5.3 References

K. Anderson, Private Communication, 1994.

F Lurmann, PT Roberts, HH Main, SV Hering, EL Avol, and SD Colome, Phase II report Appendix A: exposure assessment methodology. Final report prepared for the California Air Resources Board, Sacramento CA, by Sonoma Technology Inc, Santa Rosa CA, Aerosol Dynamics, Berkeley CA, USC School of Medicine, Los Angeles CA, and Integrated Environmental Services, Irvine CA, STI-92340-1423-FR, (October 1994).

JM Peters, "Epidemiologic Investigation to Identify Chronic Health Effects of Ambient Air Pollutants in Southern California: Phase II Draft Final Report". California Air Resources Board Contract #A033-186, Submitted May 1, 1995.

L Wallace, Indoor Particles: A Review. J. Air & Waste Manage. Assoc. 46:98-126 (1996).

Table 5-1. Summary of Air Sample Type and Number Collected During the Residential Study.

Sample Type	# Planned Samples	# Collected Samples	# Valid Samples
O3 Total	600	494	481
Inside Home (= In)	300	247	241
Outside Home (= Out)	300	247	240
Continuous In/Out Photometry	20	16	10
PM Total (= PM2.5 + PM10)***	600	340	310
PM2.5 (= PM2.5 In + PM2.5 Out)	300	148	132
PM2.5 In (teflon media only)	150	74	67
PM2.5 Out (teflon media only)	150	74	65
PM10 (= PM10 In + PM10 Out)	300	192	178
PM10 In (teflon media only)	150	96	88
PM10 Out (teflon media only)	150	96	90
HCHO Total	150	124	117
HCHO In	100	105	99
HCHO Out	50	19	18
AER Total	200	182*	161
AER	200	182*	161
AERc	not specified	182*	161
Two Week Sampler	12 homes, in & out	24	24
PM2.5 Mass	12 homes, in & out	24	24
PM2.5 Chemistry	12 homes, in & out	24	24
Acetic Acid	12 homes, in & out	24	24
Formic Acid	12 homes, in & out	24	24

* Additional AER samples were collected, but not analyzed, in some of the study homes; in these instances, the duplicate or triplicate samples were not analyzed based on cost considerations or laboratory recommendations.

*** Quartz fiber filters were used during the first few months of the study, but were found to be inadequate and were excluded from the data set; see Appendix E.

Table 5-2. Summary of Air Monitoring Results for the Residential Study.

Exposure Variable	Mean	Std Dev	Median	Minimum	Maximum	Count
Ozone (ppb)						
Inside	13	12	6	5	73	241
Outside	37	19	34	5	108	240
I/O Ratio	0.37	0.25	0.20	0.06	1.48	239
PM2.5 , ug/m3						
Inside	20.9	20.0	13.7	4.2	106.9	67
Outside	16.0	15.0	10.7	2.0	76.8	65
I/O Ratio	2.03	2.99	1.10	0.37	19.96	61
PM10 , ug/m3						
Inside	40.6	36.6	32.9	2.3	294.6	88
Outside	36.3	25.7	29.0	2.2	141.3	90
I/O Ratio	1.54	1.66	1.05	0.11	10.49	87
HCHO , ug/m3						
Inside	11.3	7.4	10.1	0.2	38.8	99
Outside	3.2	2.5	3.2	1.0	10.1	18
I/O Ratio	6.01	8.96	3.75	0.01	39.95	18
AER	0.7	0.5	0.7	0.0	2.5*	161
AERc	0.8	0.5	0.7	0.0	2.7*	161

Notes:

1) I/O Ratio refers to ratio of indoor to outdoor concentrations, and is in dimensionless units;
2) Concentrations below detection limits for ozone, PM10, and HCHO have been replaced with Limit of Detection values (5 ppb for O3, 2 ug/m3 for PM10, 0.2 ug/m3 HCHO).
3) AER Limit of Detection determined to be about 1.1/hr ; values larger than this should be considered speculative and prone to error.

Table 5-3. Percentile values for pollutants monitored in the Residential Study.

Percentile Values	Ozone, ppb			PM10, ug/m3			PM2.5, ug/m3			HCHO, ug/m3			AER per hr	AERc per hr
	Indoor	Outdoor	I/O Ratio	Indoor	Outdoor	I/O Ratio	Indoor	Outdoor	I/O Ratio	Indoor	Outdoor	I/O Ratio		
1	0	3	0	2.3	2.2	0.11	4.2	0.7	0.37	0.0	1.0	0.01	0.0	0.0
5	0	9	0	12.0	8.7	0.38	4.7	2.6	0.64	0.3	1.0	0.01	0.2	0.2
10	0	14	0	13.3	10.8	0.43	5.6	4.3	0.67	3.8	1.0	0.35	0.2	0.2
25	2	23	0.07	24.2	18.1	0.66	9.7	7.3	0.84	6.5	1.7	1.98	0.3	0.4
50	6	34	0.20	32.9	29.0	1.05	13.7	10.7	1.10	10.1	3.2	3.75	0.7	0.7
75	16	51	0.45	47.2	44.2	1.86	22.5	19.6	1.68	15.2	4.1	7.38	1.0	1.1
90	32	62	0.65	62.2	75.6	2.73	43.8	31.6	3.85	20.9	7.5	9.16	1.4	1.5
95	42	69	0.76	83.2	87.6	3.98	68.0	44.2	7.84	26.8	10.1	39.95	1.8	1.9
99	50	89	0.98	294.6	141.3	10.49	106.9	76.8	19.96	38.8	10.1	39.95	2.3	2.7

Notes:
1) Corrected ozone values less than zero (due to large blank values) were replaced with zero.
2) To avoid unreliable ratios when denominator values were very small, outdoor ozone values less than the LOD (5 ppb) were excluded;
3) I/O ratio refers to ratio of indoor to outdoor concentrations, and is in dimensionless units.

Table 5-4. Ozone Sampling Results, By Community.

Variable	Community			
	Lake Gregory	Lancaster	Riverside/Mira Loma	San Dimas
Indoor O3, by TEDS				
observations	51	70	72	47
average, ppb	17	13	12	10
std. dev.	17	11	10	9
minimum value	5	5	5	5
maximum value	73	44	52	42
Outdoor O3, by TEDS				
observations	51	70	71	47
average, ppb	45	43	31	27
std. dev.	20	19	14	16
minimum value	5	5	7	5
maximum value	108	92	79	60
Indoor/Outdoor Ratio				
observations	50	70	71	47
average	0.35	0.33	0.39	0.41
std. dev.	0.26	0.22	0.25	0.28
minimum value	0.08	0.06	0.06	0.09
maximum value	1	1	1.48	1

Table 5- 5. Ozone Sampling Results, by Reported Home Air Conditioning Type.

Variable	Air Conditioning Type			
	No A/C	Central or Room A/C	Swamp Cooler Only	Swamp + Central or Room A/C
Indoor O3, by TEDS				
observations	70	113	31	15
average, ppb	16	9	16	20
std. dev.	15	8	13	15
minimum value	5	5	5	5
maximum value	73	42	52	44
Outdoor O3, by TEDS				
observations	69	113	31	15
average, ppb	42	36	34	41
std. dev.	20	18	18	19
minimum value	5	5	9	12
maximum value	108	80	91	65
Indoor/Outdoor Ratio				
observations	68	113	31	15
average	0.37	0.32	0.46	0.44
std. dev.	0.25	0.23	0.28	0.2
minimum value	0.08	0.06	0.12	0.17
maximum value	1	1	1.48	0.76

Table 5-6. Relative Standard Deviation (RSD) of co-located pollutant samples.

Descriptive Statistics	O3, by TEDs, in ppb	PM, in ug/m3	HCHO, in ug/m3
Mean	0.28	0.18	0.09
Standard Error	0.11	0.03	0.03
Median	0.12	0.10	0.06
Standard Deviation	0.87	0.18	0.07
Minimum	-2.87	0.01	0.02
Maximum	4.70	0.70	0.20
Count	58	27	5

Table 5-7. PM2.5 Sampling Results, By Community.

Variable	Community			
	Lake Gregory	Lancaster	Riverside/Mira Loma	San Dimas
Indoor PM2.5				
observations	13	21	17	14
average, ug/m3	10.4	18.5	33.0	20.4
std. dev.	9.6	19.4	25.5	15.7
minimum value	4.2	5.5	10.3	5.5
maximum value	39.9	86.4	106.9	68.0
Outdoor PM2.5				
observations	11	22	17	14
average, ug/m3	9.6	8.3	31.6	14.1
std. dev.	12.1	3.1	19.7	6.0
minimum value	2.4	2.0	6.3	6.2
maximum value	44.2	13.2	76.8	25.5
Indoor/Outdoor Ratio				
observations	10	20	16	14
average	2.35	2.91	1.23	1.53
std. dev.	3.38	4.44	0.99	1.03
minimum value	0.37	0.58	0.43	0.76
maximum value	11.7	19.96	4.18	3.91

Table 5- 8. PM2.5 Sampling Results, by Home Air Conditioning Type.

Variable	Air Conditioning Type			
	No A/C	Central or Room A/C	Swamp Cooler Only	Swamp + Central or Room A/C
Indoor PM2.5				
observations	15	33	11	4
average, ug/m3	16.4	20.5	25.1	12.6
std. dev.	14.0	18.7	20.4	5.8
minimum value	4.2	4.2	5.5	5.9
maximum value	49.1	86.4	78.5	19.9
Outdoor PM2.5				
observations	13	34	12	4
average, ug/m3	19.1	12.7	18.7	10.0
std. dev.	21.0	10.0	10.8	3.5
minimum value	2.6	2.0	5.4	5.2
maximum value	75.0	52.0	39.8	13.4
Indoor/Outdoor Ratio				
observations	12	32	11	4
average	2	2.31	1.57	1.61
std. dev.	3.14	3.59	1.25	1.51
minimum value	0.37	0.43	0.66	0.58
maximum value	11.7	19.96	4.18	3.85

Table 5-9. PM10 Sampling Results, By Community.

Variable	Community			
	Lake Gregory	Lancaster	Riverside/Mira Loma	San Dimas
Indoor PM10				
observations	18	23	26	17
average, ug/m3	25.8	40.6	59.0	30.9
std. dev.	18.4	27.1	55.4	13.3
minimum value	2.3	13.3	23.9	8.7
maximum value	83.2	140.3	294.6	56.5
Outdoor PM10				
observations	20	24	26	16
average, ug/m3	13.9	32.3	60.3	29.6
std. dev.	6.3	13.4	31.4	13.8
minimum value	2.2	9.2	12.0	10.9
maximum value	25.1	68.9	141.3	28.6
Indoor/Outdoor Ratio				
observations	18	23	26	16
average	2.29	1.77	1.18	1.18
std. dev.	2.29	2.17	0.89	0.64
minimum value	0.11	0.26	0.33	0.43
maximum value	10.49	8.96	3.98	2.49

Table 5- 10. PM2.5 Sampling Results, by Reported Home Air Conditioning Type.

Variable	Air Conditioning Type			
	No A/C	Central or Room A/C	Swamp Cooler Only	Swamp + Central or Room A/C
Indoor PM10				
observations	21	43	6	7
average, ug/m3	30.0	33.2	92.8	54.0
std. dev.	17.8	14.2	99.2	43.1
minimum value	2.3	8.7	45.5	18.2
maximum value	83.2	62.2	294.6	140.3
Outdoor PM10				
observations	23	43	6	7
average, ug/m3	20.3	38.1	51.1	37.7
std. dev.	14.3	26.8	21.1	20.9
minimum value	2.2	5.1	26.1	9.2
maximum value	58.6	141.3	78.0	68.9
Indoor/Outdoor Ratio				
observations	21	42	6	7
average	2.08	1.16	1.73	2.92
std. dev.	2.2	0.69	1.19	3.74
minimum value	0.011	0.33	0.64	0.26
maximum value	10.49	2.73	3.98	8.96

Table 5-11. Summary of Co-Located PM10 and PM2.5 Study Data.

Home ID	Town ID	Sample Date	Inside		Outside		I/O	I/O	# Cigarettes	Inside	Outside
			PM10	PM2.5	PM10	PM2.5	PM10	PM2.5	Smoked	PM2.5/PM10	PM2.5/PM10
7	SD	82094	33.6	18.2	58.6	23.1	0.57	0.79	0	0.54	0.39
8	SD	91494	31.2	11.2	53	13.4	0.59	0.84	0	0.36	0.25
9	RIV	81694	37.3	17.4	90.2	20.8	0.41	0.84	3	0.47	0.23
12	RIV	80894	24.5	11.9	74.5	27.4	0.33	0.43	0	0.48	0.37
15	SD	81194	45.5	20.1	42.4	18.5	1.07	1.09	30	0.44	0.43
17	LAN	90294	52.3	12.5	26.1	11.5	2.00	1.09	0	0.24	0.44
18	LAN	91094	23.3	7.8	27.1	8.2	0.86	0.95	0	0.33	0.30
30	RIV	80894	35.5	20.4	82.9	31.6	0.43	0.65	0	0.57	0.38
48	RIV	92394	51	10.2	22.8	12	2.24	0.85	0	0.20	0.53
59	LAN	92194	31.5	12.7	44	10	0.72	1.27	0	0.40	0.23
62	LG	92894	---	39.9	7.4	3.4	---	11.74	40	---	0.46
69	RIV	62994	62.9	14.2	84	---	0.75	---	30	0.22	---
69	RIV	102794	162.4	106.9	113.6	76.8	1.43	1.39	40	0.66	0.67
85	SD	100594	8.7	15.2	12.7	9.1	0.69	1.67	0	1.75	0.71
100	LG	71594	2.2	4.2	20.4	---	0.11	---	---	1.85	---
100	LG	102594	26.5	10.6	17.6	10.6	1.51	1.00	0	0.40	0.60
129	LG	92894	12	4.2	5	2.4	2.40	1.75	0	0.35	0.48
129	LG	110194	14	7.8	14	6	1.00	1.30	0	0.56	0.43
130	LAN	80494	44.6	19.8	42.6	---	1.05	---	20	0.44	---
130	LAN	100894	47.5	86.3	17.4	4.3	2.73	20.07	50	1.81	0.25
131	LAN	100894	31.3	13.3	24.6	7.7	1.27	1.73	0	0.42	0.31
147	SD	92494	26.3	13.1	10.9	16.9	2.41	0.78	0	0.50	1.56
147	SD	111794	12.6	5.5	14.9	7.3	0.85	0.75	0	0.44	0.49

Notes:

PM is in units of ug/m3;

Multiple Home ID entries reflect multiple home visits;

Town ID as follows: SD = San Dimas, RIV = Riverside/Mira Loma, LG = Lake Gregory, LAN = Lancaster.

"---" represents missing data;

"# Cigarettes smoked" based on Follow-Up Survey.

Table 5-12. Formaldehyde Sampling Results, By Community.

Variable	Community			
	Lake Gregory	Lancaster	Riverside/Mira Loma	San Dimas
Indoor Formaldehyde				
observations	26	24	28	20
average, ug/m3	9.6	10.2	14.8	10.6
std. dev.	8.1	5.5	7.4	6.9
minimum value	0.0	0.0	1.4	0.1
maximum value	38.8	22.2	33.8	28.7
Outdoor Formaldehyde				
observations	3	5	5	5
average, ug/m3	1.7	3.6	3.9	4.2
std. dev.	1.2	3.8	2.1	1.7
minimum value	1.0	1.0	2.1	2.2
maximum value	3.1	10.1	7.5	6.1
Indoor/Outdoor Ratio				
observations	3	5	5	5
average	15.6	4.81	4.24	3.24
std. dev.	21.2	3.42	2.93	3.47
minimum value	1.22	0.7	0.35	0.01
maximum value	39.95	8.88	7.64	9.16

Table 5- 13. Formaldehyde Sampling Results, by Reported Home Air Conditioning Type.

Variable	Air Conditioning Type			
	No A/C	Central or Room A/C	Swamp Cooler Only	Swamp + Central or Room A/C
Indoor Formaldehyde				
observations	31	44	13	5
average, ug/m3	9.1	13.0	10.9	9.7
std. dev.	7.1	6.4	5.6	10.5
minimum value	0.0	0.0	3.5	0.1
maximum value	38.8	28.7	24.6	27.5
Outdoor Formaldehyde				
observations	3	11	2	1
average, ug/m3	2.5	3.3	3.5	5.8
std. dev.	1.3	2.7	0.2	---
minimum value	1.0	1.0	3.4	5.8
maximum value	3.4	10.1	3.7	5.8
Indoor/Outdoor Ratio				
observations	3	11	2	1
average	14.53	5.24	2.43	0.01
std. dev.	22.02	3.07	0.28	---
minimum value	1.22	0.35	2.23	0.01
maximum value	39.95	9.16	2.63	0.01

Table 5-14. Air Exchange Rate Sampling Results, By Community.

Variable	Community			
	Lake Gregory	Lancaster	Riverside/Mira Loma	San Dimas
AER, traditional*				
observations	36	46	46	33
average, 1/hr	0.9	0.7	0.7	0.7
std. dev.	0.4	0.6	0.5	0.5
minimum value	0.0	0.0	0.0	0.0
maximum value	1.8	2.5	1.9	2.0
AER, corrected ***				
observations	36	46	46	33
average, 1/hr	0.9	0.7	0.8	0.8
std. dev.	0.5	0.6	0.6	0.5
minimum value	0.0	0.0	0.0	0.0
maximum value	1.9	2.7	2.0	2.3

* traditional refers to use of conventional home volume measurement approach.
*** corrected refers to home volume correction for furniture and cabinetry (see text).

Table 5-15 . Air Exchange Results, By Reported Home Air Conditioning Type.

Variable	Air Conditioning Type			
	No A/C	Central or Room A/C	Swamp Cooler Only	Swamp + Central or Room A/C
Air Exchange Rate, Traditional*				
observations	43	84	19	8
average, 1/hr	0.9	0.6	1.2	0.6
std. dev.	0.5	0.4	0.6	0.4
minimum value	0.0	0.0	0.2	0.0
maximum value	1.9	2.3	2.5	1.2
Air Exchange Rate, Corrected***				
observations	43	84	19	8
average, 1/hr	0.9	0.6	1.3	0.6
std. dev.	0.5	0.5	0.7	0.4
minimum value	0.0	0.0	0.2	0.0
maximum value	2.0	2.7	2.7	1.3

* traditional refers to use of conventional home volume measurement approach.
*** corrected refers to home volume coorection for furniture and cabinetry (see text).

Table 5-16. Summary of Home Characteristics in which TWS was deployed.

Home ID	City	Home Type	Decade Built	Home Vol	AC Type	# Bdrms	Smokers	Dust Control	Pets	Start Date	End Date
30	RIV/ML	sfr/detached	1980's	323	central	3	no	3	none	8/18/94	9/1/94
9	RIV/ML	mobile home	1980's	248	swamp	2	no	3	cat/dog	8/18/94	9/1/94
7	SD	sfr/detached	1950's	353	none	2	no	3	cat/dog	9/1/94	9/14/94
1	SD	sfr/detached	1980's	522	central	3	no	4	none	9/1/94	9/14/94
43	LG	sfr/detached	1980's	528	none	3	yes	4	cat/dog	9/15/94	9/29/94
27	LG	sfr/detached	1970's	476	none	2	no	4	cat/dog	9/15/94	9/29/94
44	LG	sfr/detached	1970's	409	none	2	no	4	cat/dog	9/29/94	10/13/94
45	LG	sfr/detached	1960's	211	none	1	no	4	cat/dog	9/29/94	10/13/94
59	LAN	sfr/detached	1980's	468	central	3	no	5	none	10/14/94	10/27/94
83	LAN	sfr/detached	1950's	568	central	>=6	no	1	cat/dog	10/14/94	10/27/94
116	LAN	sfr/detached	1970's	415	central	2	no	2	cat/dog	10/27/94	11/10/94
56	LAN	sfr/detached	1950's	283	swamp	2	no	4	cat/dog	10/27/94	11/10/94

NOTES:

1) Home ID's are from residential study field designation.
2) City refers to Riverside/Mira Loma , San Dimas, Lake Gregory, and Lancaster.
3) "sfr/detached" refers to single family residence with detached garage.
4) Home Vol refers to calculated home volume in cubic meters.
5) Dust Control refers to technician-subjective review of home on 1-5 scale, with 1 being very cluttered and dusty, to 5 being excellent cleaning practices.

Table 5-17. Summary of Residential Two-Week Sampler Data

Particle Mass, Ions and Inorganic Acid Concentrations (µg/m3)

Dates	Site	Gravimetric Mass	PM2.5 SO4=	PM2.5 NO3-	PM2.5 NH4+	HCl	HNO3	Flag
8/19-9/1	House 030 Inside	22.7	3.0	3.7	1.7	0.1	1.9	
8/19-9/1	House 030 Inside Replicate	18.5	3.0	4.2	1.8	1.0	1.9	
8/19-9/1	House 030 Outside	24.4	4.0	10.2	4.4	1.1	9.5	
8/19-9/1	House 030 Outside Replicate	25.4	3.6	10.6	4.0	0.6	8.1	
8/19-9/1	House 009 Inside	23.7	2.8	12.0	4.8	0.0	4.7	
8/19-9/1	House 009 Outside	30.0	4.0	14.0	5.3	3.0	7.9	
9/1-9/15	House 007 Inside	16.9	3.3	8.7	3.3	0.4	5.6	
9/1-9/15	House 007 Inside Replicate	17.2	3.3	7.1	2.8	0.3	6.8	
9/1-9/15	House 007 Outside	16.1	3.5	7.7	3.3	1.1	12.8	
9/1-9/15	House 007 Outside Replicate	16.3	3.1	5.0	2.4	0.8	13.6	
9/1-9/15	House 001 Inside	18.0	2.6	5.8	2.5	0.5	3.5	1
9/1-9/15	House 001 Outside	6.6	1.3	2.8	1.1	0.6	8.0	1
9/15-9/29	House 043 Inside	32.3	1.3	4.2	1.1	<0.1	2.0	
9/15-9/29	House 043 Outside	6.5	1.3	2.4	1.1	0.2	6.1	
9/15-9/29	House 027 Inside	9.3	1.3	2.1	1.1	<0.1	1.8	
9/15-9/29	House 027 Inside Replicate	6.7	1.1	1.8	0.9	<0.1	2.6	
9/15-9/29	House 027 Outside	5.7	1.3	1.9	1.0	<0.1	4.9	
9/15-9/29	House 027 Outside Replicate	6.2	1.4	2.2	1.2	<0.1	4.6	
9/29-10/13	House 044 Inside	6.3	0.8	2.3	0.8	0.2	1.5	
9/29-10/13	House 044 Inside Replicate	6.5	0.7	2.4	0.9	0.3	1.1	
9/29-10/13	House 044 Outside	3.8	0.9	1.3	0.6	0.5	2.5	
9/29-10/13	House 044 Outside Replicate	4.1	0.8	1.3	0.7	0.9	2.5	
9/29-10/13	House 045 Inside	11.0	0.9	0.5	0.3	0.2	0.5	
9/29-10/13	House 045 Outside	6.1	0.9	1.6	0.8	<0.1	1.5	
10/14-10/27	House 059 Inside	<2.0	<0.1	2.5	0.7	0.5	1.2	2
10/14-10/27	House 059 Outside	8.8	0.7	2.2	0.9	2.0	5.1	
10/14-10/27	House 083 Inside	13.1	0.5	1.0	0.4	1.1	0.8	
10/14-10/27	House 083 Inside Replicate	10.5	0.5	1.1	0.5	1.3	1.0	
10/14-10/27	House 083 Outside	11.5	1.3	4.6	1.7	2.3	6.2	3
10/14-10/27	House 083 Outside Replicate	10.0	1.3	4.2	1.6	4.9	8.3	3
10/27-11/10	House 116 Inside	8.0	0.6	0.4	0.2	0.5	0.5	
10/27-11/10	House 116 Inside Replicate	9.2	0.9	0.5	0.2	0.4	0.8	
10/27-11/10	House116 Outside	5.8	0.8	3.4	1.2	0.4	3.2	
10/27-11/10	House116 Outside Replicate	9.0	1.1	3.1	1.0	1.0	4.5	
10/27-11/10	House 056 Inside	2.2	0.5	0.4	0.2	0.3	1.2	
10/27-11/10	House 056 Outside	5.8	0.8	2.9	1.0	0.3	3.8	

Flags:

1 = Field notes indicate power outage at house, sample time taken from elapsed time meter

2 = Teflon filter sample looks like blank.

3 = Outdoor sampler unplugged for approx. 8 days, sample volume from elapsed time meter.

Table 5-18. Summary of Residential Two-Week Sampler Data

Organic Acid Concentrations (µg/m3)

		Front Filter		Backup Filter		Sum of Filters		
Dates	Site	Acetate	Formate	Acetate	Formate	Acetate	Formate	Flag
8/19-9/1	House 030 Inside	13.9	23.1	33.3	8.5	47.2	31.6	
8/19-9/1	House 030 Outside	14.6	5.4	3.6	<0.3	18.2	5.4	
8/19-9/1	House 009 Inside	64.5	27.9	5.4	<0.3	69.9	27.9	
8/19-9/1	House 009 Inside Repliate	56.9	26.6	15.8	<0.3	72.7	26.6	
8/19-9/1	House 009 Outside	13.9	5.9	7.0	<0.3	20.9	5.9	
8/19-9/1	House 009 Outside Repliate	13.4	5.3	7.3	<0.3	20.7	5.3	
9/1-9/15	House 007Inside	16.7	16.4	17.0	4.3	33.7	20.6	
9/1-9/15	House 007 Outside	14.8	8.8	5.8	<0.3	20.6	8.8	
9/1-9/15	House 001 Inside	25.2	44.0	42.4	8.8	67.6	52.9	1
9/1-9/15	House 001 Inside Repliate	21.0	48.5	33.7	13.0	54.7	61.4	1
9/1-9/15	House 001 Outside	6.6	3.6	1.8	<0.3	8.5	3.6	1
9/1-9/15	House 001 Outside Repliate	6.2	3.3	1.1	<0.3	7.4	3.3	1
9/15-9/29	House 043 Inside	39.0	28.9	34.1	2.1	73.1	31.0	
9/15-9/29	House 043 Inside Repliate	30.9	27.9	33.4	3.0	64.3	30.8	
9/15-9/29	House 043 Outside	8.9	4.1	1.4	<0.3	10.4	4.1	
9/15-9/29	House 043 Outside Repliate	7.9	3.5	1.2	<0.3	9.2	3.5	
9/15-9/29	House 027 Inside	34.4	21.4	14.2	0.5	48.6	21.8	
9/15-9/29	House 027 Outside	8.0	4.1	3.3	<0.3	11.3	4.1	
9/29-10/13	House 044 Inside	28.5	30.8	28.5	1.8	57.0	32.6	
9/29-10/13	House 044 Outside	4.5	1.7	<0.3	<0.3	4.5	1.7	
9/29-10/13	House 045 Inside	24.7	27.1	21.1	1.0	45.8	28.1	
9/29-10/13	House 045 Inside Repliate	19.9	25.8	25.7	2.0	45.6	27.8	
9/29-10/13	House 045 Outside	4.6	1.1	<0.3	<0.3	4.6	1.1	
9/29-10/13	House 045 Outside Repliate	4.5	1.1	<0.3	<0.3	4.5	1.1	
10/14-10/27	House 059 Inside	21.6	27.6	26.2	1.5	47.9	29.1	
10/14-10/27	House 059 Inside Repliate	25.1	28.8	26.7	1.7	51.8	30.5	
10/14-10/27	House 059 Outside	7.2	3.5	1.0	<0.3	8.2	3.5	
10/14-10/27	House 059 Outside Repliate	7.2	1.3	0.6	<0.3	7.9	1.3	
10/14-10/27	House 083 Inside	28.7	18.5	32.0	1.0	60.8	19.5	
10/14-10/27	House 083 Outside	5.3	2.3	0.6	<0.3	5.8	2.3	3
10/27-11/10	House 116 Inside	22.8	25.7	27.5	1.6	50.3	27.3	
10/27-11/10	House116 Outside	5.8	2.3	0.5	<0.3	6.3	2.3	
10/27-11/10	House 056 Inside	22.6	19.9	22.3	0.6	44.8	20.5	
10/27-11/10	House 056 Inside Repliate	15.3	18.9	19.3	1.2	34.6	20.2	
10/27-11/10	House 056 Outside	6.8	3.1	0.9	<0.3	7.7	3.1	
10/27-11/10	House 056 Outside Repliate	6.9	3.4	0.9	<0.3	7.8	3.4	

Flags:

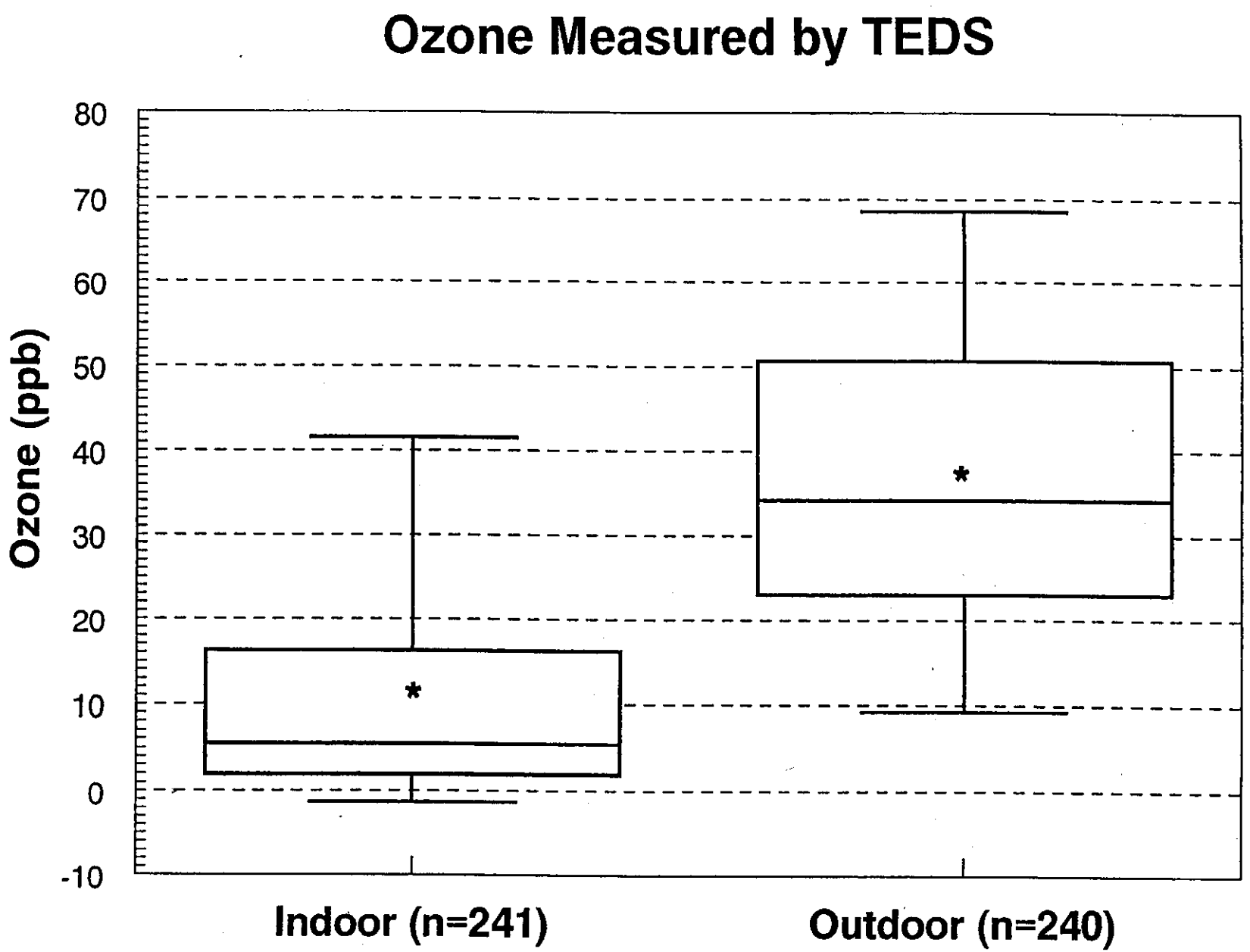
1 = Field notes indicate power outage at house, sample time taken from elapsed time meter.

2 = Teflon filter sample looks like blank.

3 = Outdoor sampler unplugged for approx. 8 days, sample time from elapsed time meter.

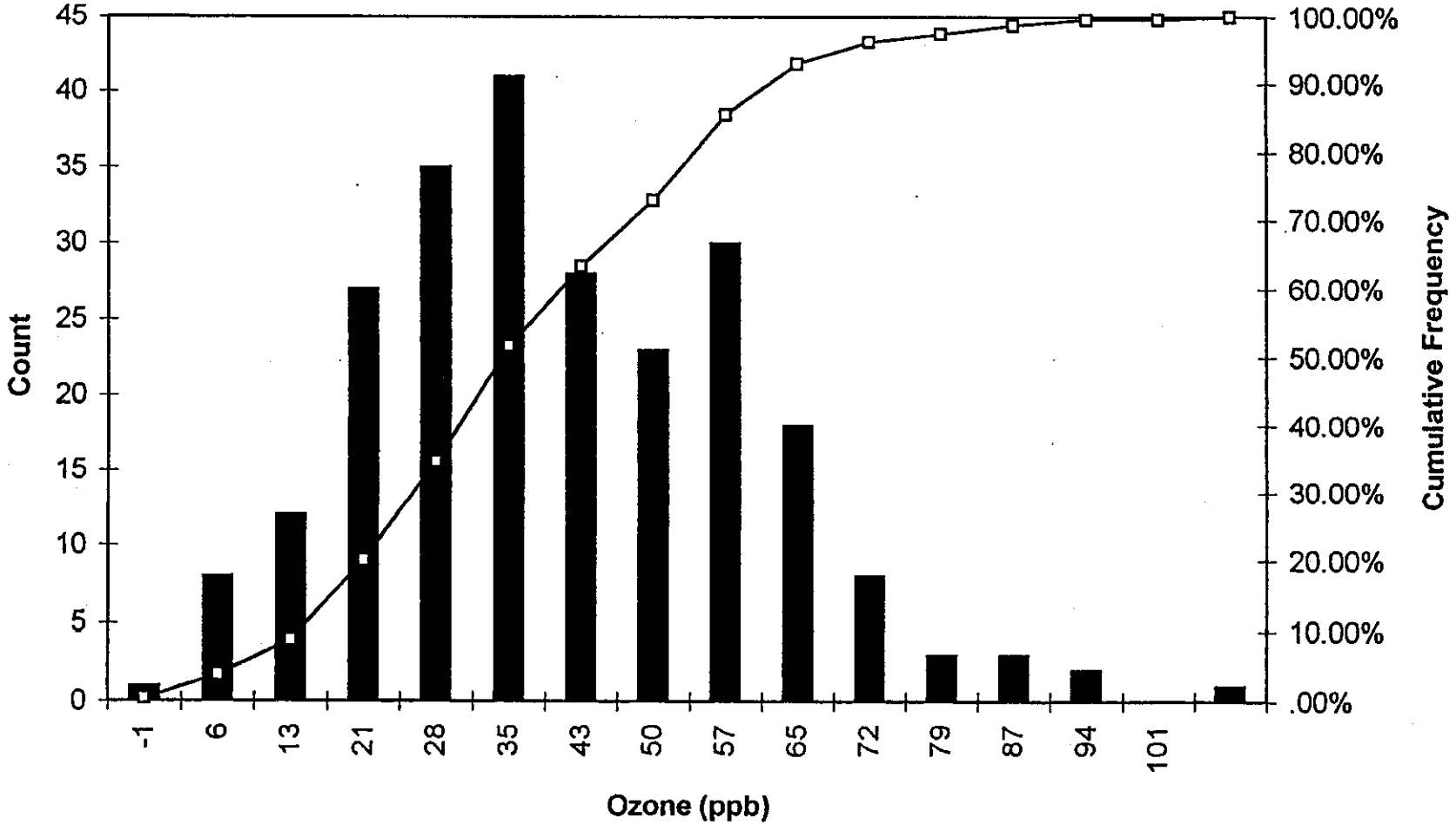
Table 5-19. Summary of coefficients of varation for Two-Week Sampler measurements.

Measurement Variable	Pooled Standard Deviation (in ug/m3)	Pooled Standard Deviation (%)
Sulfate	0.15	9
Nitrate	0.54	15
Ammonium	0.26	17
Hydrochloric Acid	0.36	57
Nitric Acid	0.57	14
Fine Particle (PM2.5) Mass	1.4	12
Acetic Acid	4	12
Formic Acid	2	11



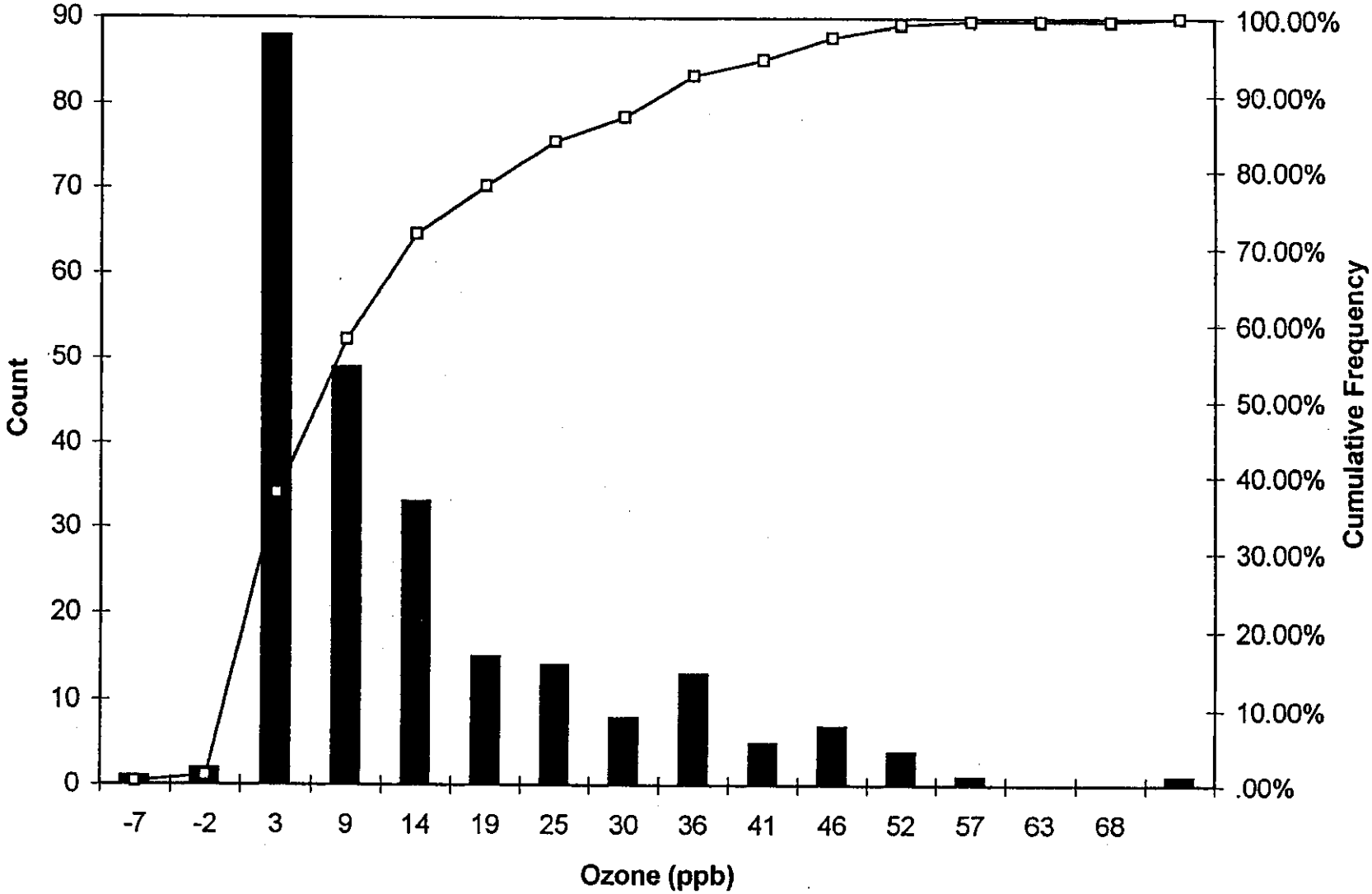
Note: 1) Negative ozone values are due to blank values exceeding measured ozone values. Measured ozone values below the LOD (5 ppb) were corrected by replacing them with a zero value in the statistical analyses.

Figure 5.2-1. Box-whisker plots of indoor and outdoor ozone measurements by TEDS.



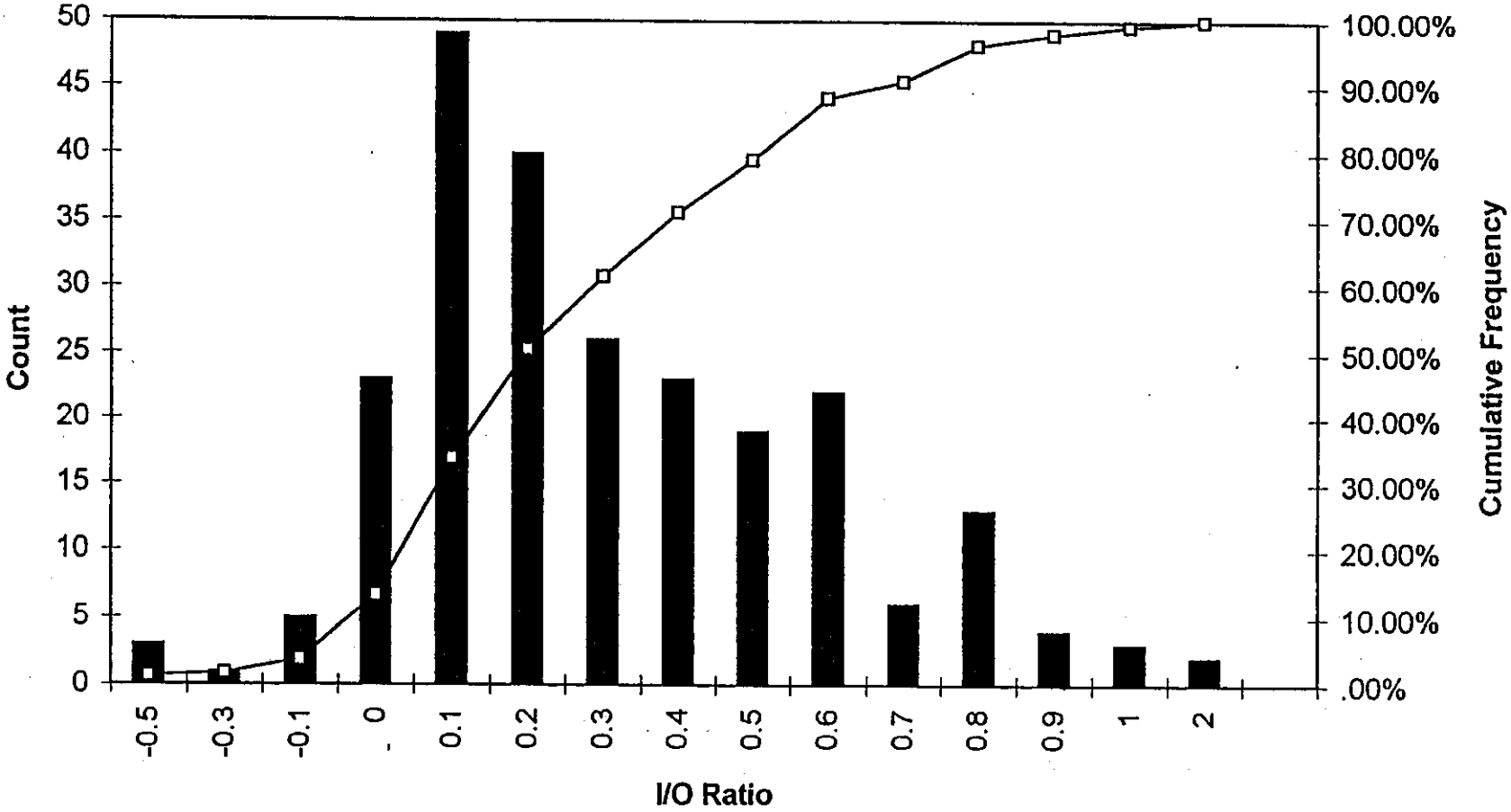
Note: 1) Negative ozone values are due to blank values exceeding measured ozone values. Measured ozone values below the LOD (5 ppb) were corrected by replacing them with a zero value in the statistical analyses.

Figure 5.2-2. Histogram of outdoor ozone measurements by TEDS.



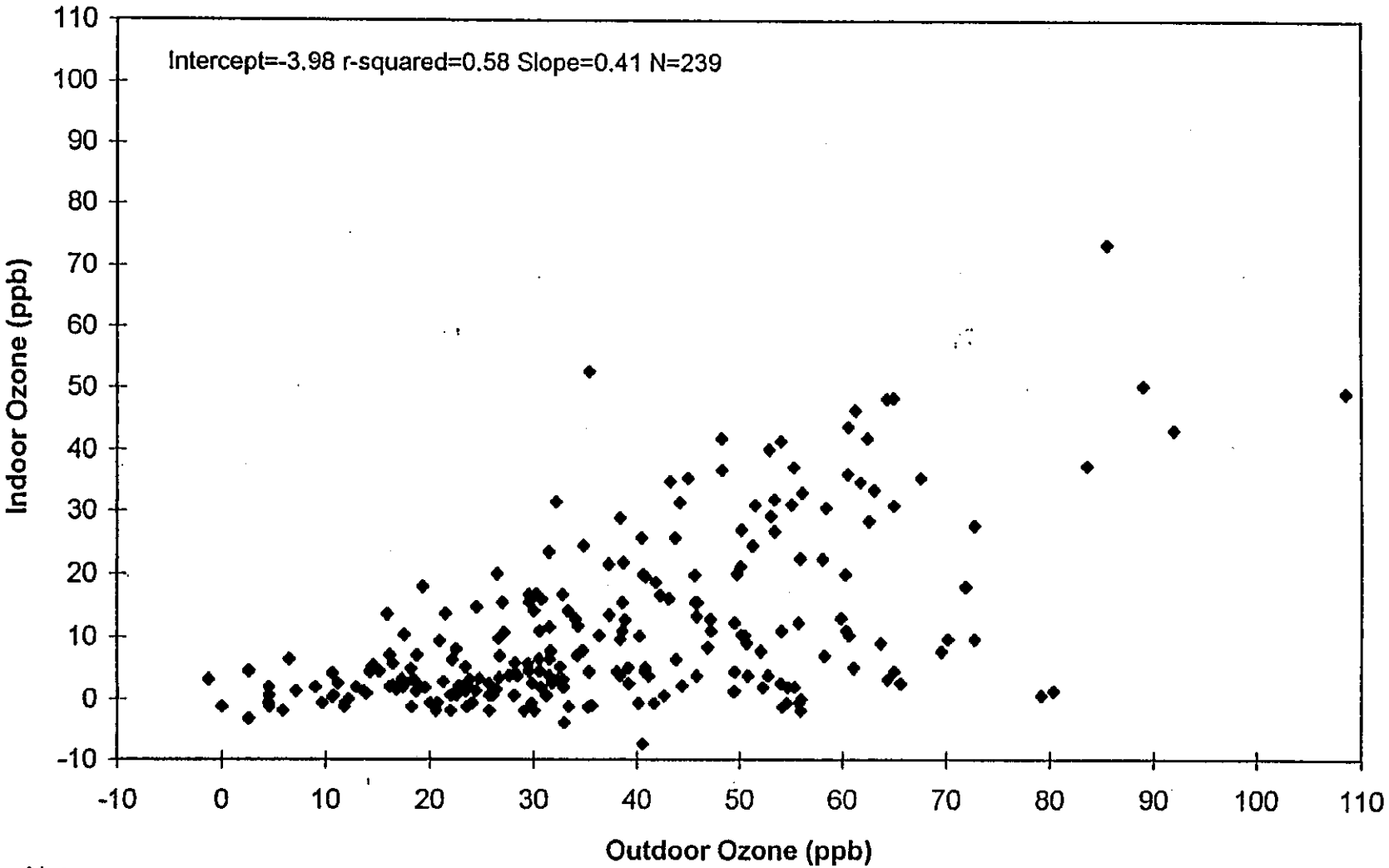
Note: 1) Negative ozone values are due to blank values exceeding measured ozone values. Measured ozone values below the LOD (5 ppb) were corrected by replacing them with a zero value in the statistical analyses.

Figure 5.2-3. Histogram of indoor ozone measurements by TEDS.



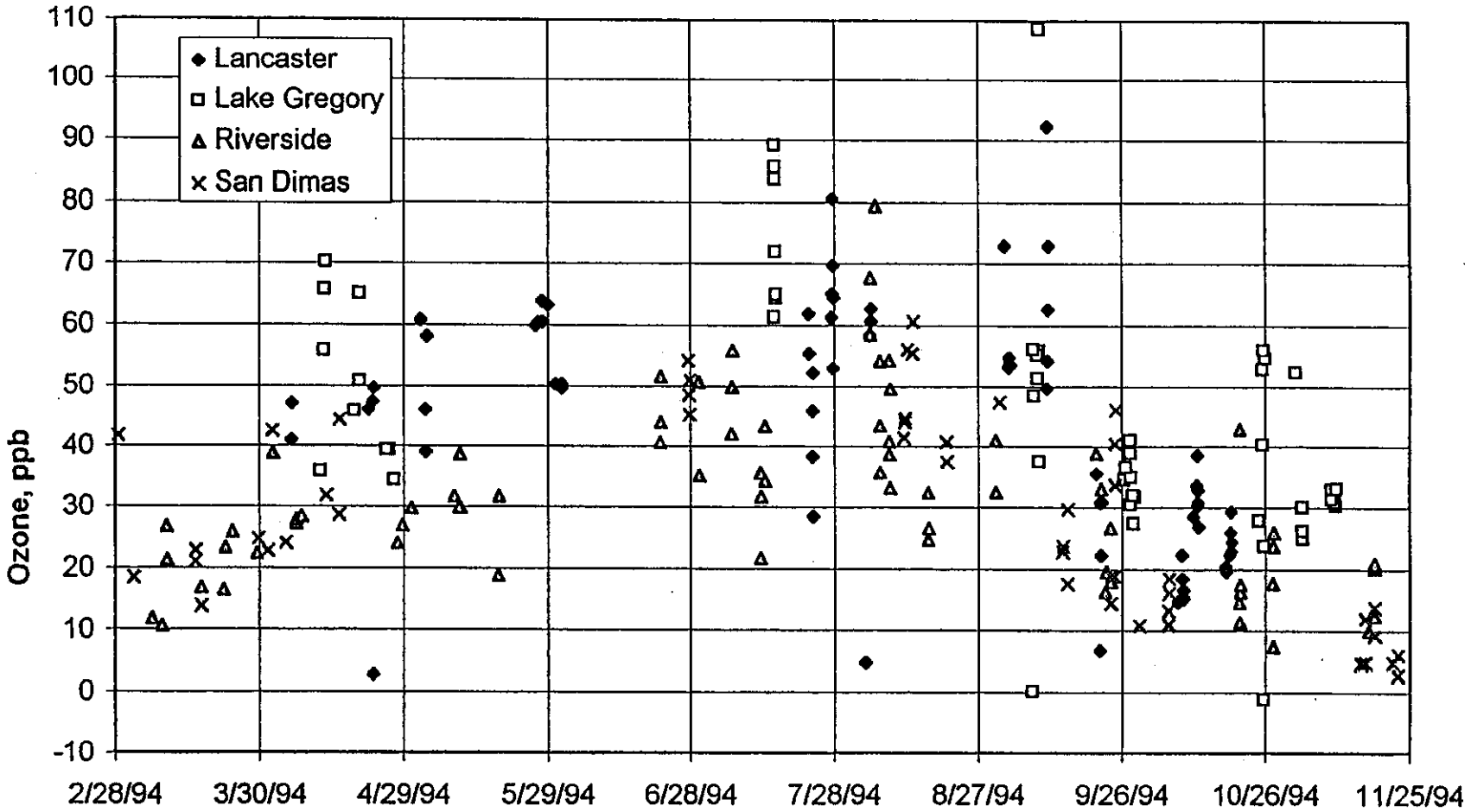
Note: 1) Negative ozone values are due to blank values exceeding measured ozone values. Measured ozone values below the LOD (5 ppb) were corrected by replacing them with a zero value in the statistical analyses.

Figure 5.2-4. Histogram of ozone I/O ratio measured by TEDS.



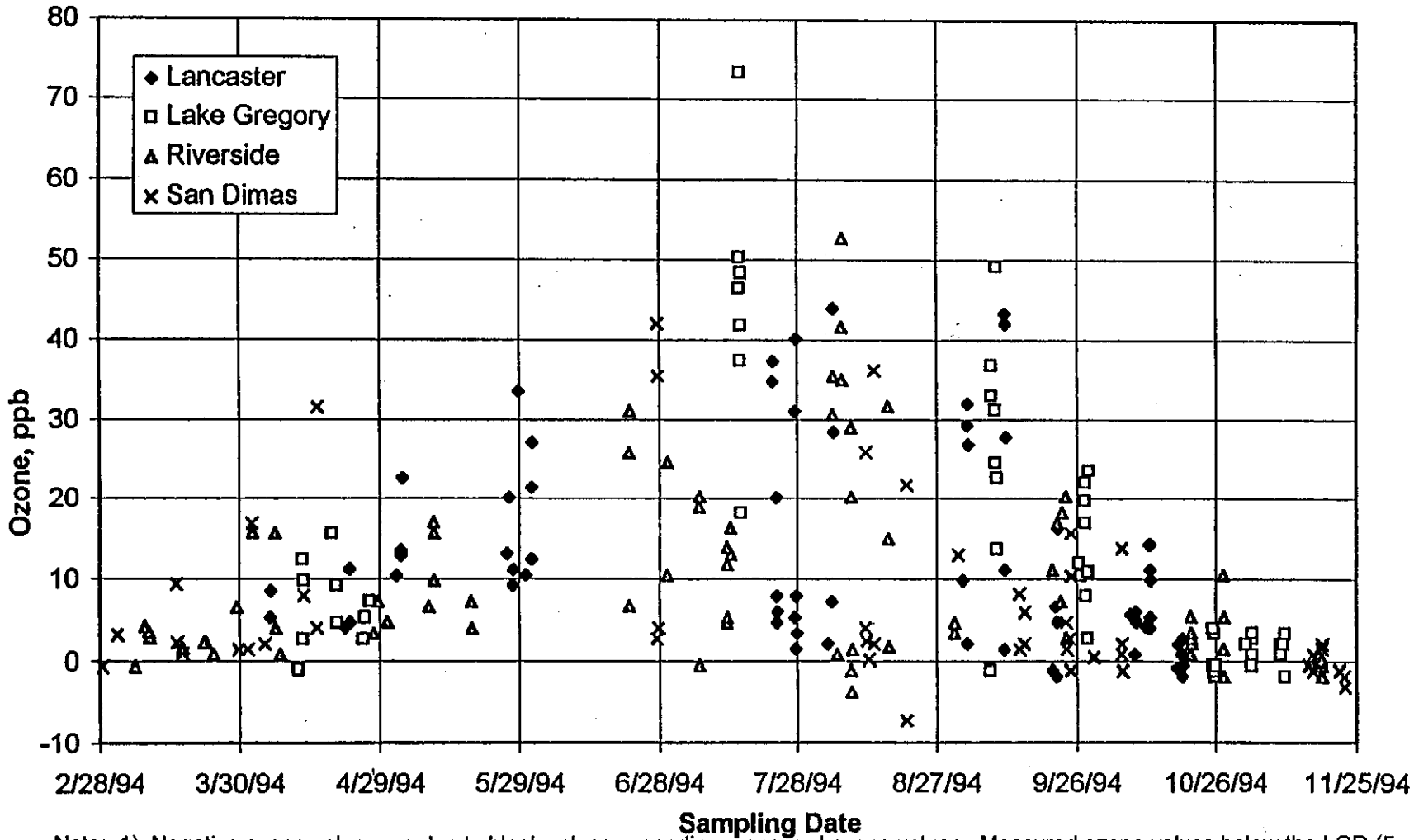
Note: 1) Negative ozone values are due to blank values exceeding measured ozone values. Measured ozone values below the LOD (5 ppb) were corrected by replacing them with a zero value in the statistical analyses.

Figure 5.2-5. Scatterplot of indoor and outdoor ozone measured by TEDS.



