and, as a result, erred in locating the centroids by at least 60% D_m. ARL PTCH-26 centroids were always to the left of the measured centroids. Since cyclonic and frontal interceptions were not factors, it is possible that ARL overemphasized the lower-level winds for this case.

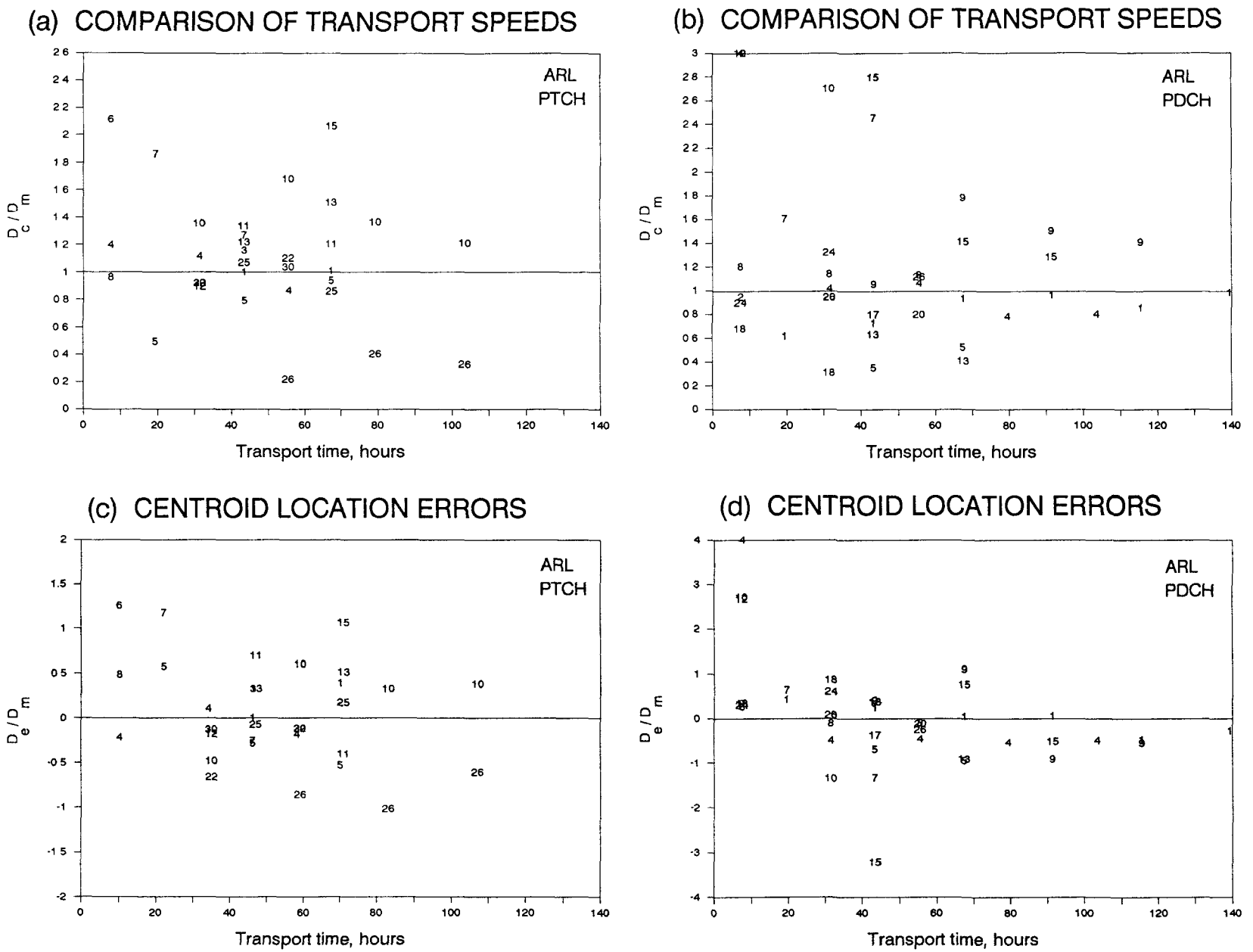


Figure 22. Errors in ARL transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

BAT--

More than half (57%) of the 61 BAT footprint-days were within ±20% of the measured transport speed (Figures 23a and 23b). When the error exceeded ±20%, BAT (like ARL) tended to overpredict--19 overpredicted footprint-days versus 7 underpredicted footprint-days. After 48 h of transport, all 12 PTCH cases and all but one PDCH case were within ±50% D_m of the measured centroids (Figures 23c and 23d). After 40 h, BAT exhibited no tendency to place its centroids to the left or right of the measured centroids; however, for the greatest errors, BAT tended to place its centroids to the right of the measured centroids. The greatest BAT errors in both transport speeds and centroid locations virtually always correspond to cyclonic and/or frontal interceptions (Table 10).

This model performed much better for PTCH than PDCH. For only one PTCH case was the calculated transport speed in error by more than ±40%, compared to 12 PDCH footprint-days. One may think that this performance difference is related to the fact that the wind patterns tend to be simpler for PTCH releases that are transported to the sampling domain. However, all but the PTCH-1 and -8 footprint-days had cyclonic and/or frontal interceptions. Therefore, the better performance could not be attributed to simpler wind patterns.

BAT appeared to have problems with PDCH-6, -10, -14, and -15 in the initial stages of transport. BAT transport speeds during the first 32 h were at least double the measured speeds; directional bias was inconsistent. All of these footprint-days but PDCH-10 were affected by cyclonic and/or frontal interceptions during the first 48 h. In contrast, even beyond 100 h, BAT performed well for PDCH-1--the case with minimal wind direction changes.

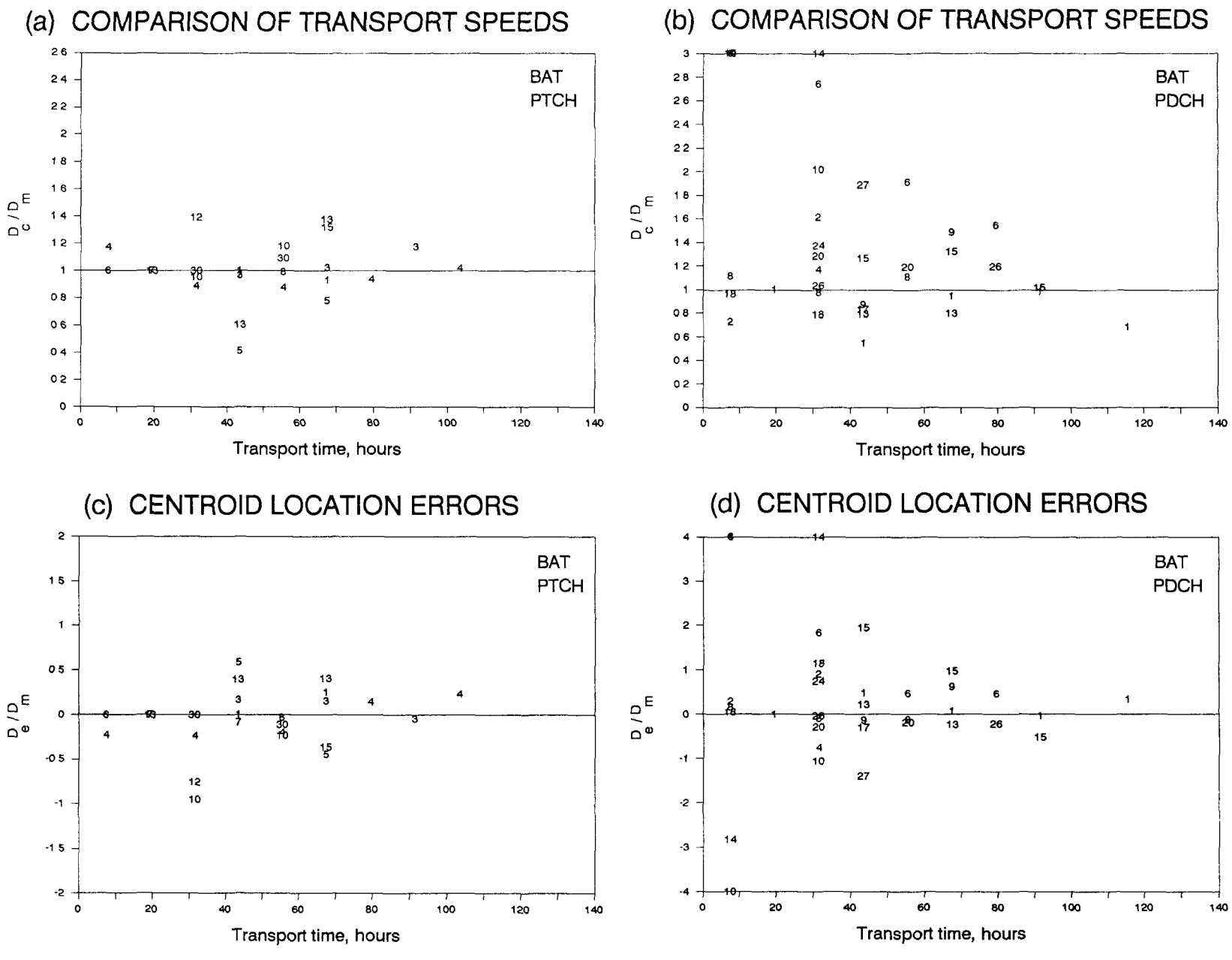


Figure 23. Errors in BAT transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

GAMUT--

Approximately half (46%) of the 61 GAMUT footprint-days were within ±20% of the measured transport speed (Figures 24a and 24b); there was no apparent tendency to under- or overestimate transport speeds (29 underestimates and 29 overestimates). After 48 h of transport, 58% of the footprint-days were within ±50% D_m of the measured centroids (Figures 24c and 24d). The GAMUT centroids showed a strong tendency to be to the left of the measured PTCH centroids, but showed no tendency for the PDCH centroids; the high frequency of PTCH footprints affected by cyclones and/or fronts (Table 9) could have been a factor. The least and greatest errors did not correspond to cyclonic and/or frontal interceptions.

GAMUT transport speeds and centroid locations compared favorably with the measurements for the PDCH-1, -5 and -8 footprint-days, even out to 100 h. For the latter two sets of footprint-days, cyclonic and/or frontal interceptions occurred after 48 h of transport. The model performed very well relative to the other models for PTCH-4, which was released at the time of a frontal passage; transport speeds differed by less than ±10% between 10 and 110 h of transport. After 100 h, the calculated centroid was within 30% D_m. However, for the case of PTCH-25, released 24 h prior to a frontal passage, the model did not perform as well.

There were five footprint-days beyond 12 h of transport when the GAMUT transport speeds were at least double the measured speeds. For four of these, cyclonic and/or frontal interceptions occurred during the first 48 h of transport. The two outlier footprint-days of PTCH-15 (released 24 h prior to a frontal passage) apparently involved improper timing of a wind direction shift; however, with time, the accumulated error decreased.

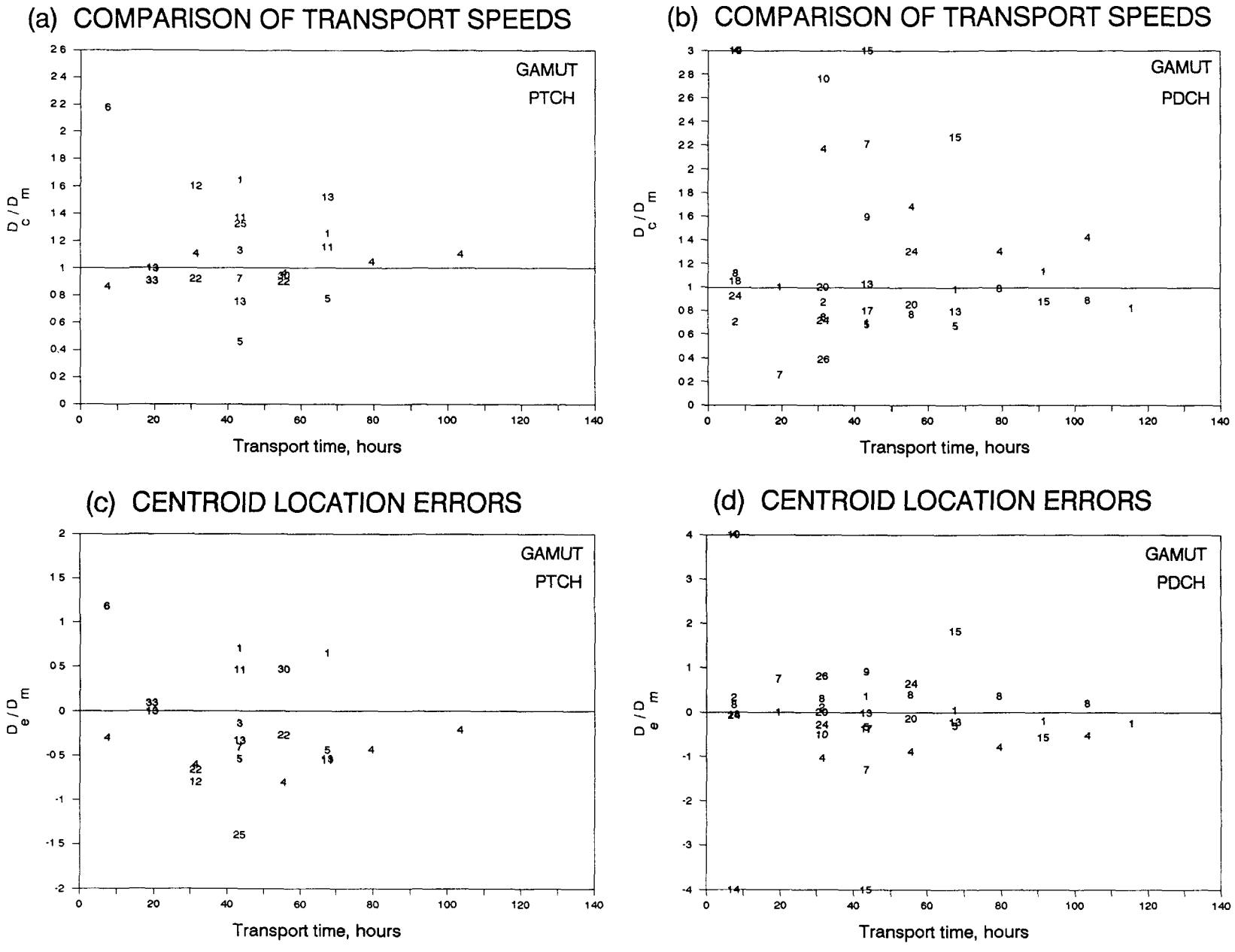


Figure 24. Errors in GAMUT transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

HY-SPLIT (Version of ARL)--

Approximately half (46%) of the 50 HY-SPLIT footprint-days were within ±20% of the actual transport speed (Figures 25a and 25b). Model performance varied significantly with the tracer; that is, for 19 of the 23 PTCH footprint-days--compared to 9 of 29 PDCH footprint-days--the calculated transport speed was in error by less than ±25%. This is further evidence that the PTCH data--by virtue of the location of the GGW release site--may be biased toward simpler flow patterns for which the models are more apt to perform well. Similar to ARL, approximately half (58%) of the HY-SPLIT footprint centroids beyond 48 h of transport were within ±50% D_m of the actual centroids (Figures 25c and 25d).

HY-SPLIT tended to overpredict PDCH transport speeds and place its centroids to the left of the actual centroids (especially PDCH-4, -10, -15, -20, and -27). However, for PTCH, this model was generally within ±25% of actual transport speeds and within ±50% D_m for centroid locations.

Determining the relationship between HY-SPLIT errors and types of meteorological scenarios was hindered by the bias of PTCH cases towards cyclonic and/or frontal interceptions (17 with and only 4 without); PDCH cases were more balanced (15 with and 20 without). The cases when HY-SPLIT transport speeds differed most from the actual transport speeds tended to occur when cyclones and/or fronts intercepted the tracer (PTCH-7, -15; PDCH-4, -5, -6, -10, -15, -27). In two other cases HY-SPLIT overpredicted for PDCH-20 and underpredicted transport speeds for PDCH-26 footprints, both within ridges of high pressure characterized by light and variable winds. For the footprints of 8 of these 10 tracer releases, HY-SPLIT transport speeds were greater than the actual transport speeds.

Similarly, HY-SPLIT location errors tended to be greatest for cyclonic and/or frontal interceptions (PTCH-7, -10, -15; PDCH-4, -6, -10, -15, -27). Location errors for PDCH-18 and PDCH-20 were also relatively large; although these cases involved no cyclonic and/or frontal interceptions, light winds prevailed. PDCH location errors tended to be the least for cases when there were no cyclonic and/or frontal interceptions (PDCH-1, -8, -13, -17, -21, -26). Since all but four PTCH footprint-days were affected by cyclones and fronts, HY-SPLIT performance could not be definitively related to types of meteorological scenarios. However, for the four footprint-days not affected by cyclones or fronts, the location errors were the lowest (PTCH-1, -5, -13 during the first 48 h of transport). After 48 h, the location errors for PTCH-5 and PTCH-13 increased as fronts intercepted the tracer. There were other cases when HY-SPLIT location errors were small, despite cyclonic and/or frontal interceptions (PTCH-4, -13, -30), demonstrating that the model could perform well under such circumstances.

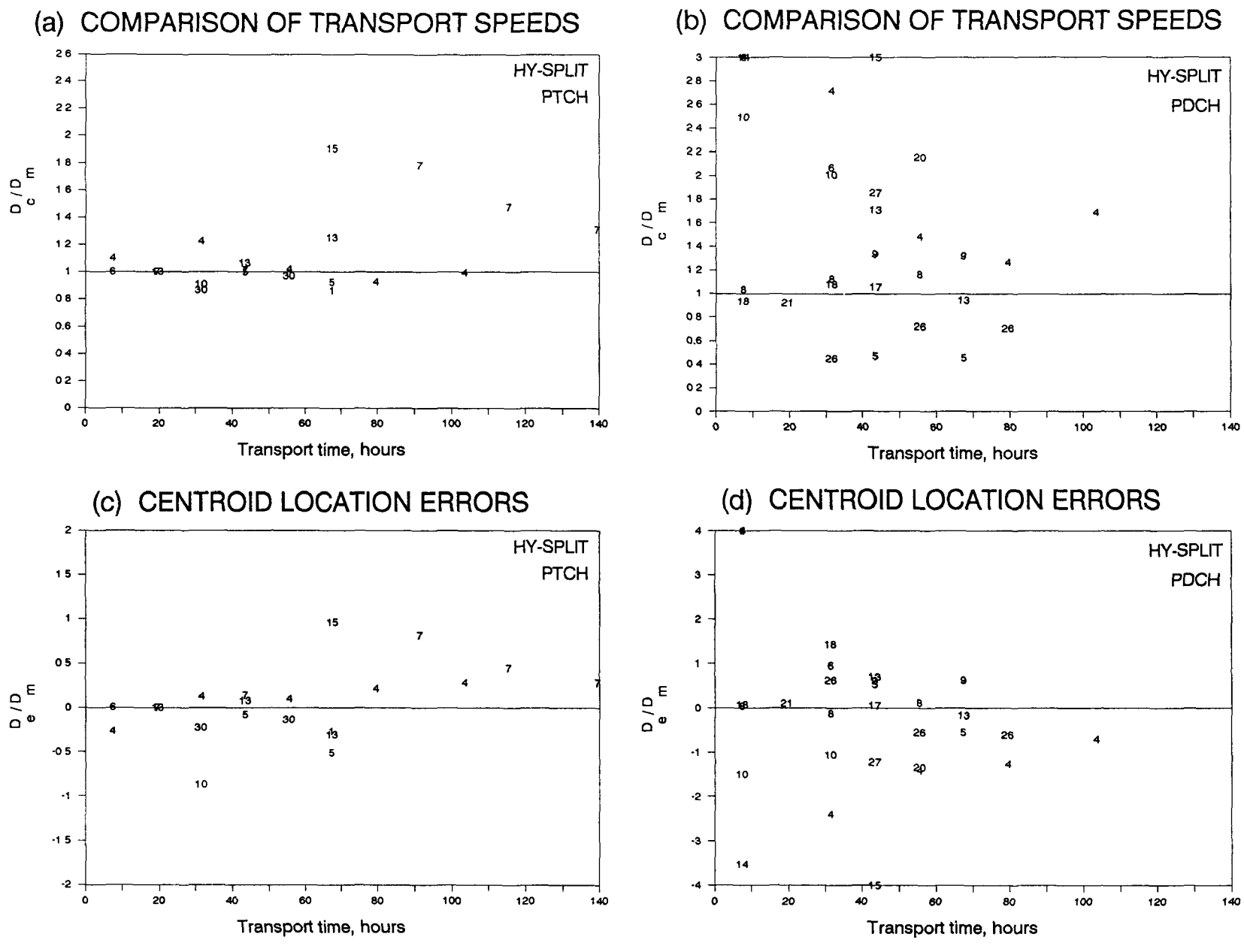


Figure 25. Errors in HY-SPLIT transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

Figures 26a and 26b compare the transport speed performances of HY-SPLIT and its sibling model ARL for only those footprint-days common to both models. The numbers accompanied by an "H" refer to the HY-SPLIT results, while those without refer to the ARL results. For the PTCH tracer, HY-SPLIT tended to outperform ARL for 10 of the 15 footprint-days; while for the PDCH tracer, neither consistently outperformed the other. Figures 26c and 26d show similar results for location errors in that HY-SPLIT tended to outperform ARL for the PTCH tracer, but for the PDCH tracer, no model consistently performed better than the other.

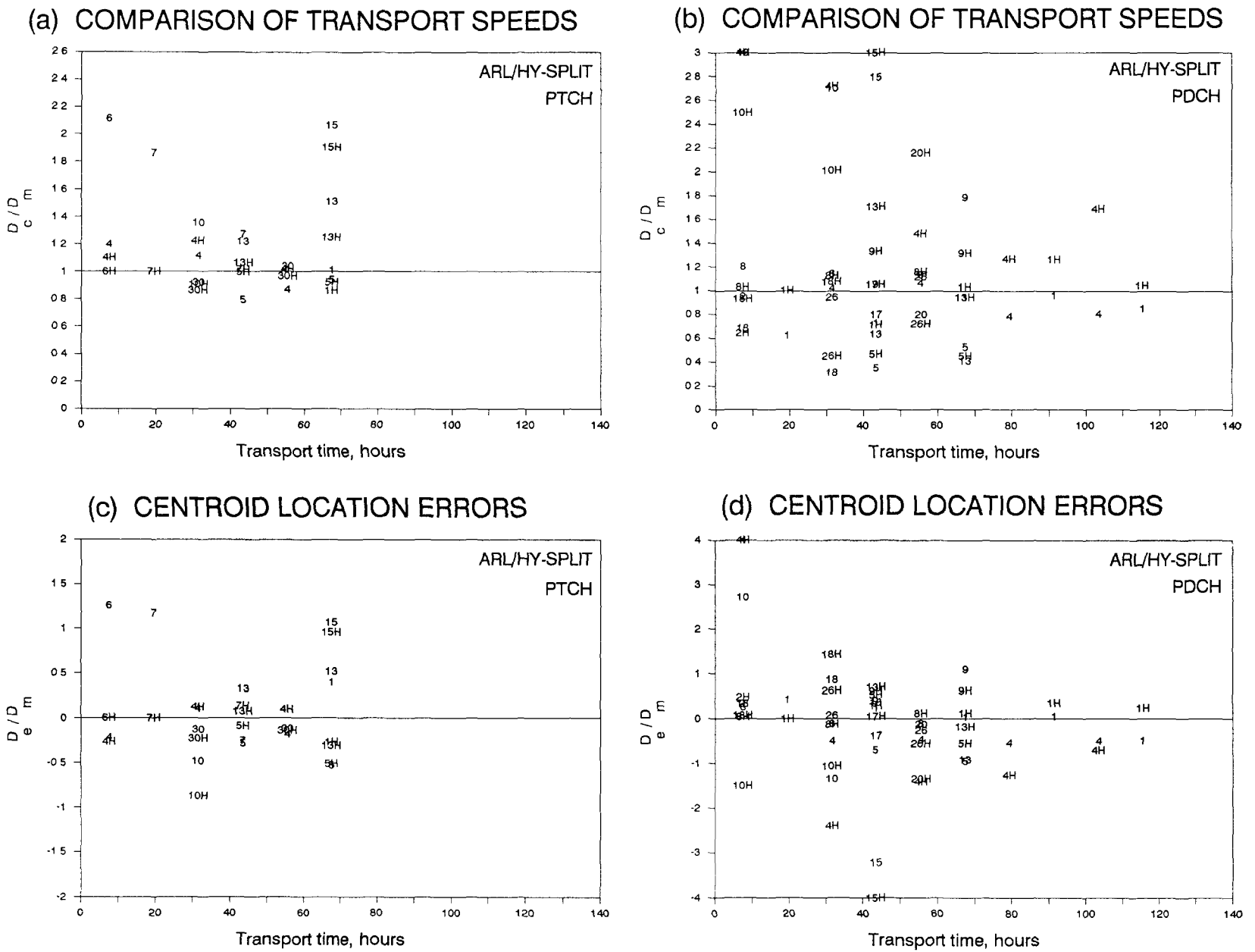


Figure 26. Comparison of ARL and HY-SPLIT (H) errors in transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints common to both sets for the designated tracer releases.

MLAM-FINE (Sibling of MLAM-COARSE)--

Of the 44 MLAM-FINE footprint-days in the first half of the ANATEX period, transport speeds for only about one-third (35%) were within ±20% of the measured transport speed (Figures 27a and 27b). However, virtually all transport speeds beyond 12 h of transport were within ±40% of the measured transport speed. The strong tendency for the model to underpredict transport speeds was apparent. Despite this tendency, after 48 h of transport, 16 of the 18 footprint-days (89%) were within ±50% D_m of the measured centroids (Figures 27c and 27d).

The success of this model in calculating centroid locations beyond 48 h was not attributable to a predominance of simple wind patterns; all but PTCH-1 and -8 and PDCH-1 footprint-days involved cyclonic and/or frontal interceptions. In fact, for the footprint-days with the least errors in transport speeds and centroid locations, the occurrences of cyclonic and/or frontal interceptions out-numbered the nonoccurrences by a 2:1 margin. This suggests that the model is, relative to the other models, well suited to simulate transport for complex wind patterns.

