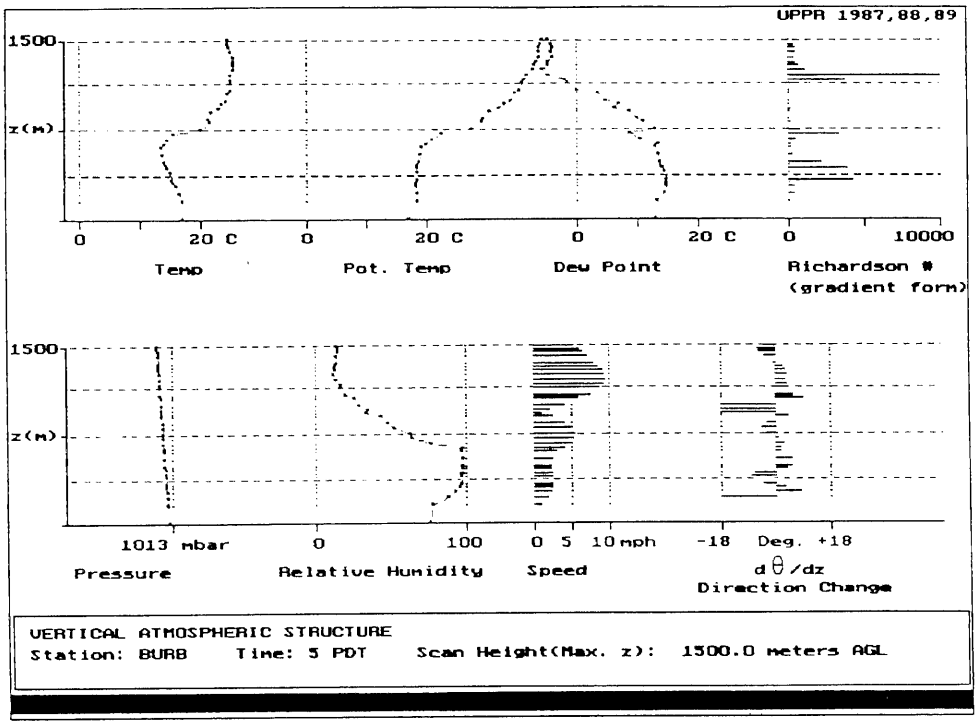
A2.2 Appendix

The collection of 68 figures to follow were generated by the program UPPR in order to tabulate mixing layer heights and provide a qualitative picture for the atmospheric structure above the surface. The figures were reduced to the smallest size which allowed them to be of value for extracting quantitative data. For exact values of the measurements, consult the California Air Resources Board data sets described earlier.



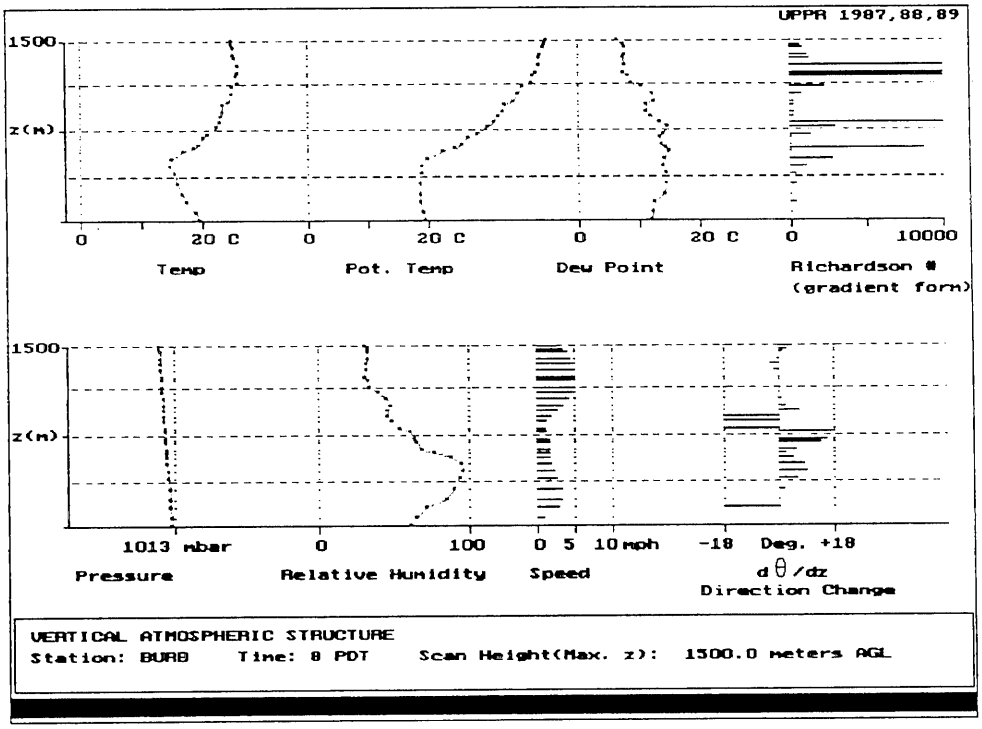
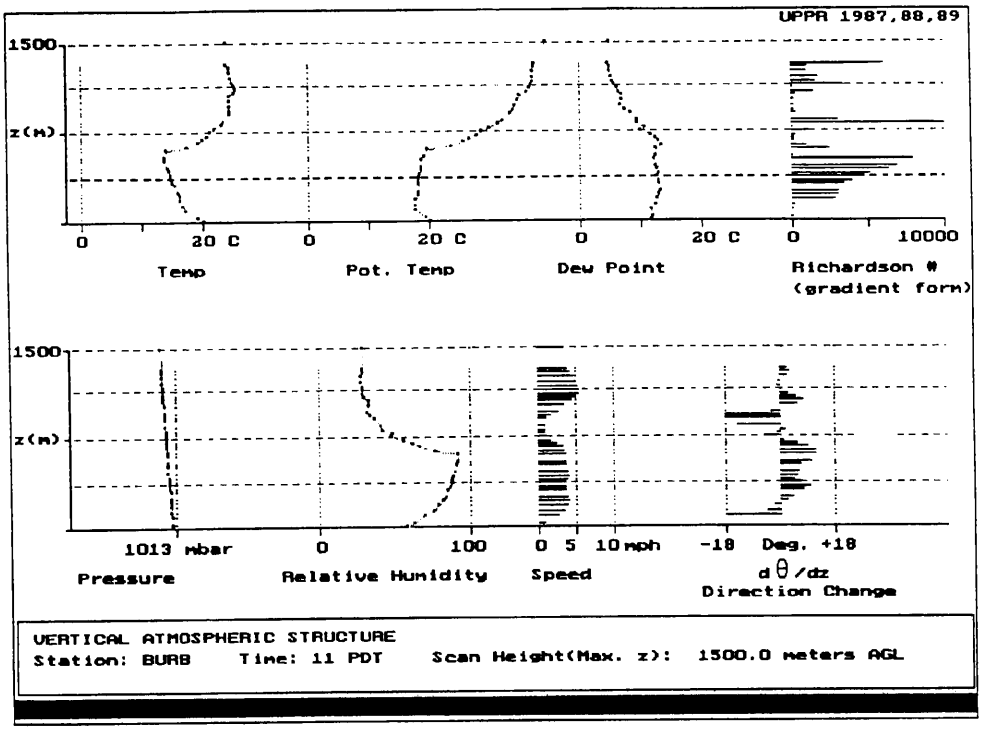


Figure 1. (Top) Vertical atmospheric structure at Burbank at 05:00 PDT, July 15, 1987.

Figure 2. (Bottom) Vertical atmospheric structure at Burbank at 08:00 PDT, July 15, 1987.



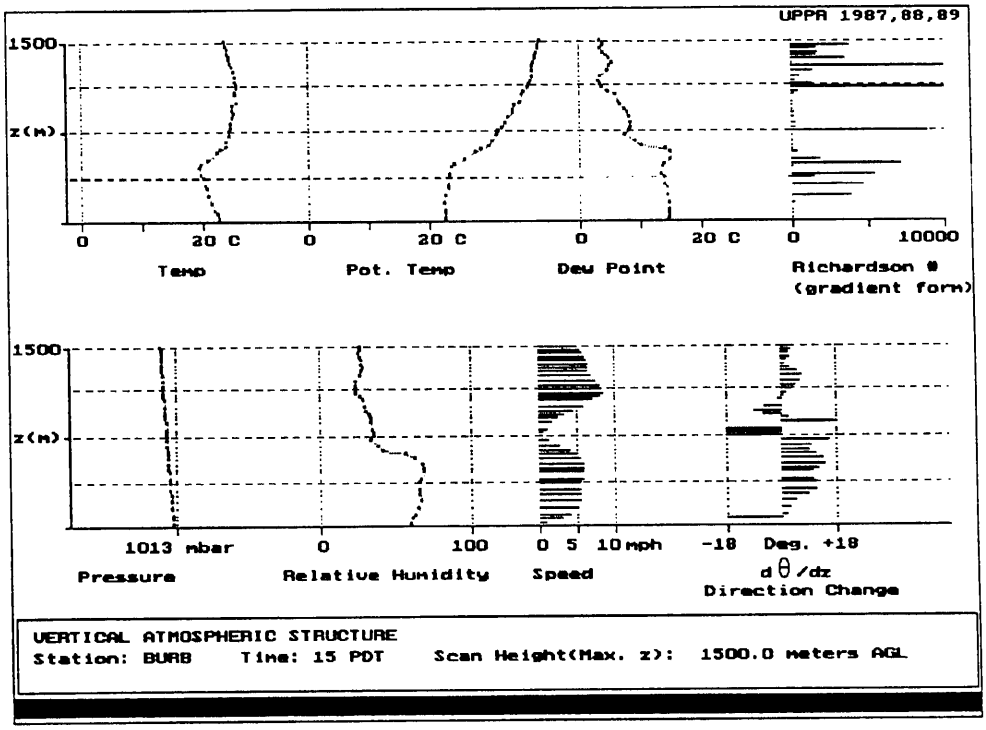
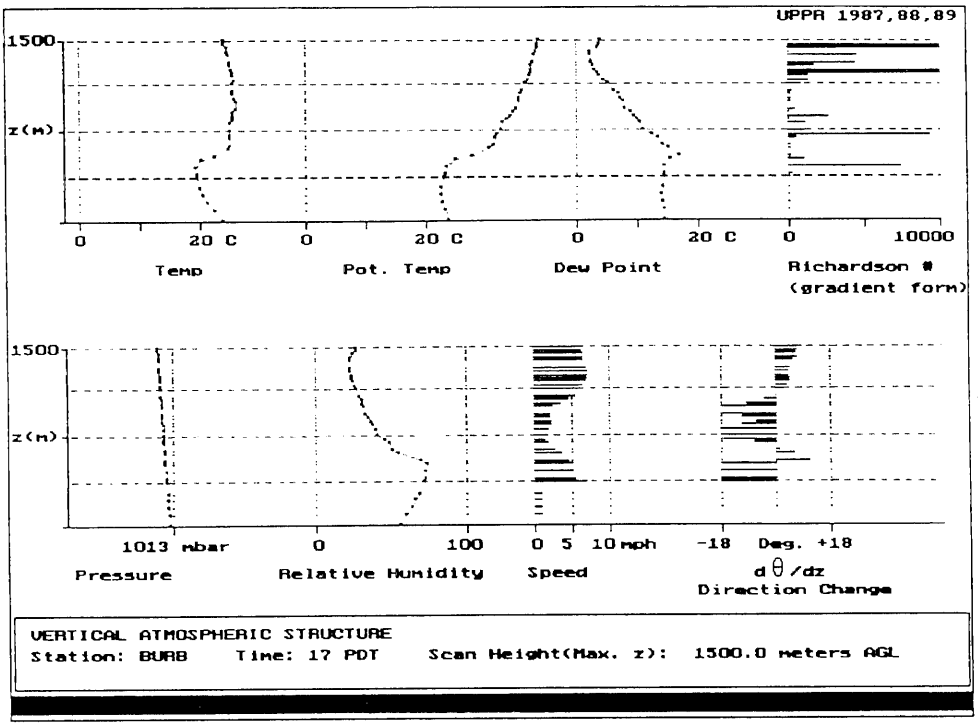


Figure 3. (Top) Vertical atmospheric structure at Burbank at 11:00 PDT, July 15, 1987.

Figure 4. (Bottom) Vertical atmospheric structure at Burbank at 15:00 PDT, July 15, 1987.



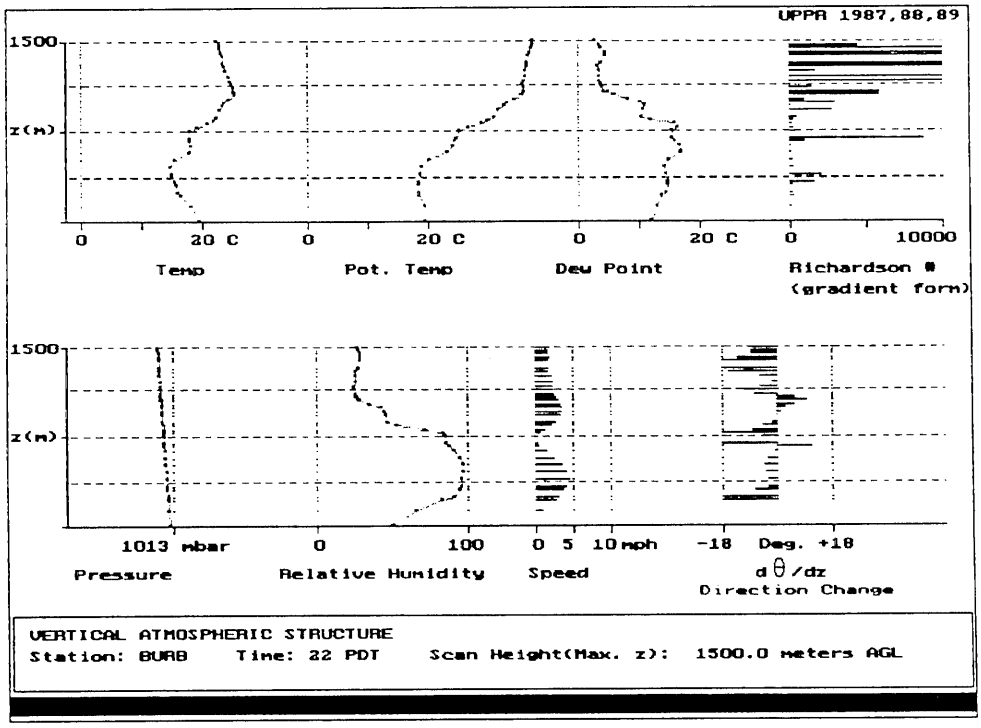
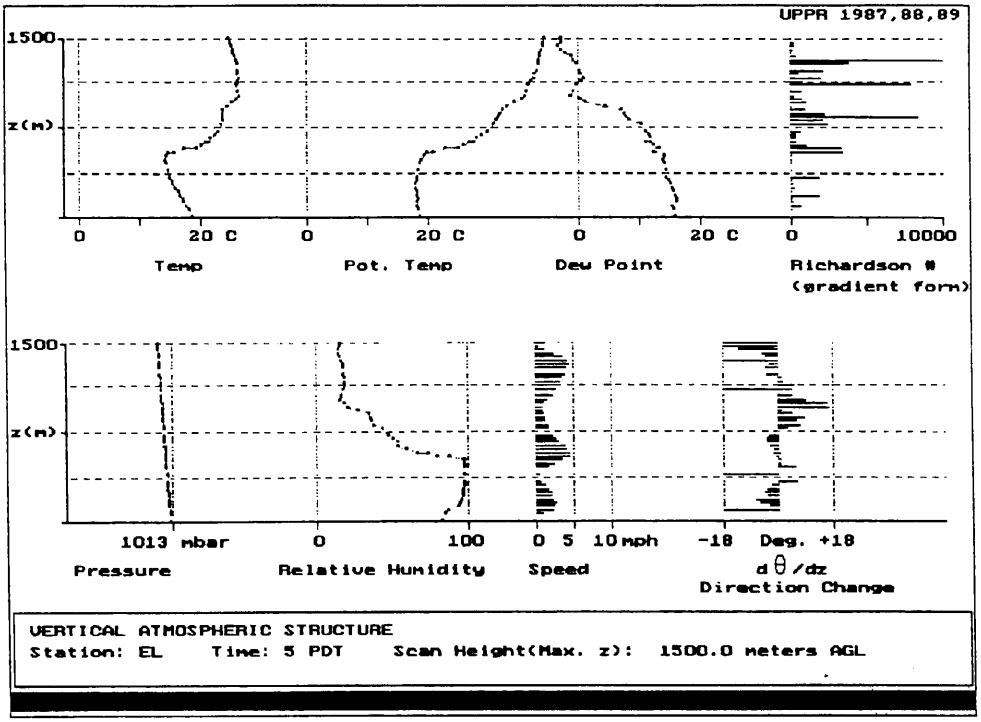


Figure 5. (Top) Vertical atmospheric structure at Burbank at 17:00 PDT, July 15, 1987.

Figure 6. (Bottom) Vertical atmospheric structure at Burbank at 22:00 PDT, July 15, 1987.



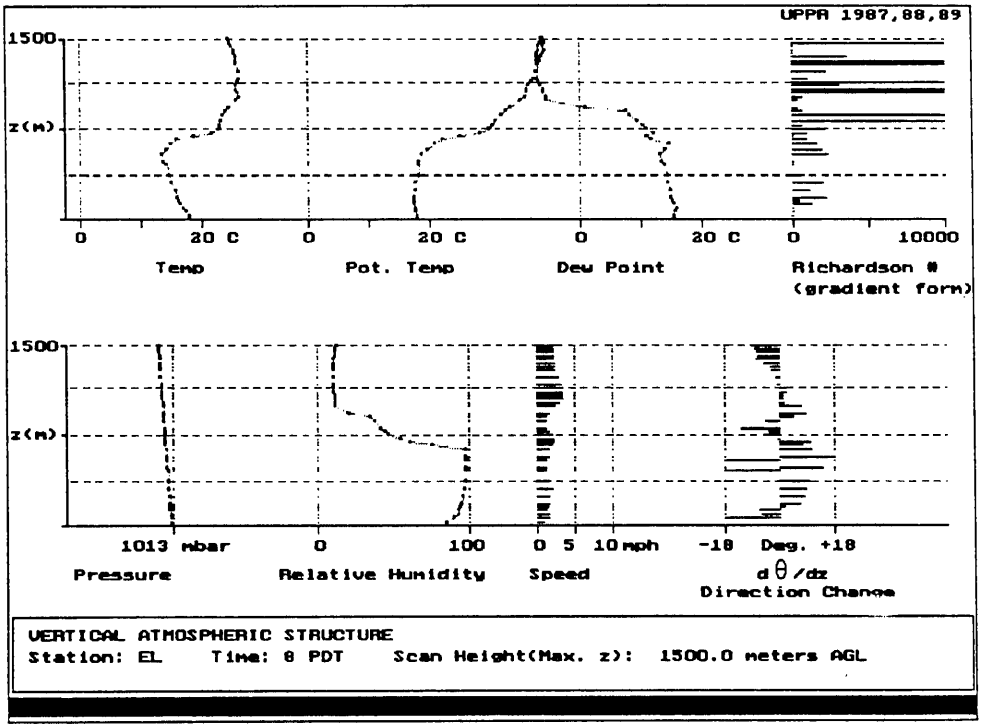
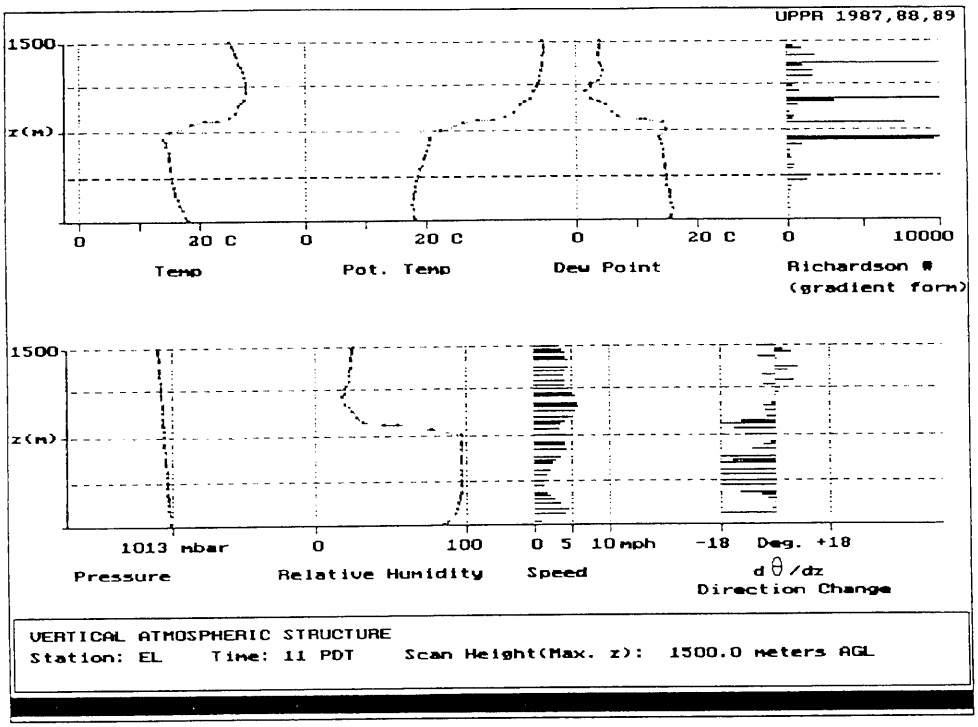


Figure 7. (Top) Vertical atmospheric structure at El Monte at 05:00 PDT, July 15, 1987.

Figure 8. (Bottom) Vertical atmospheric structure at El Monte at 08:00 PDT, July 15, 1987.



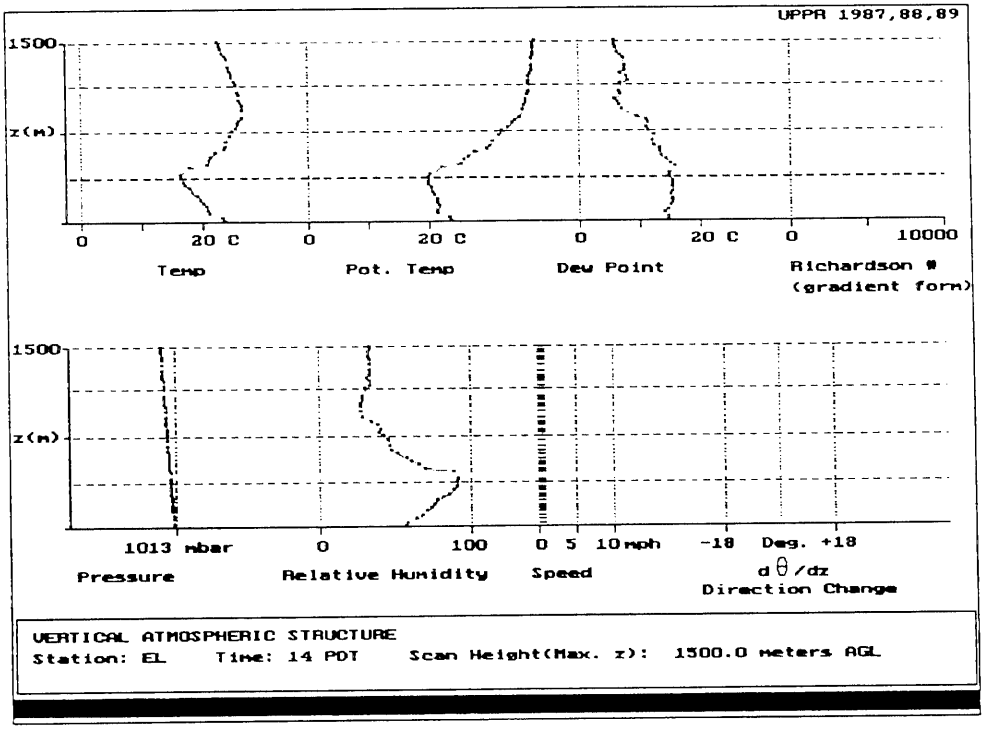


Figure 9. (Top) Vertical atmospheric structure at El Monte at 11:00 PDT, July 15, 1987.

Figure 10. (Bottom) Vertical atmospheric structure at El Monte at 14:00 PDT, July 15, 1987.

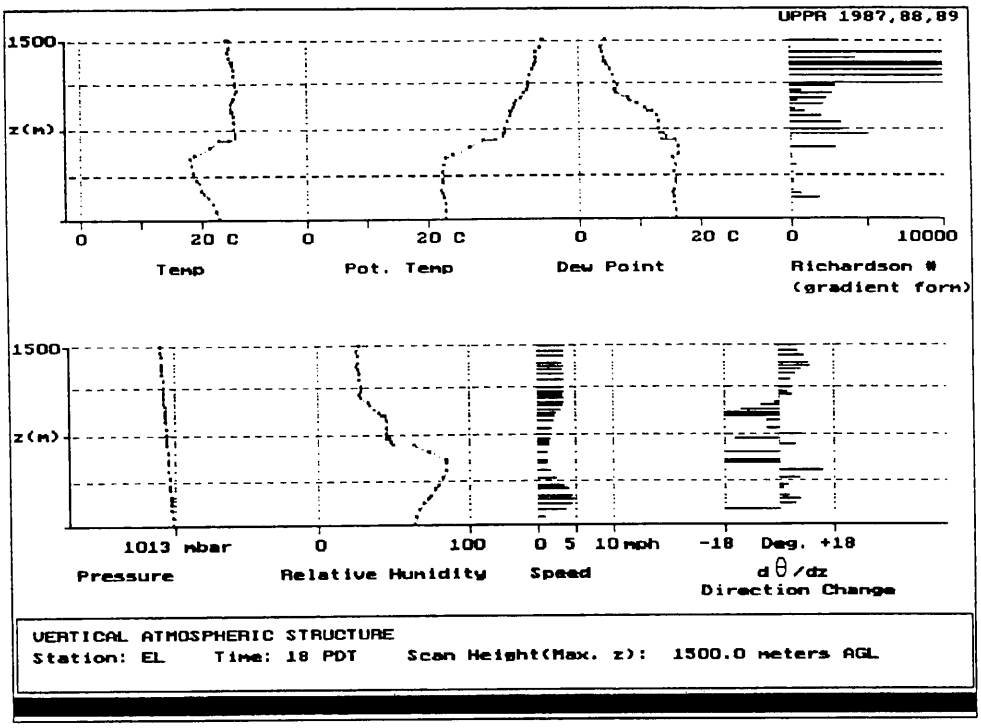
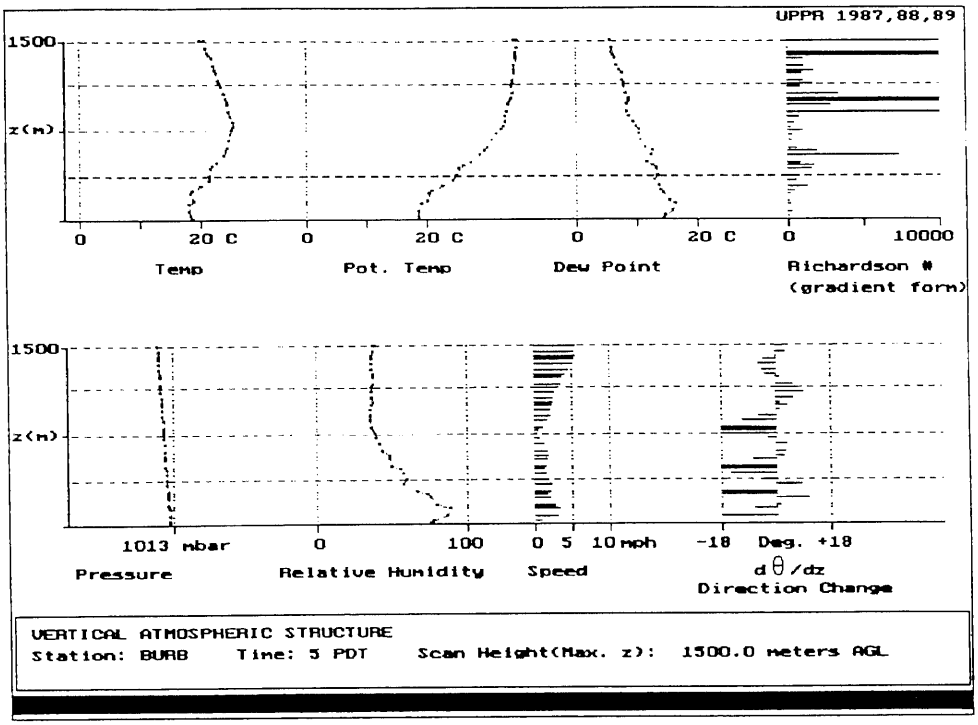


Figure 11. Vertical atmospheric structure at El Monte at 18:00 PDT, July 15, 1987.



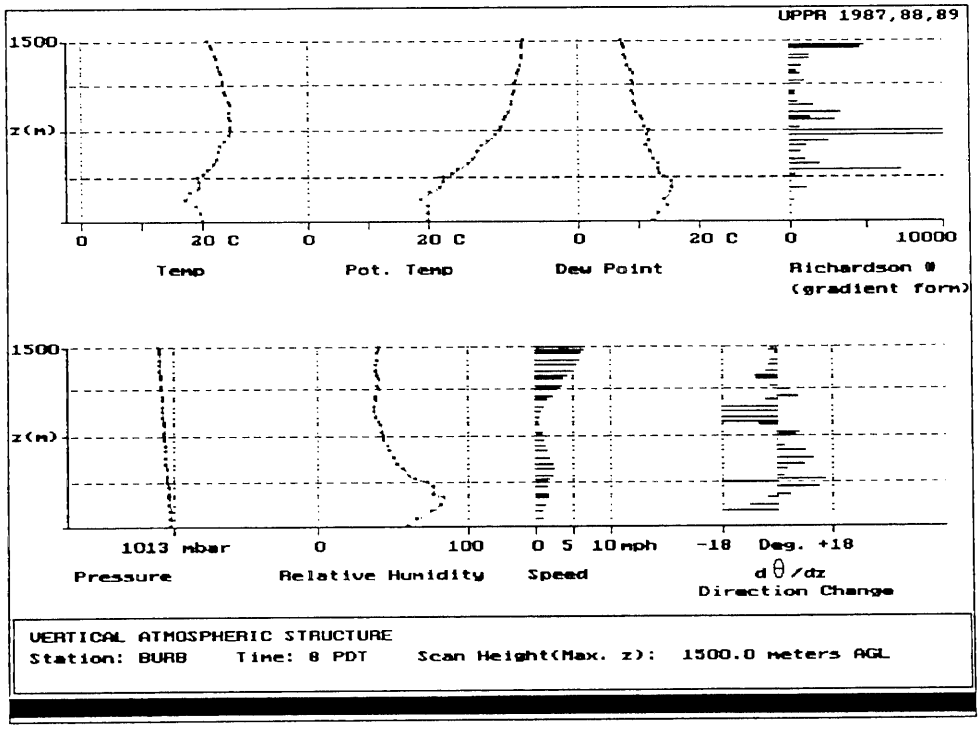
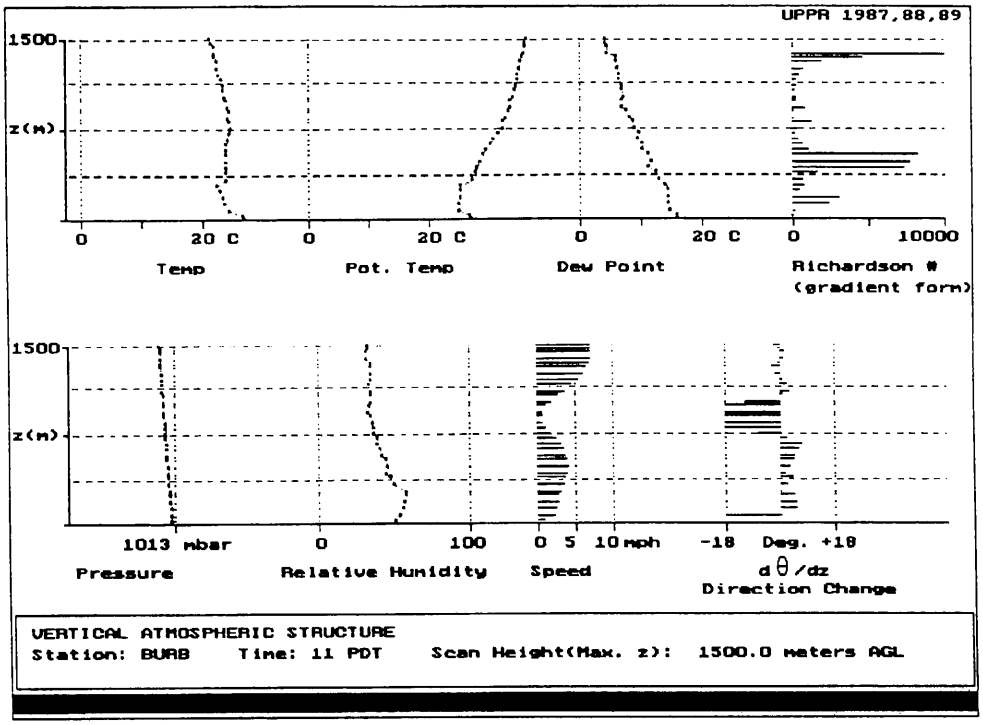


Figure 12. (Top) Vertical atmospheric structure at Burbank at 05:00 PDT, September 3, 1987.

Figure 13. (Bottom) Vertical atmospheric structure at Burbank at 08:00 PDT, September 3, 1987.