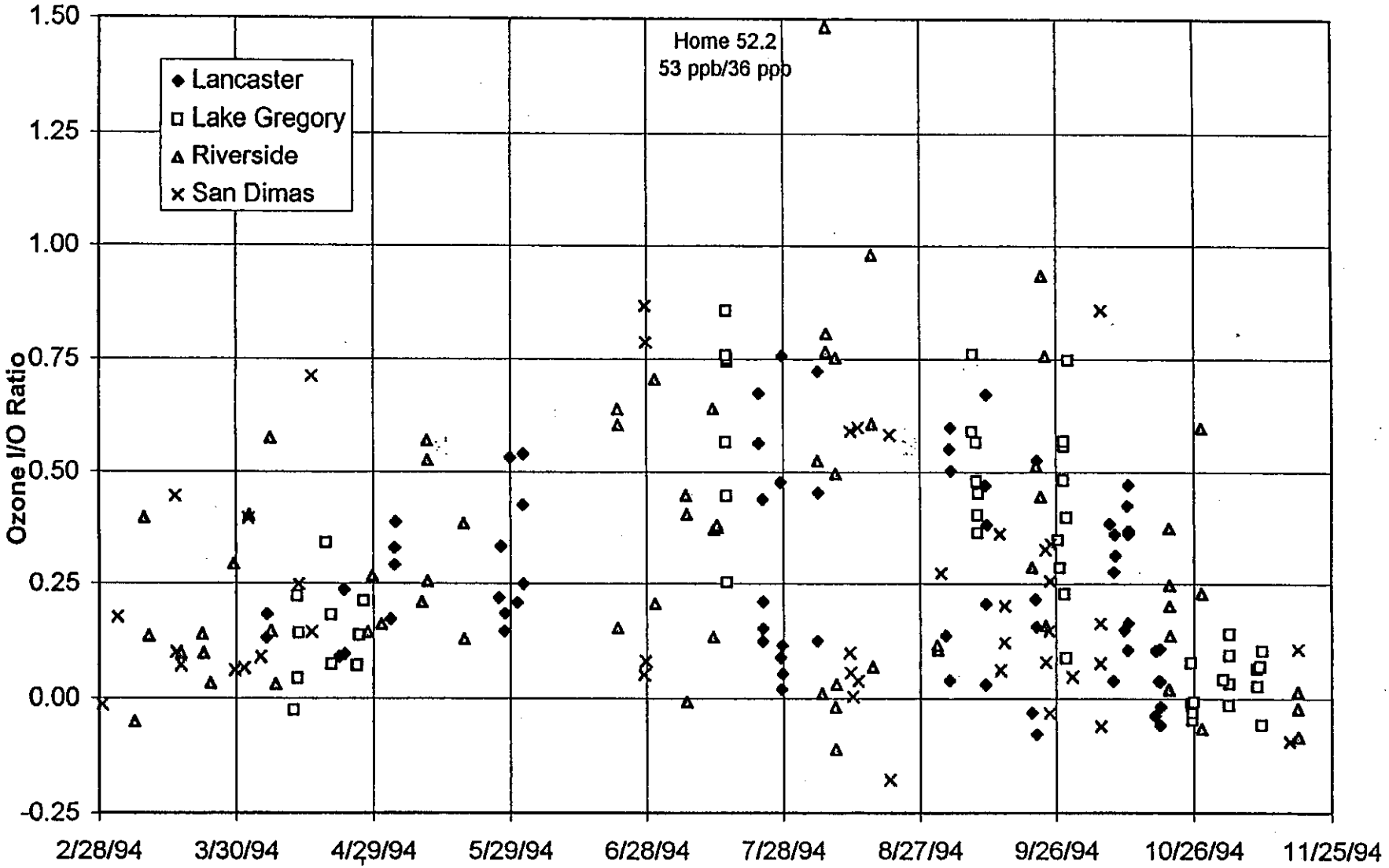
Note: 1) Negative ozone values are due to blank values exceeding measured ozone values. Measured ozone values below the LOD (5 ppb) were corrected by replacing them with a zero value in the statistical analyses.

Figure 5.2-6. Temporal pattern of outdoor TEDS ozone at home sites for each of the four communities.



Note: 1) Negative ozone values are due to blank values exceeding measured ozone values. Measured ozone values below the LOD (5 ppb) were corrected by replacing them with a zero value in the statistical analyses.

Figure 5.2-7. Temporal pattern of indoor TEDS ozone at home sites for each of the four communities.



Note: 1) Negative ozone values are due to blank values exceeding measured ozone values. Measured ozone values below the LOD (5 ppb) were corrected by replacing them with a zero value in the statistical analyses.

Figure 5.2-8. Temporal pattern of TEDS ozone I/O ratio at home sites for each of the four communities.

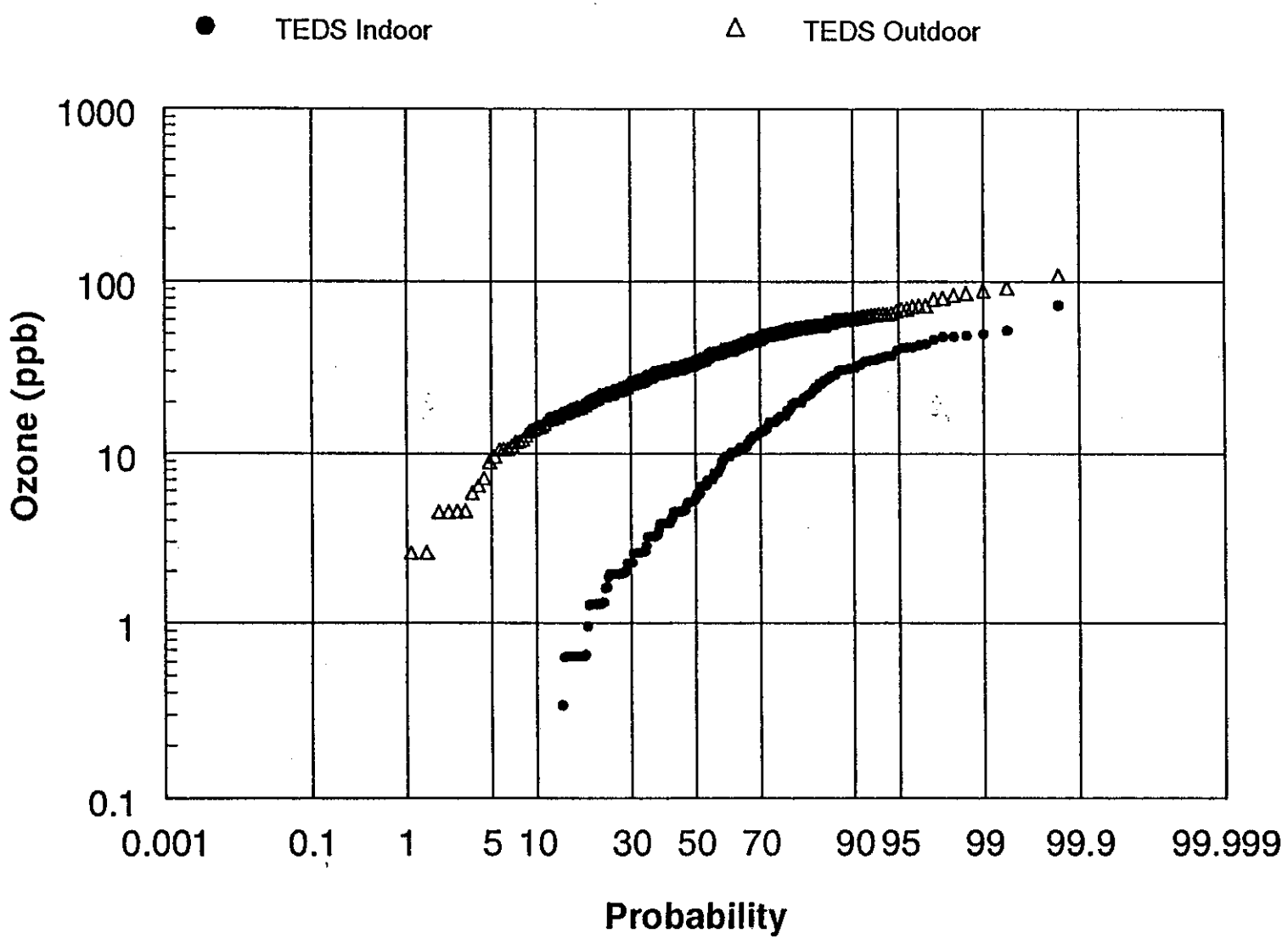


Figure 5.2-9

Figure 5.2-9. Log-probability plots of indoor and outdoor ozone measurements by TEDS.

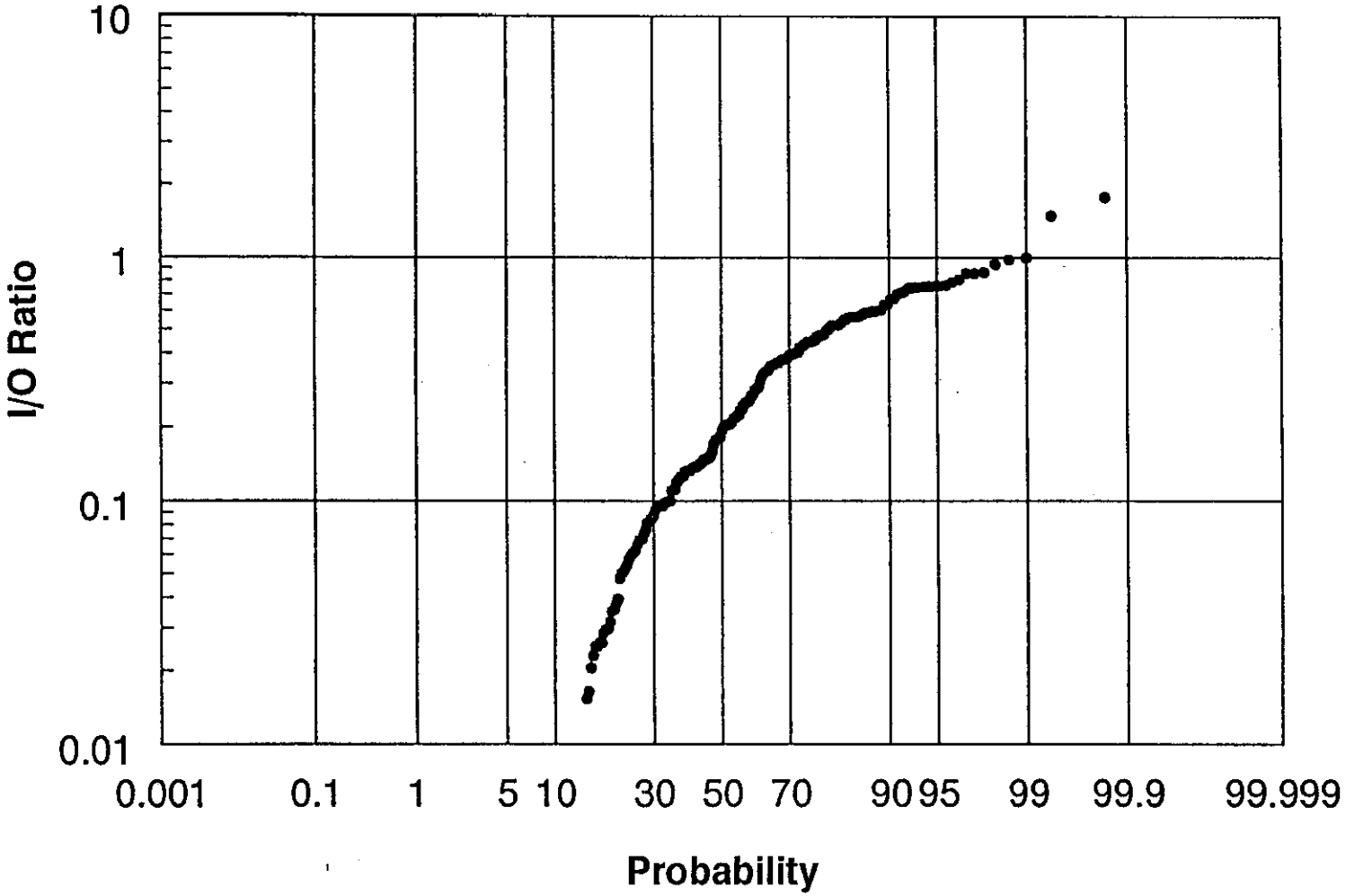


Figure 5.2-10

Figure 5.2-10. Log-probability plot of ozone I/O ratio measured by TEDS.

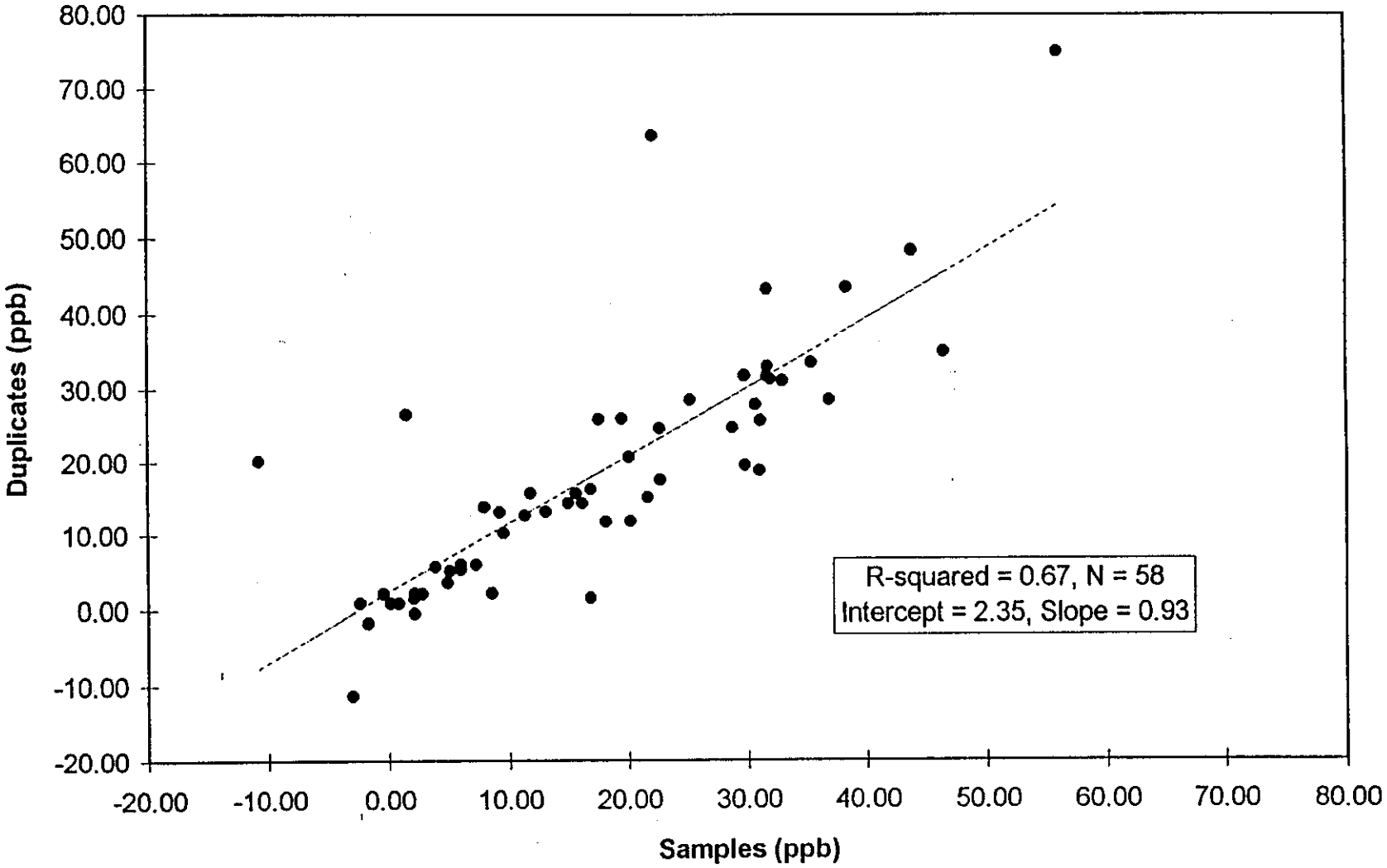


Figure 5.2-11. TEDS ozone co-located samples.

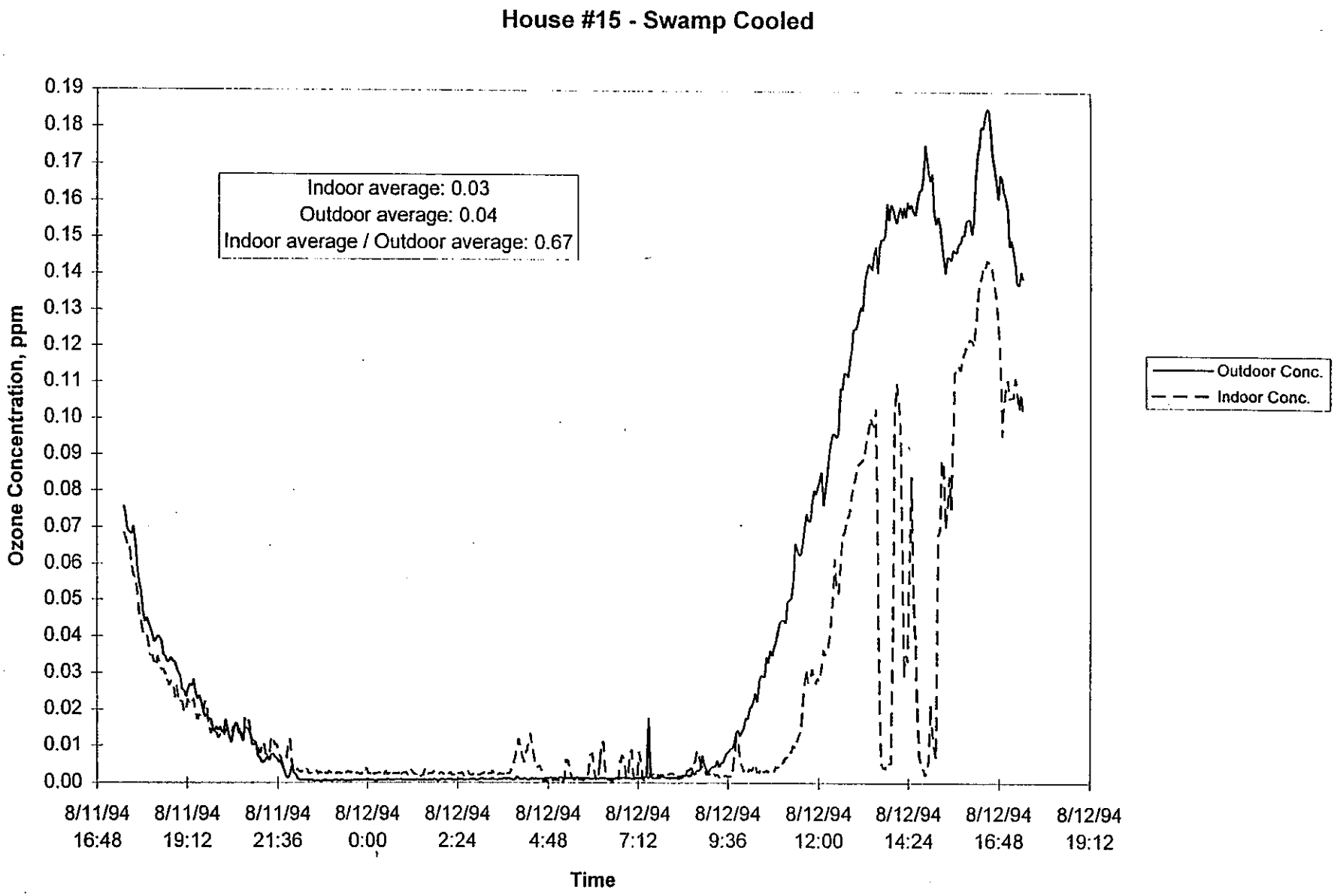


Figure 5.2-12. Continuous ozone in a swamp-cooled home - Home #15.

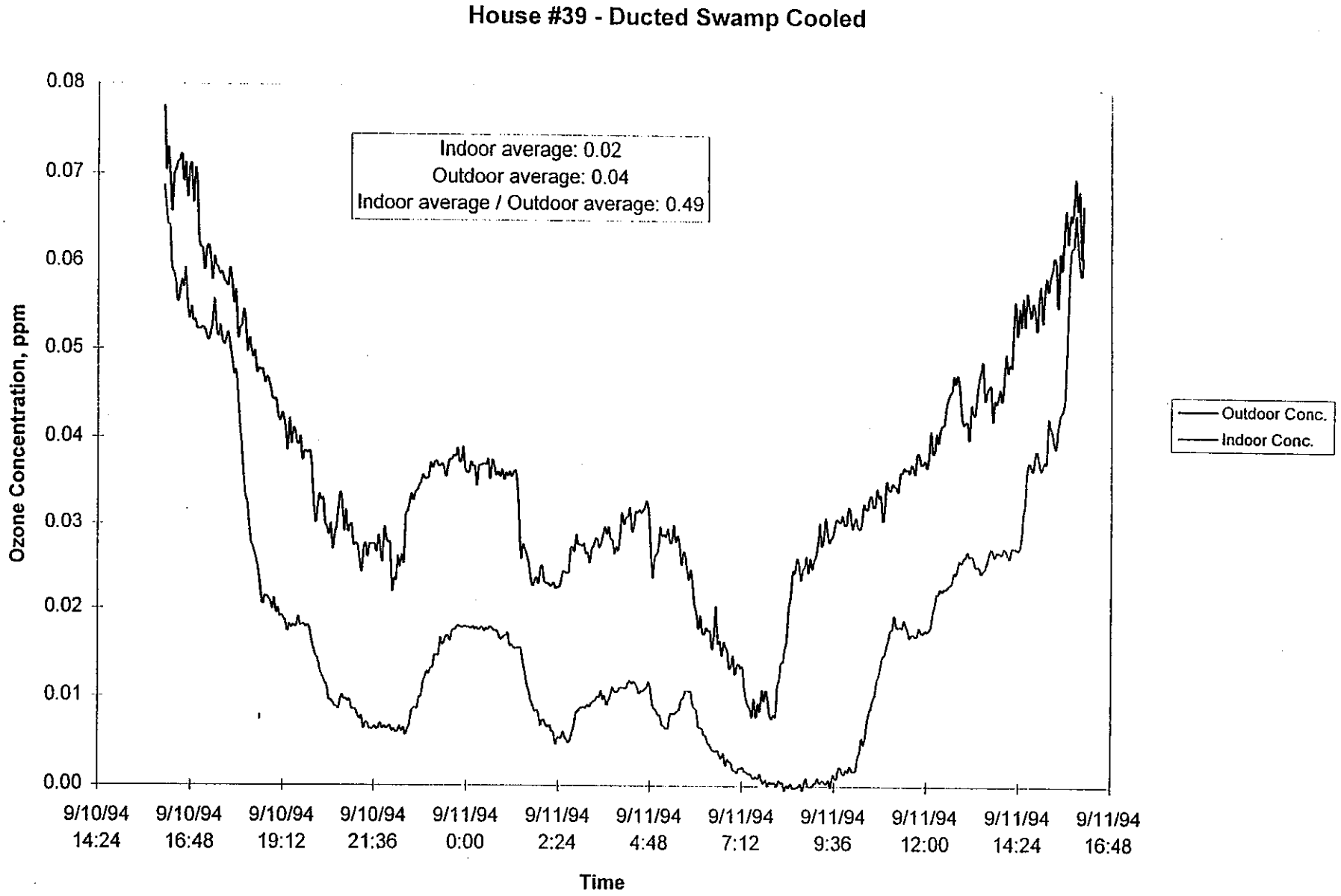


Figure 5.2-13. Continuous ozone in a swamp-cooled home - Home #39.

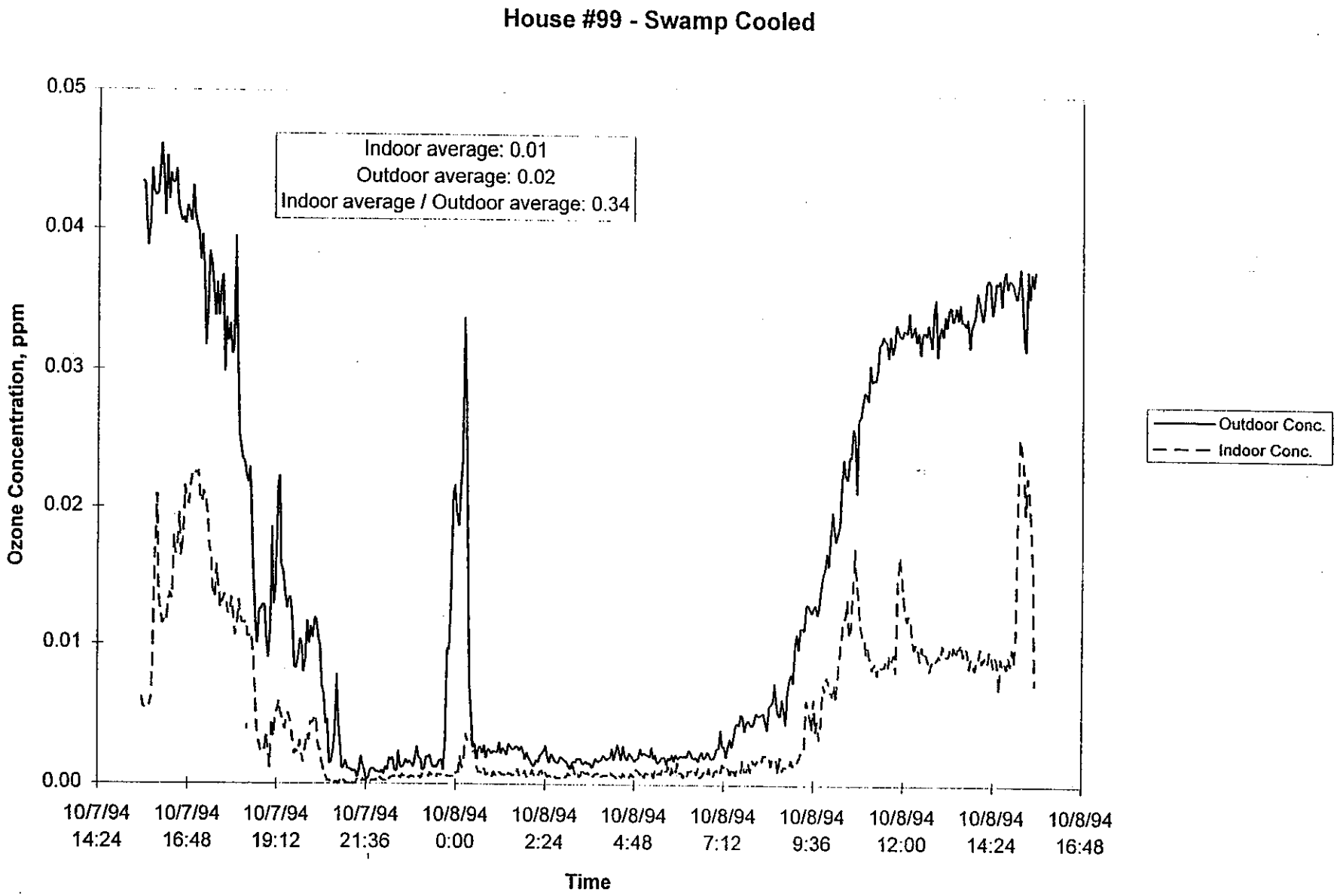


Figure 5.2-14. Continuous ozone in a swamp-cooled home - Home #99.

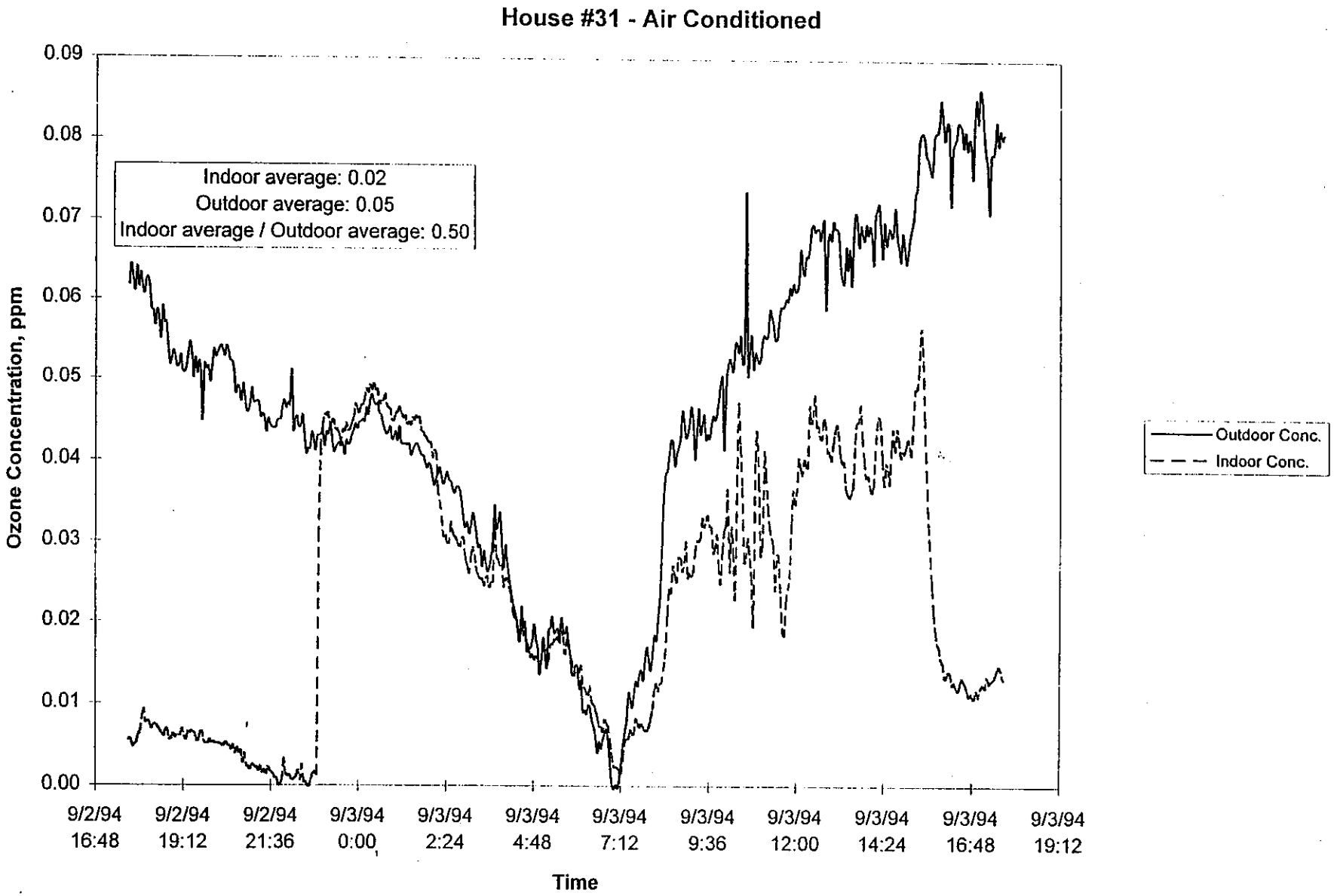


Figure 5.2-15. Continuous ozone in a refrigerant-cooled home - Home #31.

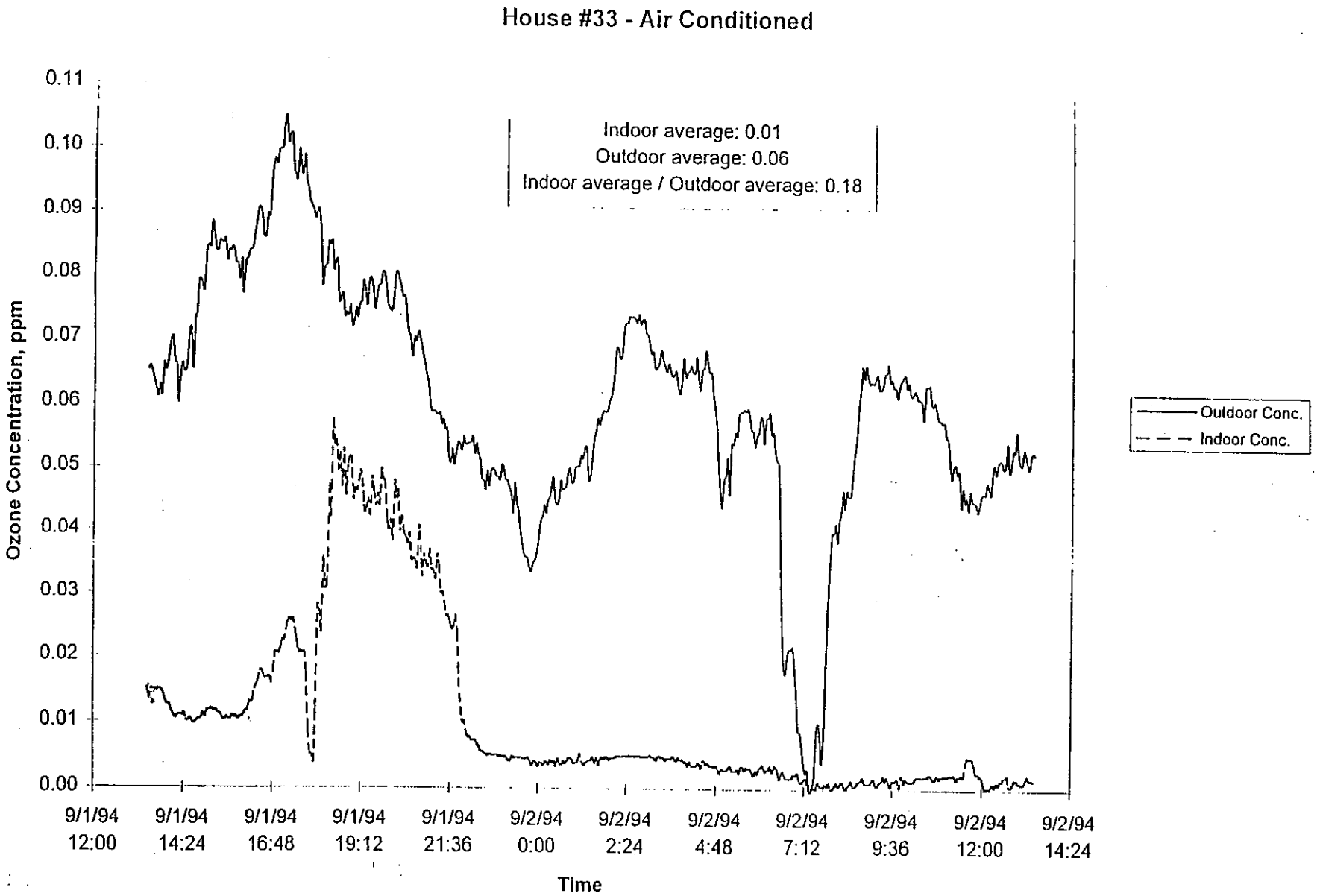
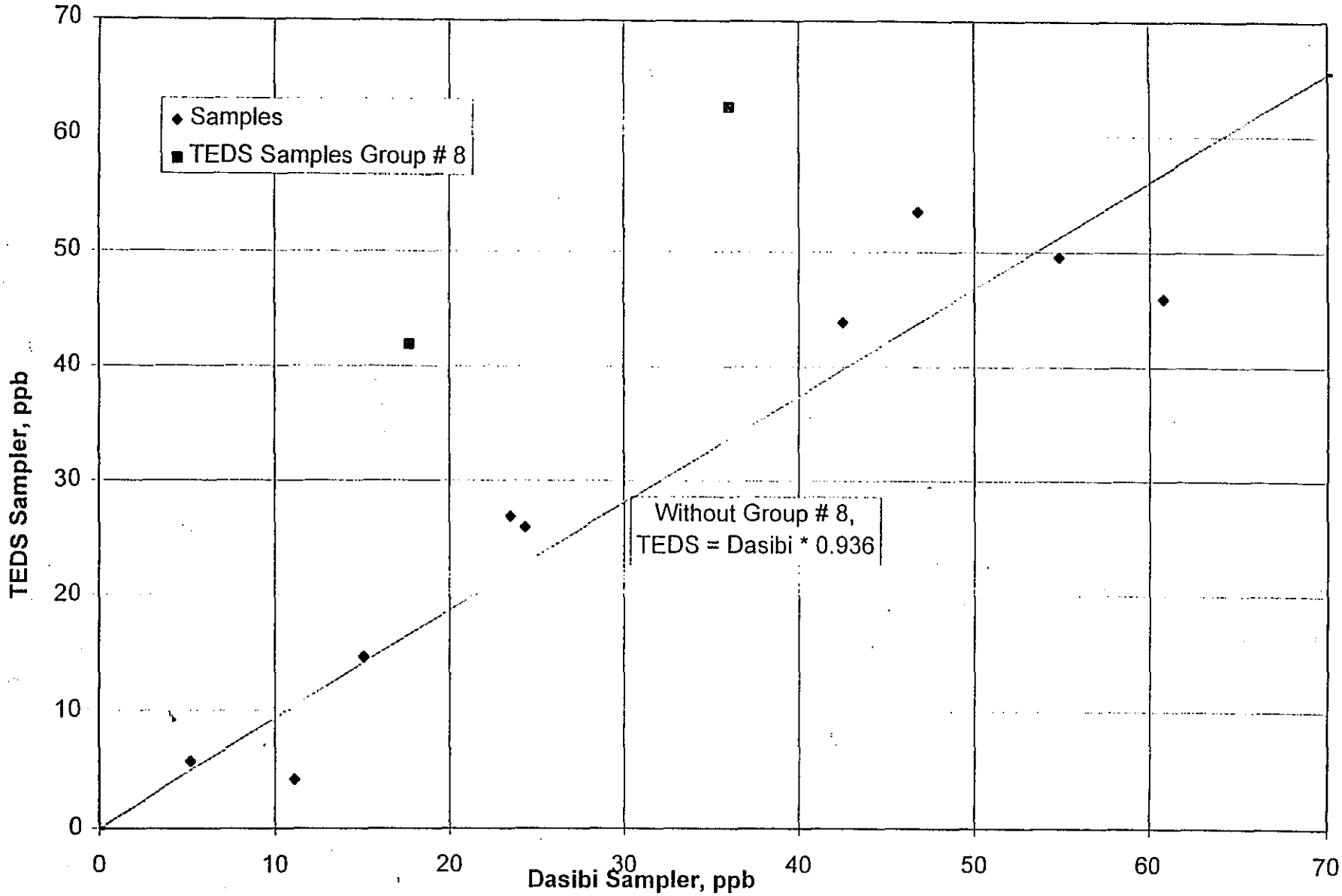


Figure 5.2-16. Continuous ozone in a refrigerant-cooled home - Home #33.



daciteds.xls

Figure 5.2-17. Comparison of ozone by Dasibi and TEDS samplers.

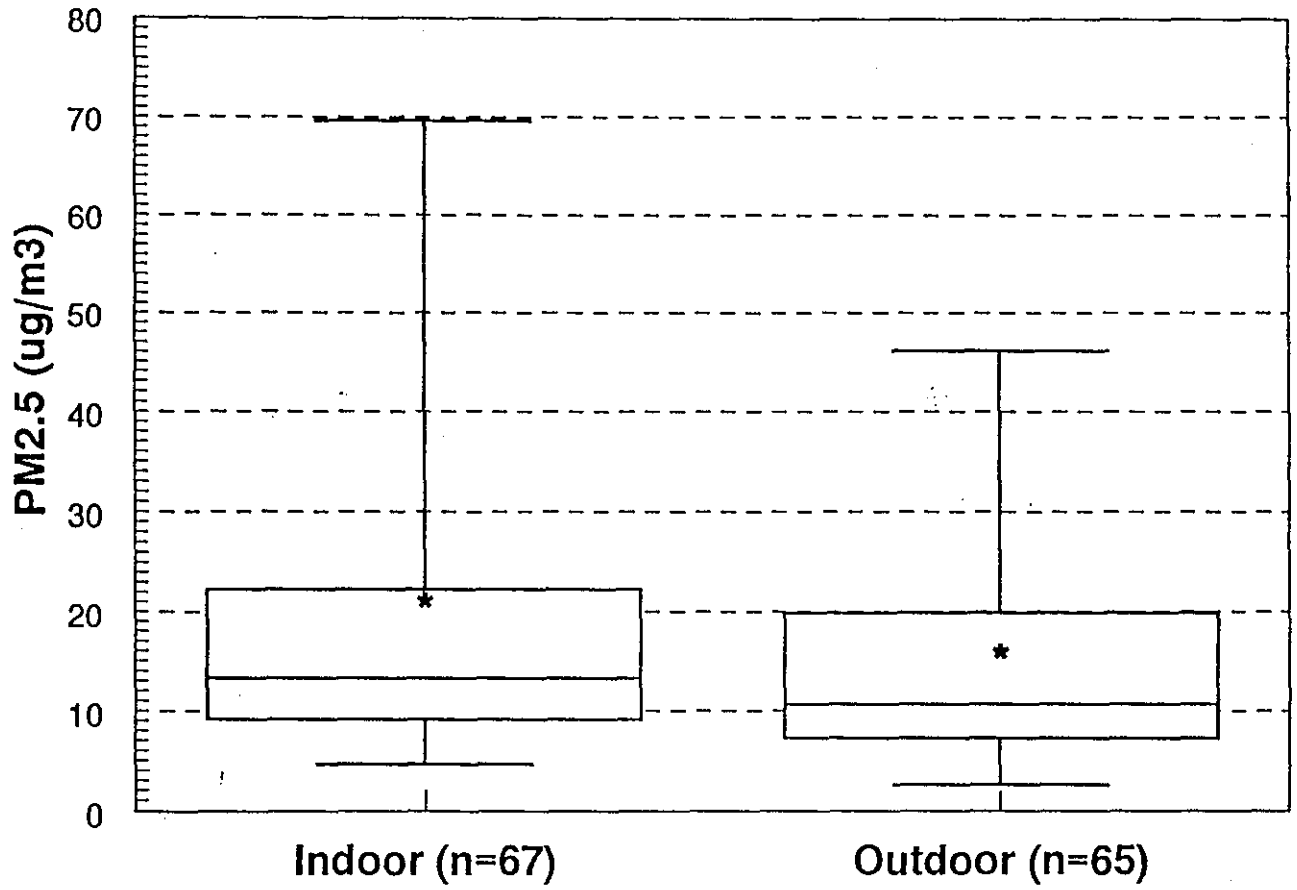


Figure 5.2-18. Box-whisker plots of PM2.5 concentrations measured with Teflon Filters.

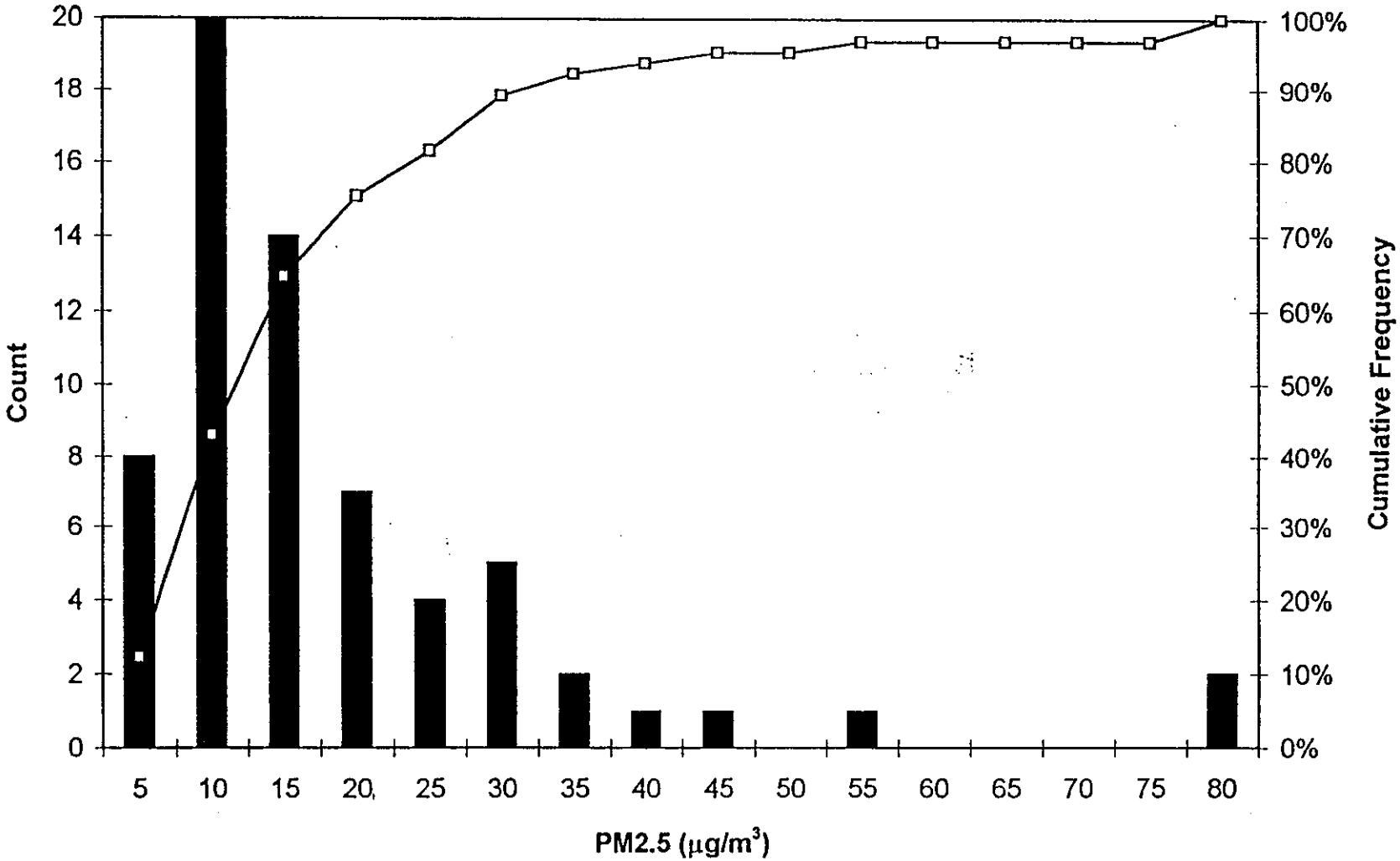


Figure 5.2-19. Histogram of outdoor PM2.5 measured with Teflon Filters.

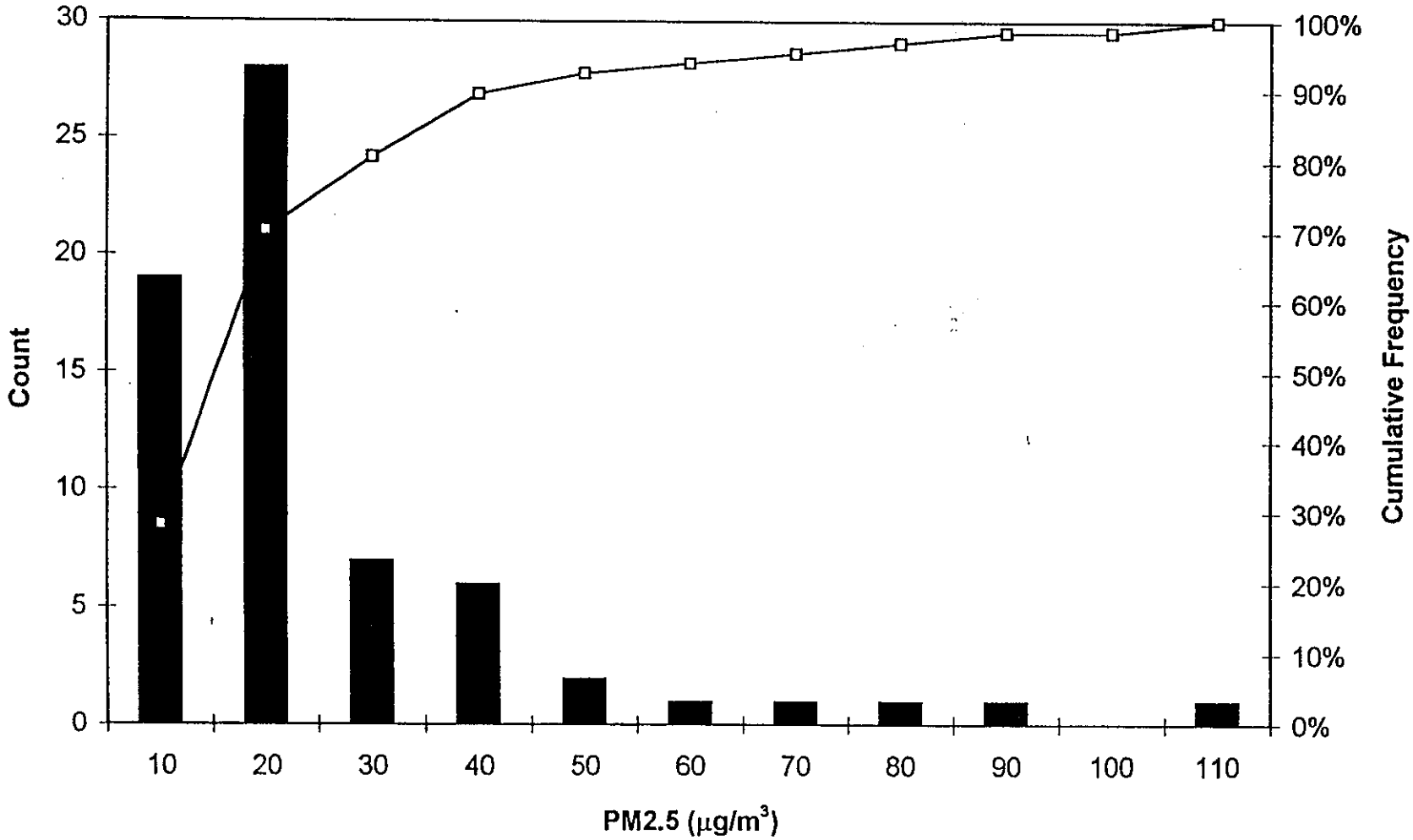


Figure 5.2-20. Histogram of indoor PM2.5 measured with Teflon Filters.

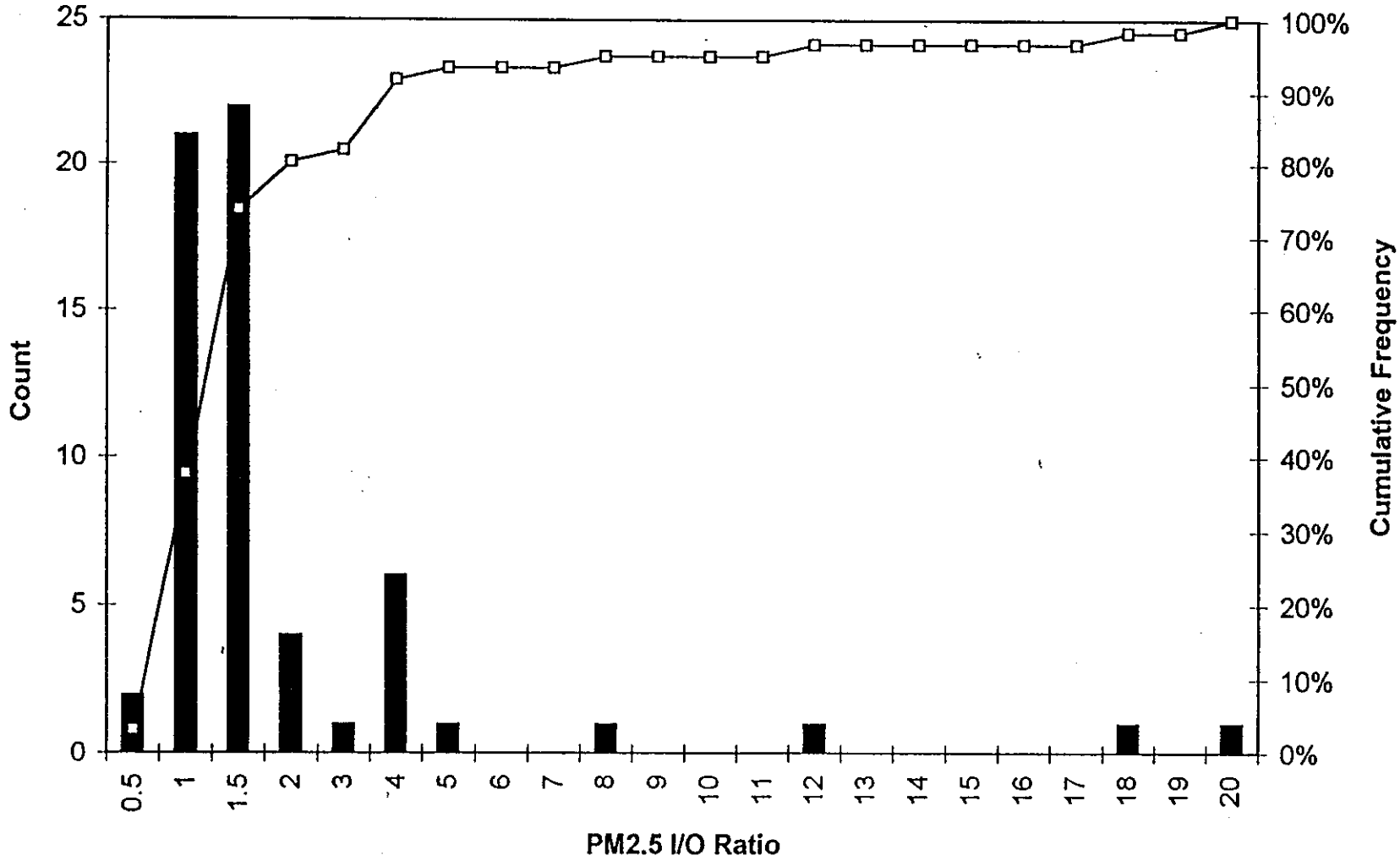


Figure 5.2-21. Histogram of PM2.5 I/O ratio measured with Teflon Filters.

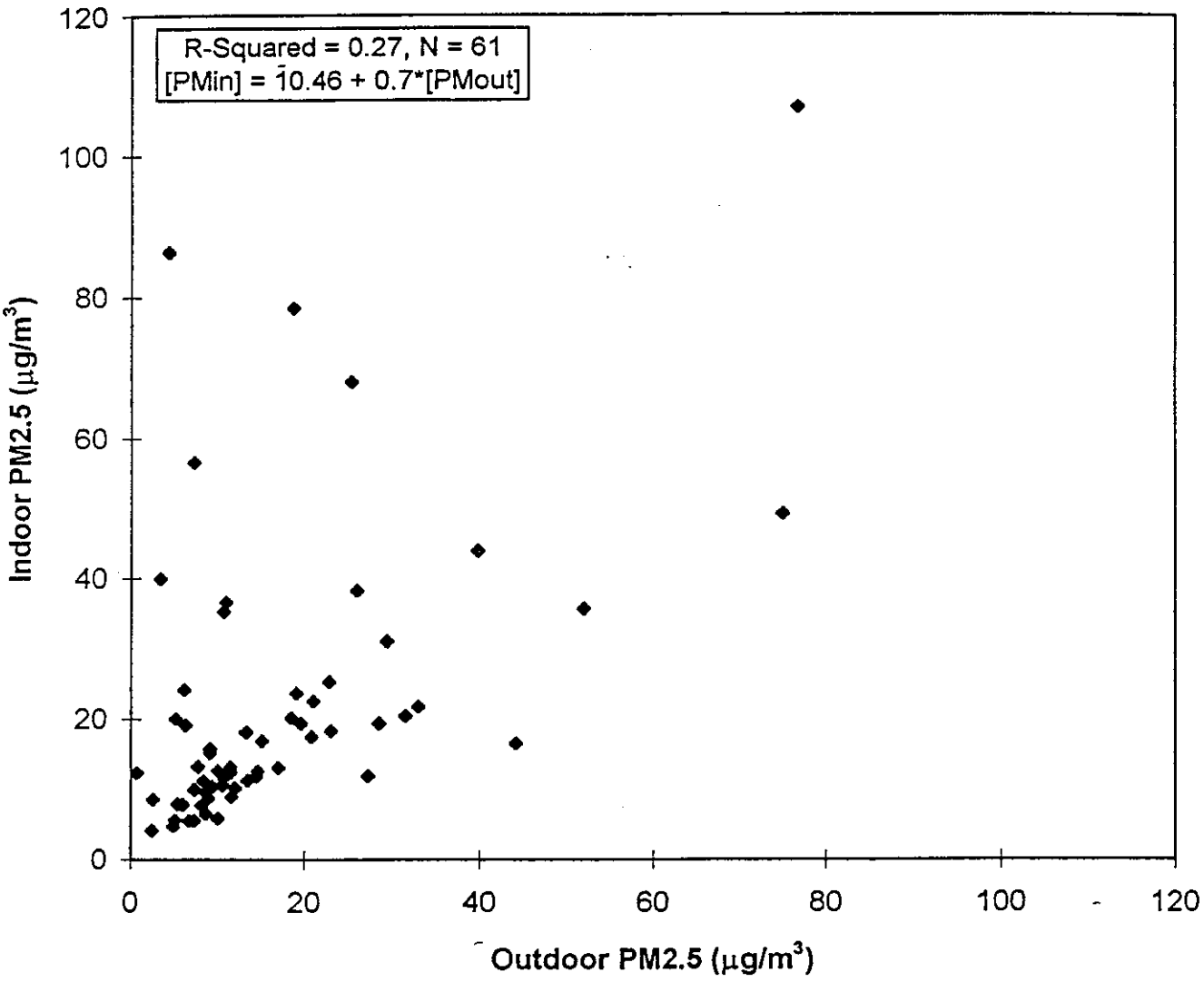


Figure 5.2-22. Scatterplot of indoor and outdoor PM2.5 measured with Teflon Filters.