MLAM-FINE appeared to have the most difficulty in simulating PTCH-4, which was affected by a cyclonic and/or frontal passage as the tracer was released. It also appeared to have difficulty in calculating the transport speeds for PTCH-5, which was released in the center of an anticyclone (a region of low wind speeds).

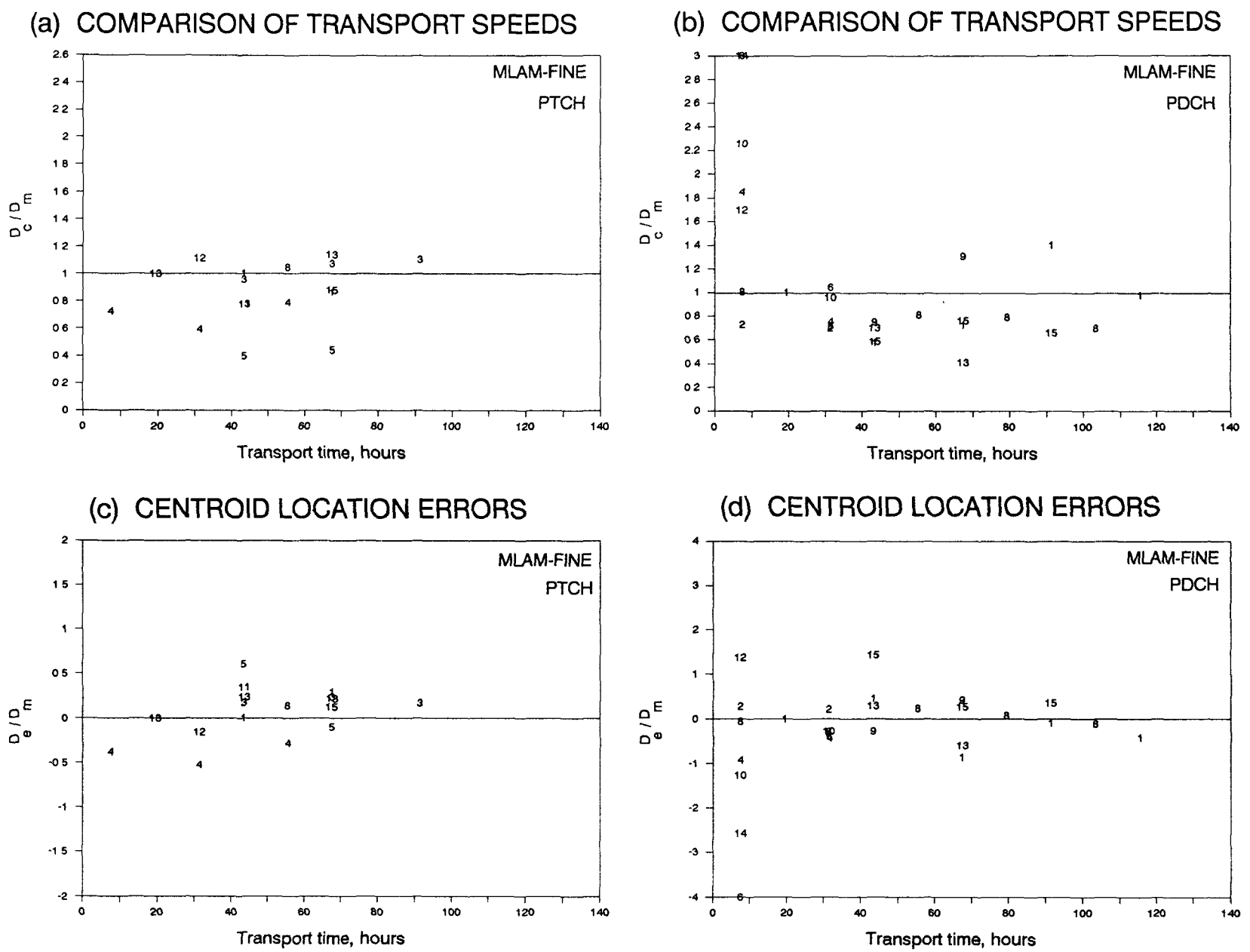


Figure 27. Errors in MLAM-FINE transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

MLAM-COARSE (Sibling of MLAM-FINE)--

Of the 44 MLAM-COARSE footprint-days in both halves of the ANATEX period, slightly more than half (59%) were within ±20% of the measured transport speed (Figures 28a and 28b). Like MLAM-FINE, MLAM-COARSE exhibited a strong tendency to underpredict transport speeds. Similarly, after 48 h of transport, 18 of the 19 footprint-days (95%) were within ±50% D_m of the measured centroids; most footprint-days were within ±25% D_m (Figures 28c and 28d). Both model versions exhibited a slight tendency to place centroids to the right of measured centroids.

For most of the footprint-days common to both sets, there was little if any difference in model performance (Figures 29a - 29d). Neither model version consistently performed better than the other for the remaining footprint-days. This indicated that the more-frequent, vertical mass redistribution in MLAM-FINE (i.e., twice daily versus once daily in MLAM-COARSE) did not improve the transport performance. Moreover, the more-frequent, vertical mass redistribution did not appear to improve the diffusion performance, as Table 5 indicates. That is, MLAM-FINE mean concentrations along bands of sites tended to differ from the mean measured concentrations more so than those of MLAM-COARSE. Relative to other sources of variation in this study, the differences between the MLAM-FINE and MLAM-COARSE code versions appear unimportant.

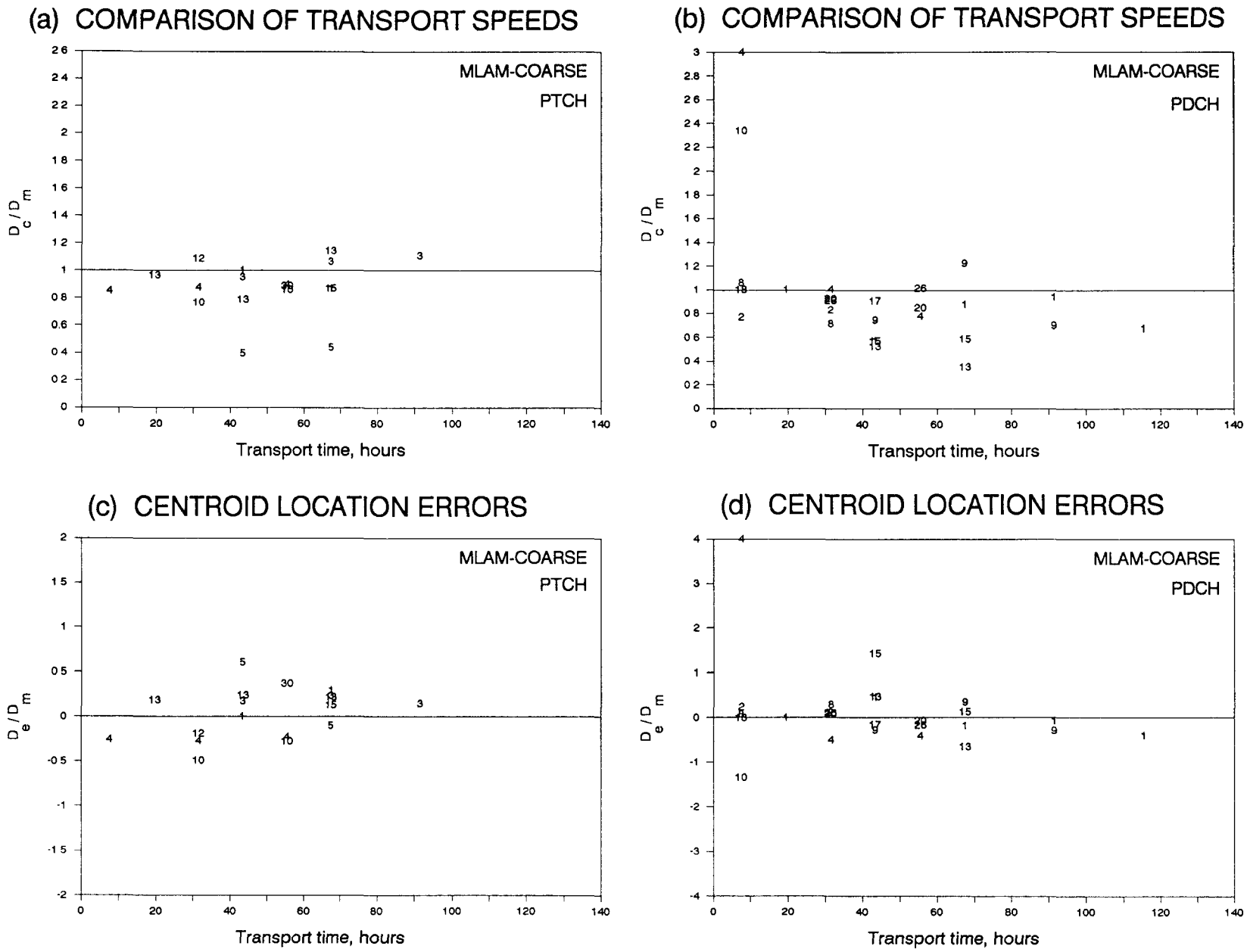


Figure 28. Errors in MLAM-COARSE transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

(This page is intentionally left blank.)

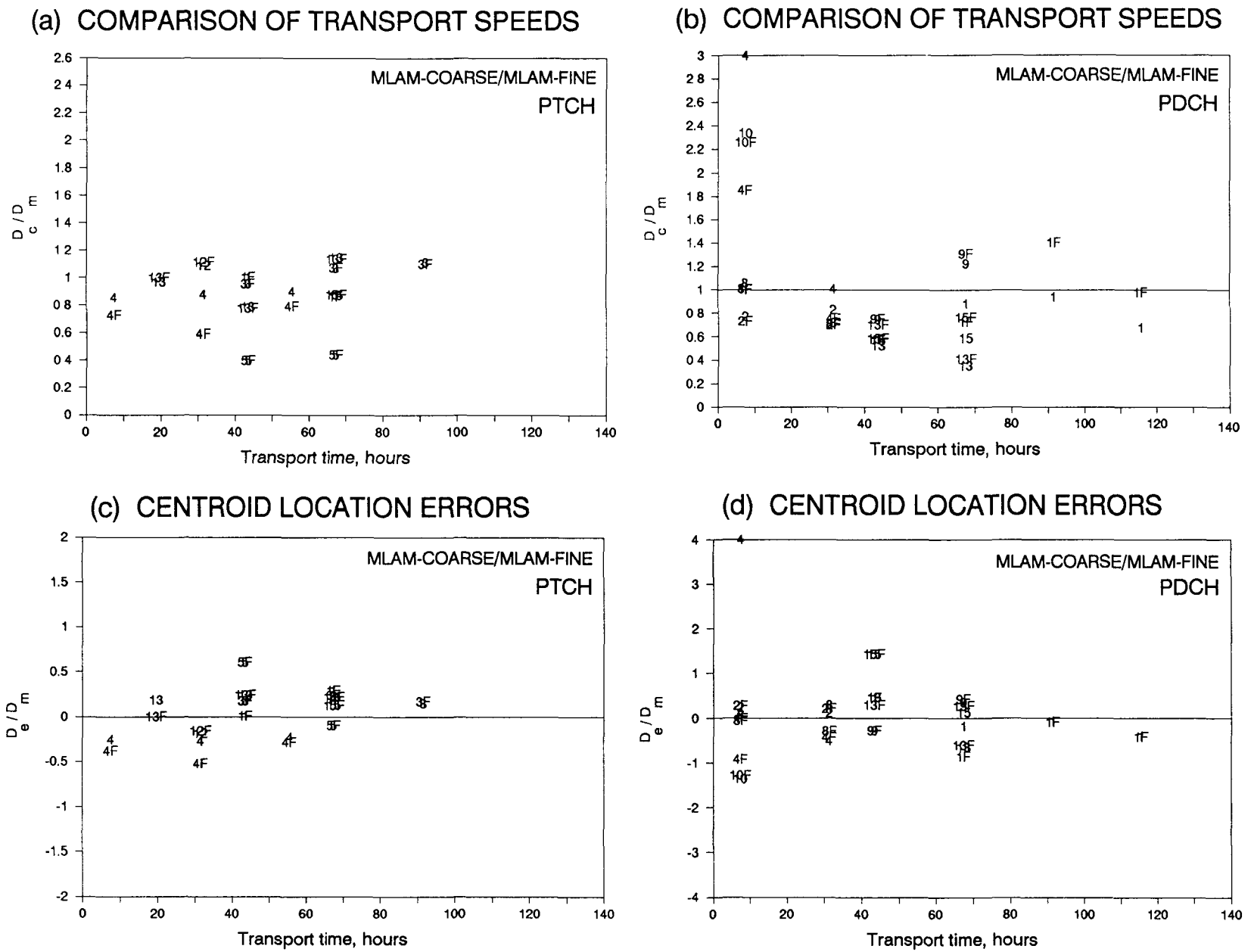


Figure 29. Comparison of MLAM-COARSE and MLAM-FINE (F) errors in transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints common to both sets for the designated tracer releases.

5.3.4 Multiple-Layer Eulerian Models

ADOM--

For less than one-third (32%) of the 28 footprint-days in the first half of the ANATEX period were the ADOM transport speeds within ±20% of the actual transport speeds (Figures 30a and 30b). Another one-third were in error by at least a factor of 2. Although its greatest errors were overpredictions, ADOM showed a strong tendency to underpredict the transport speeds. After 48 h of transport, the calculated centroids for 5 of the 7 footprint-days (71%) were within ±50% D_m of the measured centroids (Figures 30c and 30d).

ADOM transport tended to be to the right of the actual transport (19 of 28 footprint-days). Unlike many of the other models, for those cases when the separation distance was greatest, the model did not consistently underpredict or overpredict the transport speeds.

ADOM performance in predicting transport speeds tended to be related to the meteorological scenarios. The greatest ADOM transport speed errors tended to pertain to cyclonic and/or frontal interceptions when the model overpredicted (PTCH-4, -5, -15; PDCH-4, -10, -15). Exceptions were PDCH-1, which followed, but never was intercepted by, a cold front, and PDCH-8, which remained in a pressure ridge with light wind speeds. The model performed best for PTCH-1, -7, -11 and PDCH-1, -2; only PTCH-7 and PTCH-11 were intercepted by cyclones and/or fronts.

Likewise, the greatest centroid location errors tended to pertain to cyclonic and/or frontal interceptions (PTCH-15; PDCH-4, -10, -15). ADOM performed well for PTCH-1 and PDCH-1--both unaffected by cyclones and/or fronts--but also performed well for cases pertaining to interceptions (PTCH-4, -5, -7, -11; PDCH-2, -5). Therefore, ADOM location errors appeared to be less related to meteorological scenario than did its transport speed errors.

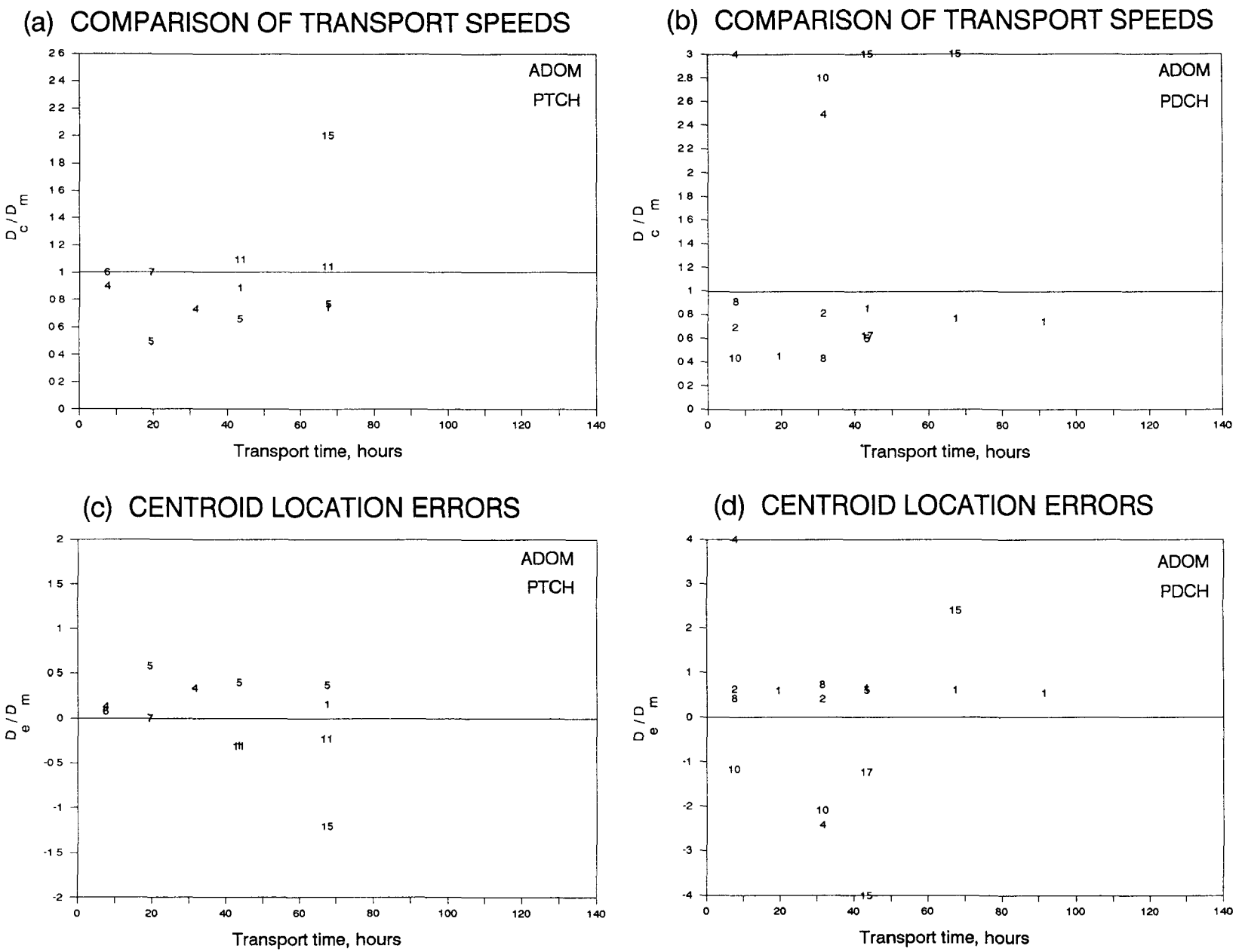


Figure 30. Errors in ADOM transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

ADPIC--

Similar to SRL, ADPIC diluted surface concentrations much slower than the measurements indicated. Consequently, there were only 27 ADPIC footprint-days for 15 releases when footprints could be related to a single release. This was the lowest total for any model--even those that were applied only during the first half of the ANATEX period [ADOM (28) and MLAM-FINE (44)]. Since there were so few footprint-days, relating ADPIC errors to type of meteorological scenario was not attempted.

As Table 12 indicates, ADPIC performance mirrored that for ADOM, the other Eulerian model. Less than one-third (30%) of the ADPIC footprint-days were within ±20% of the actual transport speed (Figures 31a and 31b). Like ADOM, ADPIC tended to underpredict the transport speeds. After 48 h of transport, 5 of the 7 footprint-days (71%) were within ±50% D_m of the measured centroids (Figures 31c and 31d), identical to ADOM. Unlike ADOM, however, ADPIC located its centroids to the right of the actual centroids for virtually all 19 PDCH footprint-days; the opposite was true for the PTCH footprint-days.

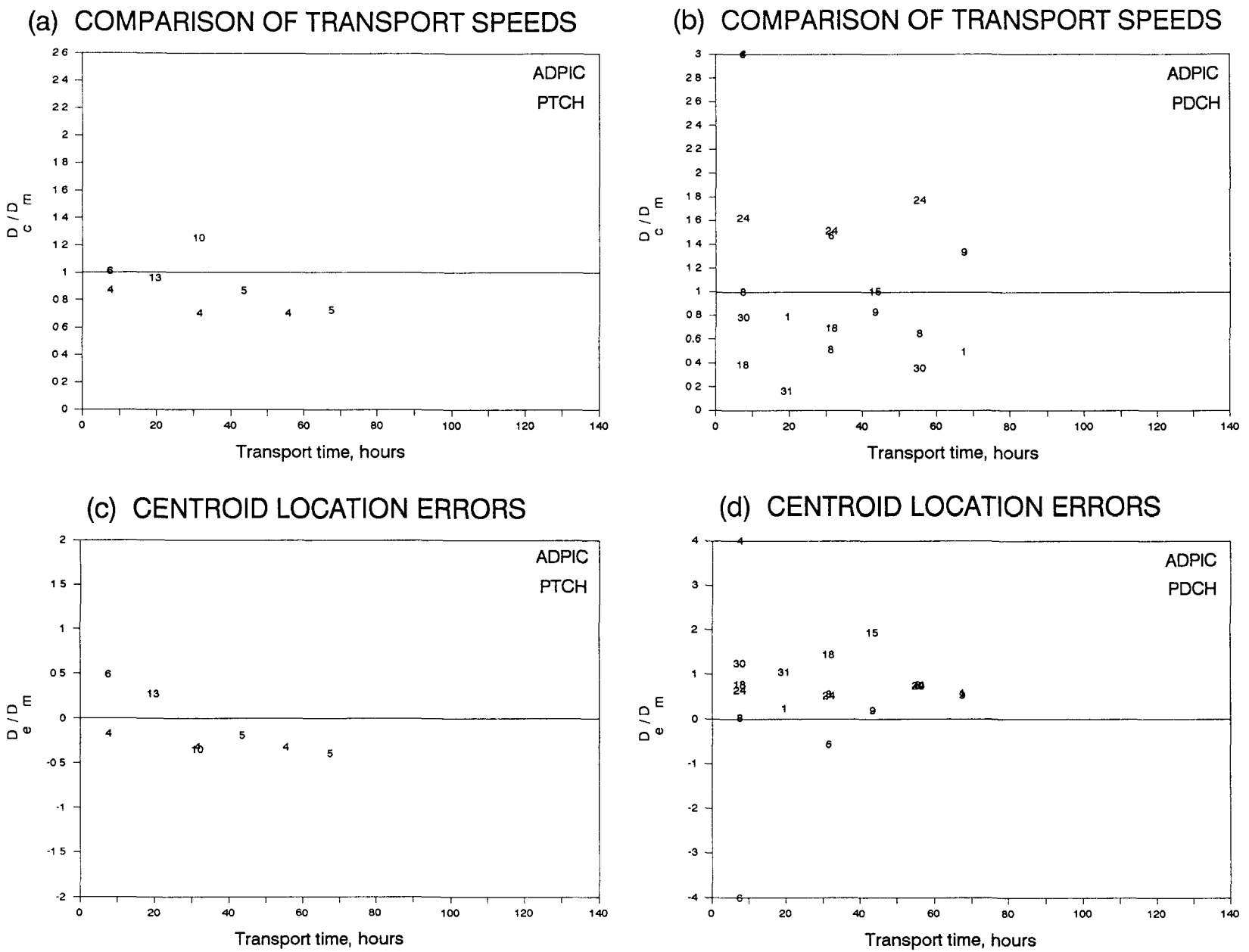


Figure 31. Errors in ADPIC transport speeds for (a) PTCH and (b) PDCH footprints; and centroid locations for (c) PTCH and (d) PDCH footprints for the designated tracer releases.

SECTION 6

SUMMARY AND CONCLUSIONS

The transport and horizontal diffusion calculated by 11 atmospheric models were compared to the transport and horizontal diffusion determined from 24-h tracer concentrations across a 77-site surface sampling network encompassing central and eastern North America. Three of these models are single-layer Lagrangian (SRL, TCAL, and VCAL), six are multiple-layer Lagrangian (ARL, BAT, GAMUT, HY-SPLIT, MLAM-FINE, and MLAM-COARSE), and two are multiple-layer Eulerian (ADPIC and ADOM). Tracer concentrations were measured during the period 5 January - 29 March 1987.

The limitations of the ANATEX data (e.g., virtually no vertical tracer distributions beyond 300 km of the release sites, the spacing between sites, and 24-h integrated sampling at surface sites), limited the scope of the model evaluation study. Firstly, evaluations based on point-to-point comparisons of simultaneous tracer concentrations were not practical. Secondly, model errors could not be related to specific model processes. Consequently, this model evaluation study focused on identifying model biases for whatever reason. When appropriate, possible problems with modeled processes were offered as explanations for the observed biases. However, a more resolved data base and additional model applications are required to reveal the actual causes of these errors.

The model performance assessment--summarized in Sections 6.1 through 6.3 for each of the three model genres--was based on seven performance measures and charts using either both halves of the entire data set or a subset of this data set relating concentrations to specific tracer releases. Table 14 succinctly summarizes the model performances via either mean errors or mean ratios of calculated-to-measured values derived from some of the figures and tables presented in preceding sections. These performance terms, appearing across the top row of Table 14, are described below.

(1) **Third quartile**--ratios calculated from the frequency distributions of nonzero concentrations (Figures 7a, 7b, 9a, and 9b).

(2) **Frequency of χ's > threshold**--ratios calculated from the percentage of concentrations exceeding the thresholds of 5 dfL/L for PTCH and 8 dfL/L for PDCH (Figures 8a, 8b, 10a, and 10b).