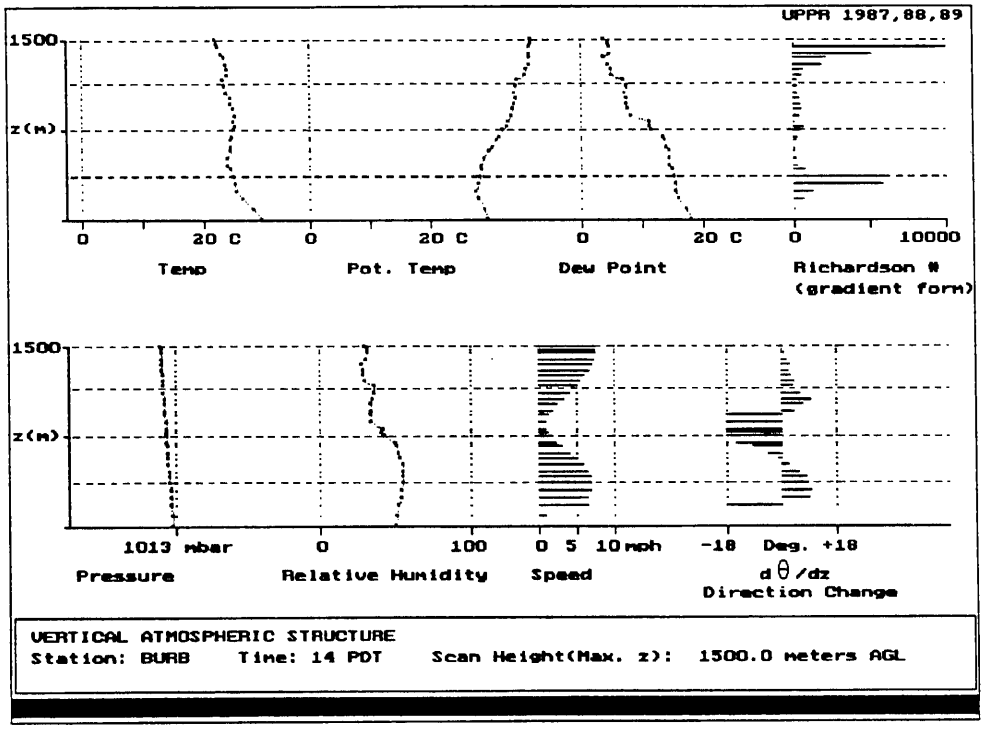
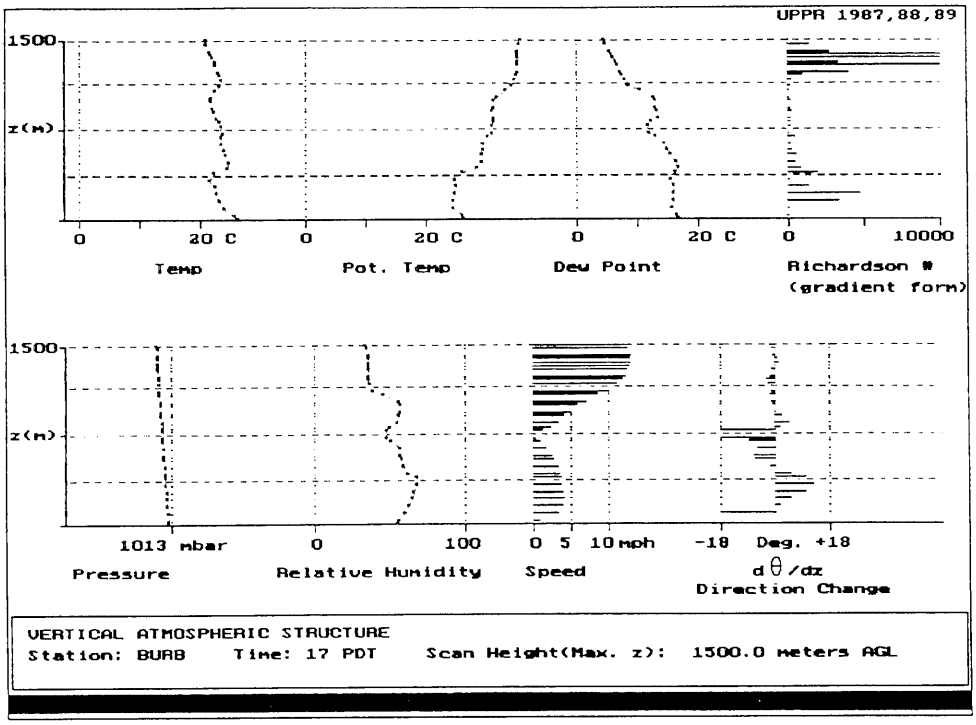


Figure 14. (Top) Vertical atmospheric structure at Burbank at 11:00 PDT, September 3, 1987.

Figure 15. (Bottom) Vertical atmospheric structure at Burbank at 14:00 PDT, September 3, 1987.



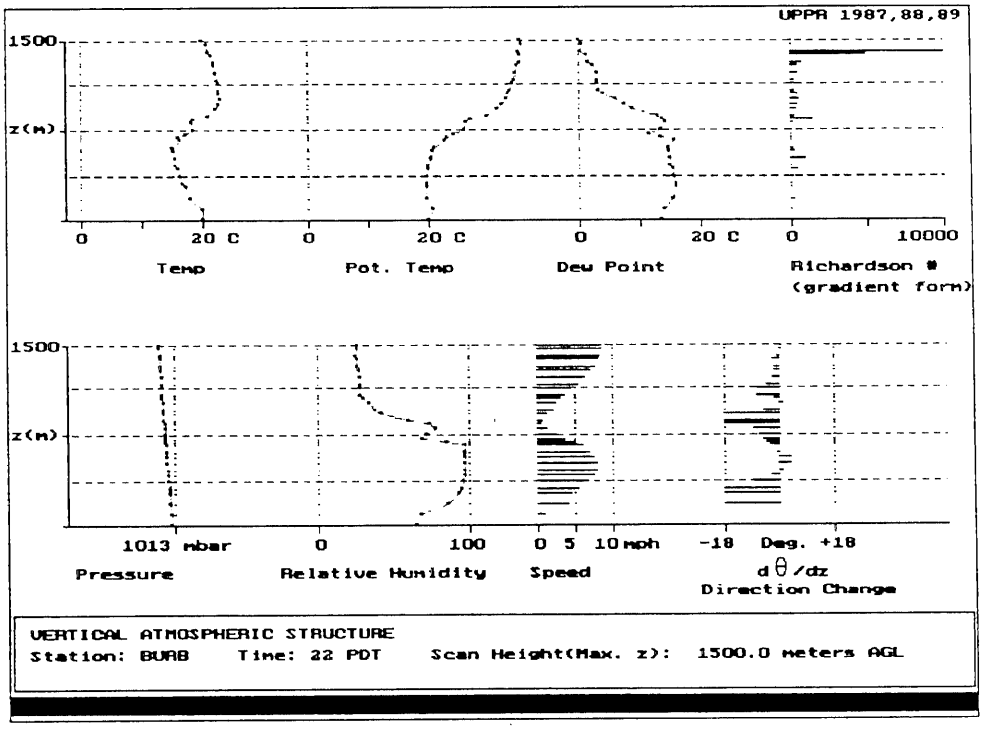
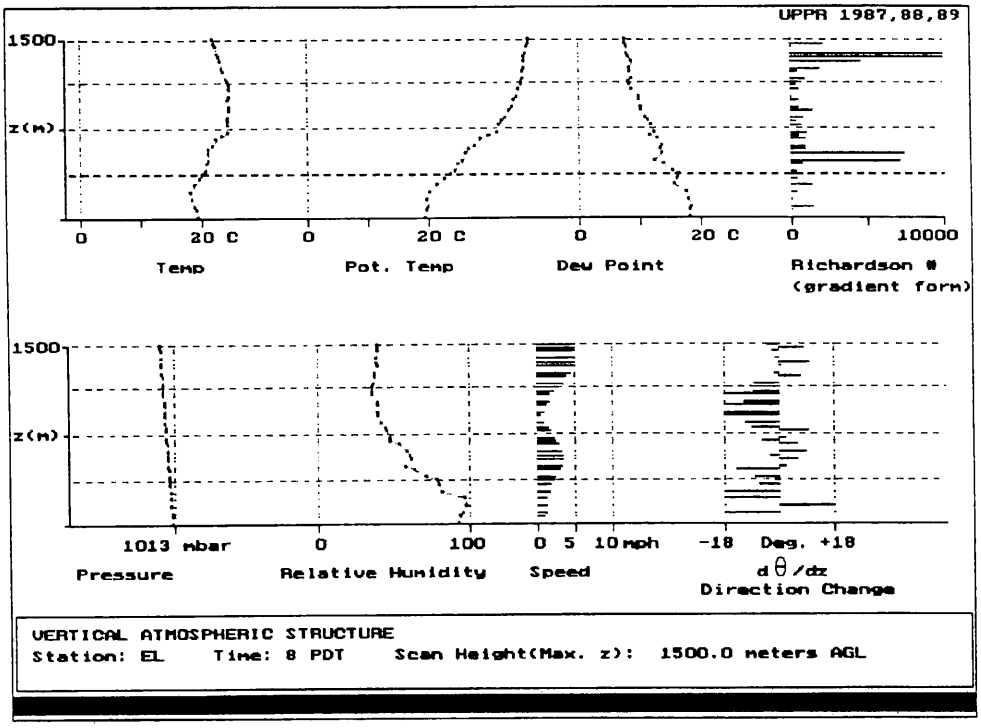


Figure 16. (Top) Vertical atmospheric structure at Burbank at 17:00 PDT, September 3, 1987.

Figure 17. (Bottom) Vertical atmospheric structure at Burbank at 22:00 PDT, September 3, 1987.



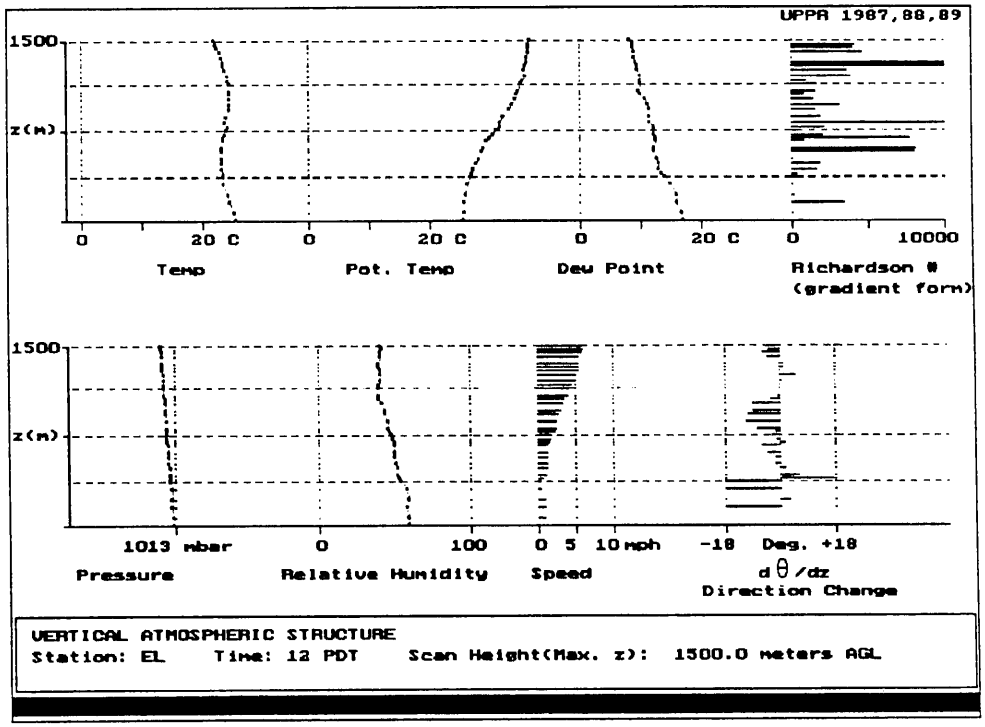
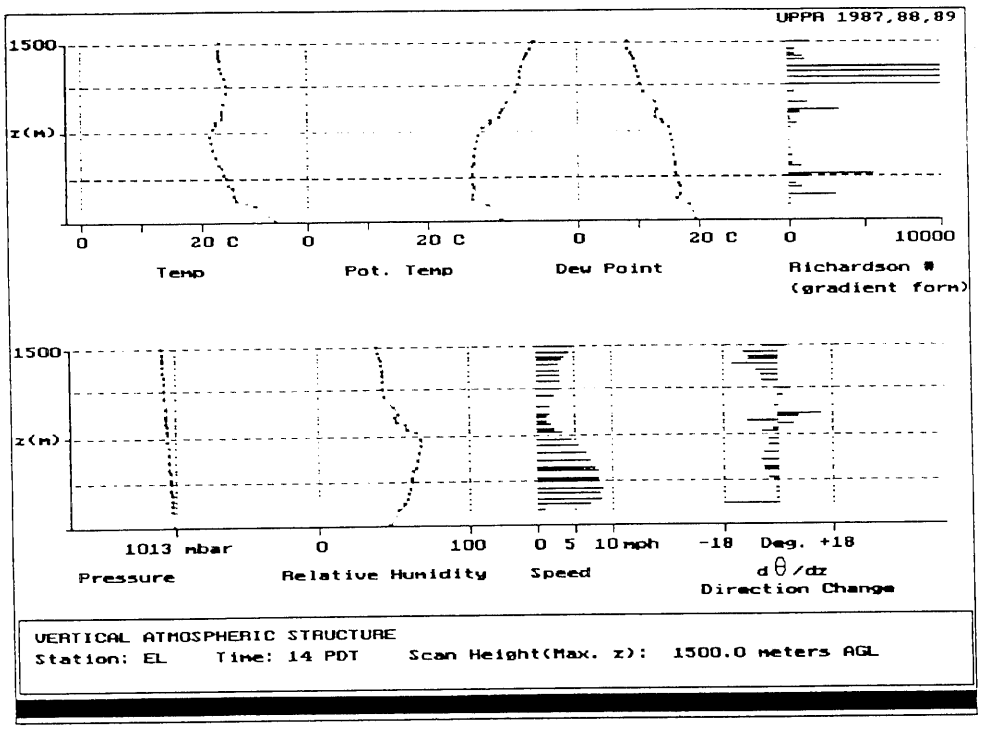


Figure 18. (Top) Vertical atmospheric structure at El Monte at 08:00 PDT, September 3, 1987.

Figure 19. (Bottom) Vertical atmospheric structure at El Monte at 12:00 PDT, September 3, 1987.



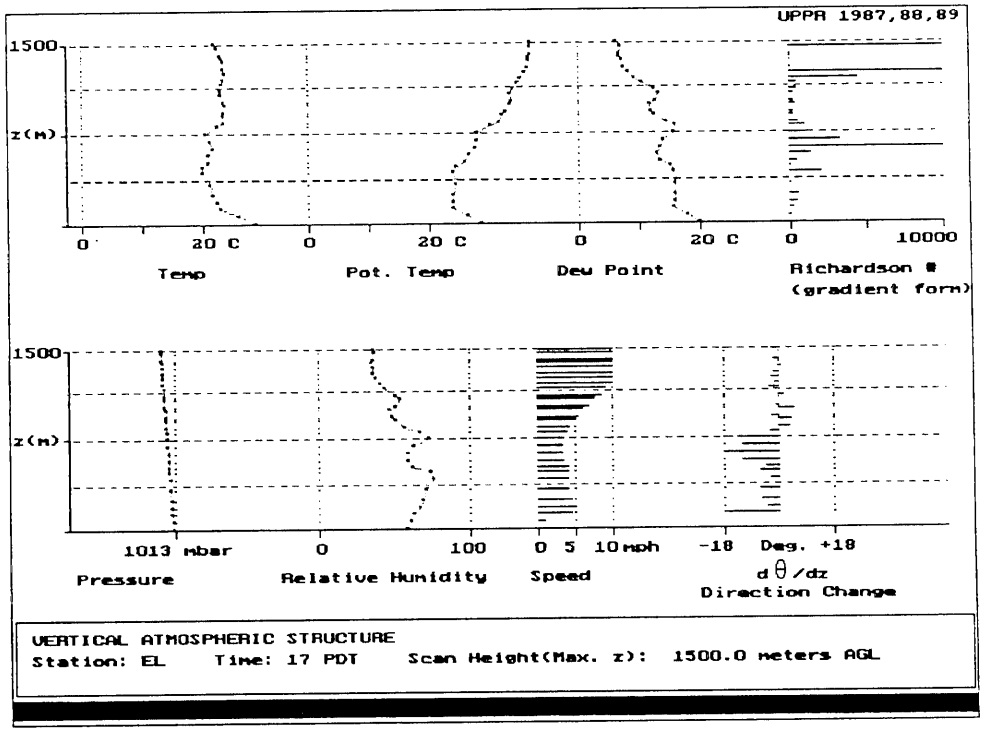
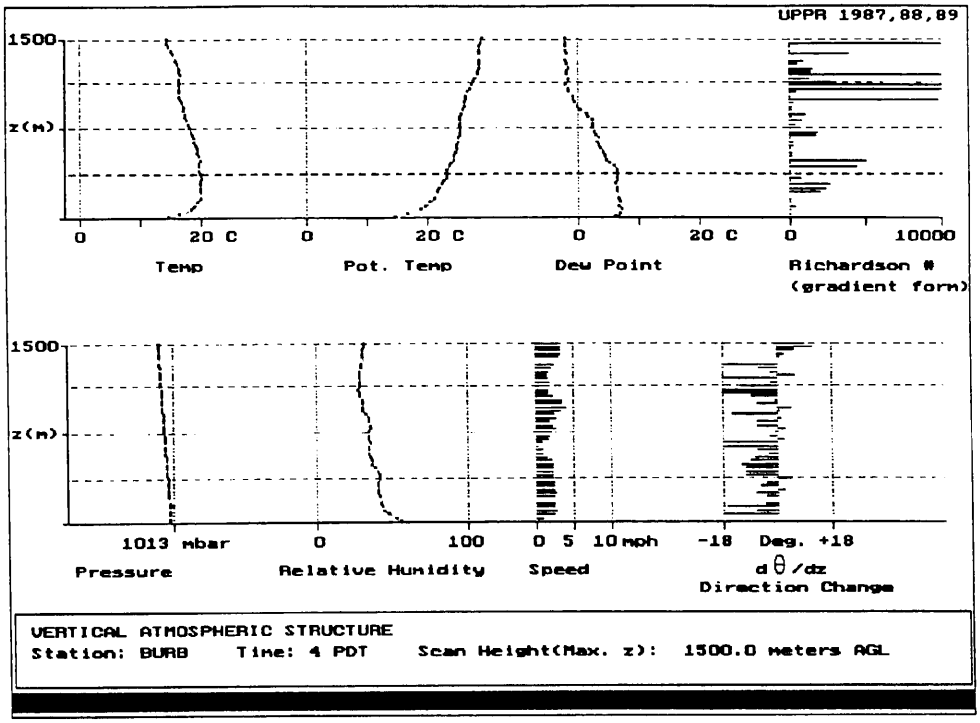


Figure 20. (Top) Vertical atmospheric structure at El Monte at 14:00 PDT, September 3, 1987.

Figure 21. (Bottom) Vertical atmospheric structure at El Monte at 17:00 PDT, September 3, 1987.



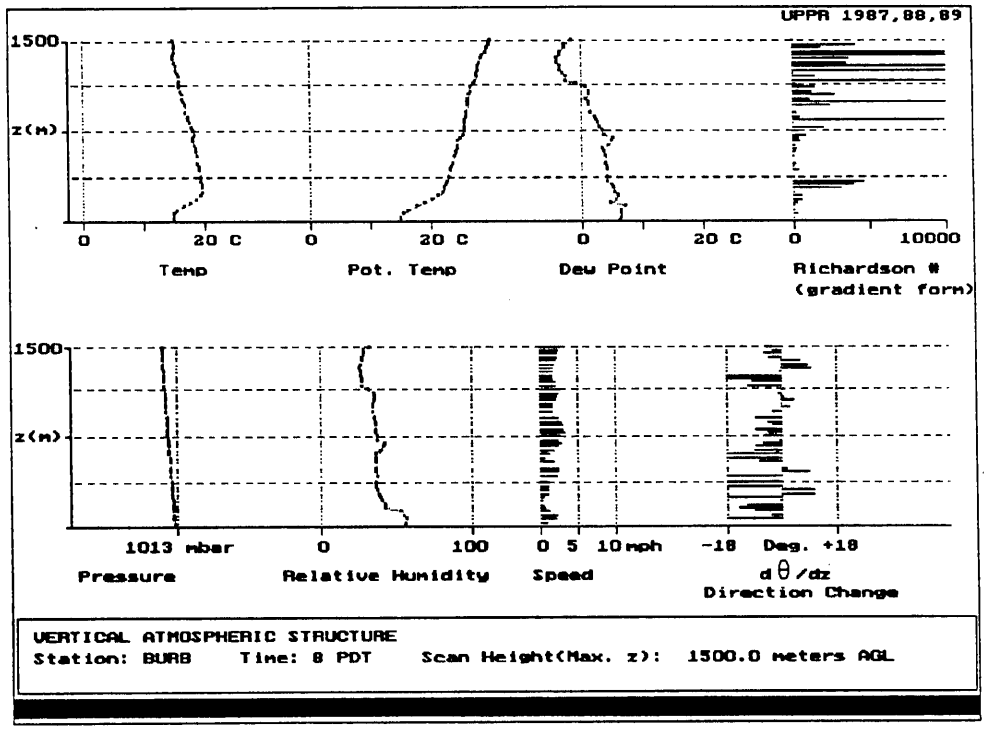
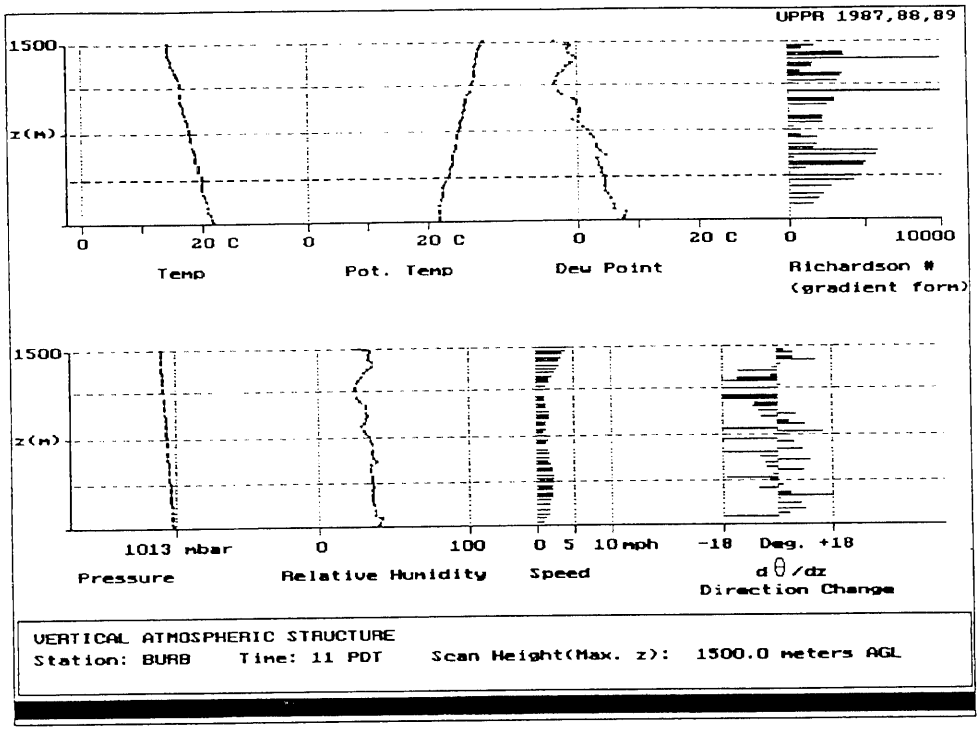


Figure 22. (Top) Vertical atmospheric structure at Burbank at 04:00 PDT, November 12, 1987.

Figure 23. (Bottom) Vertical atmospheric structure at Burbank at 08:00 PDT, November 12, 1987.



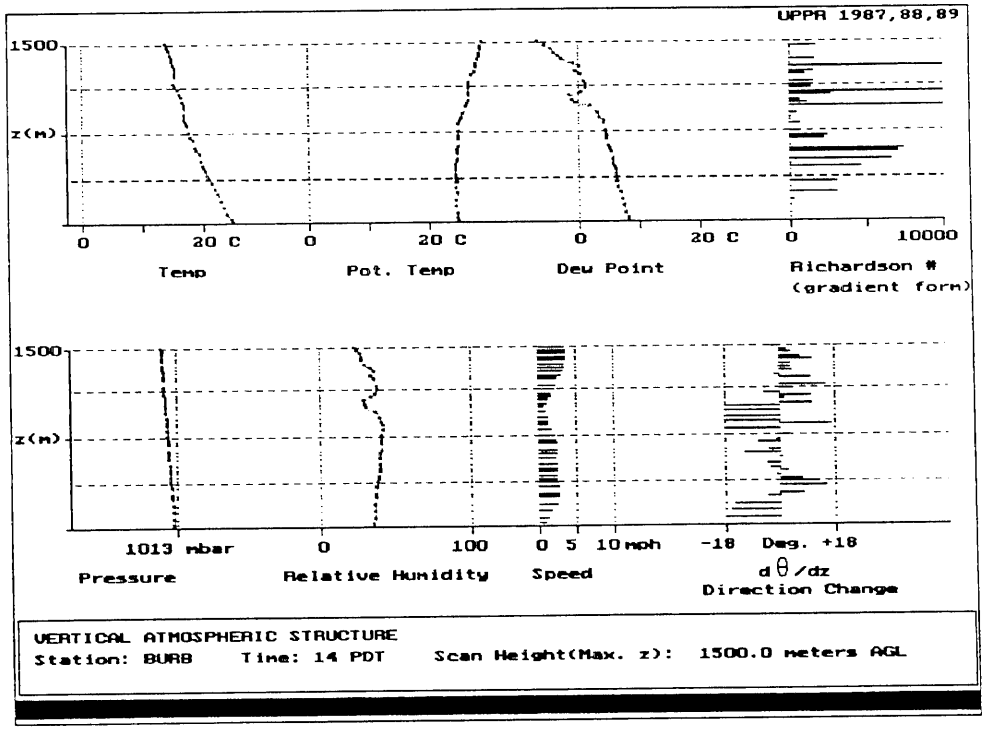
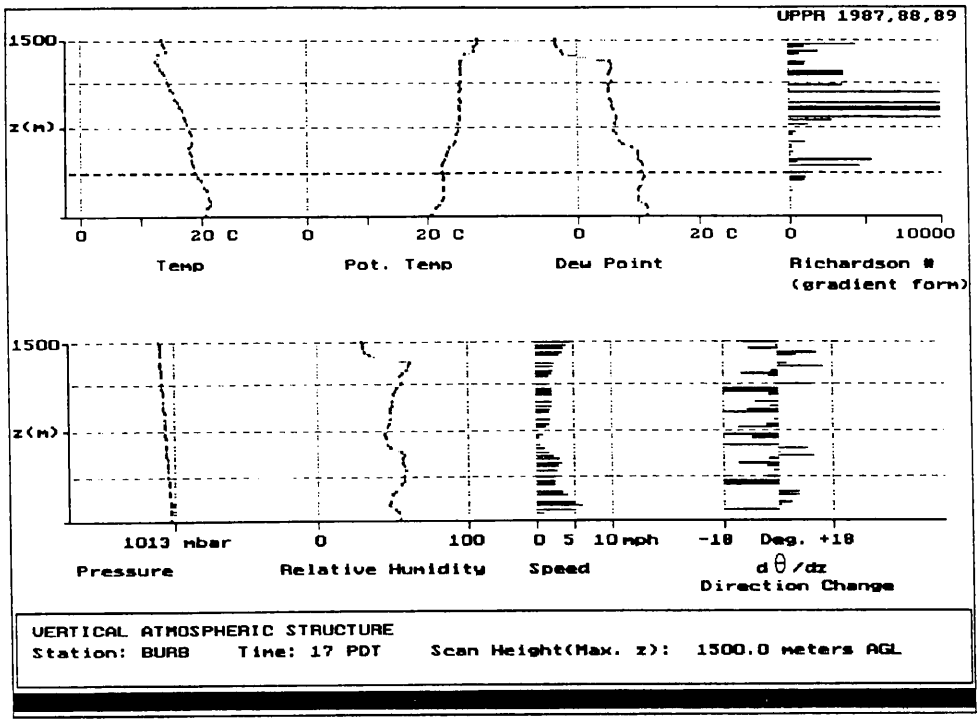


Figure 24. (Top) Vertical atmospheric structure at Burbank at 11:00 PDT, November 12, 1987.

Figure 25. (Bottom) Vertical atmospheric structure at Burbank at 14:00 PDT, November 12, 1987.



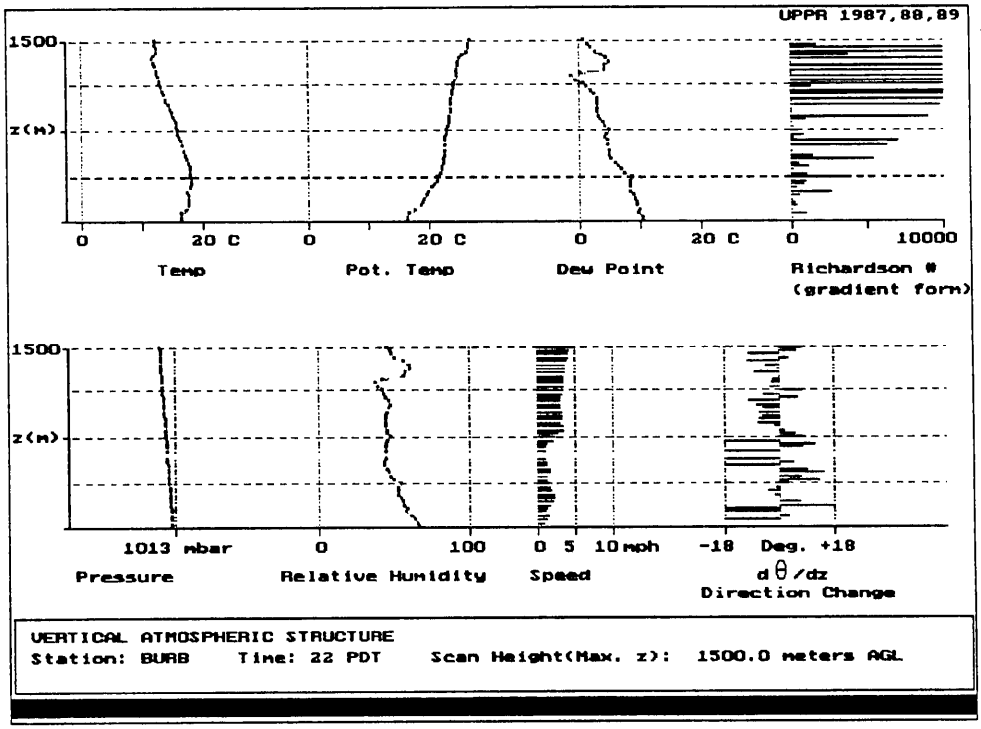
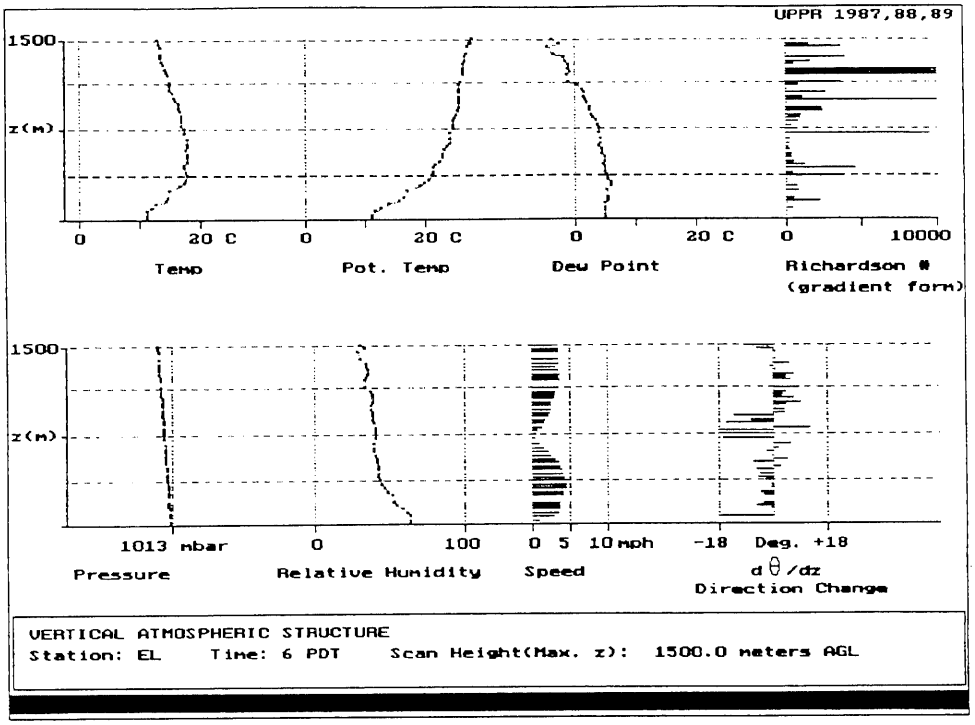


Figure 26. (Top) Vertical atmospheric structure at Burbank at 17:00 PDT, November 12, 1987.

Figure 27. (Bottom) Vertical atmospheric structure at Burbank at 22:00 PDT, November 12, 1987.