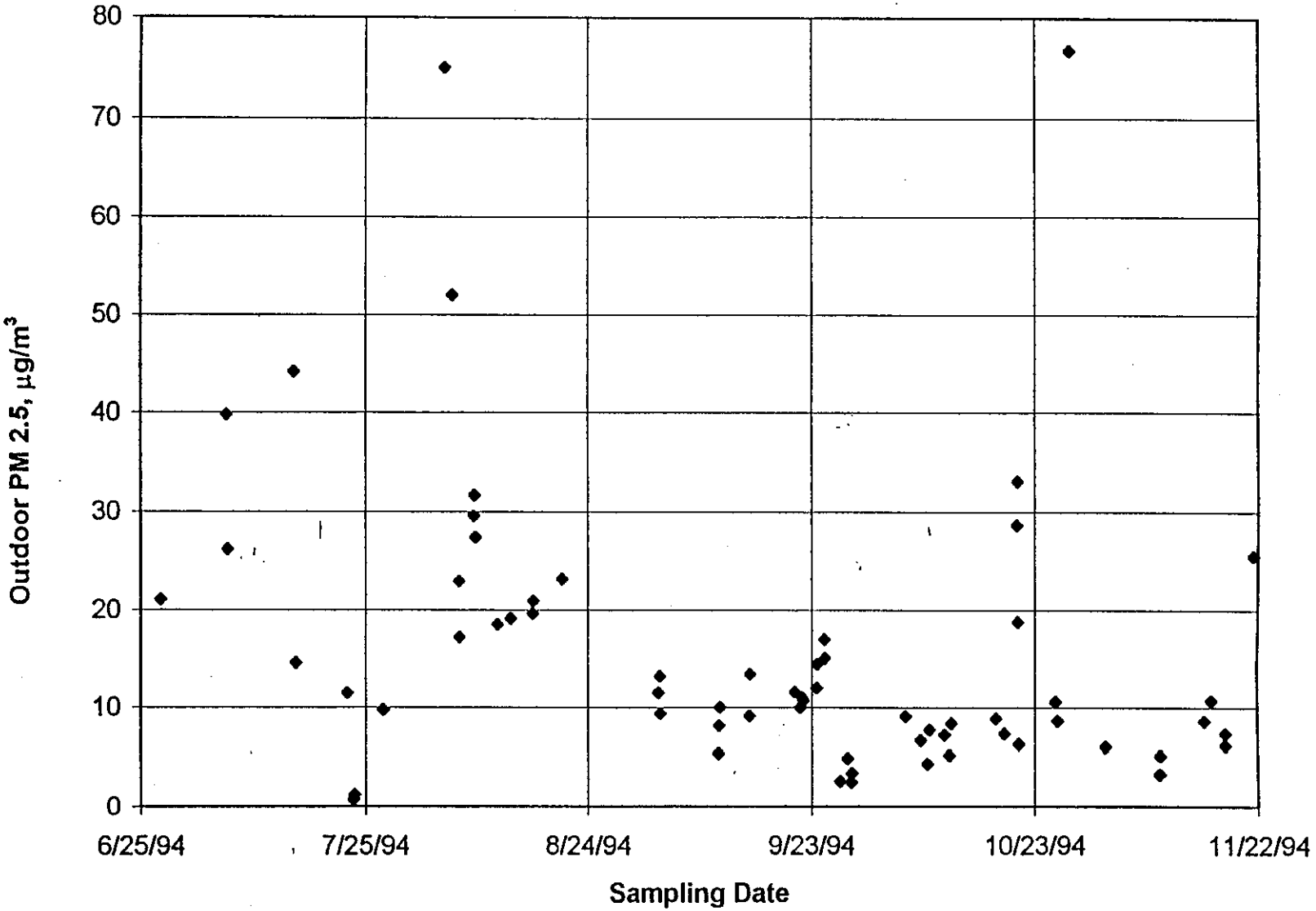


Figure 5.2-23. Temporal pattern of outdoor PM2.5.

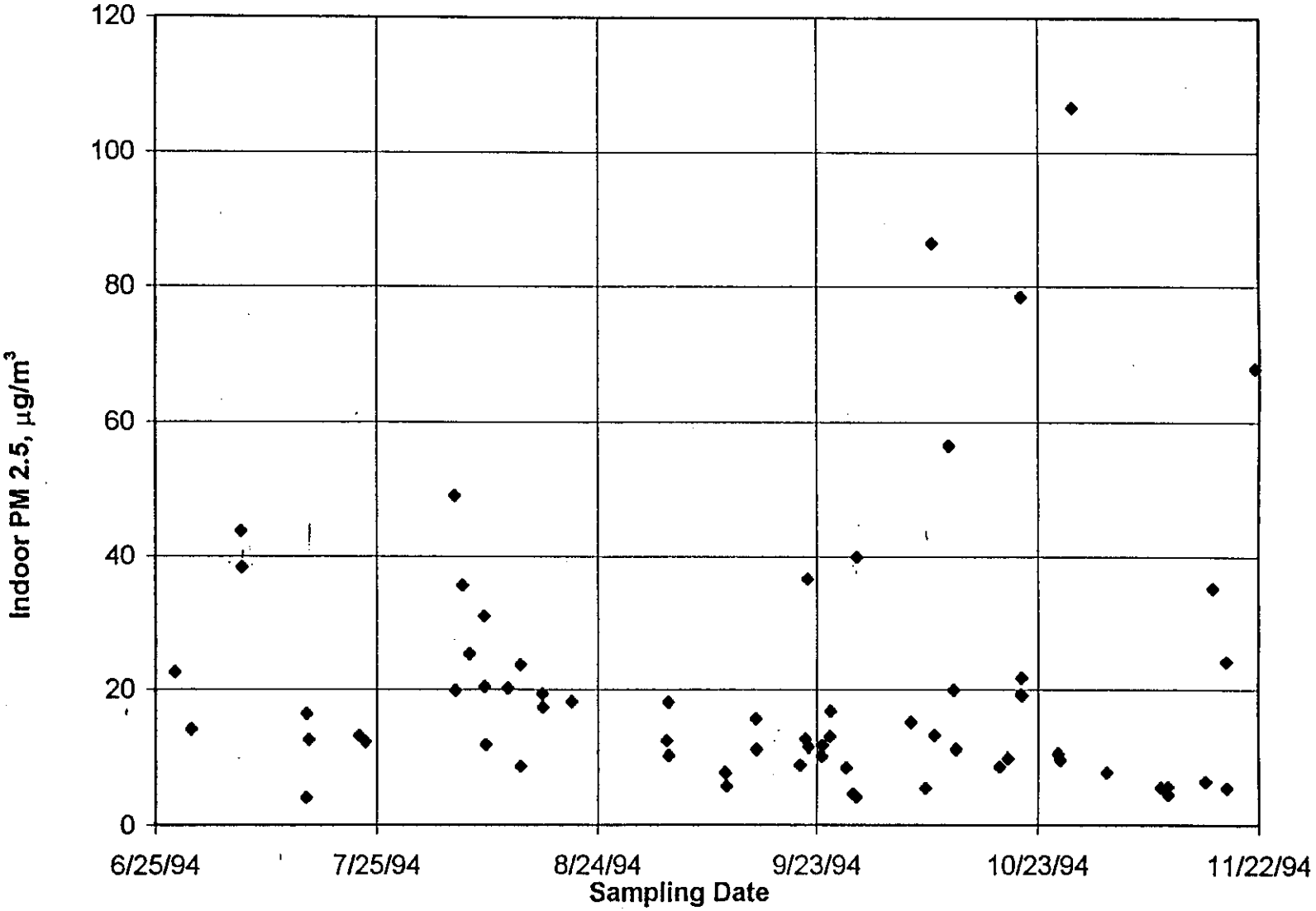


Figure 5.2-24. Temporal pattern of indoor PM2.5.

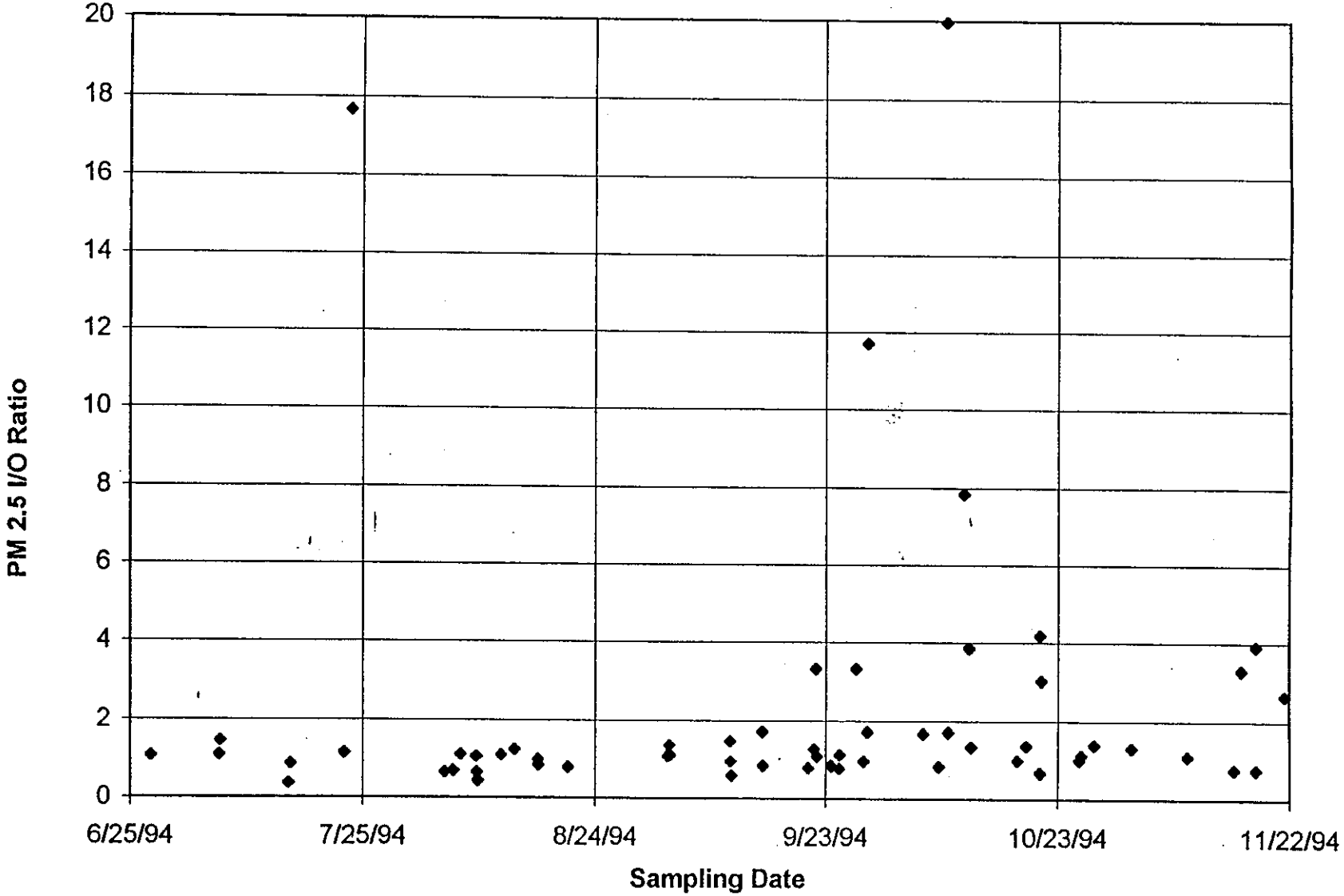


Figure 5.2-25. Temporal pattern of PM2.5 I/O ratio.

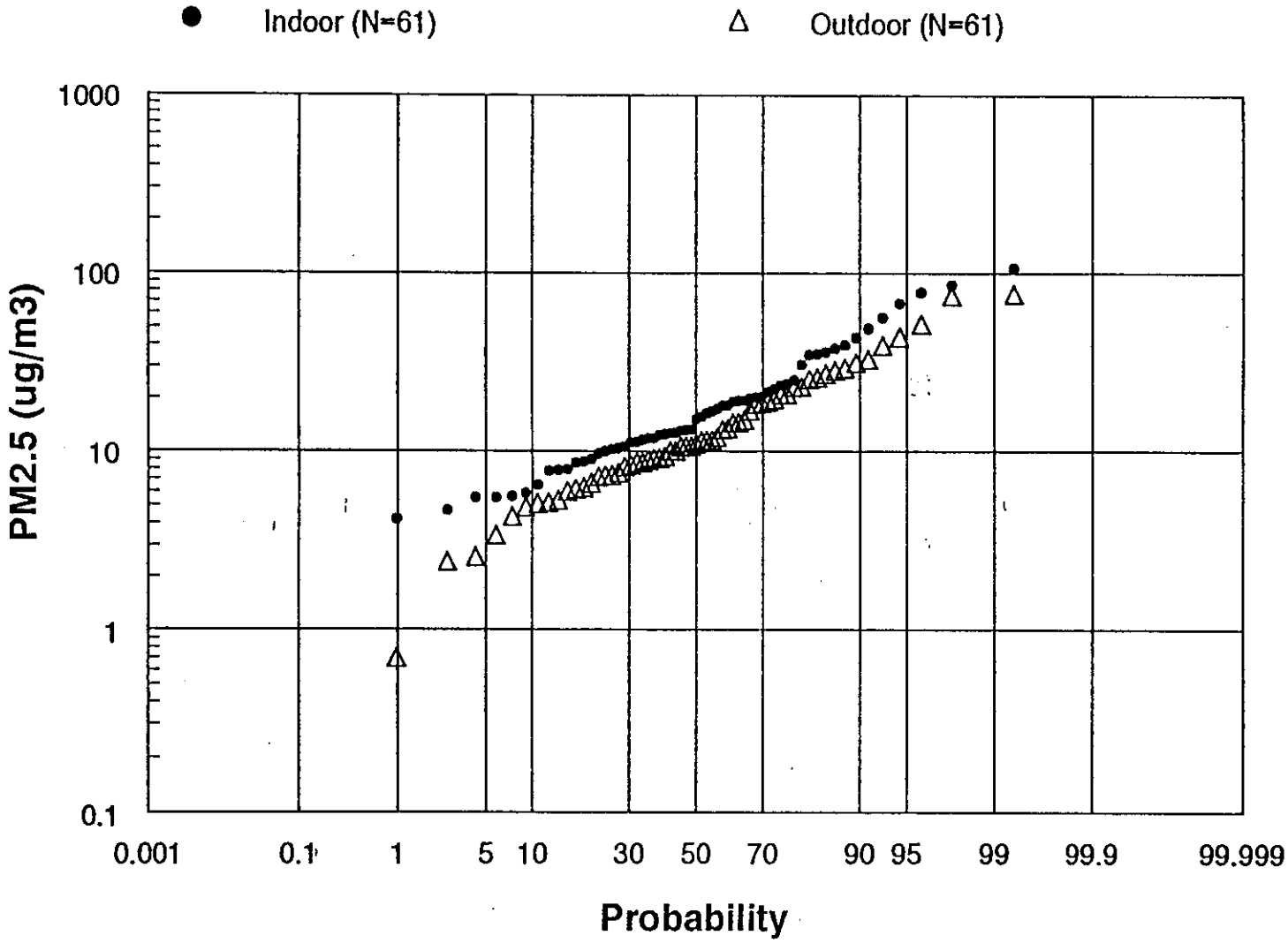


Figure 5.2-26. Log-probability plots of indoor and outdoor PM2.5.

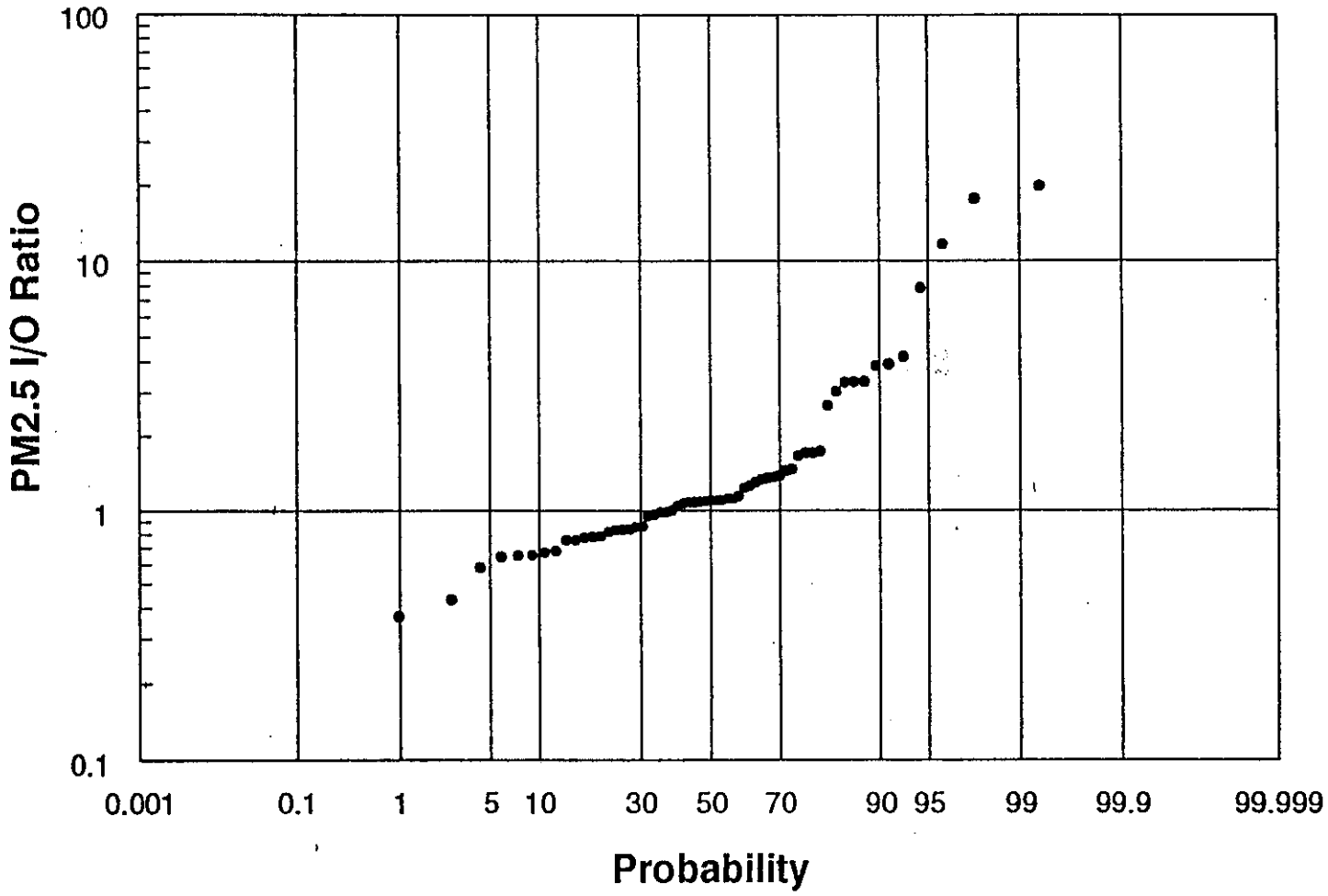


Figure 5.2-27. Log-probability plot of PM2.5 I/O ratio.

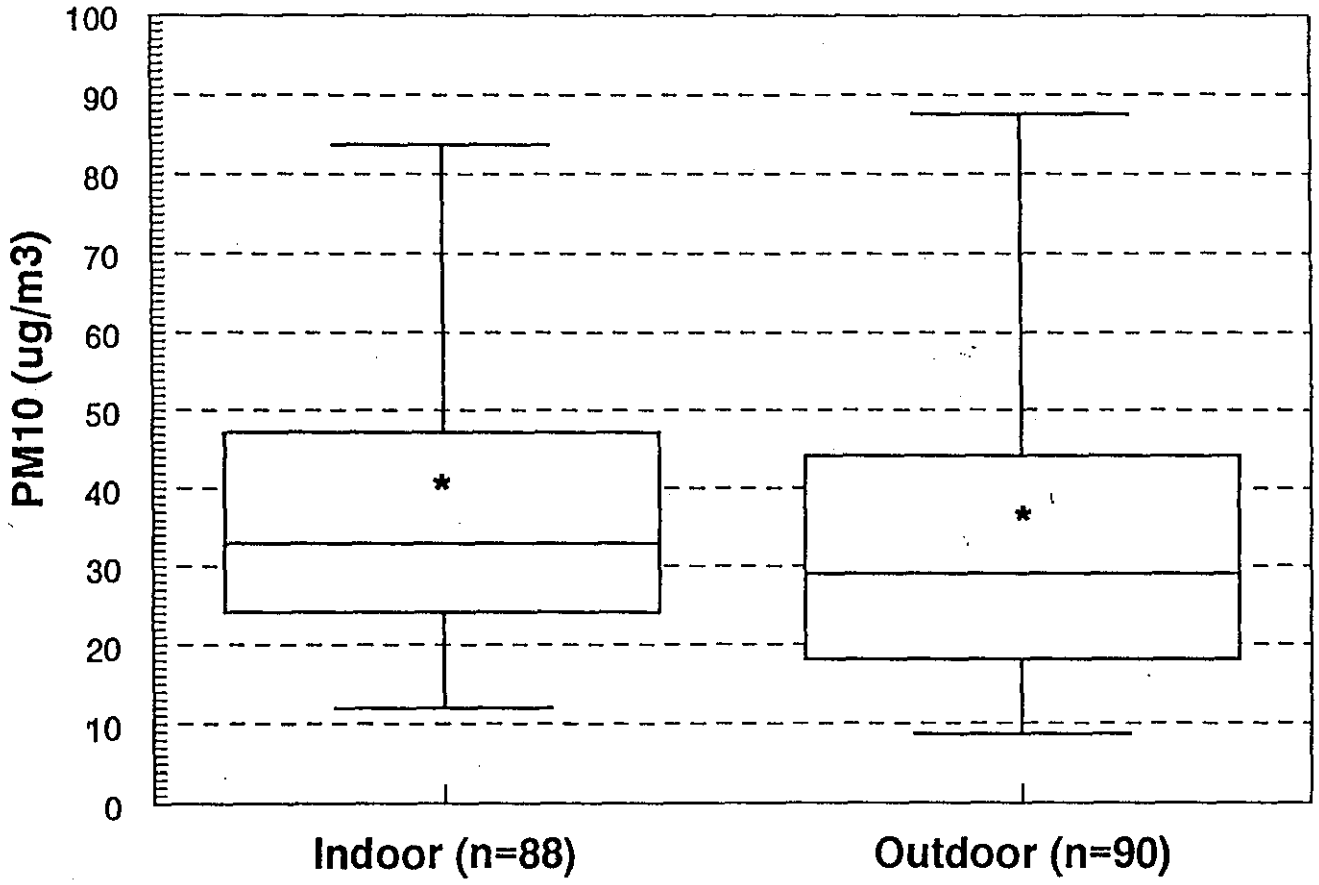


Figure 5.2-28. Box-whisker plots of PM10 concentrations measured with Teflon Filters.

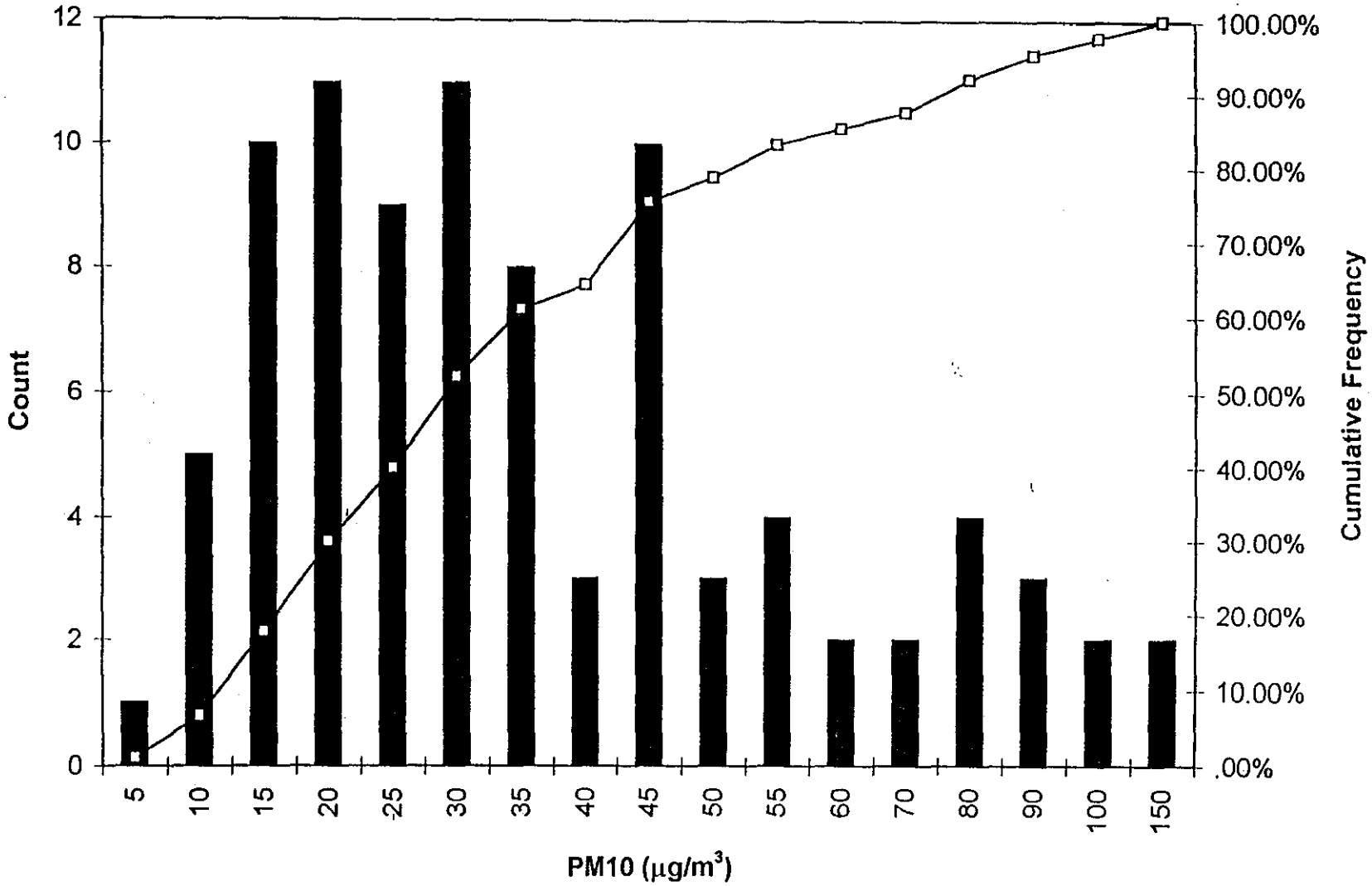


Figure 5.2-29. Histogram of outdoor PM10 measured with Teflon Filters.

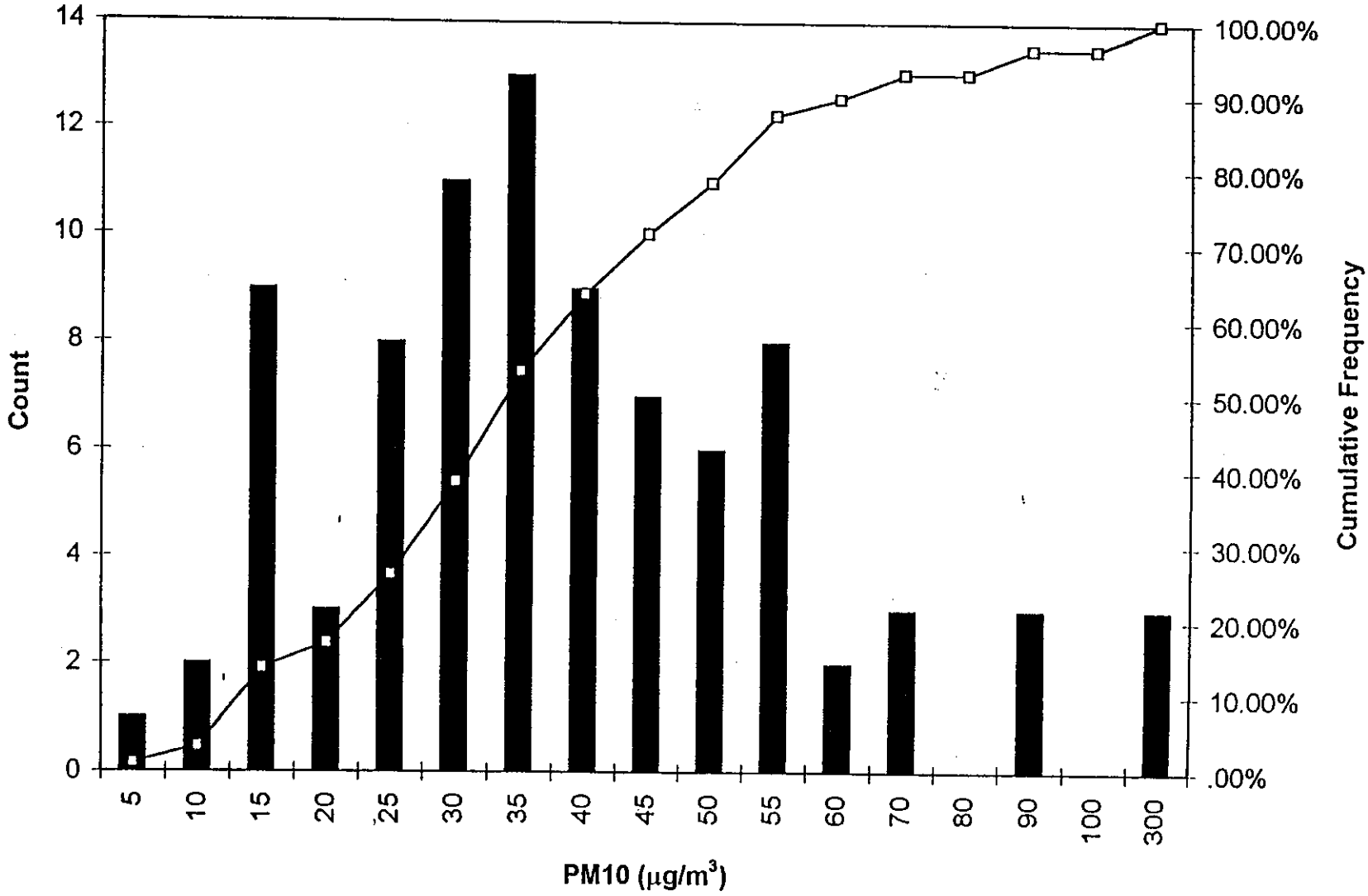


Figure 5.2-30. Histogram of indoor PM10 measured with Teflon Filters.

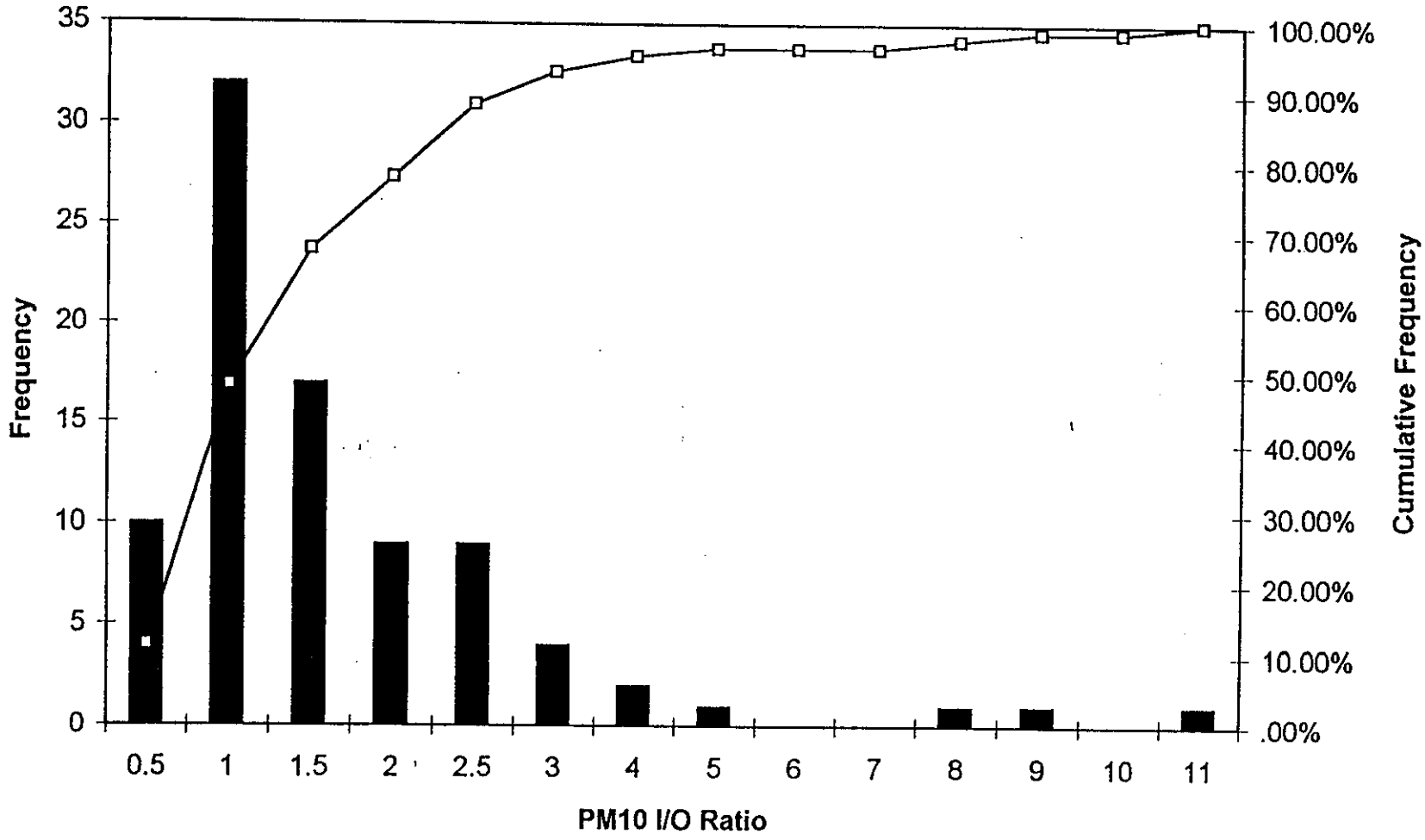


Figure 5.2-31. Histogram of PM10 I/O ratio measured with Teflon Filters.

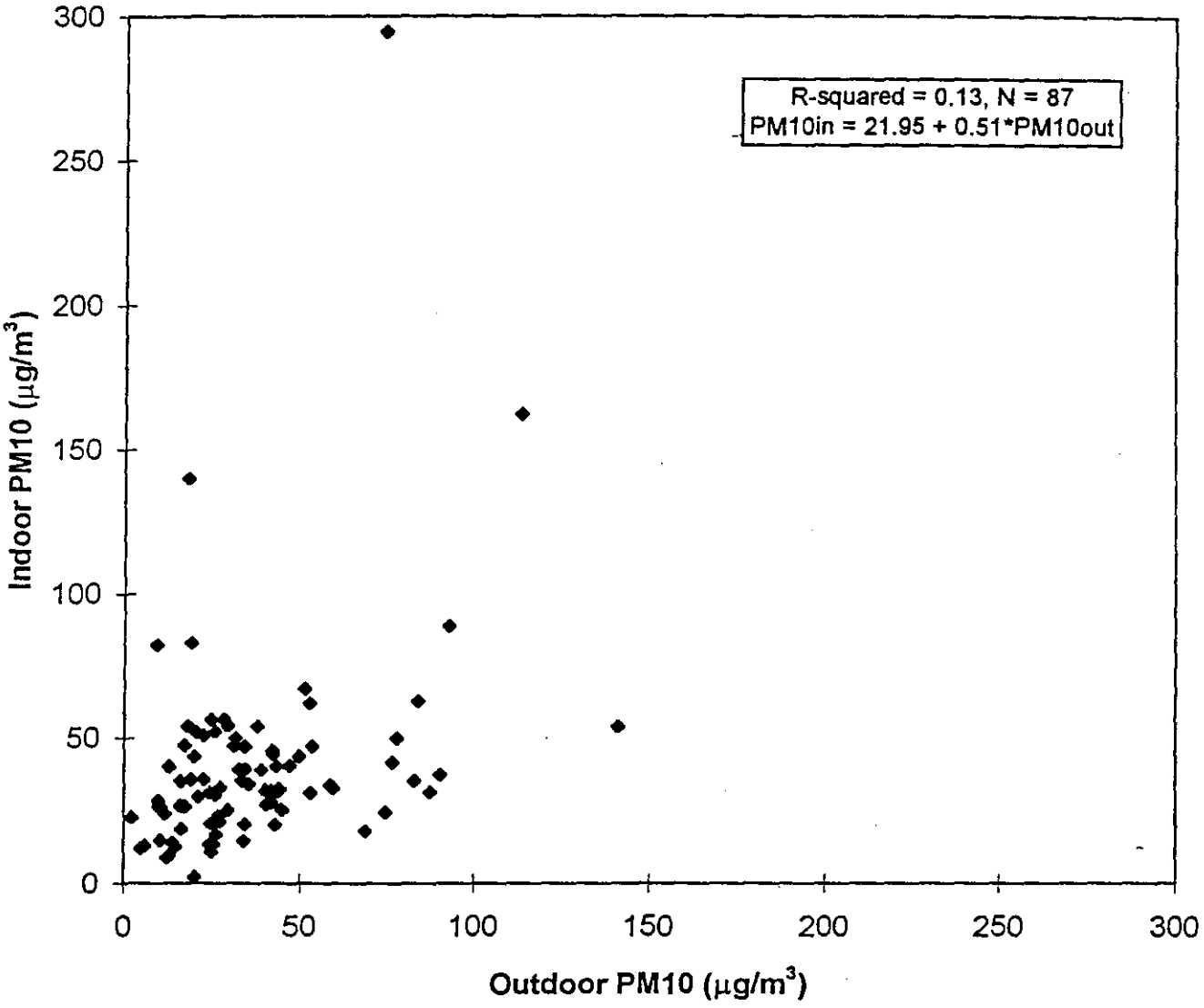


Figure 5.2-32. Scatterplot of indoor vs. outdoor PM10 measured with Teflon Filters.

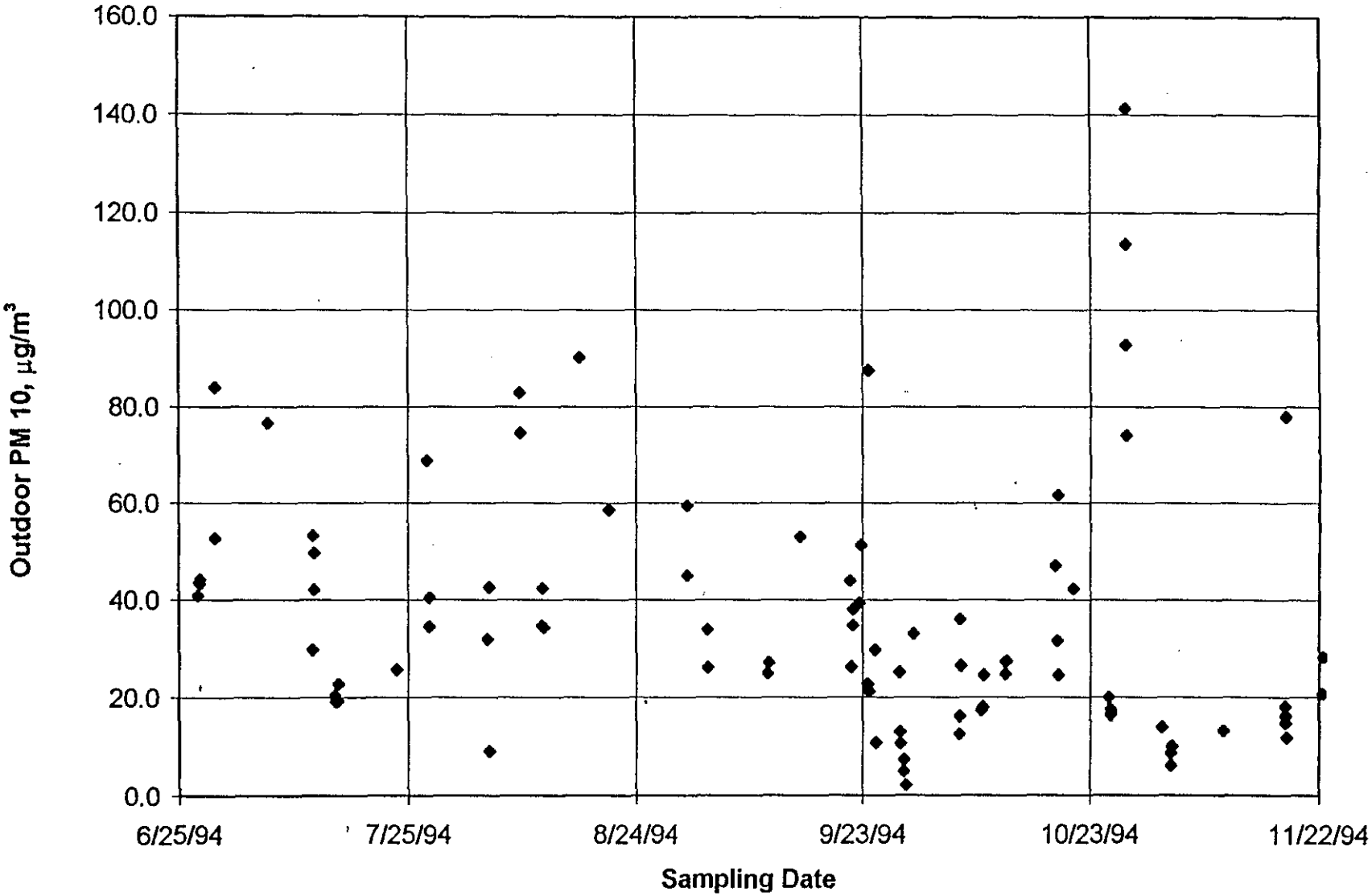


Figure 5.2-33. Temporal pattern of outdoor PM10 measured at home sites.

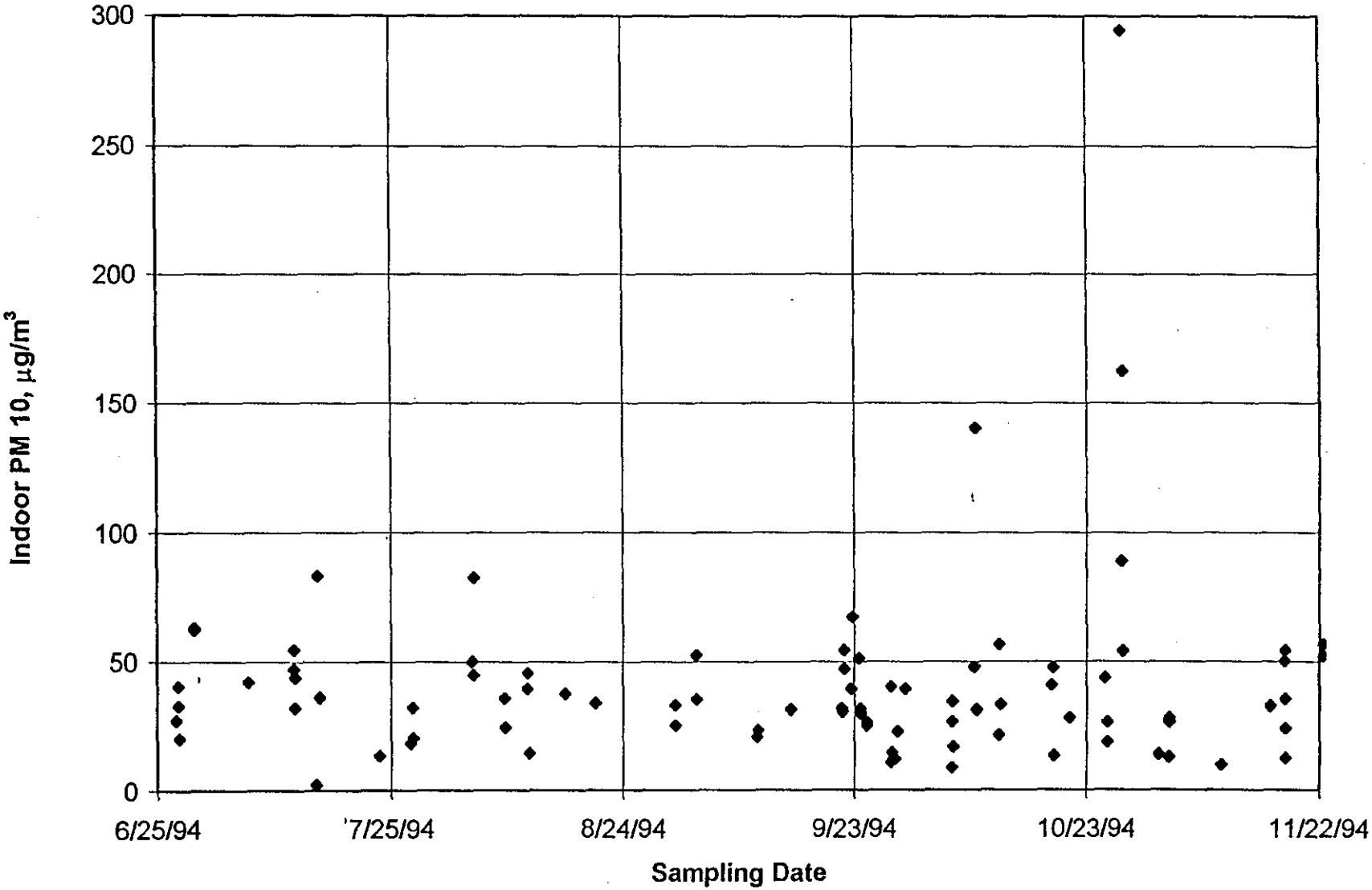


Figure 5.2-34. Temporal pattern of indoor PM10 measured at home sites.

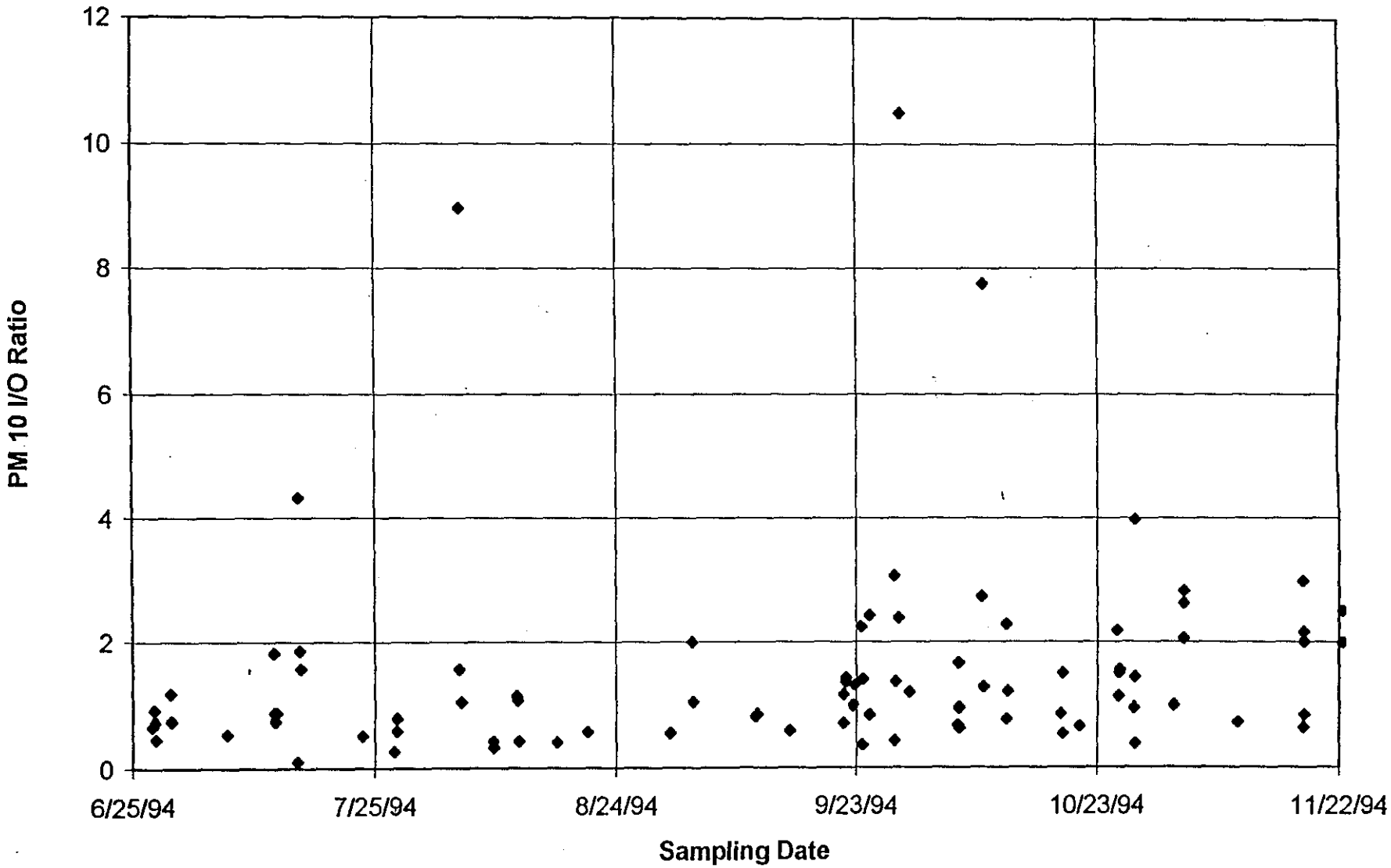


Figure 5.2-35. Temporal pattern of PM10 I/O ratio measured at home sites.

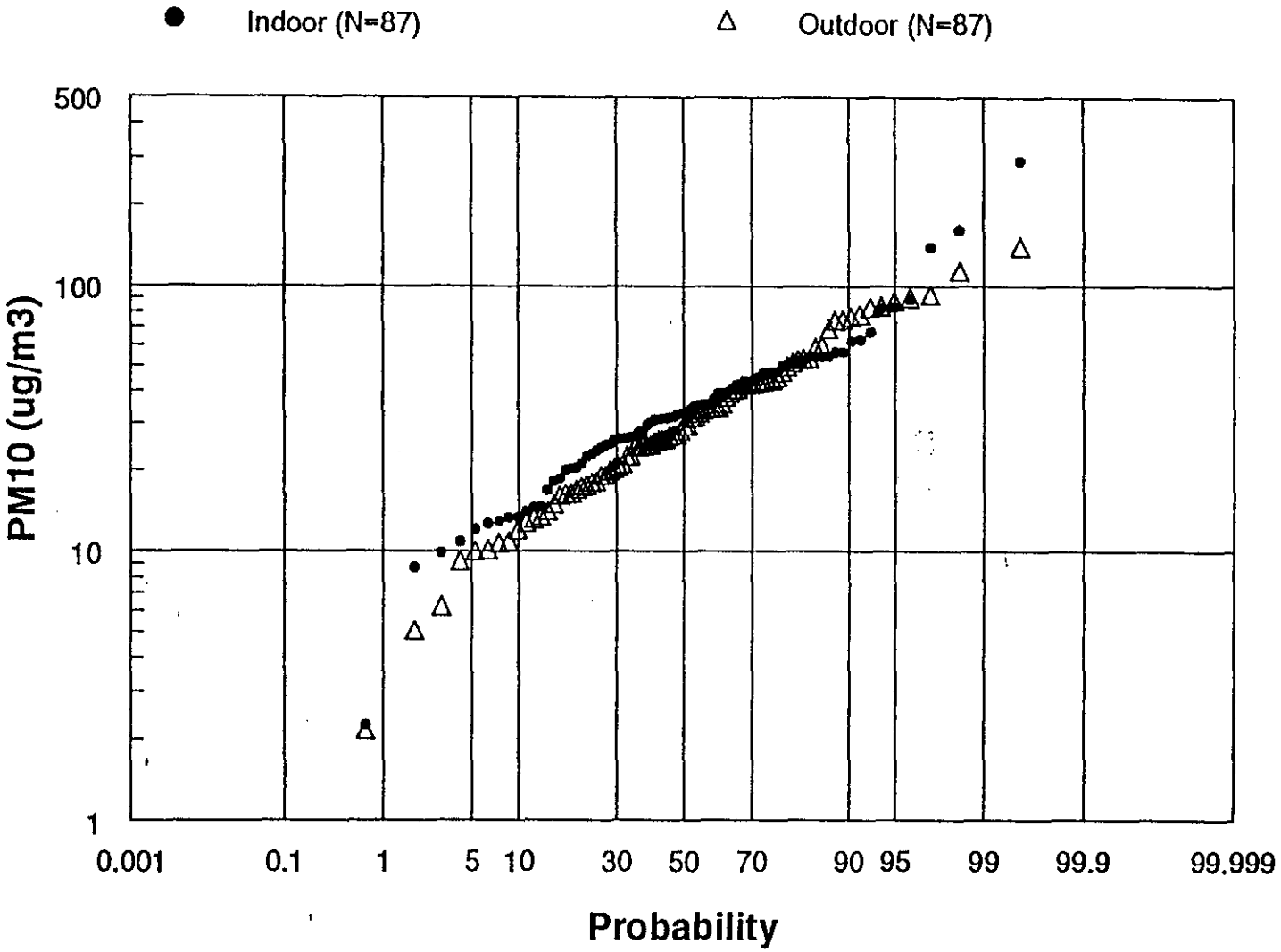


Figure 5.2-36. Log-probability plots of indoor and outdoor PM10.

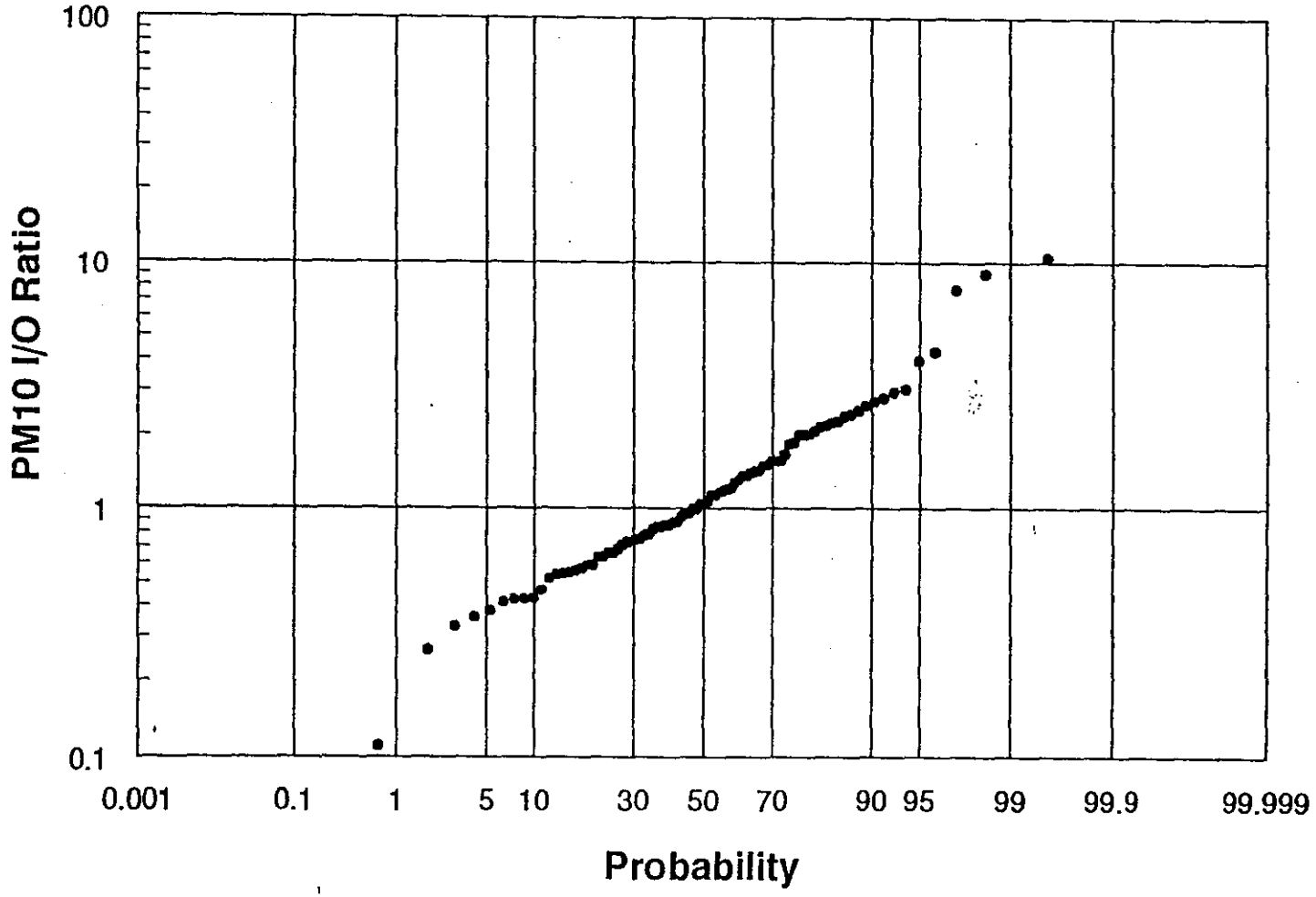


Figure 5.2-37. Log-probability plot of PM10 I/O ratio.