TABLE 14. OVERALL BIASES IN THE MODEL RESULTS AS INDICATED BY THE MODEL PERFORMANCE MEASURES

		Ensemble Measures: Ratios of c/m † (first half, second half)				Footprint Measures		
Model		Third quartile	Frequency of χ 's > threshold	Mean of $\bar{\chi}$ for 3 bands	Mean lateral diffusion for 3 bands	Mean ratio of speed (>7.5 h)	Mean ratio of location error (D_e/D_m) (>7.5 h)	Mean location error (km) (<61.5 h)‡
				Single-Layer Lagrangian:				
SRL	T	*3.7, 23.3	1.2, 0.1	*2.8, --	0.8, --	1.4	+0.44	536
	D	*6.0, 9.3	*2.0, 0.5	*5.0, 2.0	1.0, 0.2	1.4	+0.64	
TCAL	T	2.8, 0.8	0.8, 1.9	1.4, 2.8	0.6, 2.3	1.1	-0.03	570
	D	4.1, 0.8	0.9, 0.9	*2.0, 1.6	0.5, 1.2	2.8**	+0.58	
VCAL	T	*2.2, 1.6	0.9, 2.3	*1.8, 3.0	0.7, 2.1	1.1	+0.17	283
	D	*3.8, 1.8	*1.5, 2.5	*3.0, 2.4	0.7, 1.7	1.4	+0.22	
				Multiple-Layer Lagrangian:				
ARL	T	1.1, 0.8	0.6, 0.6	0.5, 0.9	0.6, 0.9	1.1	+0.03	370
	D	2.3, 1.3	1.0, 0.9	1.1, 0.9	0.6, 0.8	1.1	-0.26	
BAT	T	0.9, 0.7	*0.5, 0.4	0.7, 0.4	0.6, 0.6	1.0	-0.05	298
	D	*2.7, 1.6	0.6, 0.5	1.9, 0.6	*0.4, 0.4	1.5**	+0.44	
GAMUT	T	1.1, 1.2	0.4, 0.8	0.7, 1.2	0.4, 0.8	1.1	-0.27	417
	D	3.2, 1.4	1.0, 1.1	2.2, 1.4	0.5, 0.9	1.3	-0.27	
HY-SPLIT	T	*3.5, 2.3	0.8, 0.6	1.4, 0.7	*0.5, 0.5	1.1	+0.05	366
	D	*4.9, 2.1	0.8, 0.8	1.6, 0.9	0.4, 0.6	1.4	-0.40	
MLAM-FINE	T	0.7, --	0.8, --	0.9, --	1.0, --	0.9	+0.09	270
	D	*3.4, --	*1.9, --	*4.1, --	1.0, --	0.8	0.00	
MLAM-COARSE	T	0.9, 0.9	0.7, 1.7	0.9, --	0.7, 2.1	0.9	+0.06	241
	D	*2.8, 1.5	*1.7, 2.0	*3.8, 1.8	1.0, 1.7	0.8	+0.02	
				Multiple-Layer Eulerian				
ADOM	T	0.6, --	0.8, --	0.6, --	1.1, --	0.9	-0.02	467
	D	*1.6, --	1.3, --	*1.9, --	1.1, --	1.2	-0.93	
ADPIC	T	0.9, 1.0	1.4, 0.3	0.9, 0.4	1.4, 0.3	0.9	-0.21	457
	D	1.2, 1.6	1.8, 1.2	1.7, 0.7	1.6, 0.9	0.9	+0.68	

T: PTCH; D: PDCH

† c/m: calculated/measured.

‡ Mean location error for combined PTCH and PDCH.

* Calculated values for both half-periods differed by at least ±50% of measured values (first half for MLAM-FINE and ADOM).

** Calculated values differed by at least ±50 of measured values.

(3) **Mean of $\overline{\chi}$ for 3 bands**--mean ratios of the ensemble mean concentrations for 3 bands of sites (ranging from 300 km to 2,300 km downwind of the tracer release sites) (Figures 11a, 11b, 12a, and 12b).

(4) **Mean lateral diffusion for 3 bands**--mean ratios of the ensemble mean lateral diffusion for the same 3 bands of sites (Figures 13a, 13b, 14a, and 14b).

(5) **Mean ratio of speed**--determined from the ratios of the 24-h mean transport speeds of all footprint pairs beyond 7.5 h of transport (Figures 18-31); ratios for PTCH and PDCH footprints are provided.

(6) **Mean ratio of location error**--determined from the ratios of 24-h separation distances (D_e) normalized by actual transport distance (D_m) of all footprint centroid pairs beyond 7.5 h of transport (Figures 18-31); ratios for PTCH and PDCH footprints are provided.

(7) **Mean location error**--determined by averaging the mean absolute centroid separation distances for the combined set of PTCH and PDCH footprints at transport times of 13.5 h, 37.5 h, and 61.5 h (Figure 15).

For most of the terms, values were calculated for both PTCH and PDCH for both half periods. Those models that did not perform as well as the others (i.e., when the ratios for both half periods exceded ± 50% of measured values) are identified by an asterisk preceding the values in the table.

6.1 SINGLE-LAYER LAGRANGIAN (SLL) MODELS

6.1.1 Summary

Box Plot Distributions--Each of the three models of this genre (SRL, TCAL, and VCAL) revealed a tendency to overestimate the frequencies of higher concentrations. During the first half-period, the medians and third quartiles of each model were 2-to-6 times greater than those of the measurements. For the second half-period, the same was true for SRL only; TCAL and VCAL values were much closer (i.e., within a factor of 2) to those of the measurements.

Frequency Distributions--During the first half-period the SLL models generally approximated the frequency of concentrations above the thresholds (5 dfL/L and 8 dfL/L for PTCH and PDCH, respectively). However, SRL overestimated the frequencies of concentrations exceeding 99 dfL/L by approximately a factor of 4. TCAL, and to a lesser degree, VCAL closely approximated the distributions. During the second half-period SRL greatly underestimated the frequency of PTCH concentrations above

the threshold (<1% versus 18%), as well as the frequency of nonzero concentrations (1% versus 36%). Meanwhile, the percentage of sites with nonzero concentrations calculated by both TCAL and VCAL was much greater than that for the measurements of both tracers (55% to 80% versus approximately 40%).

Mean Concentrations as a Function of Transport Distance--For transport distances ranging from 300 to 2,300 km, the mean TCAL and VCAL concentrations during both half-periods and the mean SRL concentrations during the first half-period along several bands of sites tended to be higher than those of the other models, as well as the measurements. Deviations of factors of 2 and 3 from the measured means were common; some SRL deviations were as great as a factor of 6.

Mean Lateral Diffusion--The comparison of the model and measured mean plume widths showed an inconsistency between half-periods. For the first half-period, the model mean plume widths were generally underestimated, but within an average of 20% of the measured width for TCAL, 30% for VCAL, and 50% for TCAL. For the second half-period, TCAL and VCAL mean plume widths were greater than the measured mean plume widths especially for PTCH, by as much as 130%. SRL mean plume widths for this period (less than 250 km) were much less than the actual widths--in fact, nearly zero--indicating a serious problem with the diffusion.

Footprint Transport Speeds and Centroid Locations--Of these three models, VCAL performance was clearly best in calculating the transport speeds and centroid locations of tracer footprints. Although VCAL, as well as TCAL, demonstrated a tendency to overestimate the transport speeds (10% and 40 % for VCAL PTCH and PDCH, respectively; 10% and 180% for TCAL PTCH and PDCH, respectively), the VCAL overestimates exceeded a factor of 2 for only 6 footprint-days, compared to 24 footprint-days for TCAL. In addition, VCAL showed little bias in placing its PTCH and PDCH footprint centroids (mean ratios ~ +20% D_m). TCAL was less consistent, showing no bias for PTCH footprints, but a large positive bias (i.e., its centroids generally were to the south of the measured centroids) for PDCH footprints. TCAL also tended to place its centroids to the right of the actual centroids when the transport speeds were overestimated. SRL tended to both overestimate transport speeds by +40% and place the footprint centroids to the right of the actual centroids.

Mean Trajectory Errors--The mean centroid location errors of SRL and TCAL were greater than any other model. These errors increased linearly with transport time from approximately 350 km to 800 km after 13.5 h and 61.5 h, respectively. On the other hand, the mean centroid location errors for VCAL were among the least, half those of SRL and TCAL.

6.1.2 Conclusions

The SRL transport vectors, based on surface pressure gradients, clearly are biased: speeds tend to be overestimated and directions tend to be to the right of the actual vectors. This bias, as well as its

direction, is not surprising given the nature of geostrophic wind vectors. The high bias in the transport speed can explain SRL's tendency of overestimating the mean concentrations and the frequency of high concentrations, as well as underestimating the number of sites with nonzero concentrations and the lateral diffusion. That is, for a model that overestimates transport speed, the plumes will be narrower and the concentrations will tend to be greater for fixed distances and transport times.

The definition of the height of the mixed layer--the only difference between TCAL (fixed height at 1,500 m AGL) and VCAL (variable height based on potential temperature profiles)--has a large influence on the performance of a single-layer model. This underscores the need to choose carefully the layer through which wind vectors are to be calculated. The low mean centroid location errors for VCAL indicates that a single-layer model can perform as well as the models of the other genres.

The general tendency of TCAL, and to a lesser degree VCAL, to overestimate the transport speeds can explain their tendencies to overestimate the mean concentrations, as a consequence of slower tracer diffusion relative to transport distance. Since wind speeds generally increase with height, the higher TCAL transport speeds could have resulted from a mixed layer height that was too high; climatological data suggest that the 1,500-m height is a factor of 2 too high.

6.2 MULTIPLE-LAYER LAGRANGIAN (MLL) MODELS

6.2.1 Summary

Box Plot Distributions--In general, with the exception of HY-SPLIT, the means, medians, and third quartiles of the MLL models more closely corresponded to those of the measurements than did those of the single-layer Lagrangian models. This was especially true for the PTCH concentrations, where the means, medians, and third quartiles were within ±50% of those for the measurements. Those for the PDCH concentrations were generally greater than those for the measurements by a factor of 2. HY-SPLIT means, medians, and third quartiles tended to be greater than those of the other MLL models, greater than those of the measurements by as much as factors of 3 and 4, and more closely resembled those of the SLL models. During the first half-period, HY-SPLIT medians were comparable to the third quartiles of the measurements.

Frequency Distributions--The frequency distributions of the MLL models generally corresponded more favorably to those of the measurements than did the SLL models. With the exceptions of BAT, MLAM-FINE, and MLAM-COARSE, the MLL model frequencies of concentrations above the thresholds approximated those of the measurements; MLAM-FINE and MLAM-COARSE frequencies tended to be higher by a factor of 2. HY-SPLIT concentrations and MLAM-FINE and MLAM-COARSE PDCH concentrations exceeding 99 dfL/L during the first half-period occurred at least twice as often as those

measured; the opposite was true for ARL PTCH concentrations. ARL and GAMUT distributions of PDCH concentrations were virtually identical to those of the measurements.

Mean Concentrations as a Function of Transport Distance--Although the mean concentrations along the 300-m and 1,000-m bands for the MLL models tended to be lower than the actual means by factors of 2 to 4, the mean concentrations along the bands for the MLL models tended to more closely resemble those of the measurements than did the SLL models. This was especially true at distances farther downwind of the release sites, where the means of all MLL models but MLAM-COARSE were within ±2 dfL/L of the actual means; MLAM-COARSE means generally were high by a factor of 2 to 3 at all distances.

Mean Lateral Diffusion--During the first half-period the mean dispersions of MLAM-FINE and its sibling, MLAM-COARSE, were within ±30% of the actual dispersion. GAMUT and HY-SPLIT dispersions were the lowest of all models, factors of 2 to 3 lower than the actual dispersions. The dispersions for the other MLL models were 50% to 100% lower than the actual dispersion. During the second half-period both BAT and, once again, HY-SPLIT dispersions were the lowest of any model (lower than the actual dispersions by factors of 2 to 3). MLAM-COARSE dispersions were very high, generally double those of the measurements. The dispersions of the remaining MLL models were lower than the actual, but within 30%. For both half- periods, BAT and HY-SPLIT plume widths showed virtually no change with transport distance.

Footprint Transport Speeds and Centroid Locations--Each of the MLL models demonstrated skill in simulating the tracer transport. With two exceptions (i.e., BAT and HY-SPLIT PDCH footprints), mean transport speeds were within an average of 30% of actual mean speeds and mean relative location errors were within 30% of the actual transport distance. Furthermore, only VCAL--a SLL model--had a lower mean absolute location error than the MLL with the greatest error (GAMUT: 417 km).

The performances of half of the MLL models in simulating footprint transport speeds and centroid locations did not vary for each tracer. Performances for the three exceptions--ARL, BAT, and HY-SPLIT--were better for the PTCH data, partly because the PTCH data set tended to be dominated by simpler, northwest flows, while the PDCH data set included a wider range of flow patterns. For example, BAT showed the lowest speed and location errors and no significant biases for the PTCH footprints; however, for the PDCH footprints, BAT speeds strongly tended to be lower than actual speeds and its centroids tended to be to the right of the actual centroids. MLAM-FINE and MLAM-COARSE speeds tended to be lower while GAMUT speeds tended to be greater, but for each model more so for the PDCH footprints. Although MLAM-COARSE showed minimal location errors for both PTCH and PDCH sets, MLAM-FINE locations tended to be to the right of the actual locations for both tracers while GAMUT locations tended to be to the left of the actual PTCH footprints. Both ARL and HY-SPLIT showed the most scatter

for speeds and location errors; speeds tended to be greater than the actual speeds. ARL centroid locations for the PDCH footprints tended to be to the left of the actual centroids.

Mean Trajectory Errors--The six MLL models were divided into two types of behavior. BAT, MLAM-FINE, and MLAM-COARSE mean errors peaked to 400 km after 3.5 days of transport and showed no significant additional increase beyond that. On the other hand, ARL, GAMUT, and HY-SPLIT errors increased more sharply with transport time, reaching 530 ±30 km after 2.5 days and 920 ±160 km after 3.5 days, or about 3 times greater than the other MLL models. Errors decreased to 750 ±100 km after 4.5 days.

6.2.2 Conclusions

The MLL models clearly outperformed all others except VCAL in simulating the transport of tracer footprints. With the possible exceptions of ARL and GAMUT, none of the MLL models clearly outperformed any other in its genre. For most of these models, the majority of the performance measures indicated relatively good performance, but the remaining performance measures indicated biases in the model results. For instance, ARL showed little if any bias in its results as did GAMUT (with the exception of its high bias in the transport speeds), but the mean location errors were relatively great. This indicated that although their mean errors were rather substantial, their centroids, in general, were to the left of the actual centroids as often as they were to the right. The relatively good comparison of its distribution statistics with those of the measurements appears to indicate that ARL and GAMUT simulated rather well the lateral/vertical diffusion; however, they both appeared weak in simulating the transport.

BAT's underestimates of the mean lateral diffusion and the frequencies of occurrence of concentrations above the thresholds would appear to be related to each other. That is, a model that underestimated plume widths will show fewer cases of nonzero concentrations and concentrations above the thresholds. The tendency to overstate the PDCH footprint transport speeds and to place its PDCH footprint centroids to the right of the actual centroids indicated that BAT's vertical mixing could be overstated, effectively giving more influence to the higher-altitude winds, which tend to have greater speeds and directions to the right of the lower-altitude winds. The very low mean centroid location errors, however, indicate that BAT simulates well the transport.

HY-SPLIT's tendencies to understate the plume widths and overstate its third quartiles could be symptomatic of its algorithm for calculating atmospheric stability from the NGM results, as opposed to the algorithm of ARL (its sibling), which interpolates surface and rawinsonde data. Like ARL, HY-SPLIT mean centroid location errors were relatively great, yet no substantial biases were evident in its calculation of footprint speeds and locations. The relatively high-turbulent K_z profiles used by HY-SPLIT may have exaggerated the vertical diffusion, adversely affecting its performance.

MLAM-FINE's footprint transport speeds tended to be lower and to the right of the actual speeds and locations for the first half-period, the only period for which it was applied. These slower speeds could explain its other tendency to overestimate the frequency of concentrations exceeding the thresholds. That is, its footprint widths would be wider relative to transport distance. Furthermore, at any one site the concentrations from one release can be nonzero for two days rather than one day.

For both half-periods, MLAM-COARSE showed the same tendencies as MLAM-FINE. In addition, MLAM-COARSE tended to overstate the frequency of concentrations exceeding the thresholds as well as the mean concentrations at all distance downwind of the release sites. Especially during the second half-period, MLAM-COARSE plume widths tended to be greater than those of the actual widths. All but one of these biases could be explained by the slower MLAM-COARSE transport speeds; the high bias in the mean concentrations could be symptomatic of a low bias in the vertical mixing, causing concentrations near the surface to be biased high.

6.3 MULTIPLE-LAYER EULERIAN (MLE) MODELS

6.3.1 Summary

Box Plot Distributions--The correspondence between model and measurement distributions varied with tracer. The ADOM median and third quartile for PTCH concentrations were 30% to 60% less than those of the measurements, while the opposite was true for the PDCH concentrations. The PTCH box plot distribution for ADPIC was virtually identical to those of the measurements; the ADPIC median and third quartile for PDCH concentrations were within 60% of those for the measurements.

Frequency Distributions--The ADOM distributions for the first half-period corresponded very closely to those of the measurements; ADOM was not applied for the second half-period. The comparisons of the ADPIC with the actual distributions were inconsistent; while ADPIC distributions were quite similar to the actual distributions for PTCH during the first half-period and PDCH concentrations for the second half-period, the ADPIC frequencies for concentrations exceeding the thresholds deviated by 75% for the remaining half-periods for each tracer.

Mean Concentrations as a Function of Transport Distance--The mean PTCH concentrations for three bands of sites for both ADOM and ADPIC tended to be lower than those of the measurements. Ratios of calculated-to-predicted means for all three bands ranged from 0.4 to 0.9. However, the opposite was true for first-period PDCH concentrations; these ratios approximated 1.8 for both ADOM and ADPIC. Second-period ADPIC mean PDCH concentrations were lower by an average of 30%.

Mean Lateral Diffusion--During the first half-period the lateral dispersion of both models was generally greatest of all the models and greater than the actual dispersion an average of 10% for ADOM

and 50% for ADPIC. During the second half-period, ADPIC dispersion for PTCH was low by factors of 2 to 3 for some bands and varied little with transport distance; however, its PDCH dispersion was within ±10% of the actual dispersion at all distances.

Footprint Transport Speeds and Centroid Locations--ADOM speeds for both tracers tended to be lower than actual speeds by 20% to 40%. For its largest location errors, ADOM speeds were greater than actual speeds by at least a factor of 2 and its centroids tended to be to the left of the actual centroids. However, in general, its centroids tended to be to the right. Similarly, ADPIC speeds tended to understate the actual speeds, sometimes by as much as 60% and 80%; ADPIC centroids for the PDCH footprints were to the right of the actual centroids in nearly every case and to the left for the PTCH footprints.

Mean Trajectory Errors--Both ADOM and ADPIC mean errors were among the greatest of all models. The ADOM mean error after 1.5 days was low, approximately 200 km, but quickly increased to 600 km after 2.5 days (greatest of all models), then decreased to 500 km after 3.5 days. Similarly, the ADPIC mean error increased sharply from 300 km at 1.5 days to 850 km after 2.5 days, among the greatest.

6.3.2 Conclusions

In general, the MLE models performed quite similarly and better for the ensemble measures than they did for the footprint comparison measures. This implies that these two models performed relatively well for the average, but performed relatively poorly for individual cases. The only substantial bias observed in the ensemble measures was for lateral diffusion; both ADOM and ADPIC tended to overstate the footprint widths in the first half-period, the only period for which ADOM was applied. However, both models tended to understate the footprint speeds, which could by itself explain the high bias in lateral diffusion, as well as the large mean centroid location errors.

The strong relationship between the large ADOM centroid location errors and overestimated transport speed errors could indicate a problem with its vertical diffusion for several cases (PTCH-15, PDCH-4, -10, and -15), all of which were intercepted by cyclones or fronts. That is, the model could have overestimated vertical diffusion and, as a consequence, relied more on the faster wind speeds at higher levels.

The strong ADPIC tendency to understate the transport speeds and place footprint centroids to the right of the actual footprint centroids demonstrates its weakness in simulating the transport for the 27 footprints of this study. Additional data are needed to substantiate this conclusion. The reason for relatively few ADPIC footprints related to the fact that ADPIC concentrations very often did not return to zero days after actual tracer footprints were transported across regions.

SECTION 7

RECOMMENDATIONS

The results presented in this report are the product of the first step in evaluating these long-range transport and diffusion models. It is recommended that the developers of these models continue the evaluation process by using the ANATEX data and the AMES results to better understand the behavior of their models. This can be accomplished by sensitivity analyses in tandem with comparisons of calculated and measured concentrations. These sensitivity analyses should focus on the changes in model calculations and performances with respect to changes in such model variables as vertical mixing, horizontal diffusion, and transport speeds and directions. Publication of these latter results is strongly recommended.

It is also recommended that the ANATEX aircraft and tower data be used to evaluate the multiple-layer models with respect to the vertical distribution of tracer and the effects of vertical wind shears on the tracer plumes. Regrettably, these data were unavailable at the time this study was conducted.

Finally, it is strongly recommended that the sampling protocol for future tracer studies include aircraft sampling at greater distances downwind of the release site than the 0-300 km of ANATEX. (The problem of prescribing the sampling flight patterns in a prognostic fashion is recognized.) Vertical distribution of tracer at greater distances must be known to identify specific causes of model errors. Without these distributions supplementing the ground-base measurements we can only state that there appears to be a model error, but we would not be equipped to relate the error to a specific modeling process.

REFERENCES

Atchison, M.K. and C.R. Parks. 1986. Evaluation of a single-layer trajectory and diffusion model using the Metropolitan Tracer Experiment (METREX) data base. Fifth Joint Conference on Applications of Air Pollution Meteorology, Chapel Hill, NC, November 1986.

Businger, J.A. 1973. Turbulent transfer in the atmospheric surface layer. *Workshop in Micrometeorology*, American Meteorological Society, Boston, MA. pp. 67-100.

Businger, J.A., J.C. Wyngaard, Y. Izumi, and E.F. Bradley. 1971. Flux-profile relationships in the atmospheric surface layer. *J. Atmos. Sci.*, **28**:181-189.

Businger, J.A. and S.P.S. Arya. 1974. Height of the mixed layer in the stably stratified planetary boundary layer. *Advances in Geophysics*, **18A**, Academic Press. pp. 73-92.

Byers, H.R. 1974. *General Meteorology*, McGraw-Hill Book Company, New York, NY. 461 pp.

Davis, W.E., A.R. Olsen, and B.T. Didier. 1989. MLAM assessment of radionuclide air concentration and deposition for the Chernobyl Reactor accident. *Air Pollution Modeling and Its Application VII.* H. van Dop, Editor. Plenum Press, New York, NY. pp. 123-136.

Draxler, R.R. 1982. Measuring and modeling the transport and dispersion of Krypton-85 1,500 km from a point source. *Atmos. Environ.*, **16**:2763-2776.

Draxler, R.R. 1987. Sensitivity of a trajectory model to the spatial and temporal resolution of the meteorological data during CAPTEX. *J. of Clim. and Appl. Meteor.*, **26**:1577:1588.

Draxler, R.R. 1988a. Across North America Tracer Experiment (ANATEX) weather maps and tracer concentrations. NOAA Technical Memorandum ERL ARL-165, Air Resources Laboratory, Silver Spring, MD. 87 pp.

Draxler, R.R. 1988b. HYBRID Single-Particle Lagrangian Integrated Trajectories (HY-SPLIT): model description. NOAA Technical Memorandum ERL ARL-166, Air Resources Laboratory, Silver Spring, MD. 23 pp.

Draxler, R.R. and J.L. Heffter. 1989. Across North America Tracer Experiment (ANATEX), Volume I: Description, ground-level sampling at primary sites, and meteorology. NOAA Technical Memorandum ERL ARL-167, Air Resources Laboratory. Silver Spring, MD. 83 pp.

Eimutis, E.C. and M.G. Konicek. 1972. Derivations of continuous functions for the lateral and vertical atmospheric dispersion coefficients. *Atmos. Environ.*, **6**:859-863.

Ferber, G.J., J.L. Heffter, R.R. Draxler, R.J. Lagomarsino, F.L. Thomas, R.N. Dietz, and C.M. Benkovitz. 1986. Cross-Appalachian Tracer Experiment (CAPTEX '83) Final Report. Technical Memorandum ERL ARL-142, National Oceanic and Atmospheric Administration, Silver Spring, MD. 60 pp.

Gifford, F.A. 1987. The time-scale of atmospheric diffusion considered in relation to the universal diffusion function, f_1. *Atmos. Environ., 21:1315-1320.*

Godowitch, J.M. 1989. Evaluation and sensitivity analyses results of the MESOPUFF II model with CAPTEX measurements. EPA-600/3-89/056, NTIS No. PB 89-198253, U.S. Environmental Protection Agency, Research Triangle Park, NC. 99 pp.

Haagenson, P.L., Y-H Kuo, M. Skumanich, and N.L. Seaman. 1987. Tracer verification of trajectory models. *J. of Clim. and Appl. Meteor.*, **26**:410-426.

Heffter, J.L. 1965. The variation of horizontal diffusion parameters with time for travel periods of one hour or longer. *J. Appl. Meteor.*, **4**:153-156.

Heffter, J.L. 1980. Transport layer depth calculations. Second Joint Conference on Applications of Air Pollution Meteorology, New Orleans, LA, March 24-27, American Meteorological Society, Boston, MA. pp. 787-791.

Heffter, J.L. 1983. Branching Atmospheric Trajectory (BAT) Model. NOAA Technical Memorandum ERL ARL-121, Air Resources Laboratory, Silver Springs, MD. 19 pp.

Holzworth, G.C. 1972. Mixing heights, wind speeds, and potential for urban air pollution throughout the contiguous United States. Technical Report AP-101, U.S. Environmental Protection Agency, Research Triangle Park, NC. 118 pp.

Kahl, J.D. and P.J. Samson. 1988. Trajectory sensitivity to rawinsonde data resolution. *Atmos. Environ.*, **22**:1291-1299.

Kao, C.J. and T. Yamada. 1988. Use of CAPTEX data for evaluations of a long-range transport numerical model with a four-dimensional data assimilation technique. *Mon. Wea. Rev.*, **116**:293-306.

Lange, R. 1978. ADPIC--a three-dimensional particle-in-cell model for the dispersal of atmospheric pollutants and its comparison to regional tracer studies. *J. Appl. Meteor.*, **17**:320-329.

Lee, I.J. 1987. Numerical simulations of cross-Appalachian transport and diffusion. *Bound. Layer Meteor.*, **39**:53-66.