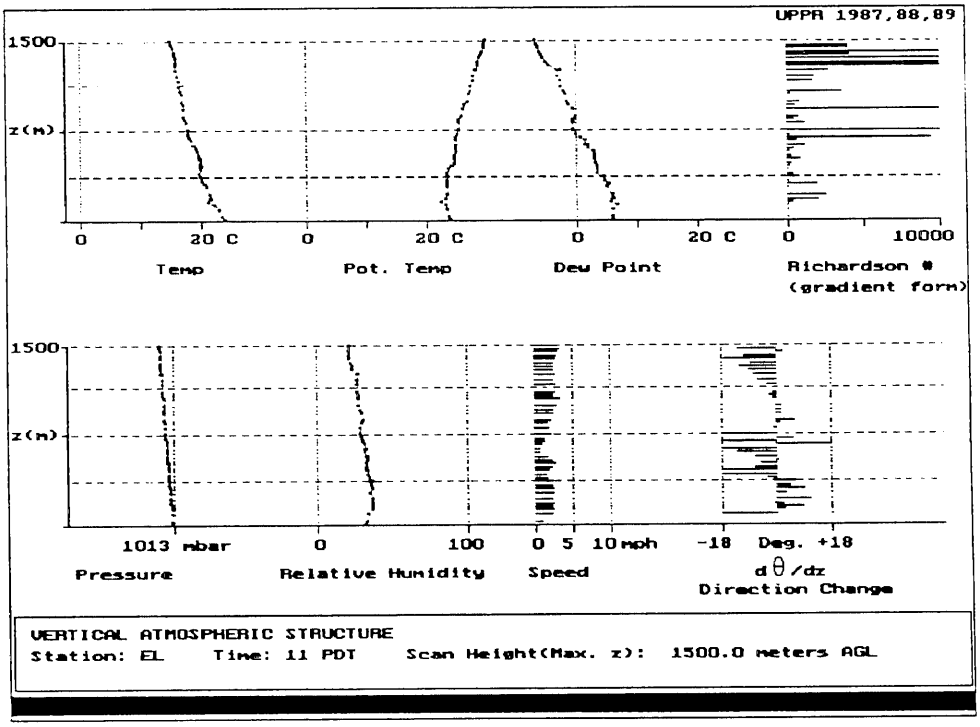
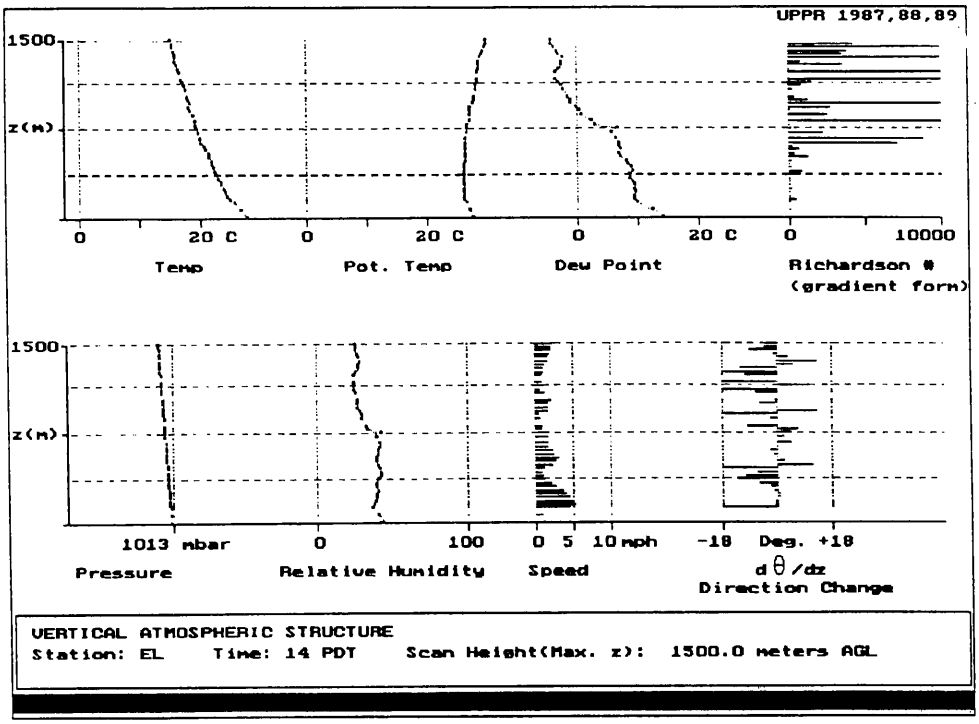


Figure 28. (Top) Vertical atmospheric structure at El Monte at 06:00 PDT, November 12, 1987.

Figure 29. (Bottom) Vertical atmospheric structure at El Monte at 11:00 PDT, November 12, 1987.



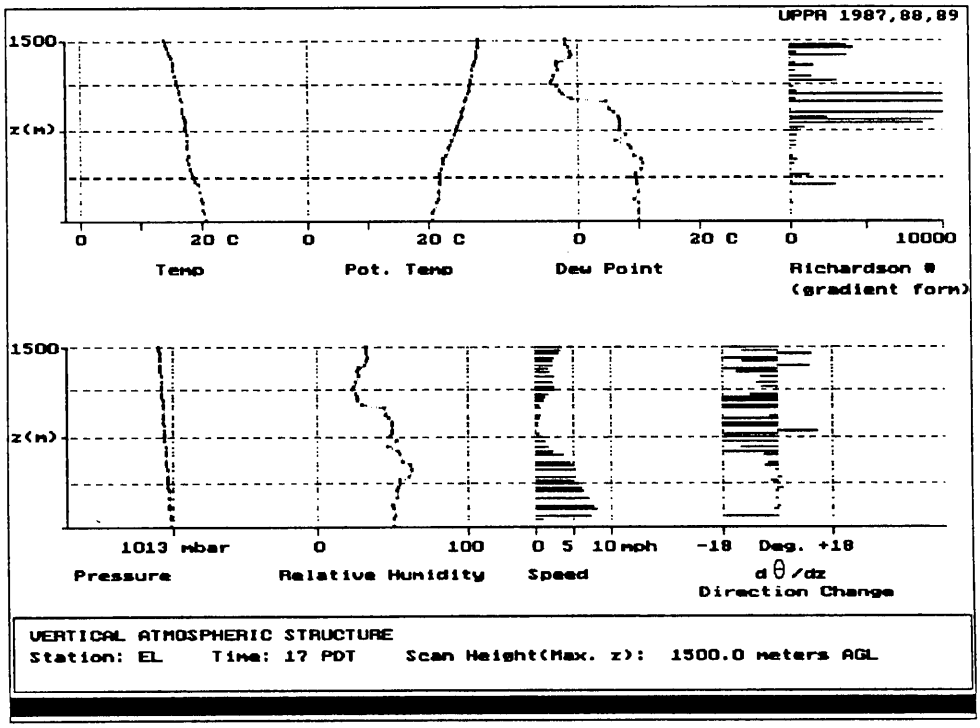
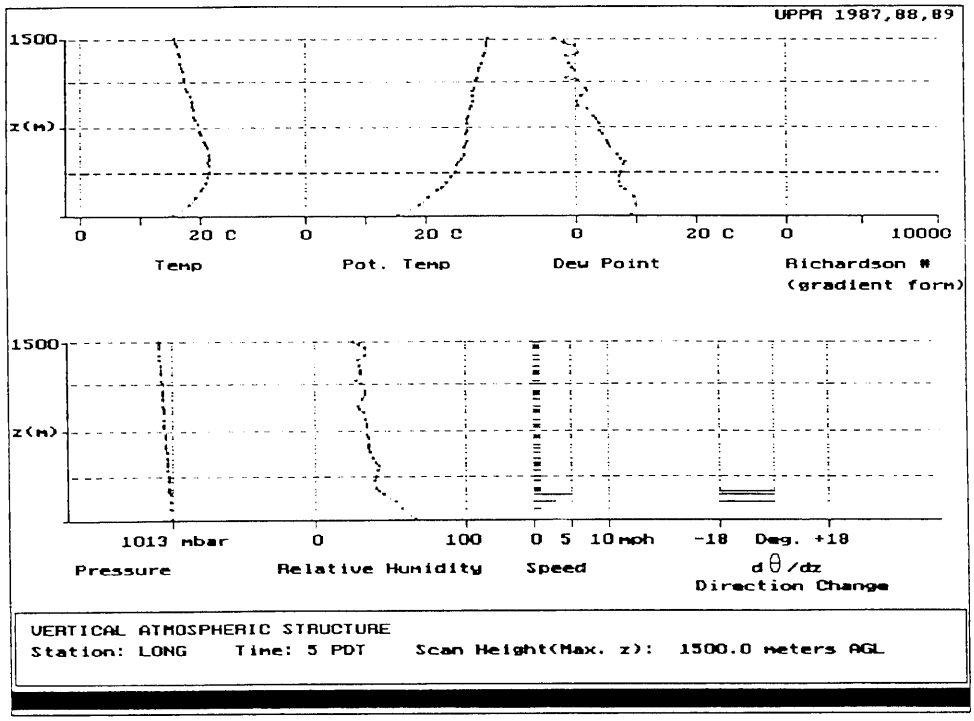


Figure 30. (Top) Vertical atmospheric structure at El Monte at 14:00 PDT, November 12, 1987.

Figure 31. (Bottom) Vertical atmospheric structure at El Monte at 17:00 PDT, November 12, 1987.



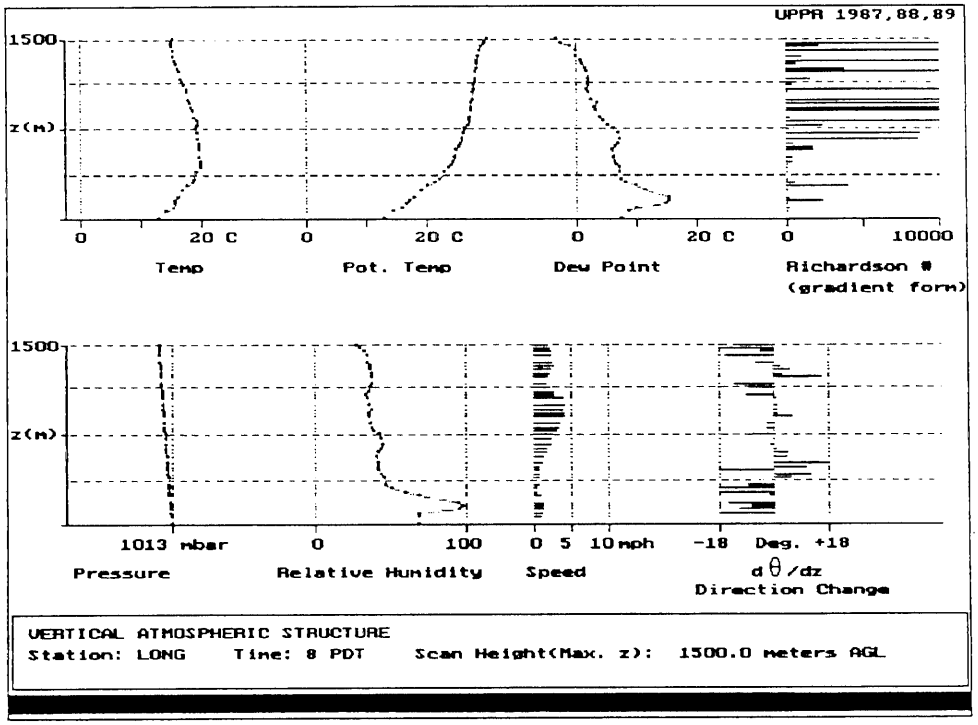
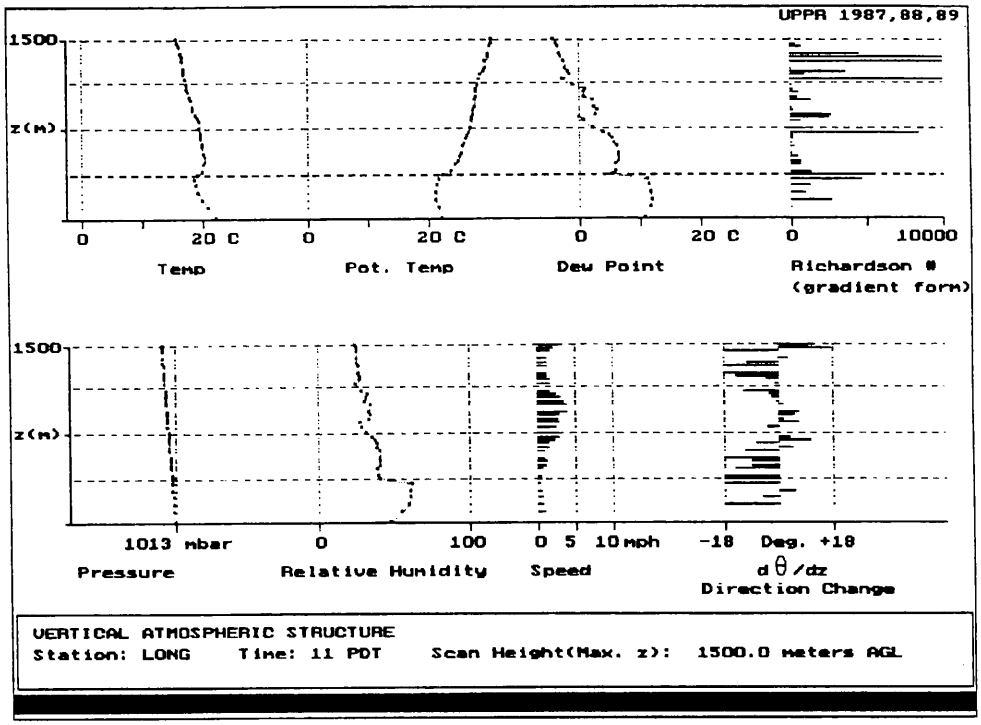


Figure 32. (Top) Vertical atmospheric structure at Long Beach at 05:00 PDT, November 12, 1987.

Figure 33. (Bottom) Vertical atmospheric structure at Long Beach at 08:00 PDT, November 12, 1987.



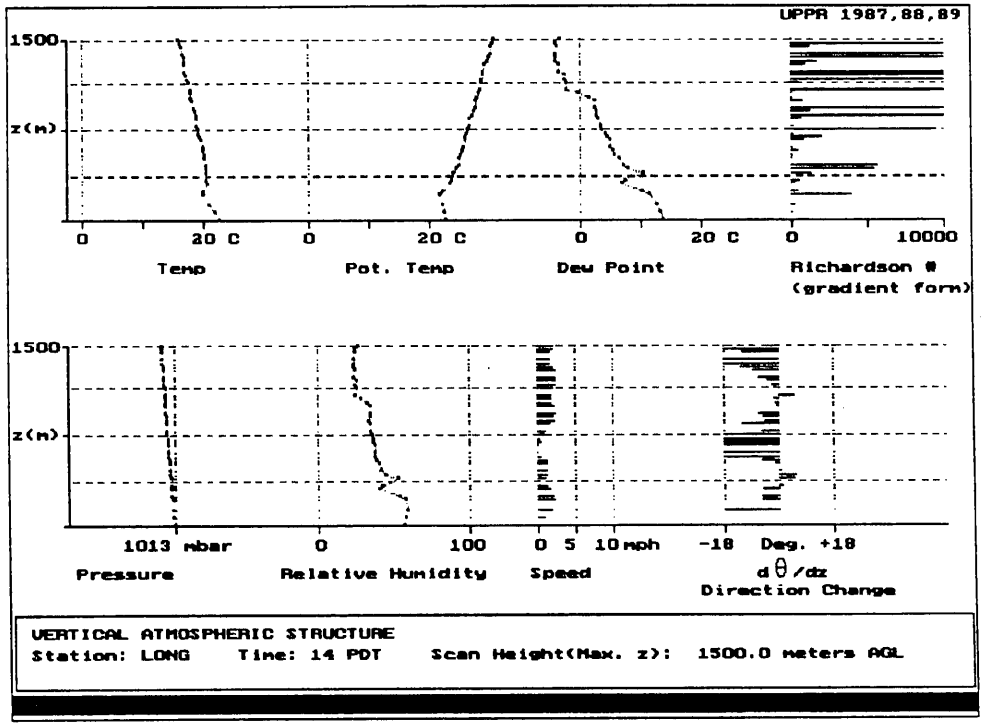
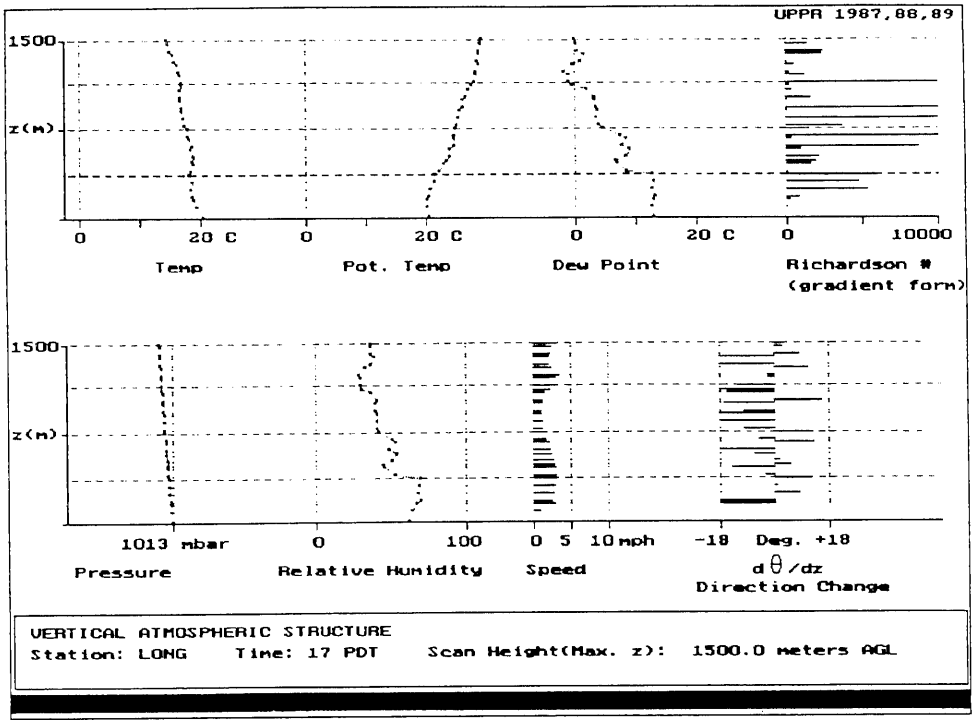


Figure 34. (Top) Vertical atmospheric structure at Long Beach at 11:00 PDT, November 12, 1987.

Figure 35. (Bottom) Vertical atmospheric structure at Long Beach at 14:00 PDT, November 12, 1987.



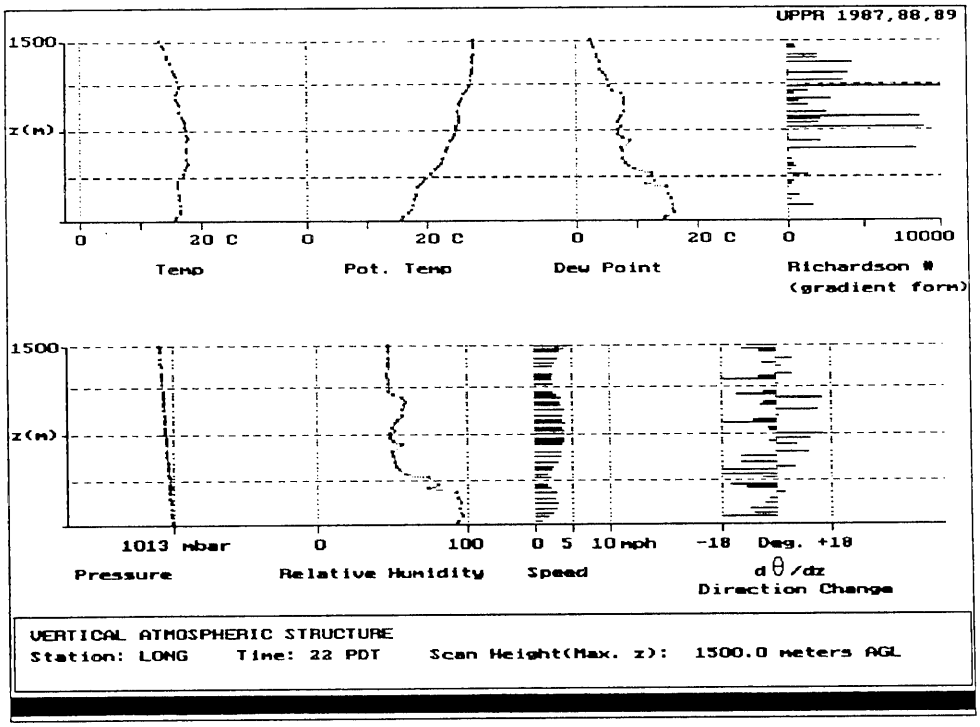


Figure 36. (Top) Vertical atmospheric structure at Long Beach at 17:00 PDT, November 12, 1987.

Figure 37. (Bottom) Vertical atmospheric structure at Long Beach at 22:00 PDT, November 12, 1987.

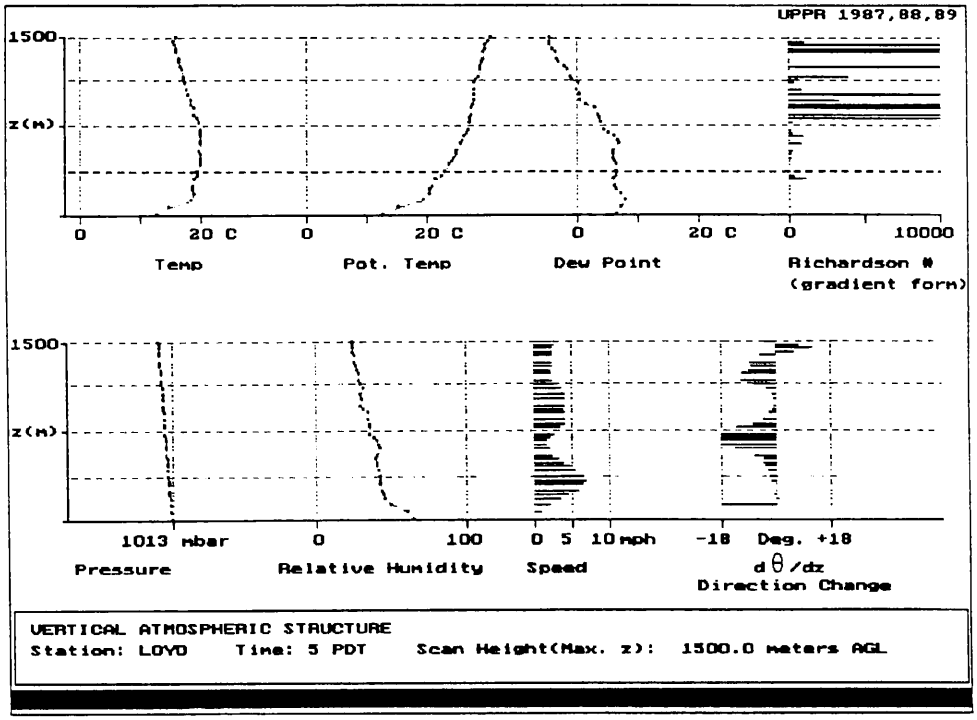
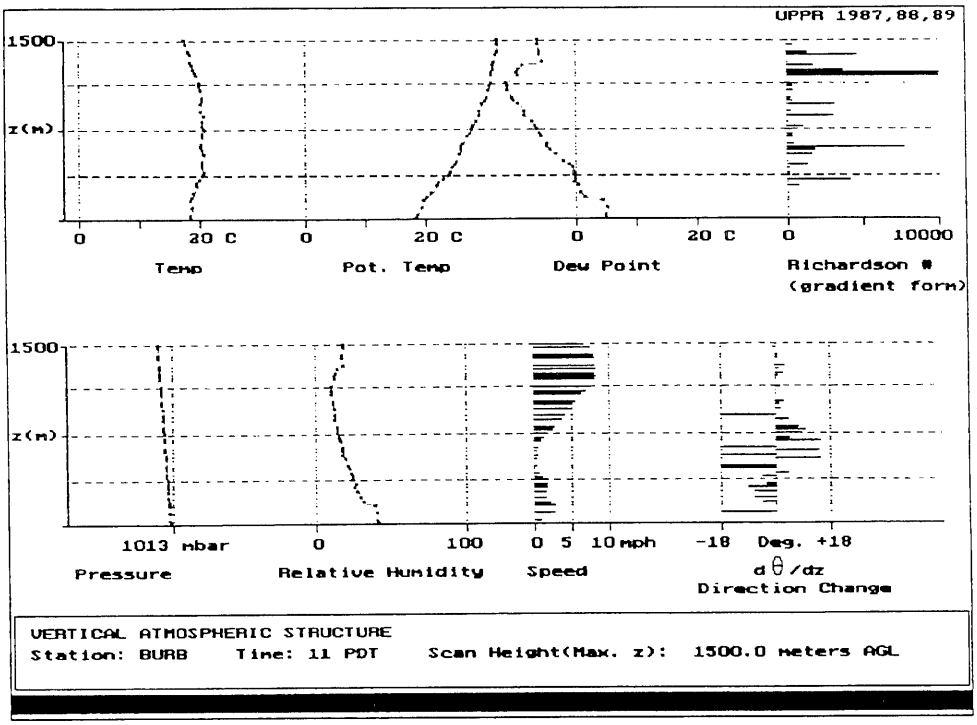


Figure 38. Vertical atmospheric structure at Loyola Marymount University at 05:00 PDT, November 12, 1987.



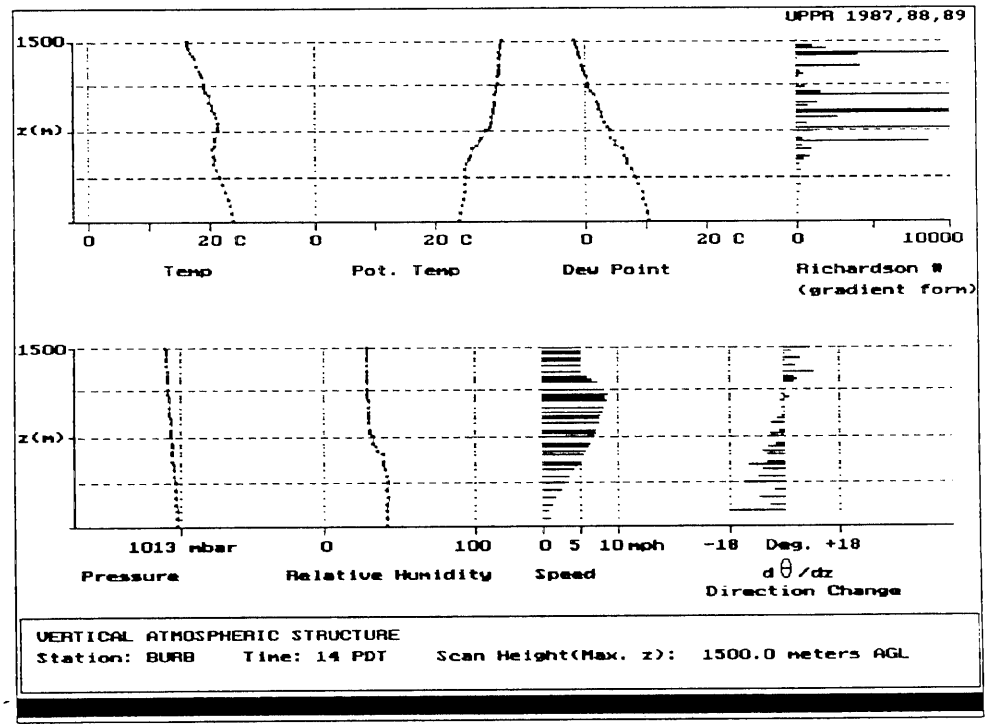
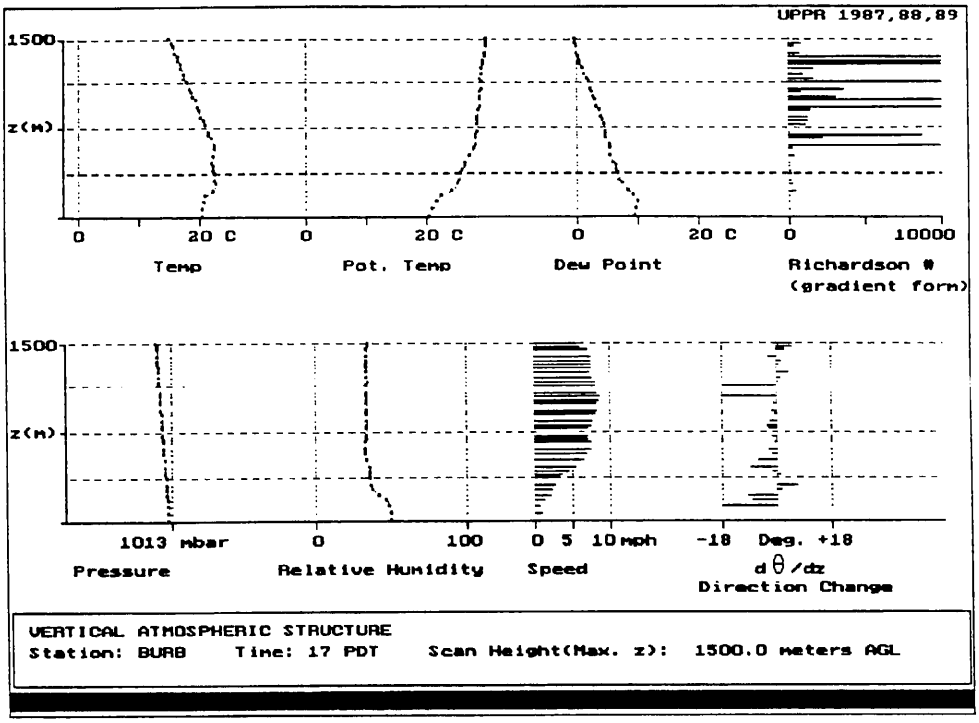


Figure 39. (Top) Vertical atmospheric structure at Burbank at 11:00 PDT, December 10, 1987.

Figure 40. (Bottom) Vertical atmospheric structure at Burbank at 14:00 PDT, December 10, 1987.



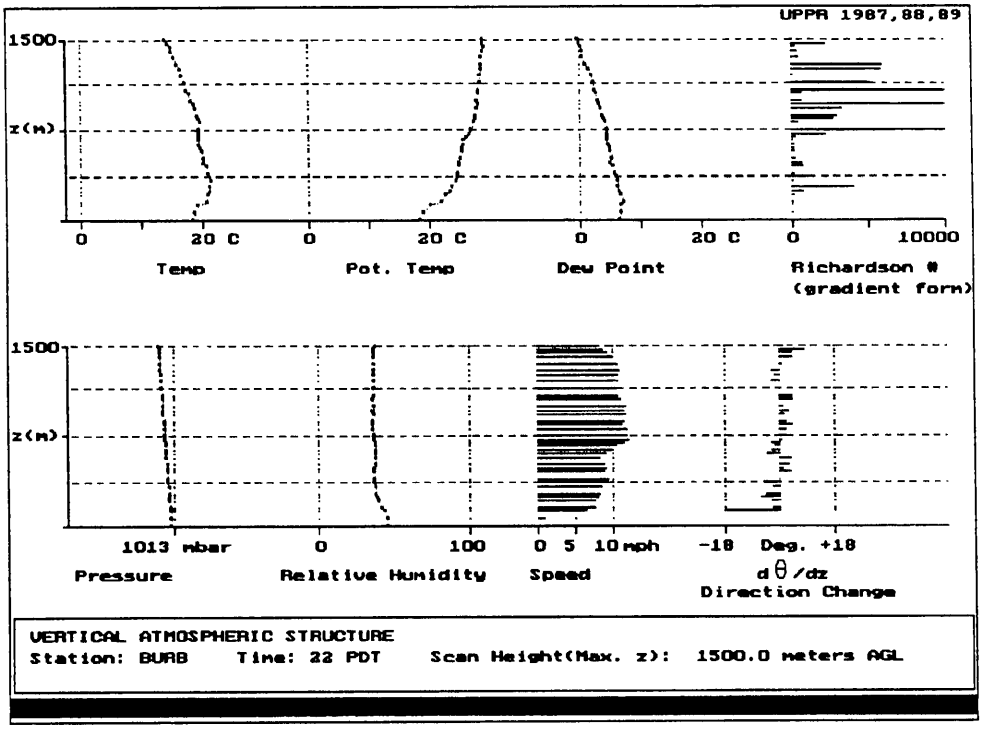
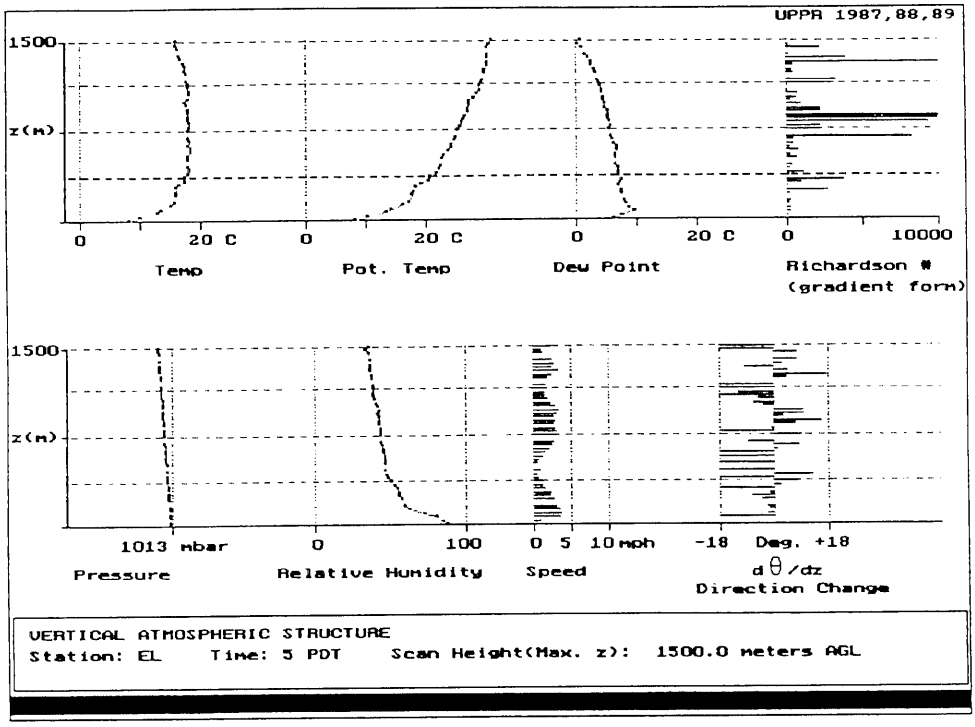


Figure 41. (Top) Vertical atmospheric structure at Burbank at 17:00 PDT, December 10, 1987.

Figure 42. (Bottom) Vertical atmospheric structure at Burbank at 22:00 PDT, December 10, 1987.