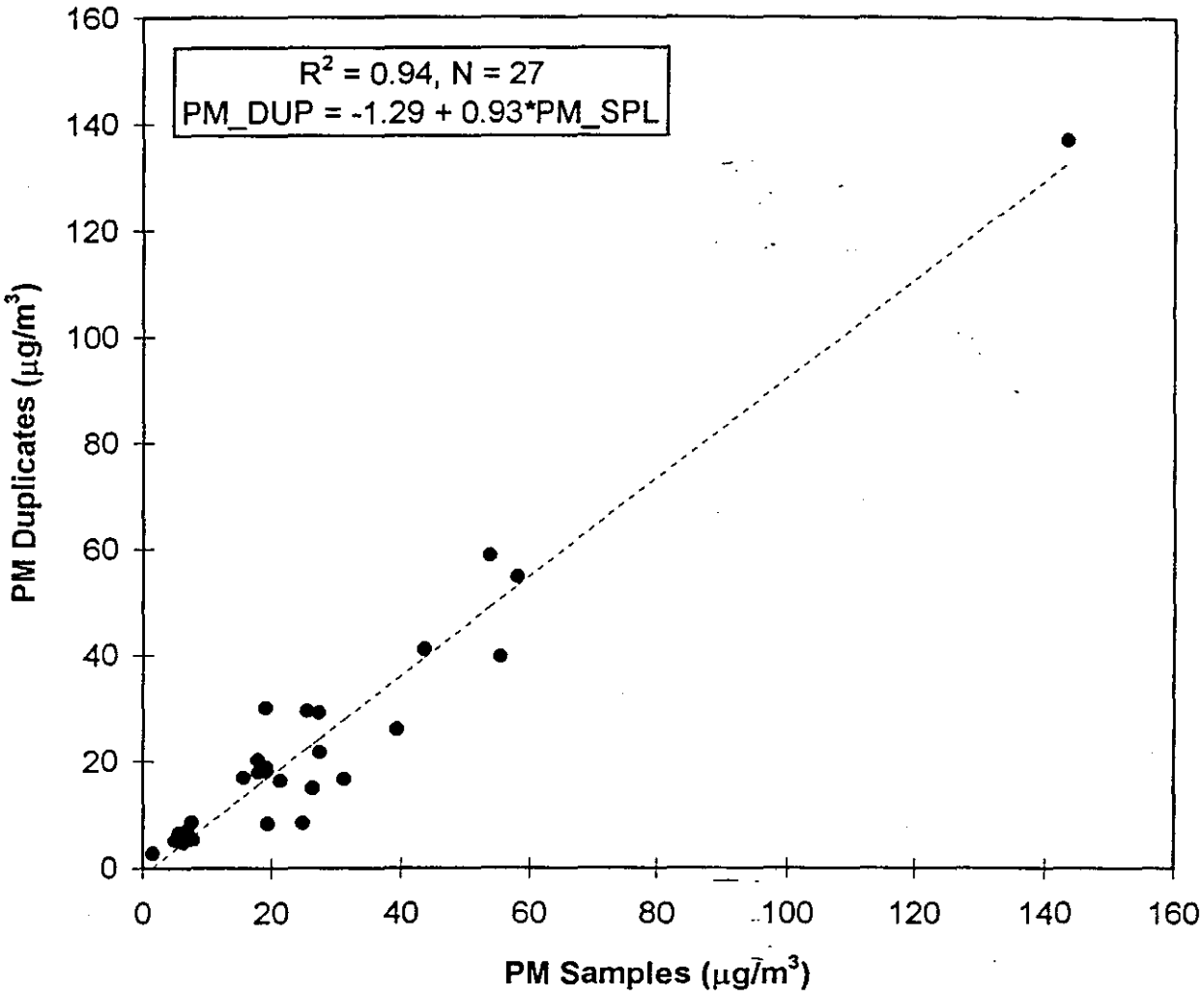


Figure 5.2-38. Scatterplot of PM co-located samples.

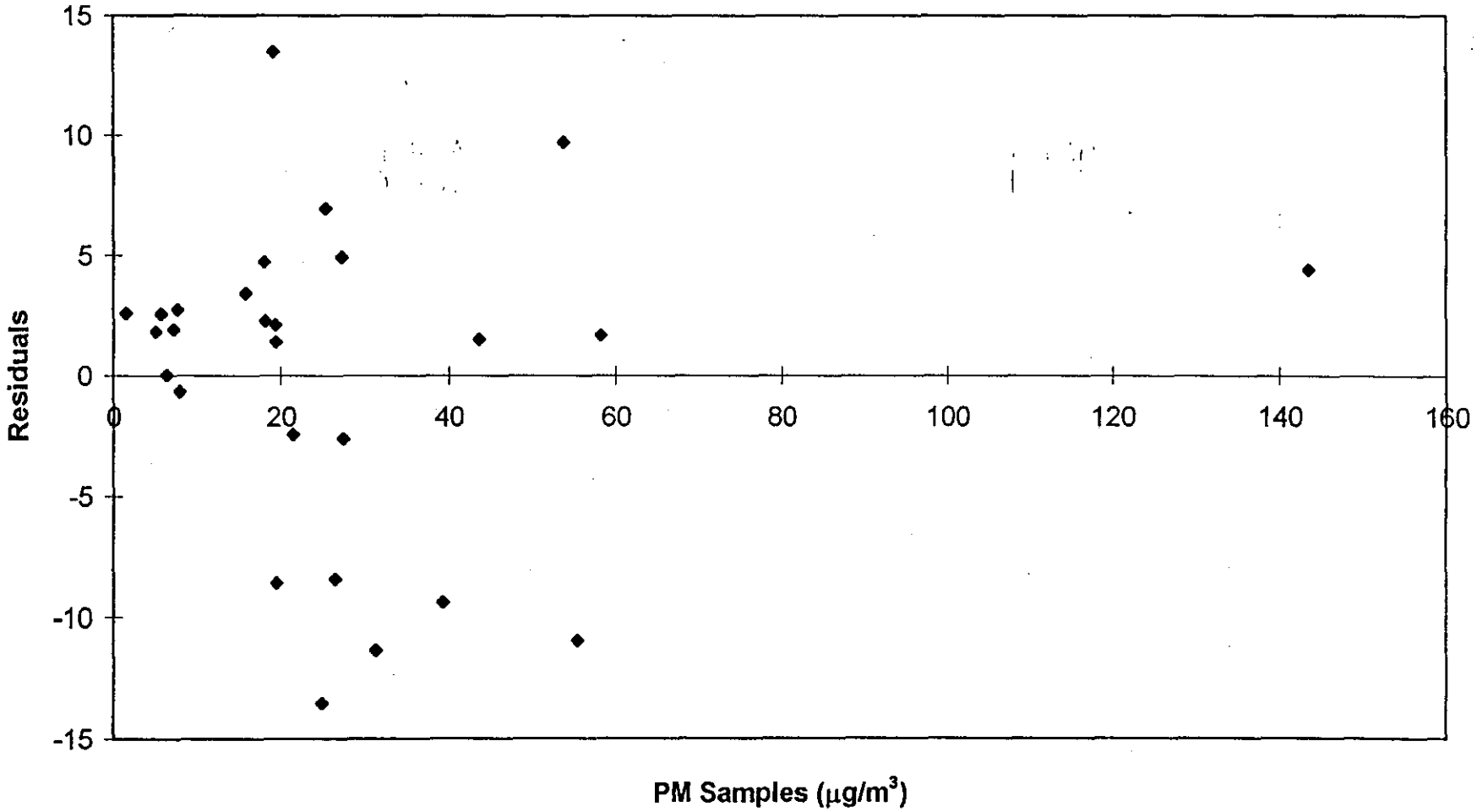


Figure 5.2-39. PM co-located samples residual plot.

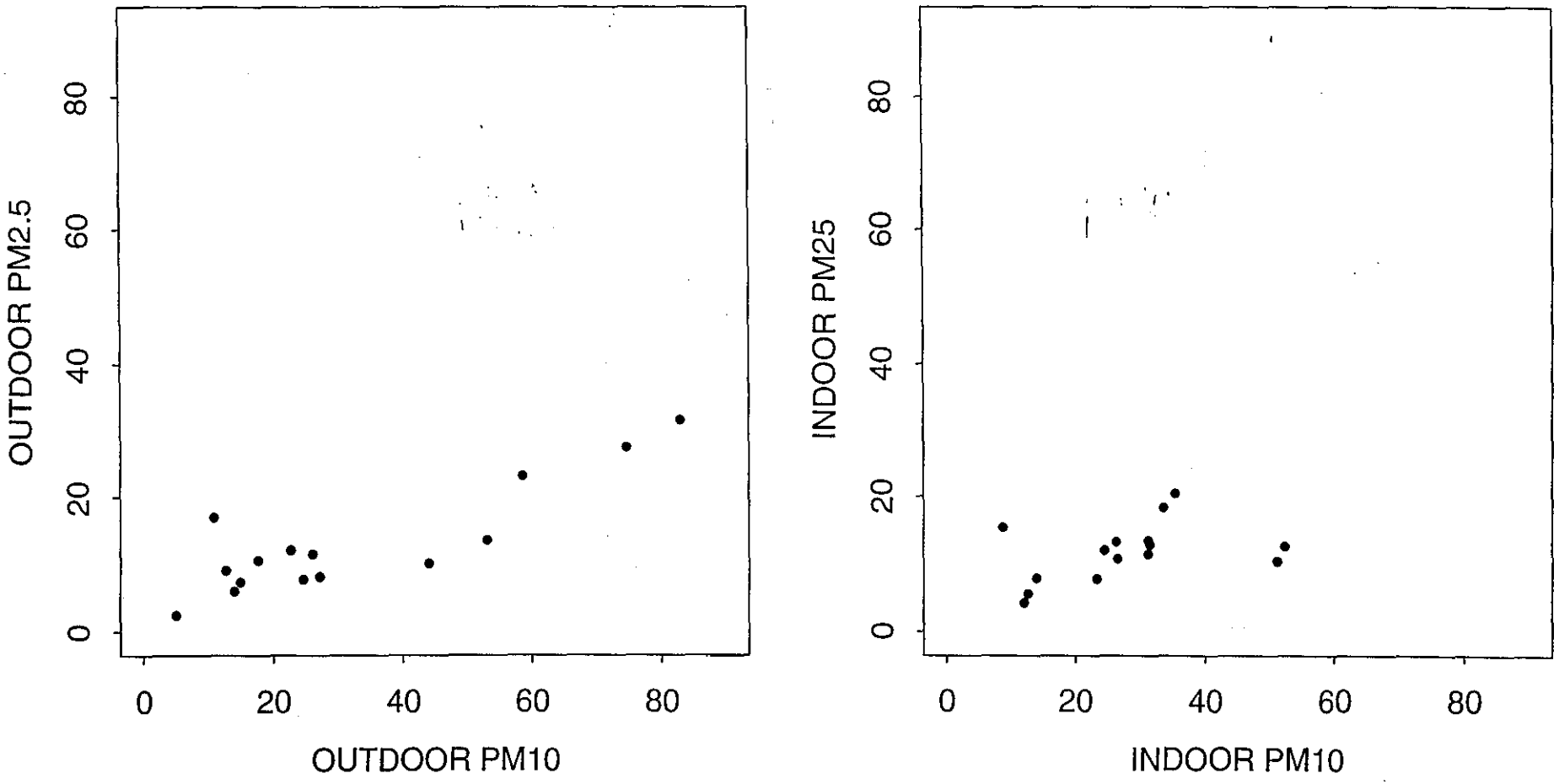


Figure 5.2-40. Co-located PM_{10} and $PM_{2.5}$ data from non-smoking homes (n=15).

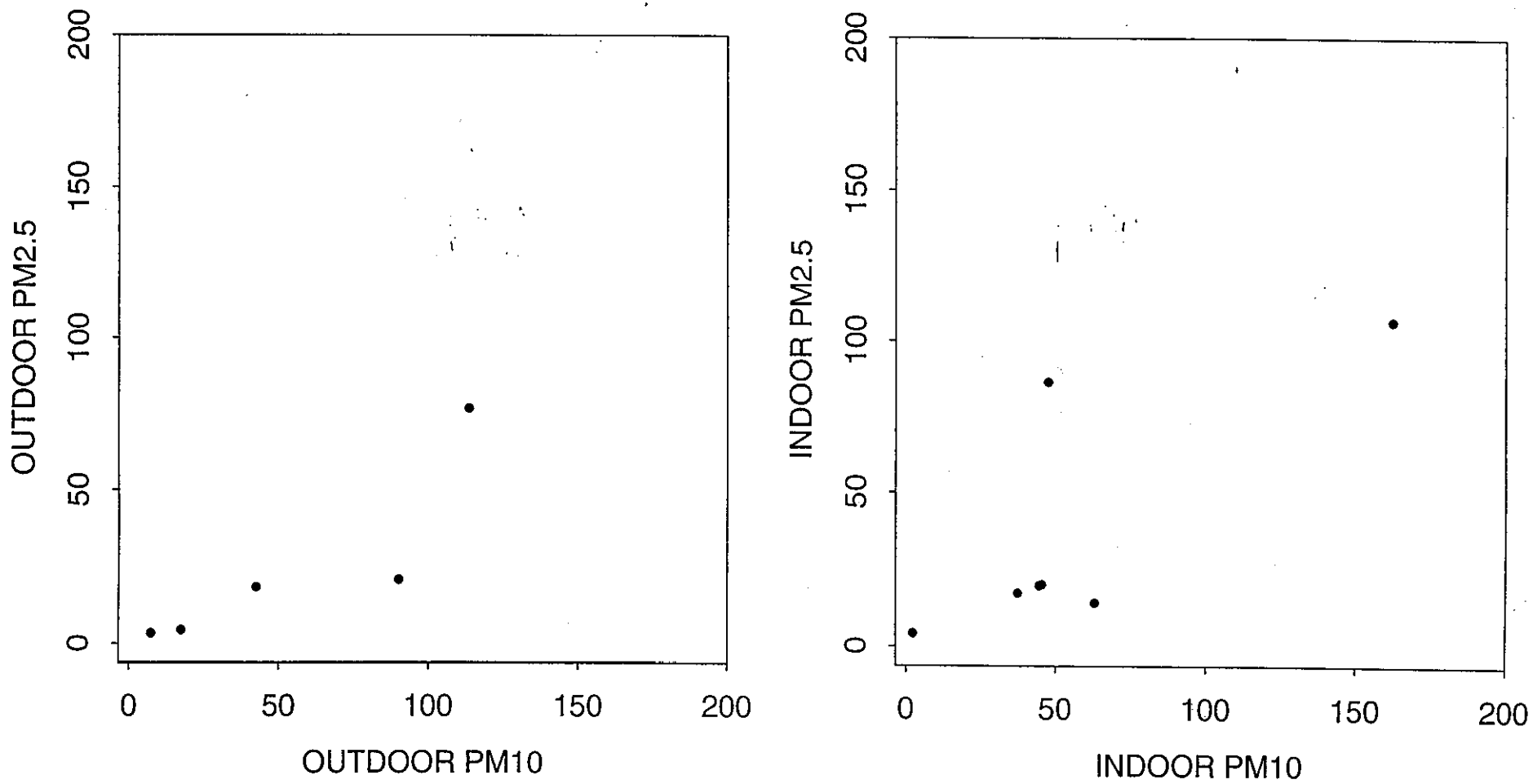


Figure 5.2-41. Co-located PM_{10} and $PM_{2.5}$ data from smoking homes (n=7).

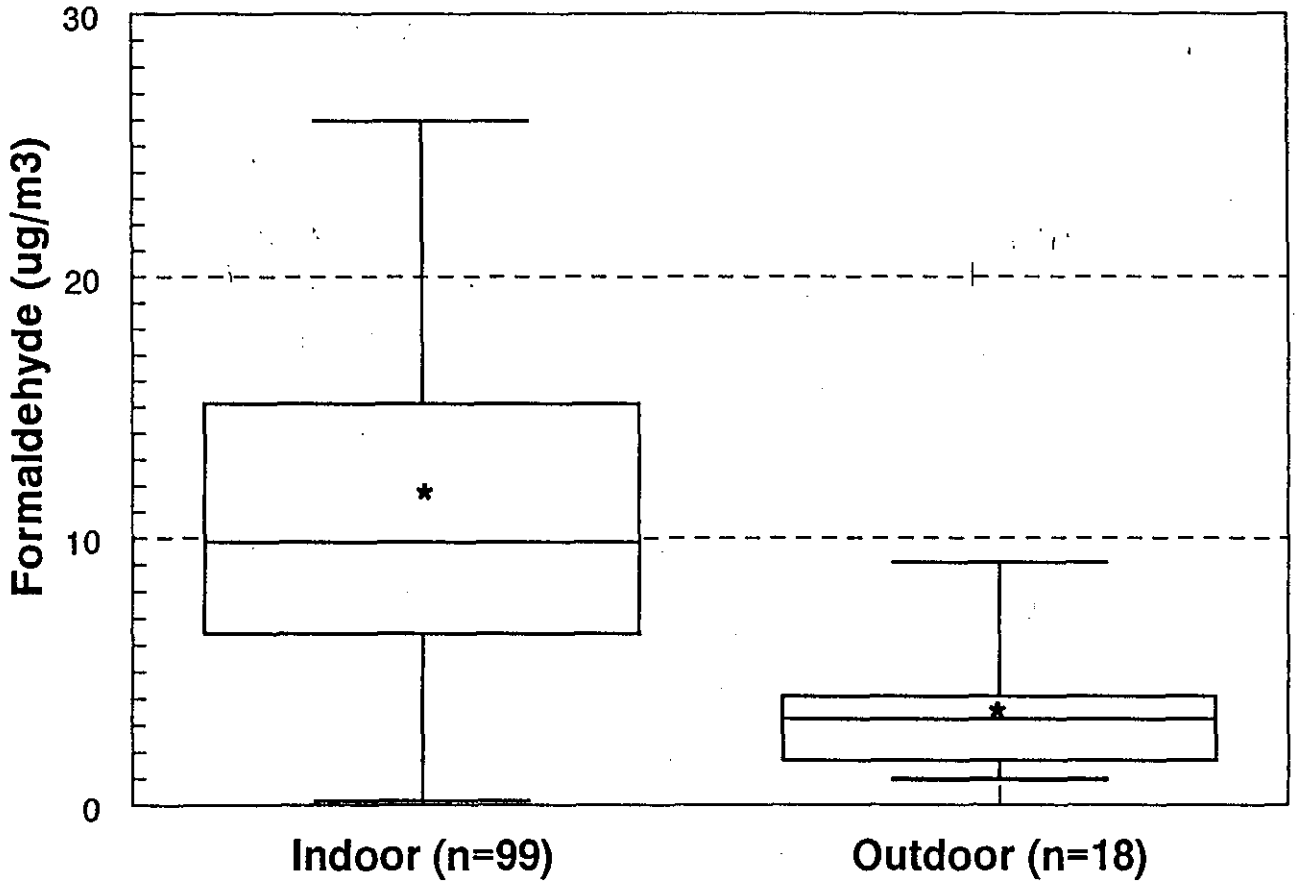


Figure 5.2-42. Box-whisker plots of indoor and outdoor formaldehyde concentrations.

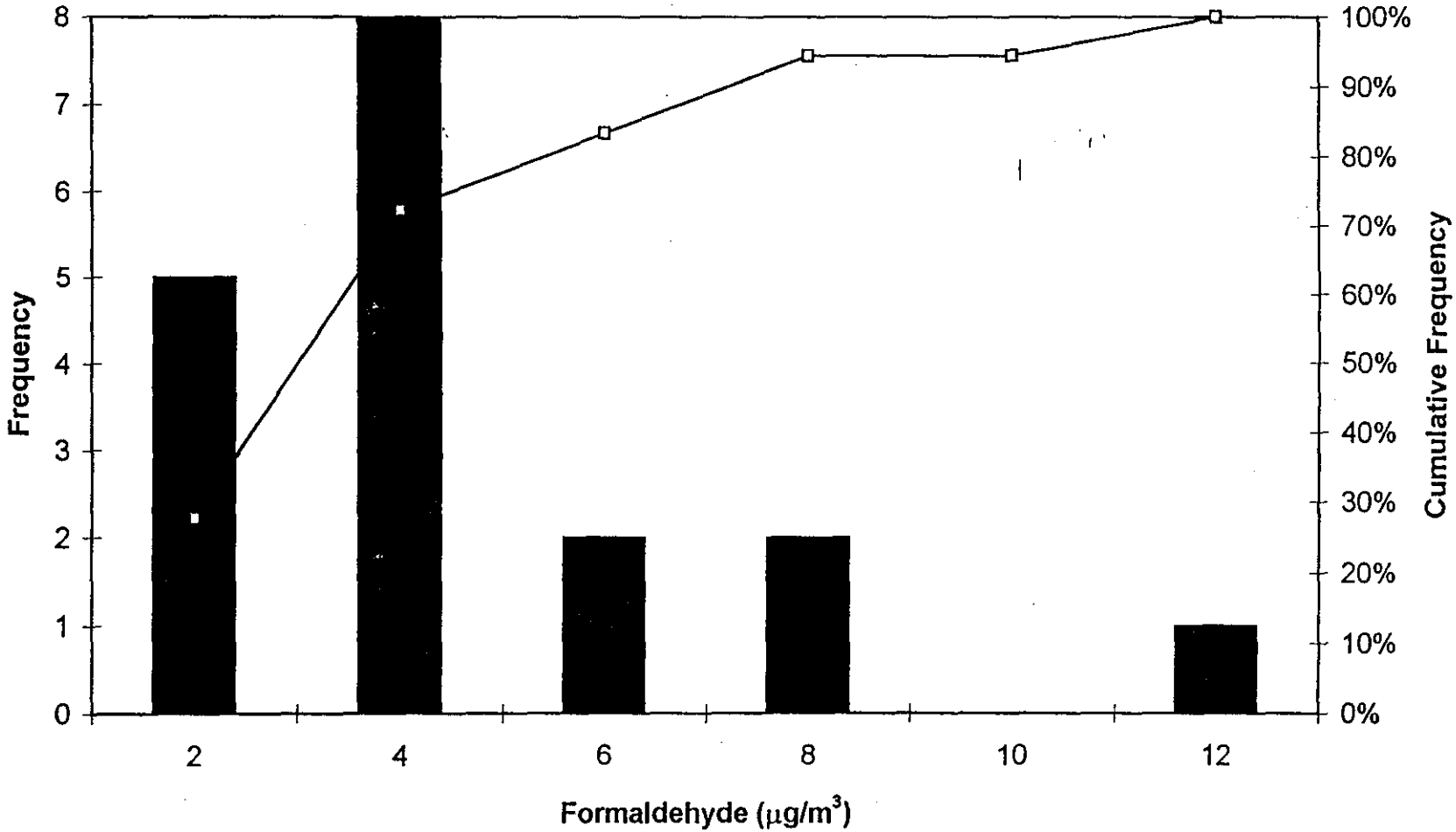


Figure 5.2-43. Histogram of outdoor formaldehyde.

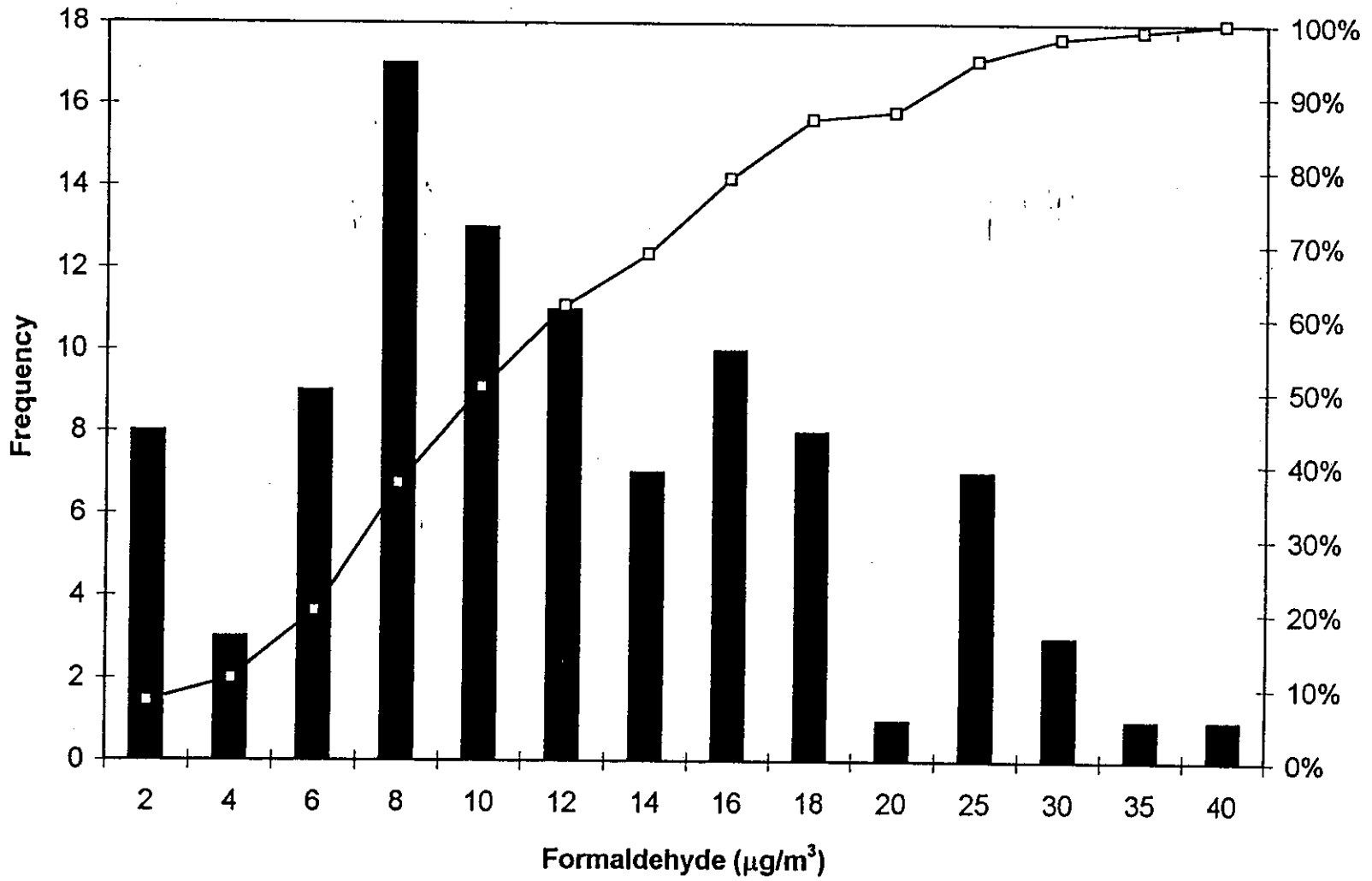


Figure 5.2-44. Histogram of indoor formaldehyde.

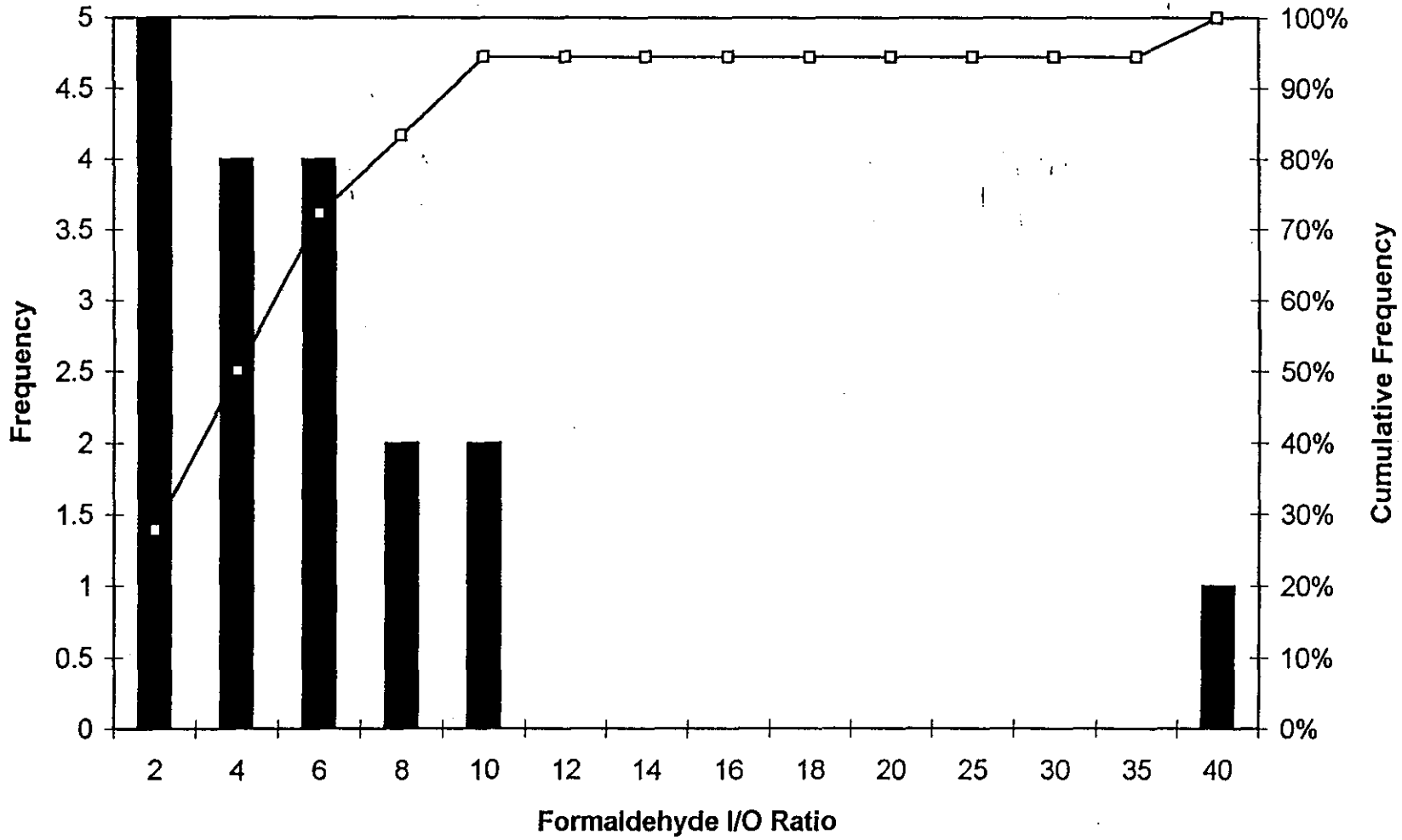


Figure 5.2-45. Histogram of formaldehyde I/O ratio.

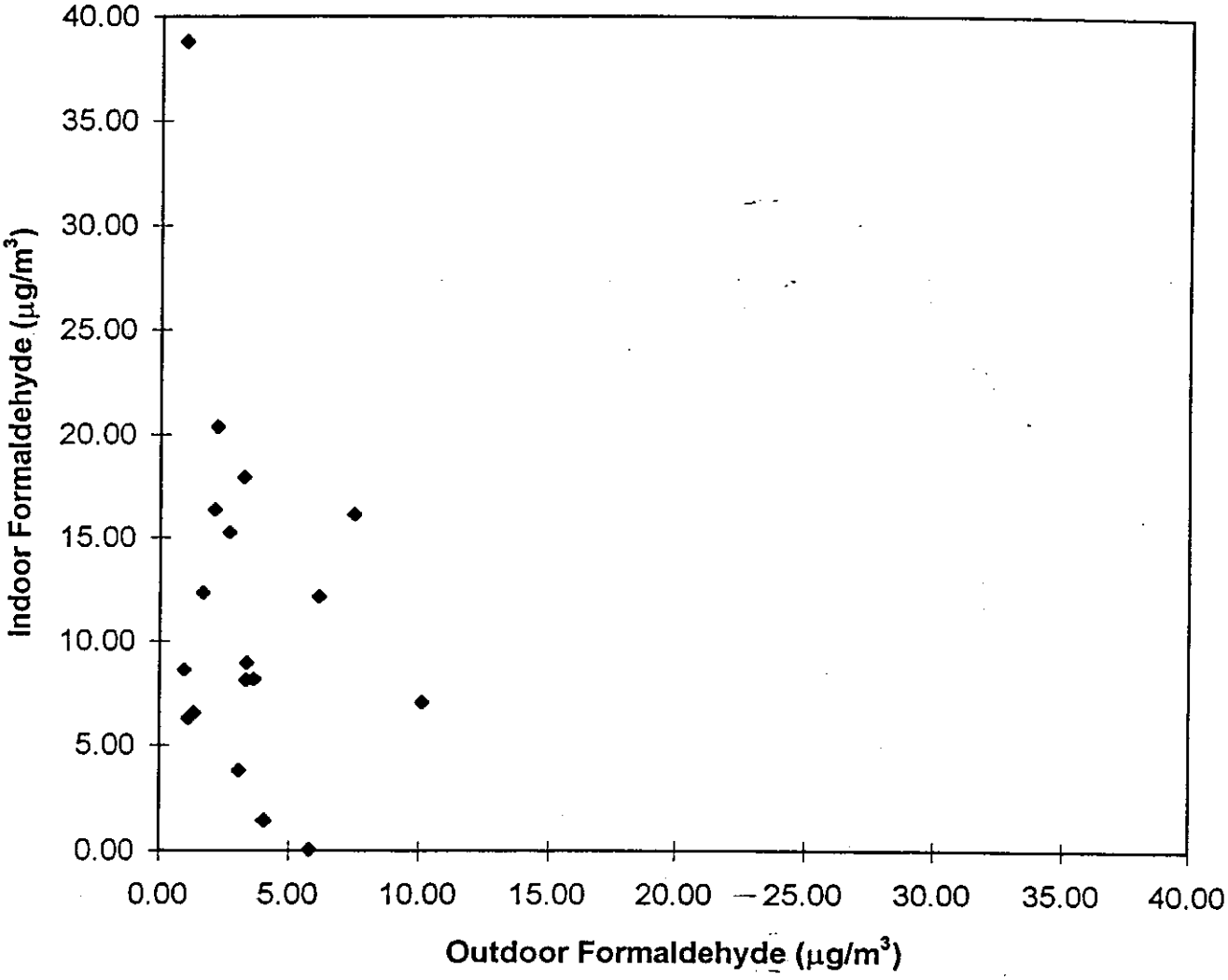


Figure 5.2-46. Scatterplot of indoor and outdoor formaldehyde.

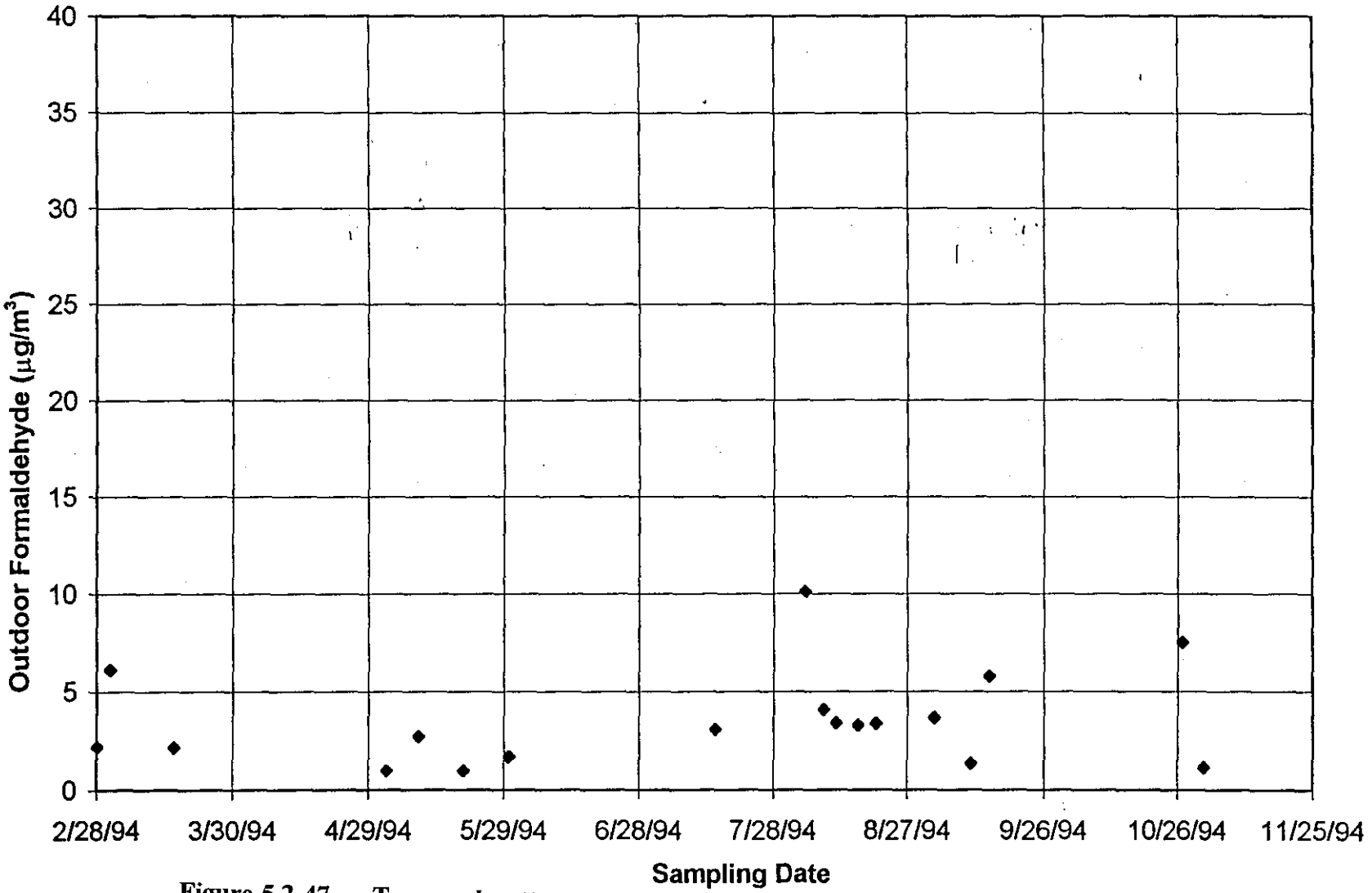


Figure 5.2-47. Temporal pattern of outdoor formaldehyde measured at home sites.

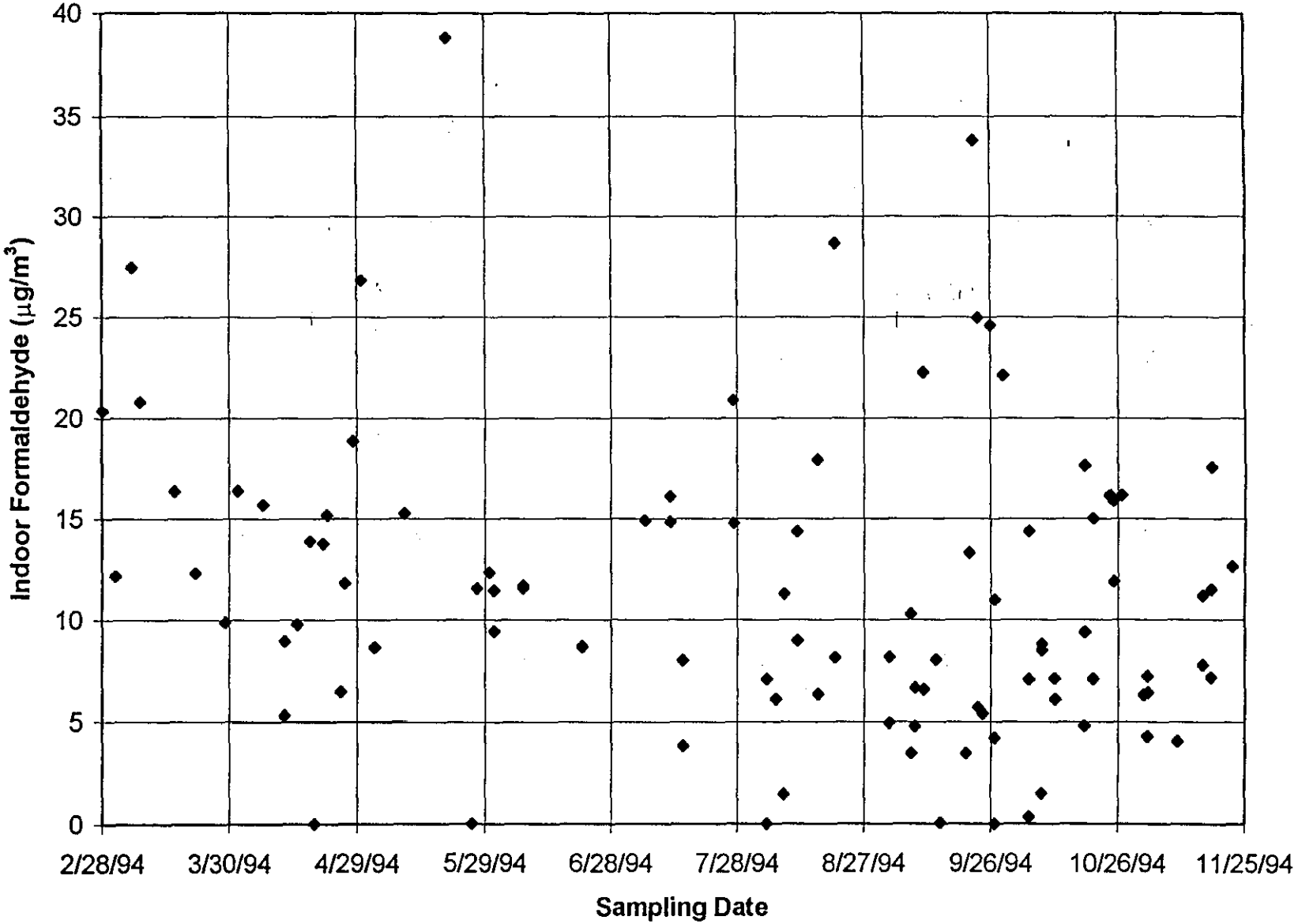


Figure 5.2-48. Temporal pattern of indoor formaldehyde measured at home sites.

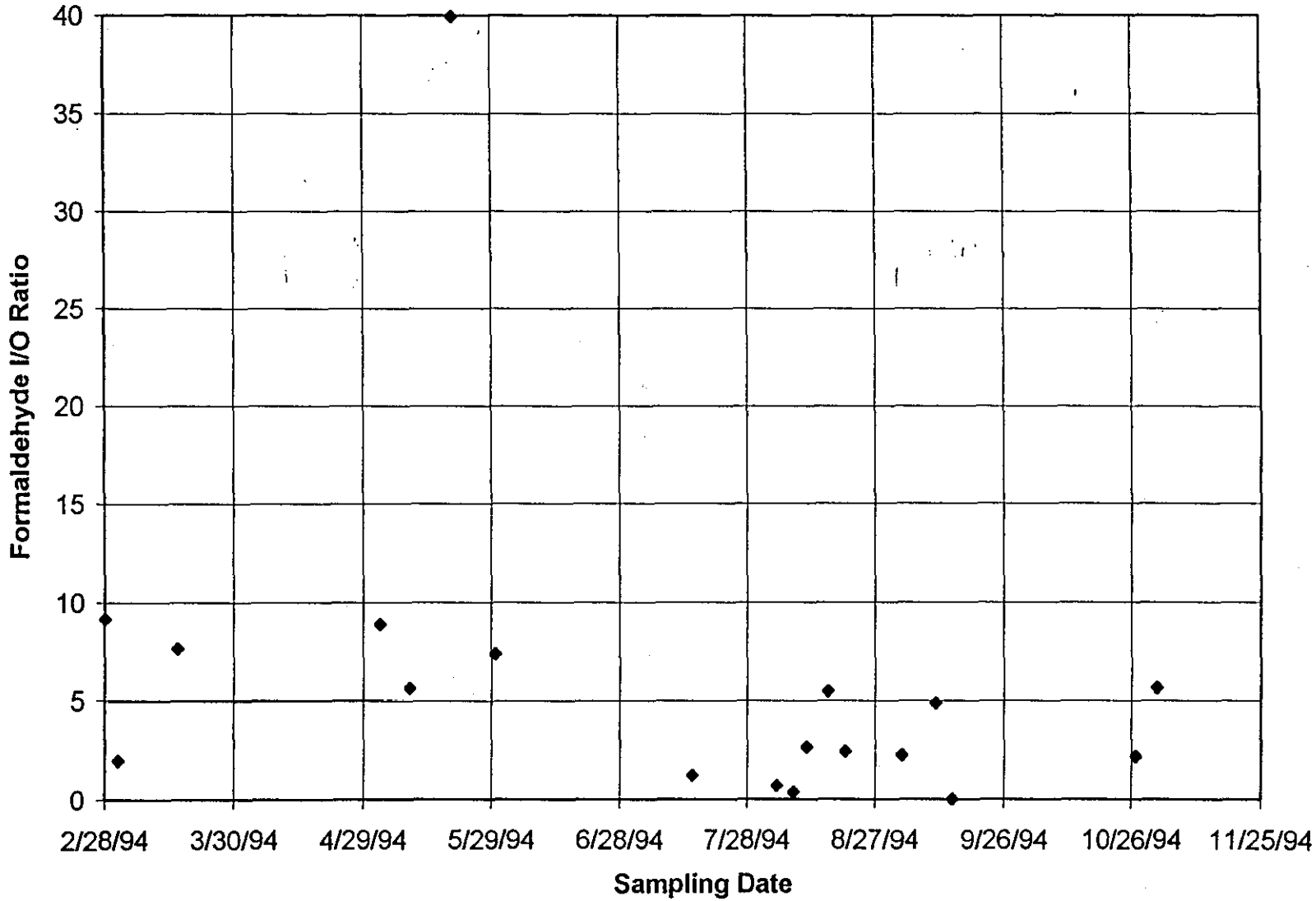


Figure 5.2-49. Temporal pattern of formaldehyde I/O ratio measured at home sites.

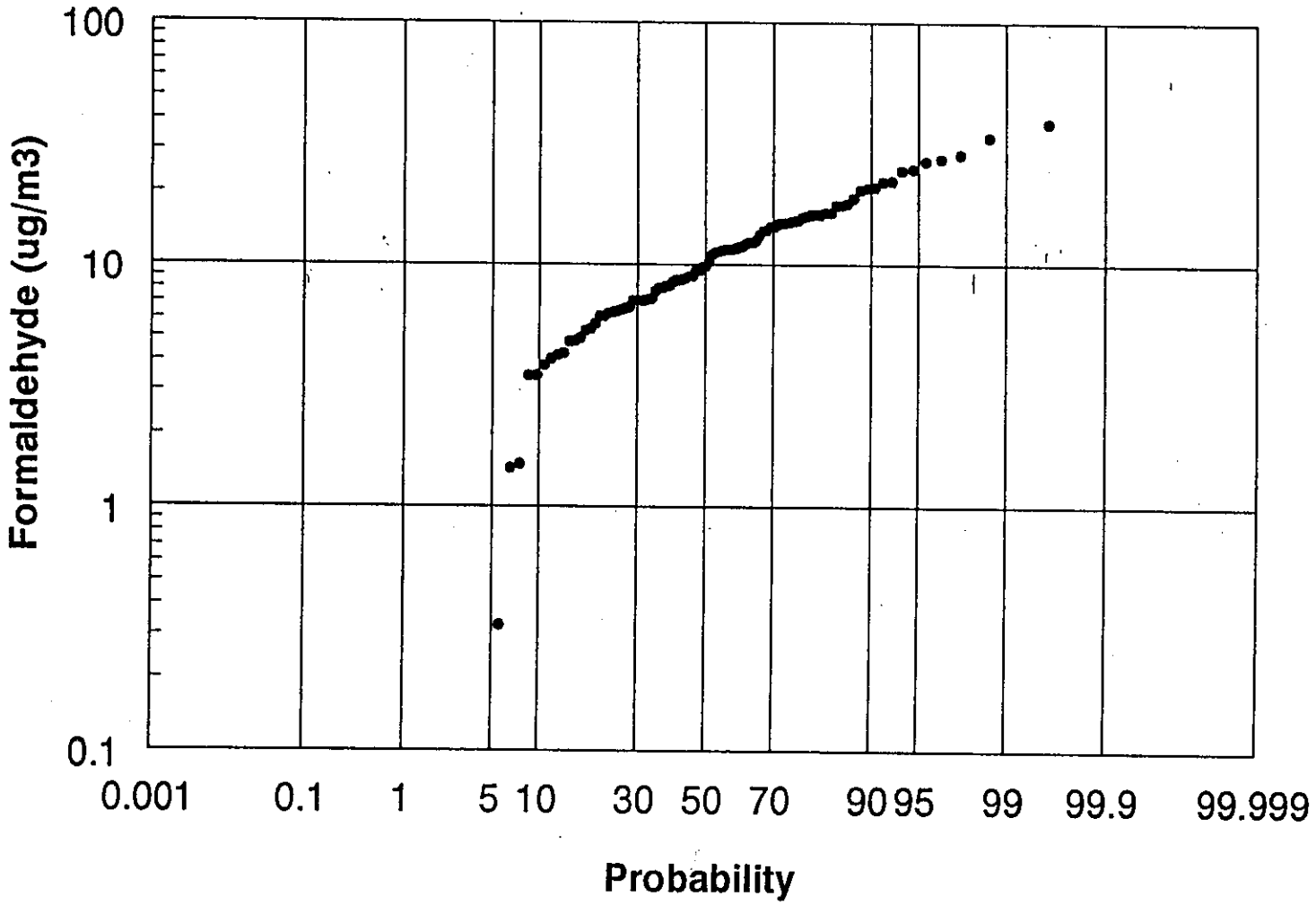


Figure 5.2-50. Log-probability plot of indoor formaldehyde.

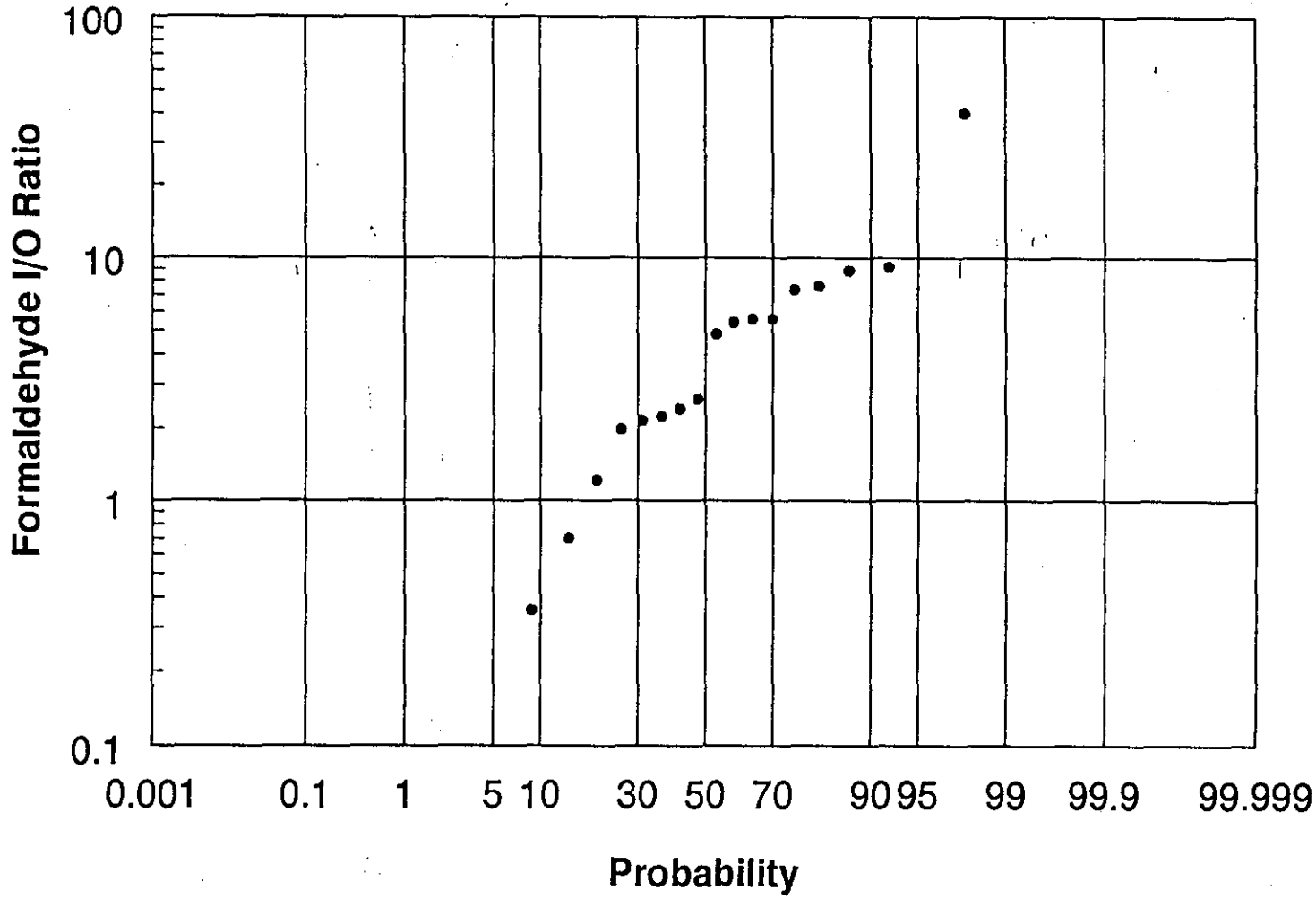


Figure 5.2-51. Log-probability plot of formaldehyde I/O ratio.

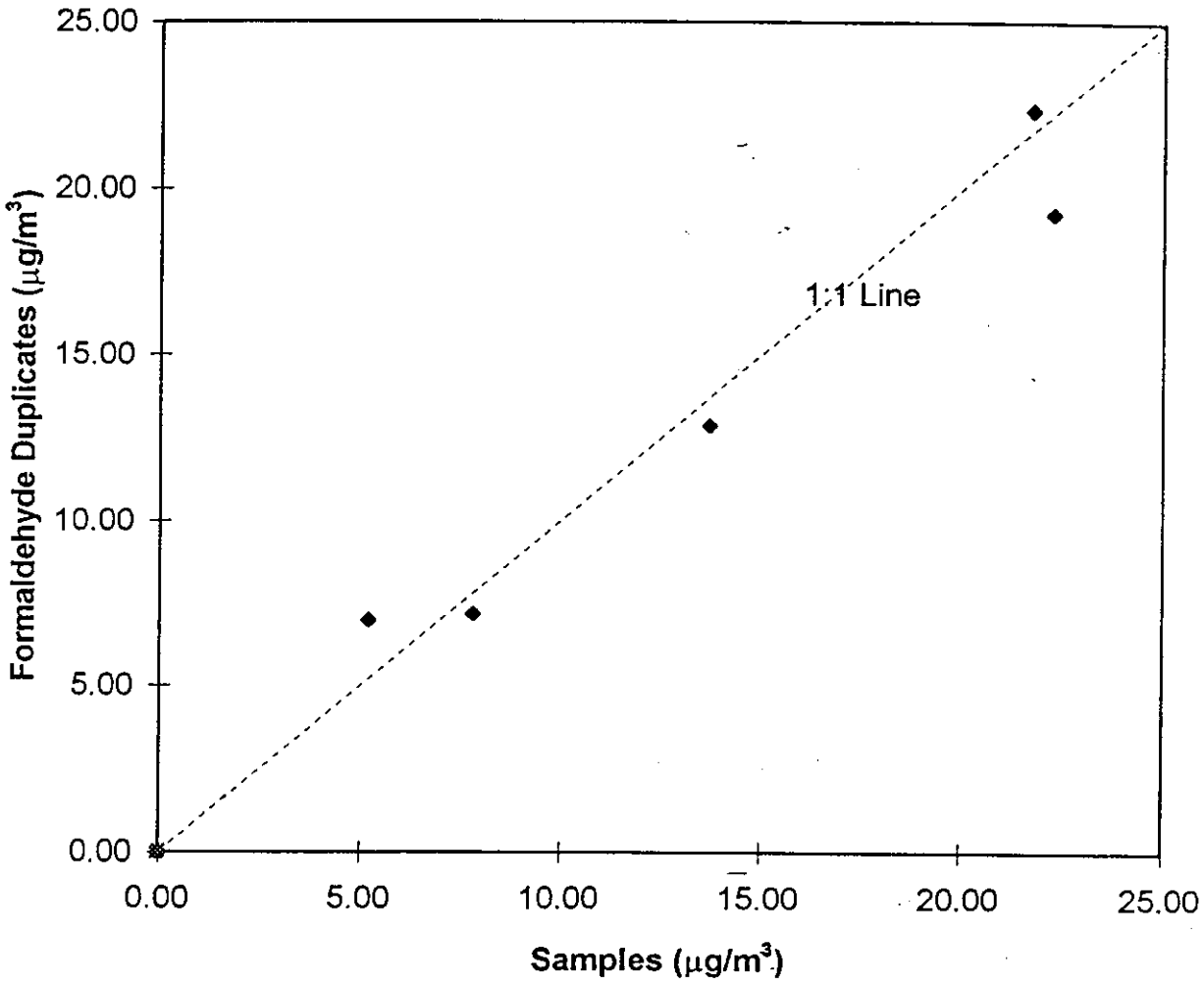


Figure 5.2-52. Scatterplot of formaldehyde samples and duplicates.