Masters, S.E. and M.A. Kienzle. 1989. The impact of vertical mixing on large-scale atmospheric transport. Sixth Joint Conference on Applications of Air Pollution Meteorology, Anaheim, CA, January 1989.

O'Brien, J.J. 1970. A note on the vertical structure of the eddy exchange coefficient in the planetary boundary layer. *J. Atmos. Sci.*, **27**:1213-1215.

Olson, M.P., K.J. Puckett, D. Davies, and K.K. Oikawa. 1989. A comparison between the ADOM model and the ANATEX tracer data. San Francisco, CA, November 1989.

Pasquill, F. 1976. Atmospheric dispersion parameters in Gaussian plume modeling: Part II. Possible requirements for change in the Turner Workbook values. EPA-600/4-76-030b, U.S. Environmental Protection Agency, Research Triangle Park, NC.

Persson, C., R. Henning, and L. DeGeer. 1986. The Chernobyl accident--a meteorological analysis of how radionuclides reached Sweden. SMHI Reports, *Meteorology and Climatology*, No. 55, X-60176, Norrköping, Sweden.

Rodriquez, D.J. and K.R. Peterson. 1989. Simulating the venting of radioactivity from a Soviet nuclear test. *Atmos. Environ.*, **23**:953-958.

Scholtz, M.T., B. Weisman, L. Mahrt, and A.D. Christie. 1986. Generation of meteorological data fields for the ADOM Eulerian regional model. Fifth Joint Conference on Applications of Air Pollution Meteorology, Chapel Hill, NC, November 18-21, American Meteorological Society, Boston, MA. pp. 145-150.

Szepesi, D.J. 1989. *Compendium of Regulatory Air Quality Models.* Akademiai Kiado es Nyomda Vallalat, Budapest, Hungary. ISBN 963 05 46752. 516 pp.

Venkatram, A., P.K. Karamchandani, and P.K. Misra. 1988. Testing a comprehensive acid deposition model. *Atmos. Environ.*, **22**:737-747.

Yamartino, R.J. and J.S. Scire. 1984. ADOM/TADAP Model Development Program Volume 3: The transportation and diffusion modules. ERT Document No. P-B980-210, ENSR Inc., 1220 Avenida Acaso, Camarillo, CA 93010.

APPENDIX A

MODEL DESCRIPTIONS

TABLE A-1. SRL ATTRIBUTES

Savannah River Laboratory Adjusted Geostrophic Model (SRL)	
Model genre:	single-layer segmented plume
Reference(s):	none
Resolution:	
Horizontal--	variable
Vertical--	single layer[a]
Temporal--	3 h
Advection:	
Diffusion:	
Horizontal--	based on curve fits to empirical and theoretical plume growth (Gifford, 1987)
Vertical--	
Met. parameters:	
Mixing depth--	function of z_0, u, and Pasquill-Turner stability class
Stability--	
Eddy diffusivity--	
Winds--	*u,v* calculated from 3-h surface *p* fields generated by an objective scheme forcing temporal continuity; *w* estimated from terrain slope wind direction, *T* advection, and positive vorticity advection

[a] Can be applied using two layers.

TABLE A-2. TCAL AND VCAL ATTRIBUTES

Trajectory Calculation Model (TCAL) and Variable Layer Trajectory Calculation Model (VCAL)	
Model genre:	single-layer Lagrangian puff
Reference(s):	Atchison, M.K. and C.R. Parks (1986)
Resolution:	
Horizontal--	rawinsonde network resolution (~ 300 km)
Vertical--	single layer
Temporal--	1 h
Advection:	layer-mean winds; TCAL: 300-1,500 m AGL; VCAL: top determined by local mixing height
Diffusion:	
Horizontal--	Gaussian
Vertical--	none
Met. parameters:	
Mixing depth--	TCAL: constant at 1,500 m AGL; VCAL: potential temperature technique based on vertical changes in the potential temperature; concentrations based on a fixed depth of 1,500 m AGL.
Stability--	N/A
Eddy diffusivity--	N/A
Winds--	Primary source: rawinsonde data; Secondary source: AFGWC HIRAS data interpolated via Cressman technique; no surface data

TABLE A-3. ARL ATTRIBUTES

Air Resources Laboratory Model (ARL)	
Model genre:	multilayer Lagrangian puff
Reference(s):	Draxler (1987)
Resolution:	
Horizontal--	200 km
Vertical--	10 layers; 0-3,000 m AGL
Temporal--	1 h for trajectories; 3 h for meteorology
Advection:	
Horizontal--	first-order approximation
Vertical--	conservation of θ_w; puffs split in the vertical dimension when they overlap more than one grid cell
Diffusion:	
Horizontal--	$d\sigma_h/dt = 1,853$
Vertical--	$(d\sigma_z)^2/dt = 2K_z$
Met. parameters:	
Mixing depth--	height of the base of first elevated inversion (Heffter, 1980)
Stability--	surface layer flux relations of Businger (1973)
Eddy diffusivity--	interpolation formula of O'Brien (1970)
Winds--	spatially and temporally interpolated surface and rawinsonde data; vertical velocity based on conservation of θ_w (Byers, 1974)

TABLE A-4. BAT ATTRIBUTES

Branching Atmospheric Trajectory Model (BAT)	
Model genre:	multilayer Lagrangian puff
Reference(s):	Heffter (1983), Szepesi (1989)
Resolution: [a]	
Horizontal--	rawinsonde network resolution (~ 300 km)
Vertical--	3 layers (300 m; critical inversion; 3,000 m)
Temporal--	3 h for advection and 1 h for diffusion
Advection:	
Horizontal--	Modified Eulerian: $\Sigma\ w_i\ \Delta t\ d_i^{-2}\ /\ \Sigma\ d_i^{-2}$
Vertical--	N/A
Diffusion:	
Horizontal--	Gaussian: $\sigma_h = 0.5\ t$
Vertical--	uniform mixing within each layer; branching puffs at day/night transitions (0300 GMT and 1500 GMT)
Met. parameters:	
Mixing depth--	where $\Delta\theta/\Delta z \geq 0.005$ K/m and $\theta_T - \theta_B \geq 2$ K
Stability--	N/A
Eddy diffusivity--	N/A
Winds--	6-h temporally interpolated rawinsonde data at sites (no spatial interpolation)

[a] No grid is used.

TABLE A-5. GAMUT ATTRIBUTES

Global Atmospheric Multilayer Transport Model (GAMUT)	
Model genre:	multilayer Lagrangian puff
Reference(s):	Masters, S.E. and M.A. Kienzle (1989)
Resolution:	
Horizontal--	280 km
Vertical--	4 layers between 0 and 3,000 m AGL
Temporal--	3 h
Advection:	
Horizontal--	layer-mean winds; predictor-corrector advection method
Vertical--	stable lofting/sinking using moist static energy as air mass tracer; vertical mixing between layers according to computed depth
Diffusion:	
Horizontal--	Gaussian, $d\sigma_h/dt$ = 1.0 nmi/h
Vertical--	trajectory splitting between layers controlled by mixing depth (may reach 100 active trajectories/release); mass loss to free troposphere by stable lofting & convective mixing
Met. parameters:	
Mixing depth--	potential instability technique; mixing depth is the level of the "critical inversion" identified by the moist-adiabatic lapse rate
Stability--	N/A
Eddy diffusivity--	N/A
Winds--	Primary source: rawinsonde data; Secondary source: AFGWC HIRAS data over most areas at 0600 and 1800 GMT; over oceans and areas/times of missing rawinsonde data, no surface data

TABLE A-6. HY-SPLIT ATTRIBUTES

Hybrid Single-Particle Lagrangian Integrated Trajectories Model (HY-SPLIT)	
Model genre:	multilayer Lagrangian--Eulerian hybrid
Reference(s):	Draxler (1988b)
Resolution:	
Horizontal--	180 km
Vertical--	8 layers up to 0.721 of surface pressure
Temporal--	1 h for trajectory calculations; 2 h for meteorological data
Advection:	first-order approximation; three dimensions; puffs split in horizontal and vertical when diffusive growth exceeds grid dimensions
Diffusion:	
Horizontal--	$d\sigma_h / dt = 1{,}853$
Vertical--	$(d\sigma_z)^2 / dt = 2 \cdot K_z$ where χ is uniform within plume of radius $1.54\ \sigma_h$ and depth of $3.08\ \sigma_z$
Met. parameters:[a]	
Mixing depth--	calculated by the NGM
Stability--	calculated from NGM boundary layer variables (u_* , T_v , θ_*)
Eddy diffusivity--	calculated from NGM boundary layer variables; $K_z = 30 \cdot (1 + \lvert R_i \rvert)$; minimum of 1 m^2/s
Winds--	near surface: normalized wind profiles (similar to Businger *et al.*, 1971); $\varphi_m = (1 - 16\ z/L)^{-0.25}$, for $L < 0$ and $\varphi_m = 1 + 5\ z/L$, for $L \geq 0$; at or above NGM σ-level 1: calculated by the NGM

[a] Gridded analytical and prognostic output from NOAA's Nested Grid Model (NGM).

TABLE A-7. MLAM (FINE AND COARSE) ATTRIBUTES

Multilayer Air Mass Model (MLAM)	
Model genre:	multilayer Lagrangian puff
Reference(s):	Davis *et al.* (1989)
Resolution:	
Horizontal--	190 km (user defined)
Vertical--	9 layers with tops of: 100; 400; 700; 1,000; 1,500; 2,000; 2,600; 3,400; and 4,400 m (user defined)
Temporal--	1 h
Advection:	three-dimensional; based on gridded fields of θ_T, u, v, w
Diffusion:	
Horizontal/ Vertical	<100 km: function of distance, σ_θ and ΔT; based on regression analysis (Eimutus and Konicek, 1972) ≥100 km: $\sigma_y^2 / 2t$ (Heffter, 1965)
Met. parameters:	
Mixing depth--	diurnal; maximum depth at 1600 LST
Stability--	hourly; function of t (LST)
Eddy diffusivity--	not used
Winds--	site measurements or gridded
Vertical redistribution:	redistribution of mass twice daily (**FINE**: 1000 LST and 1600 LST) <u>or</u> once daily (**COARSE**) during the first 3 days to create new sets of puffs

TABLE A-8. ADOM ATTRIBUTES

Acid Deposition and Oxidant Model (ADOM)	
Model genre:	multilayer Eulerian (no-cloud version)
Reference(s):	Venkatram et al. (1988); Olson et al. (1989)
Resolution:	
Horizontal--	~ 100 km
Vertical--	11 unevenly-spaced layers from 0 to 10 km with greater resolution in the planetary boundary layer
Temporal--	1 h
Advection:	numerical solution of advection-diffusion equation using Blackman cubic spline (Yamartino and Scire, 1984)
Diffusion:	numerical solution of advection-diffusion equation using the Smagorinsky scheme
Met. parameters:	derived from a diagnostic module combining information from the Canadian Meteorological Centre's large-scale numerical weather prediction model (upper-boundary conditions) with information from a high-resolution boundary layer model (Scholtz et al., 1986)

TABLE A-9. ADPIC ATTRIBUTES

Atmospheric Diffusion Particle-In-Cell Model (ADPIC)	
Model genre:	multiple-layer, Lagrangian-Eulerian, particle-in-cell
Reference(s):	Lange (1978); Rodriguez and Peterson (1989)
Resolution:	
Horizontal--	381 km
Vertical--	150 m
Temporal--	dependent on the maximum wind speed across the model domain
Advection:	flux-conservative form of the advection-diffusion equation: $\partial\chi/\partial t + \nabla_o [\chi (U_A + U_D)] = 0,$ where χ is a scalar concentration, U_A is an advection velocity, and U_D ($= -K\nabla\chi/\chi$) is a diffusivity velocity
Diffusion:	
Horizontal--	$K_h = \sigma_y \cdot (d\sigma_y/dt)$, where $\sigma_y = a\ (ut)^b$
Vertical--	$K_z = \sigma_z \cdot (d\sigma_z/dt)$, where $\sigma_z = c\ (ut)^d$ (Pasquill, 1976)
Met. parameters:	u- and v-wind components on a 47 x 51 x 15 grid based on the polar stereographic projection and vertical interpolation of AFGWC's HIRAS data

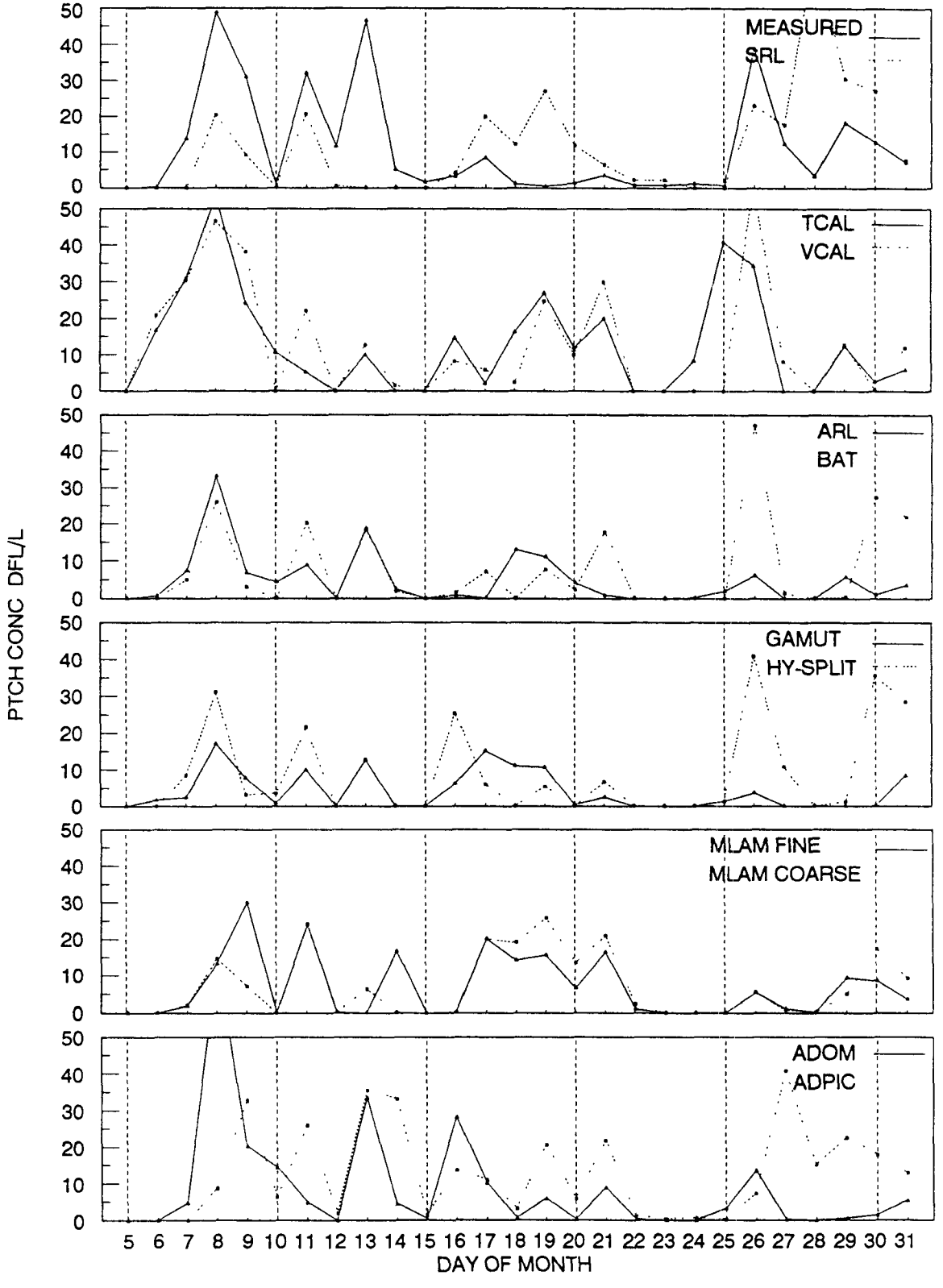
APPENDIX B

TIME SERIES PLOTS OF 24-H MEAN CONCENTRATIONS OF PTCH AND PDCH TRACERS ALONG BANDS OF SITES

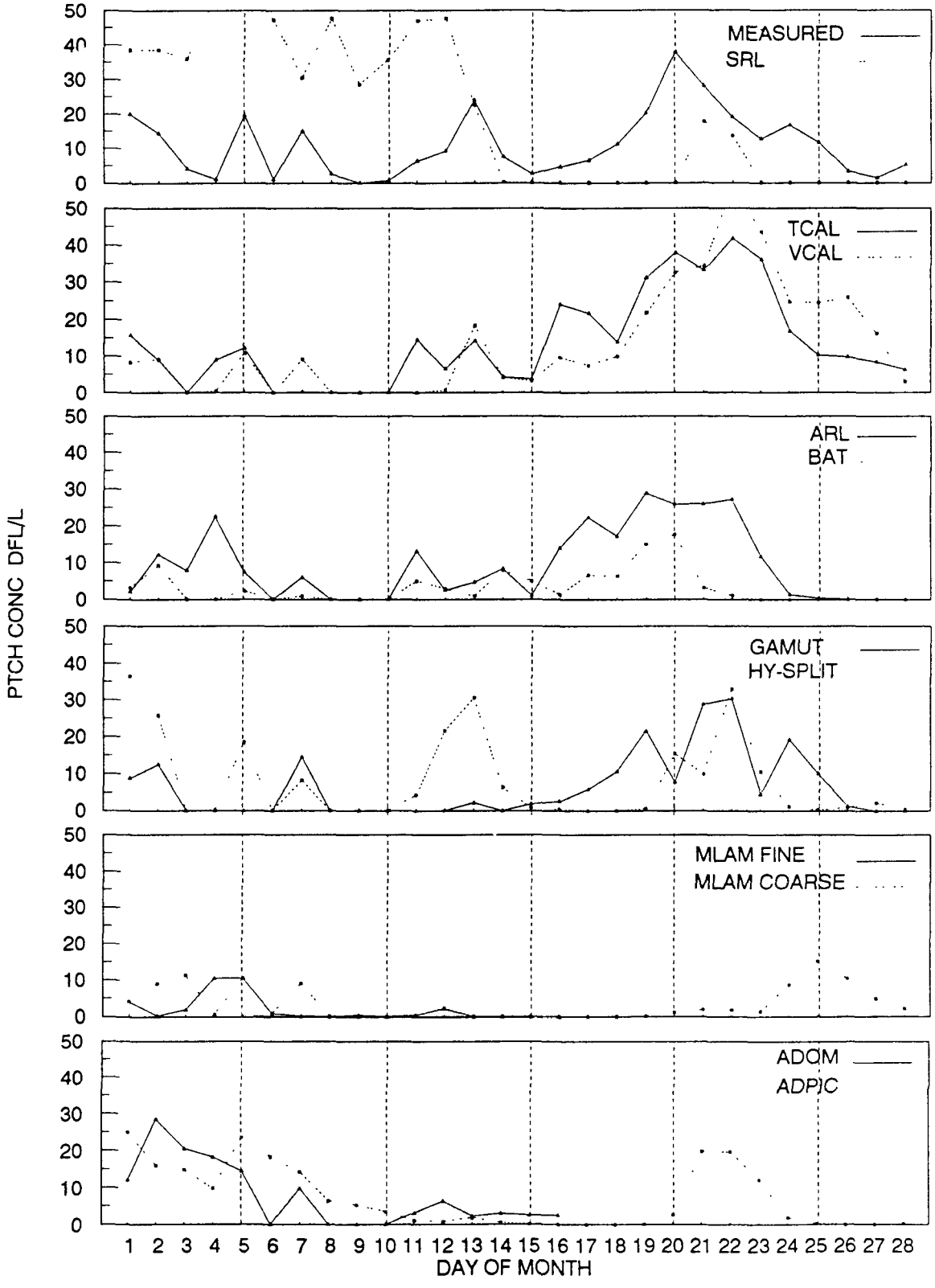
LIST OF FIGURES

Band (km)	Tracer	Month	Page no.	Band (km)	Tracer	Month	Page no.
1,000	PTCH	Jan	B-2	700E	PDCH	Jan	B-17
"	"	Feb	B-3	"	"	Feb	B-18
"	"	Mar	B-4	"	"	Mar	B-19
1,600	"	Jan	B-5	1,000W	"	Jan	B-20
"	"	Feb	B-6	"	"	Feb	B-21
"	"	Mar	B-7	"	"	Mar	B-22
2,300	"	Jan	B-8	1,000E	"	Jan	B-23
"	"	Feb	B-9	"	"	Feb	B-24
"	"	Mar	B-10	"	"	Mar	B-25
300	PDCH	Jan	B-11	1,400	"	Jan	B-26
"	"	Feb	B-12	"	"	Feb	B-27
"	"	Mar	B-13	"	"	Mar	B-28
700W	"	Jan	B-14	1,800	"	Jan	B-29
"	"	Feb	B-15	"	"	Feb	B-30
"	"	Mar	B-16	"	"	Mar	B-31

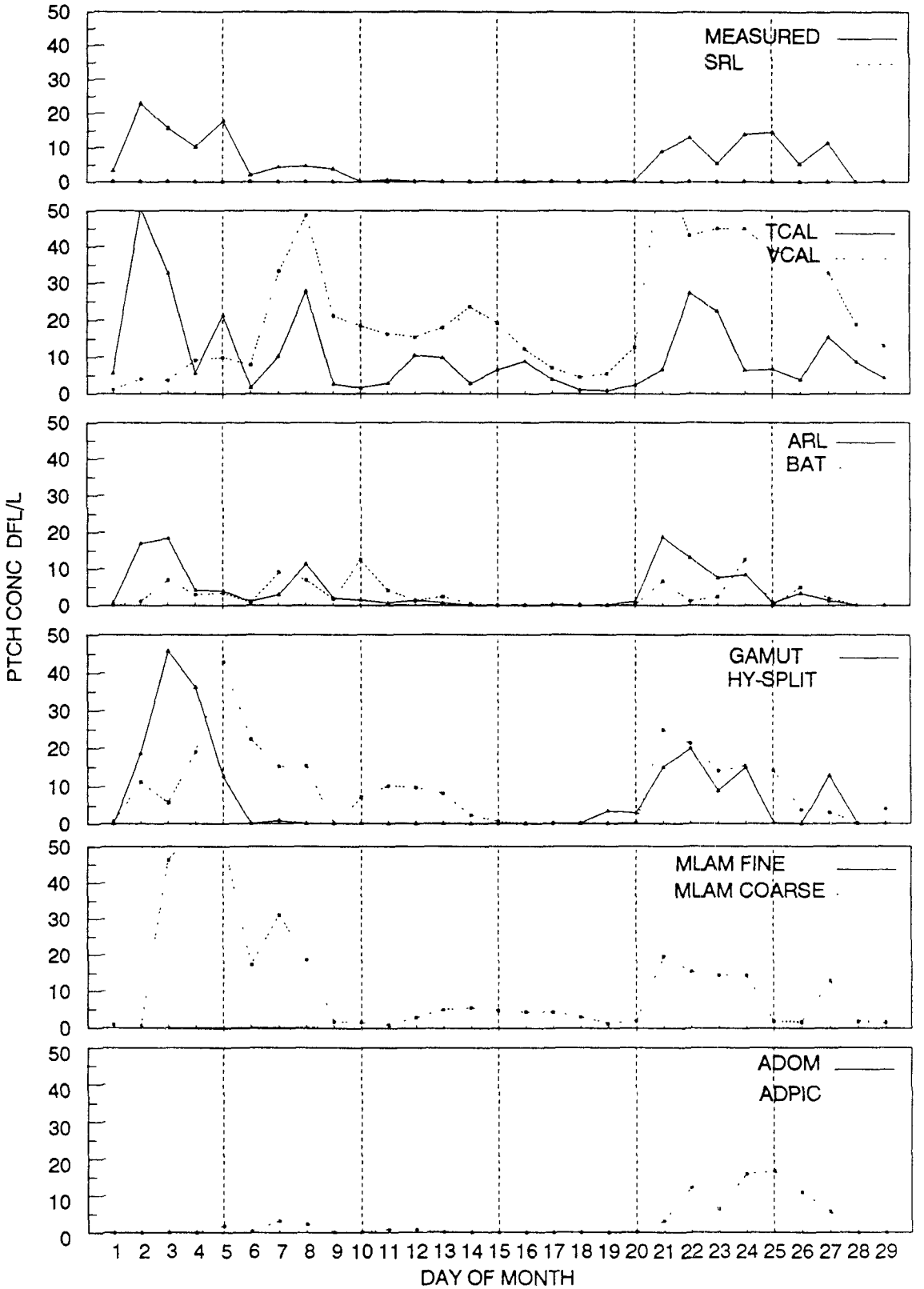
24-H AVERAGE CALCULATED AND MEASURED PTCH CONCENTRATIONS
ALONG THE 1000 KM BAND
JANUARY 1987