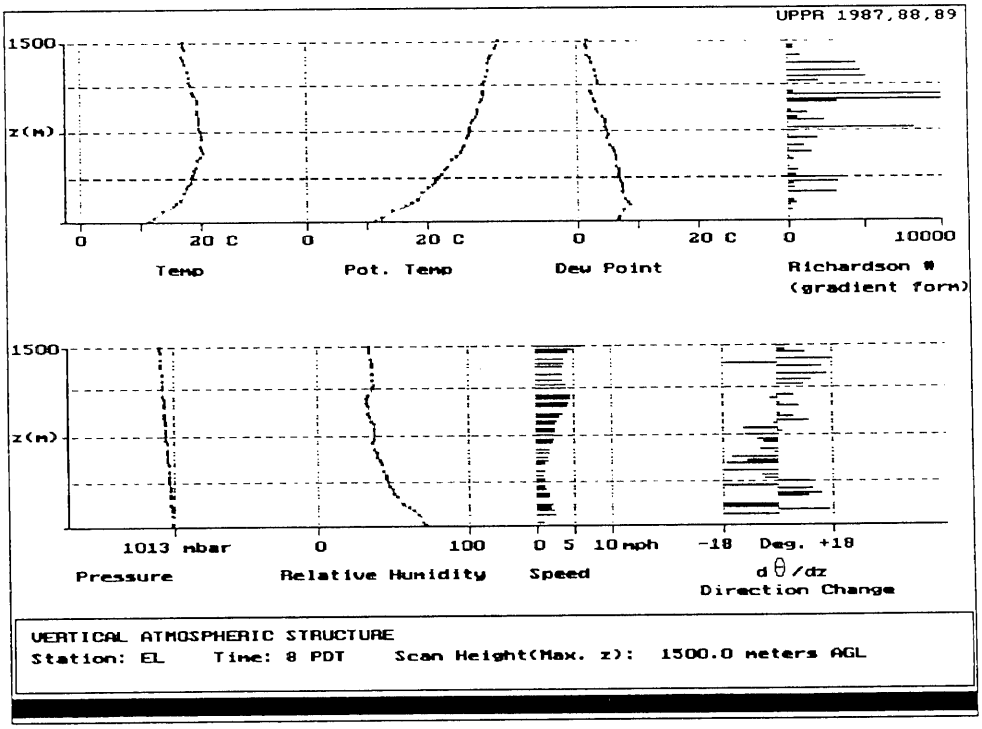
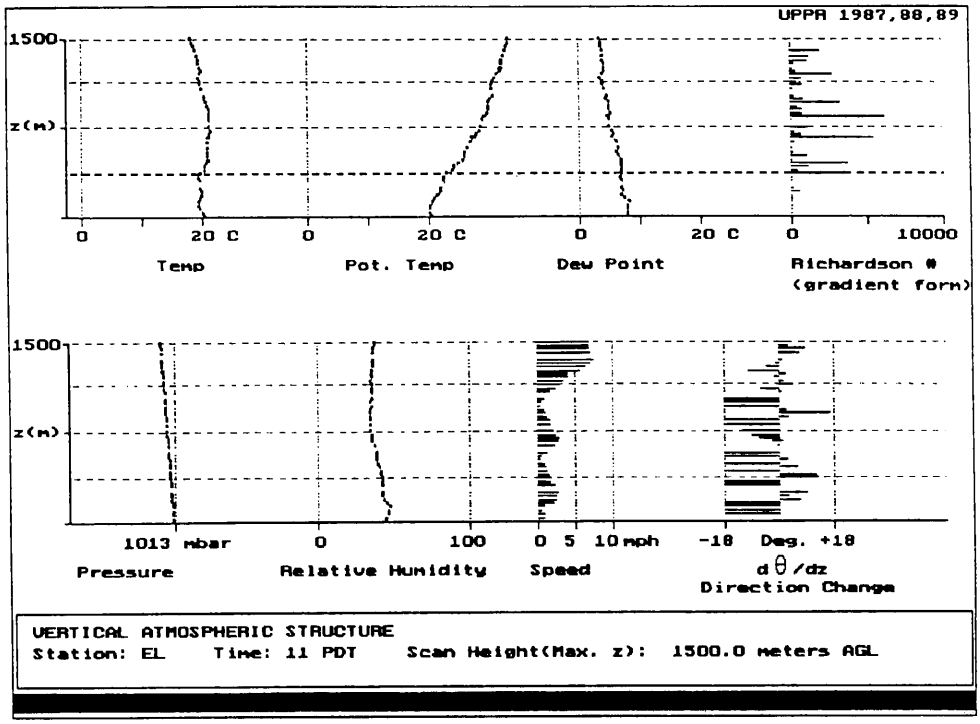


Figure 43. (Top) Vertical atmospheric structure at El Monte at 05:00 PDT, December 10, 1987.

Figure 44. (Bottom) Vertical atmospheric structure at El Monte at 08:00 PDT, December 10, 1987.



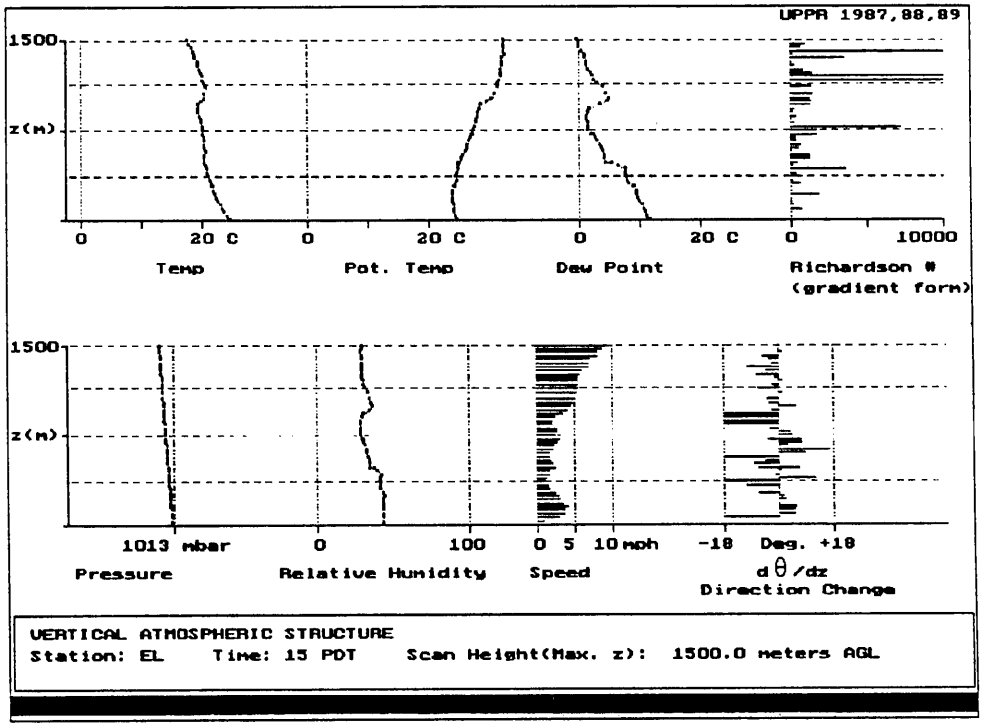
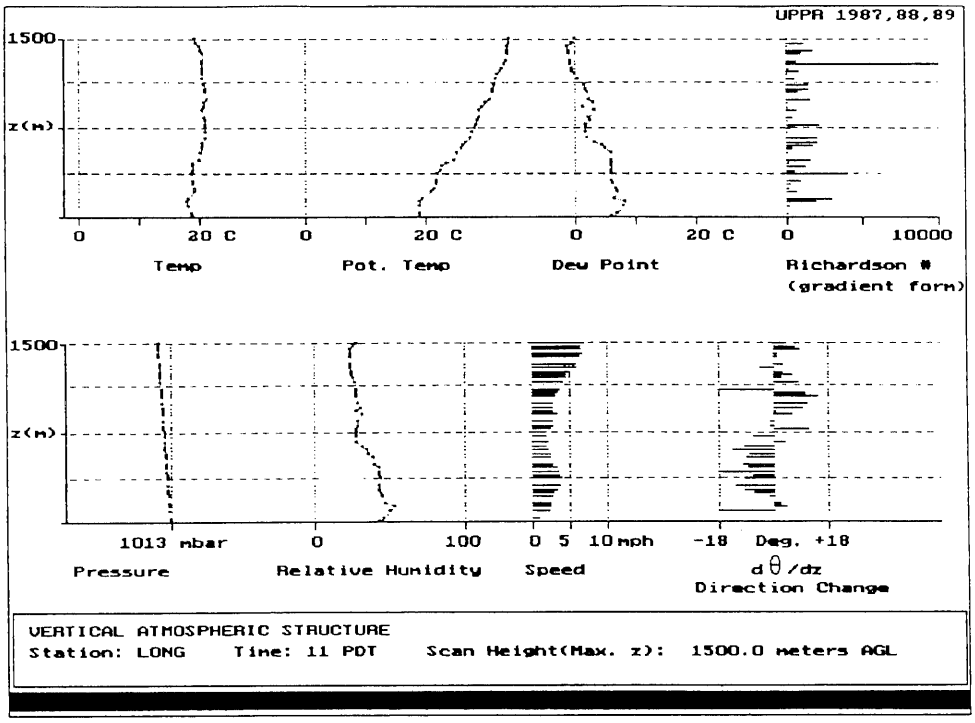


Figure 45. (Top) Vertical atmospheric structure at El Monte at 11:00 PDT, December 10, 1987.

Figure 46. (Bottom) Vertical atmospheric structure at El Monte at 15:00 PDT, December 10, 1987.



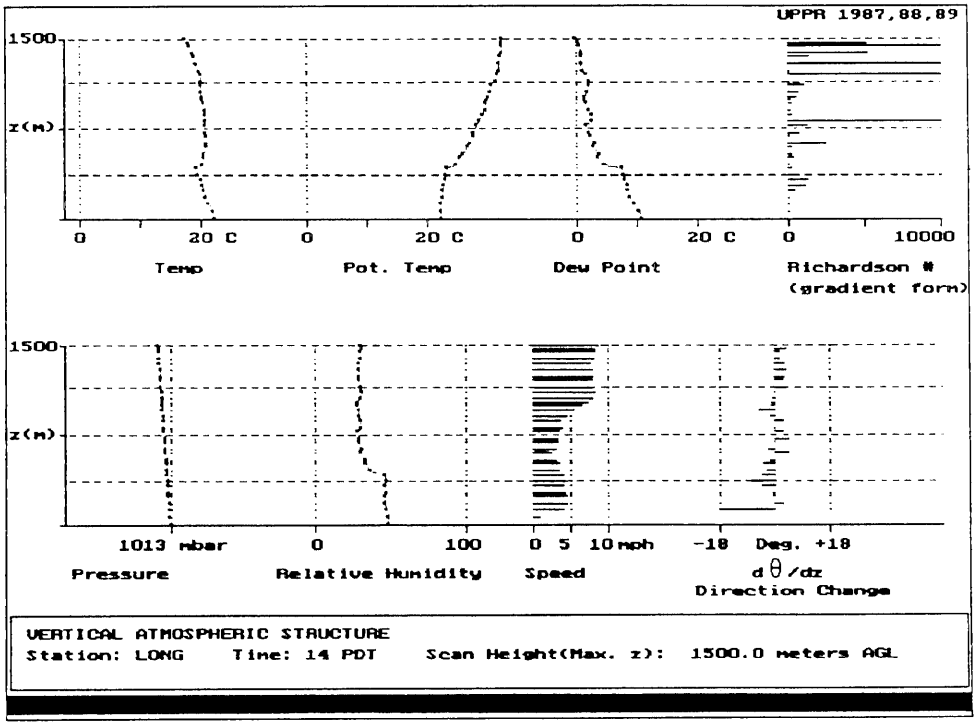
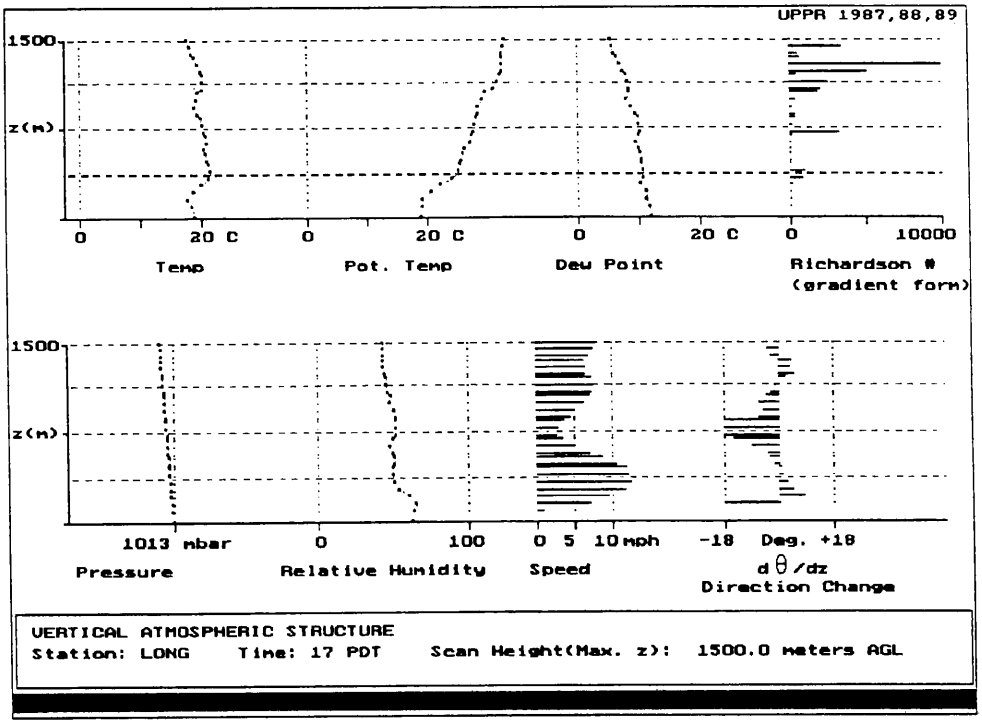


Figure 47. (Top) Vertical atmospheric structure at Long Beach at 11:00 PDT, December 10, 1987.

Figure 48. (Bottom) Vertical atmospheric structure at Long Beach at 14:00 PDT, December 10, 1987.



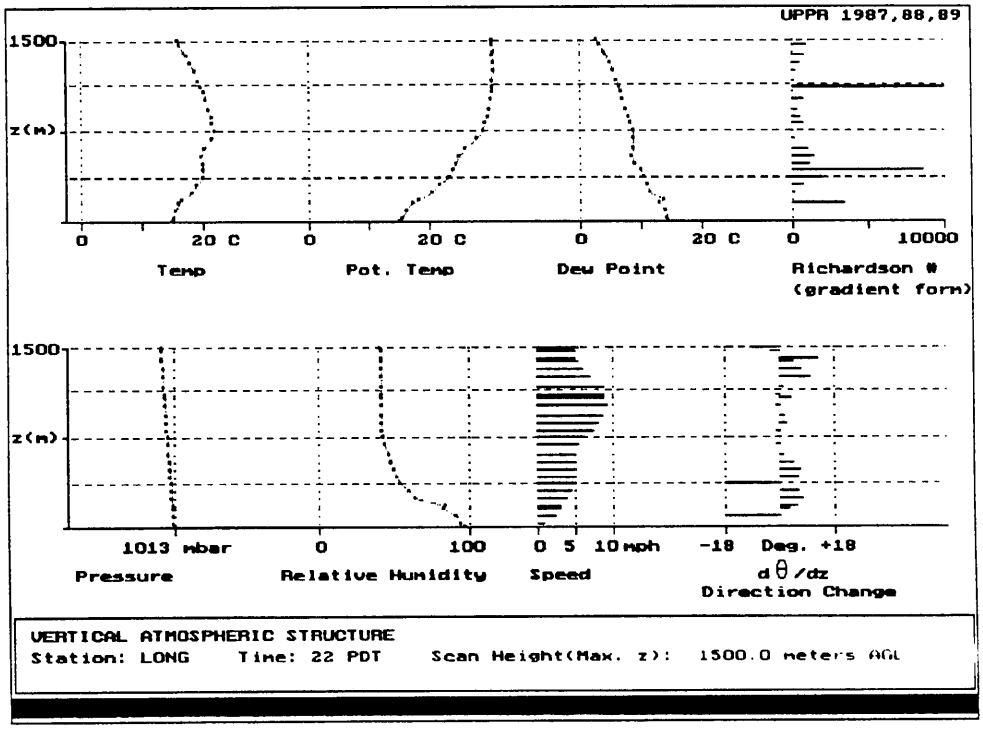
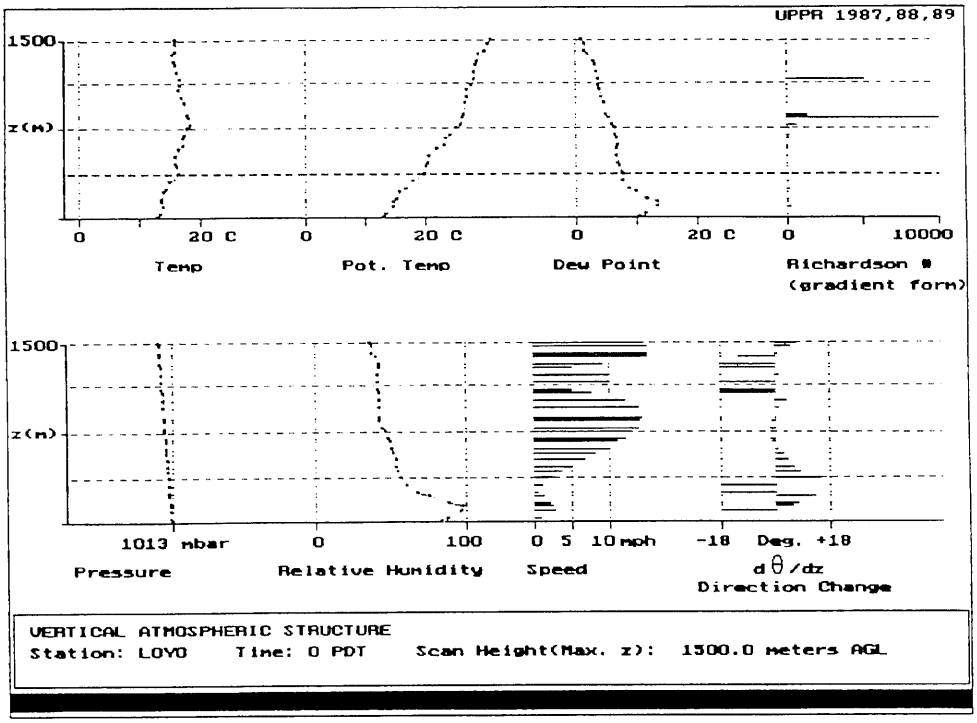


Figure 49. (Top) Vertical atmospheric structure at Long Beach at 17:00 PDT, December 10, 1987.

Figure 50. (Bottom) Vertical atmospheric structure at Long Beach at 22:00 PDT, December 10, 1987.



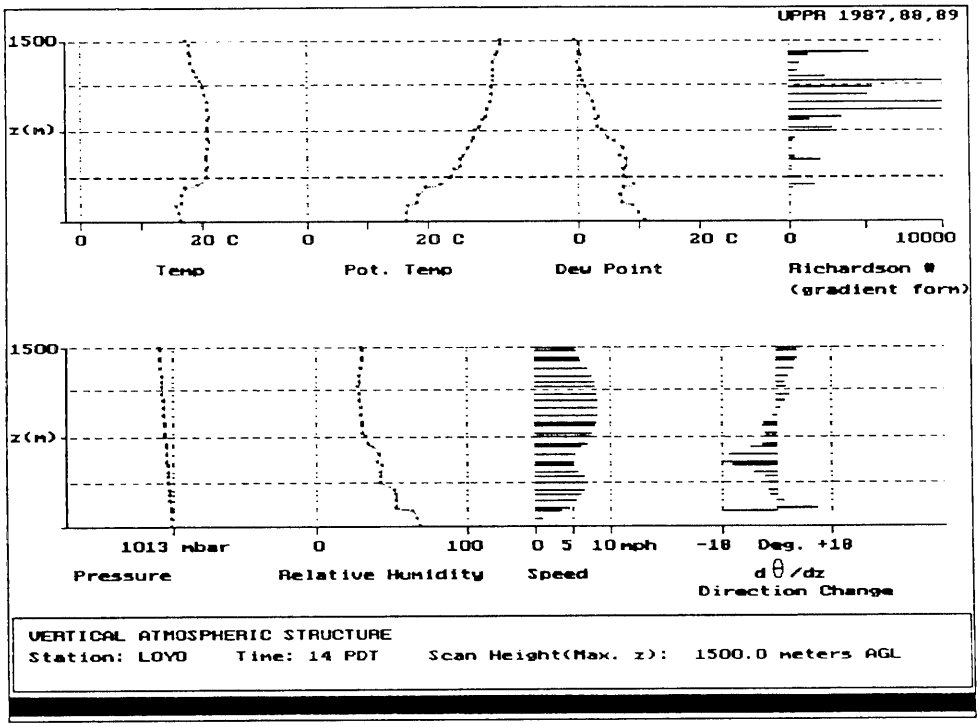


Figure 51. (Top) Vertical atmospheric structure at Loyola Marymount University at 00:00 PDT, December 10, 1987.

Figure 52. (Bottom) Vertical atmospheric structure at Loyola Marymount University at 14:00 PDT, December 10, 1987.

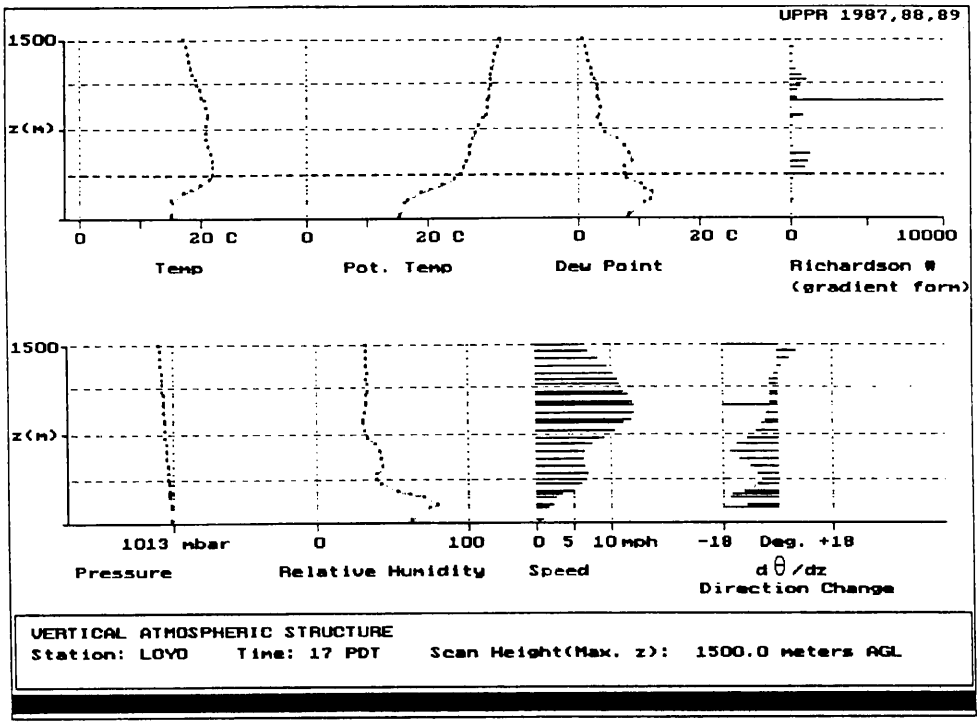
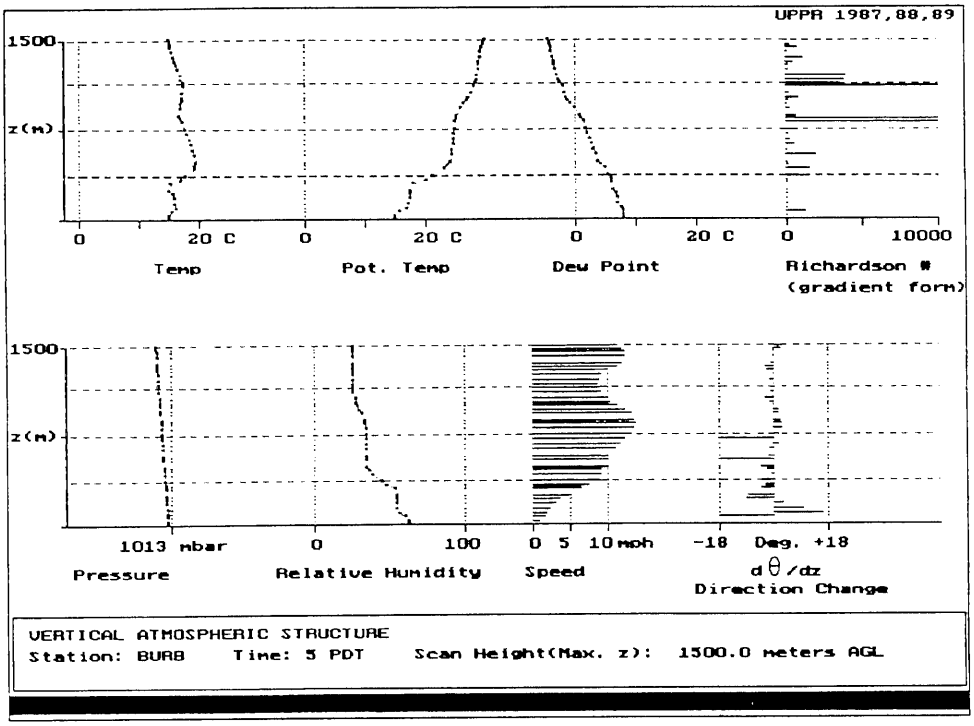


Figure 53. Vertical atmospheric structure at Loyola Marymount University at 17:00 PDT, December 10, 1987.



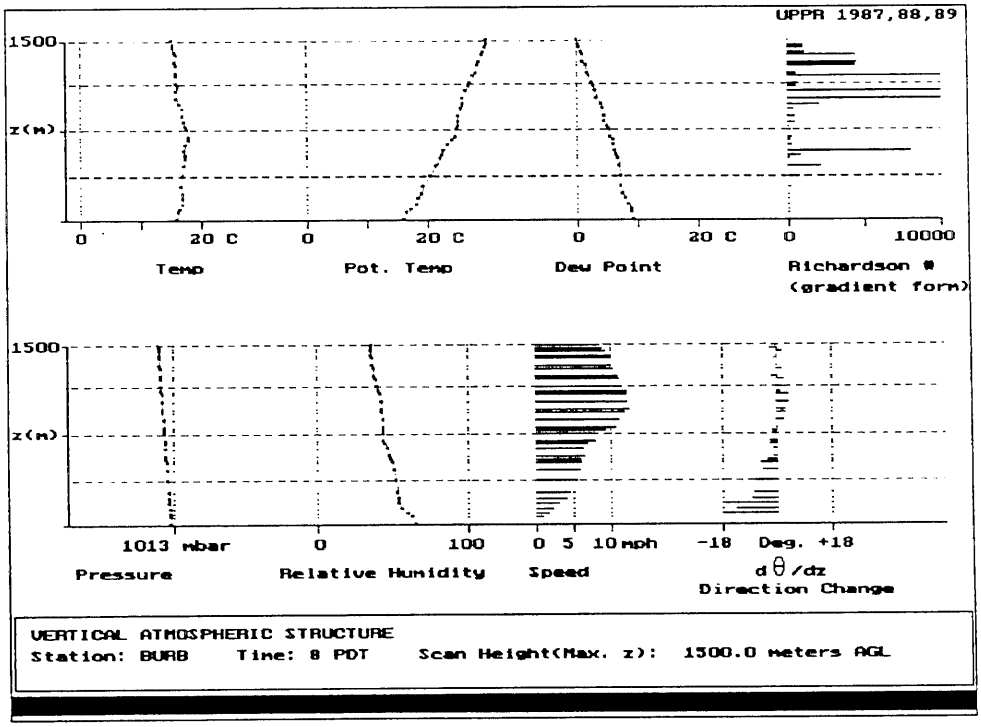
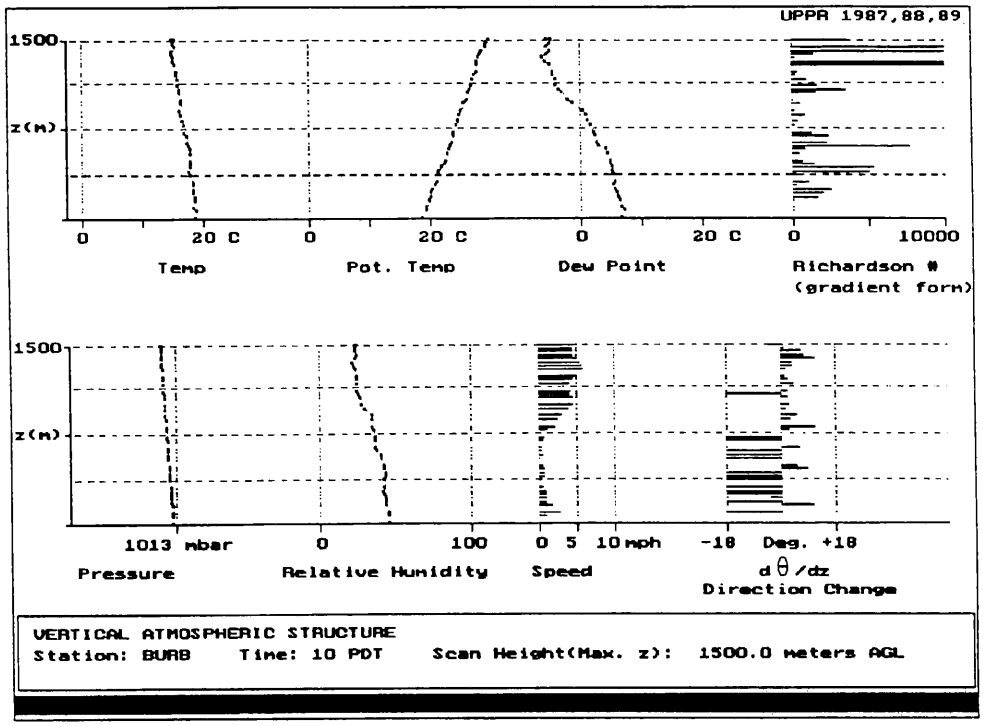


Figure 54. (Top) Vertical atmospheric structure at Burbank at 05:00 PDT, December 11, 1987.

Figure 55. (Bottom) Vertical atmospheric structure at Burbank at 08:00 PDT, December 11, 1987.



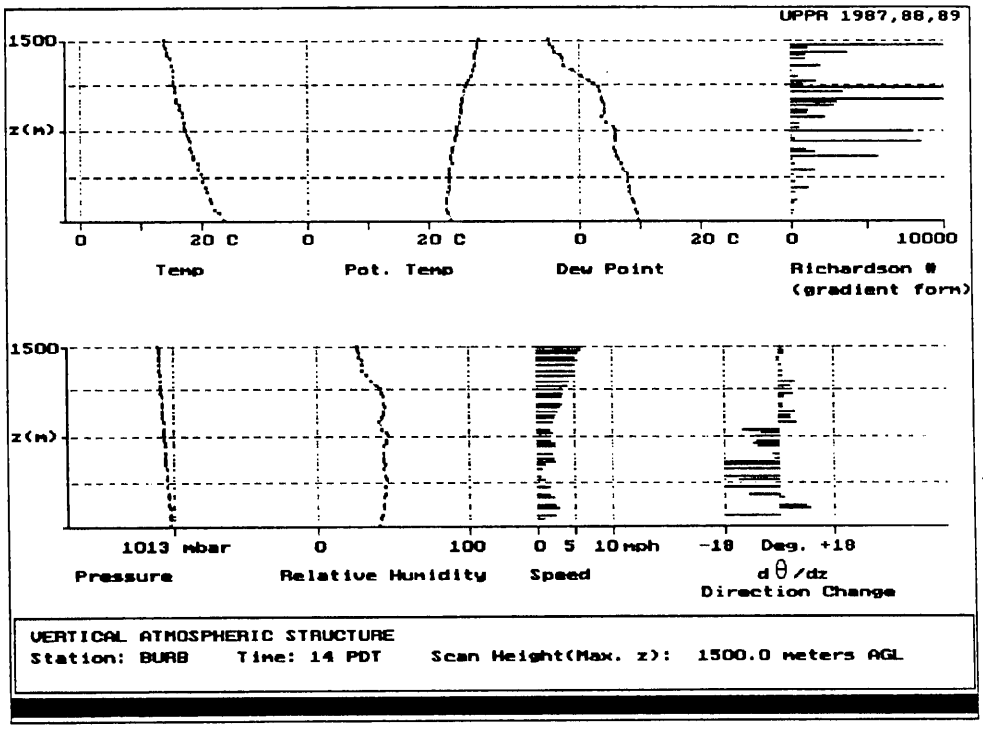
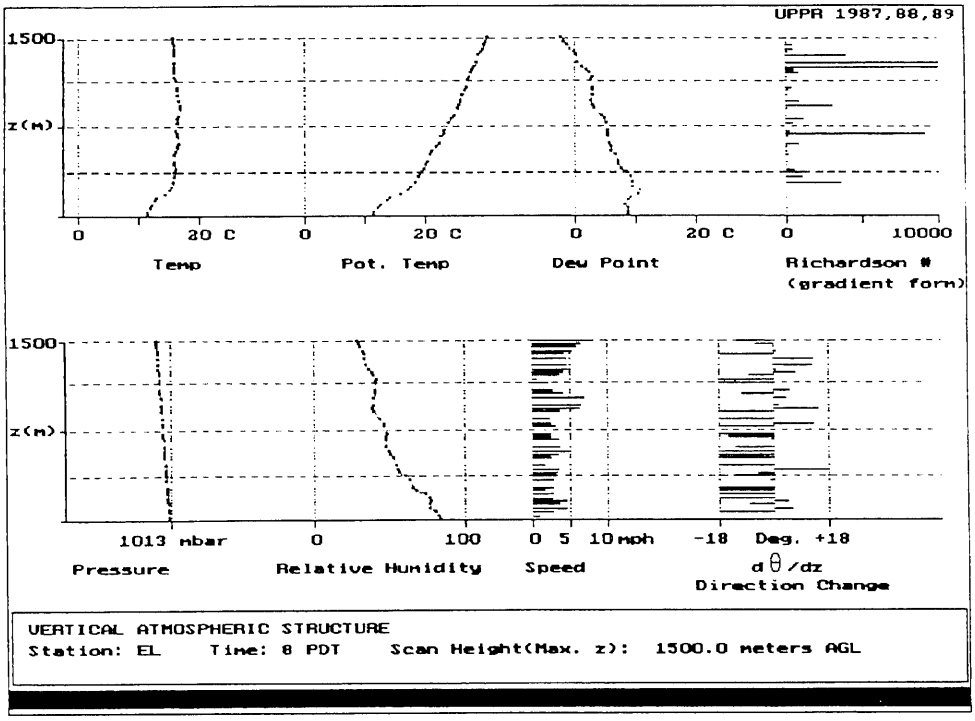


Figure 56. (Top) Vertical atmospheric structure at Burbank at 10:00 PDT, December 11, 1987.