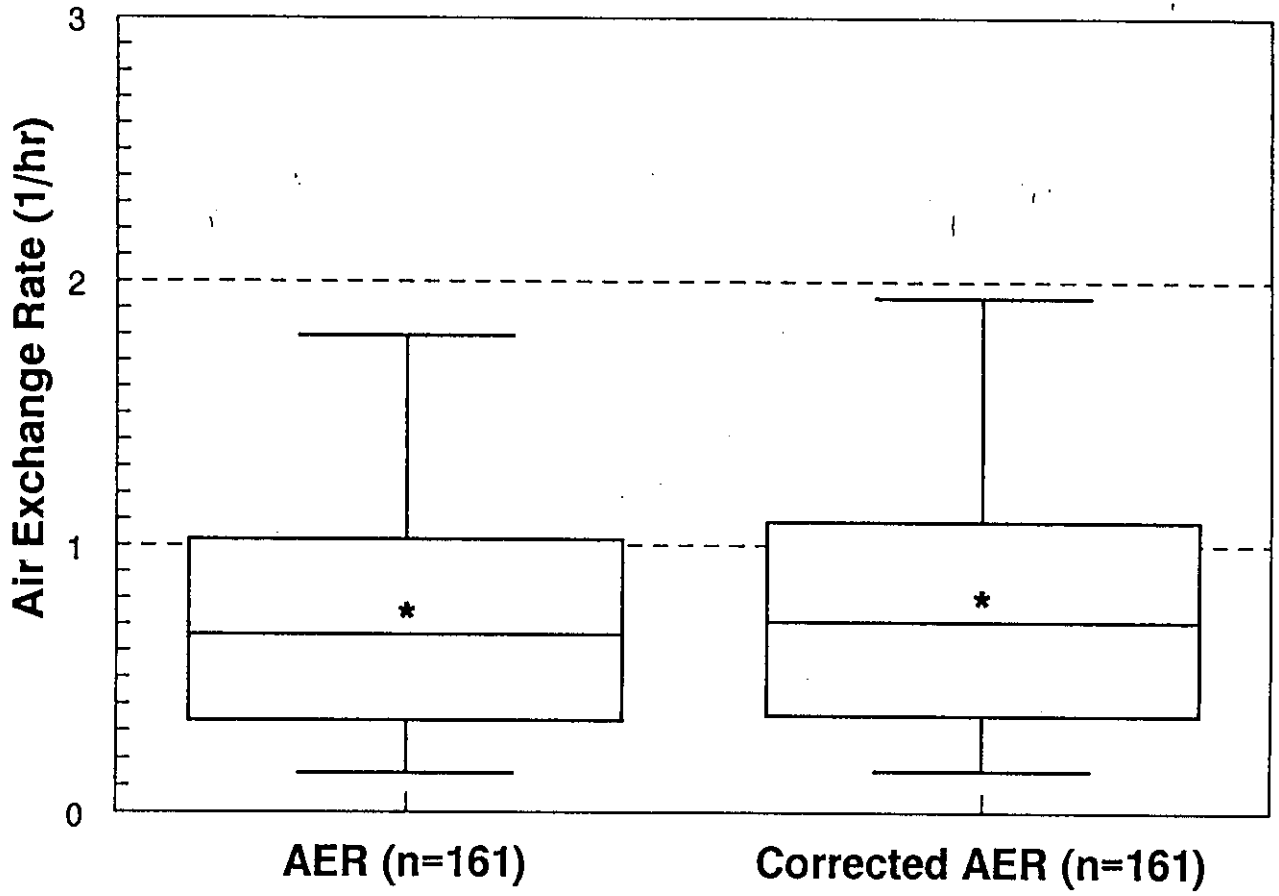


Figure 5.2-53. Box-whisker plots of AER and corrected AER.

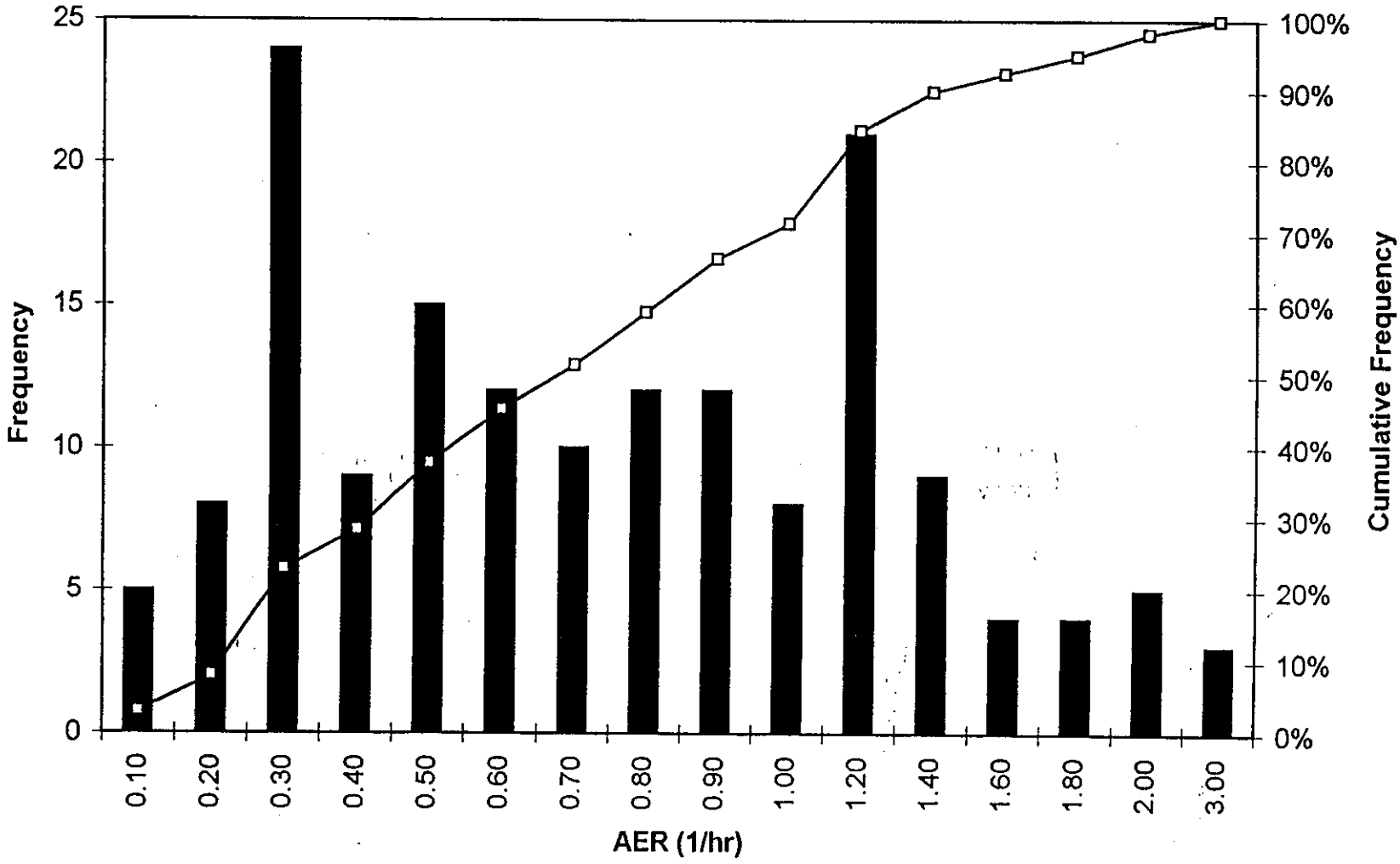


Figure 5.2-54. Histogram of AER.

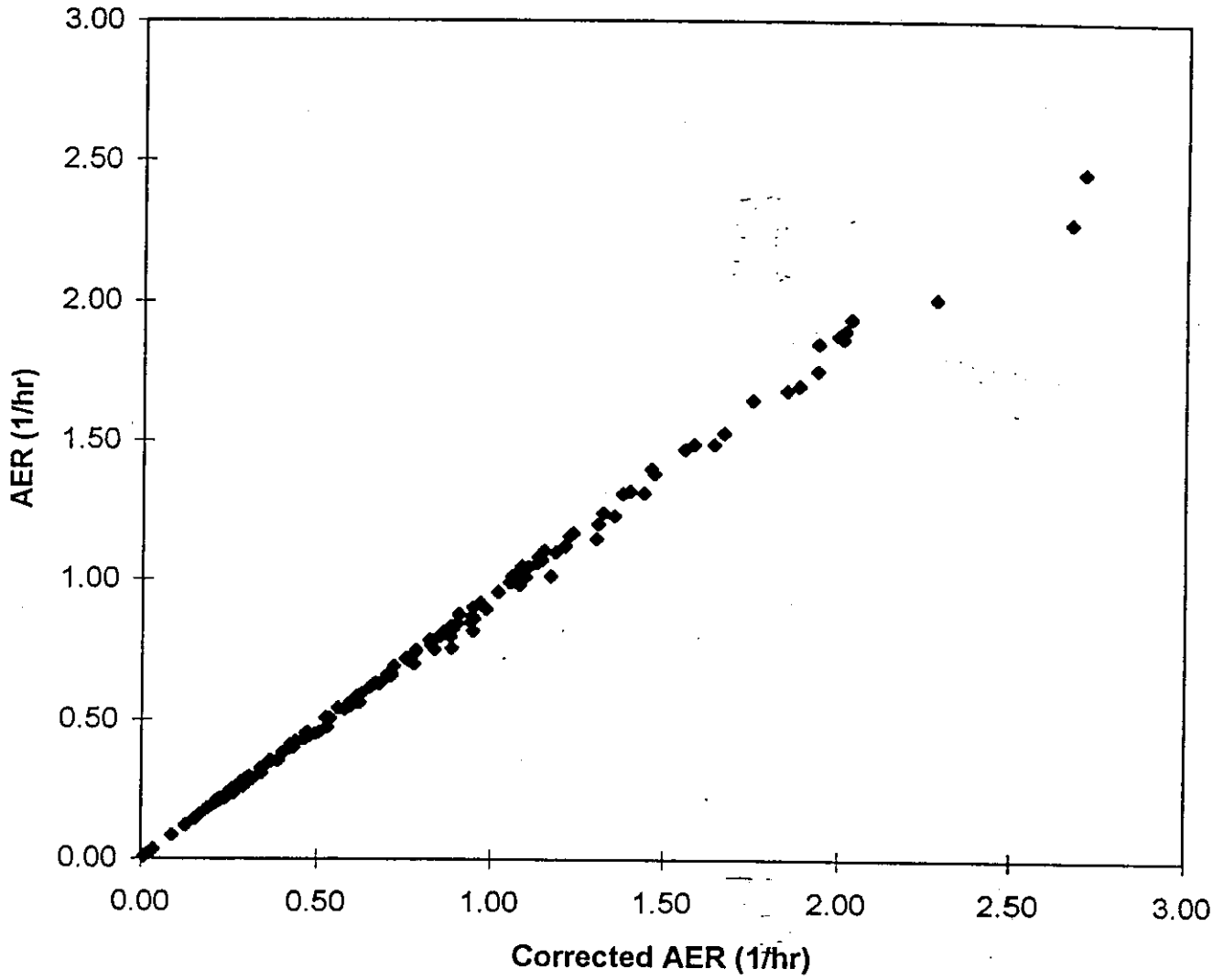


Figure 5.2-55. Scatterplot of AER and corrected AER.

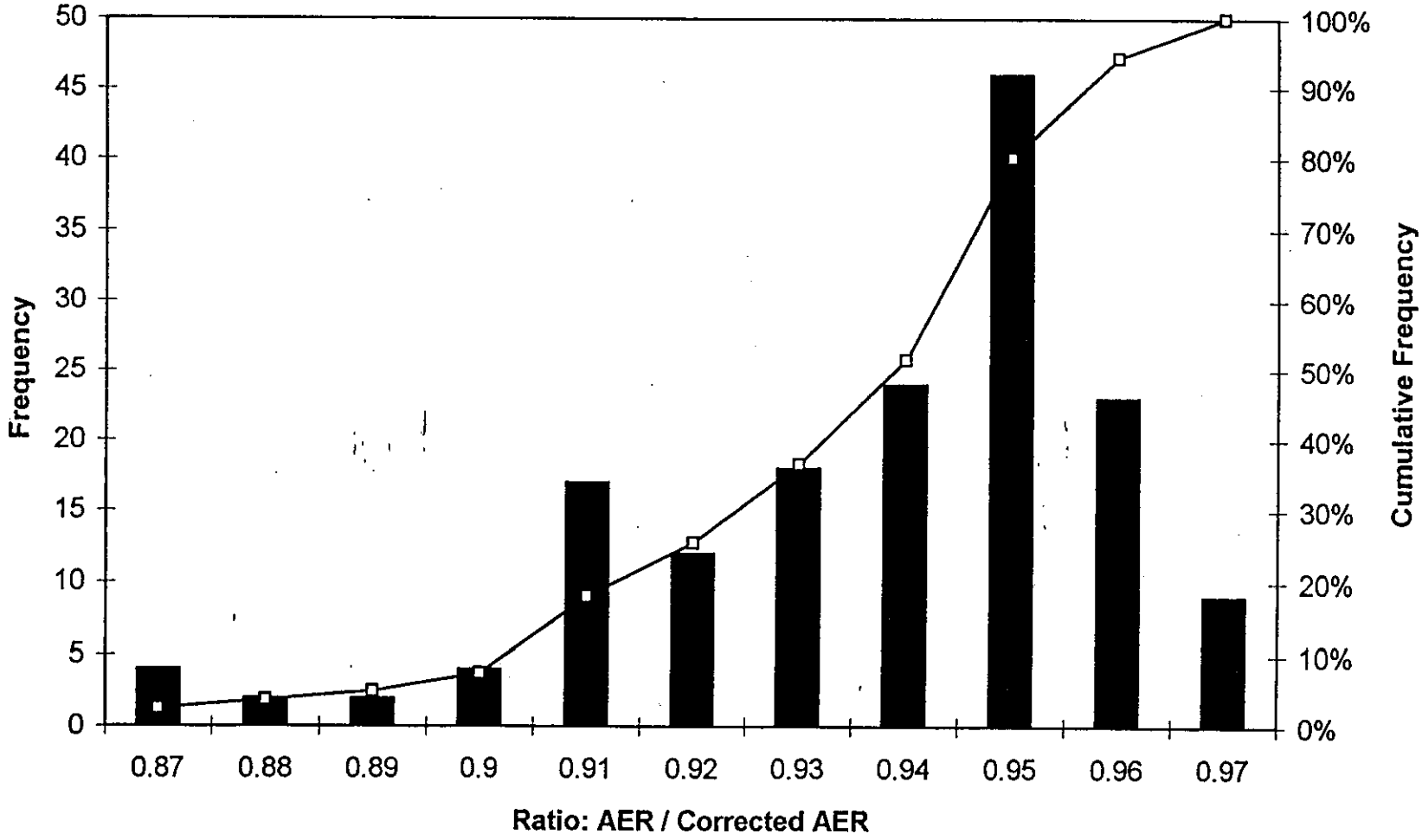


Figure 5.2-56. Histogram of the ratio of AER/corrected AER.

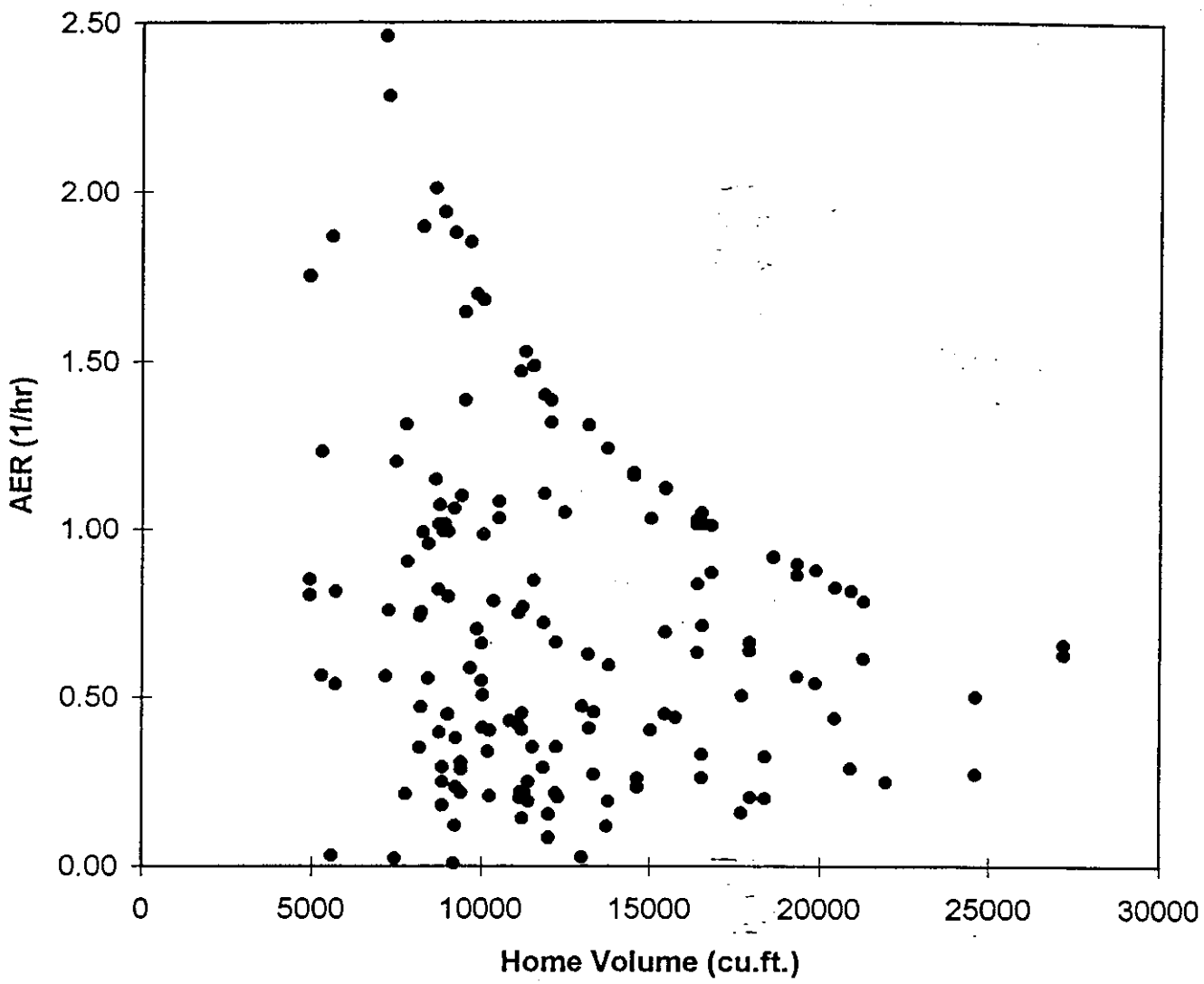


Figure 5.2-57. Scatterplot of AER and home volume.

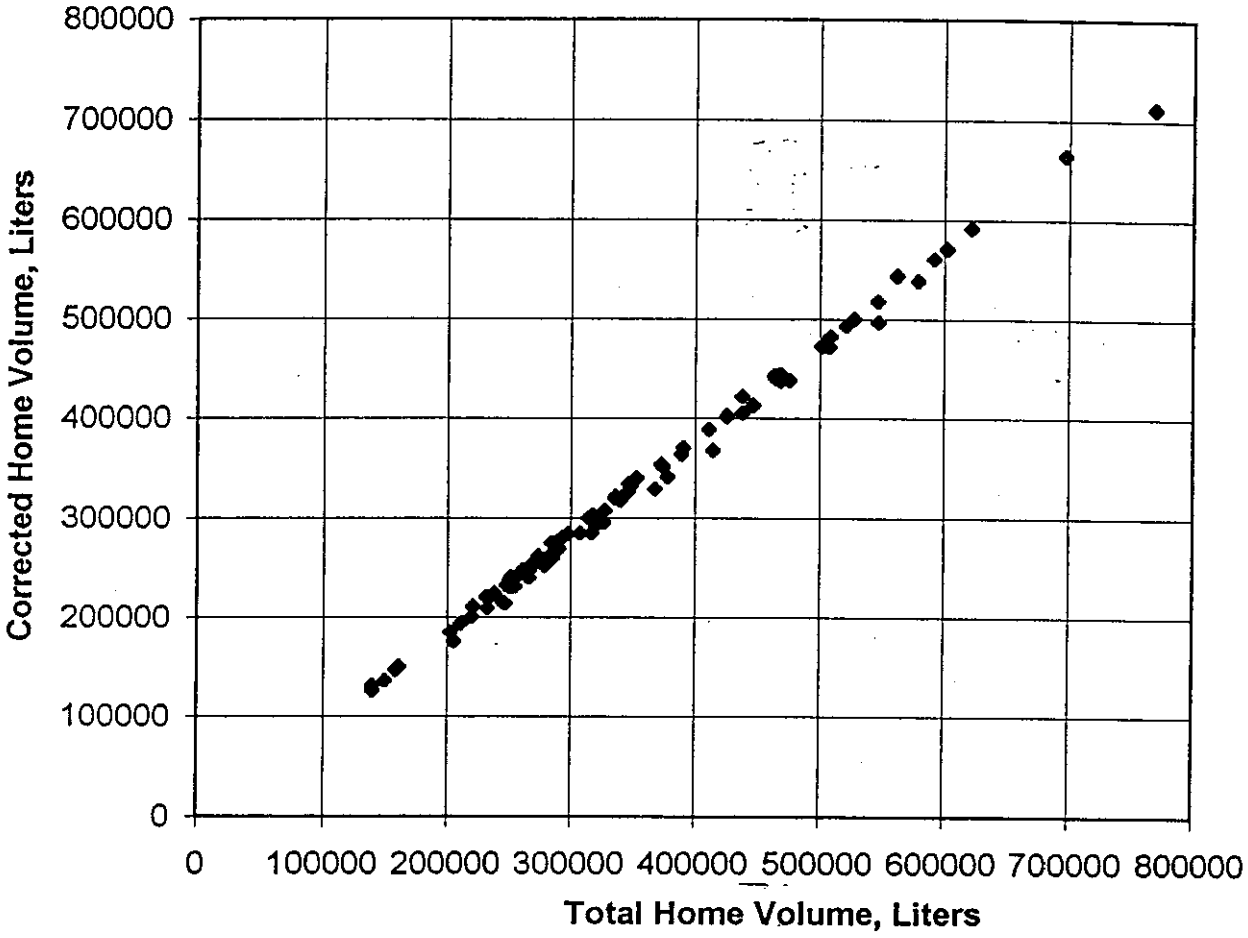


Figure 5.2-58. Scatterplot of total and effective home volume.

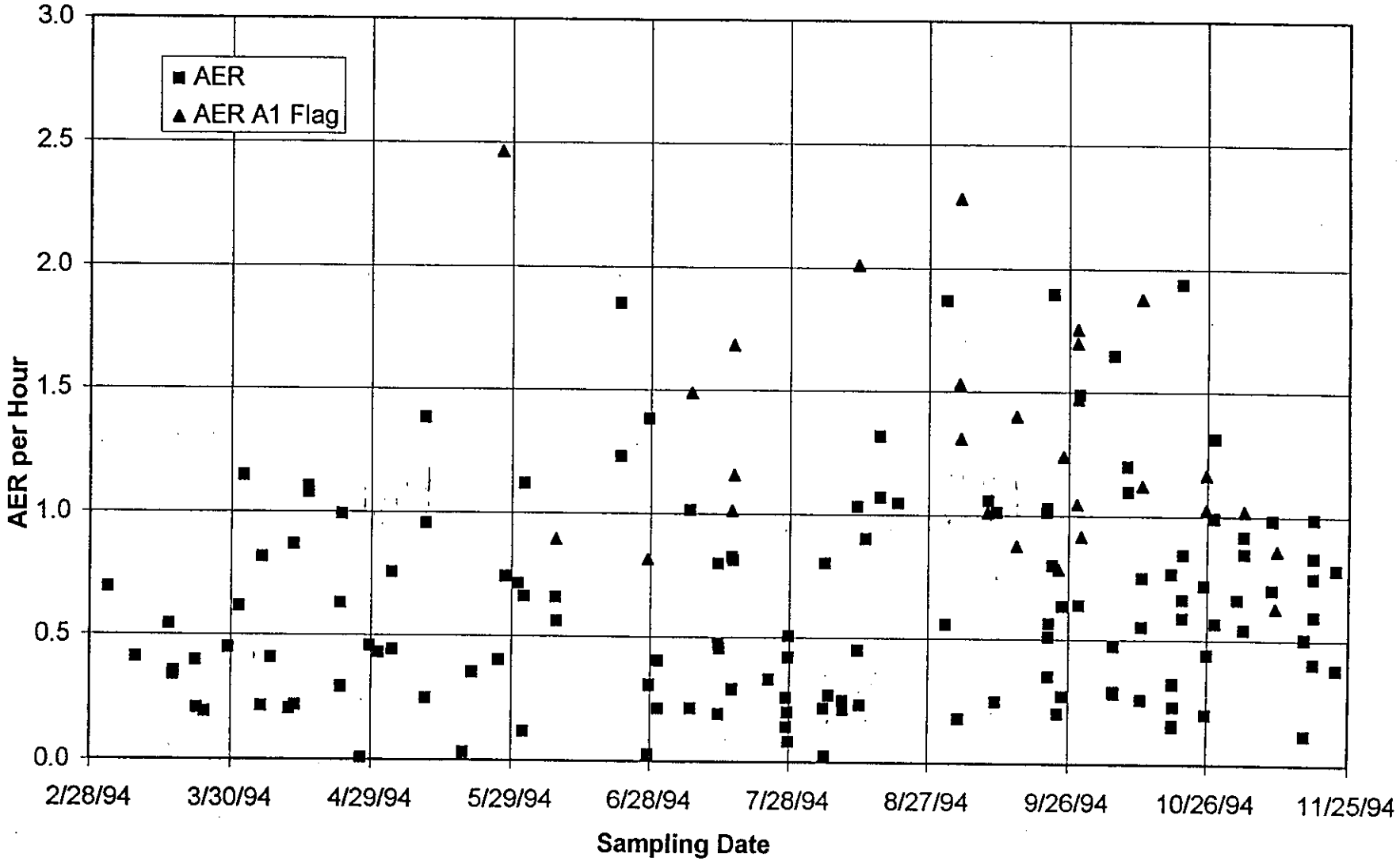


Figure 5.2-59. Temporal pattern of AER in homes.

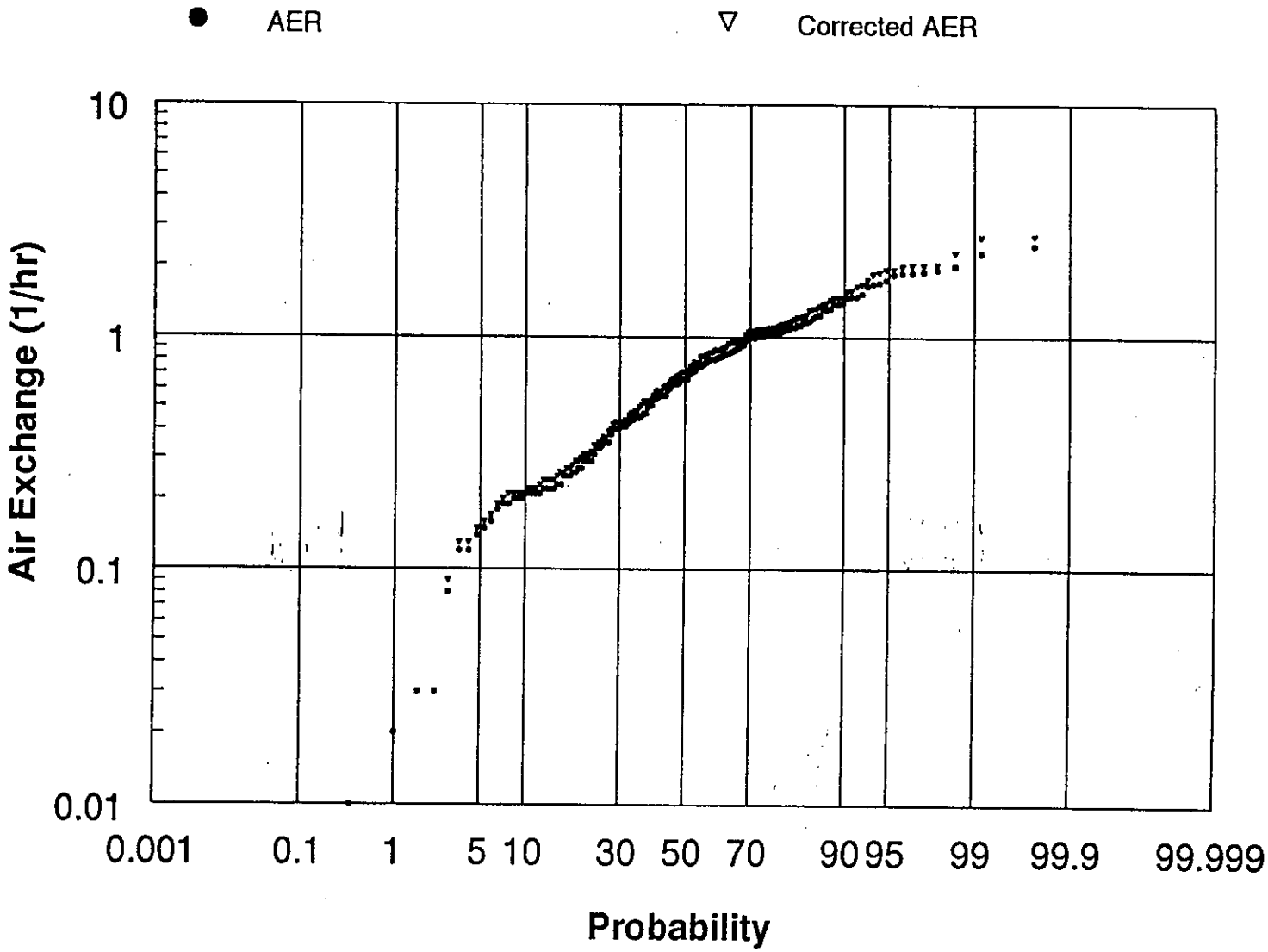


Figure 5.2-60. Log-probability plots of AER and corrected AER.

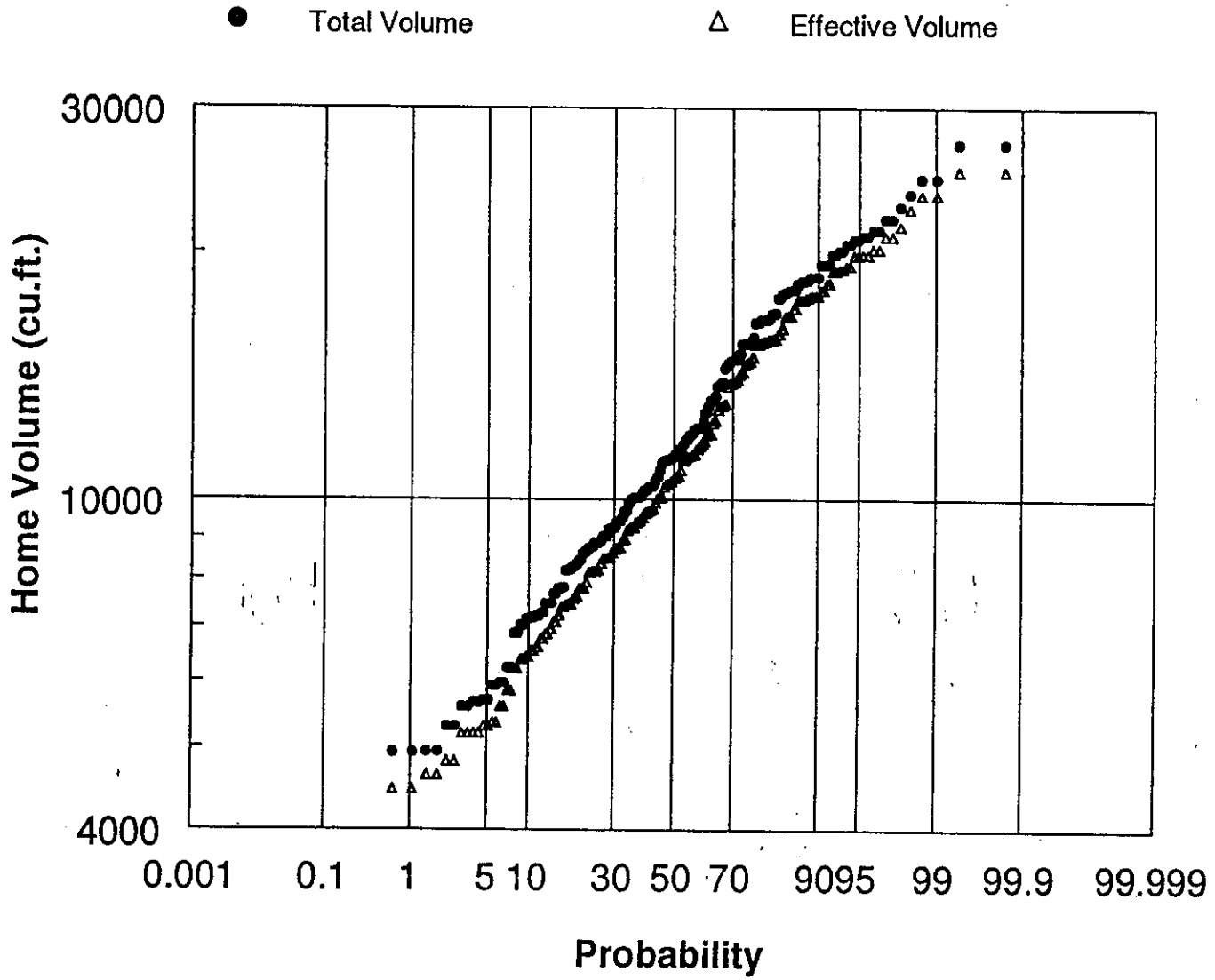


Figure 5.2-61. Log-probability plots of total and effective home volumes.

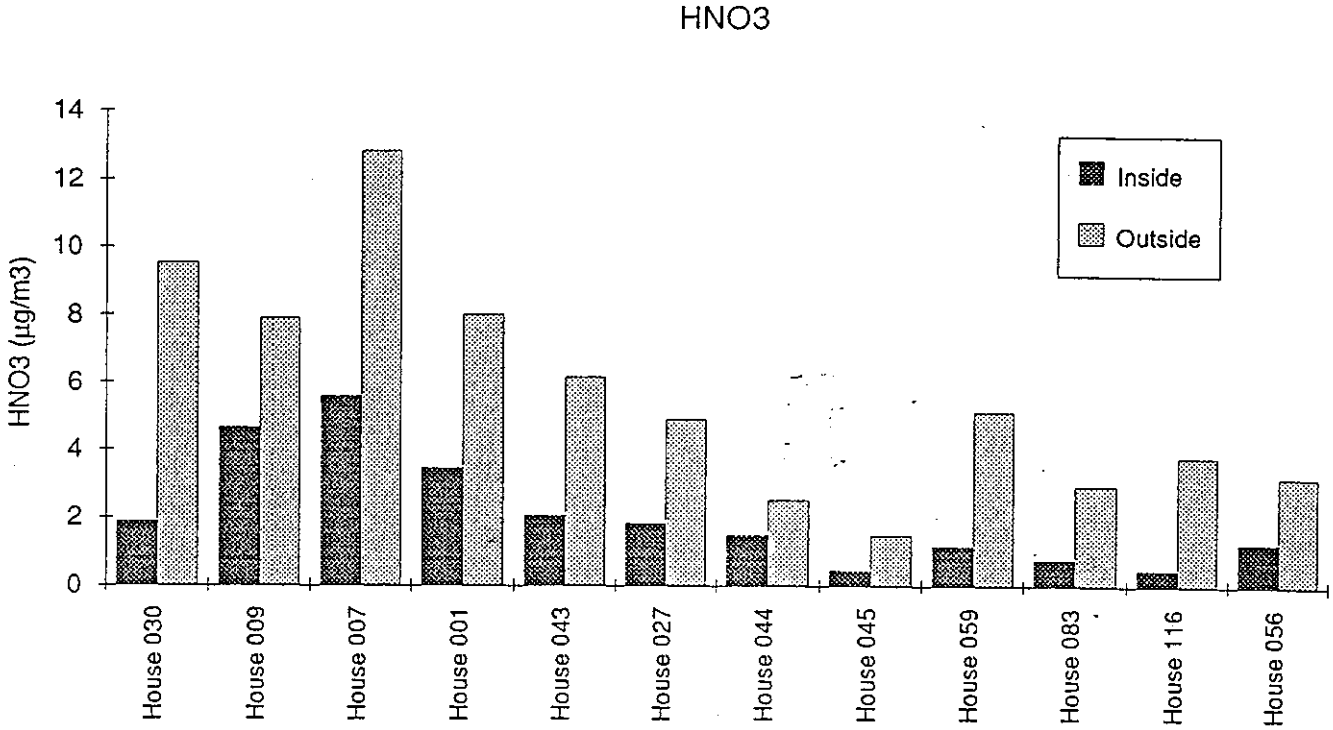


Figure 5.2-62. TWS sampling results - HNO_3, inside and outside homes.

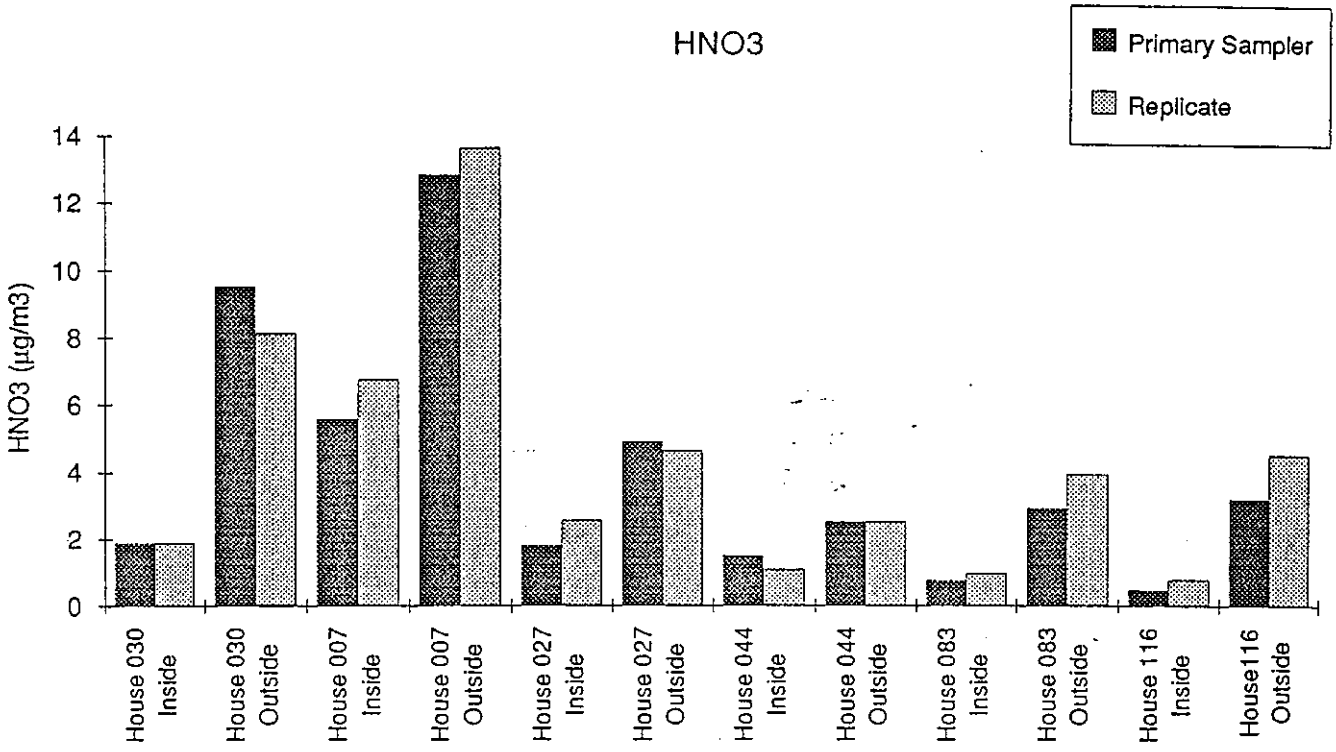


Figure 5.2-63. TWS sampling results - HNO_3, primary and replicate samples.

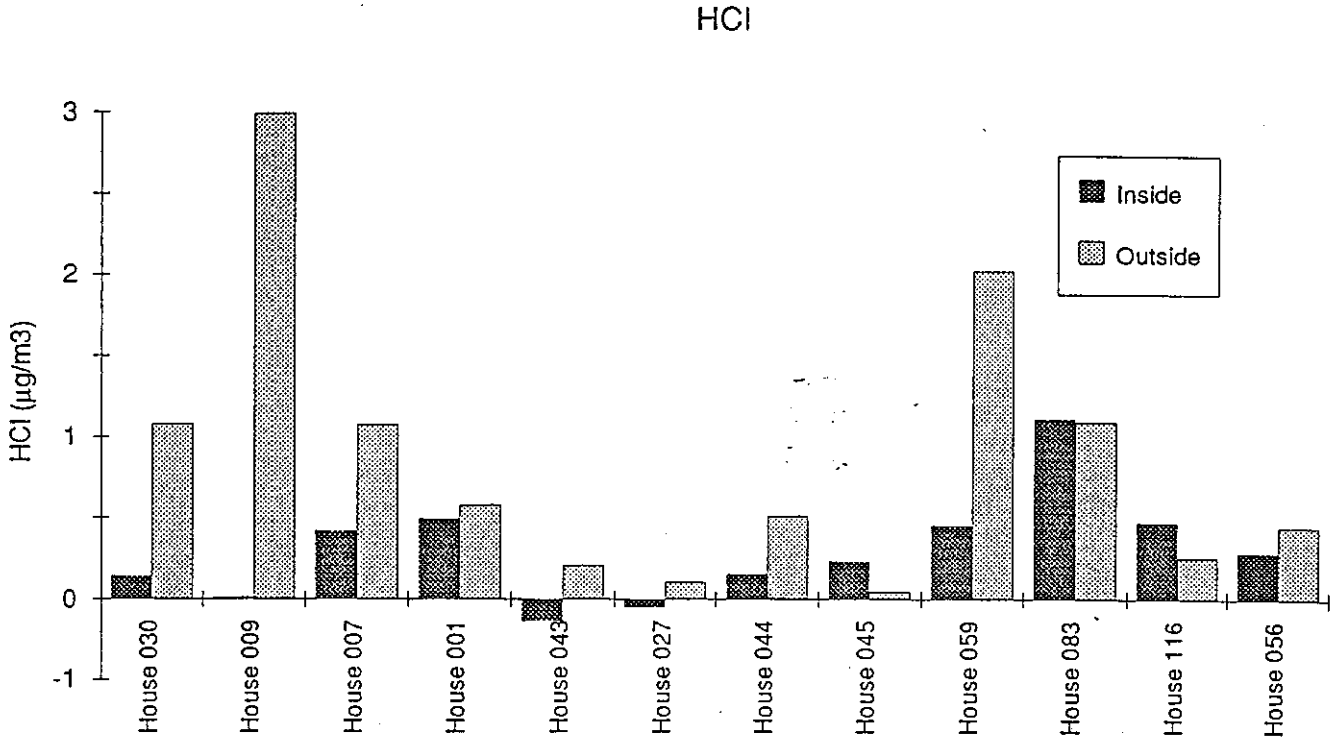


Figure 5.2-64. TWS sampling results - HCl, inside and outside homes.

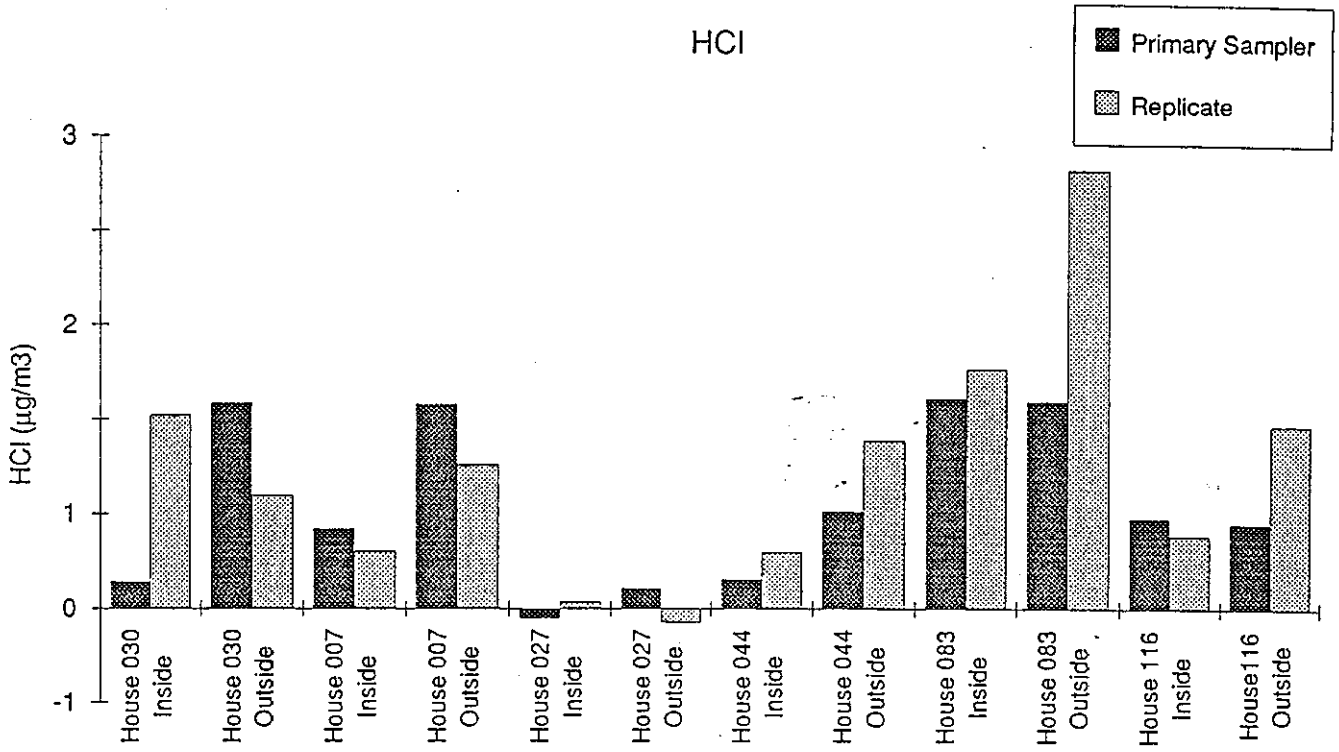


Figure 5.2-65. TWS sampling results - HCl, primary and replicate samples.

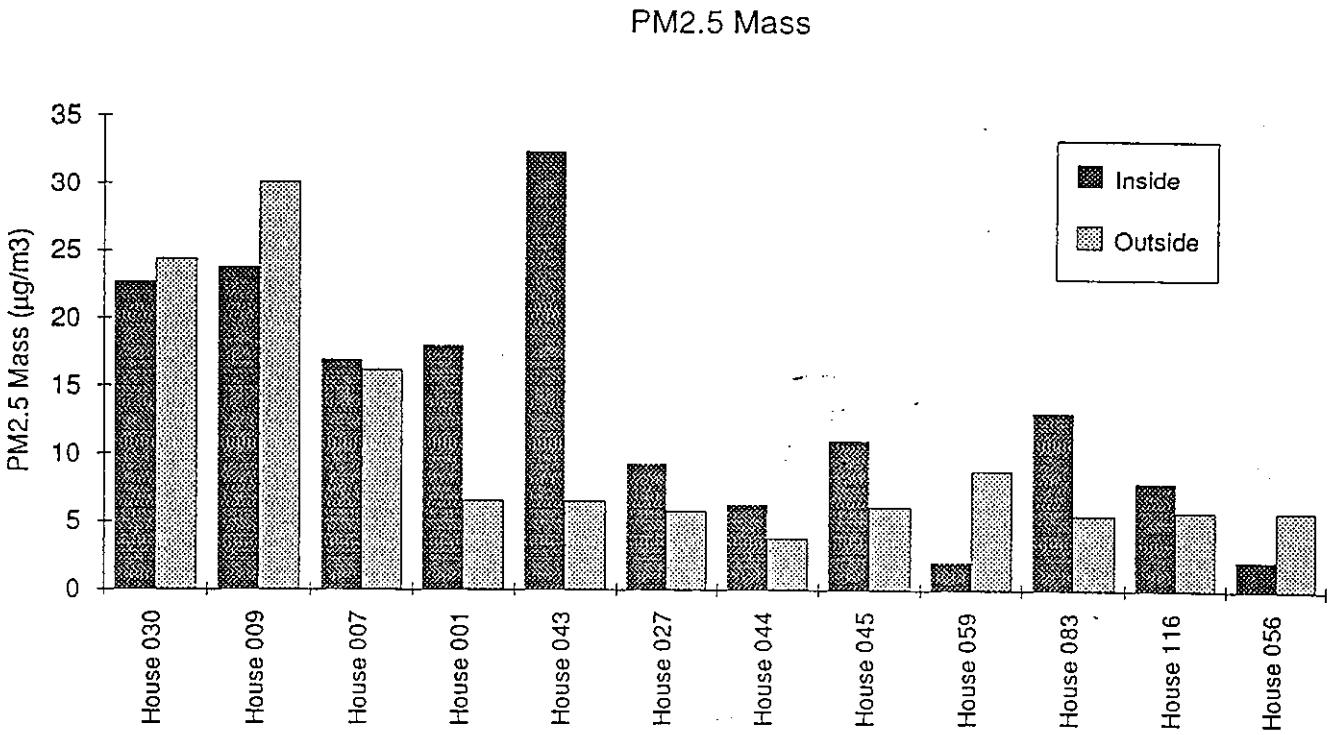


Figure 5.2-66. TWS sampling results - $PM_{2.5}$, inside and outside homes.

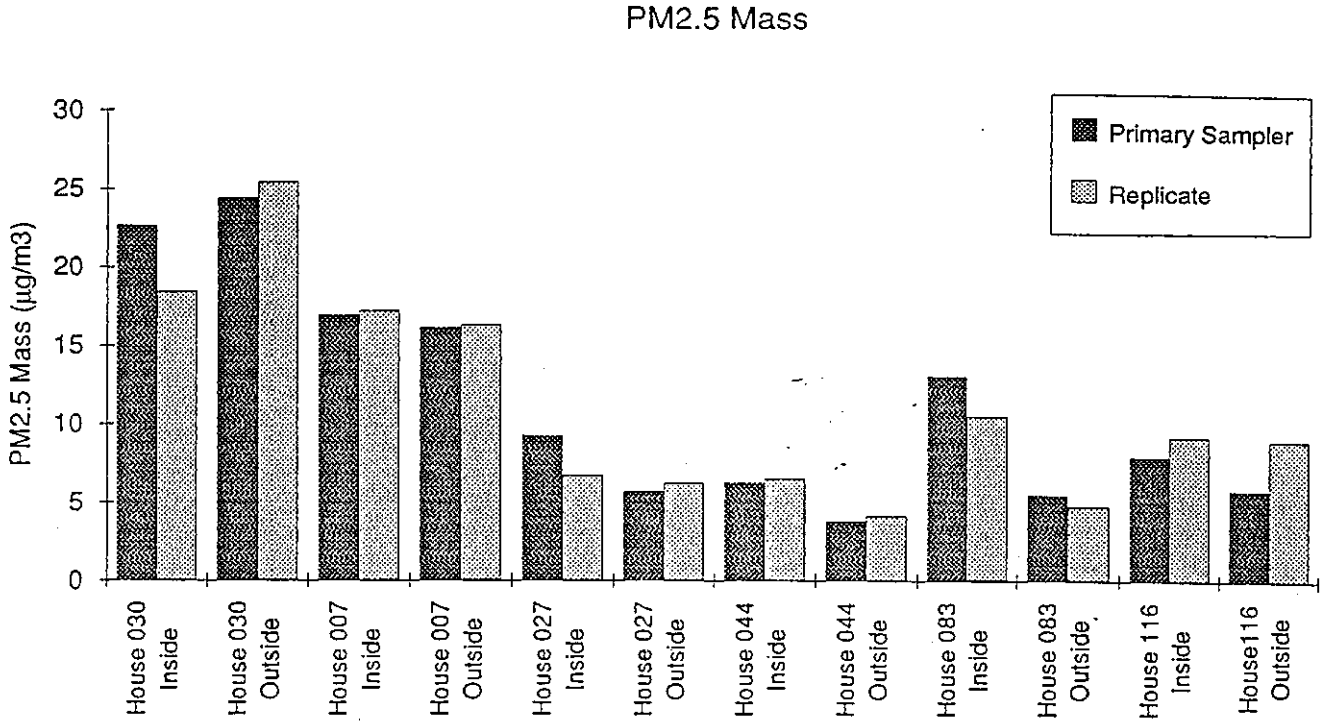


Figure 5.2-67. TWS sampling results - $PM_{2.5}$, primary and replicate samples.

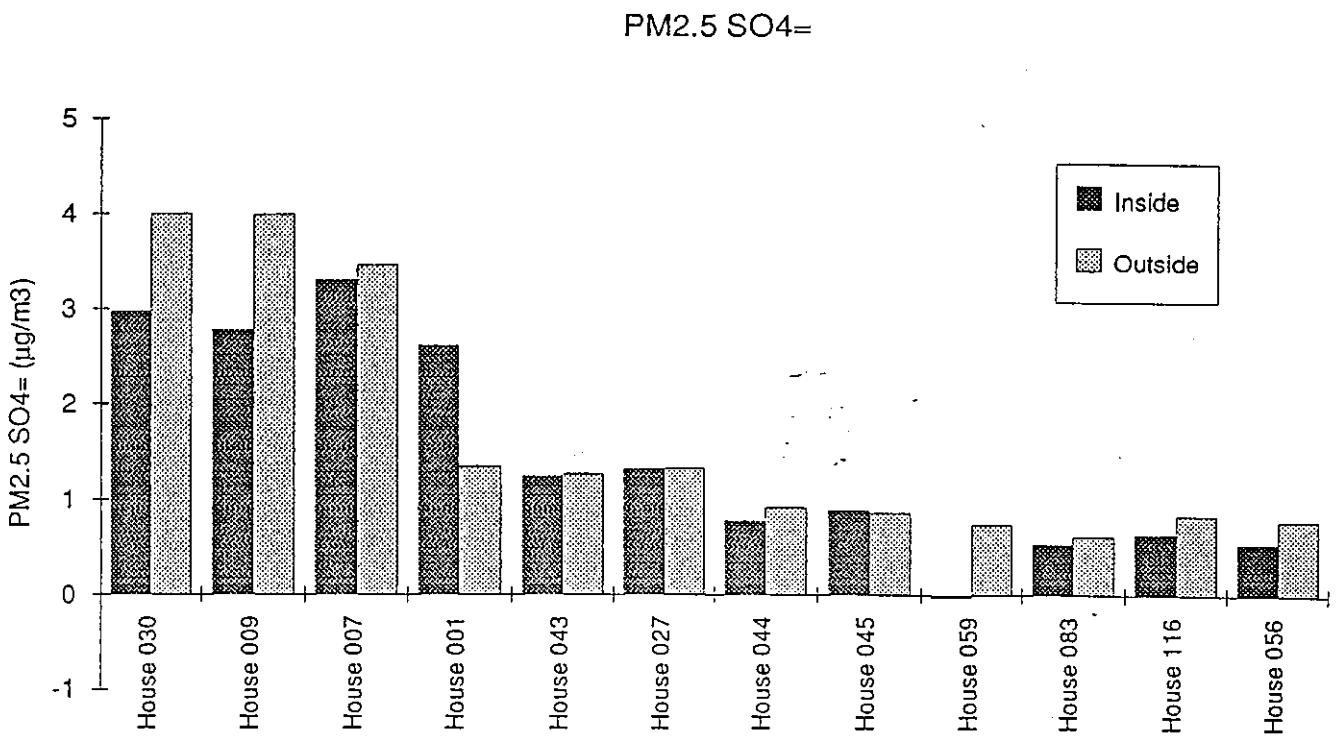


Figure 5.2-68. TWS sampling results - $SO_4^=$, inside and outside homes.