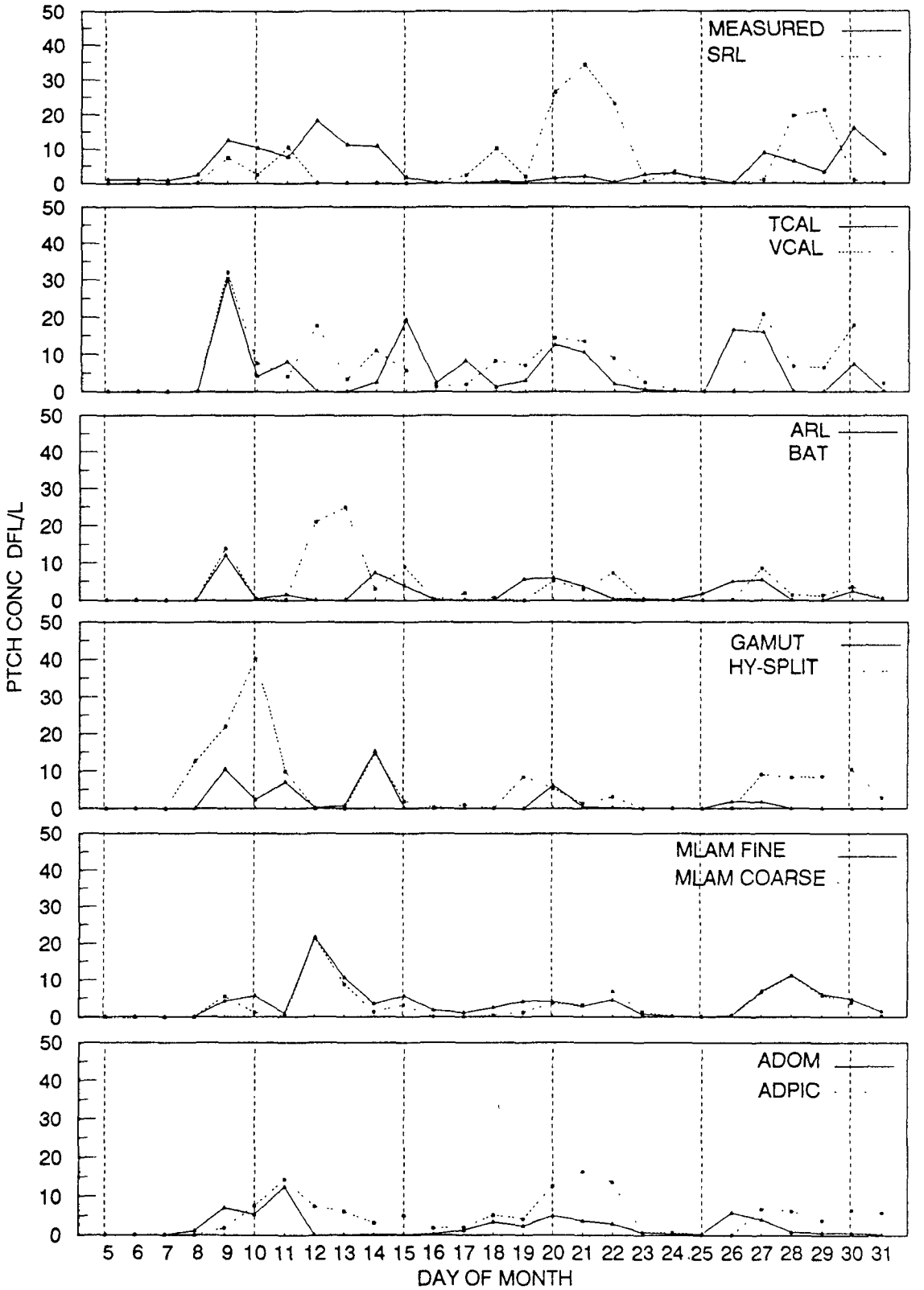
24-H AVERAGE CALCULATED AND MEASURED PTCH CONCENTRATIONS
ALONG THE 1000 KM BAND
FEBRUARY 1987



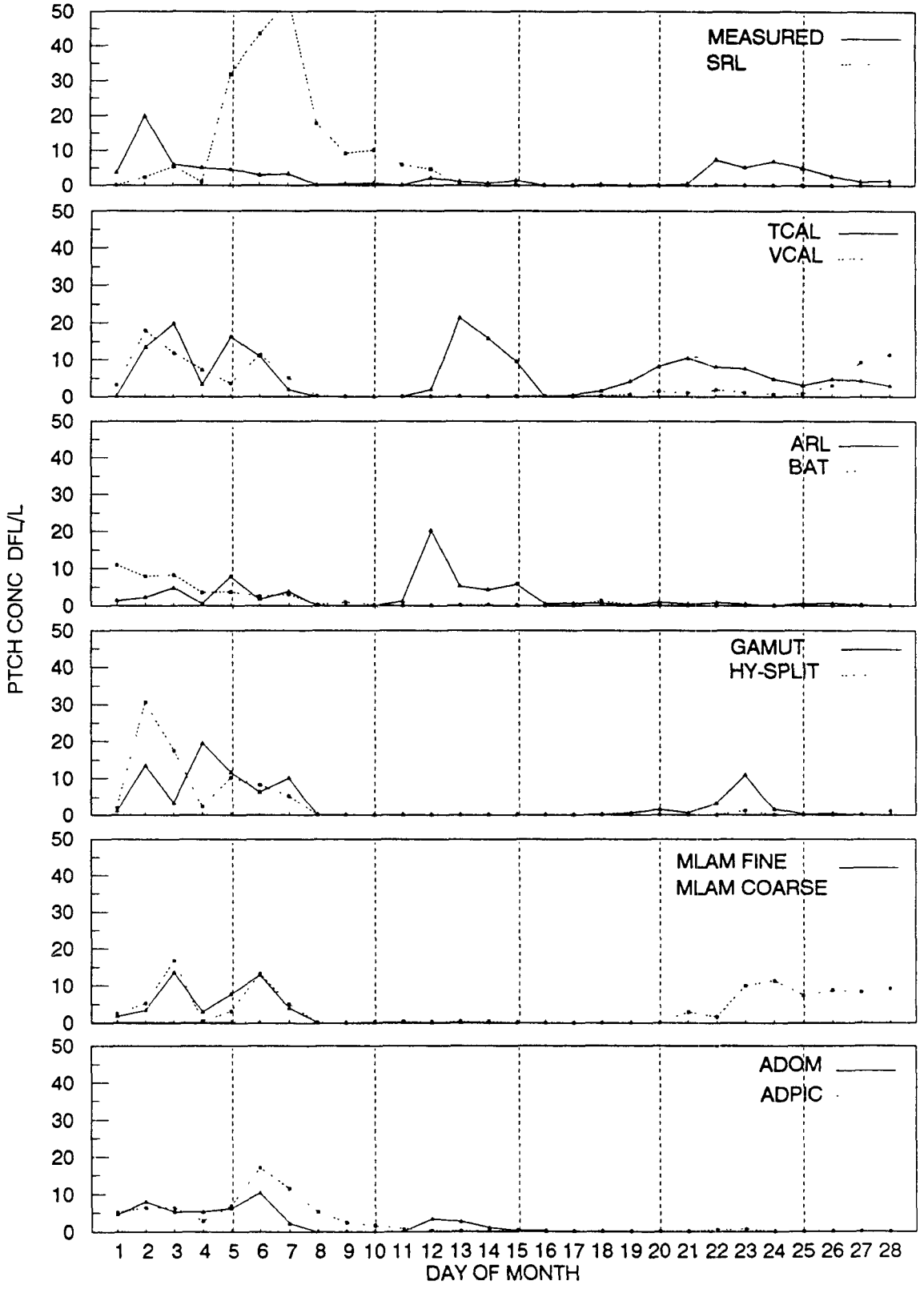
24-H AVERAGE CALCULATED AND MEASURED PTCH CONCENTRATIONS
ALONG THE 1000 KM BAND
MARCH 1987



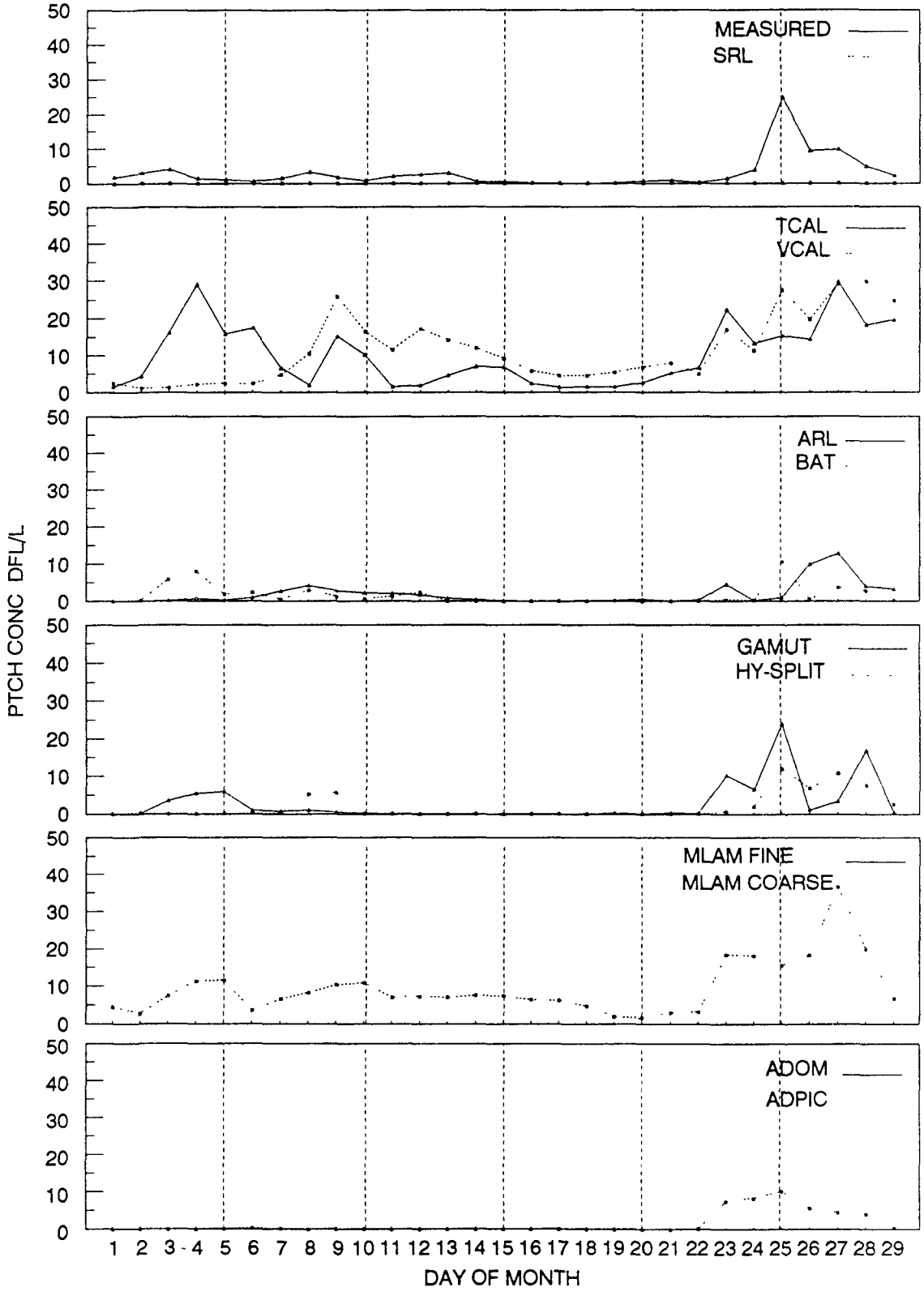
24-H AVERAGE CALCULATED AND MEASURED PTCH CONCENTRATIONS
ALONG THE 1600 KM BAND
JANUARY 1987



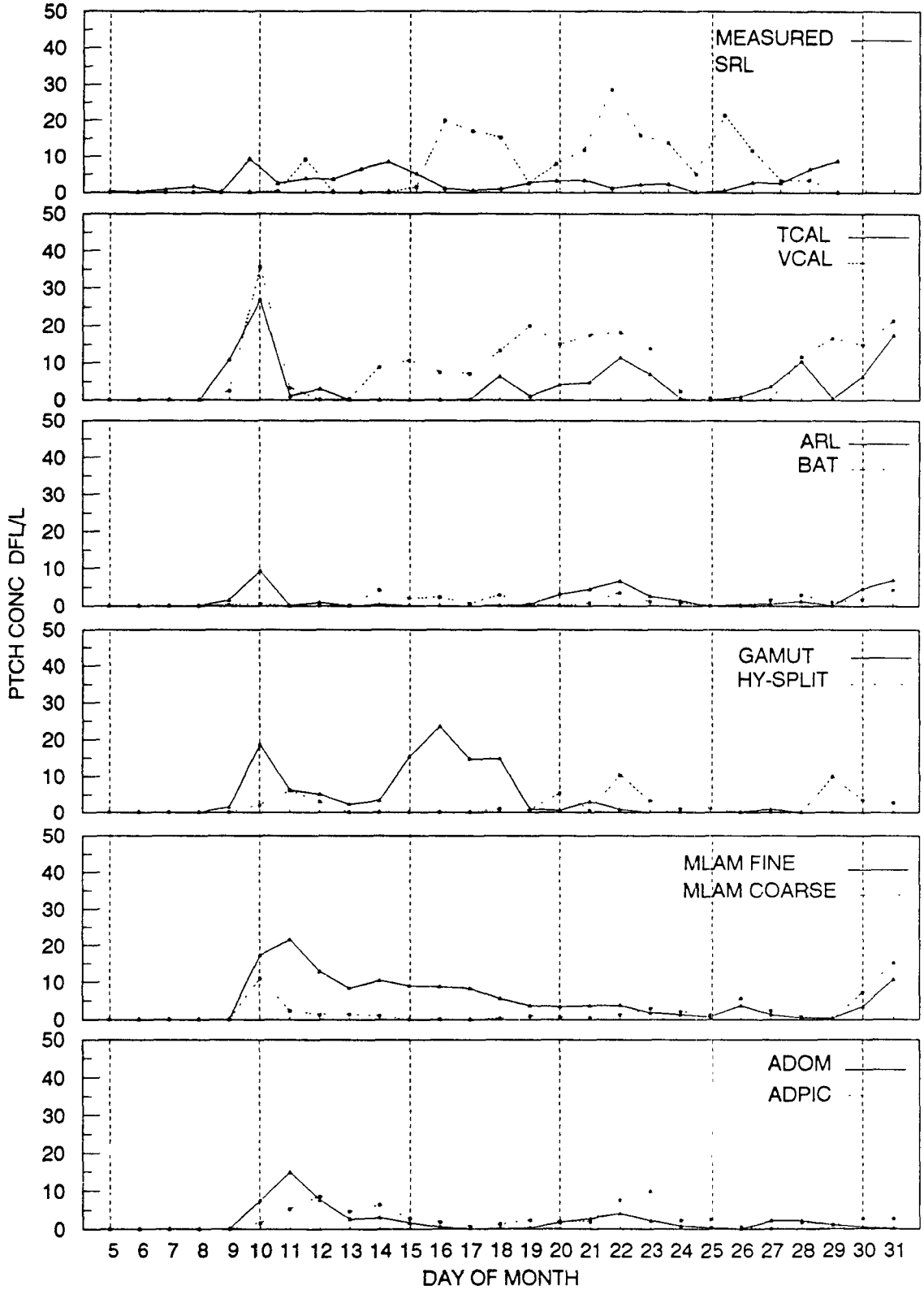
24-H AVERAGE CALCULATED AND MEASURED PTCH CONCENTRATIONS
ALONG 1600 KM BAND
FEBRUARY 1987



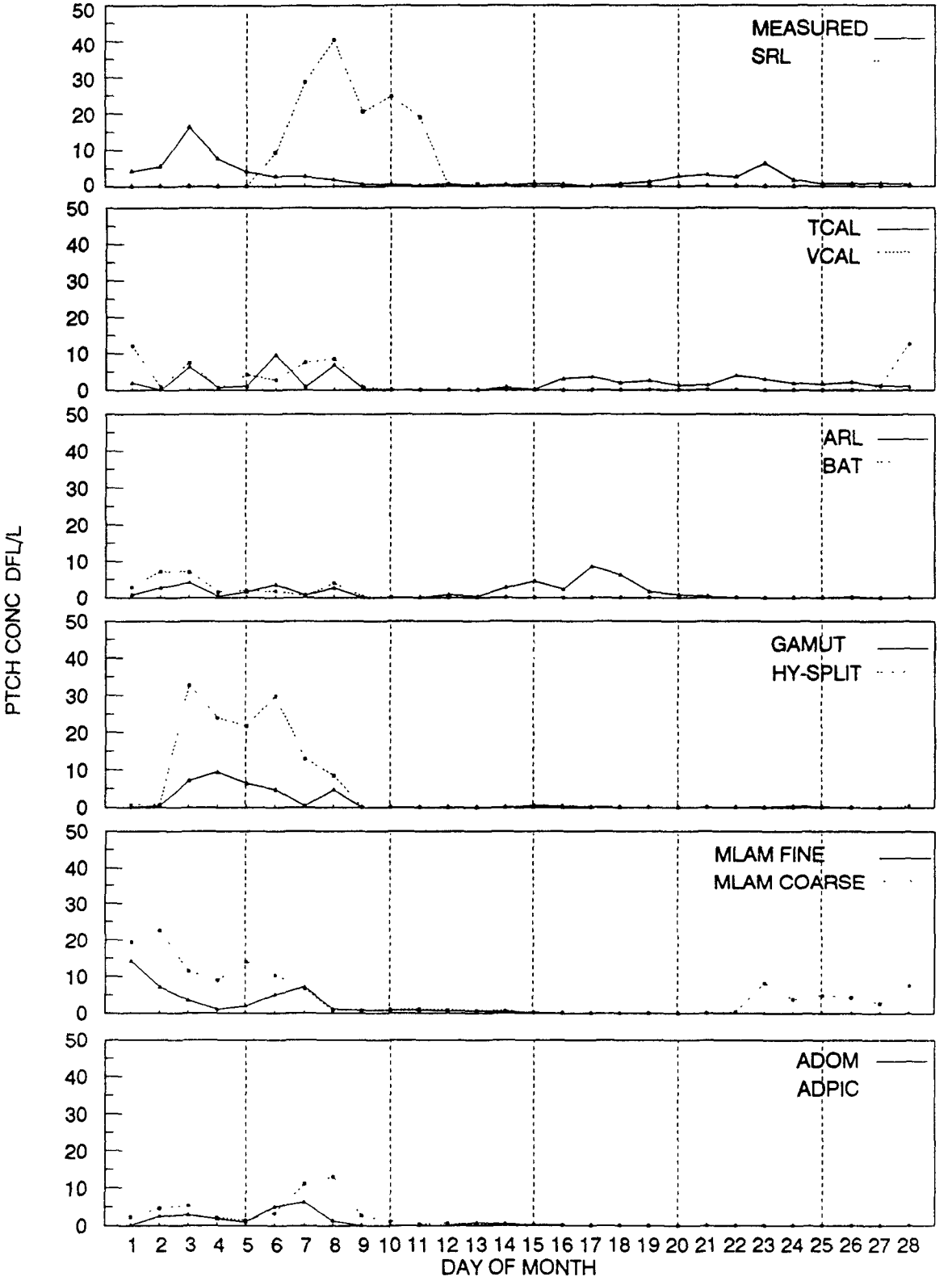
24-H AVERAGE CALCULATED AND MEASURED PTCH CONCENTRATIONS
ALONG 1600 KM BAND
MARCH 1987



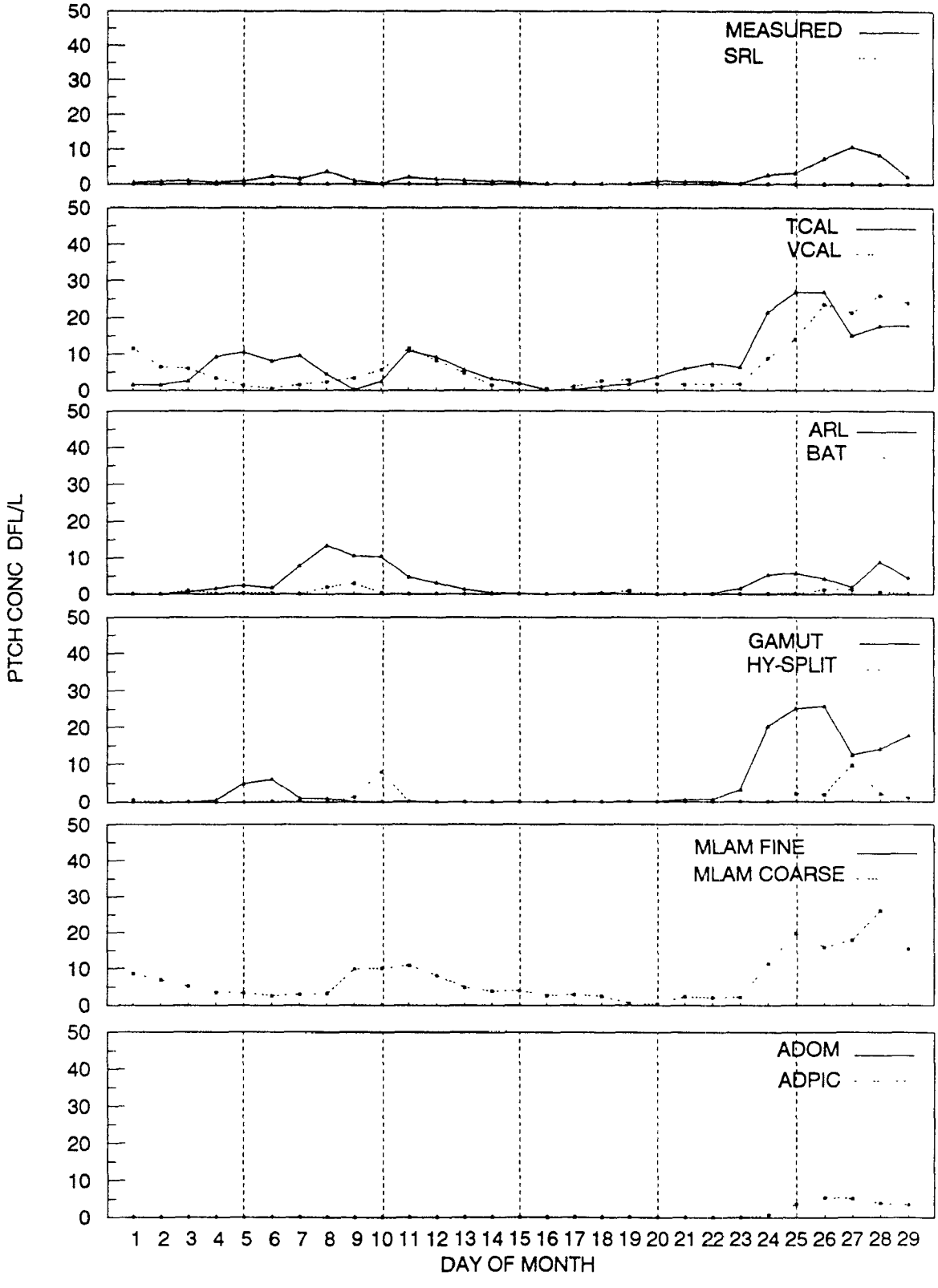
24-H AVERAGE CALCULATED AND MEASURED PTCH CONCENTRATIONS
ALONG THE 2300 KM BAND
JANUARY 1987



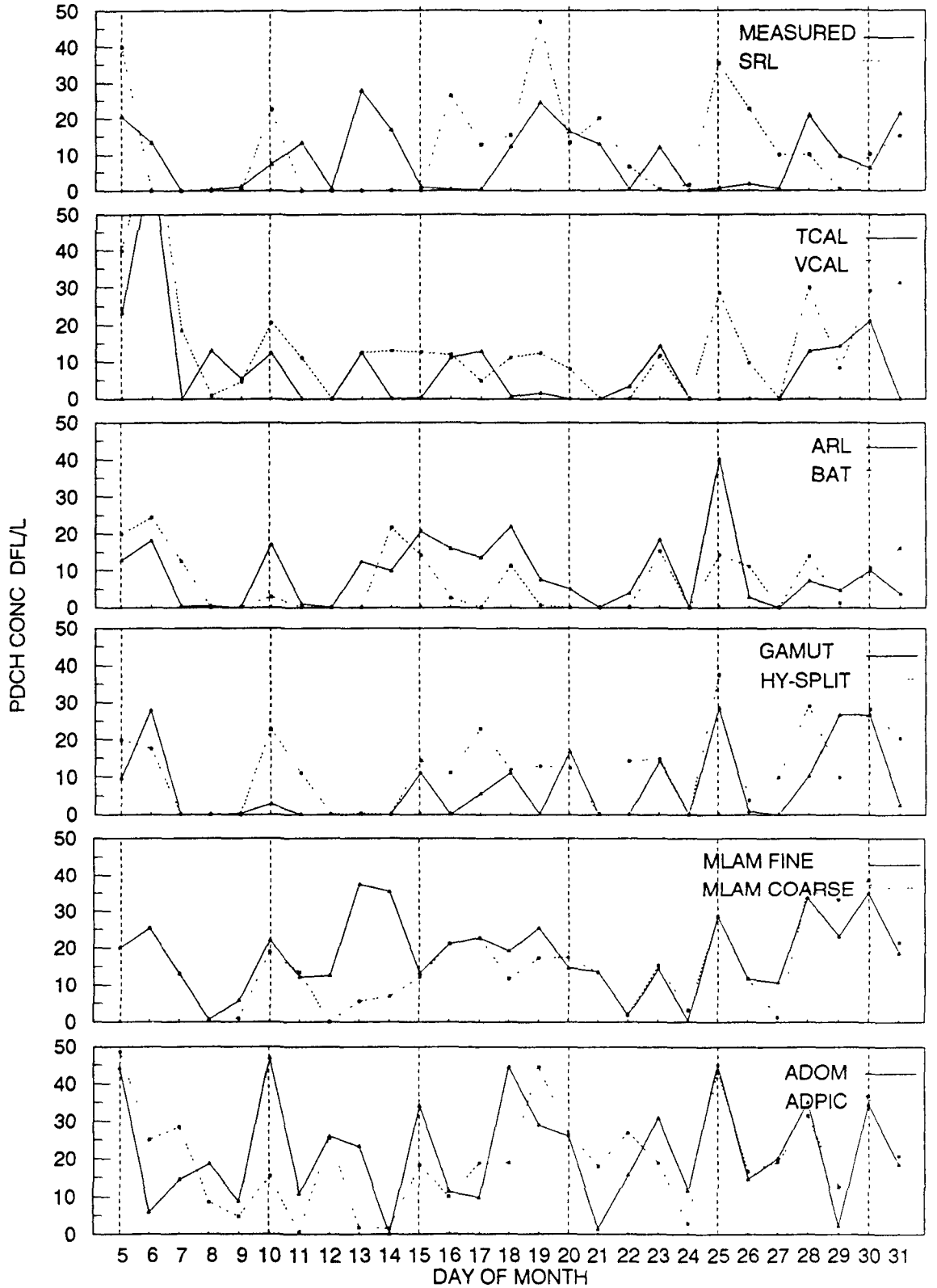
24-H AVERAGE CALCULATED AND MEASURED PTCH CONCENTRATIONS
ALONG THE 2300 KM BAND
FEBRUARY 1987