Figure 57. (Bottom) Vertical atmospheric structure at Burbank at 14:00 PDT, December 11, 1987.



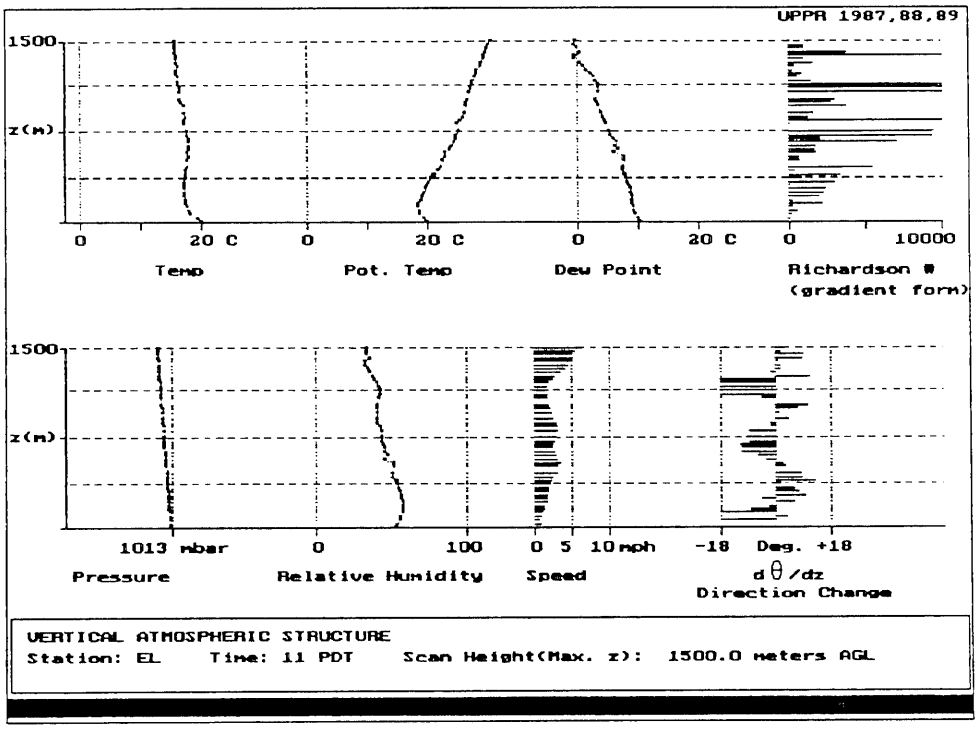


Figure 58. (Top) Vertical atmospheric structure at El Monte at 08:00 PDT, December 11, 1987.

Figure 59. (Bottom) Vertical atmospheric structure at El Monte at 11:00 PDT, December 11, 1987.

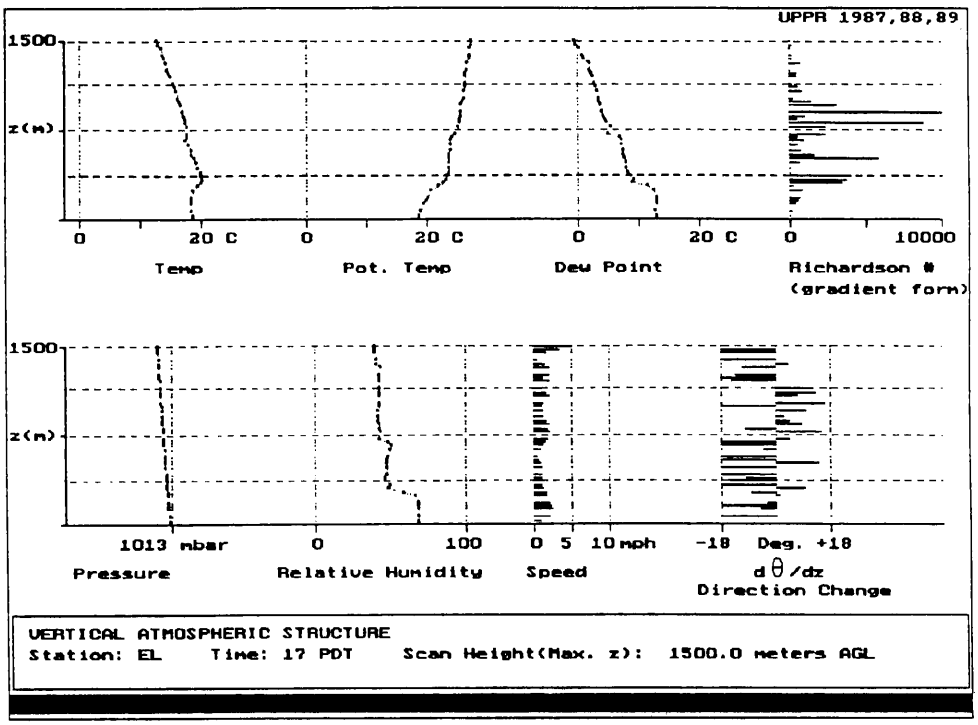
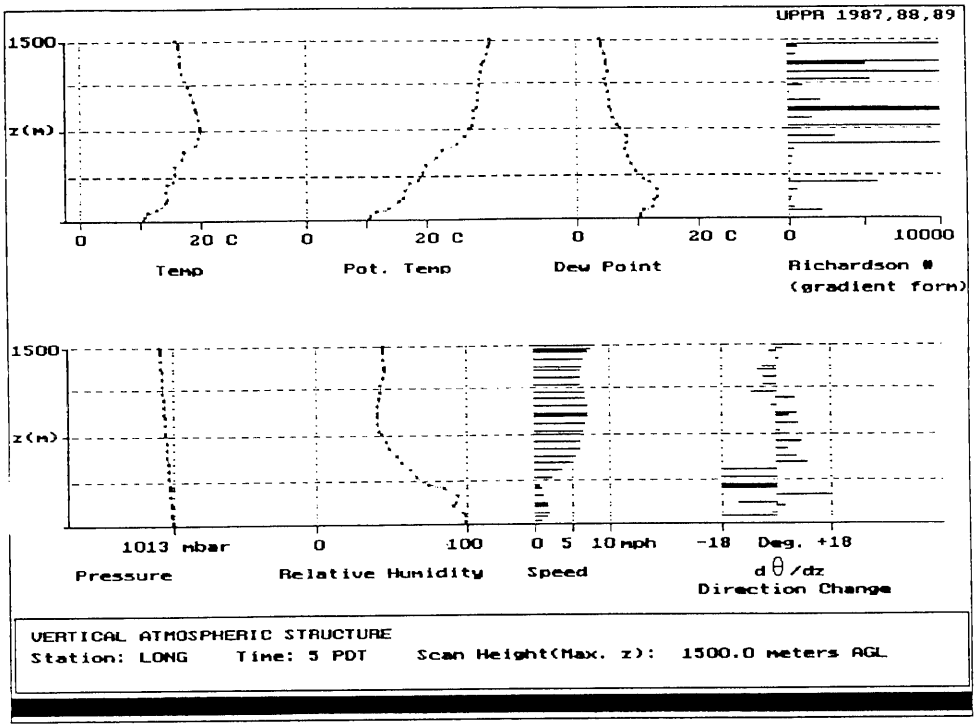


Figure 60. Vertical atmospheric structure at El Monte at 17:00 PDT, December 11, 1987.



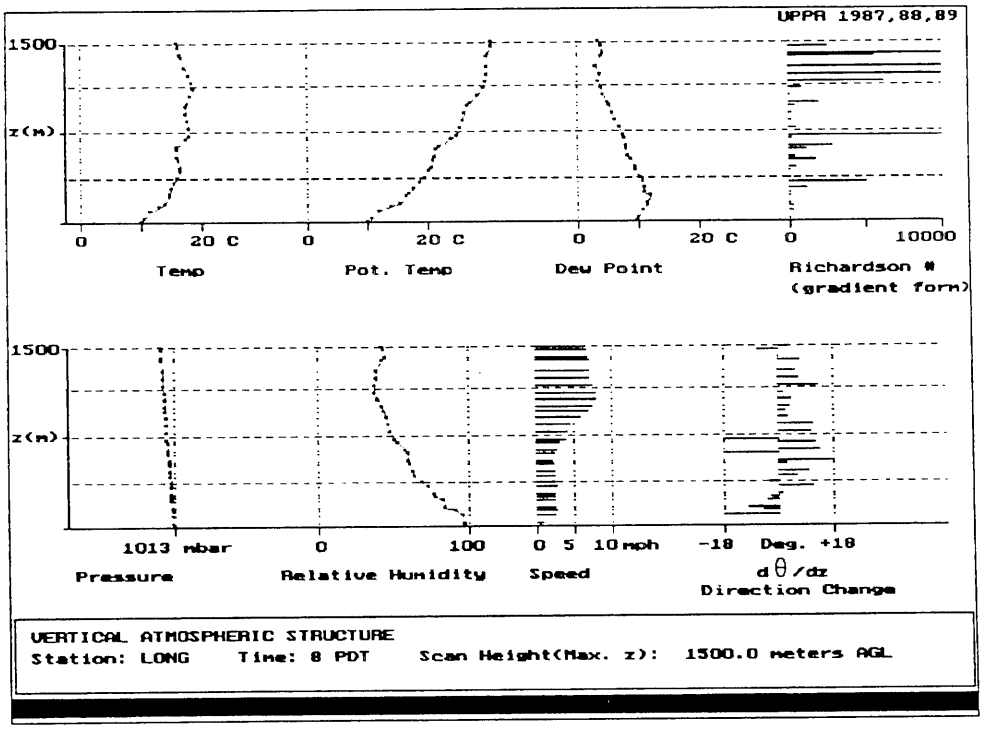
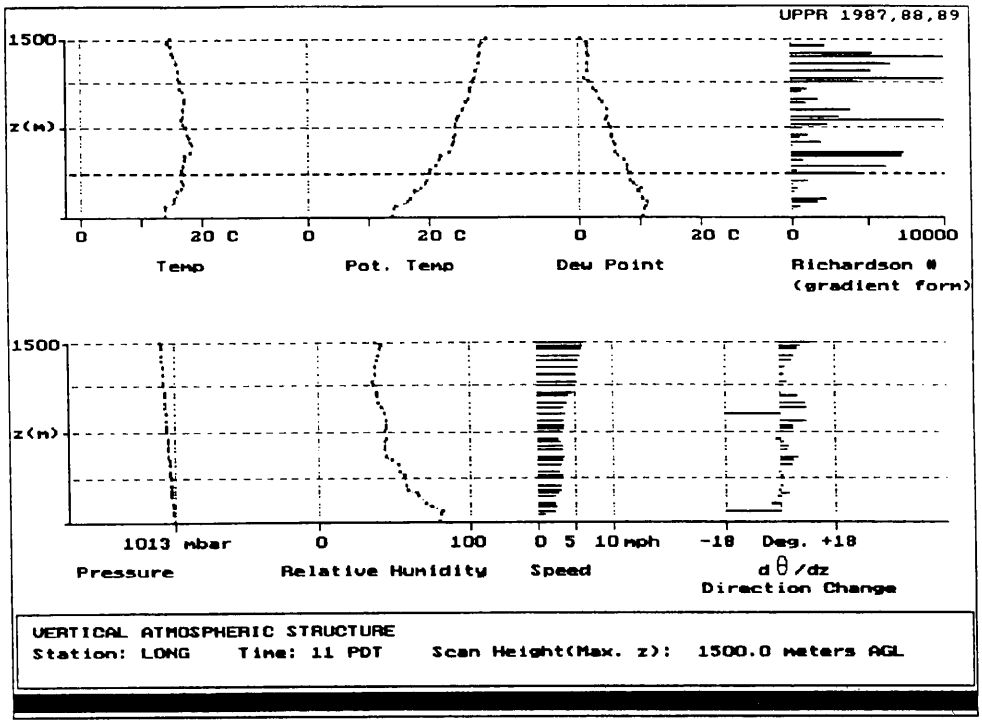


Figure 61. (Top) Vertical atmospheric structure at Long Beach at 05:00 PDT, December 11, 1987.

Figure 62. (Bottom) Vertical atmospheric structure at Long Beach at 08:00 PDT, December 11, 1987.



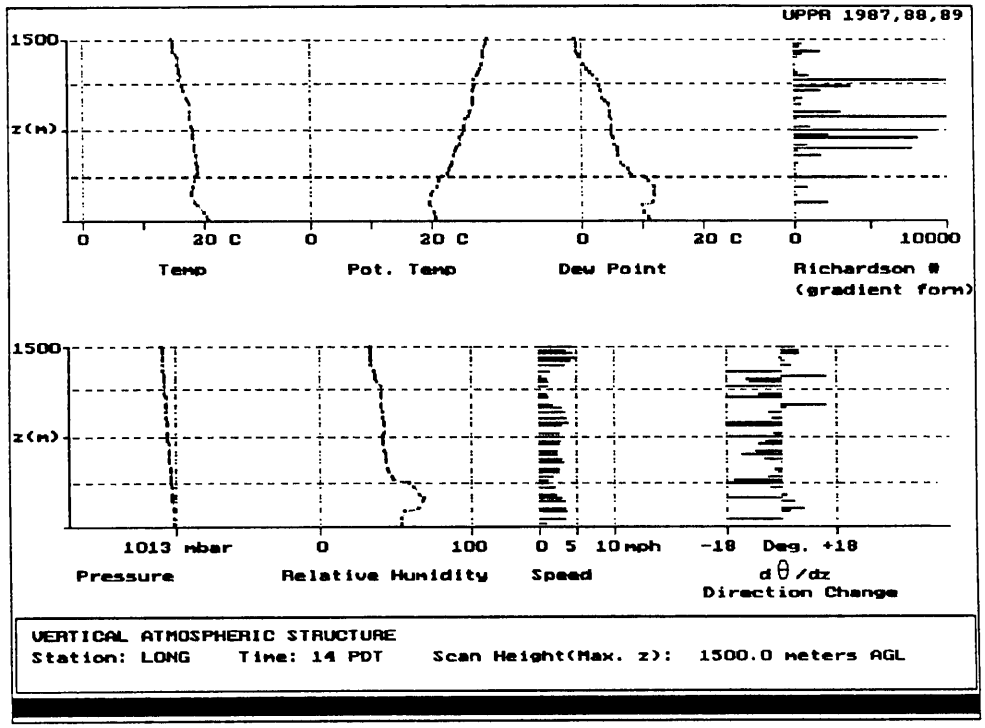
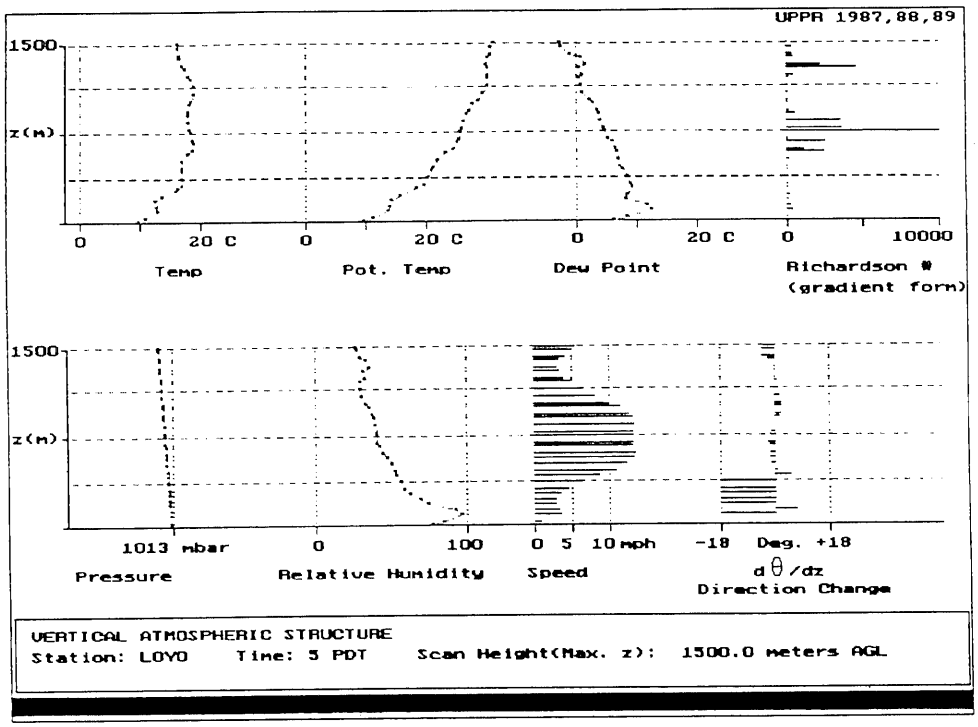


Figure 63. (Top) Vertical atmospheric structure at Long Beach at 11:00 PDT, December 11, 1987.

Figure 64. (Bottom) Vertical atmospheric structure at Long Beach at 14:00 PDT, December 11, 1987.



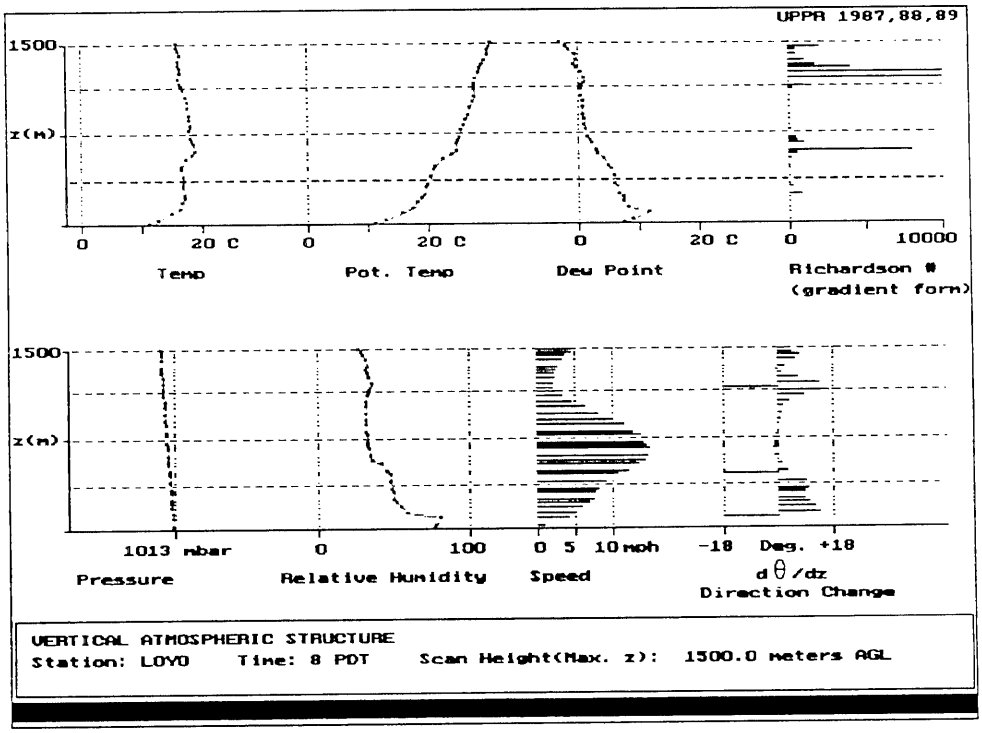
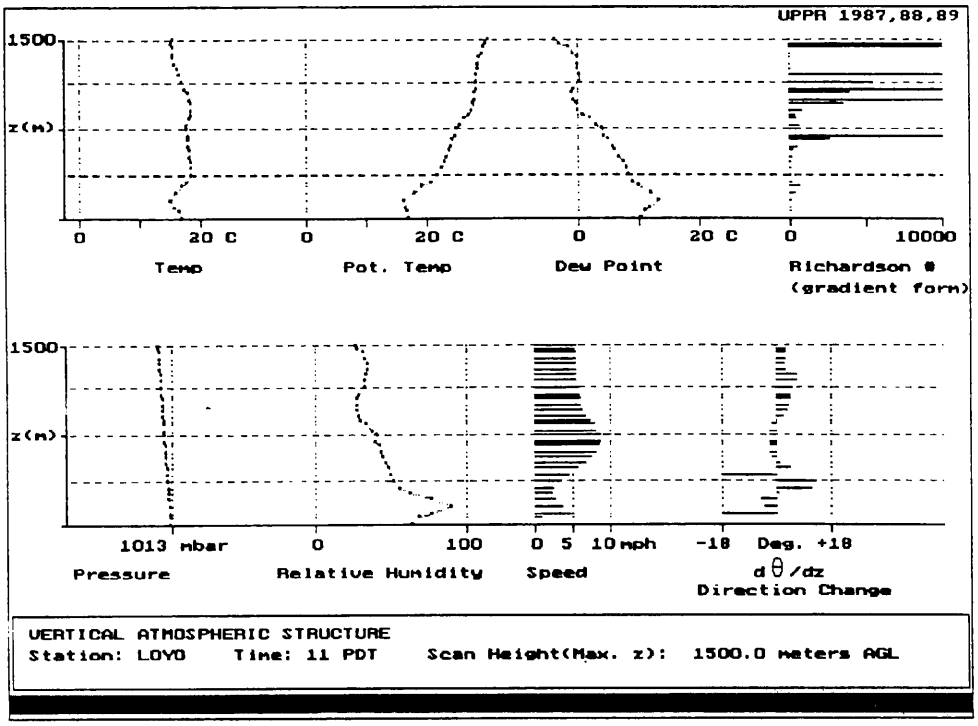


Figure 65. (Top) Vertical atmospheric structure at Loyola Marymount University at 05:00 PDT, December 11, 1987.

Figure 66. (Bottom) Vertical atmospheric structure at Loyola Marymount University at 08:00 PDT, December 11, 1987.



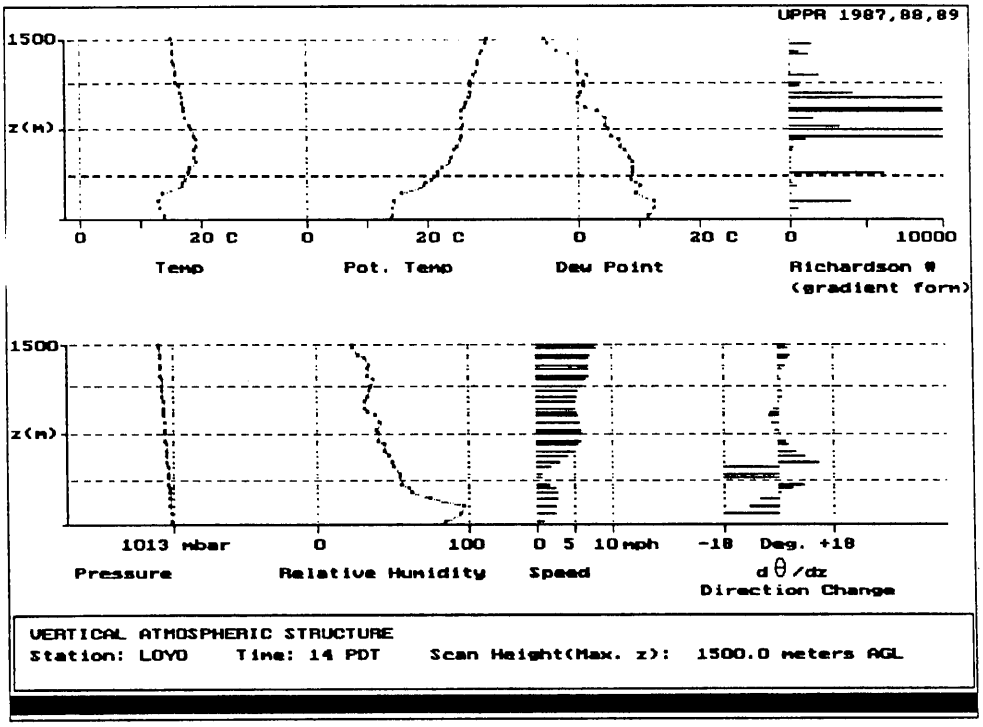


Figure 67. (Top) Vertical atmospheric structure at Loyola Marymount University at 11:00 PDT, December 11, 1987.

Figure 68. (Bottom) Vertical atmospheric structure at Loyola Marymount University at 14:00 PDT, December 11, 1987.

References

1. Beychock M. R. (1979), Fundamentals of Stack Gas Dispersion.

2. Blumenthal D. L. et al. (1978), Anatomy of a Los Angeles smog episode: pollutant transport in the daytime sea breeze regime, Atmos. Env., **12**, 893-907.

3. Edinger J. G. (1959), Changes in the depth of the marine layer over the Los Angeles Basin, J. of Met., **16** No. 3, 219-226.

4. Edinger J. G. (1973), Vertical distribution of Photochemical smog in Los Angeles Basin, **7** No. 7, 247-252.

5. EQL (1982), McRae G. J., Goodin W. R. and Seinfeld J. H., Environmental Quality Laboratory report 18, California Institute of Technology.

6. Horrell R. S., Deem M., Wyckoff P., Shair F. H. and Crawford N., (1989), Ground release SF_6 tracer experiments used to characterize transport and dispersion of atmospherice pollutants during the Southern California Air Quality Study of 1987, Air and Waste Annual Meeting paper 89-138.2.

7. Lea D.A. (1968), Vertical ozone distribution in the lower troposphere near an urban pollution complex, J. Appl. Met., **7** No. 2., 252-267.

8. Reible D. (1981), Thesis, California Institute of Technology.

9. Seinfeld J. H., 1975), Air Pollution Physical and Chemical Fundamentals, McGraw-Hill.

10. Shair et al. (1982), Transport and disperison of airborne pollutants associated with the land breeze-sea breeze system.

11. Stephens E. R. (1975), Chemistry and meteorology in an air pollution episode, J. of the Air Pollution Control Association, **25** No.5, 521-524.

Section 2.3

Submitted for Publication in Atmospheric Environment.

2.3 A Smog Season Examination of the LA – Glendale Divergence Zone

Abstract

The persistent smog season divergence zone which extends from north of the Pasadena/Glendale area through downtown Los Angeles and continues to the south west is examined by the use of surface wind data. The implications of this zone to emission transport from downtown Los Angeles and vicinity are explained in terms of trajectory analysis done for two days from the 1987 Southern California Air Quality Study (SCAQS) and for the entire smog season of 1988. The results of the trajectory analysis are used to give perspective to two summertime SF_6 tracer releases near the zone as part of SCAQS, and to clarify transport pathways starting from other locations in the western basin.

Introduction

A significant amount of the air quality research to study the Los Angeles Basin involves meteorology. The basin's complex terrain in concert with semi-permanent synoptic scale pressure gradients, and a strong tendency for a temperature inversion to exist in this area, creates wind fields which are complicated, yet in many respects similar from day to day. For the purpose of this paper, only patterns during the smog season, from May through September, will be considered. During this time, the synoptic pressure gradients resulting from the Pacific high pressure cell and the differential temperature between the ocean and land directs daytime air flow onshore (Edinger, 1967). The temperature inversion is often aloft during this time of year and exhibits a distinct diurnal variation in height (Edinger, 1959). On average, the base of the inversion ranges around one kilometer above ground level, creating a vertically restricted layer in which the winds must flow. Permanent obstructions such as hills, mountains, and valleys in this horizontal layer tend to redirect the flow horizontally and give the flow its characteristic features. Finally, the surface and upper air flow patterns, whether they be specific to a single day or summarized for an entire season, are of general interest because they describe the bulk transport within the basin.

Portrayal of the Los Angeles Basin's wind patterns has been done qualitatively by two major methods: (1) streamlines and vector maps, and (2) air parcel trajectories plots. Because the wind field is dynamic, (1) and (2) above are not the same. The flow streamlines over the basin at each hour of the day have been compiled for different circumstances from air monitoring stations within the basin. Examples are the Southern California Air Quality Management District's (see SCAQMD 1977) compilation of monthly patterns and the California Air Resources Board's (see CARB 1984) compilation of seasonal patterns in California. The trajectories of air parcels have been studied from two different perspectives: (1) indirect studies that computed trajectories from surface winds and winds

aloft, and (2) direct studies that determined air parcel movement by labeling. For example, a comparison of direct and indirect methods was done by Neiburger by using surface wind trajectories computed by several techniques and the actual transport of air parcels labeled with zinc-cadmium sulfide particles. Results were at best inconclusive because of the enormous loss of particles to unknown sinks, and the subsequent range limit (a few tens of km) to which transport could be tracked. A similar study by Calvert (1976) used radar tracked tetroons to follow air parcels in an attempt to test the theory of ozone generation in the basin. More recently, air parcels have deliberately been labeled by means of inert gaseous tracers, a technique which has proven to be superior to other labeling methods (Lamb, 1978). In the Los Angeles Basin, sulfur hexafluoride tracer experiments were done from the city of Anaheim by Drivas (1974) to examine both dispersion and transport for distances up to 100 km. More recently, the land breeze-sea breeze system of Los Angeles was studied using sulfur hexafluoride by Shair (1982). These tracer experiments are very effective for determining the actual transport trajectories for particular days, but the large amount of time and man power required to do such experiments prevents them from being done in large numbers.

Investigators have been concerned with individual recurring features of the basin's wind field that were identified by the previously discussed methods. The intra-basin divergence and convergence zones are of particular interest. As mentioned above, Edinger examined the changes in the marine layer depth over the Los Angeles Basin during the summer of 1957. Within this paper, he noted the following three basic factors that produce changes in the marine layer: (1) horizontal convergence or divergence of the wind, (2) dilution of the marine layer from above by mixing superadjacent dry (inversion) air with the marine air, and (3) advection of deeper or shallower layers of marine air into the area. It is specifically the first factor that is to be examined further in this paper. Edinger

and Helvey (1961) quantified the San Fernando Convergence Zone and examined the impact it had on channeling smog generated from downtown Los Angeles for 75 days in the summers of 1956-57. They observed the zone and compiled statistics about its location throughout the summer. In addition, Edinger provided an idealized map of the known areas of convergence and divergence within the basin. His map labels the Elsinor convergence zone southeast of the Los Angeles Basin, and notes three divergence zones located approximately near the cities of Newhall, Glendale, and San Bernardino. More recently, in an effort to suggest ways for improving air pollution simulation models, Unger (1986) suggested that the influence of divergence and convergence zones on observed pollution levels may not have been adequately investigated. This paper presents evidence to support this suggestion and presents motivation to undertake future studies of intra-basin divergence and convergence zones.

The Glendale divergence zone is particularly interesting because it apparently divides the substantial emissions from downtown Los Angeles between to the two major downwind exits from the basin to the east and northwest. Appendix A shows the hydrocarbon emission distribution of mobile sources, and the population distribution of the basin (SCAQMD,1987). The hypothesis that the trajectory of emissions from the downtown area is determined by the location of this zone is partially supported by tracer experiments done near downtown. In a recent study, as part of the 1987 Southern California Air Quality Study, four sulfur hexafluoride tracer experiments were conducted just south of downtown Los Angeles, in Vernon during the summer and fall. A detailed summary of the results prepared by Horrell et al. (1989) describes the transport of the summertime tracer out of the basin to the north and northwest. These findings conflicted with the expectations of the authors and other persons designing the SCAQS studies. At the time of the SCAQS studies, it was anticipated that the tracer would exit primarily from the eastern outlet of the basin. This study was undertaken

to place into context the unanticipated findings of the two summertime SCAQS SF_6 tracer studies, and to quantify the dynamics of the LA-Glendale divergence zone during a 24-hour cycle.

Method

The WIND simulation package described by Horrell et al. (1989) was used as the primary research tool in this study. Briefly, this package of programs provides an interactive environment for the performance of simulated tracer experiments with the measured surface winds, and to monitor simultaneously the measured surface air quality data from a randomly distributed set of reporting stations. The advanced interface of this program allows simulated tracer releases to be performed quickly, making it practical to examine data from an entire month in only a few hours. For this study, the package has been set up to look at the Los Angeles Basin from 33 degrees 26 minutes to 34 degrees 29 minutes in latitude and from 117 degrees 6 minutes to 118 degrees 42 minutes in longitude (See Figure 1). Six sets of simulations and/or data compilations were done: (1) to determine the percentage of days the LA-Glendale divergence zone was present at 6 a.m. and 3 p.m., (2) to identify the average location of the LA-Glendale zone, (3) to determine the trajectories initiated from the nine western basin sites for the average smog season day, (4) to determine the trajectories from the nine western basin sites and compile statistics on their exit route from the basin for every day in the 1988 smog season, (5) same as (4), except using the two 1987 SCAQS days, and (6) to create contour maps of the average surface values for ozone and to explain features in these maps with regard to the location of the LA-Glendale divergence zone. The source of the data and method for preparing it for input into WIND are discussed below.