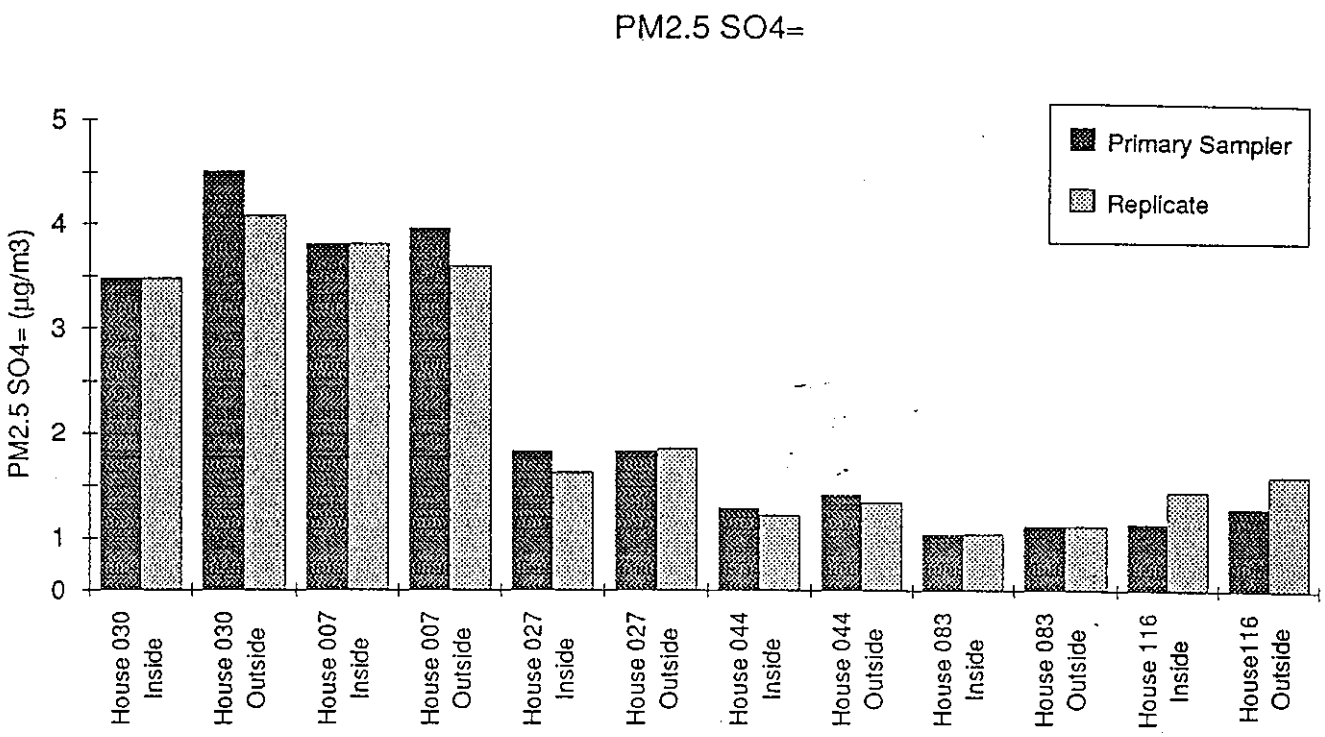


Figure 5.2-69. TWS sampling results - $SO_4^=$, primary and replicate samples.

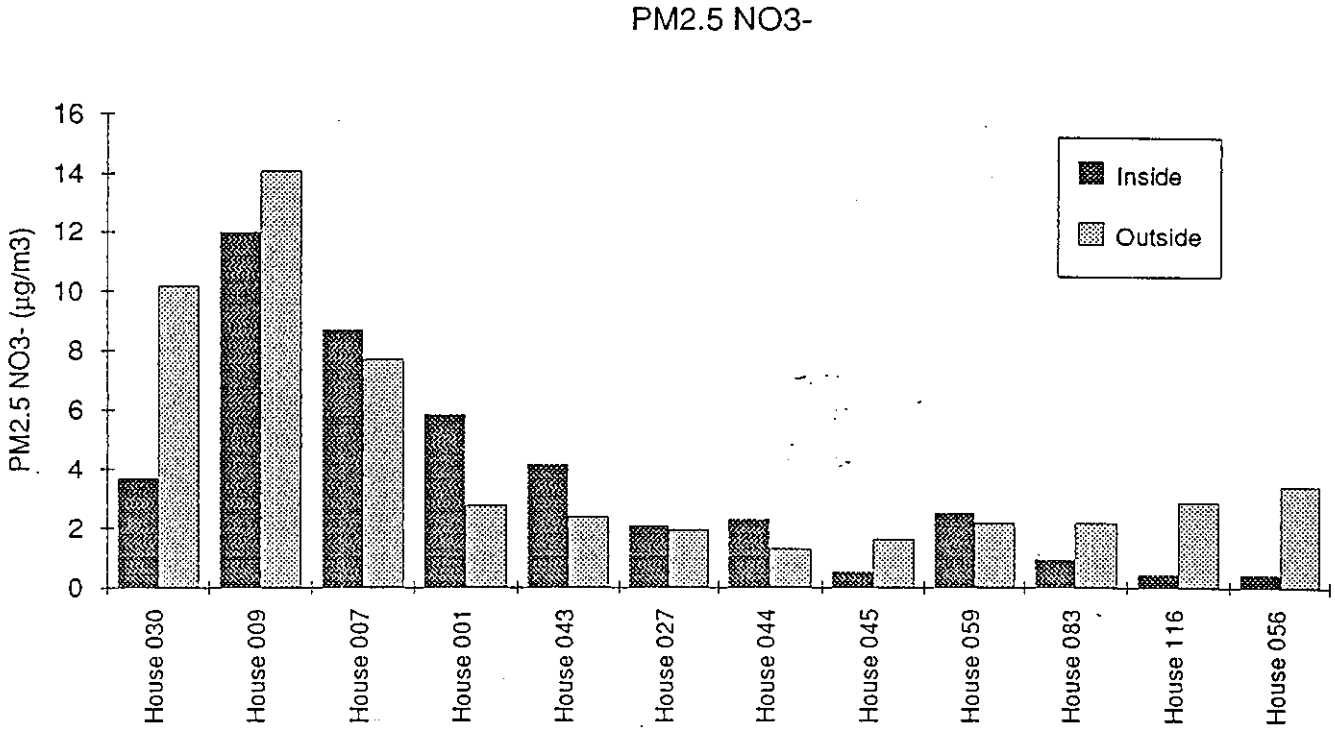


Figure 5.2-70. TWS sampling results - NO_3^-, inside and outside homes.

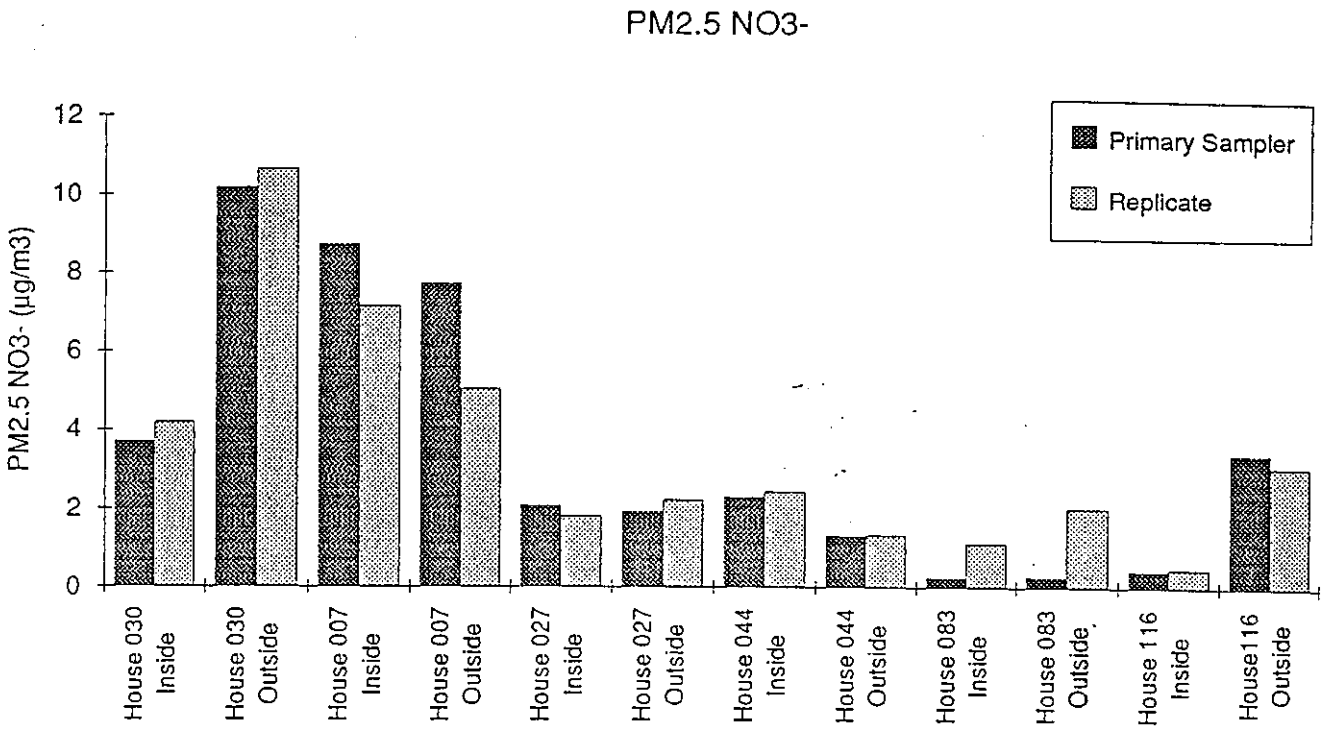


Figure 5.2-71. TWS sampling results - NO_3^-, primary and replicate samples.

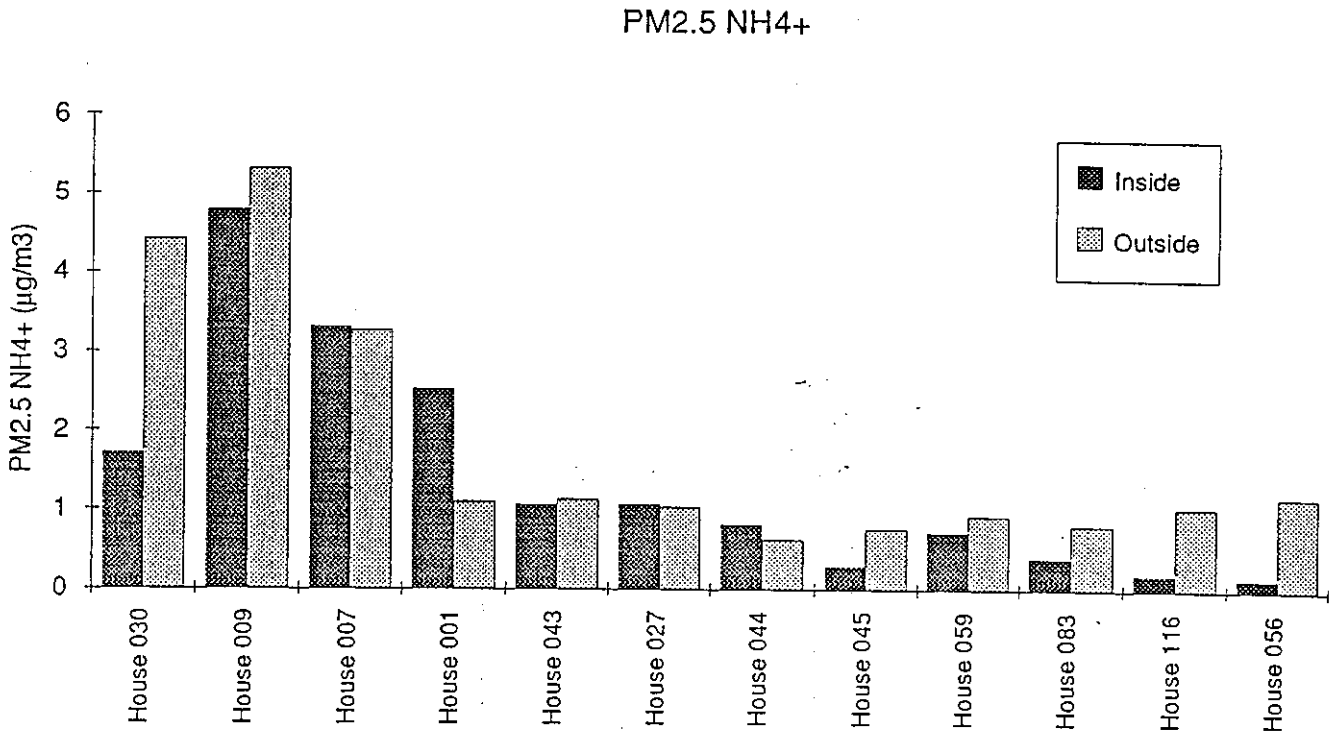


Figure 5.2-72. TWS sampling results - NH_4^+, inside and outside homes.

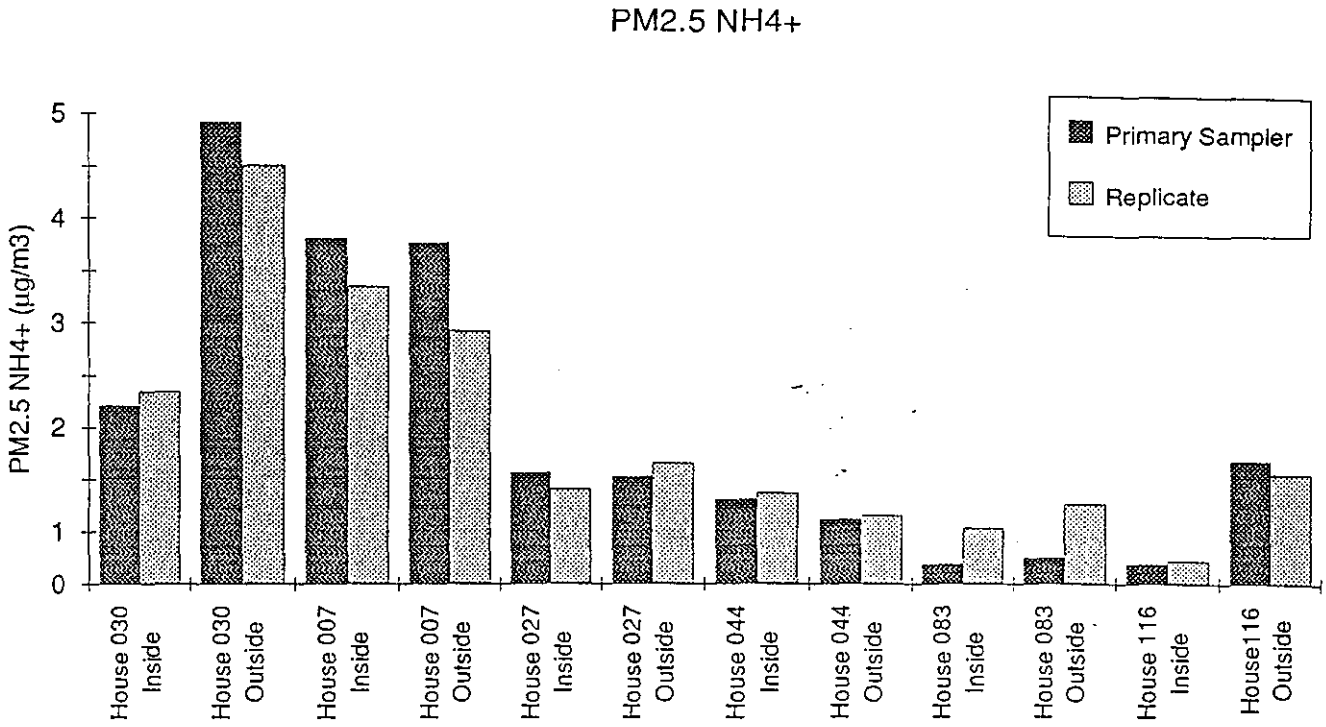


Figure 5.2-73. TWS sampling results - NH_4^+, primary and replicate samples.

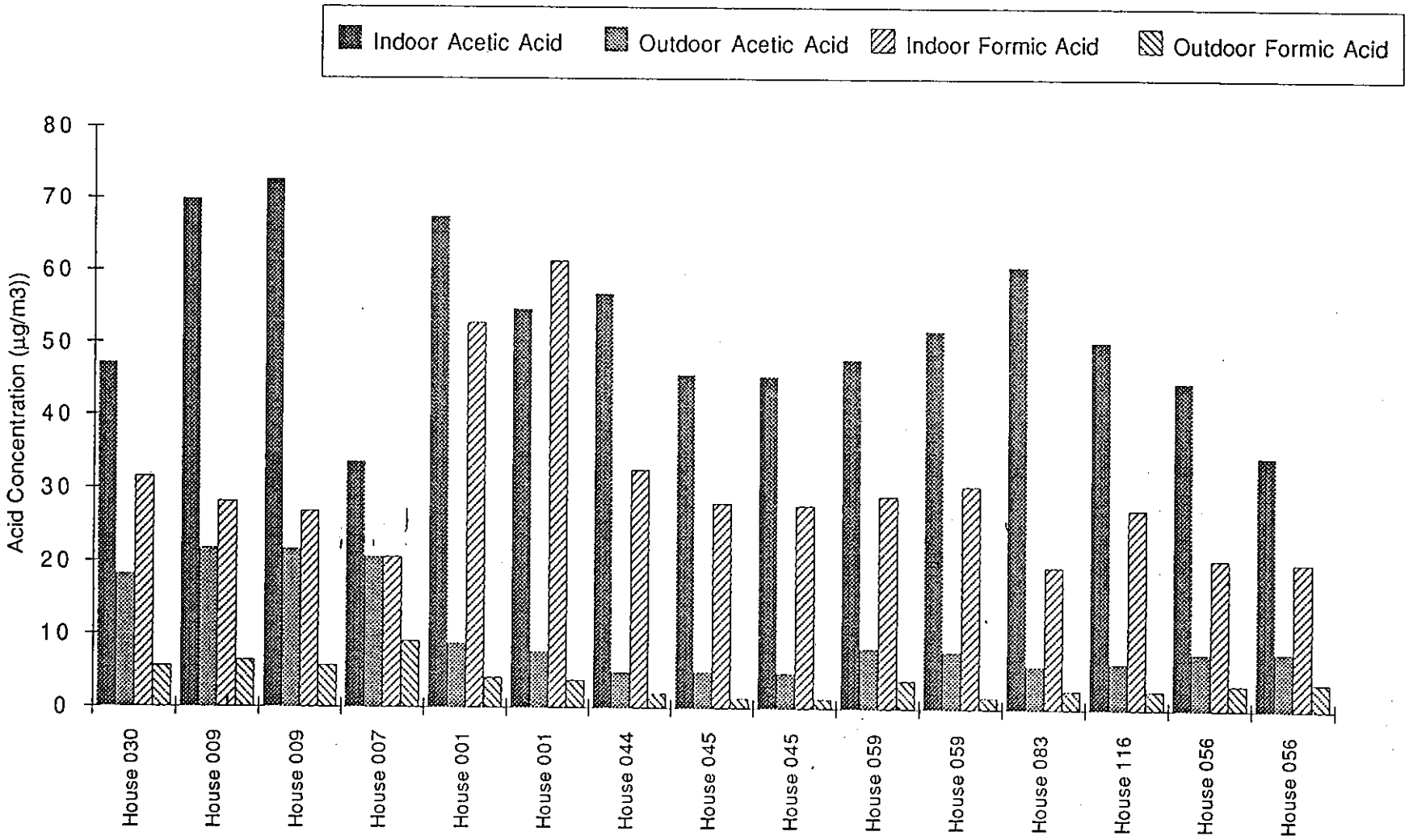


Figure 5.2-74. TWS sampling results - Organic acids, inside and outside homes.

6. MODEL DEVELOPMENT AND REFINEMENT

The collected data were used to evaluate a variety of modeling approaches to gain insights into the relationships between pollutant levels measured at community-based monitoring stations and measurements made immediately outside or inside study homes. Station and study data for ozone and PM_{10} were used to perform the present work. As explained in the respective sections, community $PM_{2.5}$ levels were derived based on previous work, since routine community-based $PM_{2.5}$ data were not available. No development of a predictive model for formaldehyde was undertaken, since routine community-based formaldehyde data was not available, and only 18 outdoor formaldehyde samples were collected in the performance of the residential study.

The following sections describe the models developed and the results obtained. The value of the intercepts in the models we have fit represent, in principle, the predicted indoor levels of pollutants in residences for which all the explanatory variables in the selected model have a value of zero. The intercept is a meaningful predictor only when it is estimated with data that contains observations around zero for each variable; otherwise, interpreting the intercept is an exercise in extrapolation. In the residential study data set presented here, a hypothetical residence with all variables equal to zero would lie outside the range of the data. For this reason, no importance should be attached to the p-value of the estimated intercepts in our models.

In principle, distributional assumptions such as normality and independence can be of concern in modeling approaches. We used the t-statistic in our initial screening process to select a group of variables for further consideration. Under certain assumptions(e.g. normality), the null distribution of this statistic is known, and tabulated p values are exact. However, it is true in general that the t-statistic measures the explanatory power of a variable. It is always true that the greater the value of the t-statistic, the greater the reduction in the mean squared error of the residuals. Therefore, the t-statistic is a valid screening device under quite general conditions, so we did not find it necessary to extensively examine the distributional assumptions.

Model selection procedures can introduce bias into the standard goodness of fit statistics, even when standard assumptions are met. For example, values of R^2 computed by the usual procedure

are often too large, because variables have been selected to maximize this quantity. To assess the degree of this bias, it is best to use a cross-validation approach. We conducted such an analysis, and we describe the procedure and its results in some detail in Section 6.1.1.1. Based on the results of that analysis, we found the bias due to model selection to be minimal.

Pollutant data often follow skewed distributions, closer to log-normal than normal representations. Transforming these variables to achieve near normality might result in the final model being somewhat different. It is not clear, however, that models constructed with transformed variables would fit the actual data better than a model constructed with the un-manipulated (or raw) variables. In fact, transformations to normality can have a deleterious impact in some cases. In the residential study data set, for example, many values observed in the indoor ozone data base were near zero. These values, by nature of their low concentrations, were less reliably measured than other larger values in other homes. However, a log transformation of the data set (to achieve near normality) would transform these small values into large negative numbers, which could exert considerable influence on the modeling results. Thus, we chose not to transform the collected data, but rather to investigate models constructed using the raw (un-manipulated) data.

Many variables have large co-variances. This can present a problem in determining causality, since pairs of correlated variables are confounded. There is less problem in building a predictive model, however. When two variables are highly correlated, they generally have roughly equal predictive power. Thus either one may reasonable be chosen for inclusion in the final model, assuming both variables are physically realistic.

Community was not considered as a variable in any model, because it would not be useful in predicting indoor levels outside the specific communities studied. We also chose not to include air exchange rate (AER) as a variable in any of the developed models. This measurement is a resource-intensive parameter to obtain, in terms of time, labor, and cost. More importantly, AER data would not typically be available in practice, so it would not be helpful in a predictive model.

6.1 Modeling Indoor Pollutant Levels

The modeling approach exploited in this study was similar regardless of specific indoor pollutant. Responses to the field technician and follow-up questionnaires, along with outdoor ambient levels and temperatures, were considered as potential explanatory variables in linear models whose outcome variables were the measured indoor levels of study interest (O_3, PM_{10}, and $PM_{2.5}$). Survey data from the baseline questionnaire, completed by the legal guardian of each subject participating in the Children's Health Study (the population from which the residential study homes were identified and selected), were also considered as potential variables in our model development (see Appendix A for specific questions assessed for this investigation).

Variable selection proceeded in a step-wise manner. We eliminated from model consideration variables with more than 24% of the values missing or with more than 90% of the responses identical. We then screened the remaining variables by the "t-test method" (Freedman, Navidi, and Peters 1988; Lysikowski 1995). In this method, each variable was tested univariately, and those found to be statistically significant at a threshold level of 15% were retained for further consideration. The decision to disqualify some variables from subsequent consideration in the modeling process was made in an objective and careful manner. In order that results be reproducible, modeling decisions must be made on the basis of specific, well-defined criteria. Our univariate criteria were designed to reduce the number of variables considered on a multivariate basis to a reasonable level, while preserving, to the greatest extent possible, all potentially relevant variables.

Having generated a list of variables to find the best model for a given number of variables by the selection approach described above, we then applied the method of best subsets (Miller 1990). Using this method, we chose, for subsequent modeling use, the largest model that provided a substantial increase in the coefficient of determination (R^2) over the model with one less variable. To determine which interaction terms to include, we once again applied the method of best subsets to the set of variables consisting of those selected in the previous round together with all of their two-way interactions. The final model was chosen according to the Mallows C_p

criterion (Mallows, 1973), which adjusts the values of R^2 for the effect of including additional variables in the model.

6.1.1 Modeling Indoor Levels of O_3

The screening approach described above was applied to the O_3 data set. In conformance with that procedure, each of the remaining variables was entered into a model whose only other explanatory variable was the ambient level of O_3 measured at the closest fixed-site monitoring station. Variables not significant at the 15% level in this univariate setting were dropped from consideration. (A listing of all variables considered and their determined p-values appears in Appendix I).

From those variables remaining, all linear models containing ambient O_3 and five or fewer additional explanatory variables were considered. Table 6-1 presents the best fitting models for each number of variables, where fit is measured by percentage of variance explained (multivariate R^2), in a weighted least squares regression with variances assumed proportional to ambient levels of O_3 measures at the monitoring station.

As the number of variables increases, the maximum attainable R^2 necessarily increases. We chose to work with the four-variable model, because improvements in R_2 were marginal for models with more than four variables. In the four-variable model, the explanatory variables chosen by this process were: ambient level of O_3; the number of hours that the home's windows were reported open; the minimum outdoor temperature; an indicator variable for the use of a central refrigerant type air conditioner with recirculating air. All models involving these four variables and their two-way interactions were considered. The best fitting model was selected on the basis of the Mallows C_p criterion (Mallows, 1973), which adjusts the values of R^2 for the effect of including additional variables in the model. In addition to the four variables previously identified, the best fitting model also contained interaction variables, between the number of hours that the windows were open and the ambient O_3 level, and the number of hours that the windows were open and the minimum temperature. This final model is presented in Table 6-2.

The resulting prediction equation for this model (R^2 of 0.55) is:

Indoor O_3 = -1.507 + 0.053 x Station O_3 (1)

- 0.419 x (# of hrs windows are open) + 0.311 x (Minimum outdoor temperature)

- 5.600 x (Central Refrigerant recirculating air conditioner use)

+ 0.012 x Station O_3 x (# of hrs windows are open)

+ 0.023 x Minimum outdoor temperature x (# of hrs windows are open)

The selected model for predicting indoor levels of O_3 suggests that the level of indoor O_3 observed in these study homes was largely determined by the product of the outdoor O_3 level and the duration of time that any windows were open (since this two-variable model alone accounted for an R^2 of 0.38). The main effect of the station level of O_3 indicates that if the home's windows were kept shut, a 1 ppb increase in the outdoor O_3 level would tend to increase the indoor O_3 level by 0.053 ppb (an amount not statistically significantly different from 0). Other things being equal, indoor O_3 levels were higher when outdoor temperatures were lower. This effect was magnified when windows were open. Finally, use of a central refrigerant recirculating air conditioner was associated with lower indoor levels of O_3 (as one might expect, given the lack of indoor sources and the reactive nature of O_3).

6.1.1.1 Evaluation of O_3 Model Robustness

To assess the utility of the model selected, several aspects of the model-building approach were reviewed. The procedure described above was performed without adjusting for measurements that were below the assigned limit of detection (LOD of 5 ppb O_3). In order to assess the impact of potential adjustment procedures, we refit the model in Table 6-2 using two extreme adjustment procedures. In the first procedure, we replaced each measurement under the assigned limit of detection (5 ppb) by 0, and in the second procedure, we replaced each measurement under the limit of detection (5 ppb) by 5. The results are shown in Table 6-3. The values of R^2 and of the coefficients were similar, whether the measured indoor levels were unadjusted, or replaced by either their lower or upper limits. Thus, the measurement error introduced by the large number of indoor measurements below the limit of detection did not affect model performance, and our assessment of the model performance may be considered

accurate regardless of where in the interval from 0 to 5 ppb the true values lie.

It is well known that when the same data set is used both to select a model and to estimate its goodness of fit, the goodness of fit tends to be overestimated. In order to assess the degree of overestimation in the procedure described above, we repeated the entire model selection procedure on the data set consisting of those homes with odd identification numbers. We then used this model to predict the indoor levels of O_3 for all homes in the study. The mean squared error of prediction for the odd-numbered homes was 86 ppb^2, while that for the even-numbered homes was 88 ppb^2. This indicates that the model will perform nearly as well on a new set of homes as on the homes in our data set. Of course, using the model to predict levels in homes that differ substantially from those in our data set may result in a more substantial degradation of performance.

To assess the advantage of measuring the outdoor O_3 level near the house rather than at the monitoring station, we fit the model presented in Table 6-2 with the ambient O_3 level measured outside the home substituted for the O_3 level observed at the closest monitoring station. The results are presented in Table 6-4. They indicate that when the ambient level of O_3 was measured just outside the house, rather than at the neighborhood monitoring station, the predictive value of the model (as measured by R^2) was only slightly increased.

We also computed correlations between the indoor level of O_3, the level outside the home, and the level at the neighborhood monitoring station. The results are presented in Table 6-5, which shows that the correlation between the indoor level and the level at the monitoring station was 0.49. Correlation between the indoor level and the level just outside the house was 0.58, which might be considered only a minor improvement, given the proximity of the measurement to the indoor reading. The correlation between the level at the monitoring station and the level just outside the house was 0.76.

The value of 0.49 for the correlation between station O_3 and indoor levels indicates that station levels are modestly related to indoor levels. The value of 0.55 for the coefficient of

determination (R^2) in the model in Table 6-2 indicates that knowledge of immediate past household use such as length of time windows were open and whether an air conditioner was used results in a somewhat better prediction than is available from the station level alone. Overall, the information we collected is adequate to enable one to predict indoor levels on a given day with modest accuracy, based on knowledge of that day's household use. Our results also indicate that O_3 levels measured just outside the house are no better for predicting indoor levels than are station ambient levels.

It should be noted that there are many models other than the model selected here that would fit the collected data essentially as well. The model selected here has the virtues of scientific plausibility and simplicity.

6.1.2 Modeling Indoor Levels of PM_{10}

As described in the initial section to this chapter (see Section 6.1), responses to the field technician questionnaire and to the follow-up questionnaire, along with outdoor ambient levels and temperatures, were considered as potential explanatory variables in linear models whose outcome variables were measured indoor levels of PM_{10}. Only measurements made with teflon filters were considered for use in the model. No collected measurements were below the limits of detection (2 $\mu g/m^3$). There were three records in which the station PM_{10} level was recorded as zero; these were not considered in our analyses.

The presence of smokers in the home was expected to be of potential importance in understanding PM levels in the home. However, the skewed distribution of the smoking variable limited our ability to assess its importance. Smoking was reported in 18% of the visits during which indoor PM_{10} was measured. The maximum reported number of cigarettes smoked was 80, but the number exceeded 40 on only three visits. In a preliminary model fit, we determined that the values of estimated coefficients were very sensitive to the number of cigarettes reported smoked in these high-smoking homes. For this reason, we decided to recode this variable as a categorical variable with categories no cigarettes smoked, 1-10 cigarettes smoked, and more than

10 cigarettes smoked.

Variable selection for model building proceeded in a step-wise manner. Variables for which 90% or more of the responses shared a common value, or for which 25% or more of the responses were missing, were eliminated from consideration. Each of the remaining variables was entered into a model whose only other explanatory variable was the ambient level of PM_{10} measured at the neighborhood fixed-site monitoring station. Variables not significant at the 15% level in this univariate setting were dropped from consideration. (A listing of all variables considered and their determined p-values appears in Appendix I).

From those variables remaining, all linear models containing ambient PM_{10} and five or fewer additional explanatory variables were considered. Table 6-6 presents the best fitting models for each number of variables, where fit is measured by percentage of variance explained (multivariate R^2), in an ordinary least squares regression. The R^2 associated with the best sets of variables for the one-variable through six-variable models ranged from 0.13 to 0.52.

The most important factor in determining indoor levels of PM_{10} was smoking, so the relative importance of this variable was considered in the analysis. Table 6-7 presents the result of fitting a model containing only the ambient PM_{10} level and the smoking category. The results show that a model including only the station level of PM_{10} and a categorical variable indicating the number of cigarettes smoked had a coefficient of determination (R^2) of 0.40. There was no significant increase in indoor PM_{10} levels associated with smoking 1-10 cigarettes during the previous 24 hours, but smoking more than 10 cigarettes was significantly associated with an increase of over 67 $\mu g/m^3$.

As the number of variables increases, the maximum attainable R^2 necessarily increases. We chose to work with the five-variable model whose explanatory variables were ambient level of PM_{10}, smoking category, existence of a dirt driveway, number of hours ceiling fan was in use, and an indicator for having been built in the 1950's. This final model is presented in Table 6-8.

Table 6-8 shows that the best five-variable model adds the existence of a dirt driveway, the duration of ceiling fan use, and an indicator for being built in the 1950's. The inclusion of these variables increases R^2 from 0.40 to 0.49, which shows that these variables are much less important than smoking and the ambient level of PM_{10}.

The model presented in Table 6-8 was based on only 78 home visits. Because this number is small, because smoking occurred in so few of these visits, and because smoking is by far the strongest factor in the level of indoor PM_{10}, the significance of the other variables in the model is likely to be due to artifactual correlations with smoking. To check this, we fit the model in Table 6-8 to those homes in which no smoking had occurred. Table 6-9 shows the results.

When the analysis was restricted to non-smoking residences, the existence of a dirt driveway and the age of the house were no longer significantly related to the indoor level of PM_{10}. In addition, the goodness of fit of the model, as measured by R^2, was only 0.15. It follows that the apparent significance of these variables could be an artifact of the small number of residences in which PM_{10} was measured, and the small percentage of those in which smoking occurred. For this reason, we have not formed conclusions about which factors other than smoking and outdoor levels affect indoor levels of PM_{10} in residences where smoking occurs.

We repeated our model selection procedure, using only those visits for which no smoking was reported. The results are shown in Table 6-10. The corresponding prediction equation is:

$$\begin{aligned}\text{Indoor } PM_{10} = {} & 20.740 + 0.402 \times \text{Station } PM_{10} \\ & + 15.284 \times \text{Odor of mold} - 15.235 \times \text{other odor} \\ & - 0.688 \times \text{number of hours of ceiling fan use} + 8.340 \times \text{cooking} \\ & - 12.255 \times \text{other smoke} + 11.785 \times \text{two bedroom house} \\ & + 2.842 \times \text{central refrigerating recirculating air conditioner} \\ & - 0.251 \times \text{station } PM_{10} \\ & \quad \times \text{central refrigerating recirculating air conditioner}\end{aligned} \qquad (2)$$

where "other odor" refers to odors other than those of mold, cigarettes, formaldehyde, or animals.

The explanatory logic for the signs associated with the variables in the above model are not altogether clear. For some of the variables (such as odor of mold, hours of ceiling fan use, and cooking), the positive value has some reasonable meaning (i.e., stronger mold odor may have to do with more airborne material present, or more hours of ceiling fan use may resuspend or deposit particles). The implication of other variable signs, such as a negative sign associated with other smoke in the house, is not immediately obvious or intuitive. The models developed here were based on a limited number of PM sampling performed in a restricted number of study homes. In developing models using a limited amount of observations and a wide array of variables, it should be noted that individual sample results may wield significant and undue influence over the model variable selection process.

In the presented model, the coefficient of "central refrigerating recirculating air conditioner" is meaningless by itself, since it refers to the effect on a hypothetical day where station PM_{10} is 0 $\mu g/m^3$. However, the coefficients of station PM_{10} and the interaction term appearing in the model are of interest. Together, they suggest that the effect of outdoor PM_{10} on indoor PM_{10} is less in homes with a central refrigerant recirculating air conditioner. This relationship might be explainable if central air conditioning units were somewhat effective in reducing indoor PM_{10} levels through inertial impaction, or some other mechanical means, compared to other homes and systems.

To assess the advantage of measuring the outdoor PM_{10} level near the house rather than at the monitoring station, we fit the model in Table 6-10 with the ambient PM_{10} level measured outside the home in place of the PM_{10} level at the monitoring station. The results are presented in Table 6-11. When the ambient level of PM_{10} was measured just outside the house, rather than at the neighborhood monitoring station, the predictive value of the model, as measured by R^2, was decreased (0.47 compared to 0.55).

We also computed correlations between the indoor level of PM_{10}, the level outside the home, and the level at the monitoring station. These results are presented in Table 6-12. The correlation between the indoor level and the level at the monitoring station was 0.355, while the correlation between the indoor level and the level just outside the house was 0.356. The correlation between the level at the monitoring station and the level just outside the house was 0.634. These results clearly indicate that PM_{10} levels measured just outside the house are no better than station levels for predicting indoor levels.

6.1.3 Modeling Indoor Levels of $PM_{2.5}$

Responses to the field technician questionnaire and to the follow-up questionnaire, along with outdoor ambient levels and temperatures, were considered as potential explanatory variables in linear models whose outcome variables were measured indoor levels of $PM_{2.5}$. Only measurements made with Teflon filters were considered. No measurements were below the limits of detection (2 $\mu g/m^3$).

For reasons described above in the section on PM_{10}, we believe that models for predicting indoor levels of $PM_{2.5}$ in smoking homes are likely to differ considerably from models for predicting indoor $PM_{2.5}$ in non-smoking homes. Since there were only 17 visits in which indoor $PM_{2.5}$ was measured with a Teflon filter during which smoking occurred, we decided to restrict our modeling to non-smoking homes.

Variable selection proceeded as follows: Variables for which 90% or more of the responses shared a common value, or for which 25% or more of the responses were missing, were eliminated from consideration. Each of the remaining variables was entered into a model whose only other explanatory variable was the ambient level of $PM_{2.5}$ measured at the monitoring station. Variables not significant at the 15% level in this univariate setting were dropped from consideration. (A listing of all variables considered and their determined p-values appears in Appendix I).

From those variables remaining, all linear models containing ambient $PM_{2.5}$ and five or fewer additional explanatory variables were considered. Table 6-13 presents the best fitting models for each number of variables, where fit is measured by percentage of variance explained (multivariate R^2), in an ordinary least squares regression. R^2 ranged from 0.46 to 0.69. As the number of variables increases, the maximum attainable R^2 necessarily increases. Table 6-13 shows that including more than three variables does not substantially improve the model fit (since the three-variable model had an R^2 of 0.64, and the R^2 associated with the four, five and six-variable models were 0.66, 0.68, and 0.69, respectively).

We therefore chose to work with the three-variable model whose explanatory variables were ambient level of $PM_{2.5}$, number of minutes a stove was in use, and an indicator for being a two-bedroom residence. This final model is presented in Table 6-14 ($R^2 = 0.61$).

The corresponding prediction equation is:

$$\begin{aligned} \text{Indoor } PM_{2.5} = & -3.893 + 0.536 \times (\text{Station } PM_{2.5}) \\ & + 13.316 \times (\text{Two-bedroom residence}) \\ & + 0.164 \times (\text{Number of minutes of stove use}) \end{aligned} \qquad (3)$$

where the Station $PM_{2.5}$ is a derived value, based on measured PM_{10} (no hourly or daily $PM_{2.5}$ measurements were made at community stations).

To assess the advantage of measuring the outdoor $PM_{2.5}$ level near the house rather than at the monitoring station, we fit the model in Table 6-14 with the ambient $PM_{2.5}$ level measured directly outside the home in place of the calculated level (based on available PM_{10} data) at the monitoring station. The results are presented in Table 6-15. We also computed correlations between the indoor level of $PM_{2.5}$, the level outside the home, and the level at the monitoring station. The results are presented in Table 6-16.

The value of 0.64 in Table 6-13 for the correlation between station $PM_{2.5}$ and indoor levels indicates that station levels were reasonably well related to indoor levels. However, the correlation between station PM_{10} and the indoor level of PM_{10} in non-smoking homes was only 0.29. This indicates that, at least in non-smoking residences, ambient levels of $PM_{2.5}$ were much more predictive of indoor levels than were ambient levels of PM_{10}.

The best model we could find for estimating indoor levels of $PM_{2.5}$ included the duration of stove use and an indicator for a two-bedroom residence. Table 6-13 shows that the inclusion of these additional explanatory variables increases the coefficient of determination (R^2) from 0.46 to 0.64. However, the model selection was based on only 47 visits in which $PM_{2.5}$ was measured in non-smoking homes. Because of the small sample size, the model selection procedure may be somewhat unstable, and the coefficient of determination may be somewhat overestimated.

The results in Table 6-15 indicate that when the ambient level of $PM_{2.5}$ was measured just outside the house rather than at a monitoring station, the predictive value of the model, as measured by R^2, only improved from 0.61 to 0.67 (a marginal improvement).

We also computed correlations between the indoor level of $PM_{2.5}$, the level outside the home, and the level at the monitoring station. Table 6-16 shows that the correlation between the indoor level and the level at the monitoring station was 0.64, while the correlation between the indoor level and the level just outside the house was 0.59. The indoor levels thus seem to be slightly better related to the station levels than to the levels just outside, but this may well be due to sampling error, since sample sizes were small. The correlation between the level at the monitoring station and the level just outside the house was 0.72. These results clearly indicate that $PM_{2.5}$ levels measured just outside the house are no better than station levels for predicting indoor levels.

6.1.4 Effect of Season on Indoor Pollutants

We studied the effect of season (apart from temperature, which was already included in some of the previously considered models) on the indoor levels of ozone, PM_{10}, and $PM_{2.5}$ by

fitting the models presented in Tables 6-2, 6-10, and 6-14, respectively. We defined the period before May 16 as "spring", the period between May 16 and September 30 (inclusive) as"summer", and the period after September 30 as "fall".

The effect of season was statistically insignificant in each case, with p-values ranging from 0.26 to 0.89. From this, we conclude that season, apart from temperature, did not appear to be related to the indoor levels of ozone, PM_{10}, and $PM_{2.5}$ observed in this study. However, it is important to note that PM sampling data (using teflon filters) was only available from this study for the time period late June through November, so that PM data could only be analyzed for the summer and fall periods as defined above.

6.1.5 References

Freedman DA, WC Navidi, SC Peters, On the impact of variable selection in fitting regression equations, In Model Uncertainty and its Statistical Implications, TK Dykstra, Editor, Springer-Verlag, Berlin, pp1-16 (1988).

Lysikowski J, Bootstrapping Variable Selection Procedures in Linear Models, PhD Thesis, Division of Biostatistics, University of Southern California (1995).

Mallows CL, "Some comments on C_p," Technometrics 15:661-675, (1973).

Miller AJ, Subset Selection in Regression, Chapman and Hill, New York (1990).

6.2 Predicting Residential Exposures From Community Monitors

6.2.1 Introduction

Residents spend a large amount of time in and near their homes. As a result, the concentrations of ozone and PM outside their homes are expected to have a large influence on their total exposure to these pollutants. In the absence of significant indoor sources, outdoor concentrations may have a large influence on indoor pollutant concentrations. For exposure assessment, it is important to accurately characterize outdoor concentrations near residences. However, measurements of these concentrations cannot be made for all subjects during all times in large-scale epidemiologic studies. Commonly, the ambient concentrations at the nearest community monitoring station are used as surrogates for residential outdoor concentrations in epidemiologic studies. The residential data collected in this study provided an opportunity to evaluate this approach and to develop an improved understanding of within-community variations in ambient concentrations.

Comparisons of 24hr average outdoor residential concentrations of ozone, PM_{10}, and $PM_{2.5}$ measured at study homes were made with ambient data collected from nearby community monitoring stations. Residential $PM_{2.5}$ measurements were compared to calculated $PM_{2.5}$ estimates, since $PM_{2.5}$ was not routinely measured at any of the community monitoring sites. Because the community monitoring data was measured hourly for both ozone and PM_{10}, but the residential monitoring data was collected over an integrated 24hr sampling period, the hourly community data was averaged over the time period of the residential measurements.

6.2.2 The Community Monitoring Database

During the residential study, outdoor concentrations were measured at community monitoring stations operated by the regional regulatory monitoring agency (SCAQMD) or a project subcontractor (STI) in the four communities included in the study. In San Dimas, ozone and PM_{10} were measured hourly at the San Dimas monitoring station, and ozone was measured hourly at the nearby SCAQMD Glendora and Azusa monitoring stations. In Riverside, measurements of both hourly ozone and PM_{10} were made at the UC Riverside, Mira Loma, and Rubidoux monitoring sites. In Lancaster, hourly ozone and PM_{10} were measured at only one community

monitoring station; no other community monitors were in close proximity. In Lake Gregory, ozone data were collected at two community monitoring stations (the Lake Arrowhead and Lake Gregory stations), while PM_{10} measurements were made at one site (Lake Arrowhead).

The distances between the community monitoring stations within each community are shown in Table 6-17. The within-community distance between community-station-based monitors ranged from 4 to 18 km. The correlation between ambient concentrations measured at community monitors within the same communities are also shown in Table 6-17. Ozone data measured at stations within the San Dimas and Riverside communities were well correlated ($0.84 \leq R^2 \leq 0.94$). Ozone data collected in Lake Gregory ($R^2 = 0.50$) and PM_{10} data collected at Mira Loma and Rubidoux ($R^2 = 0.63$) were moderately well correlated. PM_{10} data collected at UC Riverside were not well correlated ($R^2 \approx 0.2$) with those collected at Mira Loma or Rubidoux. (The explanation for the poor correlation in PM measurements in the Riverside area stations is currently unknown; TEOM data from the UC Riverside monitoring station are under review, due to questions about instrument performance and lack of station comparability with other data; this data review does not affect the modeling results reported in this project, since, as the following sections explain, data from this site was not selected for use in subsequent modeling activities).

Residential PM data was collected on a 24hr basis using 37mm filter samplers. The hourly PM_{10} measurements obtained at the community monitoring stations were made with Tapered Element Oscillating Microbalances (TEOMs), which do not retain volatile PM_{10} constituents well (since the TEOM monitors volatilize ammonium nitrate and some organic material when the sample is heated to 50°C, as part of the instrument's automated sampling protocol). Data from collocated TEOM and residential PM_{10} samplers were not collected in this study, which made direct comparison difficult. Information on potential sampler bias from previous studies was reviewed in order to select adjustment procedures for use in this project.

In a previous study (Thomas et al 1993), the residential PM_{10} sampler used in the current study was directly compared to High Volume Air (HiVol) and dichotomous sampler performance using collocated measurements. The researchers found that the residential PM_{10} sampler was

biased high compared to both the HiVol and the dichotomous samplers. For $PM_{2.5}$, the residential sampler was biased high compared to the dichotomous sampler (Thomas et al 1993). In contrast, the TEOM PM_{10} monitors are known to be biased low compared to HiVol samplers (Allen et al 1995; Peters 1995).

Based on the limited sampler inter-comparison data available, adjustments were made to the residential study data for the purposes of data analysis and model development. Site-specific regression equations were developed to adjust TEOM data to simulate HiVol data (Peters 1995). The TEOM adjustment equations were based on regressions of collocated or nearby TEOM monitors and HiVol samplers (such as at nearby stations in Rubidoux, Glendora, and Lake Arrowhead). Since no other community station was close enough to compare with the Lancaster site, the regression equation for Lancaster was developed using collocated HiVol and TEOM measurements at another community monitoring location (Atascadero) with similar PM chemistry. (Atascadero is one of the twelve Children's Health Study communities [ARB Contract #A033-186], but not a residential study community; the monitoring instrumentation and operating protocol at the Atascadero site were similar to and in conformance with the other community monitoring locations from which data was used in the analyses discussed).

The adjustments in residential PM_{10} concentrations were made to normalize residential sampler data to the HiVol data for regression analyses and data presentations. Thomas et al reported regressions of collocated sampling between the HiVol and the residential sampler for PM_{10}, between the HiVol and the dichotomous sampler for PM_{10}, and between the residential sampler and the dichotomous sampler for $PM_{2.5}$. A regression based on the slope and intercept approach of Thomas et al was used to develop the adjustments. This approach was chosen based on simplicity and the observation that its use does not result in negative values at low concentrations (as did some other regression approaches). No adjustment was made in the data set for variable wind conditions.

For PM_{10}, the slope for the regression of the residential sampler (dependent variable) and the HiVol (independent variable) was 1.25. Thus, the adjustment for PM_{10} is 0.80 (1/1.25). For

$PM_{2.5}$, there was no direct correlation between the HiVol and the residential sampler, because the HiVol does not measure $PM_{2.5}$. However, it is still desirable to develop a factor that normalizes the $PM_{2.5}$ concentrations to a commensurate level for the HiVol PM_{10} measurement. For $PM_{2.5}$, the slope of the regression between the residential sampler (dependent variable) and the dichotomous sampler (independent variable) was 1.2. Thus, to normalize the $PM_{2.5}$ concentrations to the dichotomous sampler, the concentrations were multiplied by 0.83 (the result of 1/1.2). To normalize the dichotomous sampler concentration to the HiVol, the regression between the HiVol and the dichotomous sampler for PM_{10} was used. The slope for the regression between the HiVol (dependent variable) and the dichotomous sampler (independent variable) was 0.93. Thus, the normalization factor for $PM_{2.5}$ was 0.77 (the product of 0.83 and 0.93). In summary, this study included no collocated comparison sampling of the residential sampler and other types of particle samplers (such as HiVol, TEOM, or dichotomous samplers). Adjustment factors were developed for this analysis from other previous work and may not be completely applicable.

Daily $PM_{2.5}$ concentrations were not measured at the community monitoring stations during the residential study; only two-week average $PM_{2.5}$ concentrations were measured at the community monitoring stations, using the Two-Week Sampler. In order to obtain an approximate comparison of residential and community $PM_{2.5}$ concentrations, $PM_{2.5}$ concentrations at the community monitors were estimated from the TEOM and two-week sampler $PM_{2.5}$ data. Estimates of $PM_{2.5}$ concentrations on specific days were calculated by applying the appropriate quarterly average ratio of $PM_{2.5}$ to adjusted TEOM PM_{10} at each station to the adjusted TEOM PM_{10} for that day. The quarterly average $PM_{2.5}$ to PM_{10} ratios were based on the 1994 data from San Dimas, Mira Loma, UC Riverside, Lancaster, and Lake Arrowhead stations for the following periods: (1) February-April, (2) May-July, (3) August-October, and (4) November-December. These periods were thought to roughly correspond to periods of similar PM chemistry in Southern California. The $PM_{2.5}/PM_{10}$ ratios used in the analysis are shown in Table 6-18.

6.2.3 Correlation of Outdoor Concentrations Measured at Community Monitors and Residences

An analysis was performed to determine the correlation between the concentrations measured by the community monitors and the residential samplers. Because there was more than one community monitor in several communities, an analysis was performed to determine which of the monitors in proximity to the residences best correlated with the residential concentrations. Residential concentrations were compared to data from each community monitor and to spatially interpolated data. When data were available from more than one community monitor in a community, the data were spatially interpolated to the location of the residence using an inverse distance squared weighting procedure.

The results of the analysis are shown in Tables 6-19 through 6-21. Scatter plots of residential and community monitor concentrations of ozone, PM_{10} and $PM_{2.5}$ are shown in Figures 6-1 through 6-12. For PM_{10} and $PM_{2.5}$, these figures show the adjusted TEOM and residential sampler concentrations.

6.2.3.1 Ozone

As shown in Table 6-19, the residential and community monitoring data for ozone agreed fairly well in San Dimas, Riverside, and Lancaster. The correlation coefficients (R^2) were above 0.60, and mean 24-hr average concentrations were within 5 ppb in all three communities. The residential and community monitoring ozone data in Lake Gregory did not agree as well. Perfect agreement was not expected between residential and community monitoring data, since ozone is known to have considerable spatial variability within relatively small areas as a result of NO titration and dry deposition. Ambient monitors are cited so as to minimize the effects of these processes, but the location of the sampling homes was beyond the control of study investigators. In addition, previous evaluations have shown that imprecisions of ±5 and ±12 percent are associated with community and residential monitoring devices, respectively (Lurmann et al 1994). A six percent bias between the sampling devices (with ozone concentrations derived from TED filter samples averaging about six percent higher than collocated UV photometer measurements) has been previously reported (Lurmann et al 1994).

The mean residential ozone concentration observed in San Dimas was slightly lower than the mean concentrations observed at the San Dimas, Glendora, and Azusa community monitors (28 ppb compared to 34, 36, and 29 ppb, respectively). This may be a result of NO titration at the residences. The correlation coefficients (R^2) between the individual community monitors' ozone data and the residential data were 0.79, 0.78, and 0.80 for San Dimas, Glendora, and Azusa, respectively. The mean concentration (28 ppb) and correlation of the interpolated community monitoring data ($R^2 = 0.82$) were comparable to those determined from the individual community monitors' data in San Dimas. Figure 6-1 shows that almost all of the residential ozone concentrations were within 10 ppb of the San Dimas monitor's data, and the slope of the regression line was close to unity.

The residential ozone data in Riverside was in reasonable agreement with the community monitoring data. The mean residential ozone concentration in Riverside (31 ppb) was lower than both the mean interpolated concentration of 36 ppb and the individual monitor mean concentrations of 38 ppb at the UC Riverside and Rubidoux stations, but comparable to the mean reported Mira Loma station value (31 ppb). Again, the generally lower residential concentrations may be a result of NO titration of ozone at the residences. The correlation coefficients (R^2) between the individual community monitors' ozone data and the residential data ranged from 0.65 to 0.75. The scatter plot of the residential and Rubidoux ozone data, shown in Figure 6-2, showed good correlation and few outliers. This suggests that even though Riverside is a large community, the community monitoring data for ozone were representative of levels observed outside Riverside residences.

The ozone concentrations measured near residences in Lancaster (43 ppb) were comparable to ozone concentrations observed at the Lancaster station (42 ppb). The correlation was 0.60, which was lower than the correlation in Riverside and San Dimas, yet still indicative of a moderate relationship. However, as shown in Figure 6-3, there were at least four cases with extremely poor agreement (e.g., 5 ppb at a home versus 66 ppb at the station, and 92 ppb at another home versus 53 ppb at the station), which clearly reduced the correlation.

The mean ozone concentration measured outside residences in Lake Gregory was 46 ppb, which was slightly lower than the data from the Lake Gregory monitor (53 ppb) and substantially lower than the data from the Lake Arrowhead monitor (67 ppb). The correlation between the residential ozone and the community monitors' ozone was 0.34 and 0.43 for Lake Arrowhead and Lake Gregory, respectively, which was low. The ambient data at the two sites were moderately correlated (R^2 = 0.50). Figure 6-4 shows there was substantial scatter in the residential and community monitor ozone data in the Lake Gregory study area. The presence of a significant forest tree canopy in and around homes in this community (and around the Lake Gregory station itself) may partially explain the significant within-community ozone variation relative to other communities.

6.2.3.2 PM_{10} and $PM_{2.5}$

The residential and community comparisons for PM_{10} and $PM_{2.5}$ were not as closely correlated as those for ozone. However, like those for ozone, the comparisons were better in San Dimas and Riverside than in Lancaster and Lake Gregory (suggesting that community monitors are located at more representative locations in San Dimas and Riverside than in Lancaster and Lake Gregory). Figures 6-5 through 6-12 show considerable scatter in the $PM_{2.5}$ and PM_{10} data. As shown in Table 6-20, the unadjusted mean residential PM_{10} concentrations were significantly lower (23 to 45 percent lower) than the adjusted TEOM community monitor PM_{10} in San Dimas, Riverside, and Lake Gregory. The adjusted residential PM_{10} concentrations were 38 to 56 percent lower. In Lancaster,the mean unadjusted residential PM_{10} concentration was only 11 percent lower than the community monitor PM_{10}, and the adjusted concentration was 29 percent lower. The correlation coefficients (R^2) between the residential and community monitor PM_{10} data were 0.43 in San Dimas and from 0.35 to 0.49 in Riverside, which indicate a moderate degree of correlation. The interpolated PM_{10} data in Riverside showed the best correlation with the residential data (R^2 = 0.56), and this correlation value reflected only a moderate relationship. The residential PM_{10} data collected in Lancaster and Lake Gregory had little correlation with the community monitor's PM_{10} (R^2 = 0.13 to 0.20), but this may be due to a number of factors, including an insufficient range of PM levels observed in community monitoring data or inappropriate siting of the community monitor.

Comparisons of the residential $PM_{2.5}$ concentrations to the approximate $PM_{2.5}$ concentrations at the community monitor were similar to those for PM_{10}. The mean unadjusted residential $PM_{2.5}$ concentrations were 5 to 23 percent lower than the estimated community monitor $PM_{2.5}$ levels, and the adjusted residential concentrations were 27 to 41 percent lower. The residential $PM_{2.5}$ data showed moderate correlation ($R^2 = 0.36$ to 0.54) with the approximate $PM_{2.5}$ concentrations at the San Dimas, Mira Loma, Rubidoux, and Lake Arrowhead monitors. Little correlation (R^2 ranging from less than 0.01 to 0.07) was found between the residential $PM_{2.5}$ data and the approximate $PM_{2.5}$ concentrations at the Lancaster and UC Riverside monitors. These values are similar to the PM_{10} comparison, which was not surprising, since the $PM_{2.5}$ ambient concentrations were estimated from the PM_{10} concentrations. Reasons for the observed lack of correlation may include location of the Lancaster monitor in a potentially non-representative area (a dusty industrial area) or data of questionable validity from the UC Riverside TEOM monitor.

Differences between the residential study PM sampler and the TEOM monitor probably contributed to the moderate to low correlation for PM in all of the communities. In particular, the amount of material volatilized by the TEOM is expected to vary with meteorological conditions and is not fully accounted for by the TEOM adjustment procedures.