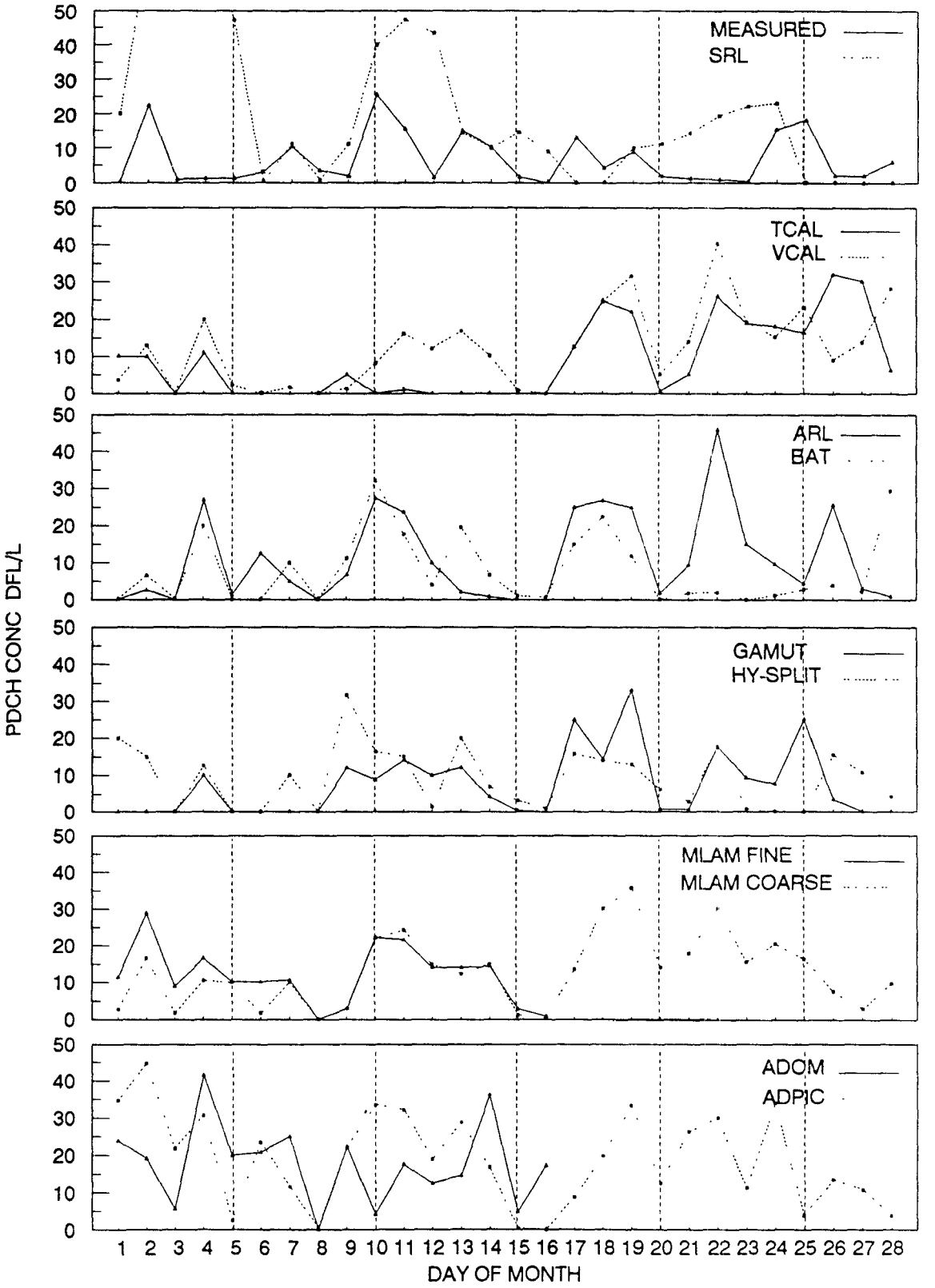
24-H AVERAGE CALCULATED AND MEASURED PTCH CONCENTRATIONS
ALONG THE 2300 KM BAND
MARCH 1987



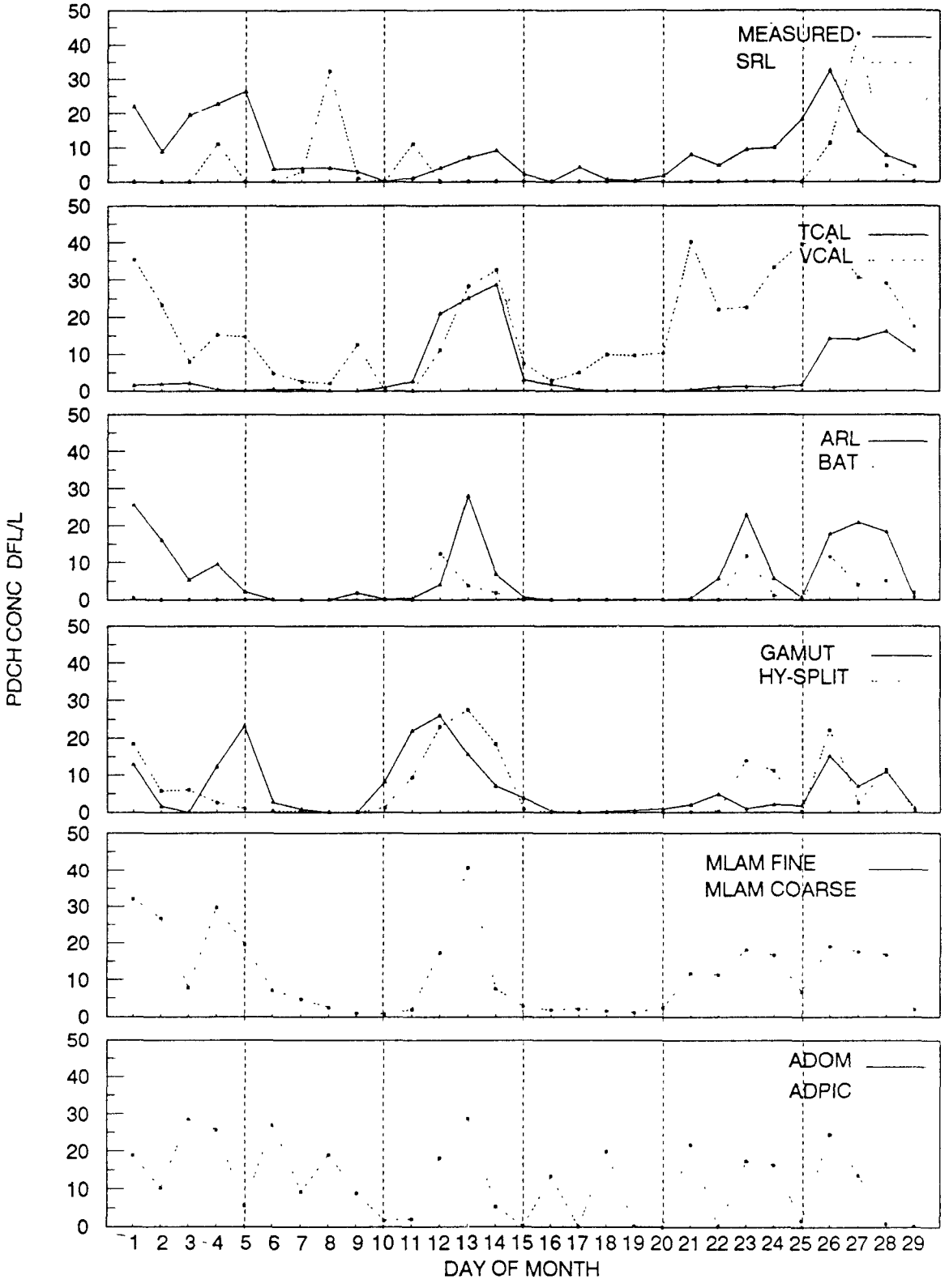
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 300 KM BAND
JANUARY 1987



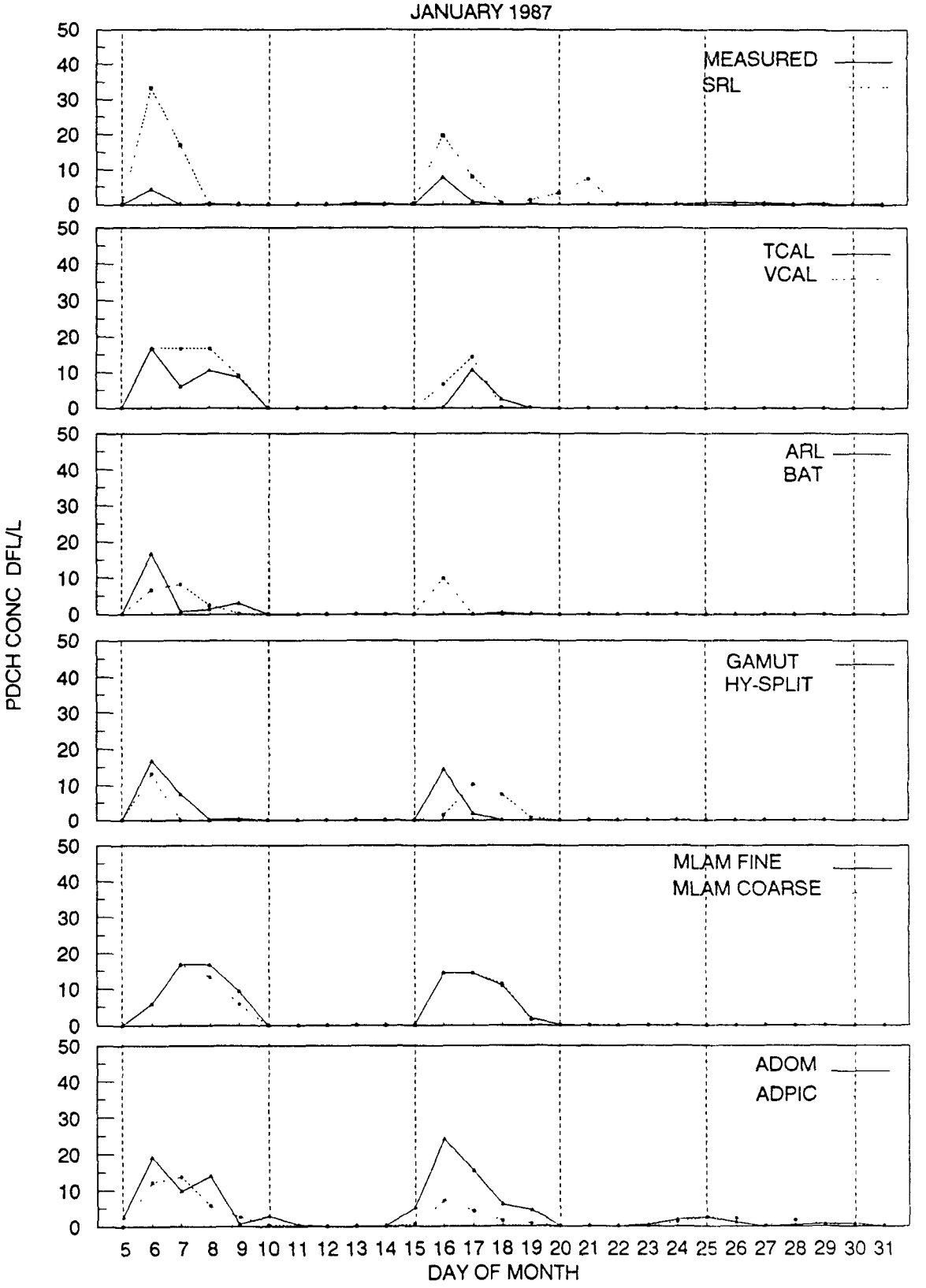
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 300 KM BAND
FEBRUARY 1987



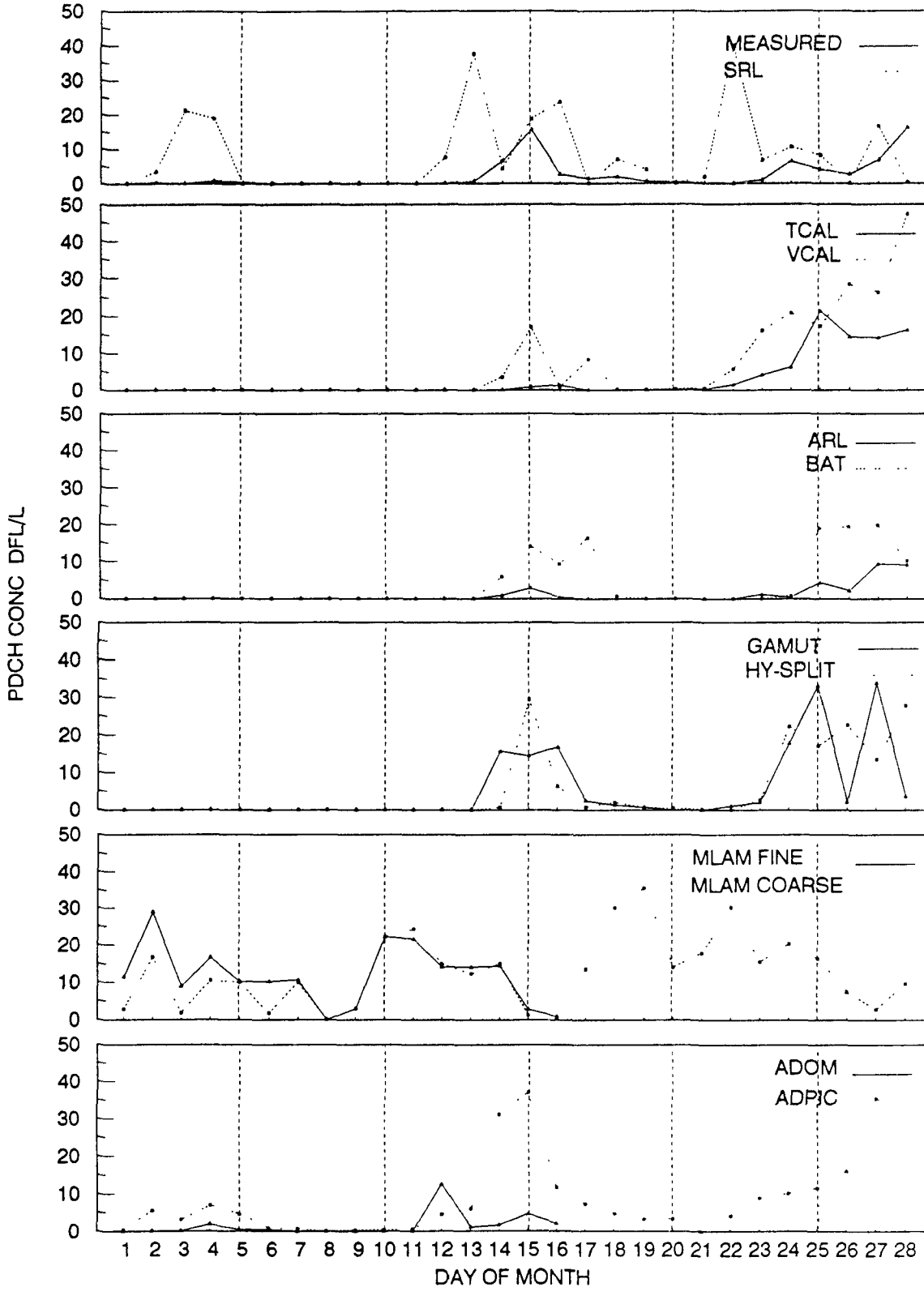
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 300 KM BAND
MARCH 1987



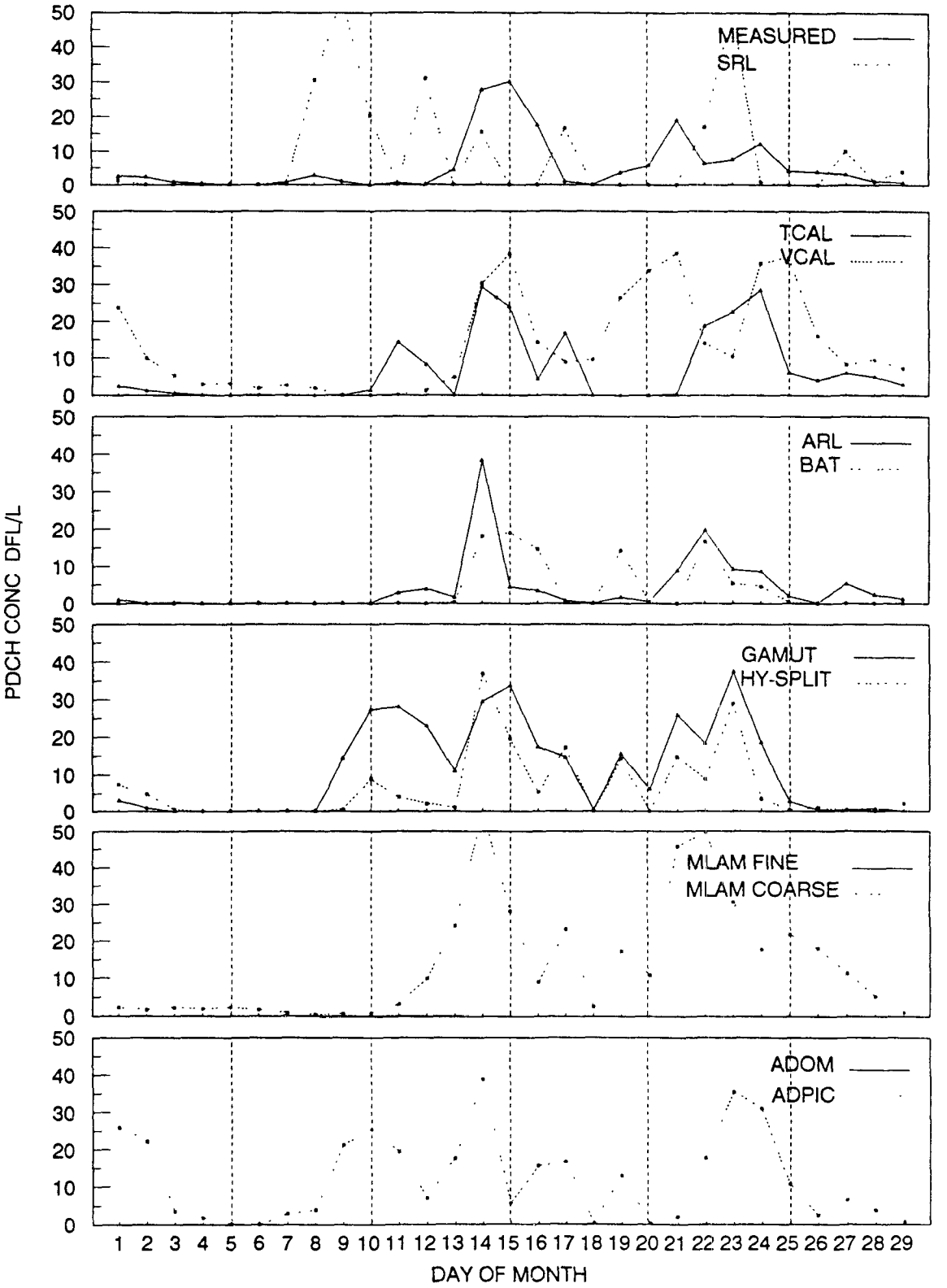
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 700W KM BAND
JANUARY 1987



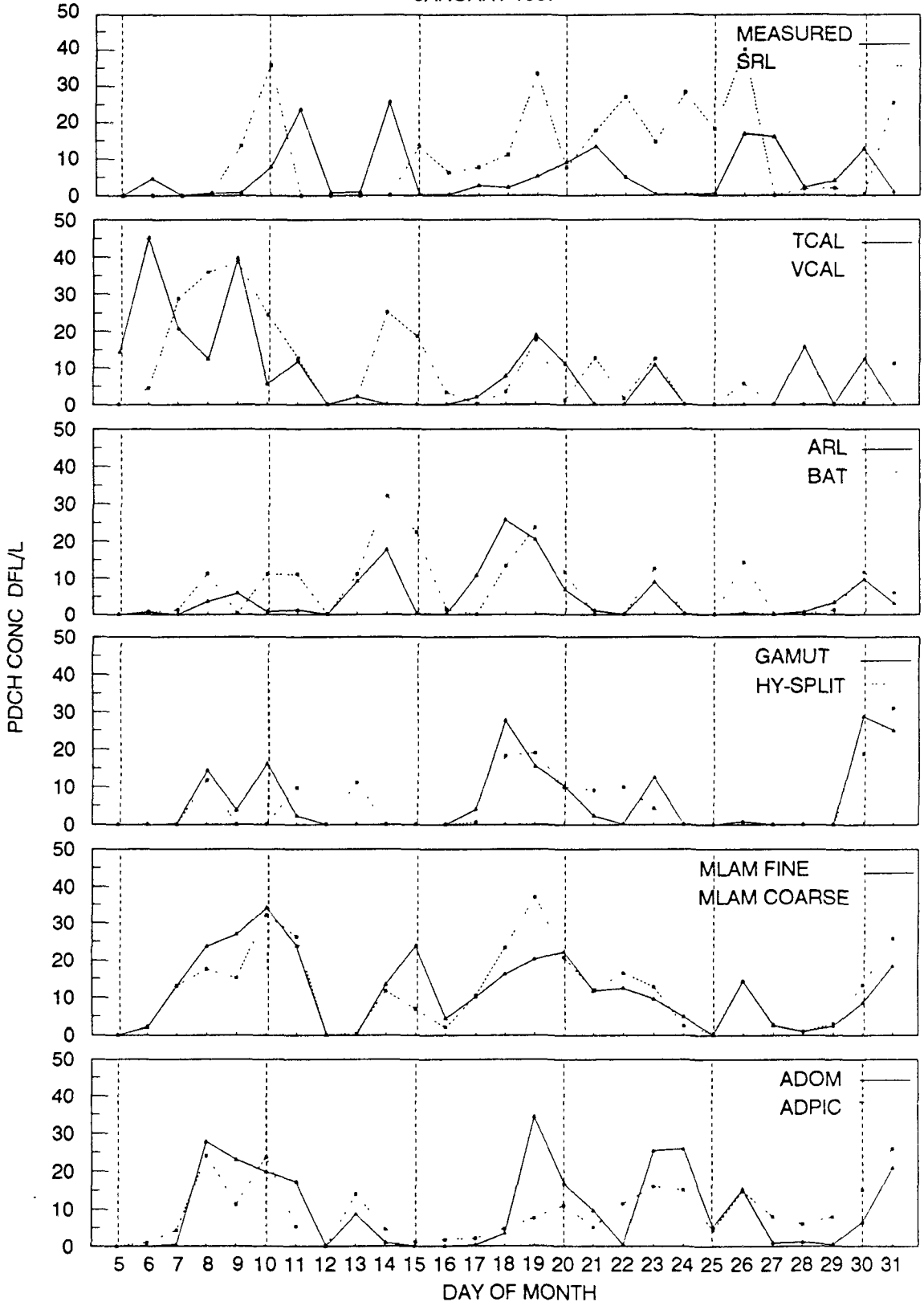
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 700W KM BAND
FEBRUARY 1987



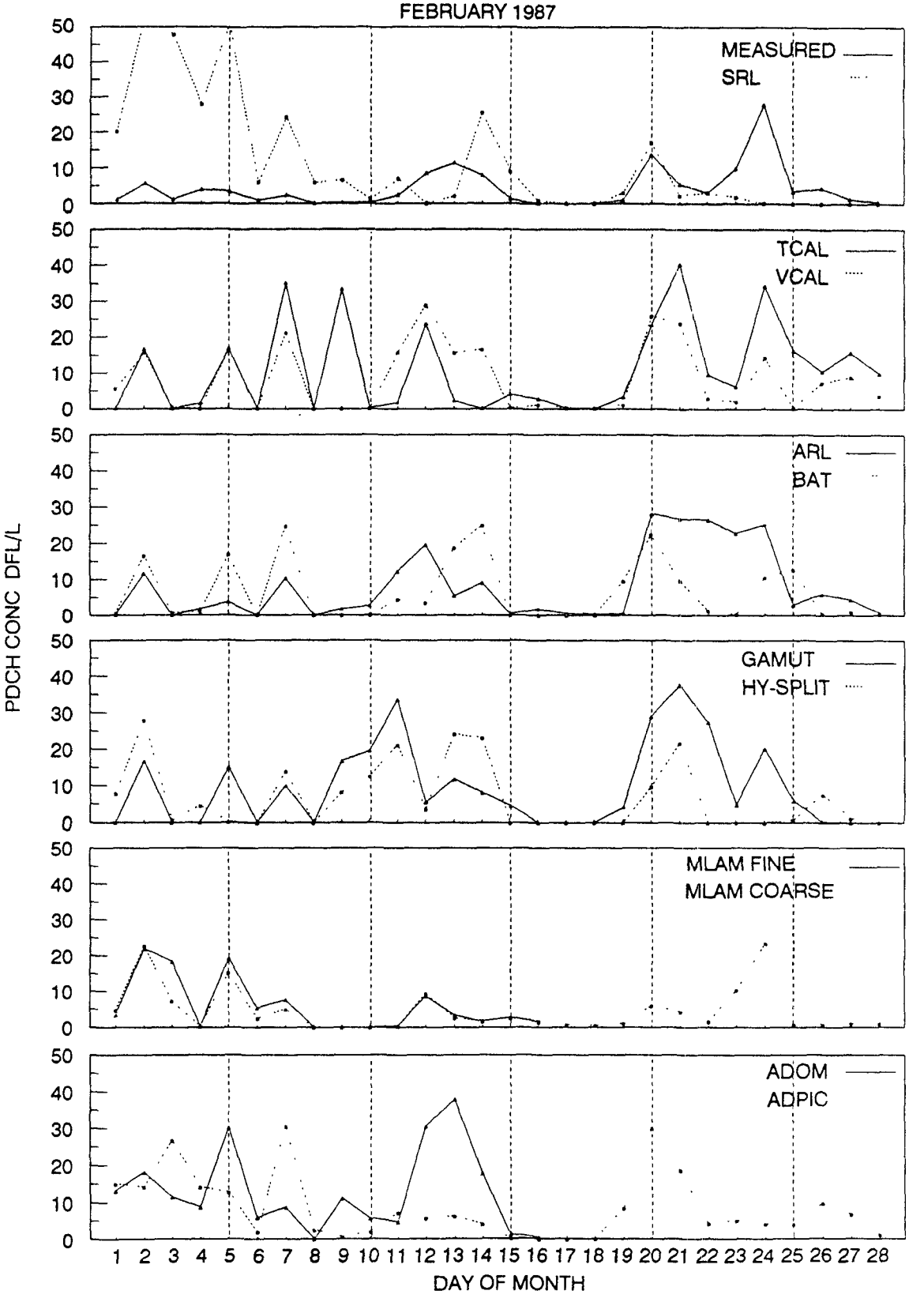
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 700W BAND
MARCH 1987