Surface wind data came from the AQMD data telemetry system for: (1) the two days of the summer SCAQS period of 1987 (July 15 and September 3) and (2) all days during the months of May through September 1988. Of the 153 days in the 1988 smog season, 151 were used; June 3 and July 8 were missing from the data set. The surface wind data is received as hourly averaged data and is reported as a speed in miles per hour and a direction in degrees off north. The

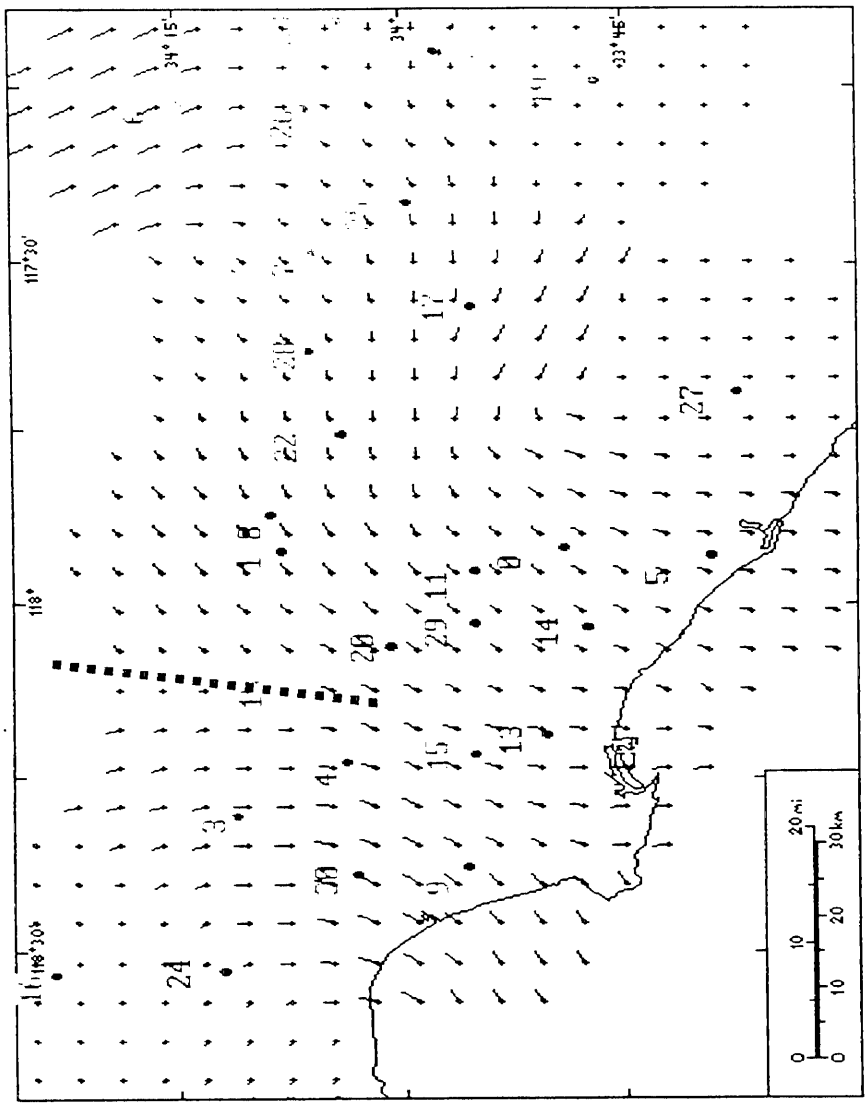


Figure 1. Hourly averaged wind vectors for the 1988 composite average day during the smog season at 8:00 a.m. Vectors are shown 5 km apart, their length proportional to the wind speed, and their direction away from their origin is the direction of flow. The bold dashed line through Pasadena (site 18) and east of downtown LA (site 4) shows the location of the divergence zone at this time.

data is grouped as single days, one value for each of the 24 hours at each of the reporting stations. Thirty-one data reporting sites were operational during this study (see Table 1 for a detailed list of the site names and their location). In addition to the surface wind data, hourly averaged ozone data was obtained from the AQMD for all 31 stations in a similar format.

The surface wind data was then prepared as input data for the WIND simulator. The data was converted from its raw text format into binary files each containing a block of data for a single month. From the wind data, one additional data set was generated containing the vector-averaged hourly winds for the entire period from May through September of 1988 and the scalar averaged ozone for the same period. This created a composite day (31 sites by 24 hours, where each hour is the average for that hour for all 1988 smog season days). In this study, the composite day was used to do preliminary assessments of the data set, to look for patterns and trends in the wind fields and simulated tracer releases, and to compare with individual days from 1988 and other years. The WIND program treats this composite day just like any other day.

Trajectories and Vector Maps

WIND was set up to use a linear interpolation between reporting stations to develop the wind field as described in Horrell (1989). Though inverse square law relationships have been recommended by other authors (Seinfeld et al., 1982) for interpolating wind data, testing with various power laws for the set of simulations failed to establish additional accuracy for any particular power. Consequently, a linear interpolation was chosen because of the associated ease of calculation and interpretation. These experiments were conducted from nine points distributed in the western basin (see Table 2). These sites were chosen for two reasons. First, they surround the sulfur hexafluoride tracer release site. Second, they bound the major emission region associated with the downtown Los Angeles area. Simulated releases were conducted at 6 a.m. and 3 p.m. Pacific Daylight

Station Number	Abbrev.	Station Name	Latitude	Longitude
0	ANAH	Anaheim	33.818	117.919
1	AZUS	Azusa	34.133	117.923
2	BANN	Banning	33.9	116.83
3	BURK	Burbank	34.184	118.304
4	CELA	Central LA	34.063	118.228
5	COST	Costa Mesa	33.652	117.929
6	CRES	Crestline	34.243	117.274
7	FONT	Fontana	34.096	117.491
8	GLEN	Glendora	34.186	117.936
9	HAWT	Hawthorne	33.926	118.376
10	HEME	Hemet	33.733	116.950
11	LAHB	La Habra	33.919	117.951
12	LANC	Lancaster	34.667	118.133
13	LGBH	Long Beach	33.837	118.205
14	LSAL	Los Alamitos	33.792	118.033
15	LYNN	Lynwood	33.917	118.215
16	NEWL	Newhall	34.388	118.533
17	NORC	Norco	33.921	117.570
18	PASA	Pasadena	34.132	118.126
19	PERI	Perris	33.778	117.244
20	PICO	Pico Rivera	34.013	118.059
21	PLSP	Palm Springs	33.8	116.60
22	POMA	Pomona	34.066	117.754
23	RDLD	Redlands	34.066	117.159
24	RESE	Reseda	34.199	118.528
25	RIVR	Riverside	33.993	117.418
26	SNBO	San Bernardino	34.104	117.283
27	TORO	El Toro	33.621	117.691
28	UPLA	Upland	34.103	117.634
29	WHIT	Whittier	33.918	118.026
30	WSLA	West LA	34.051	118.389

Table 1. Air Quality and Meteorological Reporting Stations

Time	Percent
8 a.m.	50
3 p.m.	88

Table 2. Percentage of Days where the LA – Glendale Divergence Zone is Identifiable at a Particular Time

Time (PDT). These times were chosen to label emissions from the morning and evening rush hours.

The trajectory simulations all had the same protocol. At each site, tracer was released for three hours. This means one trajectory was started at each hour, and it steps ahead in time at one hour intervals. The trajectory away from the release site was plotted as a non-diffusing point carried by the resultant winds for a period of 12 hours. Twelve hours was chosen because a subset of test experiments showed that the residence time of morning emissions from this region was on the order of 12 hours, and a similar value of 10 hours was obtained from the actual SF_6 tracer experiments. Trajectories were computed using the recursive trajectory method described in Horrell et al. (1989) and carried out via its implementation in the WIND program.

Visualization Method for Locating Divergence Zones

The WIND package computes and displays maps that highlight regions of angular spreading in the surface wind fields and assist in identifying divergence zones. The method for this will be discussed in this chapter. The method was developed to provide a quick, objective, visual way of identifying regions of angular spreading in the surface wind vector field.

For the maps to be displayed below, the region of interest was divided into grid squares 2.5 km on a side. A value was computed for every grid square by using the following rule. The angular spreading associated with a particular grid square was determined by computing the dot product of the velocity vectors from the four corners of that grid square. Six combinations of vectors result when all corners are dotted with their neighbors and diagonals (see below).

$$a \cdot \quad \cdot b$$

$$c \cdot \quad \cdot d$$

The angle between the two vectors was determined by using Equation (1).

$$\theta = cos^{-1}\left[\frac{\mathbf{A}\cdot\mathbf{B}}{|A||B|}\right] \tag{2.3.1}$$

The pair yielding the largest angle was chosen to be the value for a given grid square. The relative direction of the two vectors was determined geometrically, i.e., whether the vectors pointed towards, away, or parallel to one another. The sign of the angle was assigned negative if the two vectors pointed away from one another or were parallel and was assigned positive if they pointed towards one another. The positive angle differences indicate the presence and location of divergence zones. WIND graphically displays the results of these calculations by dividing each individual grid square's value by the maximum observed in the entire region. Then, a false color image of those values that are larger than the fractions 0.08, 0.11, 0.14, and 0.17 is created. These fractions were found to highlight most clearly the divergence zones in the basin when the method was developed.

Unfortunately, some of the visual contrast of the images is lost when they are converted to black and white. The darkest squares are experiencing the strongest divergence, whereas the lightest squares show the domain of the calculation. Those shades of gray in between are lower levels of divergence. The divergence maps should be compared with the vector maps to assist in reading them.

One alternative method for identifying regions of high angular spread in the surface wind field was considered. This method is based on the fact that locations of divergence would have negative accumulation when a line integration was performed around them. When first considered, this approach seemed promising since large regions (10 km edge squares) where the surface winds diverged nominally showed negative accumulation. However, when attempted on smaller scale grid squares, 2.5 km edges, this approach was unsuccessful for two main reasons. First, negative accumulation occurs more commonly where the flow accelerates in the direction of flow than where it diverges. Second, the flow

decelerates near stagnation regions, which gives positive accumulation even in regions of angular spreading. Both problems tend to obscure the identification of the actual divergence zones.

Results

Wind Fields

The first use of the WIND program was to create and display the wind fields at each hour from the 1988 composite average day for the purpose of locating the divergence zone (Simulation 1). Two hours from the composite day, 8 a.m. and 3 p.m., are shown in Figures 1 and 2, respectively. Figures 1 and 2 show the LA-Glendale divergence zone marked with a bold dashed line.

Divergence

As part of this preliminary investigation into the location of the LA-Glendale divergence zone, the two-dimensional angular divergence in the surface winds was computed at each hour over the region of interest. Similar to the wind vectors, the computation was done for the entire region under consideration and at each of the 24 hours. These images were converted into the pictures, examples of which are seen in Figures 3 and 4. These figures correspond to the same times as the vector maps in Figures 1 and 2.

Observed Frequency of the LA-Glendale Divergence Zone

The percentage of time the LA-Glendale divergence zone was present at 8 a.m. and at 3 p.m. for the 151 days of the 1988 data set was determined from WIND generated wind vector maps. (See Table 2.) The wind vector maps were studied visually to determine if the zone was present. In cases where the zone barely appeared, the angular divergence was computed, displayed as discussed previously, and used as an objective means of deciding whether the zone should be counted. If the zone appeared on the map as clear line, several grid squares thick, it was counted as being present. If it did not appear clearly, or if it appeared as a scattered pattern of grid squares, it was counted as absent. The presence or absence of a zone was clarified in all marginal cases using this technique.

Simulation 2 revealed that the zone was much more likely to be identified in

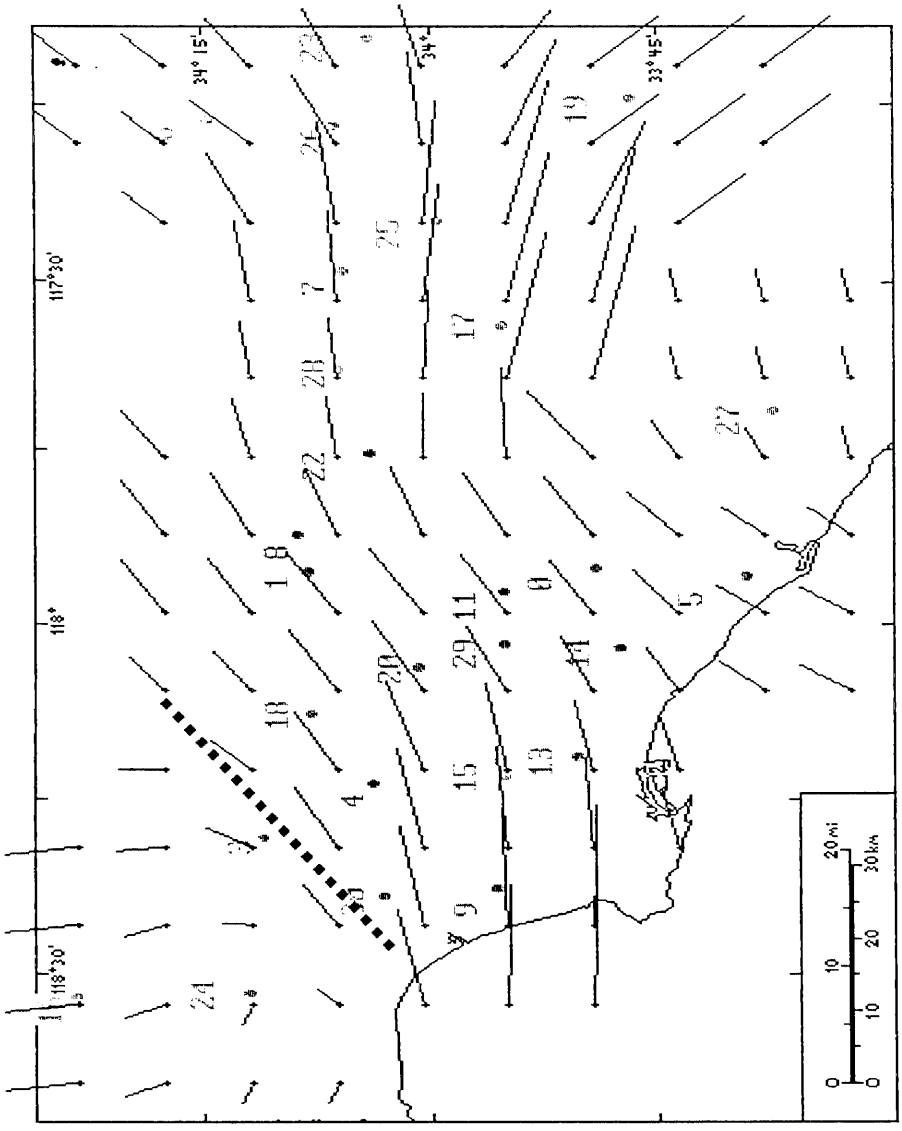


Figure 2. Hourly averaged wind vectors for the 1988 composite average day during the smog season at 3:00 p.m. Vectors are shown 10 km apart, their length proportional to the wind speed, and their direction away from their origin is the direction of flow. The bold dashed line to the west of Pasadena (site 18) and downtown LA (site 4) shows the location of the divergence zone at this time.

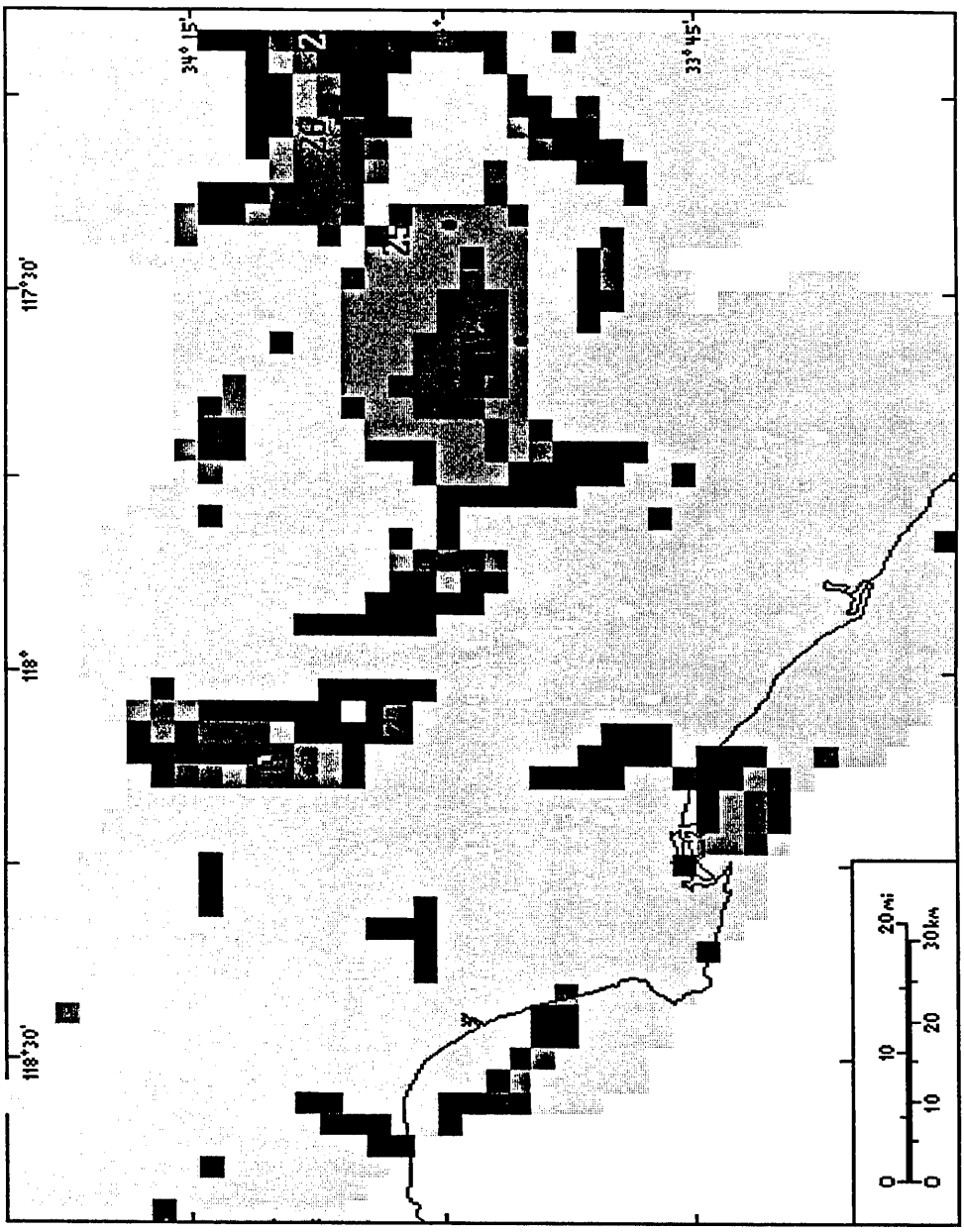


Figure 3. Angular divergence map of the western Los Angeles Basin for the 1988 composite average day during the smog season of 1988 at 8:00 a.m. Grid size is 2.5 km on an edge. The black regions represent locations of maximum angular divergence, the lightest shading indicates the total region in which angular divergence is calculated, shades in between the lightest shading and the black indicate areas of slight angular divergence.

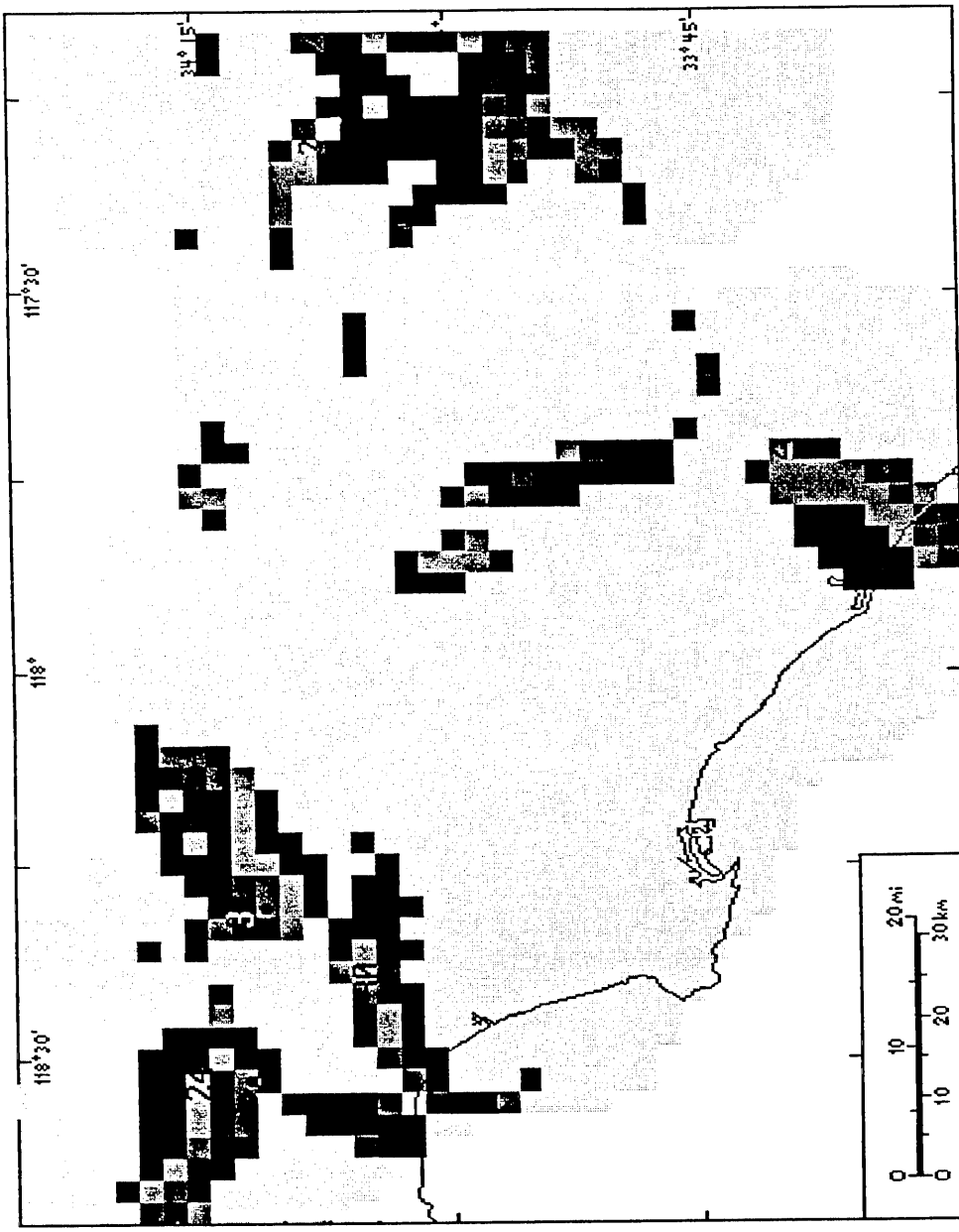


Figure 4: Angular divergence map of the western Los Angeles Basin for the 1988 composite average day during the smog season of 1988 at 3:00 pm. Grid size is 2.5 km on an edge. The black regions represent locations of maximum angular divergence, the lightest shading indicates the total region in which angular divergence is calculated, shades in between the lightest shading and the black indicate areas of slight angular divergence.

the afternoon than in the morning. The zone was identifiable around 50 percent of the time at 8 a.m., but was seen clearly 90 percent of the time by 3 p.m. At 3 p.m., the zone had often moved to be centered over station 3 or even to be between station 24 and 3 (and from there extended down to station 30). In contrast, if the zone was observed during the morning, it was most often located over station 18. An occasional day put the zone to the east of station 18, but the zone was never identified west of station 4 in the morning. The information provided by these statistics explains why the observed SF_6 transport might have been unexpected. Prior to the SCAQS, the morning period near downtown was relatively uncharacterized, while the afternoon period was more well understood. If the flow patterns for the afternoon were used as a guideline, one would conclude tracer would flow to the eastern basin.

Surface Trajectories

Average 1988 Day

For comparison, simulated tracer experiments to elucidate source-receptor relationships were performed using the average 1988 day (Simulation 3). As mentioned, two sets of releases were done, the first began at 6 a.m. and the second at 3 p.m.. The resulting trajectories are shown in Figures 5 (a)-(d) and 6 (a)-(d) for 6 a.m. starts and Figures 7 (a)-(d) and 8 (a)-(d) for the 3 p.m. starts.

Individual Days of 1988 and 1987

To understand how the dynamics of the LA-Glendale divergence zone's movement affects the trajectory of downtown LA emission on individual days, a set of simulated tracer experiments was done on each day of the 1988 smog season (Simulation 4) and on all of the 1987 days (Simulation 5) mentioned above. Four sites were selected that created a square region that both surrounded the average mid-morning location of the divergence zone and encompassed the region of massive morning emissions associated with downtown LA. These are a subset of the sites shown in Table 3 and correspond roughly with data reporting stations 4, 30, 9, and 15.

Statistics were then compiled for these 151 days by performing a simple data reduction. Each simulated release was assigned to 1 of 16 bins according to the direction, relative to the release point, the trajectory exited from the basin. The 16 bins correspond to the 16 points of the compass centered in the 16 equipartioned pie slices of a circle. Ambigous cases were resolved on the basis of the direction of the tracer at its last hour or two. It should be noted that less than one percent of the cases were ambigous, and that not a single trajectory failed to exhibit a strong directional sense by which it could be characterized. From the three releases associated with each site, a composite bin value was determined for that site by averaging.

Further processing of the binned data was done to assist in interpreting the

Release Point Description	SiteCode	Latitude	Longitude
Near Site 30	s1	34 4'	118 23'
Site 4	s2	34 4'	118 14'
N of Site 20	s3	34 4'	118 4'
Near Site 9	s4	33 57'	118 23'
Near Site 15	s5	33 57'	118 14'
NW of Site 29	s6	33 57'	118 4'
Coast Line	s7	33 48'	118 23'
SW of Site 13	s8	33 48'	118 14'
W of Site 14	s9	33 48'	118 4'

Table 3. Release Sites in the Western Los Angeles Basin

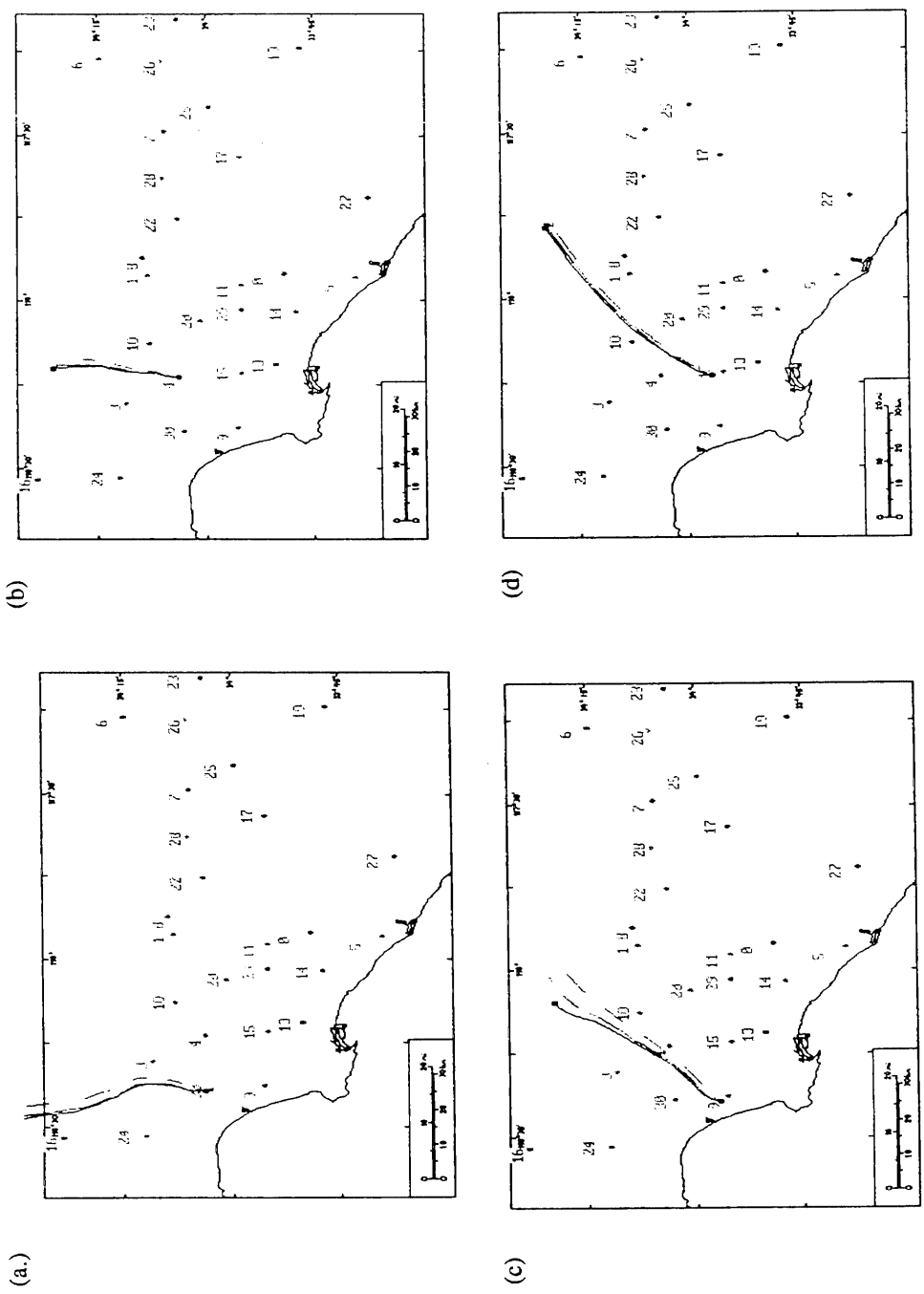


Figure 5. Simulated 12 hour tracer releases initiated at 6,7, and 8 a.m. for 4 sites in the western Los Angeles Basin using the 1988 average smog season winds, (a) West Los Angeles near site 30, (b) downtown LA site 4, (c) Hawthorne near site 9, and (d) Lynwood near site 15. Only the trajectory lines are depicted.

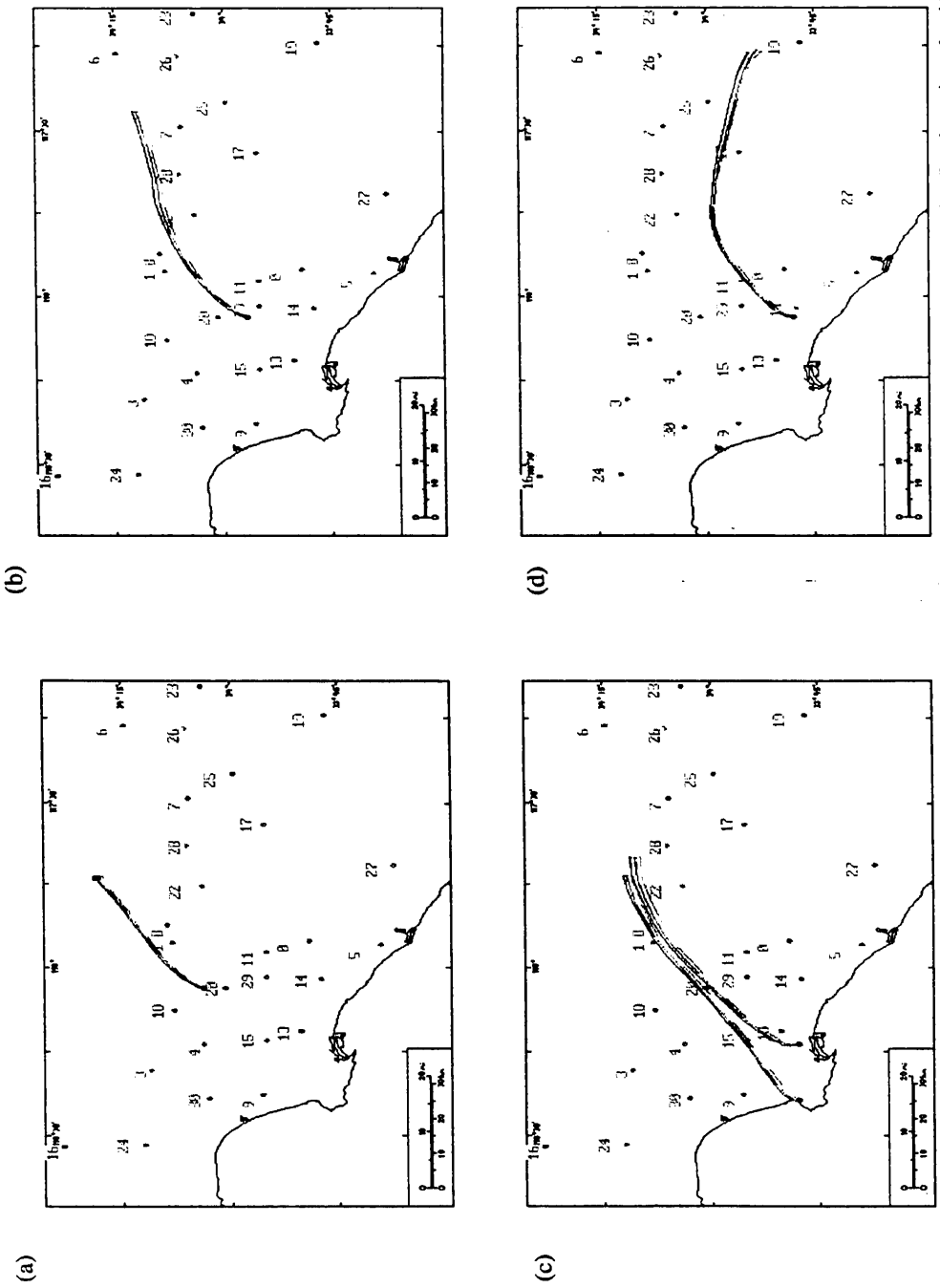


Figure 6. Simulated 12 hour tracer releases initiated at 6, 7, and 8 p.m. for 5 sites in the western Los Angeles Basin using the 1988 average smog season winds, (a) Pico Rivera near site 20, (b) Whittier near site 29, (c) southwest coastline and Long Beach near site 13, and (d) Los Alamitos near site 14. Only the trajectory lines are depicted.