However, while the TEOM monitor characteristics may partly explain the low correlations, it does not explain why the residential concentrations were consistently lower than the community monitors. This may be a result of differences in the siting criteria for PM monitors at residences and community monitors, or perhaps an inherent bias in the residential PM sampler relative to TEOM or HiVol samplers. For example, the ammonium nitrate volatilization losses from the residential sampler are not adequately characterized. Also, at residences, the PM samplers were usually placed in the back of the house about one meter above the ground. Some were placed a few meters from the house, while others were placed on back porches or patios. This siting criteria was considered representative of the actual human exposure outside the residence. However, the influence of building wake and vegetation effects, near and around residences, may result in differences in PM concentrations between residences in the same community.

Conversely, community monitors are typically cited so that the sampling inlet is away from vegetation and above the top of the shelter (several meters above the ground) so that the sampling inlet has generally unobstructed airflow. These siting differences probably contributed significantly to the PM concentration differences observed as well as to lower correlation between station and residential data, than those observed for ozone. Future studies need to examine in more detail the difference between residential exposure measurements and community monitor measurements.

6.2.3.3 Interpolated Concentrations

The interpolated concentrations did not correlate significantly better than those from the highest correlated site for any of the pollutants in any of the communities. In most communities with multiple sites, the correlations were similar for each of the individual sites. One exception was for Lake Gregory. For this area, the Lake Gregory site had a considerably better correlation coefficient than the Lake Arrowhead site for ozone. The Lake Arrowhead ozone concentration was much higher than the ozone at the residences, despite the fact that most of the residences were closer to the Lake Arrowhead site. For the other sites, the mean ozone concentrations at the residences and the community monitors were similar. The Riverside community actually had a higher interpolated PM_{10} correlation than for any of the individual sites, but not by a wide margin.

One reason why the interpolation did not considerably improve the correlation is because the ambient sites within communities generally correlated well (see Table 6-17). For ozone, the exception was Lake Gregory, where the Lake Arrowhead and Lake Gregory sites only had an R^2=0.47. This moderate correlation suggested that there was high spatial variance of ozone in the area, which would explain the low correlation between the residential and ambient sites. For PM_{10}, the Mira Loma and Rubidoux sites correlated fairly well (R^2=0.69), but the nearby UC Riverside site did not correlate well with either of these sites (perhaps because of the suspect TEOM data). This may explain the lower correlation of the UC Riverside site with the

residential concentrations.

For locations with multiple monitors, the best overall individual community monitor (the one with the highest correlation combined with the closest mean concentration compared to the residences) was selected for exposure assessment purposes. For ozone, the best overall monitoring sites in each of the four study communities were San Dimas, UC Riverside, and Lake Gregory (Lancaster only had one site to choose from). For PM, the best overall monitoring site in the Riverside/Mira Loma area was Rubidoux.

6.2.4 Investigation of Factors Influencing Ambient and Residential Differences

The residential questionnaire was examined to determine if any of the collected information was potentially useful in explaining the observed variance in near-residential outdoor concentrations. For ozone, the second question in the Technician Walk-Through Survey was deemed potentially useful. This question asked, "Is the house located within 100 yards of a busy roadway?" This could be an indicator of the effect of NO titration at the residences. For PM, the third question in the technician Walk-Through Survey was deemed potentially useful. This question asked, "Are any of the following sources of exposed dirt located within 100 yards of this house?: (a) Dirt drive (e.g., dirtway or road), (b) Other (specify)."

Two tests were performed to determine if these variables increased the predictability of the model. The first test was a simple comparison of mean outdoor concentrations for those residences answering yes or no to the question. For both questions, the comparison showed that the means were not statistically different. For the second test, a new variable for each question was defined, where each residence was assigned a "0" or "1" based on their response to the questions. This variable was then added to the regression models described above. The covariates of each of these new variables were not statistically significant, indicating that they did not increase the correlation coefficient. Based on these analyses, it was determined that these particular questions do not improve the predictive power of the model.

6.2.5 Implications for Exposure Assessment

To analyze how well ambient monitoring station data represented the pollutant concentrations immediately outdoor of the study residences, histograms of the differences and percent differences between the two measurements (ambient minus residential) were plotted for each of the pollutants. For the percent difference plots, only ozone concentrations above 20 ppb and PM concentrations above 20 $\mu g/m^3$ were included to lessen the effect of the high estimation errors associated with measurements of low concentrations. Figures 6-13 through 6-18 show the histograms for each of the pollutants. For ozone, the distribution was biased towards higher community monitor concentrations, with 70 percent of the residuals greater than zero. This observed bias is probably the result of greater ozone titration (by NO) and dry deposition near residences (more vegetation) than at community monitoring stations. The percentage difference histogram shows considerably more variance, with only 59 percent of the values within ±20 percent. Still, ozone residential measurements agreed sufficiently well with the ambient station measurements to be useful for exposure assessment purposes. For PM_{10} and $PM_{2.5}$ there was a large positive bias, with 92 and 90 percent of the residuals exceeding zero, respectively. These differences, as mentioned earlier, were probably a result of sampler differences or different siting criteria for the residential and community monitor measurements. The bias and moderate to poor correlation in these PM data illustrates potential difficulties in using community monitoring data for PM exposure assessment purposes.

6.2.6 Conclusions

Several conclusions can be drawn from the above comparison of outdoor residential concentrations of ozone, PM_{10}, and $PM_{2.5}$ with community monitor concentrations. First, ozone outdoor residential concentrations correlated reasonably well with regional monitoring data, but had a lower mean concentration compared to the corresponding community monitor concentrations. The lower residential concentrations may be a result of NO titration and/or dry deposition. Nonetheless, the high correlation suggests that the community monitors can be used to predict outdoor levels of ozone immediately around residences with some confidence.

The results for PM, are less encouraging than those for ozone. Outdoor residential PM concentrations did not correlate as well as those for ozone. PM had a consistently lower mean concentration immediately outside the study homes compared to the corresponding community monitor concentrations. The observation of lower mean residential PM concentration, compared to regional PM data, may be due to differences in PM monitor siting criteria at the two types of sampling locations, to systematic differences between the samplers (the analyses were based on data generated at the community stations by TEOMs, and filter-based information collected at the homes on single-stage impactors driven by personal sampling pumps), or to some other unidentified consideration. These differences need to be investigated before the data can be used for exposure assessment purposes.

Other attempts to improve the observed correlation between home measurements and community monitors, by interpolation using nearby community station data or through the use of questionnaire-collected data about housing factors and home operation, did not result in any significant improvement. Ambient sites within study communities typically agreed well with each other, so interpolated measurements across them were not very different from individual site observations. Most of the survey questions completed by residents and observations made by field personnel were directed towards characterization of housing factors of potential importance for understanding indoor pollutant levels. In that regard, the importance of any housing factor would likely be muted by the strength of the housing variable in predicting the indoor/outdoor pollutant ratio, which might link the in-home pollutant level to pollutant observations made at the community monitoring station.

Therefore, knowledge of regional ozone concentrations may be of value in predicting residential exposure to ozone, but several complicating factors currently preclude useful prediction of PM levels outside residences on the basis of information collected at regional or community monitoring stations.

6.2.7 References

G Allen, P Koutrakis, R Reiss, FW Lurmann, PT Roberts, R Burton,W Wilson, "Evaluation of the TEOM method for continuous measurement of ambient particle mass in large urban areas," Paper presented at the Air & Waste Management Association Conference on Particle Matter: Health and Regulatory Issues, Pittsburgh, PA,April 4-6, (1995).

FW Lurmann, PT Roberts, HH Main, SV Hering, EL Avol, SD Colome, Phase II report Appendix A: exposure assessment methodology. Final report prepared for the California Air Resources Board,Sacramento CA, by Sonoma Technology, Inc., Santa Rosa CA, Aerosol Dynamics, Inc., Berkeley CA, USC School of Medicine, Los Angeles CA, and Integrated Environmental Services, Irvine CA, STI-92340-1423-FR, (October 1994).

J.M. Peters, "Epidemiologic Investigation to Identify Chronic Health Effects of Ambient Air Pollutants in Southern California: Phase II Draft Final Report". California Air Resources Board Contract #A033-186, Submitted May 1, 1995.

KW Thomas, ED Pellizzari, CA Clayton, DA Whitaker, RC Shores, J Spengler, H Ozkaynak, SE Froehlich, LA Wallace, "Particle Total Exposure Assessment Methodology (PTEAM) 1990 Study: Method Performance and Data Quality for Personal, Indoor, and Outdoor Monitoring," J. Expos. Analysis Environ. Epidem., 3, 203-226,(1993).•

6.3 Improvements in Exposure Modeling

6.3.1 Introduction

All modern exposure models use the microenvironmental approach where the time integrated exposure is estimated as the sum of the exposures in each microenvironment occupied by the individual for the time period of interest (Sexton and Ryan, 1988). The generalized exposure equation may be written as follows:

$$E_{ij} = \sum_{m=1}^{M} C_{ijm} \, \Delta t_{ijm} \qquad (1)$$

where:

E_{ij} = integrated exposure of the ith individual in the jth community

C_{ijm} = concentration in the mth microenvironment when it is occupied by the ith individual in the jth community

Δt_{ijm} = amount of time the ith individual in the jth community spent in the mth microenvironment

In order to express exposure in familiar (engineering) units of concentration, rather than in units of concentration-time, the time-weighted-exposure (*TWE*) is commonly used. The TWE is calculated from

$$TWE_{ij} = \frac{1}{\sum_{m=1}^{M} \Delta t_{ijm}} \sum_{m=1}^{M} C_{ijm} \, \Delta t_{ijm} \qquad (2)$$

This approach is used in the Regional Human Exposure (REHEX) model (Lurmann et al., 1989; Lurmann and Korc, 1994) and the NAAQS Exposure Model (McCurdy and Johnson, 1989), which are designed to estimate the distribution of exposures to urban air pollutants for the general population.

The central issue in exposure modeling is not the structure of the model (because they are almost all the same), but rather how to estimate the model inputs from available data. The two key elements of exposure modeling are the methodologies for the assignment of people to microenvironments and the estimation of microenvironmental concentrations. Generally, the methodologies involve extrapolating from small data sets to characterize time-activity and microenvironmental concentrations for large numbers of people, hours, and locations.

One of the objectives of the residential study was to improve the REHEX exposure model, through analysis of the residential data and through model evaluation using personal monitoring data. Clearly, the analysis of residential indoor concentrations of ozone and PM_{10}, presented in Section 6.1, provides an improved methodology for estimating indoor concentrations, which should improve the REHEX model's estimates of personal exposure. However, the study's ability to evaluate predictions of personal exposure was greatly reduced when it was determined that personal ozone monitoring data was could not be collected in the performance of this study, due to the lack of an acceptable validated personal ozone sampler. Nevertheless, considerable work was performed during the cross-sectional portion (Phase II) of the Southern California Children's Health Study (CARB Contract #A033-186, abbreviated herein as CHS) to adapt the REHEX model for generalized use in epidemiologic studies and to specifically use the data from the CHS project. The structure of the model was not modified significantly, but a means of providing data for individual subjects was implemented. The new features provide the capability to estimate personal exposures using time activity and microenvironmental factors specific to the individual participants in the epidemiological study. The residential data collected in the current project are useful for improving the estimates of indoor concentrations at residences. This new version is referred to as REHEX-III. The microenvironment assignment procedures and microenvironmental concentration estimation procedures used in REHEX-III are described below.

6.3.2 REHEX-III Improvements

The REHEX-III model was applied to estimate exposure to ozone and nitrogen dioxide for 3000 students for an entire year using community ambient air monitoring data, individual student time-activity survey data, individual school indoor/outdoor ozone ratios, and housing information for the individual students (Peters et al., 1995). The model was not applied to particulate matter in this phase of the Children's Health Study. (Personal exposure to particulate matter remains an important but poorly understood exposure variable).

6.3.2.1 Assignment of Microenvironments

The initial assignment of microenvironments was based on student time-activity questionnaire data collected in 1993. The questionnaire asked parents of students to estimate the amount of time their students spent outdoors for various periods during the day for non-summer weekdays (i.e., school days), non-summer weekends, and summer days. This division of types of days was used for the modeling.

Our knowledge of specific school schedules and hourly diary information collected on a small subset of students allowed us to consider three microenvironments: (1) indoors at home, (2) indoors at school, and (3) outdoors. The 1993 questionnaire retrospectively asked for best estimates of temporal and spatial information for the weekdays, weekends, and summer days for the last 2 weeks for the non-summer period and for typical 2-week period during the summer. Students' responses were expressed as frequencies. For example, the students reported the number of days (e.g., 0 to 10 weekdays in the 2-week period) that they were outdoors for a particular time period. A probabilistic methodology was developed to estimate microenvironmental assignments for each day of the year based on the survey data and numerous assumptions for use in the modeling effort. The questionnaire data only covered selected hours of the day; assumptions were necessary to assign microenvironments for the non-surveyed hours of the day. When the survey data spanned multiple hours, the microenvironmental assignments for specific time increments were randomly assigned. The procedures and principal assumptions for the microenvironmental assignments are summarized below:

Non-Summer Weekday

- (i) Midnight to 8:00 a.m. and (ii) 6:00 p.m. to midnight hours were assigned as "indoors at home";
- 8:00 a.m. to 2:00 p.m. hours were assigned as "indoors at school";
- Two randomly chosen hours between 6:00 a.m. and 2:00 p.m. were assigned as "outdoors;
- 2:00 p.m. to 6:00 p.m. hours were assigned probabilistically, based on the survey data.

Non-Summer Weekend

- (i) Midnight to 9:00 a.m. and (ii) 9:00 p.m. to midnight hours were assigned as "indoors at home";
- 9:00 a.m. to 6:00 p.m. hours were assigned probabilistically, based on the survey data;
- 6:00 p.m. to 9:00 p.m. hours were assigned as 40 minutes "outdoors", randomly selected.

Summer

- (i) Midnight to 9:00 a.m. and (ii) 9:00 p.m. to midnight hours were assigned as "indoors at home";
- 9:00 a.m. to 9:00 p.m. hours were assigned probabilistically, based on the survey data.

Microenvironmental assignments were made in 20-minute increments for each day of the year. The survey typically asked if the student spent at least half of the time outdoors for each time period (e.g., 2 hours out of a 3-hr period, or 30 minutes out of an hour). Thus, for each hour in the survey reported as principally outdoors, 40 minutes were assigned to outdoors activities. The other 20 minutes were assigned as indoors at home or indoors at

school, as appropriate.

In contrast to the REHEX-I and REHEX-II models where detailed time activity patterns were input to the models, the REHEX-III model was pre-loaded for this application with time-activity survey data and with internal sub-routines to probabilistically estimate time-activity patterns for each child on each day of the year. This pre-loading approach greatly reduced data storage requirements and computer processing time.

6.3.2.2 Microenvironmental Concentration Estimates

The basis for the initial concentration estimates for each microenvironment was the hourly ambient ozone and NO_2 data collected in each community. For the outdoor microenvironment, the exposure was assigned the ambient concentration for that hour. During the school monitoring program, indoor and outdoor ozone concentrations were measured at 48 of the 50 CHS Phase II schools in 1993-94. The average indoor to outdoor ozone ratio over the sampling periods for each school was calculated and used in the model for students attending each different school. For schools where no measurements were available, the average from all of the schools was used. For nitrogen dioxide, there were no school measurements, so a nominal indoor/outdoor ratio of 0.5 was chosen based on the literature (Spengler et al., 1994). The exposure estimation was then made by multiplying the community monitoring site concentration by the indoor/outdoor ratio.

At the time the modeling was performed for the Phase II of the Children's Health Study, data were not available for indoor concentrations or the indoor/outdoor ratios at individual student's residences. Literature estimates of the ozone indoor/outdoor ratio suggested a value of about 0.2. This value was used for the first version of REHEX-III contingent on the residential study results. For nitrogen dioxide, there was also no student-specific information on the indoor/outdoor ratio. However, the housing questionnaire did ask the students if their homes had gas stoves and/or gas pilot lights, which are known to increase indoor nitrogen dioxide concentrations. Spengler et al. (1994) has found in the Los Angeles Basin that the typical indoor/outdoor ratio for nitrogen dioxide is about 0.4 when there is no

gas stove or pilot light. When there is a gas stove, the indoor concentration is typically 12 ppb higher, and when there is also a pilot light the indoor concentration is about 15 ppb higher. Thus, for nitrogen dioxide there are two additional microenvironments: (1) indoors at home with gas stove without a pilot light and (2) indoors at home with gas stove with pilot light.

6.3.3 Improved Models for Estimation of Indoor Residential Concentrations

One of the purposes of the residential study was to develop improved models for predicting the concentration of ozone and PM_{10} in Southern California homes. The regression analysis presented in Section 6.1 describes 6-variable models which can explain approximately half of the variance in indoor ozone and PM_{10} concentrations, given outdoor concentrations and other information collected in the residential study. These models have much more predictive power than models which use only a constant I/O ratio, as was used in most previous applications of the REHEX model. The 6-variable models could easily be incorporated into the REHEX III model given proper input data for the subjects and their homes. However, as a practical matter, day-specific information on the operating parameters of individual homes is not readily available.

It would be difficult to collect sufficient data to use the six-variable models in long-term exposure assessment like that needed for characterization of exposure in the Children's Health Study. Specifically, the indoor ozone model requires the number of hours windows are open and whether or not an air conditioner is used. The number of hours windows are open is the second most important variable in the indoor ozone model and it is likely to vary on a daily basis. Air conditioner use is also likely to vary daily, although temperature may be a reasonable surrogate for air conditioner use in homes that have air conditioners. Also, the predictive power of the indoor PM_{10} model is improved with the inclusion of daily data for heater use or the duration of ceiling fan use.

A more practical approach for exposure modeling is to use indoor concentration estimation procedures that do not depend on day-specific operating parameters. Several

regression models were evaluated as part of this effort. Model development was based on the same basic modeling approach previously described in Section 6.1, using a screening procedure to identify candidate variables followed by a "best subset" approach to find the best model for a given number of variables. The primary difference between the modeling approach employed here and the approach described in Section 6.1 was the exclusion, in this application, of day-specific home operating factors (such as the length of time windows were kept open and the use of home air conditioning) in order to identify a model not dependent on day-specific operating parameters. As before, we chose to work with the largest model that provided a substantial increase in R^2 over the model with one less variable. Following model selection by this approach, the variables selected in the above procedure, and all of their two-way interactions, were considered for use in the model. As described previously in Section 6.1, the methods of best subsets (Miller 1990) and the Mallows C_p criterion (Mallows 1973) were then applied to choose the final model.

The simplest model for ozone is the one that estimates indoor levels from outdoor data alone. Using the data collected in this project, the following relationship can be found:

$$[O3]_{indoor} = 0.25\,[O3]_{ambient} \qquad (3)$$

This model, however, only explains 24 percent of the variance. Using the approach described above, a model that predicts indoor ozone levels from community station-monitored ozone concentration ($[O3]_{ambient}$), 24-hr minimum temperature ($T_{24\text{-hr min}}$), building age (Built before 1960), and whether or not a residence has central air conditioning equipment (AC Equipment) was able to explain 49 percent of the variance in the indoor ozone data collected in this study. This model,

$$[O3]_{indoor} = -14.43 + 0.27\,[O3]_{ambient} + 1.13\,[T]_{24\text{-}hr\ \min} + 3.99\,[Built\ before\ 1960] + 10.74\,[AC\ Equipment] - 1.05\,[T]_{24\text{-}hr\ \min}[AC\ Equipment] \qquad (4)$$

where ozone is in ppb, temperature in °C, and the "built before 1960" and "AC equipment" parameters have values of zero for false and one for true, showed almost as much predictive power ($R^2 = 0.49$) as the 6-variable model that uses day-specific window opening data and air conditioner use data ($R^2 = 0.55$). The model described by Equation (4) could readily be implemented in REHEX-III and is probably the best model for this purpose. However, it should be recognized that models like REHEX use 1-hr average estimates of indoor concentrations and these indoor models were derived from 24-hr average ozone concentrations. The accuracy and applicability of the indoor models for shorter averaging times is unknown. It is likely that ozone models derived from 24-hr data will under-predict 1-hr maximum daytime levels because the dynamic range of the 24-hr ozone data is so much lower than that for the 1-hr maximum ozone data.

For PM_{10}, it is essential to treat homes with nonsmokers and smokers separately, because smoking usually makes a large contribution to indoor PM levels (Wallace 1996). In the analyses attending this study, we considered only non-smoking homes, because there were too few homes with active smokers during study sampling to provide useful data for the development of improved models. As was the case with ozone, the simplest model that can be proposed is one based on outdoor data alone. Using the PM data collected in this study, the following relationship was proposed:

$$[PM_{10}]_{indoor} = 0.95\,[PM_{10}]_{ambient} \qquad (5)$$

This simplistic approach was not especially informative, explaining about 10 percent of the variance in indoor PM_{10} in non-smoking homes. An improved model, based on the collected data and presented in Section 6.1, predicted indoor PM_{10} levels from ambient PM_{10} concentration (monitored at the community station) and the existence of odor or mold (assigning a model factor value of 0 if absent, and 1 if present), the presence of house pets (0 or 1, for absence or presence), and the location of the heating duct in the home, as shown below:

$$[PM10]_{indoor} = 22.04 + 0.19[PM10]_{ambient} + 11.11[Odor\ of\ Mold] + 5.68[Pets] - 6.04[Heat\ Duct\ in\ Interior\ Closet\ or\ Wall] \quad (6)$$

This model explained 33 percent of the variance in indoor PM_{10} levels in the homes of nonsmokers in this study. This equation does not use day-specific data (other than for ambient PM_{10}, monitored at the community station) and could readily be implemented in the REHEX model.

In principle, a similar approach could be followed to develop a $PM_{2.5}$ model, if ambient $PM_{2.5}$ data were routinely available. At the present time, however, PM_{10} data is routinely collected at community stations (for regulatory purposes); $PM_{2.5}$ information is only sporadically collected and does not really exist on any routine or regular basis for the communities in which this study was performed. In a previous section (see Section 6.2), interpolated concentrations of $PM_{2.5}$ were developed and used, based on PM_{10} measurements and Hi-Vol or dichotomous sampler data collected in a earlier study by other investigators. The modeling results were unimpressive, but this outcome may have been partially due to the lack of valid community $PM_{2.5}$ data to use as model input. Therefore, while a useful $PM_{2.5}$ model was not developed or presented here, future development and refinement of REHEX-III, in the context of the CARB-supported Children's Health Study, may promote progress in obtaining community $PM_{2.5}$ data, and may lead to improved $PM_{2.5}$ modeling.

Successful model development and application rely on a number of factors. In this application, the potential for success must be tempered by the realization that, in the modeling efforts undertaken here, regression equations have been developed from data collected over a 24-hr average collection period, but that the ultimate (REHEX-based) use for this data is in the context of far shorter (1-hr) averaging times. As stated earlier, the accuracy and applicability of these developed models for shorter averaging times is unknown. Since the dynamic range of 24hr averaged data is generally more limited than for 1hr data, one might expect the model to underestimate 1hr maximum daytime levels. However, no shorter

duration data was available from the current study.

6.3.4 References

Lurmann F.W. (1993) User's guide to the regional human exposure (REHEX) model. Draft report prepared for the South Coast Air Quality Management District, Diamond Bar, CA by Sonoma Technology, Inc., Santa Rosa, CA, STI-90070-1363-D.

Lurmann F.W., Winer A.M., and Colome S.D. (1989) Development and application of a new regional human exposure (REHEX) model. In *Proceedings from the U.S. Environmental Protection Agency and Air & Waste Management Association Conference on Total Exposure Assessment Methodology: New Horizons, Las Vegas, NV, November, 27-30,* Air & Waste Management Association, Pittsburgh, PA.

McCurdy T. and Johnson T. (1989) Development of a human exposure model for ozone. Paper presented at the *Air & Waste Management Association and U.S. Environmental Protection Agency Joint Conference on Total Exposure Assessment Methodology: New Horizons, Las Vegas, NV, November 27-30.*

Mallows CL, "Some comments on C_p," Technometrics 15:661-675, (1973).

Miller AJ, Subset Selection in Regression, Chapman and Hill, New York (1990).

Sexton K. and Ryan P.B. (1988) Assessment of human exposure to air pollution: methods, measurements and models. In *Air Pollution, the Automobile and Public Health,* Watson A.Y., Bates R.R., and Kennedy D., eds., National Academy Press, New York, NY, pp. 207-238.

Spengler J., Schwab M., Ryan P.B., Billick, I., Colome S., Wilson, A.L., and Becker, E. (1994) Personal exposure to nitrogen dioxide in the Los Angeles Basin, *J. Air Waste Manage. Assoc.*, **44**, 39-47.

Wallace, L. (1996) Indoor Particles: A Review. *J. Air & Waste Manage. Assoc.* **46**, 98-126.

Table 6-1. Best Sets of Variables for Predicting Indoor O_3 among Those Containing Station Measurement of O_3 and Five or Fewer Other Explanatory Variables

Number of Variables	Best Set of Variables	R^2
1	Station O_3	0.24
2	Station O_3, Number of Hours Windows Open	0.38
3	Station O_3, Number of Hours Windows Open, Minimum Temperature	0.43
4	Station O_3, Number of Hours Windows Open, Minimum Temperature, Use of Central Refrigerant Recirculating Air Conditioner	0.46
5	Station O_3, Number of Hours Windows Open, 24 Hour Average Temperature*, Use of Central Refrigerant Recirculating Air Conditioner, Odor of Mold	0.48
6	Station O_3, Number of Hours Windows Open, Minimum Temperature, Use of Central Refrigerant Recirculating Air Conditioner, Odor of Mold, Use of Central Evaporative Air Conditioner	0.49

R^2 is for the model containing the main effects of the variables only, without interactions.
* Replacing 24 hour average temperature with minimum temperature reduces R^2 by 0.001.

Table 6-2. Selected Model for Predicting Indoor O_3, with Coefficients Estimated from Measured Indoor Levels Uncorrected For Limit of Detection

Variable	Coefficient	P–value
Intercept	-1.507	0.50
Station O_3	0.053	0.37
Number of Hours Windows Open	-0.419	0.00
Minimum Temperature	0.311	0.02
Use of CRRAC	-5.600	0.00
Station O_3 × No. of Hrs. Win. Open	0.012	0.00
Minimum Temperature × No. of Hrs. Win. Open	0.023	0.01

CRRAC: Central Refrigerant Recirculating Air Conditioner. Units for Indoor O_3 and Station O_3 are ppb. Units for temperature are °C. $R^2 = 0.55$.

Table 6-3. Selected Model for Predicting Indoor O_3, with Coefficients Estimated from Measured Indoor Levels Where Those Below the Limit of Detection are Replaced With Their Lower or Upper Limits

	Lower Limit (0 ppb)		Upper Limit (5 ppb)	
Variable	Coefficient	P–value	Coefficient	P–value
Intercept	-3.043	0.19	2.810	0.17
Station O_3	0.063	0.31	0.045	0.41
Number of Hours Windows Open	-0.429	0.01	-0.467	0.00
Minimum Temperature	0.303	0.03	0.173	0.16
Use of CRRAC	-5.864	0.00	-4.164	0.02
Station O_3 × No. of Hrs. Win. Open	0.013	0.00	0.012	0.00
Minimum Temperature × No. of Hrs. Win. Open	0.025	0.01	0.025	0.00

CRRAC: Central Refrigerant Recirculating Air Conditioner. Units for Indoor O_3 and Station O_3 are ppb. Units for temperature are °C. For lower limit, $R^2 = 0.55$, for upper limit, $R^2 = 0.52$.

Table 6-4. Selected Model for Predicting Indoor O_3, with O_3 Level Outside the House Replacing Station O_3. Coefficients are Estimated from Measured Indoor Levels Uncorrected For Limit of Detection

Variable	Coefficient	P–value
Intercept	-4.012	0.02
O_3 Outside House	0.227	0.00
Number of Hours Windows Open	-0.279	0.03
Minimum Temperature	0.127	0.31
Use of CRRAC	-6.748	0.00
O_3 Outside House × No. of Hrs. Win. Open	0.007	0.00
Minimum Temperature × No. of Hrs. Win. Open	0.027	0.01

CRRAC: Central Refrigerant Recirculating Air Conditioner. Units for Indoor O_3 and O_3 Outside House are ppb. Units for temperature are °C. $R^2 = 0.61$.

Table 6-5. Correlations Between Indoor O_3, Outdoor O_3, and Station O_3.

	Indoor O_3	Outdoor O_3	Station O_3
Indoor O_3	1.00	0.58	0.49
Outdoor O_3	0.58	1.00	0.76
Station O_3	0.49	0.76	1.00

Table 6-6. Best Sets of Variables for Predicting Indoor PM_{10} among Those Containing Station Measurement of PM_{10} and Five or Fewer Other Explanatory Variables

Number of Variables	Best Set of Variables	R^2
1	Station PM_{10}	0.13
2	Station PM_{10}, Smoking Category	0.40
3	Station PM_{10}, Smoking Category, Existence of Dirt Driveway	0.44
4	Station PM_{10}, Smoking Category, Heat Used, Other Smoke in Home	0.46
5	Station PM_{10}, Smoking Category, Existence of Dirt Driveway, Duration of Ceiling Fan Use, Built in 1950s Indicator	0.49
6	Station PM_{10}, Smoking Category, Existence of Dirt Driveway, Duration of Ceiling Fan Use, Built in 1950s Indicator, Other Activities Indicator	0.52

Other activities are those that produce smoke, dust, or pollen besides vacuuming, dusting or sweeping, carpet cleaning, lawn mowing, gardening, burning leaves or wood, outdoor cooking, grilling, frying, or barbecueing, indoor cooking, grilling, or frying, or using a clothes dryer.

Table 6-7. Model for Predicting Indoor PM_{10} from Ambient PM_{10} and Smoking Level, with Coefficients Estimated from Measured Indoor Levels.

Variable	Coefficient	P–value
Intercept	17.069	0.01
Station PM_{10}	0.398	0.00
Smoking 1–10 Cigarettes	9.646	0.45
Smoking More Than 10 Cigarettes	67.223	0.00

Units for PM_{10} are $\mu g\ /m^3$. Baseline is no smoking. $R^2 = 0.40$.

Table 6-8. Selected Model for Predicting Indoor PM_{10}, with Coefficients Estimated from Measured Indoor Levels.

Variable	Coefficient	P–value
Intercept	12.458	0.07
Station PM_{10}	0.463	0.00
Smoking 1–10 Cigarettes	11.282	0.36
Smoking More Than 10 Cigarettes	63.126	0.00
Dirt Driveway	22.203	0.02
Number of Hours of Ceiling Fan Use	-0.927	0.07
Built in 1950s	21.097	0.05

Units for PM_{10} are $\mu g\ /m^3$. Baseline is no smoking, no dirt driveway, house built other than in the 1950s. $R^2 = 0.49$.

Table 6-9. Model for Predicting Indoor PM_{10}, Selection Based on All Residences, with Coefficients Estimated from Measured Indoor Levels in Non-Smoking Residences

Variable	Coefficient	P–value
Intercept	24.885	0.07
Station PM_{10}	0.241	0.01
Dirt Driveway	0.331	0.95
Number of Hours of Ceiling Fan Use	-0.496	0.08
Built in 1950s	4.827	0.47

Units for PM_{10} are $\mu g\ /m^3$. Baseline is no dirt driveway, house built other than in the 1950s. $R^2 = 0.15$.

Table 6-10. Model for Predicting Indoor PM_{10}, Selection and Coefficient Estimation Based on Measured Indoor Levels in Non-Smoking Residences

Variable	Coefficient	P–value
Intercept	20.740	0.00
Station PM_{10}	0.402	0.00
Number of Hours of Ceiling Fan Use	-0.688	0.01
Odor of Mold	15.284	0.00
Other Odor	-15.235	0.00
Other Smoke	-12.255	0.01
Cooking	8.340	0.01
Two Bedroom House	11.785	0.02
Central Refrigerating Recirculating Air Conditioner	2.842	0.76
Station $PM_{10}\times$ Central Refrigerating Recirculating Air Conditioner	-0.251	0.09

Units for PM_{10} are $\mu g\ /m^3$. Other odor refers to odors other than those of mold, cigarettes, formaldehyde, or animals. Baseline is no odor. $R^2 = 0.55$.

Table 6-10a. Best Sets of Variables for Predicting Indoor PM_{10} among Those Containing Station Measurement of PM_{10} and Eight or Fewer Other Explanatory Variables

Number of Variables	Best Set of Variables	R^2
1	Station PM_{10}	0.09
2	Station PM_{10}, Other Odor	0.21
3	Station PM_{10}, Other Odor, Odor of Mold	0.29
4	Station PM_{10}, Other Odor, Odor of Mold Duration of Ceiling Fan Use,	0.34
5	Station PM_{10}, Other Odor, Odor of Mold Duration of Ceiling Fan Use, Cooking	0.40
6	Station PM_{10}, Other Odor, Odor of Mold Duration of Ceiling Fan Use, Cooking, Other Smoke in Home	0.45
7	Station PM_{10}, Other Odor, Odor of Mold Duration of Ceiling Fan Use, Cooking, Other Smoke in Home Central Refrigerating Recirculating Air Conditioner	0.48
8	Station PM_{10}, Other Odor, Odor of Mold Duration of Ceiling Fan Use, Cooking, Other Smoke in Home Central Refrigerating Recirculating Air Conditioner Two-Bedroom House	0.53
8	Station PM_{10}, Other Odor, Odor of Mold Duration of Ceiling Fan Use, Cooking, Other Smoke in Home Central Refrigerating Recirculating Air Conditioner Two-Bedroom House, Wall Unit Air Conditioner	0.54

Table 6-11. Selected Model for Predicting Indoor PM_{10}, with Coefficients Estimated from Measured Indoor Levels in Non-Smoking Residences with PM_{10} Level Outside the House Replacing Station PM_{10}. Coefficients are Estimated from Measured Indoor Levels in Non-Smoking Residences

Variable	Coefficient	P–value
Intercept	31.253	0.00
Outdoor PM_{10}	0.208	0.01
Number of Hours of Ceiling Fan Use	-0.716	0.01
Odor of Mold	12.965	0.01
Other Odor	-14.054	0.00
Other Smoke	-10.790	0.05
Cooking	8.741	0.01
Two Bedroom House	11.785	0.02
Central Refrigerating Recirculating Air Conditioner	-8.051	0.41
Station $PM_{10}\times$ Central Refrigerating Recirculating Air Conditioner	0.041	0.76

Units for PM_{10} are $\mu g\ /m^3$. Other odor refers to odors other than those of mold, cigarettes, formaldehyde, or animals. Baseline is no odor. $R^2 = 0.47$.

Table 6-12. Correlations Between Indoor PM_{10}, Outdoor PM_{10}, and Station PM_{10}.

	Indoor PM_{10}	Outdoor PM_{10}	Station PM_{10}
Indoor PM_{10}	1.000	0.356	0.355
Outdoor PM_{10}	0.356	1.000	0.634
Station PM_{10}	0.355	0.634	1.000

Table 6-13. Best Sets of Variables for Predicting Indoor $PM_{2.5}$ among Those Containing Station Measurement of $PM_{2.5}$ and Five or Fewer Other Explanatory Variables

Number of Variables	Best Set of Variables	R^2
1	Station $PM_{2.5}$	0.46
2	Station $PM_{2.5}$, Duration of Stove Use	0.57
3	Station $PM_{2.5}$, Duration of Stove Use, Two-Bedroom Residence	0.64
4	Station $PM_{2.5}$, Duration of Stove Use, Two-Bedroom Residence, Minimum Temperature	0.66
5	Station $PM_{2.5}$, Duration of Stove Use, Two-Bedroom Residence, Dusting or Sweeping, Use of Heater	0.68
6	Station $PM_{2.5}$, Duration of Stove Use, Two-Bedroom Residence, Dusting or Sweeping, Use of Heater, Return Duct to Air Handler Located in Garage	0.69

Table 6-14. Model for Predicting Indoor $PM_{2.5}$, Selection and Coefficient Estimation Based on Measured Indoor Levels in Non-Smoking Residences

Variable	Coefficient	P–value
Intercept	-3.893	0.29
Station $PM_{2.5}$	0.536	0.00
Number of Minutes of Stove Use	0.164	0.00
Two Bedroom Residence	13.316	0.05

Units for $PM_{2.5}$ are $\mu g / m^3$. $R^2 = 0.61$.

Table 6-15. Selected Model for Predicting Indoor $PM_{2.5}$, with Coefficients Estimated from Measured Indoor Levels in Non-Smoking Residences with $PM_{2.5}$ Level Outside the House Replacing Station $PM_{2.5}$. Coefficients are Estimated from Measured Indoor Levels in Non-Smoking Residences

Variable	Coefficient	P–value
Intercept	-4.971	0.15
Outdoor $PM_{2.5}$	0.612	0.00
Number of Minutes of Stove Use	0.187	0.00
Two-Bedroom Residence	19.037	0.00

Units for $PM_{2.5}$ are $\mu g\ /m^3$. $R^2 = 0.67$.

Table 6-16. Correlations Between Indoor $PM_{2.5}$, Outdoor $PM_{2.5}$, and Station $PM_{2.5}$.

	Indoor $PM_{2.5}$	Outdoor $PM_{2.5}$	Station $PM_{2.5}$
Indoor $PM_{2.5}$	1.000	0.592	0.641
Outdoor $PM_{2.5}$	0.592	1.000	0.720
Station $PM_{2.5}$	0.641	0.720	1.000

Table 6-17. Distance and correlation of pollutant concentrations between community monitors.

Community	Community Monitor #1	Community Monitor #2	Distance (km)	Ozone Correlation (R^2)	PM_{10} Correlation (R^2)
San Dimas	San Dimas	Glendora	3.8	0.84	-
	San Dimas	Azusa	8.7	0.88	-
	Glendora	Azusa	6.8	0.92	-
Riverside/ Mira Loma	UC Riverside	Rubidoux	8.8	0.94	0.22
	Mira Loma	Rubidoux	9.7	0.89	0.63
	UC Riverside	Mira Loma	18.0	0.89	0.20
Lake Gregory	Lake Gregory	Lake Arrowhead	8.2	0.50	-

Table 6-18. Ratios of $PM_{2.5}$ to PM_{10} mass concentrations* in the Children's Health Study at selected locations.

Community	Community Monitoring Site	February-April	May-July	August-October	November-January
San Dimas	San Dimas	0.45	0.46	0.48	0.67
Riverside/ Mira Loma	UC Riverside	0.79	0.56	0.43	0.46
	Mira Loma	0.52	0.45	0.33	0.43
	Rubidoux	0.52	0.49	0.37	0.42
Lancaster	Lancaster	0.26	0.29	0.23	0.46
Lake Gregory	Lake Arrowhead	0.62	0.54	0.34	0.22

* Ratio of 1994 Two-Week Sampler $PM_{2.5}$ mass to adjusted TEOM PM_{10} mass.

Table 6-19. Summary of residential and community monitor ozone concentrations and correlations.

Community	Community Monitoring Site	Mean Community Monitor Ozone (ppb)	Mean Interpolated Ozone (ppb)	Mean Residential Ozone (ppb)	Correlation (R^2) With Individual Monitor Data	Correlation (R^2) With Interpolated Data
San Dimas	San Dimas	34.1	33.9	28.0	0.79	0.82
	Glendora	36.0			0.78	
	Azusa	29.5			0.80	
Riverside/ Mira Loma	UC Riverside	37.7	35.6	31.3	0.75	0.71
	Mira Loma	31.5			0.65	
	Rubidoux	37.6			0.73	
Lancaster	Lancaster	42.1	-	42.9	0.60	-
Lake Gregory	Lake Arrowhead	67.1	64.3	46.2	0.39	0.47
	Lake Gregory	52.5			0.43	

Table 6-20. Summary of residential and community monitor PM_{10} concentrations and correlations.

Community	Community Monitoring Site	Mean Community Monitor TEOM PM_{10} ($\mu g/m^3$)	Mean Community Monitor Adjusted TEOM PM_{10} ($\mu g/m^3$)	Mean Residential PM_{10} ($\mu g/m^3$)[a]	Mean Adjusted Residential PM_{10} ($\mu g/m^3$)[a]	Correlation (R^2) With Individual Monitor Data	Correlation (R^2) With Interpolated Data
San Dimas	San Dimas	29.3	44.7	33.0	26.5	0.43	-
Riverside/ Mira Loma	UC Riverside	46.9	61.8	47.7	38.2	0.35	0.56
	Mira Loma	61.8	79.0			0.45	
	Rubidoux	59.3	78.0			0.49	
Lancaster	Lancaster	33.6	35.2	31.3	25.0	0.01[b]	-
Lake Gregory	Lake Arrowhead	24.0	24.5	13.4	10.7	0.13	-

[a] Quartz-fiber filter data with negative mass were not included in the analysis.
[b] Outliers were removed.

Table 6-21. Summary of residential and estimated community monitor $PM_{2.5}$ concentrations and correlations.

Community	Community Monitoring Site	Mean Community Monitor Estimated $PM_{2.5}$ ($\mu g/m^3$)	Mean Residential $PM_{2.5}$ ($\mu g/m^3$)[a]	Mean Adjusted Residential $PM_{2.5}$ ($\mu g/m^3$)[a]	Correlation (R^2)
San Dimas	San Dimas	22.3	18.1	13.9	0.54
Riverside/ Mira Loma	UC Riverside	29.8	28.3	21.8	0.07
	Mira Loma	30.3			0.36
	Rubidoux	32.3			0.45
Lancaster	Lancaster	10.8	8.3	6.4	<0.01
Lake Gregory	Lake Arrowhead	9.0	7.1	5.5	0.45[b]

[a] Quartz-fiber filter data with negative mass were not included in the analysis.
[b] Outliers were removed.

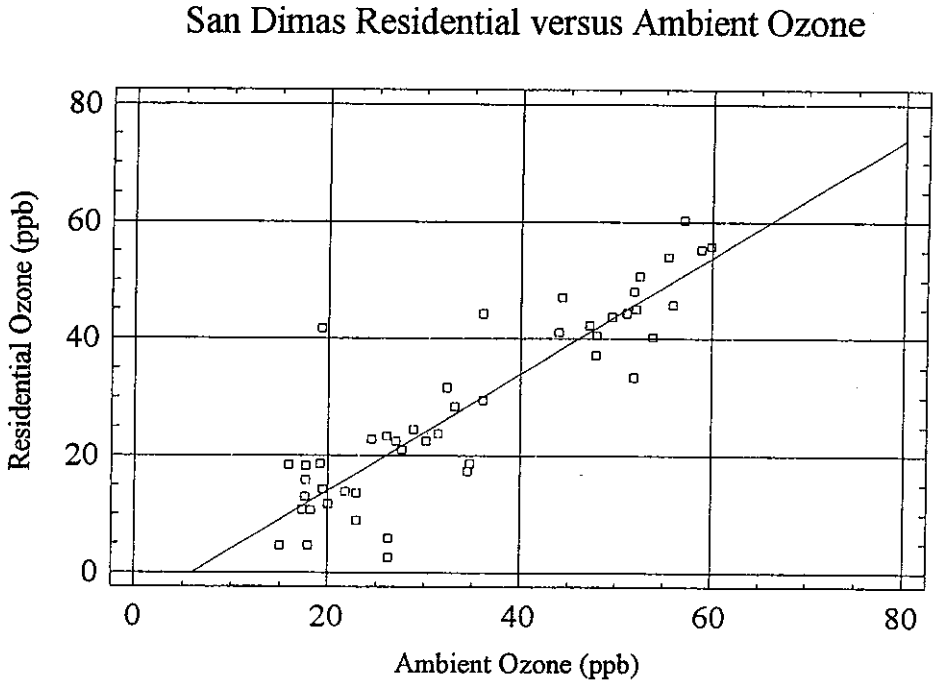


Figure 6-1. Ozone concentration at San Dimas residences versus the San Dimas community monitoring station.

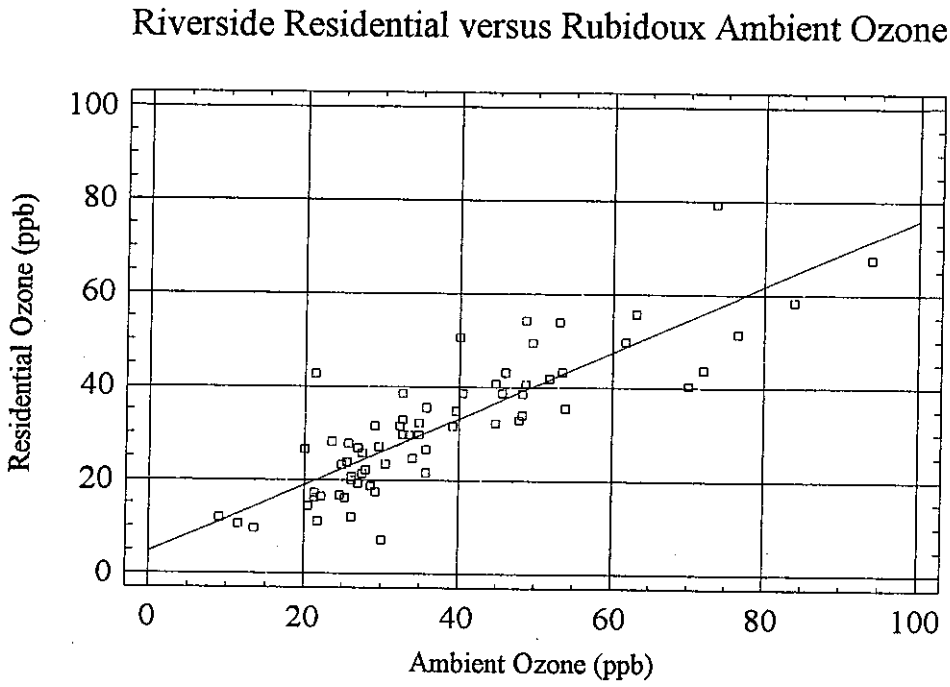


Figure 6-2. Ozone concentration at Riverside residences versus the Rubidoux community monitoring station.

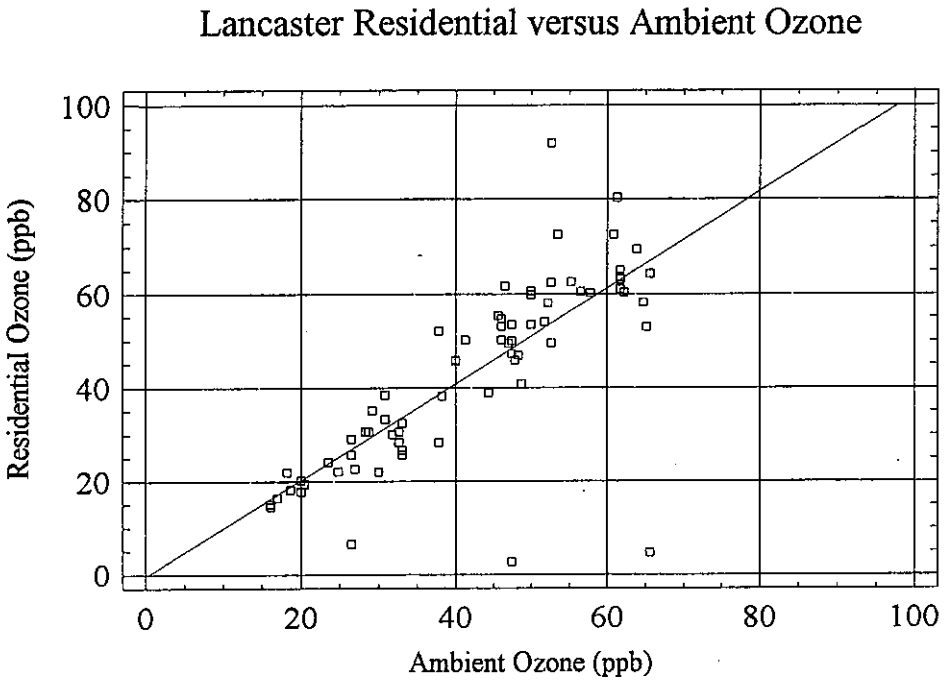


Figure 6-3. Ozone concentration at Lancaster residences versus the Lancaster community monitoring station.

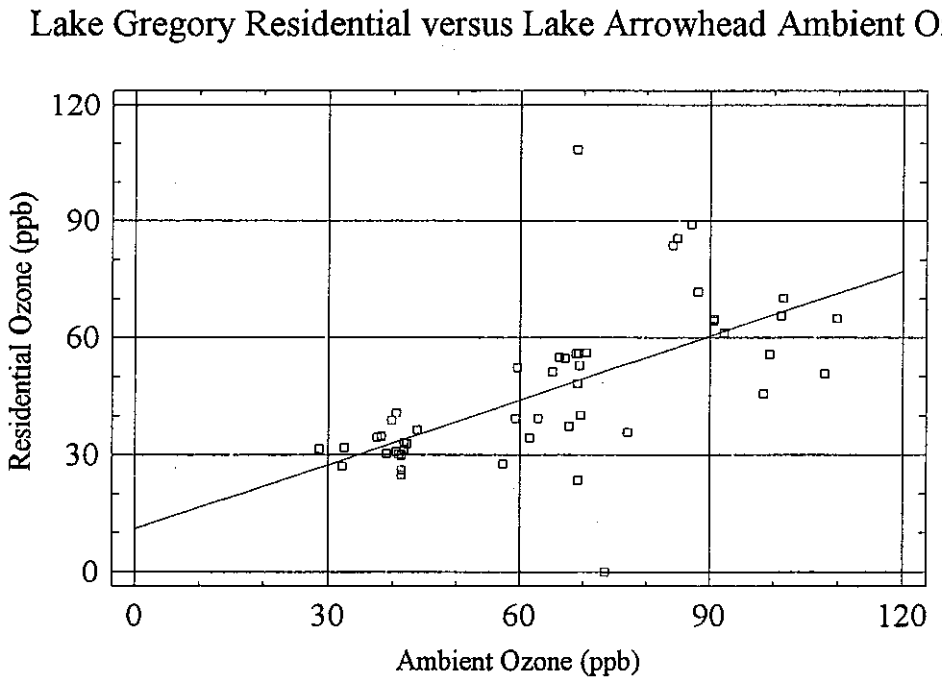


Figure 6-4. Ozone concentration at Lake Gregory residences versus the Lake Arrowhead community monitoring station.

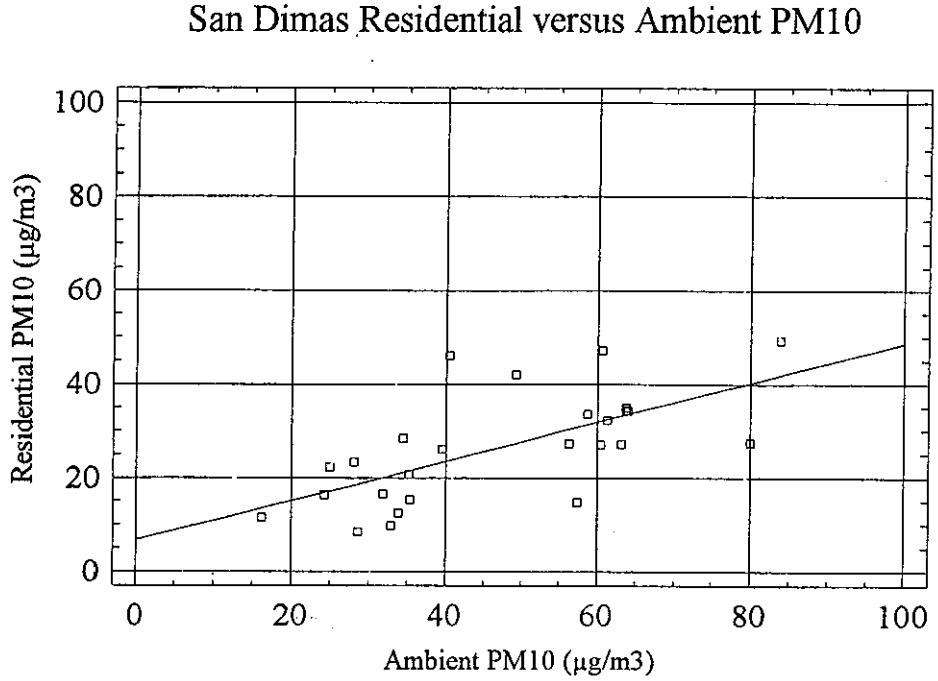


Figure 6-5. PM_{10} concentration at San Dimas residences versus the San Dimas community monitoring station.

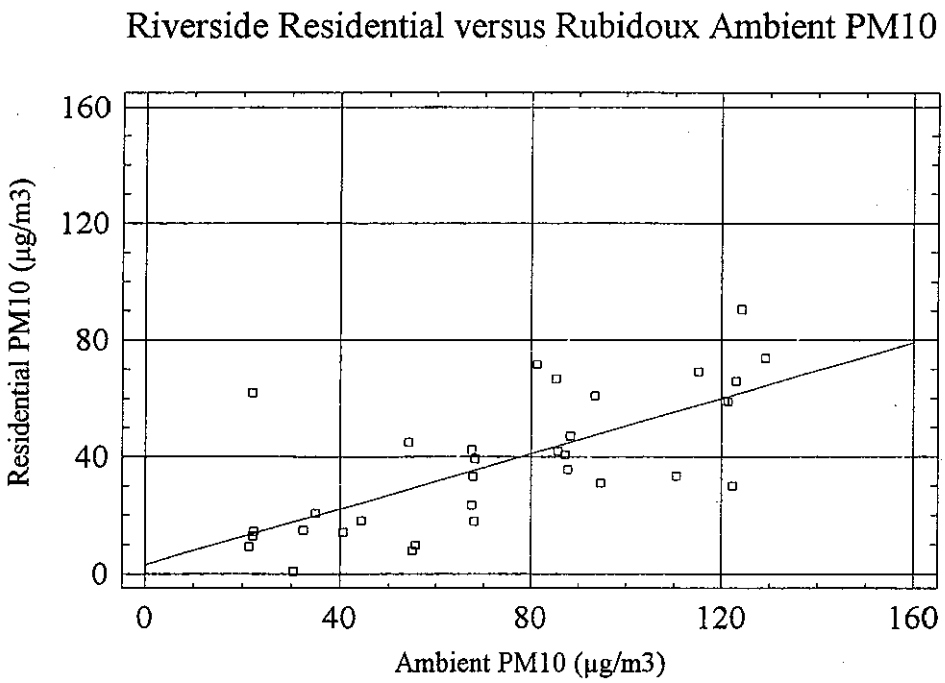


Figure 6-6. PM_{10} concentration at Riverside residences versus the Rubidoux community monitoring station.

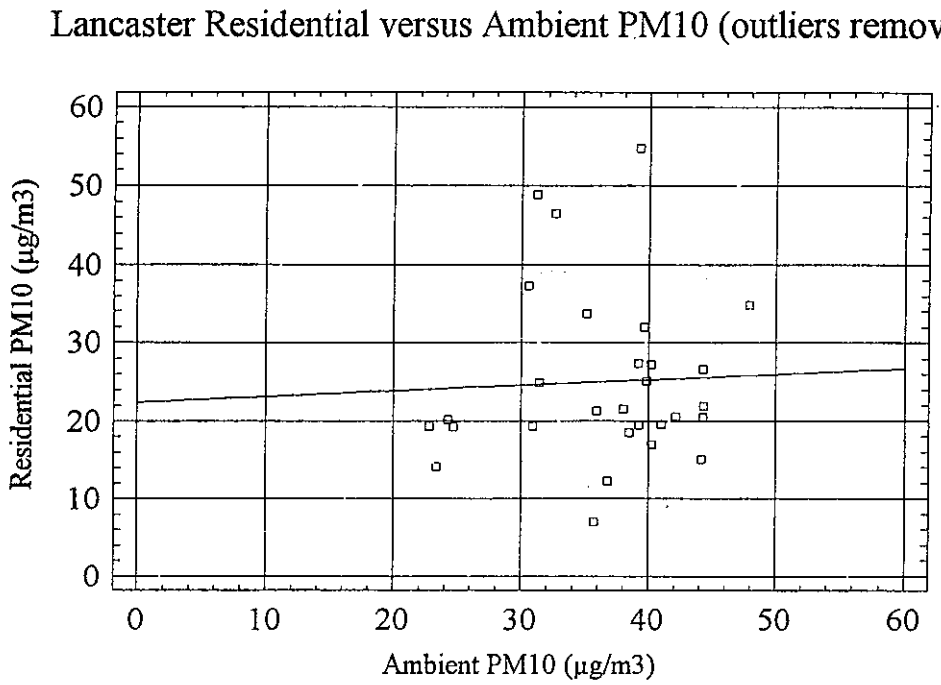


Figure 6-7. PM_{10} concentration at Lancaster residences versus the Lancaster community monitoring station.