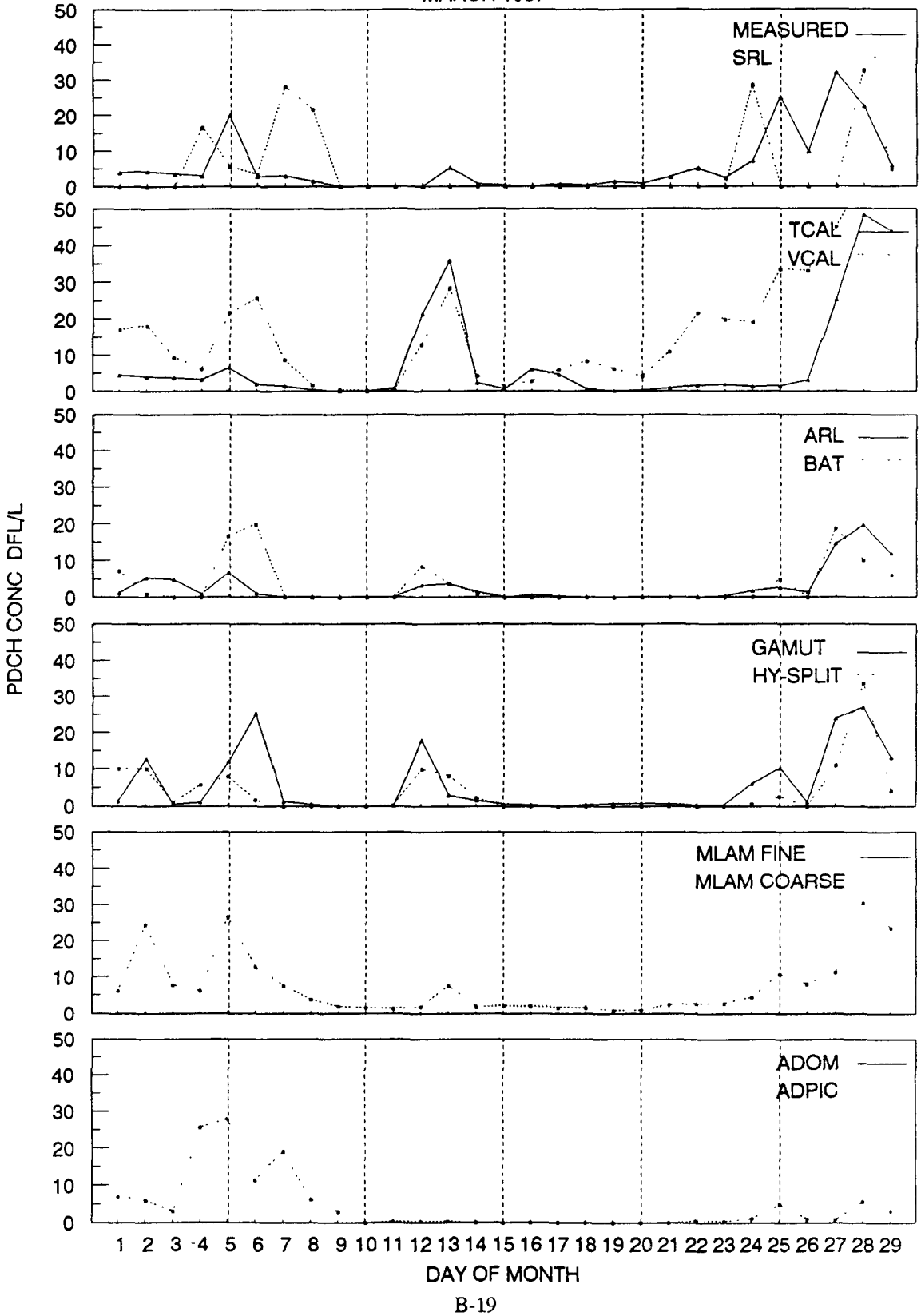
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 700E KM BAND
JANUARY 1987



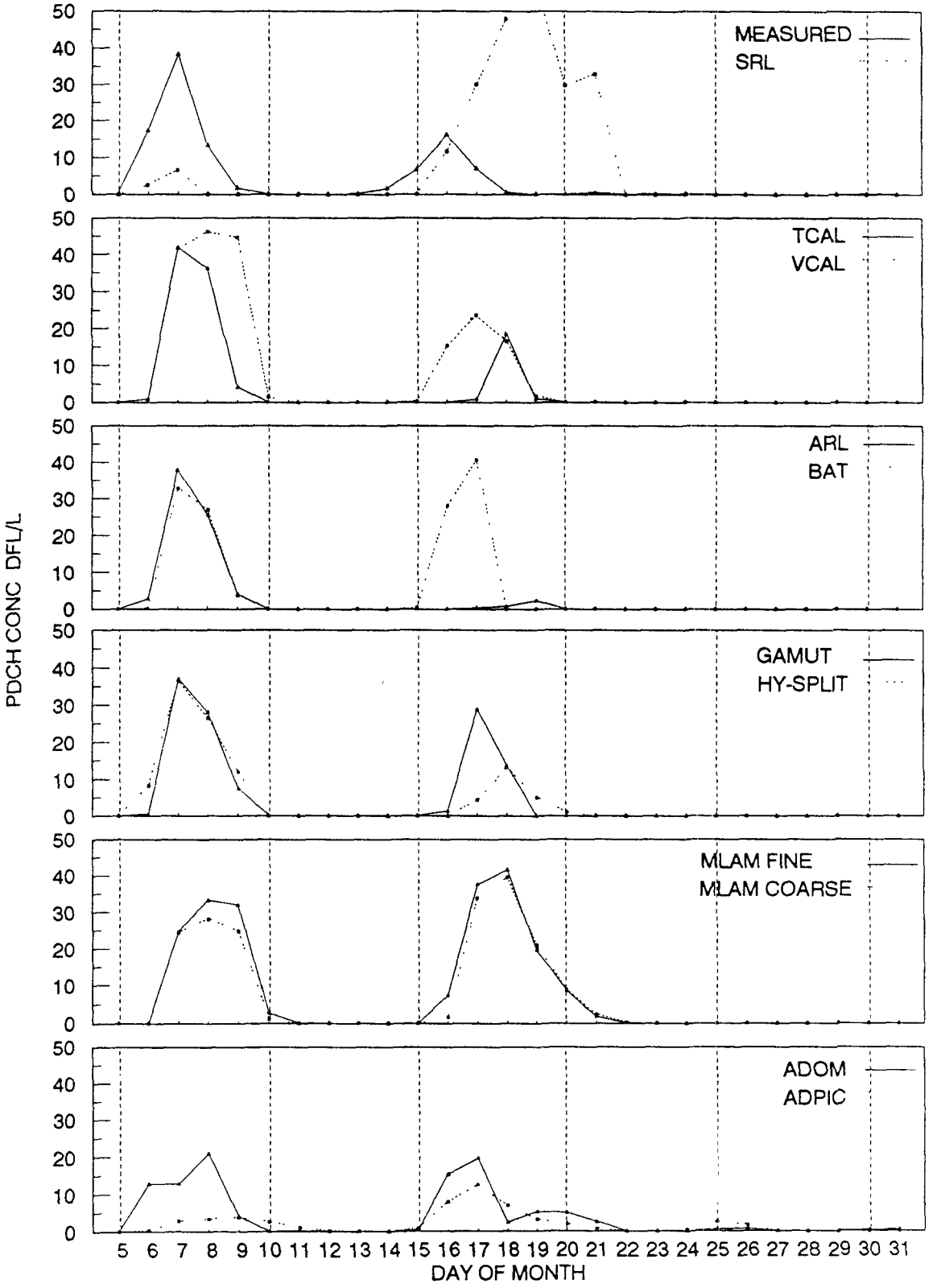
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 700E KM BAND
FEBRUARY 1987



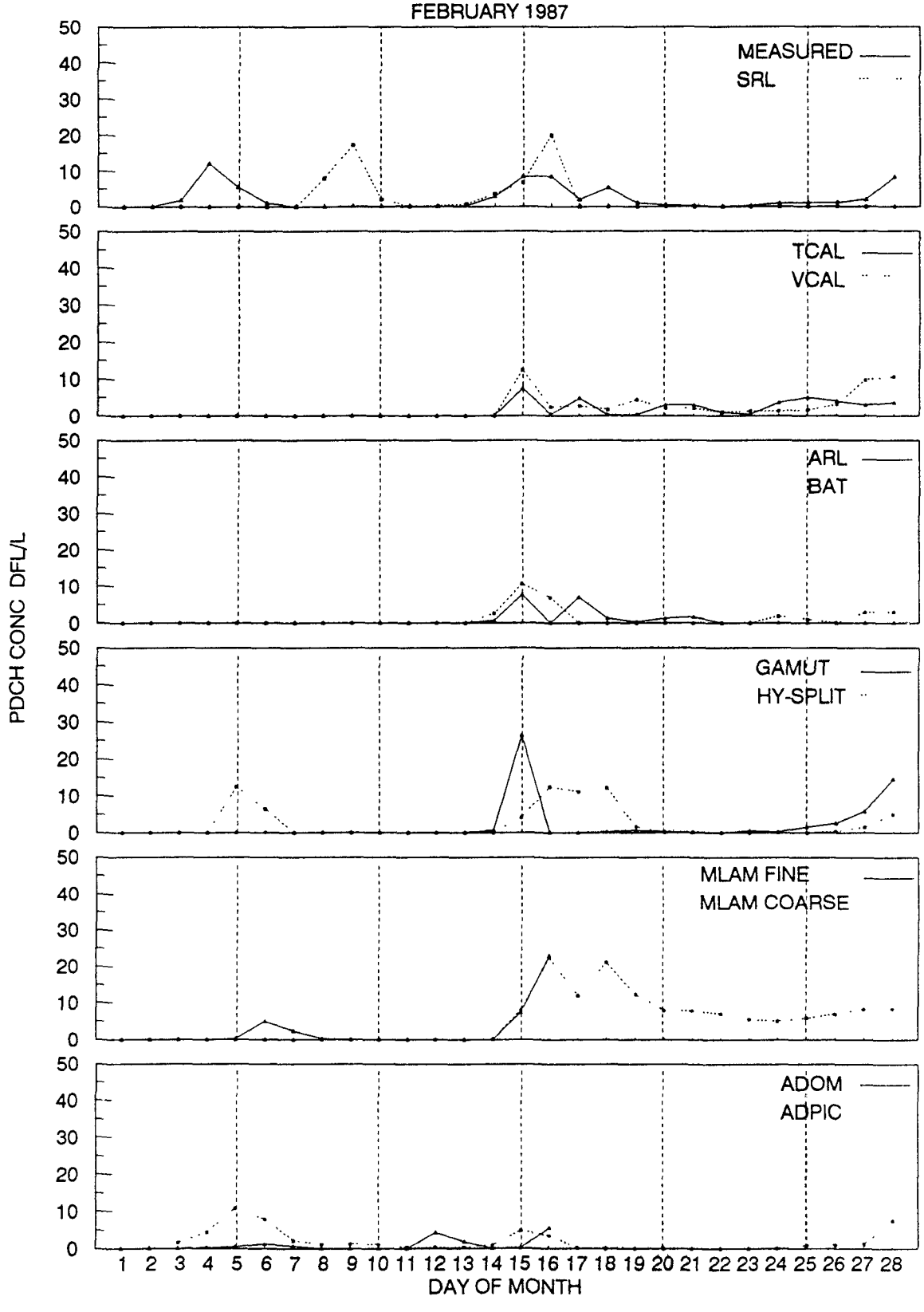
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 700E KM BAND
MARCH 1987



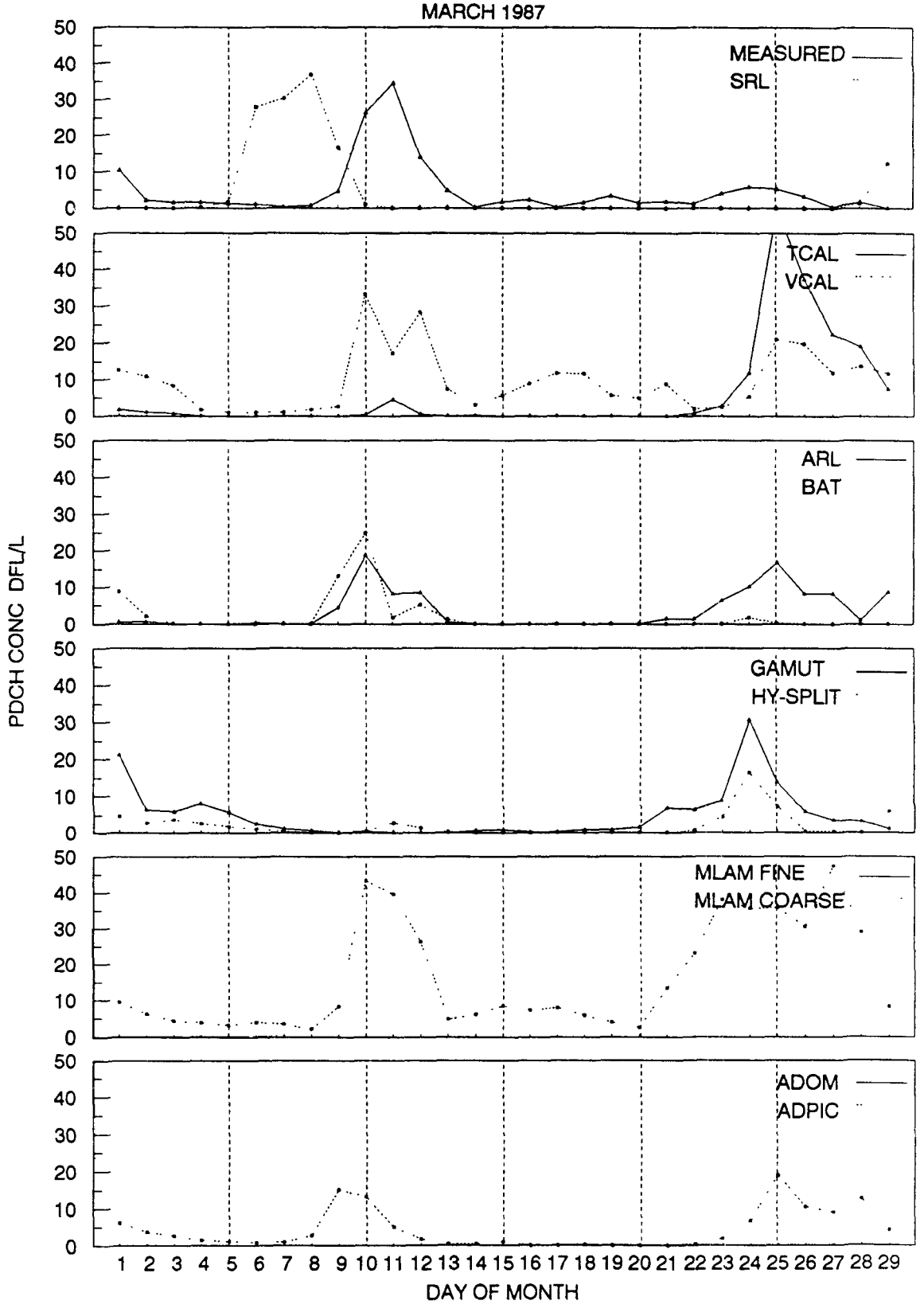
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1000W KM BAND
JANUARY 1987



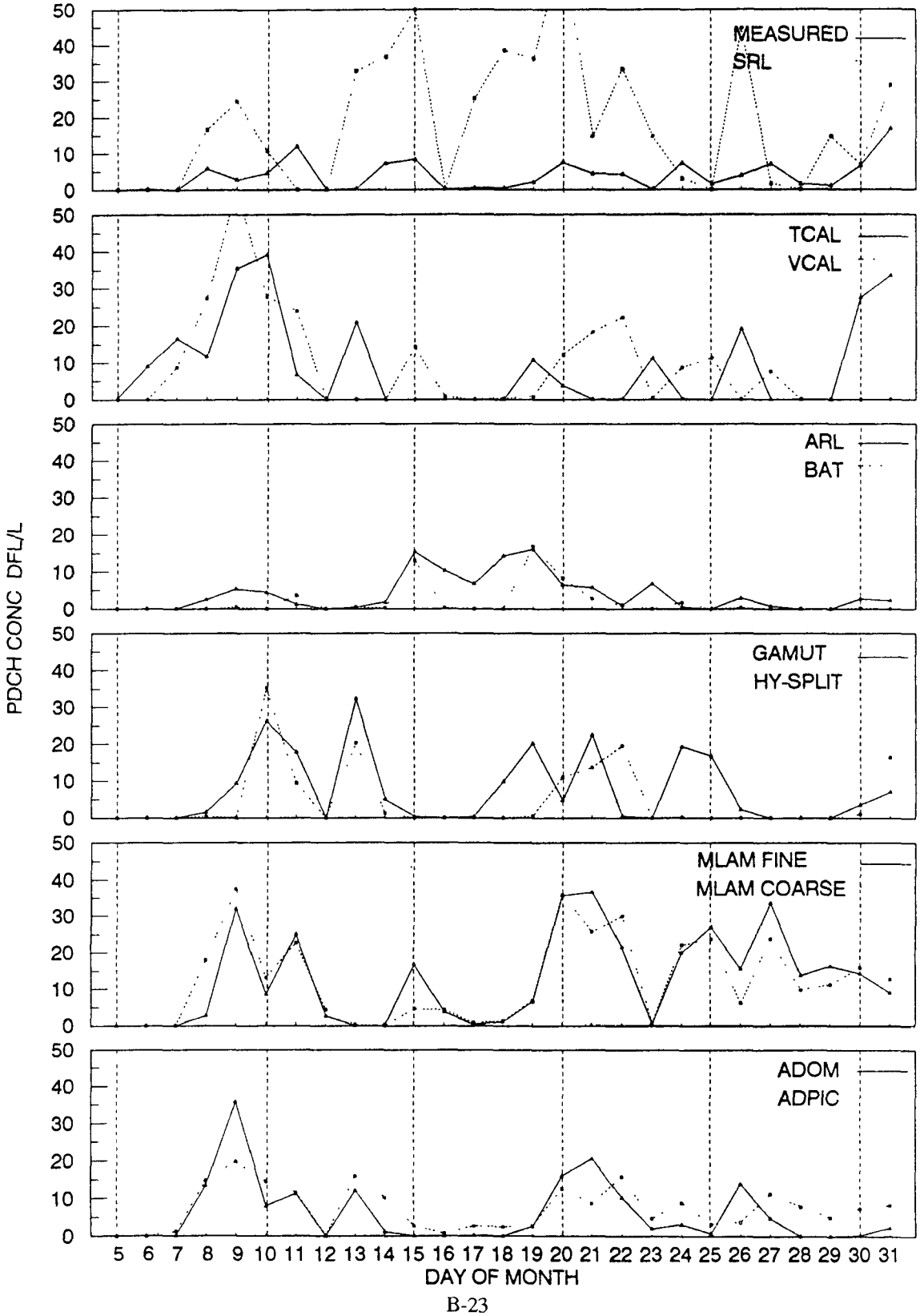
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1000W KM BAND
FEBRUARY 1987



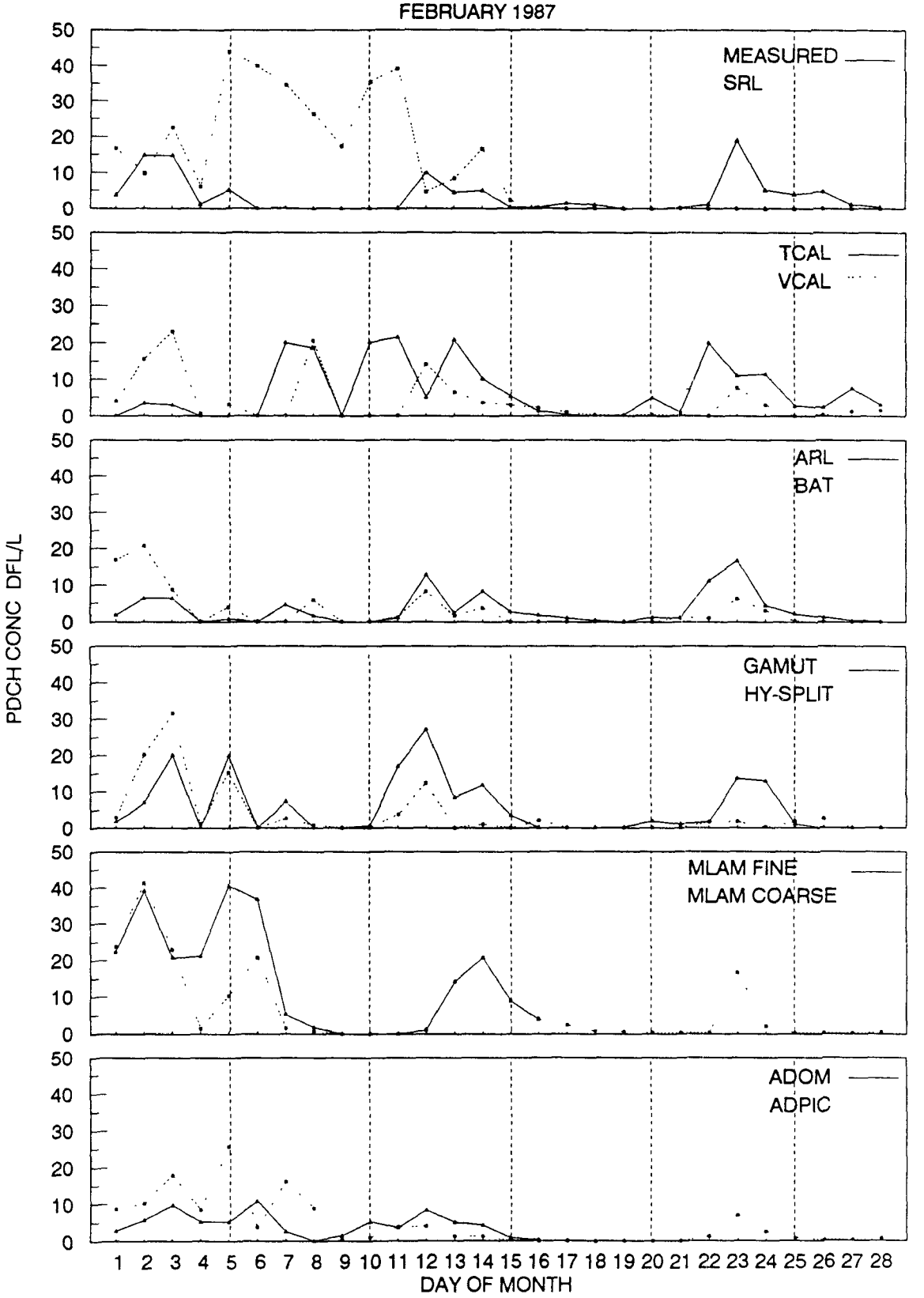
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1000W KM BAND
MARCH 1987



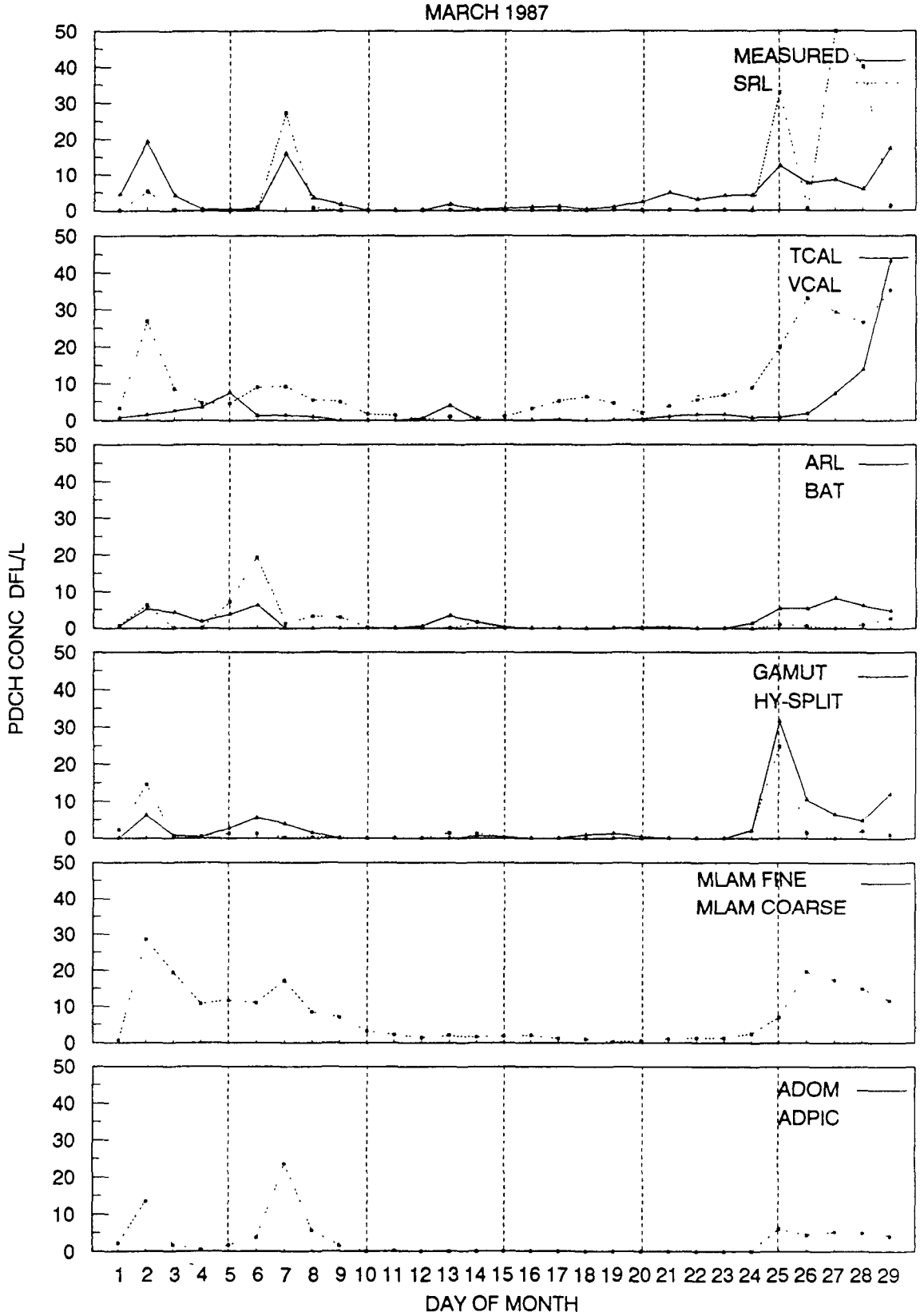
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1000E KM BAND
JANUARY 1987