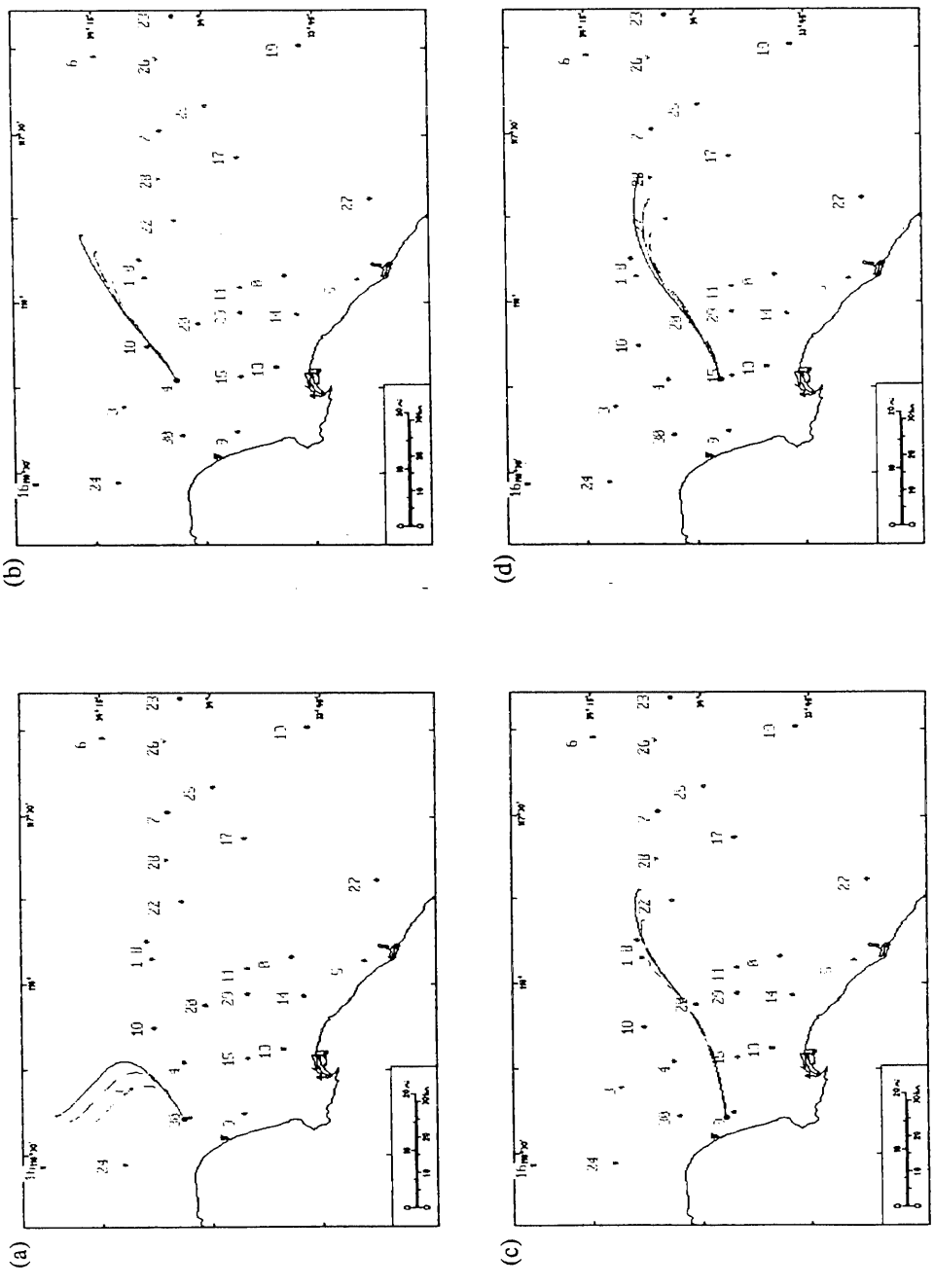


Figure7. Simulated 12 hour tracer releases initiated at 3,4, and 5 p.m. for 4 sites in the western Los Angeles Basin using the average smog season winds, (a) West Los Angeles near site 30, (b) downtown LA site 4, (c) Hawthorne near site 9, (d) Lynwood near site 15. Only the trajectory lines are depicted.

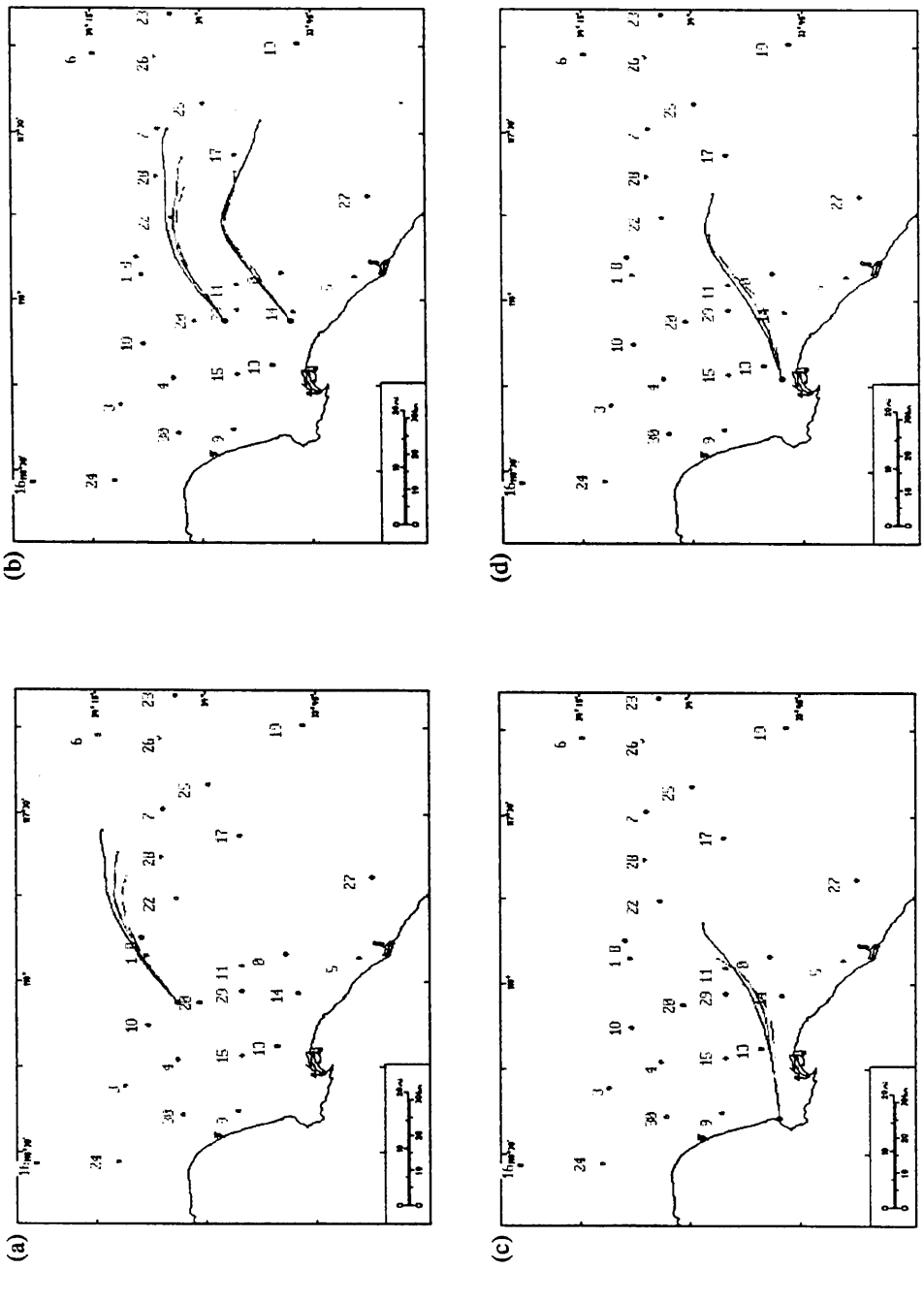


Figure 8. Simulated 12 hour tracer releases initiated at 3,4, and 5 p.m. for 5 sites in the western Los Angeles Basin using the 1988 average smog season winds, (a) Pico Rivera near site 20, (b) Los Alamitos near site 14 and Whittier near site 29, (c) southwest coastline, and (d) Long Beach near site 13. Only the trajectory lines are depicted.

results. The composite bin value for each site was aggregated into six groups, which reflect the distribution of the four corners in the sixteen bins on a single day. One category is the days when zero of the four corners are transported to the NE, E/NE, or E. The second category is the days when one of the four corners are transported to NE, E/NE, or E, and that one corner is one of the two southernmost sites, 9 or 15. The third category is the days when the two southernmost corners (sites 9 and 15) are transported NE, E/NE, or E. The fourth category is the days when three corners (except site 30) are transported to the NE, E/NE, or E. The fifth category is the days when all four corners are transported to the NE, E/NE, or E. The sixth category is comprised of all the days that do not fit in any of the above categories. Figure 9 shows the the results from simulation 4 as a percentage of the 151 1988 days examined.

In Simulation 5, the two days of the 1987 SCAQS had trajectories simulated for them in the same fashion as was done for the individual days of 1988. Both days were found to be in category one above, where none of the four corners was transported into the eastern basin.

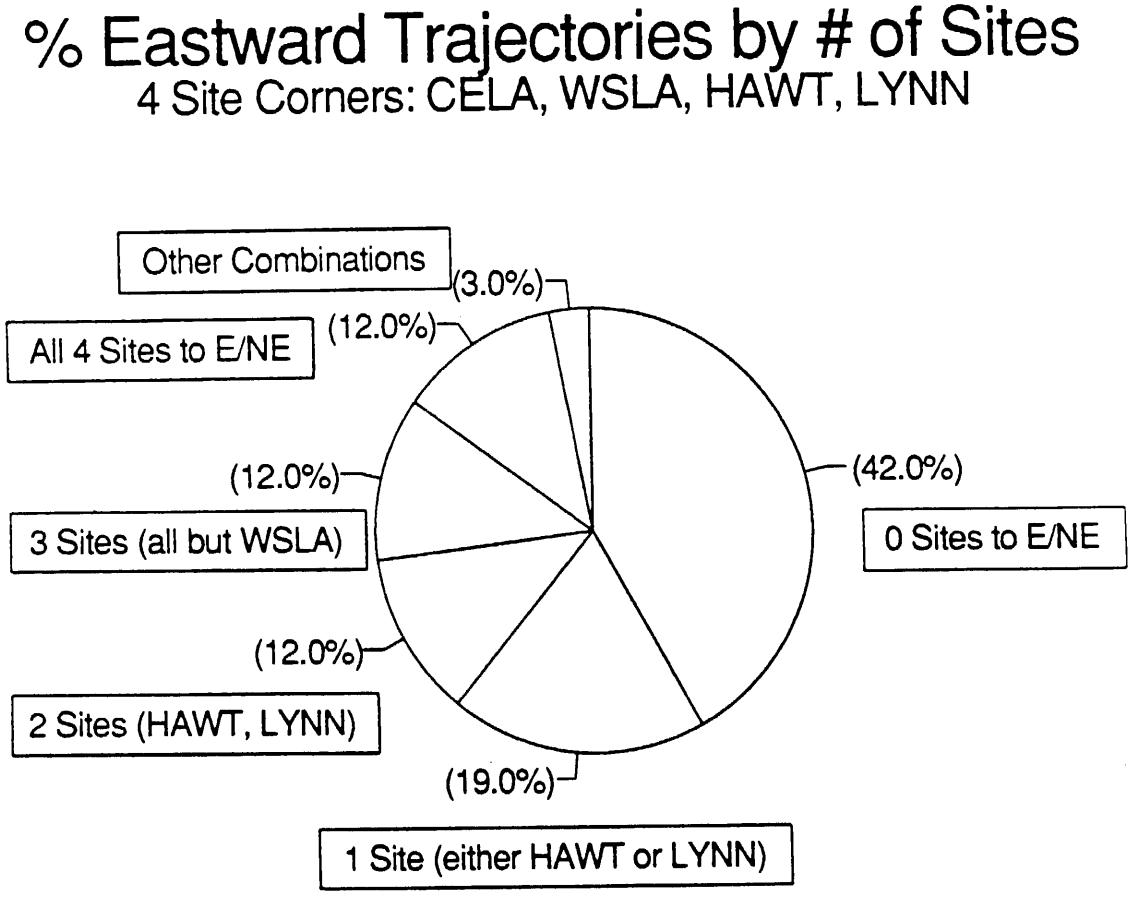


Figure 9. Percentage of simulated western Los Angeles Basin trajectories exiting the basin to the E/NE from 4 selected sites (downtown Los Angeles, West Los Angeles, Hawthorne, and Lynwood). Each trajectory started at 6 a.m. The 4 sites are grouped by the number of sites which produce trajectories that exit the basin to the E\NE. The data is 151 days of the 1988 smog season from May to September.

Dynamics of the LA-Glendale Divergence Zone

Based on examination of the divergence maps created from the composite average smog season day of 1988 (Simulation 1), it can be said that LA-Glendale divergence zone is a feature of the basin's wind field for about 21 hours in a day. The zone is also dynamic, moving 30 km east to west and back over the course of the composite day.

Figure 10 depicts the dynamics of the zone and labels various positions the zone occupies. In the average day the zone showed the following pattern of movement. It appeared at 7 a.m., at which time it lay along a nearly perfect north-south line originating between station 4 and station 20, position (a). The zone remains at this location until around 10 a.m.. At this time, the zone begins to elongate to the southwest stretching from station 4 towards stations 9 and 30 passing through positions (b) and (c) to position (d). At 11 a.m. the zone extends from the coast around station 9 inland to station 18 (Pasadena) and then north beyond Pasadena stabalizing in position (d). At around 1 p.m., the zone begins to migrate west at all locations so that it extends from the coast north of station 30 to north of station 18, position (e). Between 2 and 3 p.m., the zone moves to a new location which lies north of station 30 and extends through station 3 through to a point about 10 km north of station 18, position (f). A broad double peak occurs in the divergence map from between station 24 and station 18, indicating the zone is equally likely in any number of places between station 3 and station 18. The zone remains in this position until 6 p.m. when it starts to drift back eastward at all points about 5 km, between position (e) and (f). The length of the zone begins to shorten at this time also, as it recedes from the coastline. At 8 p.m., the zone extends from station 30 south of station 3, north of station 4 to a point north of station 18, between position (e) and (d). The zone slowly moves eastward, at about 1 to 2 km an hour until 3 a.m. at which time it spans from station 30 to station 4 then to station 18. At 4 a.m.

the zone begins to disappear and from 5-6 a.m., the zone is gone entirely.

Analogy

A simple analogy for the basin is proposed. Under this proposition, the western basin is considered analogous to a Y-shaped pipe which has a hinged door at the branch of the Y. Offshore flow comes up the base of the Y, gathers emissions, and is split by the divergence zone into either of the northwestern or eastern outlet branches. Tracer experiments have demonstrated that transport is restricted across the zone, suggesting that the hinged door be considered a solid door, not a screen door. The fractional splitting of the downtown emissions into the two outlet streams is sensitive to the position of the hinged door. Subsequent smog development later in the day results, at least in part, from the fractional splitting. To complete the analogy, the position of the hinged door must be described. Its most likely position is due to factors such as the pressure gradients across the basin that result from the synoptic scale weather and differential heating at the outlets. Thus, the zone's position at any time represents the relative difference between the pressure at both outlets and the inlet.

The mechanism by which the LA-Glendale zone partitions the massive western basin emissions has a distinct influence on receptor locations for these emissions. On the average day, the LA-Glendale divergence zone sweeps across an area of approximately 800 square kilometers. On particular days, the zone appears to be able to sweep across as much as twice this area as shown by the demarcation lines for the extremes in the observed locations of the zone, Figure 10 positions (a') and (f'). This means that emissions may be routed to the northwestern basin in the morning while the zone occupies position (a) of Figure 10, but as the day proceeds the zone moves into position (f) and emissions are routed into the eastern basin. Because the zone directs a large portion of the western basin's emissions into two vastly different directions, a complete understanding of the zone is necessary to describe accurately the source-receptor relationships

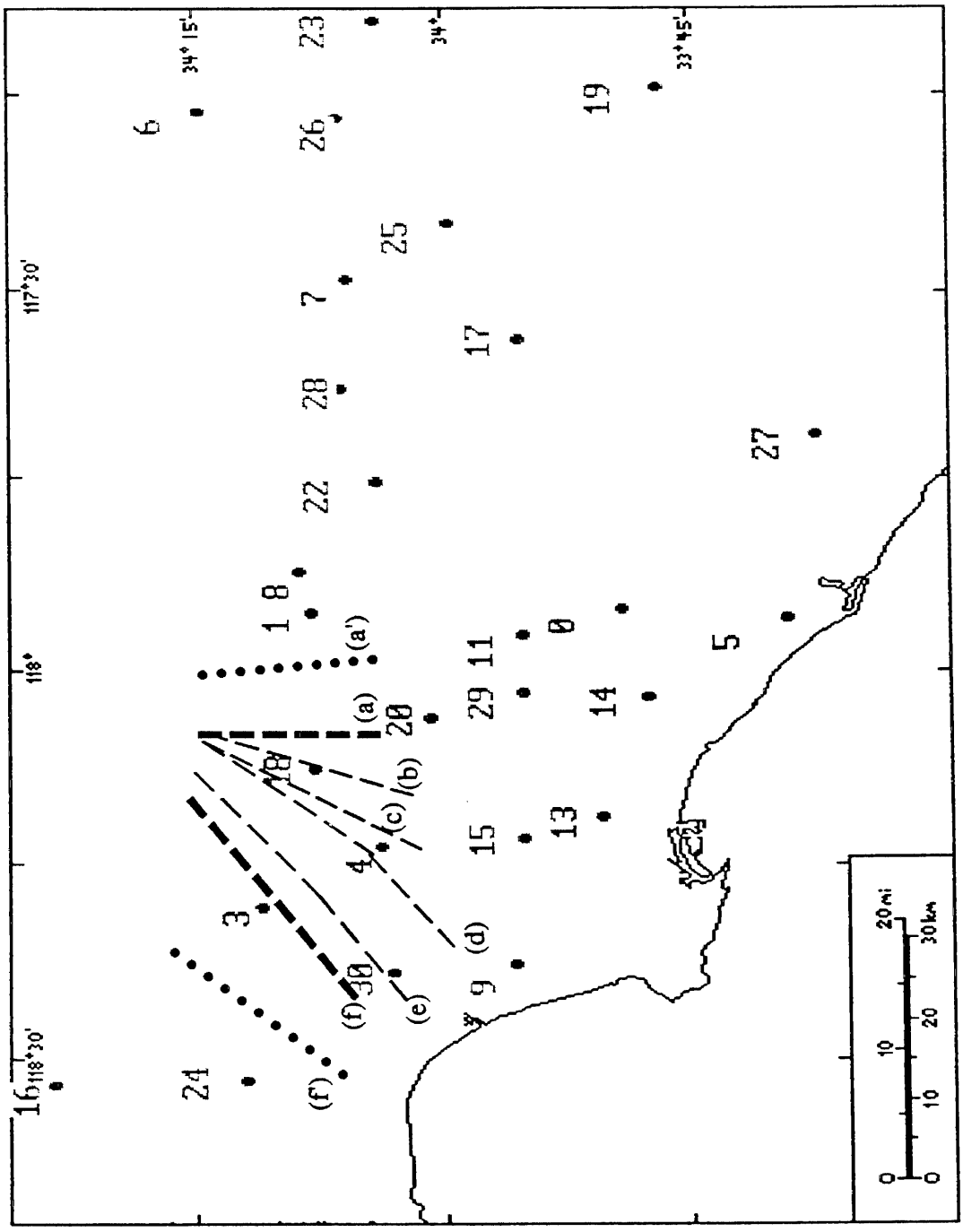


Figure 10. Positions of the LA-Glendale divergence zone based on observations made using the 1988 composite average day, sites (a) through (f), and total range of the zone's movement based on individual days (a') through (f').

between the western basin and the northwestern and eastern portions of the basin.

Two plausible conclusions can be drawn based on the above analogy and the trajectory analysis of the 1988 average smog season day. On average the higher observed ozone levels in the eastern portion of the Los Angeles Basin result from emissions made during the evening of the previous day. These are present when the emissions from the current morning are produced and then the mixture undergoes reaction. Elevated afternoon ozone levels in the San Fernando Valley (sites 3, 16, and 24) depend in part on the position of the divergence zone. Evening rush hour emissions are not generally directed to the northwest into the San Fernando Valley, and therefore do not impact morning ozone levels in that locale.

The trajectories generated in simulation 3 provide practical evidence of how the LA-Glendale divergence zone affects emission trajectories from sources in the western basin on average. Trajectories in Figures 5(b) and 6(a) demonstrate that on average the zone lies between the two release points depicted in these figures at 8 a.m.. This is seen by comparing the different directions the trajectories take away from the two sites. Examination of the same two sites in Figures 7(a) and 7(b), and the site depicted in Figure 8(a), shows that by 3 p.m. the LA-Glendale divergence zone has moved westward, and is now routing a majority of western basin emissions into the eastern basin. The results of simulation 4 where individual days are examined in the proximity of the zone will complement these results.

The results from simulation 4 show just how infrequently the emissions from the entire area encompassed by the four points in the northwestern basin had no impact on the eastern basin. Figure 9 summarizes the findings for the simulated morning releases. Forty-two percent of the time four corners fit the first category above. The second category accounted for 19 percent of the observations. In

contrast to the first category, only 12 percent of the time did the four corners fit the third, fourth, and fifth category. It should be noted that only three percent of the observations did not fit any of the categories.

It is interesting to note the route the trajectories took when they did not exit the basin through the E to NE directions. In excess of 95 percent of the trajectories went out the NNE, N, NNW and NW. This route sends the air parcels into the San Gabriel Mountains or into the San Fernando Valley, exactly where the SF_6 was observed in the 1987 studies.

Simulation 5 clearly demonstrated the observations of the summer SCAQS SF_6 tracer experiments were, in fact, not anomalous but representative of most smog season days. Both days fell into category one, which represents 42 percent of the total days. Furthermore, from the Vernon release site, a day could fall into categories two or three and still have a trajectory like those actually observed. This means up to 73 percent of the time emissions from the Vernon tracer release site are destined to exit the basin to the ENE, N, NNW, or NW.

Association Between Observed Ozone Levels and the LA-Glendale Divergence Zone

The ozone contour maps produced as part of simulation 6 clearly depict the geometrical relationship between the LA-Glendale divergence zone and the ambient ozone levels. These maps were generated using the 1988 smog season average ozone levels described previously. Twenty-four maps were generated, one for each hour of the average day, and two are presented here. At 8 a.m., the only region of the basin that is below the 0.1 pphm ozone level is the region right around site 3 (See Figure 11). This is approximately 15 km east of the divergence zone. At 3 p.m., a minimum band in the ozone lies along the line formed by the sites 9, 30, and 3. (See Figure 12.) This also corresponds to the location of the divergence zone. The position of this depressed ozone band and the divergence zone are both the result of the geography, but the minimum ozone band seems

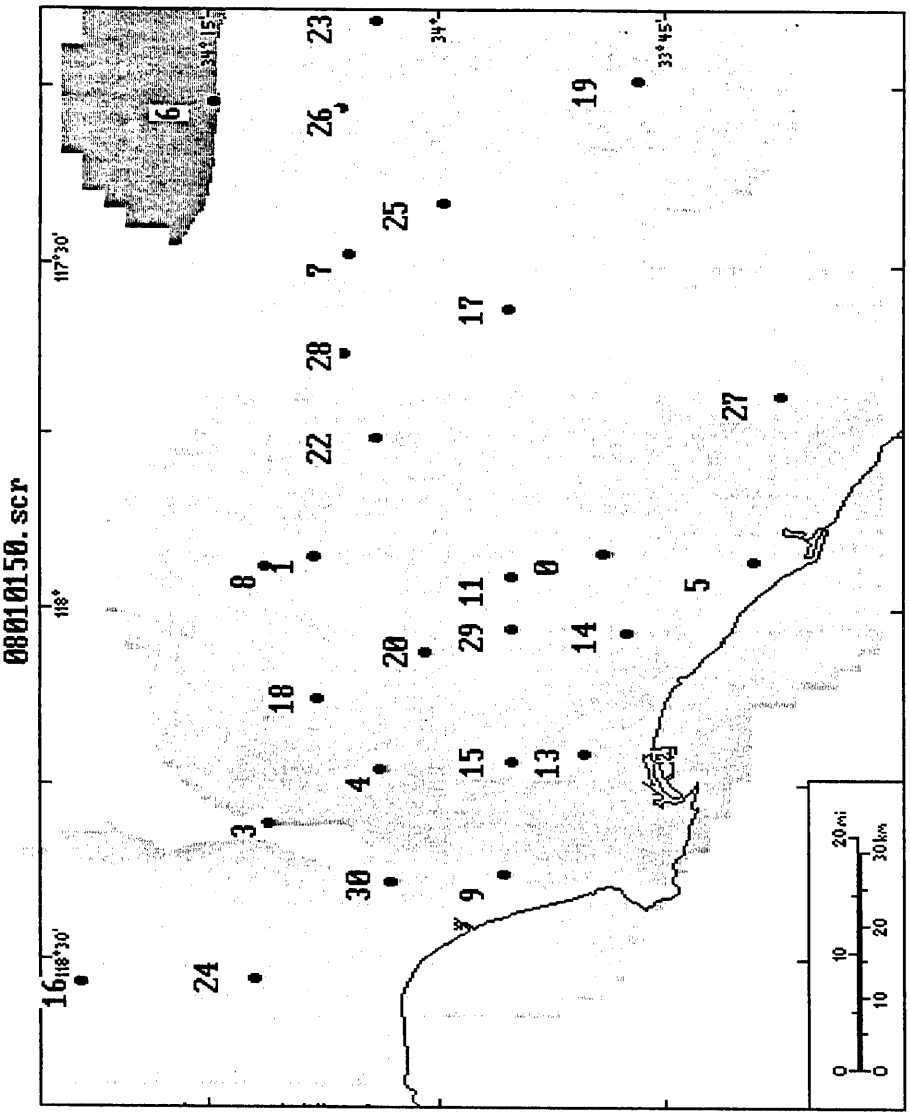


Figure 11. Hourly averaged ozone contour map for the 1988 composite averaged day during the smog season at 8 a.m. The contoured regions are, from lightest shading to darkest: (i) 0.1 -4.1 pphm, (ii) 4.1-8.1 pphm, (iii) 8.1-12.1 pphm, and (iv) 12.1-16.1 pphm. As a reference station 5 is in (i), and station 6 is in (ii). No higher levels are reported. Note the position of the ozone minimum just west of the location of the Los Angeles-Glendale divergence zone.

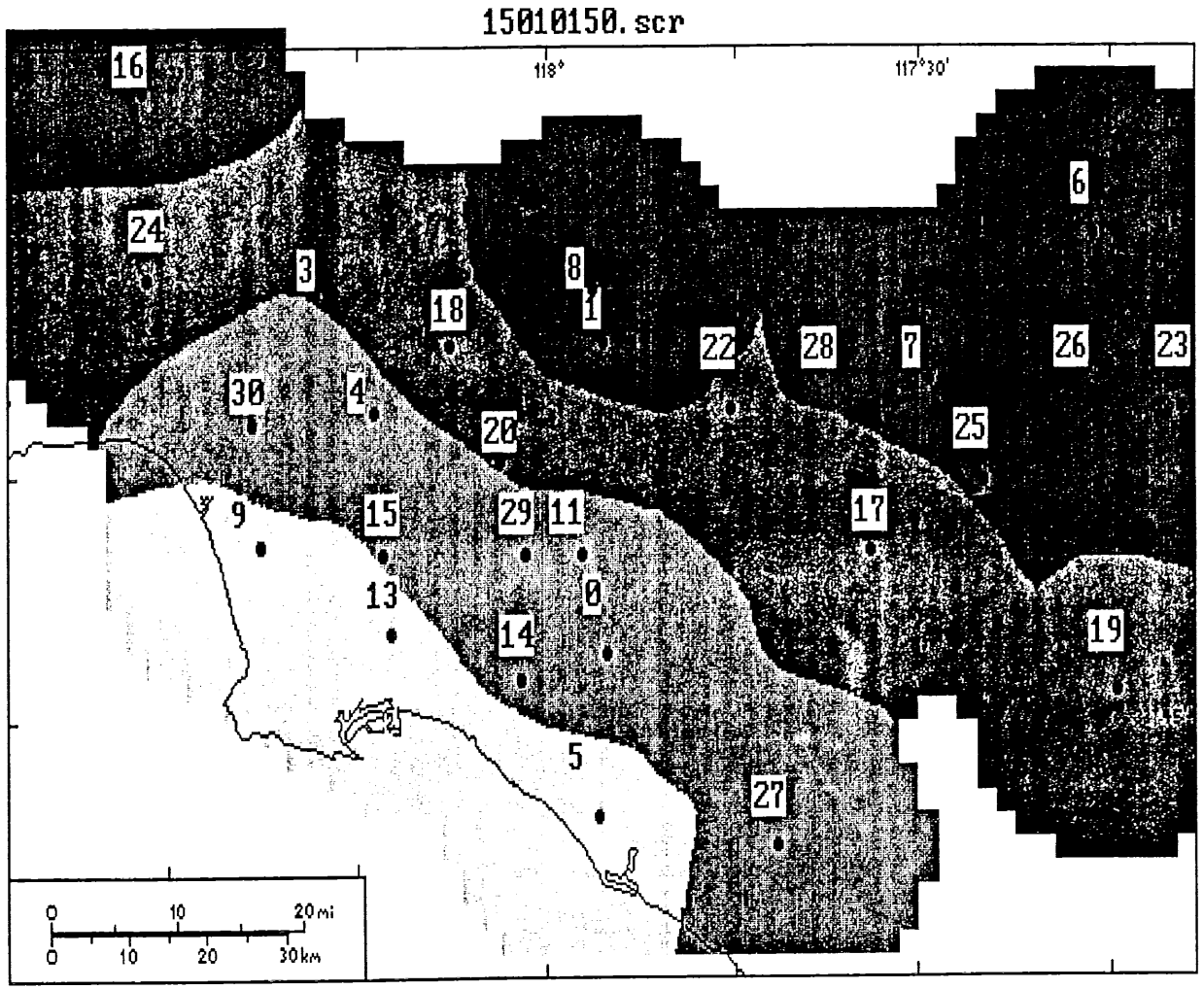


Figure 12. Hourly averaged ozone contour map for the 1988 composite averaged day during the smog season at 3 p.m. The contoured regions are, from lightest shading to darkest: (i) 0.1 -4.1 pphm, (ii) 4.1-8.1 pphm, (iii) 8.1-12.1 pphm, and (iv) 12.1-16.1 pphm. As a reference station 5 is in (i), station 27 is in (ii), station 19 is in (iii), and station 7 is in (iv). No higher levels are reported. Note the position of the ozone minimum just west of the location of the Los Angeles-Glendale divergence zone.

to be stationary and does not follow the divergence zone. This has been verified using the contour and wind field maps for the rest of the hours of the average day which were not depicted. Notice that the diminished ozone band is roughly at the same place on both maps independent of the location of the LA-Glendale divergence zone. Possibly the geography around site 3 and the lower population density of this area are responsible for these observations.

One additional observation that can be made from the ozone contour map of Figure 12, and from other maps not depicted, is the presence of an ozone maximum peak to the east of the divergence zone. This maximum is a persistent feature of the 1988 average ozone map for the hours of 10 a.m. to 6 p.m., and is located near sites 1 and 8 on the maps. One hypothesis that explains the presence of this maximum peak is that the three-dimensional structure of the eastern side of the divergence zone is a wake, or recycling cell, that is trapping emissions along the mountains that bound that portion of the basin to the north. Surface wind data alone is not sufficient to resolve definitively whether the above hypothesis is correct. The apparent association between the location of the divergence zone and the location of this maximum ozone peak is intriguing. Currently, insufficient data exists to confirm the observations, but perhaps future investigations will unravel any connection between them.

Summary

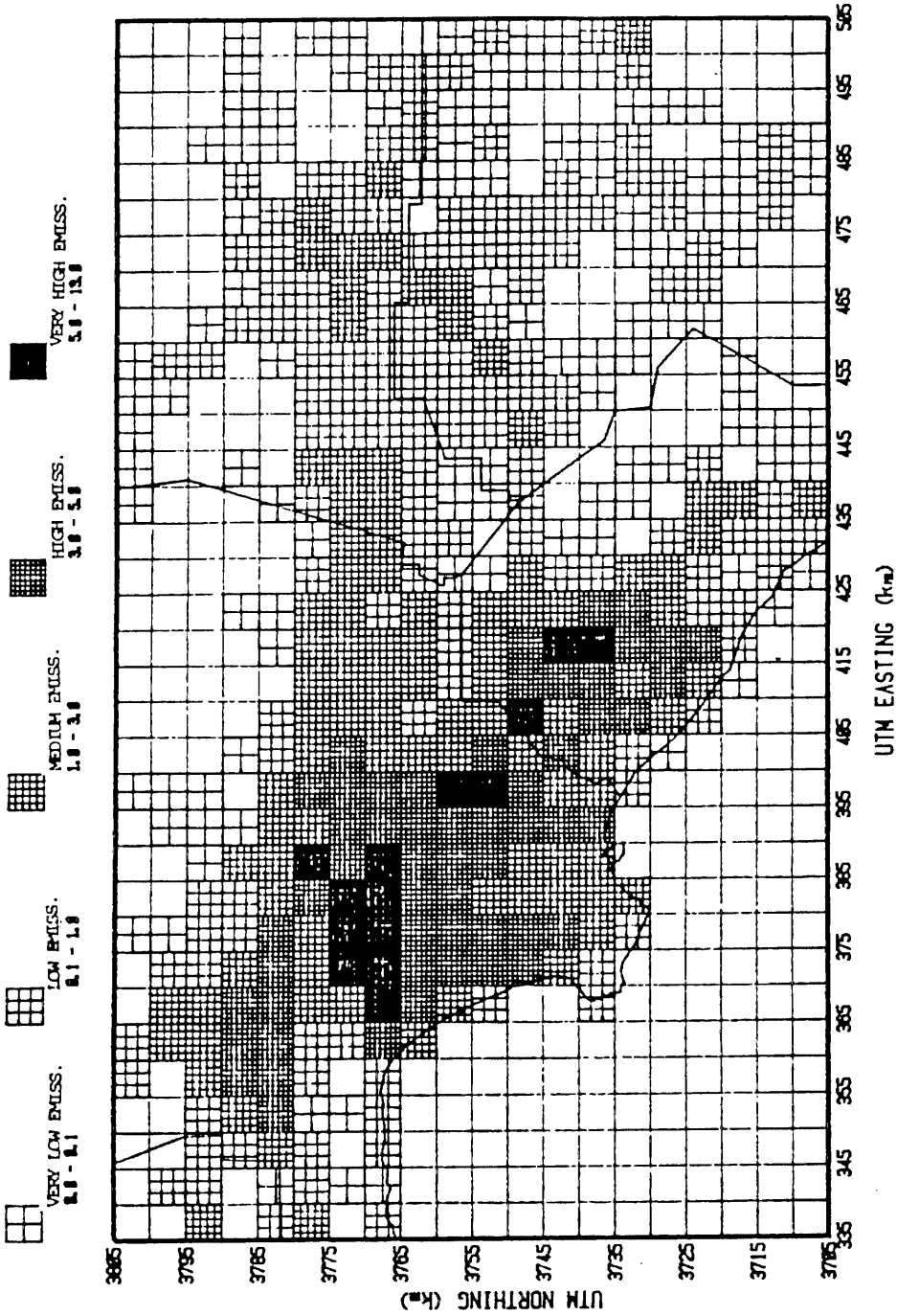
The six sets of simulations and compilations performed using the 1987 and 1988 wind and ozone data have elucidated the dynamic nature of the LA-Glendale divergence zone and established its persistence. Simulations 4 and 5 clarify that the transport observed during the two summer SCAQS studies were not anomalous observations, but in fact the norm for the smog season. The comparison of ozone data with the location of the zone in simulation 6 suggests a link between the ozone peaks around sites 1 and 8 and the eastern edge of the zone. The physics of the area around sites 1 and 8 has not been adequately studied so as to clarify the interaction between the divergence zone, the mountains, ozone precursors, and any other unknown factors.