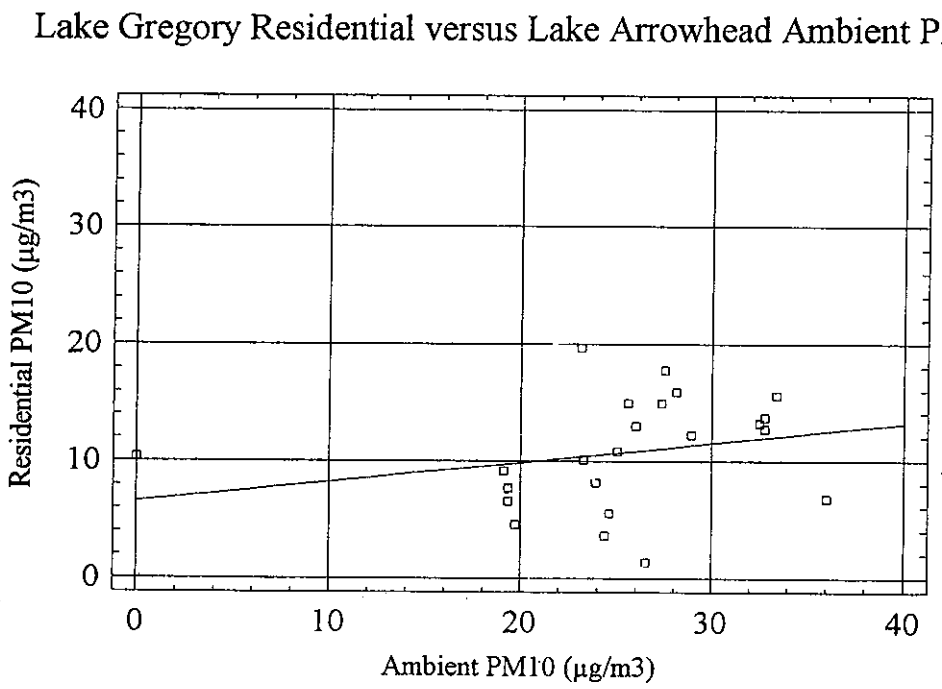


Figure 6-8. PM_{10} concentration at Lake Gregory residences versus the Lake Arrowhead community monitoring station.

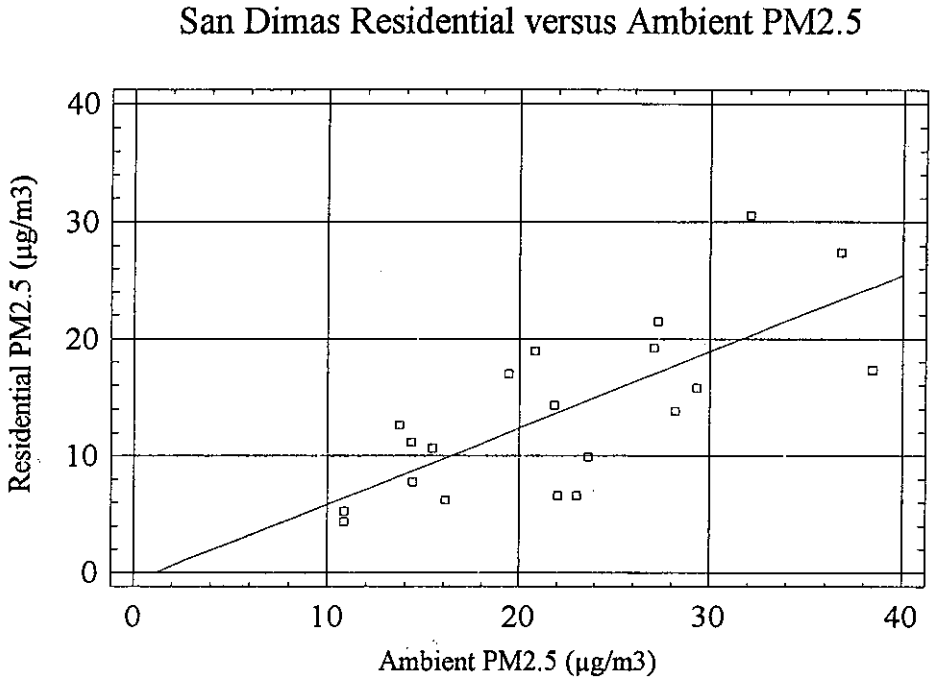


Figure 6-9. $PM_{2.5}$ concentration at San Dimas residences versus the San Dimas community monitoring station.

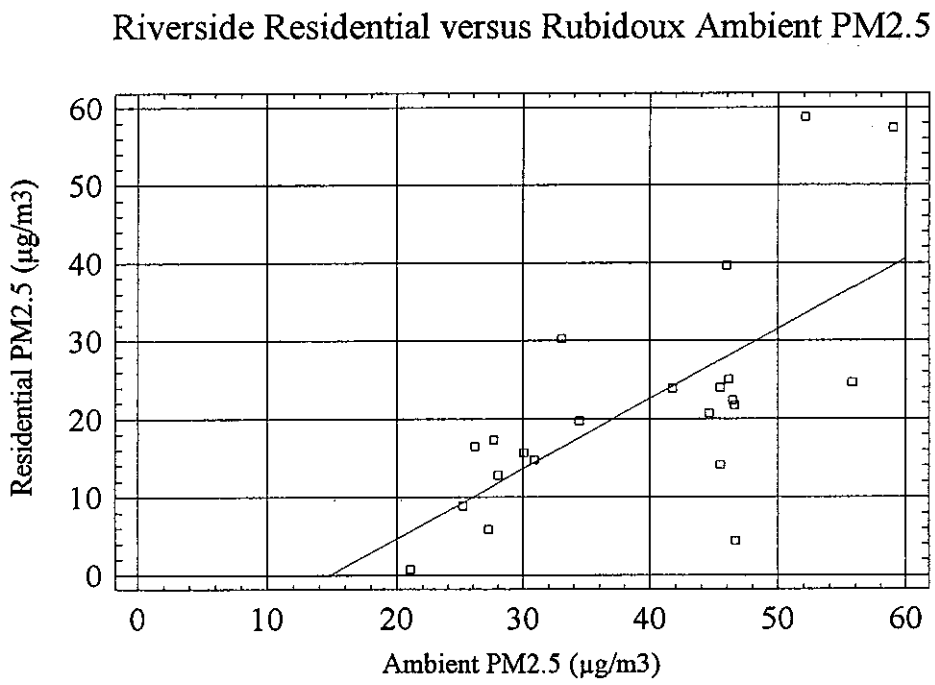


Figure 6-10. $PM_{2.5}$ concentration at Riverside residences versus the Rubidoux community monitoring station.

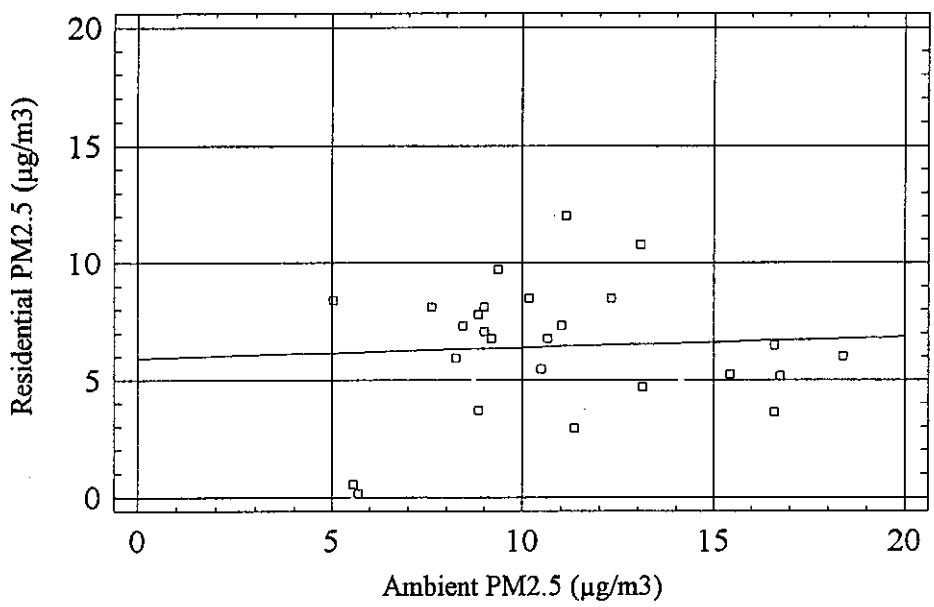


Figure 6-11. $PM_{2.5}$ concentration at Lancaster residences versus the Lancaster community monitoring station.

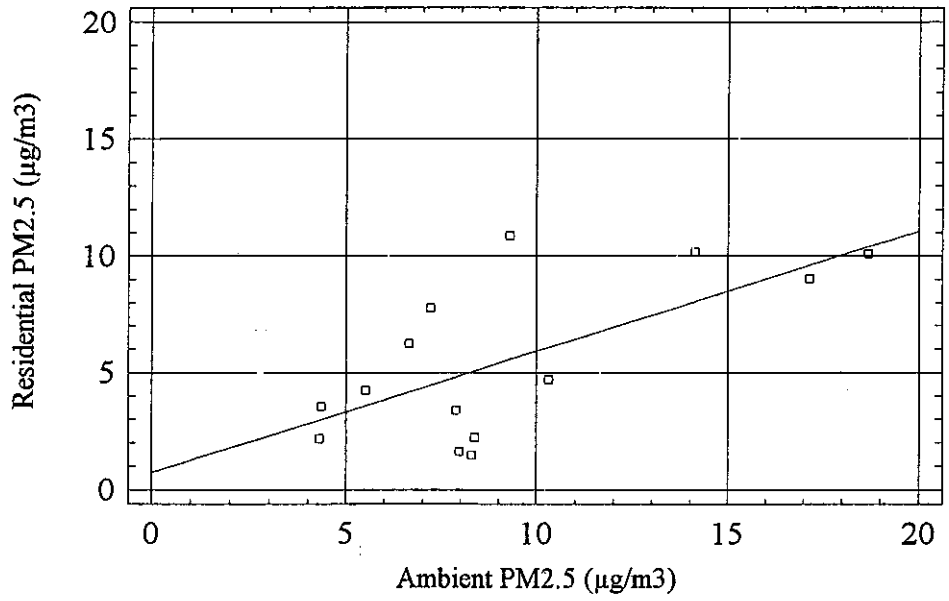


Figure 6-12. $PM_{2.5}$ concentration at Lake Gregory residences versus the Lake Arrowhead community monitoring station.

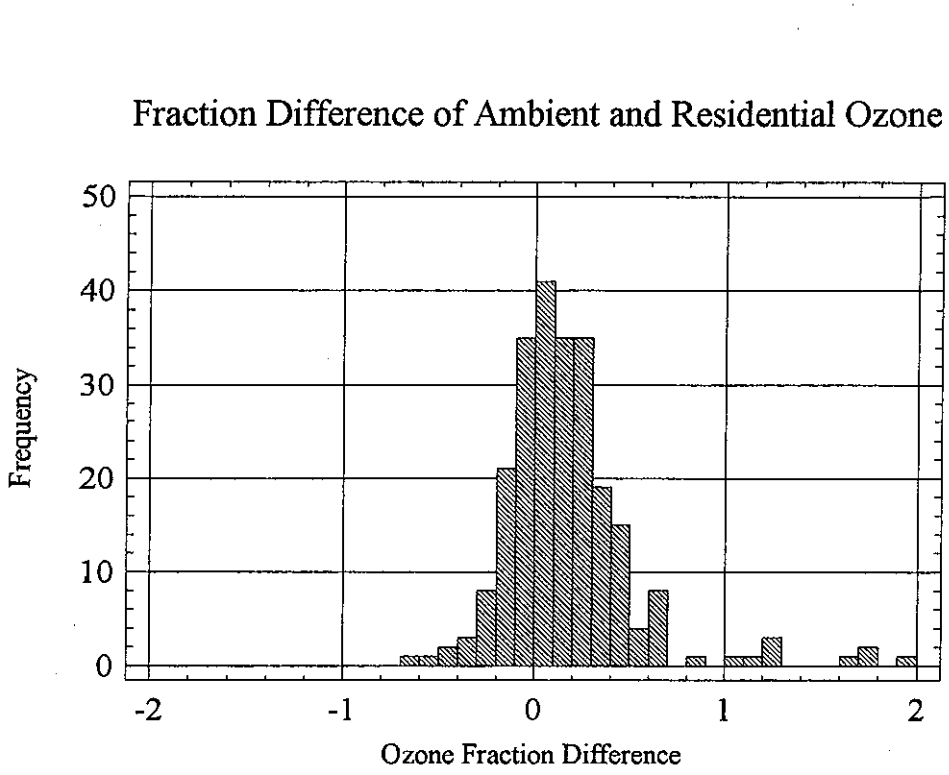


Figure 6-13. Histogram of the difference between the ambient and residential ozone concentration for all communities.

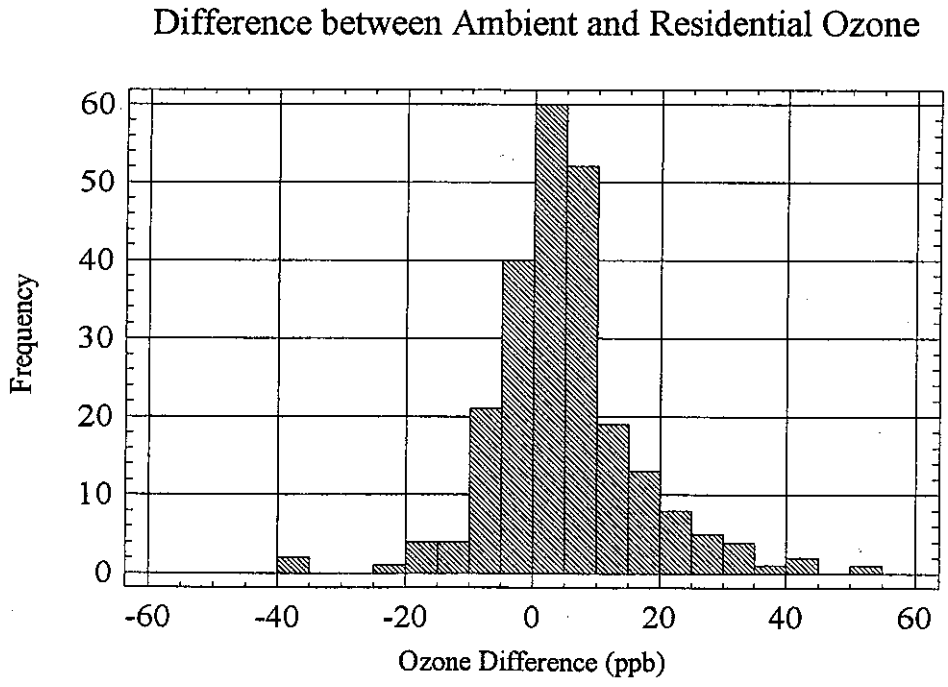


Figure 6-14. Histogram of the fraction difference between the ambient and residential ozone concentration for all communities.

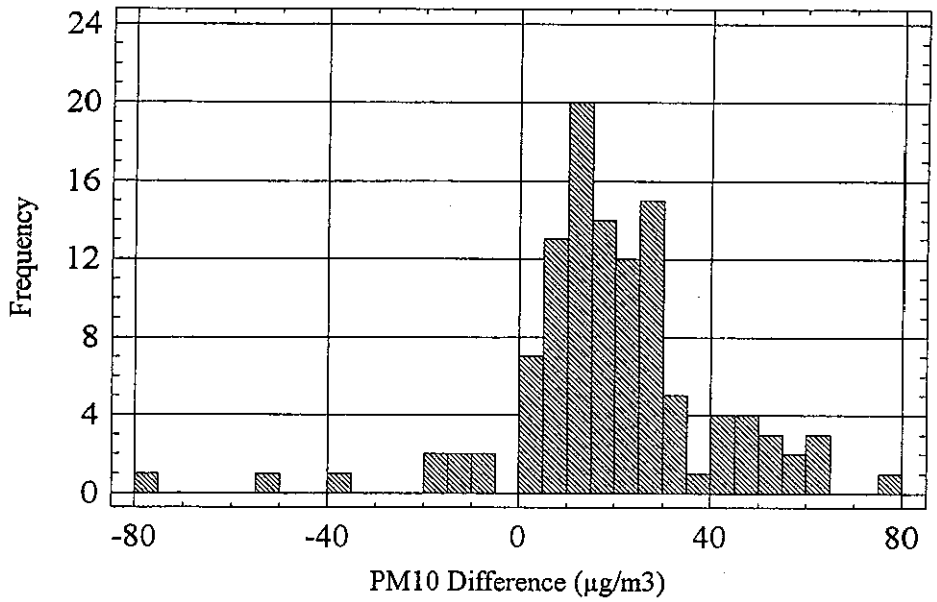


Figure 6-15. Histogram of the difference between the ambient and residential PM_{10} concentration for all communities.

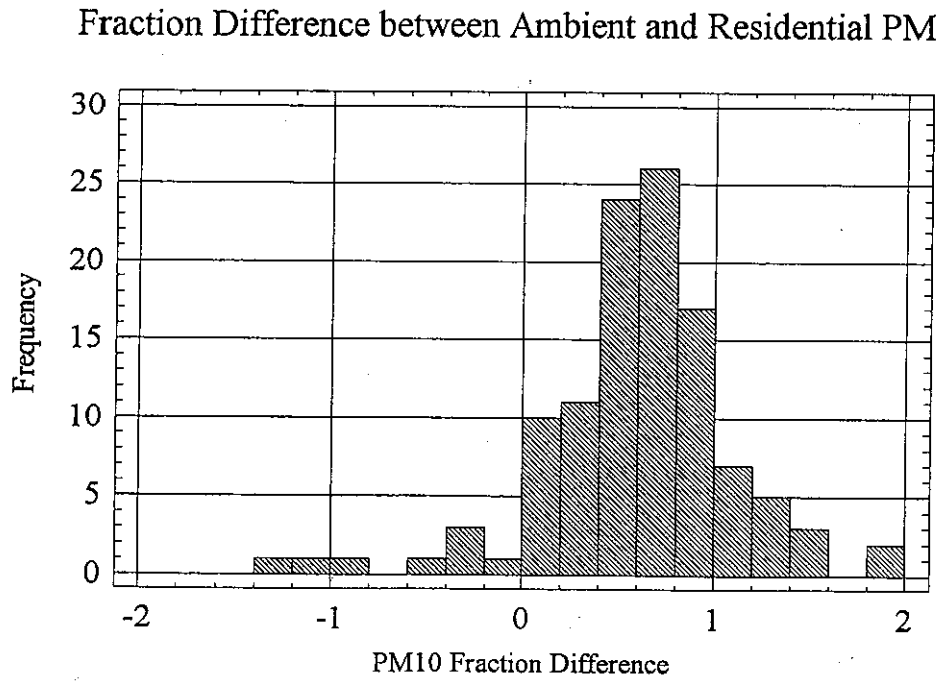


Figure 6-16. Histogram of the fraction difference between the ambient and residential PM_{10} concentration for all communities.

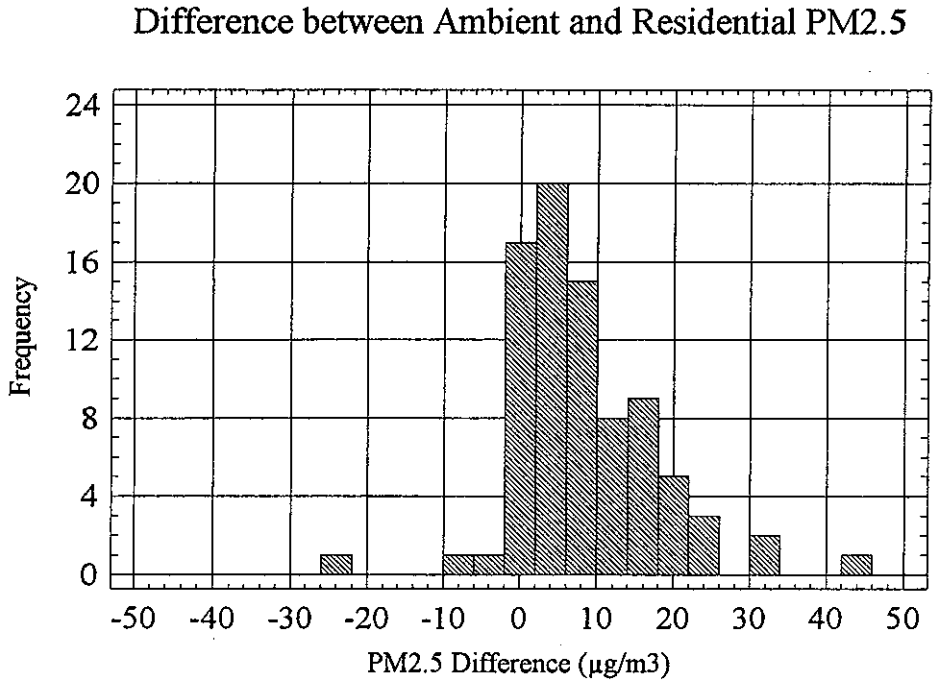


Figure 6-17. Histogram of the difference between the ambient and residential $PM_{2.5}$ concentration for all communities.

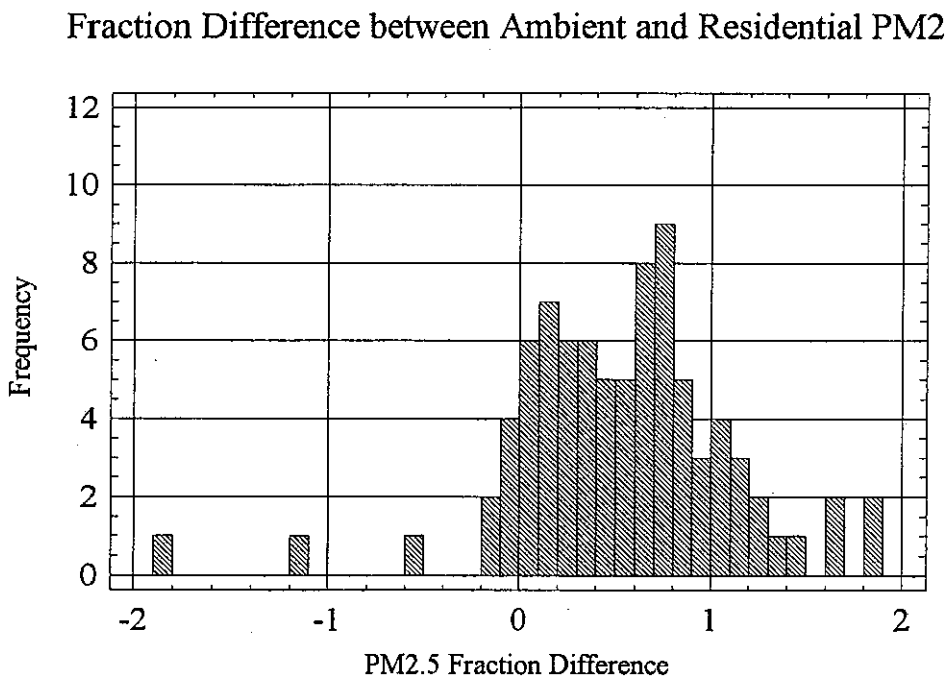


Figure 6-18. Histogram of the fraction difference between the ambient and residential $PM_{2.5}$ concentration for all communities.

7. DISCUSSION

This project measured the 24hr integrated concentrations of ozone, PM_{10}, $PM_{2.5}$, and formaldehyde in a selection of Southern California homes between February and December 1994. The information collected in this project provides a rich data base to assess the relationships between indoor and outdoor pollutants, and the housing factors that may affect indoor and outdoor pollutant levels. In the following sections, some of the comparative information provided by this data set with other previous studies are discussed.

7.1 Comparison of Study Data to State Ambient Standards and Guidelines

Table 7-1 compares the pollutant monitoring information collected in this study with existing California ambient air quality standards or CARB-recommended guidelines for exposure.

7.1.1 Ozone

Comparison of the ozone data collected in this investigation to the California Ambient Air Quality Standard was problematic, due to an inconsistency in exposure metrics. Whereas the State ozone standard is based on a one-hour averaging time, the samples from this study were collected over an integrated 24hr measurement period (due to detection limit concerns). However, continuous ozone monitors were operated in five study homes. Of the five homes, recorded O_3 values exceeded the 0.09 ppm 1Hr Standard outdoors at one residence (Home #33, see Figure 5.2-16), but not indoors over the same time interval. (For this home, the ratio of indoor to outdoor ozone average 0.18). Both outdoors and indoors Home #15 (see Figure 5.2-12), ozone levels exceeded the 1hr standard for several hours during the sampling day afternoon.

7.1.2 PM_{10} and $PM_{2.5}$

For PM_{10}, study samples were compared to the California Ambient Air Quality Standard of 50 $\mu g/m^3$. PM_{10} measurements inside 16 of the homes (or 27% of the homes in which PM_{10} was measured) and outside 17 of them (28%) exceeded the standard. Six of the sixteen homes with elevated indoor PM_{10} levels also had simultaneously measured outdoor

levels higher than the standard. Eight of the remaining ten homes in this subset (i.e., those homes with indoor levels exceeding the 50 $\mu g/m^3$ ambient standard, but with simultaneous outdoor levels lower than 50$\mu g/m^3$) had measured indoor levels of 51 to 56 $\mu g/m^3$, only marginally higher than the ambient standard. Of the remaining two homes (with measured PM_{10} levels of 140 and 83 $\mu g/m^3$, respectively), two smokers lived in the higher PM_{10} home.

The collected questionnaire data was reviewed to identify the possible factors that contributed to elevated PM_{10} levels in these homes. The results of that analysis are summarized in Table 7-2. A number of factors were significantly associated with homes in which PM_{10} levels exceeded 50 $\mu g/m^3$, including: smoking; the presence of moldy odors; combustion activities, such as the use of a wall furnace or stove; air-moving (dust-raising) activities, including dusting and the use of a window fan or air conditioning (with swamp-cooled homes primarily responsible for the association with air conditioning type); home construction-related factors, such as the type of housing foundation, recent remodeling, or the presence of an attached garage. Some of the apparent associations, such as with housing foundation, may be surrogates for other factors or community-specific situations.

Evaluation of study $PM_{2.5}$ data in this context was not possible, since no State ambient standard currently exists for $PM_{2.5}$.

7.1.3 Formaldehyde

For the formaldehyde samples collected in this project, comparisons were made to CARB-recommended "action" and "target" levels of 0.10 ppm and 0.05 ppm indoor formaldehyde. No sample collected, either indoor or outdoors, was observed to be in excess of either recommended level.

7.2 Comparison of Study Data with Other Studies

7.2.1 Ozone

Several studies have investigated the ratios between indoor and outdoor ozone concentrations in homes, operated under a variety of conditions. Stock et al noted that indoor O_3 levels were very low (less than 10 ppb) in twelve homes monitored in the Houston area, regardless of outdoor diurnal variation (Stock et al 1985). In a study in which ventilation conditions (the use of mechanical air conditioning or opening and closing of windows) and indoor combustion (use of gas stove) were carefully controlled, Zhang and Lioy reported that mean ratios of indoor to outdoor ozone concentrations ranged from 0.22 to 0.62 (Zhang and Lioy 1994a).

These observations, and the range of observed ratios, agree well with our study data, which showed a mean indoor to outdoor ozone ratio of 0.37, with a standard deviation of 0.25. The direct comparison of our reported indoor and outdoor ozone concentrations with these other studies was not directly possible. In the other studies, continuous monitors were used to collect ozone information for portions of the day, while in our study, an integrated 24hr sampler was used. Continuous monitoring, when used in our study, did show trends consistent with those observed in other researchers' work. Unique to our continuous monitoring ozone data set (albeit on a very small sampling set) was the captured observations of the impact of fan cycling in air-conditioned homes (see Figure 5.2-12) and the influence of opening windows on indoor ozone levels (for example, see Figure 5.2-15).

7.2.2 Particle Data

In recent years, several studies have focused attention on outdoor or indoor contributions to PM_{10} or $PM_{2.5}$ concentrations. With respect to this study, the Particle Total Exposure Assessment Methodology (PTEAM) Study, performed in Riverside CA from late September 1990 through early November 1990, is of special interest (Clayton et al 1993, Thomas et al 1993).

As part of that investigation, 12hr daytime and nighttime PM_{10} and $PM_{2.5}$ concentrations were measured inside and outside the homes of 178 individuals. Indoor daytime PM_{10} and $PM_{2.5}$ levels were similar to those measured outdoors, while nighttime concentrations tended to be slightly lower indoors. Smoking and cooking were associated with elevated PM levels indoors. The central site monitors were reported to be in fairly good agreement (Spearman correlations of 0.8 to 0.85) with measurements made outside the study homes.

In the current investigation, indoor PM_{10} and $PM_{2.5}$ levels were found to be slightly higher than those collected concurrently immediately outside of the homes. Unlike the PTEAM study results, the observed correlations in the present study between PM_{10} data monitored at community stations and the values measured immediately outside resident's homes were poor.

There are several differences between the two studies that may account for the observed disparity in PM_{10} agreement. In our study, PM_{10} measurements made at the community station were collected using TEOM monitors to collect hourly PM data, with 24hr data calculated as an aggregate of the individual sampling hours. At the study homes, PEM monitors, using filters with small sampling pumps, were used to collect 24hr aggregate data. The direct relationship between TEOM and PEM monitors was not explicitly established within the context of the study. In PTEAM, sampling was performed in 12hr daytime/nighttime sampling segments, using the PEM monitor at the community station (in addition to other larger particle sampling instruments). Secondly, PTEAM field operations were performed over a much more restrictive time period (September through November), encompassing one season, in one Southern California community. The PM data collected in the current study were measured in homes in four different communities (including forested, varying elevation, and high desert locations) between from late June through late November (sampling during at least two seasons of the year).

As has been found in other investigations, the importance of smoking and cooking on indoor pollution levels was confirmed by the data collected in the current project. As in several other studies (Koutrakis et al 1992, Clayton et al 1993, Spengler et al 1985, Wallace 1996), smoking and cooking were found in the residential study to be associated with increased PM levels indoors.

7.2.3 Formaldehyde

The formaldehyde levels reported in this study (0 to 39 $\mu g/m^3$ [32 ppb] indoors, and 1 to 10 $\mu g/m^3$ [1 to 8 ppb] outdoors) were on the lower end of the 20 to 150 ppb range previously reported for conventional (non-manufactured) California homes (Ota 1990), and generally lower than those reported in homes sampled in the northeastern United States by Zhang and co-workers (Zhang et al 1994b). Our study homes were predominately single family detached structures of conventional construction, with half of the homes built during the 1970's or 1980's. This may explain the lower levels, since considerable time would have passed for outgassing of originally applied resins and building materials. while we did inquire about recent remodeling or renovation activities and did factor these variables into our analyses, there were no homes identified with elevated levels of formaldehyde.

Outdoor concentrations of formaldehyde reported in this investigation were similarly low and unremarkable. The ratio of indoor to outdoor levels observed in our study homes (6.01, with a range of 0 to 40) was in general agreement with the range of values reported by other investigators in several cities across the United States (Zhang et al 1994b, Ota 1990).

7.2.4 Air Exchange Rates

The AER measurements reported in this study (mean value of 0.7 hr^{-1}, or 0.8 hr^{-1} using the corrected home volume measurement approach, with an observed range from 0.0 to 2.5 hr^{-1}) were in the range of previously reported AERs for homes (Suh et al 1994, Ozkaynak et al, 1994). However, the measured median AER reported for all homes measured in the current study (0.7 hr^{-1}) was slightly lower than that reported for Riverside homes in the PTEAM

investigation. During PTEAM, AER measurements were collected for 12 hr intervals, to investigate potential differences between daytime and nighttime home use and exposure. In that investigation, median daytime AER values were reported to be 0.9 hr^{-1} with nighttime values of 0.7 hr^{-1}. These values agree somewhat better with the AERs measured in the current study in our Riverside/Mira Loma sub-set (0.7 hr^{-1}).

7.2.5 Indoor Acids

Although the indoor acids data collected in the course of this study was clearly a demonstration pilot effort, performed in 12 study homes, some comparison to previously published data is possible. Suh and co-workers reported on airborne acidity measurements collected in 47 Pennsylvania homes (Suh et al 1994). Their reported indoor nitric acid (HNO_3) concentrations (mean of 0.2 ppb, with a maximum of 2.4 ppb HNO_3) and outdoor levels (mean of 1.7 ppb, maximum of 5.2ppb HNO_3) were lower than those reported in our pilot sampling efforts. Outdoor levels measured in our study varied from 1 to 13 $\mu g/m^3$ (0.4 to 5 ppb), while indoor levels were typically less than half of the outdoor reading. The differences in outdoor levels between the two study data sets can likely be attributed to the importance of nitrate chemistry in the oxidizing atmosphere of Southern California.

The formic and acetic acid levels measured in our twelve pilot study homes were in the same range as reported values by investigators in the Eastern United States (Reiss et al 1995, Zhang et al 1994c). In greater Boston homes measured by Reiss and co-workers, indoor formic and acetic acid concentrations averaged 16 ppb (30 $\mu g/m^3$) and 10 ppb (24 $\mu g/m^3$) in the winter, and 29 ppb (55 $\mu g/m^3$) and 18 ppb (44 $\mu g/m^3$) in the summer, respectively. In six homes studied by Zhang and co-workers, formic acid levels averaged 8.8 ppb (17 $\mu g/m^3$), and acetic acid concentrations averaged 24 ppb (59 $\mu g/m^3$). These results were in reasonable agreement with the range of values observed in our study homes, even though the collection techniques were somewhat different across studies.

7.3 References

Clayton CA, RL Perritt, ED Pellizari, KW Thomas, RW Whitmore, LA Wallace, H Ozkaynak, JD Spengler, Particle Total Exposure Assessment Methodology (PTEAM) Study: Distributions of Aerosol and Elemental Concentrations in Personal, Indoor, and Outdoor Air Samples in a Southern California Community, J. Exp. Anal. Environ. Epi. 3,2:227-250 (1993).

Koutrakis P, SLK Briggs, BP Leaderer, Source Apportionment of Indoor Aerosols in Suffolk and Onondaga Counties, New York, Environ. Sci. Technol, 26:521-527 (1992).

Ota E, E Mulberg, "Exposure to Formaldehyde from Indoor Air", California Air Resources Board Report ARB/RD-90-01, July 1990.

Ozkaynak H, J Xue, R Weker, D Butler, P Koutrakis, J Spengler, The Particle Team (PTEAM) Study: Analysis of the Data, Draft Final Report, Volume III, Contract Number 68-02-4544, United States Environmental Protection Agency, May 1994.

R Reiss, PB Ryan, SJ Tibbetts, and P Koutrakis, Measurement of Organic Acids, Aldehydes, and Ketones in Residential Environments and Their Relation to Ozone. J. Air & Waste Manage. Assoc. 45:811-822 (1995).

Stock TH, DJ Kotchmar, CF Contant, PA Buffler, AH Holguin, BM Gehan, LM Noel, The Estimation of Personal Exposures to Air Pollution for a Community-Based Study of Health Effects in Asthmatics - Design and Results of Air Monitoring, JAPCA 35:1266-1273 (1985).

Suh HH, P Koutrakis, JD Spengler, The relationship between airborne acidity and ammonia in indoor environments, J. Exp. Anal. & Environ. Epi., 4,1: 1-23 (1994).

Thomas KW, ED Pellizari, CA Clayton, DA Whitaker, RC Shores, Particle Total Exposure Assessment Methodology (PTEAM) 1990 Study: Method Performance and Data Quality for Personal, Indoor, and Outdoor Monitoring, J. Exp. Anal. Environ. Epi., 3,2:203-226 (1993).

Spengler JD, RD Treitman, TD Tosteson, DT Mage, ML Soczek, Personal Exposures to Respirable Particulates and Implications for Air Pollution Epidemiology, Environ. Sci. Technol. 19:700-707 (1985).

Suh HH, P Koutrakis, JD Spengler, The Relationship Between Airborne Acidity and Ammonia in Indoor Environments, J. Exp. Anal. Environ. Epi., 4,1:1-23 (1994).

Wallace LA, Indoor Particles: A Review. J. Air & Waste Manage. Assoc. 46:98-126 (1996).

Zhang J, PJ Lioy, Ozone in Residential Air: Concentrations, I/O Ratios, Indoor Chemistry, and Exposures, Indoor Air, 4:95-105 (1994a).

Zhang J, Q He, PJ Lioy, Characteristics of Aldehydes: Concentrations, Sources, and Exposures for Indoor and Outdoor Residential Environments, Environ. Sci. Technol. 28:146-152 (1994b).

Zhang J, WE Wilson, PJ Lioy, Sources of Organic Acids in Indoor Air: A Field Study, J. Exp. Anal. Environ. Epi. 4:25-47 (1994c).

Table 7-1. Comparison of Study Data with California Ambient Standards and Indoor Guidelines.

Pollutant Variable	Standard' or Guideline"	Levels Observed In This Study	Number (and %) of Home Visits Exceeding Guidelines	Number (and %) of Homes Exceeding Guidelines
PM10				
indoor	50 ug/m3	2 to 294 ug/m3	19 (22%)	16 (27%)
outdoor	50 ug/m3	2 to 141 ug/m3	19 (21%)	17 (28%)
PM2.5'*'				
indoor	---	4 to 107 ug/m3	---	---
outdoor	---	2 to 77 ug/m3	---	---
O3				
indoor*	---	0.005 to 0.073 ppm*	---	---
outdoor*	0.09 ppm*	0.005 to 0.108 ppm*	---	---
HCHO				
indoor	0.10 ppm	0 to 0.032 ppm	0	0
outdoor	---	0.001 to 0.008 ppm	---	---

Notes:
'' California Ambient Air Quality Standards
" CARB recommended action level for indoor formaldehyde is 0.10 ppm, while recommended target level (goal to strive for) is 0.05 ppm.
'*' No California ambient air quality standard currently exists for PM2.5.
* O3 Standard is based on 1hr average, whereas this study's measurements were based on 24hr averages, so no direct comparisions are possible; However, 24hr O3 levels above 0.09 ppm were measured outside of 2 study homes.

Table 7-1. Comparison of Study Data with California Ambient Standards and Indoor Guidelines.

Pollutant Variable	Standard' or Guideline"	Levels Observed In This Study	Number (and %) of Home Visits Exceeding Guidelines	Number (and %) of Homes Exceeding Guidelines
PM10				
indoor	50 ug/m3	2 to 294 ug/m3	19 (22%)	16 (27%)
outdoor	50 ug/m3	2 to 141 ug/m3	19 (21%)	17 (28%)
PM2.5				
indoor	---	4 to 107 ug/m3	---	---
outdoor	---	2 to 77 ug/m3	---	---
O3				
indoor*	---	0.005 to 0.073 ppm*	---	---
outdoor*	0.09 ppm*	0.005 to 0.108 ppm*	---	---
HCHO				
indoor	0.05 ppm	0 to 0.032 ppm	0	0
outdoor	---	0.001 to 0.008 ppm	---	---

Notes:
' = California Ambient Air Quality Standards
" CARB recommended action level for indoor formaldehyde is 0.10 ppm, while recommended target level (goal to strive for) is 0.05 ppm. Because formaldehyde is a carcinogen, the ARB guideline recommends keeping indoor formaldehyde levels as low as possible.
= No California ambient air quality standard currently exists for PM2.5.
* O3 Standard is based on 1hr average, whereas this study's measurements were based on 24hr averages, so no direct comparisions are possible; However, 24hr O3 levels above 0.09 ppm were measured outside of two study homes, so the air quality outside two of the five homes monitored exceeded the State one-hour O3 standard.

Table 7-2. Housing and Activity Factors Associated with Indoor PM10 Levels exceeding 50 ug/m3.

Housing Factor	Survey Identifier*	Statistical Result"
Smoking		
# cigarettes smoked	FQ8X	0.005
resident smoker	C91	0.001
non-resident smoker	C93	0.001
Odors		
mold odors	TQ8A	0.008
cigarette odors	TQ8B	0.005
Combustion Activities		
cooking	FQ1H	0.006
wall furnace	FQ4A	0.039
stove use (other than cooking)	FQ12.4	0.063
Air-Moving Activities		
dusting	FQ1B	0.059
window fan	FQ2N	0.063
dryer vented into home	FQ16	0.03
air conditioning type	C69	0.015
Building Construction		
foundation type	TQ6	0.022
recent remodel	FQ17	0.038
attached garage	C74	0.054

Notes:

* Refers to specific question in particular survey, as follows:
  TQ = Technician Questionnaire
  FQ = Follow-Up Questionnnaire
  C = Baseline Questionnaire

" Chi Square p-value results

8. SUMMARY AND CONCLUSIONS

This project collected monitoring and survey information in 126 Southern California homes to learn about the potential importance of various housing factors (including air exchange rates) and indoor/outdoor levels of ozone, PM_{10}, $PM_{2.5}$, and formaldehyde. In addition to direct measurement of O_3, PM_{10}, $PM_{2.5}$, formaldehyde, and air exchange rate, three types of survey questionnaires were collected to characterize the building construction, air handling, and operational factors of study residences.

Homes from the communities of Lake Gregory/Lake Arrowhead, Riverside/Mira Loma, San Dimas, and Lancaster (all in the greater Los Angeles metropolitan area) participated in the study. Eligibility to participate was based upon prior enrollment of children living in the participating home in a concurrent CARB-sponsored epidemiologic investigation of chronic respiratory health effects of air pollution in Southern California school children. In any given community, selection of actual study homes was based upon stratification of air conditioning, willingness to participate, and home accessibility. Each home was studied on two separate occasions, separated by at least 45 days, to permit sampling of the home under different ambient seasonal conditions. Home studies were conducted between February 1994 and November 1994.

8.1 Ozone

Observed indoor 24hr average concentrations of ozone ranged from below the detection limit of 5 ppb to over 73 ppb (median value of 6 ppb, with an inter-quartile range of 2 to 16 ppb), as measured by a nitrite-coated filter method. Outdoor 24hr average ozone levels were observed from below the detection limit of 5 ppb to over 108 ppb. Observed 24hr average ozone concentrations indoors varied among the four communities from 0.010 ppm in San Dimas homes to 0.017 ppm in Lake Gregory homes, a reflection of the higher outdoor levels also observed in Lake Gregory. Homes with central or room air conditioning had slightly lower indoor ozone levels (averaging 0.009 ppm over 24hrs, compared to 0.016 ppm) than homes with swamp cooling or no air conditioning, but there was a overlapping range of observed values.

Indoor/outdoor ozone ratios in homes studied varied over a range of values (median percentile value of 0.20, inter-quartile range of 0.07 to 0.45). Indoor/outdoor ozone ratios varied with season, with I/O ratios increasing with increasing summertime ambient ozone levels. Based on these data, the strategy of staying indoors during elevated outdoor ozone episodes does offer a reasonable means of protection from exposure, since observed indoor concentrations of ozone were lower than outdoor levels. Based upon the measurements made in this study, indoor ozone levels were likely to be 10% to 50% of simultaneously measured outdoor levels.

The lack of any notable indoor ozone sources made prediction of residential ozone concentrations based on community station data a technical possibility. Analysis of the collected data yielded somewhat mixed results. In many sampling locations, where regional ozone levels were more uniform and community stations were appropriately sited, correlation between ozone measurements made immediately outside the home and at the community station were relatively good (R^2 of 0.65 to 0.75). However, in non-metropolitan sampling locations, such as in mountain resort areas (Lake Gregory) or high desert communities (Lancaster), correlations were lower (R^2 of 0.43 to 0.60). Clear explanations for the observed poor correlation in these non-metropolitan communities (including problems with community station siting and potential measurement interference caused by vegetation canopies or building wake effects around residences) are suspected but unconfirmed at the present time.

The model that was developed, based on the collected data, to predict indoor levels of ozone indicated that indoor ozone was largely determined by the outdoor concentration and the duration of time that any home windows were left open. Use of a central refrigerant recirculating air conditioner was associated with lower indoor levels of ozone. Community monitoring station ozone data were modestly correlated ($R^2 = 0.49$) with indoor O_3 levels. Information about immediate past household use (such as length of time windows were left open, or whether an air conditioner was used) only marginally improved the correlation (from an R^2 of 0.49 to 0.55). Ozone levels measured just outside the home were no better for predicting indoor levels than were concentrations reported from the community air monitoring

station.

Based on the information collected in this study, residential ozone exposures have likely been overestimated if central station data have been used without adjustment for the probable reduction in indoor levels compared to outdoor values. Furthermore, based on the relationships observed in this study between community station ozone measurements and measurements made immediately outside residents' homes, the agreement between measured community station ozone levels and ozone concentrations in the surrounding residential neighborhoods may vary subtantially.

8.2 PM_{10}

Measurements of PM_{10} inside study homes (collected as 24hr averages) ranged from a few $\mu g/m^3$ to almost 300 $\mu g/m^3$ (median value of 32.9 $\mu g/m^3$, inter-quartile range of 24.2 to 47.2 $\mu g/m^3$). Measured outdoor PM_{10} levels were observed from a few $\mu g/m^3$ to 141 $\mu g/m^3$ (median value of 29 $\mu g/m^3$, inter-quartile range of 18.1 to 44.2 $\mu g/m^3$.

Indoor levels of PM_{10} were comparable to, and often higher than, corresponding outdoor levels. No seasonal relationship was detected in the collected PM_{10} data for the portion of the year (summer to fall) sampled. On average, homes in the Riverside/Mira Loma area had both higher indoor PM_{10} measurements and a higher range of indoor values (As a group, Riverside/Mira Loma homes also had higher outdoor PM_{10} values as well). The median I/O PM_{10} ratio observed during this study was 1.05 (inter-quartile range of 0.7 to 1.9). This result suggests that for many homes, in terms of particulate mass, indoor levels of PM_{10} are equal to or larger than corresponding outdoor PM_{10} concentrations.

However, a small number of homes were found with significant indoor sources, often related to smoking in the home. In fact, the most important factor for determining indoor levels of PM_{10} was smoking. Knowledge about the community station PM_{10} level and a categorical variable indicating the number of cigarettes smoked were the two most important variables in modeling home PM_{10}, yet this model was only partially successful in describing

the observed data (R^2 of 0.40). Additional variables did not especially enhance modeling results. Modeling efforts suggested that smoking more than ten cigarettes in the home in the previous 24hrs was associated with a 67 $\mu g/m^3$ increase in indoor PM_{10}.

In non-smoking homes, a multi-variable model was also developed to predict indoor PM_{10}. The selected model was moderately successful in predicting indoor PM_{10} levels (R^2 = 0.55).

Correlations between 24hr PM_{10} concentrations observed at community station monitors and samples collected immediately outside residents' homes were generally low, and ranged from 0.13 in Lancaster to 0.56 in Riverside. Mean PM_{10} concentrations measured immediately outside homes were consistently lower than those levels simultaneously recorded at the corresponding community monitoring stations.

Possible explanations for the apparent uncoupled behavior between station and residential PM levels include potential sampling artifact complications caused by localized wall effects around the home or vegetation canopies surrounding the home, and differences in this study between measurement methodology at community and residential sites. Even if community and outdoor residential PM correlations were higher, the relative importance and variability of indoor PM sources in the assessment of overall human PM burdens around the home remain an issue to be addressed. Therefore, prediction of residential PM exposure levels based on community monitoring information, at this time, does not appear to be reliably possible.

In approximately one-fourth of the homes in which PM_{10} sampling was performed, observed levels exceeded the CARB Indoor Air Quality Guidelines and the California Ambient Air Quality Standard of 50 $\mu g/m^3$. The activity and housing factors associated with these elevations included smoking, the presence of mold odors, combustion activities in the home (such as cooking or wall furnace use), air-moving activities (such as dusting or air conditioning), and building construction factors (foundation type, recent remodeling, an

attached garage).

8.3 $PM_{2.5}$

Indoor levels of $PM_{2.5}$ (median of 13.7 $\mu g/m^3$, inter-quartile range of 9.7 to 22.5 $\mu g/m^3$) appeared to be related to outdoor levels (outdoor $PM_{2.5}$ median value 10.7 $\mu g/m^3$, inter-quartile range of 7.3 to 19.6 $\mu g/m^3$) for a large fraction of homes, but a number of homes were also observed with significant indoor $PM_{2.5}$ sources. In this study, the median indoor/outdoor $PM_{2.5}$ ratio was 1.10 (inter-quartile range of 0.84 to 1.68), suggesting the presence of significant indoor sources.

On average, $PM_{2.5}$ levels in Riverside/Mira Loma homes were higher than those in other communities, and outdoor $PM_{2.5}$ levels were also higher. As was the case with PM_{10}, no seasonal relationship was observed in the collected $PM_{2.5}$ data for the portion of the year (summer to fall) sampled.

The $PM_{2.5}$ data set was divided into smoking and non-smoking homes, based on the demonstrated influences of smoking as a model variable of importance. Only 17 sample pairs were collected in smoking homes, so modeling efforts were focused on the non-smoking subset. In addition to ambient $PM_{2.5}$ concentration data, the duration of stove use and an indicator for a two-bedroom residence were the explanatory variables selected by the model (R^2 =.64). Use of $PM_{2.5}$ data collected immediately outside the home, rather than station data, did not improve the model. However, the limited number of $PM_{2.5}$ data in non-smoking homes (47 sampling pairs) may have made the model selection procedure somewhat unstable and overestimated the coefficient of determination (R^2).

Attempts to correlate observations collected immediately outside of residential sampling sites with community monitoring data were largely unsuccessful. The likely reasons for this uncoupled behavior include those proposed above for PM_{10} and the fact that, since community station $PM_{2.5}$ monitoring information is not usually collected, the data used to perform the correlation analyses in this investigation were generated from available PM_{10}

information (which, as described in Section 7.2 above, correlated poorly).

8.4 Formaldehyde

Formaldehyde levels were quite low in most homes (median value of 10.1 $\mu g/m^3$, inter-quartile range of 6.5 to 15.2 $\mu g/m^3$) but higher than measured outdoor levels (median value of 3.2 $\mu g/m^3$, inter-quartile range of 1.7 to 4.1 $\mu g/m^3$). Formaldehyde levels did not vary very much across communities, although the data suggested that on average, homes in the Riverside/Mira Loma area had slightly higher levels than homes elsewhere (15 $\mu g/m^3$, compared to about 10 $\mu g/m^3$ in the other communities). No conclusive differences were observed with regard to formaldehyde as a function of reported air conditioning type.

Although formaldehyde sampling was admittedly limited in scope in this effort, the uniformly low indoor concentrations which were observed suggested that formaldehyde exposure to elevated levels indoors was a rare event for the housing stock distribution evaluated. The median indoor/outdoor ratios of formaldehyde observed in this study, based on the 18 comparative measurements made, was 3.75 (with an inter-quartile range of 1.98 to 7.38).

8.5 Housing Factors

The median air exchange rate observed was about 0.7 hr^{-1} (using the traditional home volume calculation approach or the corrected home volume calculation approach) across a wide range of Southern California homes. As might have been predicted, AERs in homes with swamp coolers were higher than in homes with central or no air conditioning.

Some evidence was present to suggest that AERs were lower during the spring and fall compared to the summer portion of the year (no measurements were made during the winter). AER was more variable and higher during the summer, perhaps reflecting resident preference for natural ventilation (opening windows) over the use of home air conditioning.

Air exchange rate contributed little in the way of predictive indoor pollutant levels, if information about housing characteristics and the presence of smoking in the home was available. This is a useful observation, since the documentation of housing operation is potentially more feasible than the costs associated with the collection of air exchange rate information. The air exchange rate measurement itself contains inherent assumptions that limit its operational usefulness, especially with regard to assumptions or attainment and maintenance of equilibrium.

8.6 Pilot Assessment of the Two-Week Sampler

Sampling in residences for airborne levels of acids and fine particle chemistry, using the Two-Week Sampler, was demonstrated to be feasible and provided interesting observations for future resolution. The small number of homes sampled in this pilot demonstration necessarily limit the implication of specific measurement results. Nevertheless, the suggestion of elevated indoor concentrations of organic acids in some homes may warrant future investigation.

The potential for predicting residential acid exposure levels, based upon data collected at community monitoring sites, remains to be addressed, pending the collection of a larger (and thus, more representative) residential data set, with concurrent community monitoring data, in some future investigation.

8.7 General Comments

The attempt to develop a model that accurately predicts indoor concentrations of pollutants, based on fixed-site ambient monitoring data, surveyed housing factors, and interviews with residents about activities in the home and operation of the home during sampling, met with moderate success. A wide range of housing factor information was collected in the performance of this study, of which only a portion appeared to be useful for model refining purposes.

The information collected in the performance of this project provides residential exposure data for the exposure assessment of study subjects participating in a long-term epidemiological study of Southern California school children. Improved exposure characterization of the residential microenvironment is a necessary component of any successful human exposure assessment effort, since most people, especially children, spend much of their time indoors. Due to the lack of a validated personal ozone sampler, however, this project could not address issues of personal exposure. Information from that component of human exposure assessment would allow progress in the validation and refinement of human exposure models, such as REHEX, for use in large-scale epidemiological investigations. Improvements in human exposure models such as REHEX can be made using the data collected in this study, but is limited by model requirements for one hour time-resolved ozone data and the data set's applicability of the PM modeling discussion to non-smoking residences.

9. RECOMMENDATIONS

Based on this study and the conclusions drawn from it, the following recommendations are made:

1) **Personal O_3 and PM measurements are needed to provide information for use in modeling efforts.** Information of this type is critical for the improvement and validation of regional exposure modeling, would provide important information to compare personal exposure levels to indoor and outdoor concentrations, and could provide insights into the microenvironmental exposure of children and adults living in homes with high PM or ozone levels, or helping to explain significant changes in health indices in children participating in the CARB-supported Children's Health Study.

2) **Additional study is needed to account for the observation of lower PM loadings reported immediately outside homes compared to those measurements reported from community monitoring sites.** Assessment of the relative importance of potential explanatory factors, such as vegetation canopies or localized wall effects around the home, should be resolvable with a focused sampling study designed to unravel this apparent difference.

3) **The apparent disparity in community monitoring information between and among specific community stations (such as Lake Gregory and Lake Arrowhead, Riverside and Rubidoux, and Lancaster) should be investigated.** If confirmed, corrective action, up to and including possible relocation of existing stations, should be considered.

4) **The archived filters collected in the course of PM monitoring for this study should be chemically analyzed.** Gravimetric comparisons of indoor and outdoor levels may accurately reflect relationships, but chemical speciation of sulfur and other constituents could help to more definitively establish the contribution of indoor sources to observed PM levels.

5) **Information about the strength, nature, and source of the relationship between outdoor and indoor PM_{10} and $PM_{2.5}$ levels, in varying modes of home operation (such as the home sealed during the day, or with windows left open) is needed to improve human PM exposure assessment.** A sampling study indoor and outside of a population of homes operated in a limited and directed representative number of modes, would provide the needed information.

5) **The Two-Week Sampler (TWS) is a viable and cost-effective platform for the collection of information about the distribution of airborne acids and fine particle chemistry in and around homes.** Based on the pilot demonstration sampling performed in this project, a full-scale study should be planned and performed to evaluate the potential for prediction of longer-term residential acid exposures based on community monitoring information. The suggestion of elevated indoor levels of formic and acetic acids, based on the limited pilot sampling data collected in this study, could also be addressed in the course of such an investigation.

10. GLOSSARY

24hr - twenty-four hour

37mm - thirty-seven millimeter

AER - air exchange rate

CARB - California Air Resources Board

CATs - capillary adsorption tubes (for AER)

CHS - Children's Health Study

DNPH - 2,4-dinitrophenylhydrazine

EDMI - electronic distance measuring instrument

EPA - United States Environmental Protection Agency

GRI - Gas Research Institute

HCHO - formaldehyde

HCl - hydrochloric acid

HiVol - high volume air sampler

HNO_3 - nitric acid

HPLC - high performance liquid chromatography

hr - hour

hr^{-1} - 1/hr, the unit of air exchange rate

I/O - indoor/outdoor ratio

inter-quartile - values between the 25th and 75th percentiles

l^3 - volume, in cubic liters

LAN - Lancaster

LG - Lake Gregory

LOD - limit of detection

ML - Mira Loma

MMAD - mass median aerodynamic diameter

mm - millimeter

m^3 - volume, in cubic meters

NAS - National Academy of Science

NH_4^+ - fine particle ammonium

NO - nitric oxide

NO_2 - nitrogen dioxide

NOx - oxides of nitrogen

NO_3^- - fine particle nitrate

O_3 - ozone

PAN - peroxyacetyl nitrate

PEM - personal exposure monitor (a device for particle sampling)

PFT - perfluorocarbon tracer

PM - particulate matter

PM - clock time, between noon and midnight

PM_{10} - particle mass with an MMAD of 10 microns or smaller

$PM_{2.5}$ - particle mass with a MMAD of 2.5 microns or smaller

PMCH - perfluoromethylcyclohexane, a tracer gas

ppb - parts per billion

PTEAM - Particle Total Exposure Assessment Methodology

REHEX - Regional Human Exposure Model

RIV/ML - Riverside/ Mira Loma

RMSE - root mean square error

RSD - relative standard deviation

SCAQMD - South Coast Air Quality Management District

SD - standard deviation

$SO_4^=$ - fine particle sulfate

STI - Sonoma Technology Incorporated

TEAM - Total Exposure Assessment Methodology

TED - Timed Exposure Diffusion (sampler for ozone)

TEOM - Tapered Element Oscillating Microbalance

TWS - Two-week sampler for fine particulate matter and acids

VAC - voltage on alternating current; house line voltage

VOCs - Volatile Organic Compounds

$\mu g/m^3$ - micrograms per cubic meter

11. APPENDICES

A. Baseline Questionnaire (Printout of Selected Sections)

B. Technician Questionnaire and Printout

C. Follow-Up Questionnaire and Printout

D. Data Reduction

E. Quality Control

F. Convenience Home Sampling Report

G. Ten-Home Sampling Report

H. Summary of P-Values for Predicting Indoor O_3, PM_{10}, and $PM_{2.5}$