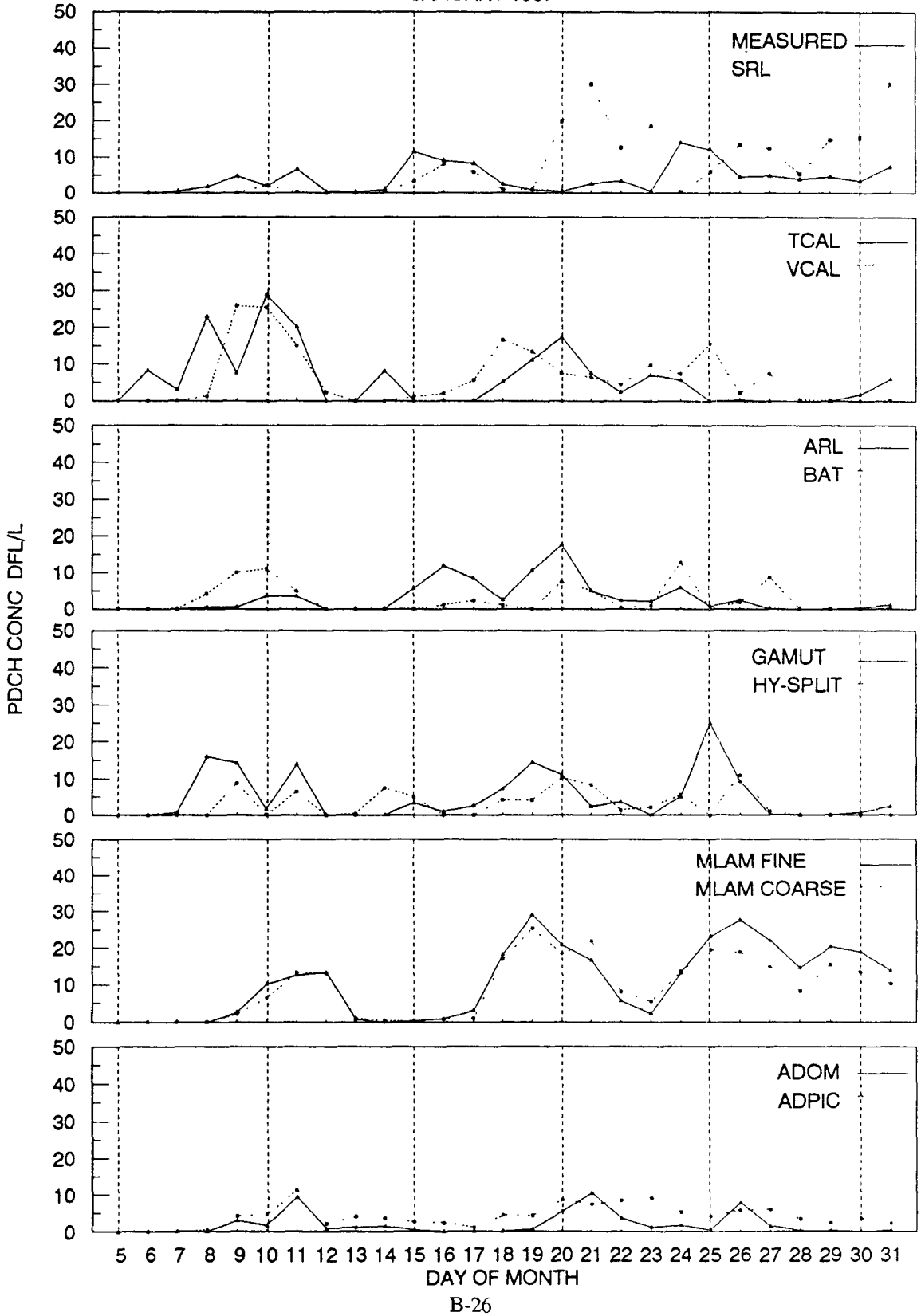
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1000E KM BAND
FEBRUARY 1987



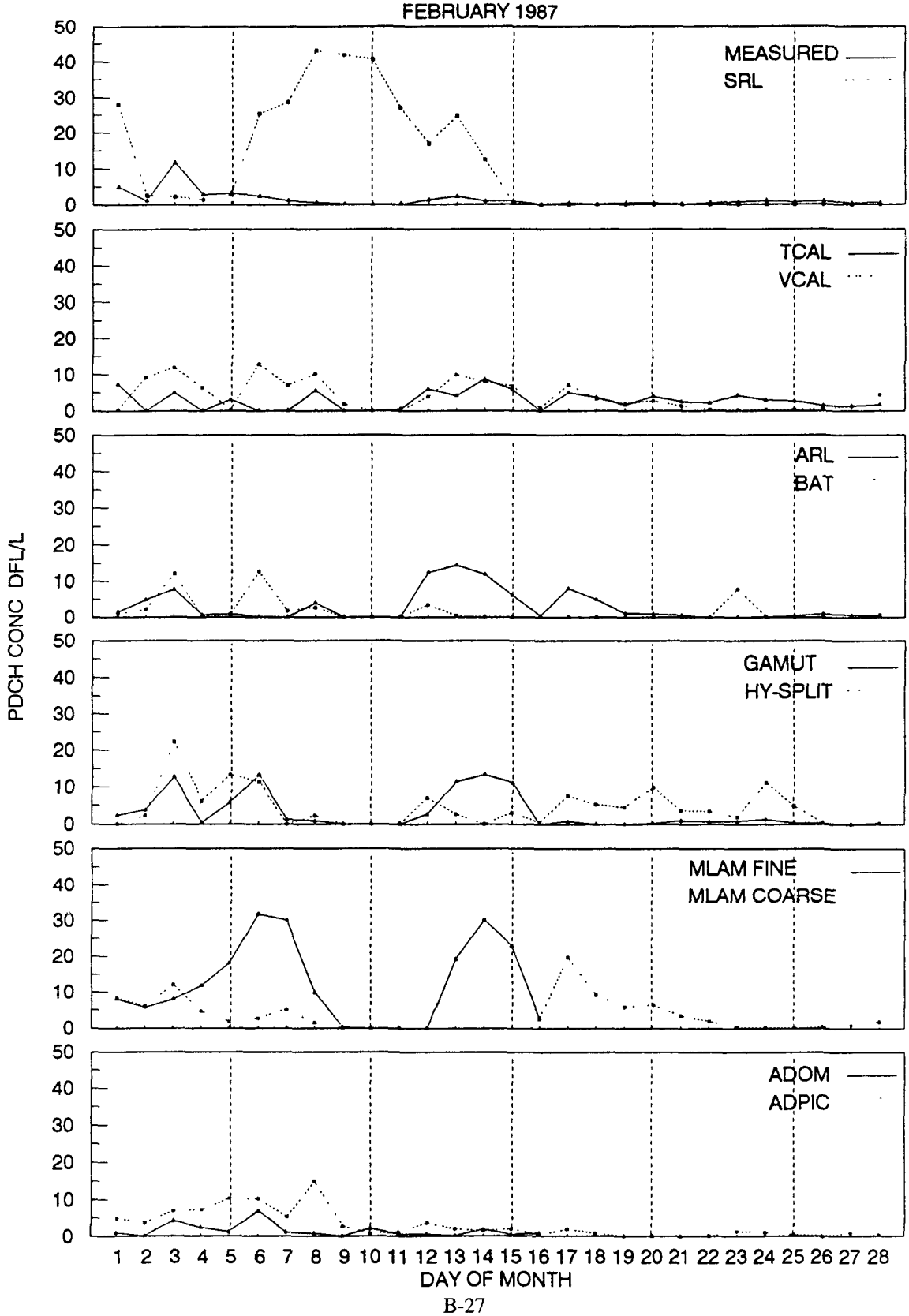
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1000E KM BAND
MARCH 1987



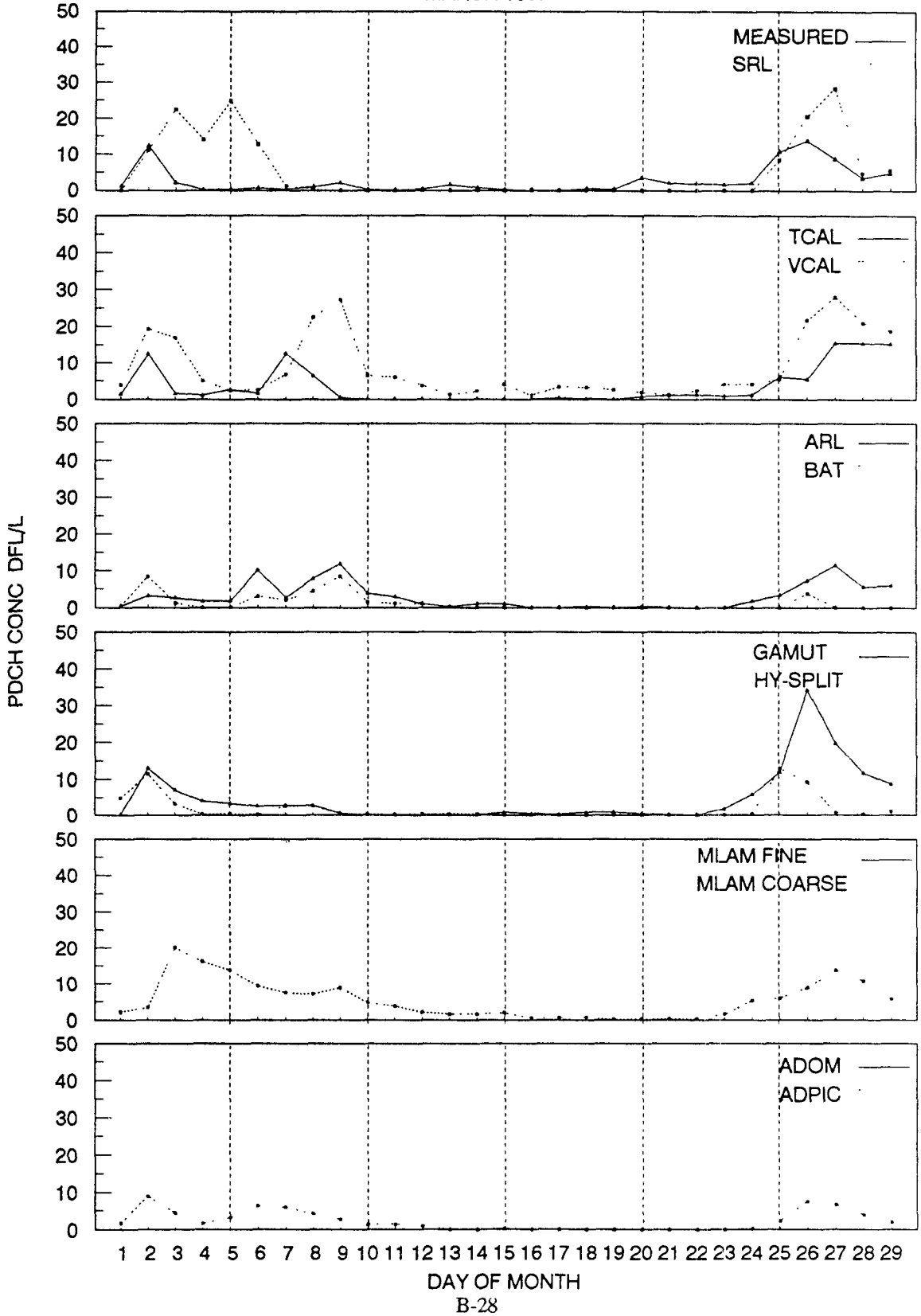
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1400 KM BAND
JANUARY 1987



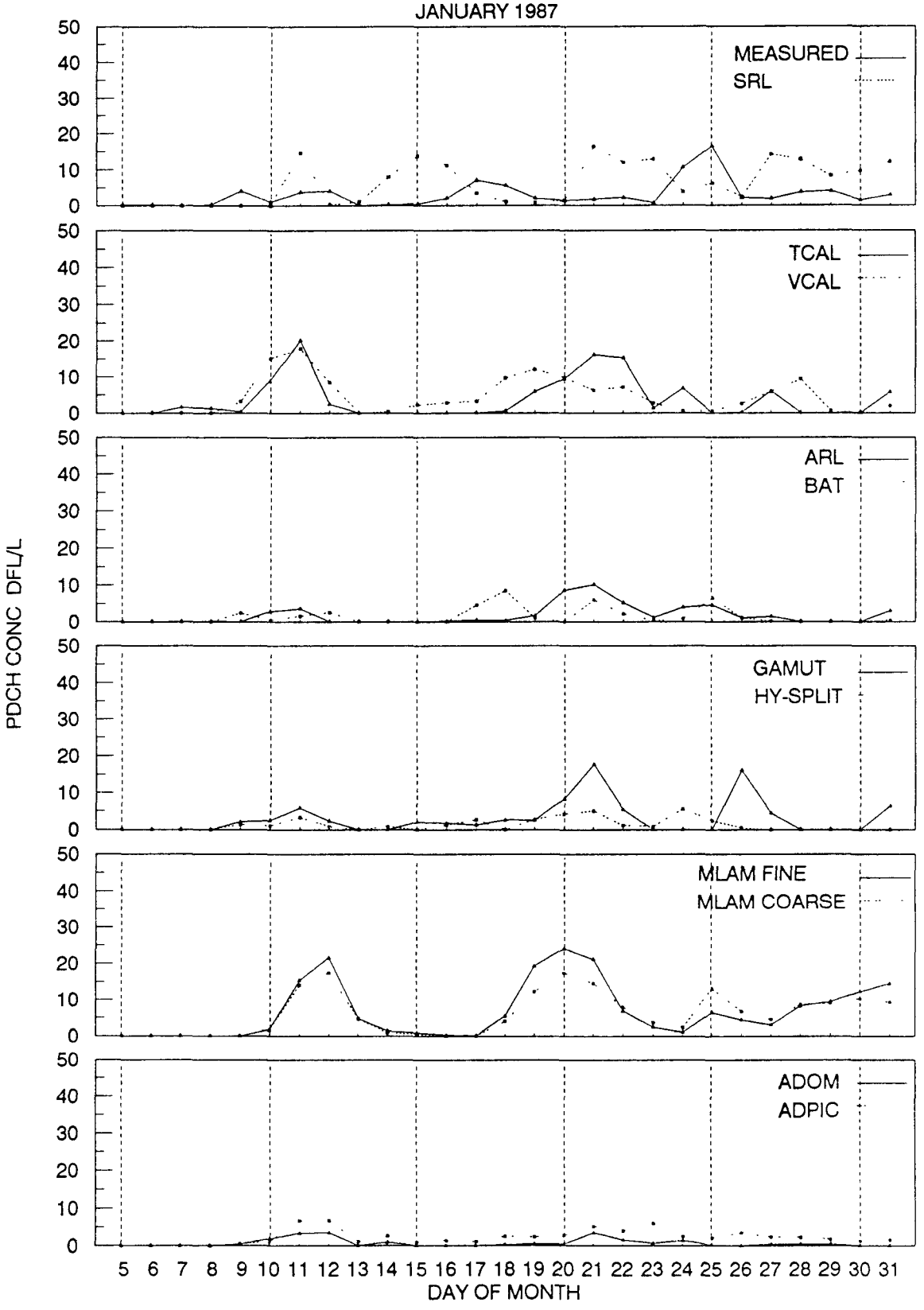
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1400 KM BAND
FEBRUARY 1987



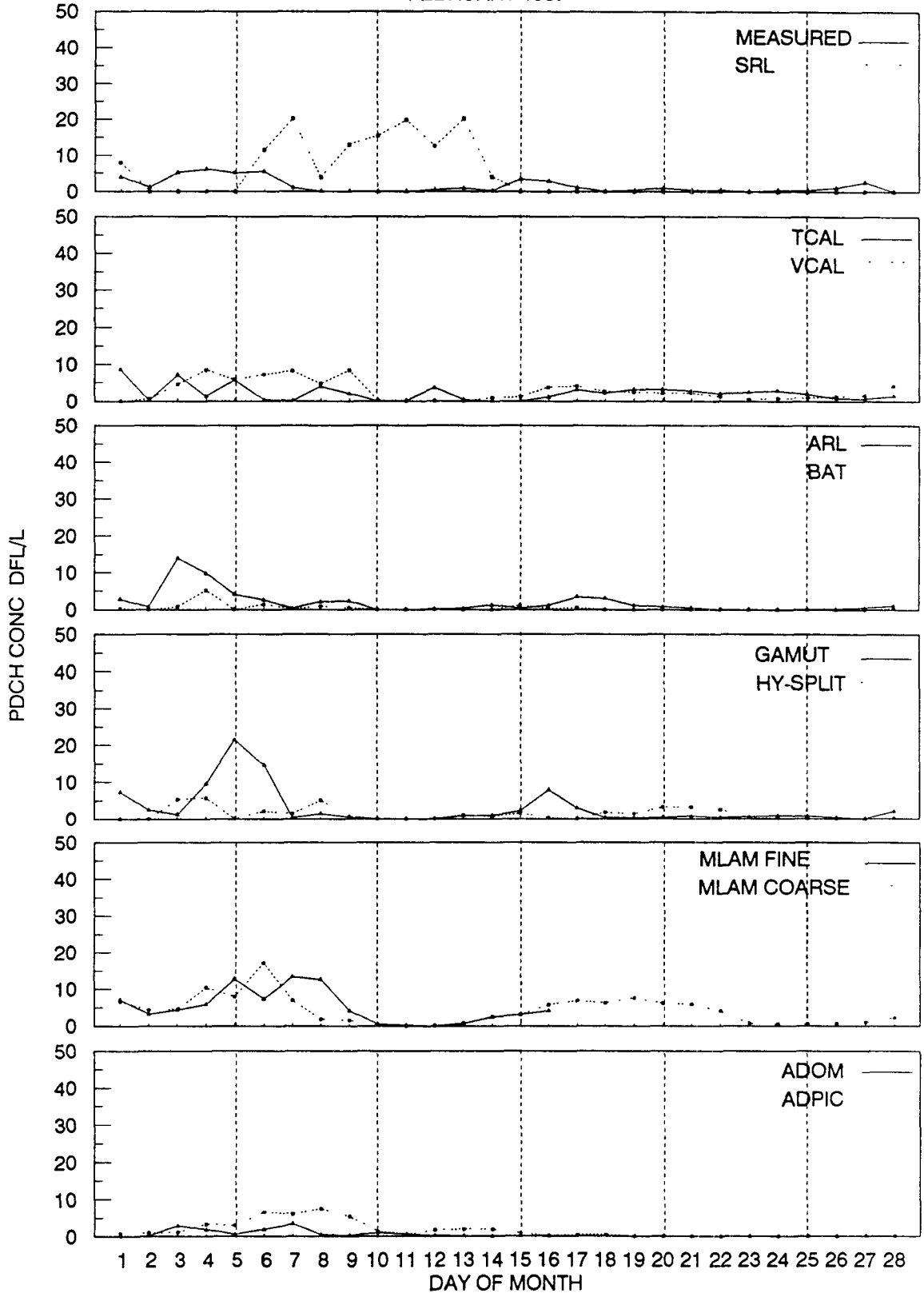
24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1400 KM BAND
MARCH 1987



24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1800 KM BAND
JANUARY 1987



24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1800 KM BAND
FEBRUARY 1987



24-H AVERAGE CALCULATED AND MEASURED PDCH CONCENTRATIONS
ALONG THE 1800 KM BAND
MARCH 1987

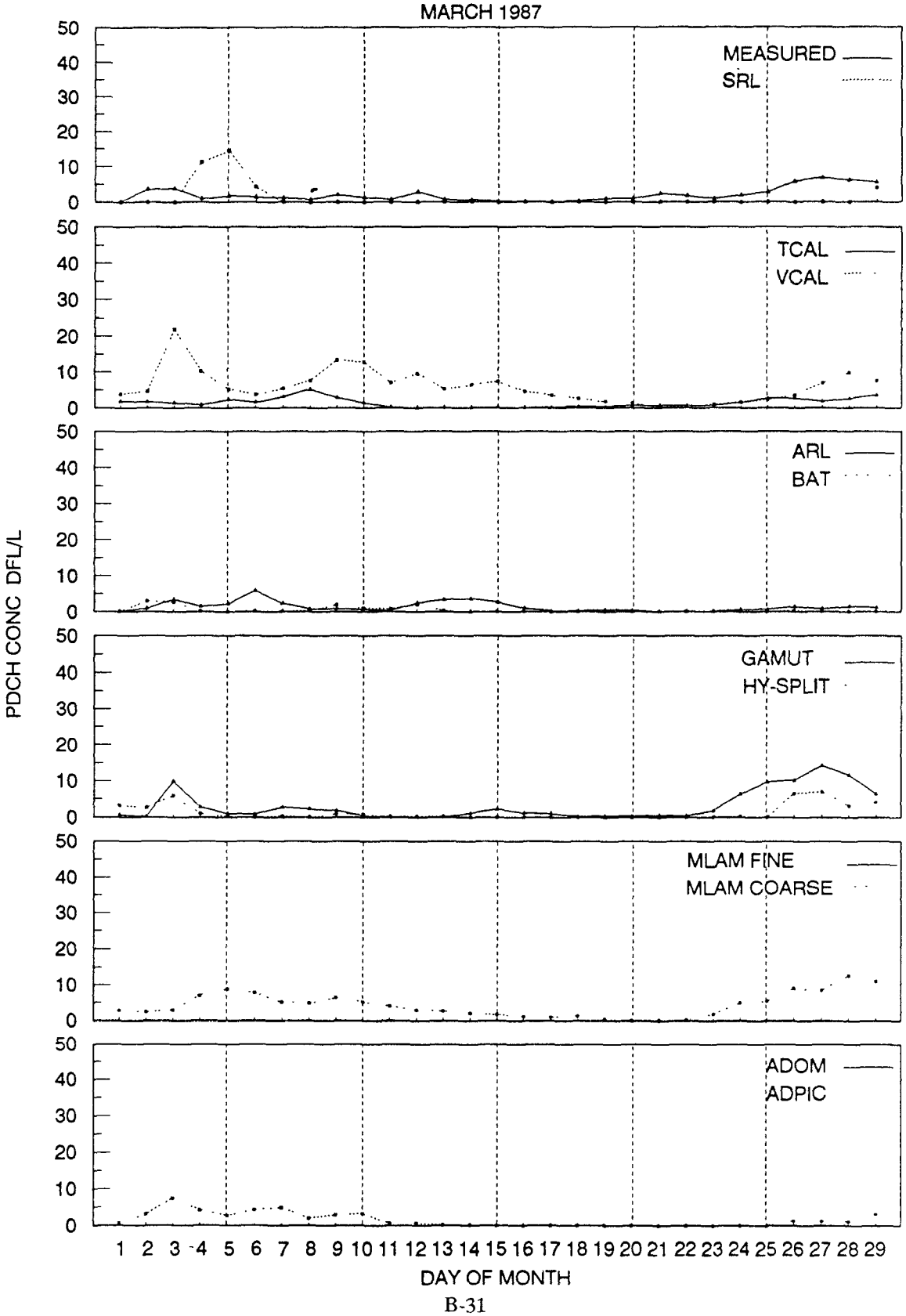


TABLE B-1. LIST OF 24-H MEAN CONCENTRATIONS EXCEEDING 50 dfL/L

Band (km)	Tracer	Page no.	Date	Model SRL	TCAL	VCAL	MLAM-COARSE	ADOM	ADPIC
1,000	PTCH	B-2	Jan 8		53.0			70.9	
"	"	"	Jan 26			55.9			
"	"	"	Jan 28	65.8					
"	"	B-3	Feb 4	54.6					
"	"	"	Feb 5	57.5					
"	"	"	Mar 22			56.0			
"	"	B-4	Mar 2		50.5				
"	"	"	Mar 4				55.0		
"	"	"	Mar 5				51.9		
"	"	"	Mar 21			62.8			
1,600	"	B-6	Feb 7	53.4					
300	PDCH	B-10	Jan 6		64.8	78.1			
"	"	B-11	Feb 2	69.3					
"	"	"	Feb 3	76.0					
"	"	"	Feb 4	73.2					
700W	"	B-14	Feb 28						51.4
"	"	B-15	Mar 9	54.7					
"	"	"	Mar 14				55.0		
"	"	"	Mar 23	51.0					
700E	"	B-17	Feb 2	54.0					
"	"	"	Feb 5	50.8					
"	"	B-18	Mar 28			58.0			
"	"	"	Mar 29			51.8			
1,000W	"	B-19	Jan 19	57.6					
"	"	B-21	Mar 25		56.8				
1000E	"	B-22	Jan 9			56.8			
"	"	"	Jan 20	66.7					

★U.S. GOVERNMENT PRINTING OFFICE: 1990-748-149/20454