This more complete understanding of the dynamics of the LA-Glendale divergence zone is likely to improve the overall understanding of the source-receptor relationship that exists the western basin and the receptors to the east and northwest. Finally, the exact location of this zone can be used as a sensitive diagnostic for confirming wind field models, because any model that does not accurately position the LA-Glendale zone is not correctly capturing the physics of the western Los Angeles Basin.

Acknowledgements

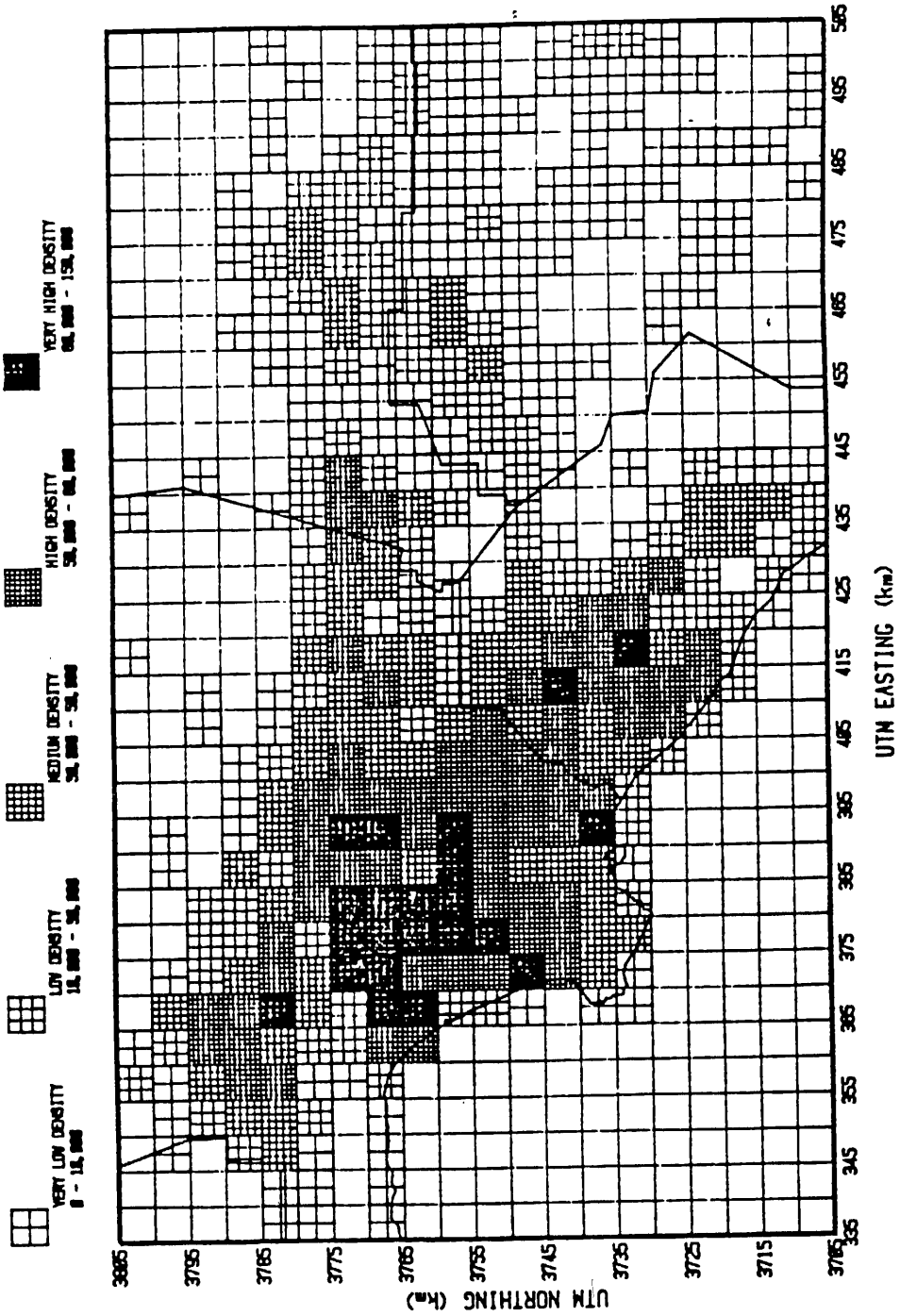
The authors wish to express their thanks to the technical staff of the California Air Resources Board, especially Chuck Bennet. Thanks are also made to the technical staff of the South Coast Air Quality Management District (SCAQMD) for supplying the data used in this analysis. Special thanks are made to Joseph Casmassi and Dr. Julia Lester, both of the SCAQMD, for their insightful suggestions. Partial Funding for this project came from the California Air Resources Contract A932-051.

A2.3 Appendix



DISTRIBUTION OF MOBILE SOURCE EMISSIONS OF REACTIVE ORGANIC GASES

Figure A.1. Taken from: The Magnitude of Ambient Air Toxics Impacts from Existing Sources in the South Coast Air Basin (SCAQMD 1987).



SPATIAL DISTRIBUTION OF POPULATION DENSITY

Figure A.2. Taken from: The Magnitude of Ambient Air Toxics Impacts from Existing Sources in the South Coast Air Basin (SCAQMD 1987).

References

1. Calvert J. G. (1976), "Test of the Theory of Ozone Generation in Los Angeles Atmosphere," Env. Sci. and Tech., **10** (3), pp. 248-255.
2. CARB (1984), *California Surface Wind Climatology*, State of California Air Resources Board Aerometric Data Division.
3. Edinger J. G. (1959), "Changes in the Depth of the Marine Layer over Los Angeles Basin," J. of Met., **16** (3), p. 219.
4. Edinger J. G. and Helvey R. A. (1961), "The San Fernando Convergence Zone," Bull. Am. Met. Soc., **42** (9), p. 626.
5. Edinger J.G. (1967) , Watching for the Wind, Doubleday and Company, Inc., NY.
6. Horrell R. S., Deem M., Wyckoff P., Shair F. and Crawford N. (1989), "Ground Release SF_6 Tracer Experiments Used to Characterize Transport and Dispersion of Atmospheric Pollutants During the Southern California Air Quality Study of 1987," Air and Waste Association 82nd Annual Meeting, paper 89-138.2.
7. Lamb B.K. (1978), Development and Application of Dual Atmospheric Tracer Techniques for the Characterization of Pollutant Transport and Dispersion, Doctoral Thesis, California Institute of Technology, Pasadena, CA 91125.
8. McRae G. J., Goodin W. R. and Seinfeld J. H. (1982), EQL Report Number 8, California Institute of Technology, Environmental Air Quality Laboratory.
9. Neiburger M. (1955), "Tracer Tests of the Accuracy of Trajectories Computed from the Observed Winds in the Los Angeles Area," Air Pollution Foundation,.
10. SCAQMD (1977), "California South Coast Air Basin Hourly Wind Flow Patterns," South Coast Air Quality Management District, El Monte, CA.
11. SCAQMD (1987), "The Magnitude of Ambient Air Toxics Impacts from Existing Sources in the South Coast Air Basin," South Coast Air Quality

Management District, El Monte, CA.

12. Shair et al. (1982), "Transport and Dispersion of Airborne Pollutants Associated with the Land Breeze-Sea Breeze System," Atmospheric Environment, **16** (9), pp. 2043-2053.

13. Unger C. D. (1986), "Meteorological Factors Which Affect Pollution Concentration," Research Division Draft Report, California Air Resources Board, Sacramento, CA 95812.

2.4 Examination of Wind Field Dynamics Using Cross Differencing

Abstract

A new visualization technique for rapidly assessing the dynamics in the wind flow at a single location is presented. This method is useful for making comparisons of the horizontal wind speed and direction of one day with (1) another day at the same location, (2) an average day (monthly, seasonally, yearly, etc.) at the same location, (3) the same day at different locations, and (4) the same day at the same location where each set of data comes from different sources such as raw, smoothed, and computer generated. The technique highlights similarities in the dynamics of different days which might otherwise be overlooked when analyzed on the basis of static wind fields alone. It can also be used to compare (rapidly and concisely) the accuracy of results generated by numerical models with field data. The technique is implemented to show a single day's winds which were collected in 24 one hour averaging periods from wind reporting sites in Los Angeles. The method is used to quantitatively examine wind direction transitions at 31 sites in Los Angeles and its surroundings during the smog season from May through September 1988.

Introduction

The wind field dynamics can be as important as the static features of the flow in determining transport patterns and source-receptor relationships. While this may seem evident, methods for studying wind flow fields are firmly rooted in visualization techniques which emphasize static conditions. This stems primarily from the fact that surface wind data is commonly and easily visualized by means of a two-dimensional vector map, an example of which is shown in Figure 1. This method of presentation places its emphasis on describing what the state of the wind flow field is at a specific time at all positions in the region of interest.

At least one additional step is required to go from the static vector map to the dynamics of the wind flow field. For many investigators, this additional step may not be an attractive option. For instance, a cumbersome option is the manual examination of a time series of vector maps such as those in California Surface Wind Climatology generated by the California Air Resources Board, (CARB,1984). However, where computing resources are available, it is common to plot air-parcel trajectories by advancing the parcel in accordance with a set of interpolation rules in position and time. This approach is instructive when considering the fate of a given parcel of air, but precludes the possibility of examining the dynamics at a specific location in the wind field. An additional problem with using trajectories is the steady accumulation of errors in position along the trajectory at each time step. These are only mitigated when the flow field is known to a high spatial and temporal resolution. Animation of the maps is very demonstrative, but it directly translates into the need for more sophisticated computer equipment.

However, an alternative method for visualizing wind dynamics exists which is derived directly from the time dependent variations in the flow. The technique is based on computing a signed difference between a given time series and a reference time series between all hours, and at a single location in the flow

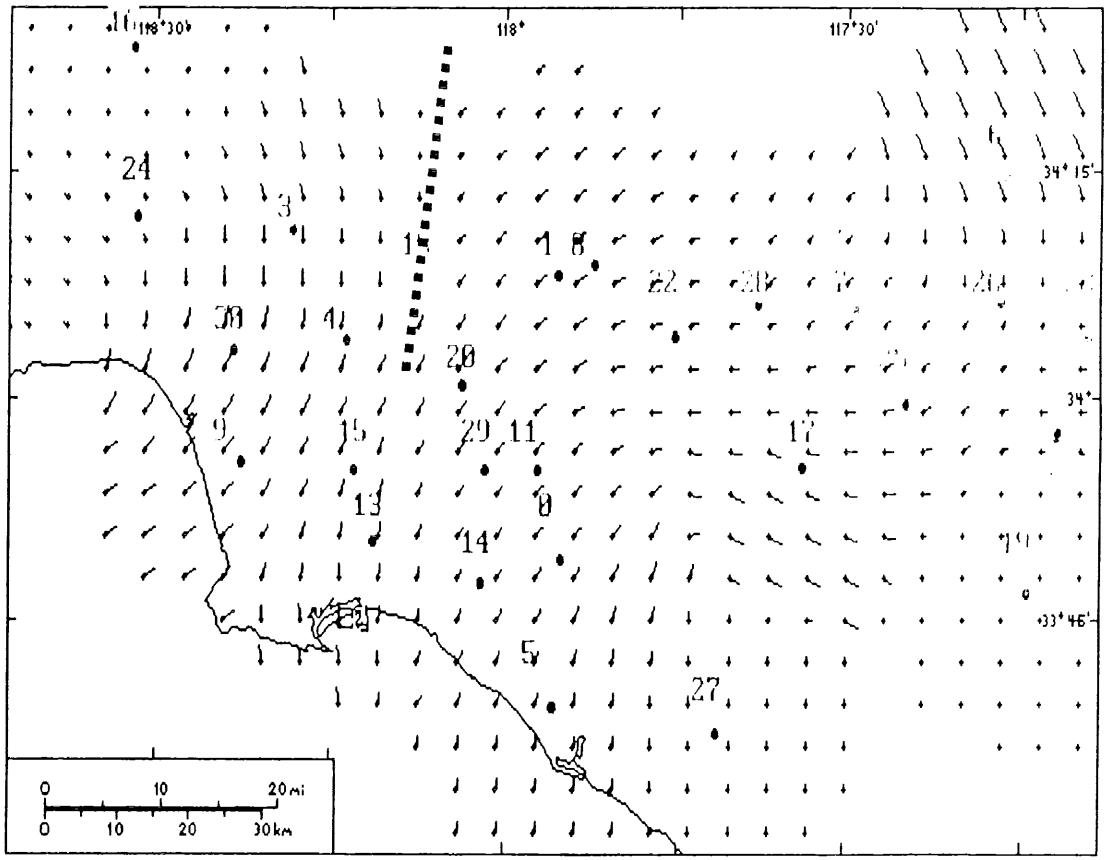


Figure 1. Hourly averaged wind vectors for the 1988 composite average day during the smog season at 8:00 a.m. Vectors are shown 5 km apart, their length proportional to the wind speed, and their direction away from their origin is the direction of flow. The bold dashed line through Pasadena (site 18) and east of downtown LA (site 4) shows the location of the divergence zone at this time.

field. The two types of times series which are considered are wind speed and direction. Graphically displaying the two-dimensional array which is created provides a geometrical way to quickly interpret the wind dynamics at that single location. From here on, this technique will be referred to as cross differencing. Cross differencing has the advantage that the location being examined can be a place where data is actually being taken, thus minimizing errors to those which correspond to the data acquisition methods. Cross differencing requires minimal calculation and can be implemented in any number of ways using general purpose computer equipment which is likely to be available, such as a PC or a workstation. The details of performing and displaying results from a cross differencing are the subject of the first part of this paper.

The second part of this paper focuses on using the cross differencing method to study the average smog season dynamics of 31 sites in the Los Angeles Basin, including the San Fernando Valley and the eastern desert in locations as distant from the coastline as Palm Springs. A detailed study of downtown Los Angeles' dynamics is required to clarify the complex interaction between emissions released during the morning rush hour and the timing of sharp wind direction changes that generally occur during the summertime morning hours and at some sites in the afternoon and evening. Such sharp directional changes are attributed to the interactions between the macroscale synoptic scale pressure differences which result from the earth's spin and differential heating of the earth's surface and mesoscale phenomena due to local geography and energy transport. The mesoscale phenomena include the well-studied land breeze-sea breeze system, e.g., Defant (1951), Scorer (1978), Shair (1982), the mountain valley system e.g., Schroeder (1960), and Blumen (1984), slope wind e.g., Banta (1981), Doran (1983), Horrell (1987, 1990), Kondo (1988), and the less well-studied convergence and divergence zones Edinger (1959) and Horrell (1990). These mesoscale systems are a persistent feature in the coastal regions of California and exhibit

considerable influence over the wind flows during certain times of the day. This is due in part to the orientation and close proximity of the coastal mountain ranges to the ocean.

The third portion of this paper focuses on identifying the unique features of the surface flow at three locations in the basin (Downtown Los Angeles, Burbank, and Upland) on two days of the 1987 Southern California Air Quality Study (July 15, and September 3). These were days on which the authors released sulfur hexafluoride tracer. The two sulfur hexafluoride tracer experiments performed from the downtown area (Horrell 1989) provided quantitative evidence on how the low speed morning transition winds disperse and route rush hour emissions from this release site out of the basin. The results from these two experiments provide supporting evidence for the theory that the time when the daily flow field establishes itself in the downtown area is as important as the specific features of the wind field in determining source-receptor relationships on certain days. The reason for studying the detailed dynamics of downtown Los Angeles, Burbank, and Upland is to further validate this theory and refine current understanding of the mechanism by which transport and dispersion occur within this area of the basin.

Method

Prior to mathematically describing the cross differencing technique, a simple intuitive discussion of the method is provided. The starting point is a time series, having 24 elements, each element representing the one hour averaged wind speed or direction at every hour of the day. A second time series having the same number of elements as the first is used as a reference. The question is how do these two time series compare to one another on the basis of shifts in time, the difference in magnitudes of individual elements at identical times, and the similarity of elements at different times. A simple method to establish the answers to all these questions simultaneously is to subtract one element of the first time series from each element of the reference time series. Thus, from two time series a two-dimensional array is created, in which each element is the difference described above. The diagonal elements of the matrix are a measure of the difference between the two series at identical times, and the rest of the elements describe the difference between a single series element and all the other elements in the reference series. If one examines the matrix, regions of similarity have the lowest absolute magnitude, whereas regions of high absolute difference are dissimilar. The orientation of these regions of small and large difference provides information about the relative timing of particular features between the two series. Finally, to avoid the generally undesirable task of examining a matrix of numbers, a graphical technique is used to display the results. The rest of this section will be used to explain mathematically how to take the difference, normalize the differences for plotting, graphically present the results, and interpret various patterns in the plots. This combination of tasks is referred to as cross differencing.

Mathematical Description

Cross differencing is described mathematically by the following relationship.

$$\Delta V[q(t1),r(t_2)] \;=\; (Vq(t_1) \;\triangle\; Vr(t_2) \;=\; \frac{[Vq(t_1) - Vr(t_2)]}{\mathrm{Max}|\mathrm{V[q,r]}|}\Bigg|_{t_1=0\;\;t_2=0}^{t_1=n\;\;t_2=n} \tag{2.4.1}$$

Where $\Delta V[q(t_1),r(t_2)]$ is the n by n element cross difference matrix, $Vq(t_1)$ is the time series being compared, and $Vr(t_2)$ is the reference series. Cross differencing between two time series will be denoted by the operator $\triangle$, and will be written in the above format. The series being compared will always appear first (left of the operator) and the reference series will appear second (right of the operator). The index variables t_1 and t_2 will independently range from some initial time, generally $t_1 \;=\; t_2 \;= 0$ to some final time in distinct intervals of seconds, minutes, hours, etc. The largest absolute difference between any element in $V[q(t_1),r(t_2)]$ is denoted by

$$\mathrm{Max\ V[q,r]} \;=\; \mathrm{Max\ Vq(t_1) - Vr(t_2)}\Big|_{t_1=0\;\;t_2=0}^{t_1=n\;\;t_2=n}. \tag{2.4.2}$$

Normalizing the raw difference between series elements by the absolute maximum guarantees all elements of the cross difference are in the interval [-1,1]. This rescaling provides an invariant size range when presenting the data in graphical form.

Several other definitions are made which will be used later. The maximum diagonal difference for a pair of variables $Vq(t_1)$ and $Vr(t_2)$ is defined as:

$$\mathrm{DiaMax\ V[q,r]} \;=\; \mathrm{Max\,|Vq(t_1) - Vr(t_2)|}\Big|_{t_1=t_2=0}^{t_1=t_2=n} \tag{2.4.3}$$

Similarly, the average diagonal difference for a pair of variables $Vq(t_1)$ and $Vr(t_2)$ is defined as:

$$\overline{\mathrm{Dia\ V[q,r]}} \;=\; \sum_{t_1=t_2=0}^{t_1=t_2=n} \frac{\mathrm{Max\,|Vq(t_1) - Vr(t_2)|}}{n} \tag{2.4.4}$$

Cross differencing for hourly averaged data, for a single day, using wind speed S and angle from the north T takes on the following form:

$$(Sq\ \Delta\ Sr) = \frac{[Sq(t_1) - Sr(t_2)]}{\text{Max S[q,r]}} \Bigg|_{t_1=0\ \ t_2=0}^{t_1=23\ t_2=23} \tag{2.4.5}$$

$$(Tq\ \Delta\ Tr) = \frac{[Tq(t_1) - Tr(t_2)]}{\text{Max T[q,r]}} \Bigg|_{t_1=0\ \ t_2=0}^{t_1=23\ t_2=23} \tag{2.4.6}$$

Note that $(Sq\ \Delta\ Sr)$ and $(Tq\ \Delta\ Tr)$ are 24 by 24 matrices. For the examples which follow, it is assumed that the series consist of hourly averaged data. The averaging process was done forward in time such that the index $t_1 = 0$ signifies the average from midnight to 1 a.m. taken at one minute intervals, $t_1 = 1$ the average from 1 to 2 a.m., and so forth all the way up to $t_1 = 23$. The same holds true for t_2.

Two time constants are defined and will be used below. The first is the duration time constant, t_d. It is the length of time over which a series remains at a constant value. The second is the transient time constant, t_t. It is the length of time over which a series exhibits change and falls between two sub-series of constant value. Typically, every element of the wind speed and direction time series to be examined below can be assigned to a t_d or t_t sub-sequence of the main sequence. The cross diffferencing plots highlight the boundaries between the constant and transient sub-sequences.

Examples of Cross Differencing

At this point in the discussion, the best way to become familiar with cross differencing and the selection process for references relevant to a particular task is by example. The examples will build in complexity, starting with some very idealized conditions, and will end with more complex cases which are analogous to the real cases to be encountered further in this paper. The examples are designed to educate the reader about how to spot the salient features which occur in cross difference plots involving wind speed and direction. The features which are presented may help the reader to perceive the following relationships between series: when the series (1) are identical, (2) are the same at certain times, (3) are similar but shifted in time from one another, (4) are similar in dilated or compressed time, (5) consist of similar discrete regions of constant value and transients between them, and (6) the series have coinciding extrema.

It is necessary to keep in mind, while going through these examples, how a sensible choice of a reference is really the key to obtaining useful information. In particular, a sensible reference has the property of selectively highlighting the information being searched for in the comparison series. In this respect, cross differencing acts as a selective filter, and the examples will be referred to by the type of filtering which they perform. At the same time, the reference must be simple enough for the information which is highlighted to be identified.

Method of Data Presentation

Before proceeding, it should be noted that the graphs are presented in pairs; the left one is the speed cross difference and the right one the direction (angle) cross difference. The magnitude of each element of the cross difference array is displayed as a circle; the bigger the size of the element, the larger the size of the circle. There are six circle sizes, each one represents one sixth of the range between zero and one. Six is not a special number, but it was chosen because of graphical display constraints. Recall large array elements mean a big

difference exists between the two series elements being compared. Conversely, the small elements are regions of small difference. A very noteworthy observation within the following plots is the case where the diagonal contains small elements. This means the two sequences are the same. The method of plotting means the diagonal runs from the lower left corner of the axis to the upper right corner. More will be said about this below.

This method has several inherent advantages for graphically displaying the magnitude of each element. First, this type of plot conveys information clearly when limited only to black print and single color page reproduction. It is also possible to represent magnitude using a uniform sized region by varying the shade and/or color of the region. However, in this case it is currently difficult to have the color images published and, even supposing that they were, future reproduction of the pictures with conventional copying techniques leads to pictures with indiscernable shades. Finally, the graphical method has precedent in current literature. Analogs of this method are being used in the scientific literature in fields such as neurocomputing and electrical engineering to represent system states.

Opened and filled circles are used in the plots to denote the sign of the difference. For the plots presented herein, only the speeds are signed. The angle differences are given as an absolute value. Positive differences are denoted with black. Thus, the angle differences and the positive speed differences are all black. Negative values are denoted by white, which has a black rim to distinguish it from the background. (The key on each plot says blue, again a remnant of the computer screen image from which these pictures are taken.) Finally, the diagonal is denoted by a gray edge in speed plots and a solid gray image in angle plots. (On copying this may tend toward black, which presents no risk of information loss.)

In these plots, the axis is always oriented the same way with the reference series time along the abscissa and the comparison series along the ordinate. This

provides a simple rule for analyzing the series being compared. For a given hour m of the ordinate, one reads the plot horizontally along a line of circles. If the circle at a particular abscissa location n is large, then the comparison series differs from the reference series at the time found by vertically scanning down to the abscissa. Similarly, a small circle at (m,n) means the comparison series at time m is like the reference at time n. For the purpose of this paper, moving from left to right and down to up on the axis corresponds to going from early morning to late evening.

Visually, the method provides only relative information about the two data sets being compared. Thus, it is useful to display simultaneously absolute numbers along with the plots. In order to gauge the size of the largest circles, the maximum difference between any two elements in the matrix is displayed (as defined by Equation 2.4.2). This value is also a measure of the dynamic range for that variable. The maximum diagonal difference, Equation 2.4.3, and the average diagonal difference, Equation 2.4.4, are also displayed. These provide a measure of the difference between the two series. Each of these three numbers is reported below the plots.

Filtering for Constant Flow Speed and Direction

The first example, Figure 2, is a cross difference performed between two series which have uniform properties in time. Refer to Table 2a (wind speed) and 2b (wind direction) which show the numerical values for each of the examples. The simplest cross reference is a uniform data set with itself, in this case two wind speeds, ($S1 \triangle S1$). Similarly, a cross difference between two different but constant in time data sets produces a uniform but non-zero matrix; in this case it is two wind directions, ($T1 \triangle T2$). The results are shown on the right half of Figure 2. These two cases demonstrate the two extremes in the cross differencing

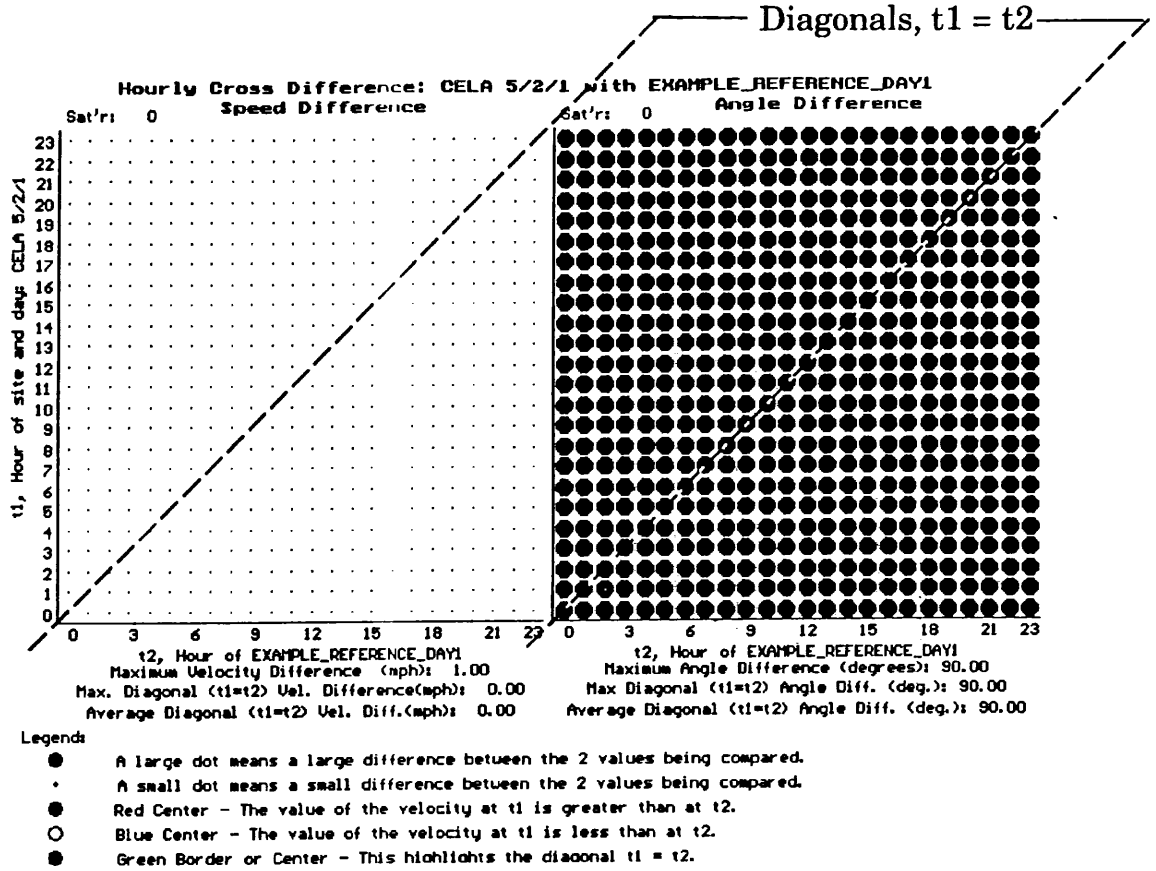


Figure 2. Cross differencing example 1, (left) speed: uniform for all 24 hours [S1 table 2] cross differenced with itself [S1 table 2] as a reference, (right) direction: uniform for all 24 hours at 180 degrees [T2 table 2] cross differenced with a uniform direction at 90 degrees [T1 table 2] at all 24 hours as a reference. (Note: CELA 5/2/1 is the label given to the comparison day consisting of speed S1 and direction T2. EXAMPLE_REFERENCE_DAY1 is speed S1 and direction T1.)

Station Number	Abbrev.	Station Name	Latitude	Longitude
0	ANAH	Anaheim	33.818	117.919
1	AZUS	Azusa	34.133	117.923
2	BANN	Banning	33.9	116.83
3	BURK	Burbank	34.184	118.304
4	CELA	Central LA	34.063	118.228
5	COST	Costa Mesa	33.652	117.929
6	CRES	Crestline	34.243	117.274
7	FONT	Fontana	34.096	117.491
8	GLEN	Glendora	34.186	117.936
9	HAWT	Hawthorne	33.926	118.376
10	HEME	Hemet	33.733	116.950
11	LAHB	La Habra	33.919	117.951
12	LANC	Lancaster	34.667	118.133
13	LGBH	Long Beach	33.837	118.205
14	LSAL	Los Alamitos	33.792	118.033
15	LYNN	Lynwood	33.917	118.215
16	NEWL	Newhall	34.388	118.533
17	NORC	Norco	33.921	117.570
18	PASA	Pasadena	34.132	118.126
19	PERI	Perris	33.778	117.244
20	PICO	Pico Rivera	34.013	118.059
21	PLSP	Palm Springs	33.8	116.60
22	POMA	Pomona	34.066	117.754
23	RDLD	Redlands	34.066	117.159
24	RESE	Reseda	34.199	118.528
25	RIVR	Riverside	33.993	117.418
26	SNBO	San Bernardino	34.104	117.283
27	TORO	El Toro	33.621	117.691
28	UPLA	Upland	34.103	117.634
29	WHIT	Whittier	33.918	118.026
30	WSLA	West LA	34.051	118.389

Table 1. Air Quality and Meteorological Reporting Stations

(a) SPEED [S]

Example: #	Site	Time: Hour 0	1	2	3	4	5	6	7	8	9	10	11
S1	CELA	7	7	7	7	7	7	7	7	7	7	7	7
S2	CELA	0	1	2	3	4	5	6	7	8	9	10	11
S3	CELA	0	1	2	3	4	5	6	7	8	9	10	11
S4	CELA	0	1	2	3	4	5	6	7	8	9	10	10
S5	CELA	0	2	4	6	8	10	12	10	8	6	4	2
S6	CELA	0	1	2	3	4	5	6	7	7	7	7	7
S7	CELA	4	5	6	7	8	9	10	11	12	11	10	9

#	Site	Hour 12	13	14	15	16	17	18	19	20	21	22	23
S1	CELA	7	7	7	7	7	7	7	7	7	7	7	7
S2	CELA	12	13	14	15	16	17	18	19	20	21	22	23
S3	CELA	12	11	10	9	8	7	6	5	4	3	2	1
S4	CELA	10	10	10	9	8	7	6	5	4	3	2	1
S5	CELA	1	1	1	1	1	1	1	1	1	1	1	0
S6	CELA	7	7	7	7	7	7	6	5	4	3	2	1
S7	CELA	8	7	6	5	4	3	2	1	1	1	1	1

(b) DIRECTION [T]

Example: #	Site	Time: Hour 0	1	2	3	4	5	6	7	8	9	10	11
T1	CELA	90	90	90	90	90	90	90	90	90	90	90	90
T2	CELA	180	180	180	180	180	180	180	180	180	180	180	180
T3	CELA	90	90	90	90	90	90	90	90	90	90	90	90
T4	CELA	90	90	90	90	90	90	90	90	180	180	180	180
T5	CELA	90	90	90	90	90	180	180	180	180	180	180	180
T6	CELA	90	90	90	90	90	90	90	90	90	90	90	180
T7	CELA	0	15	30	45	60	75	90	105	120	135	150	165
T8	CELA	90	90	90	90	90	90	105	120	160	180	180	180

#	Site	Hour 12	13	14	15	16	17	18	19	20	21	22	23
T1	CELA	90	90	90	90	90	90	90	90	90	90	90	90
T2	CELA	180	180	180	180	180	180	180	180	180	180	180	180
T3	CELA	180	180	180	180	180	180	180	180	180	180	180	180
T4	CELA	180	180	180	180	90	90	90	90	90	90	90	90
T5	CELA	180	90	90	90	90	90	90	90	90	90	90	90
T6	CELA	180	180	180	180	180	180	180	90	90	90	90	90
T7	CELA	180	195	210	225	240	255	270	285	300	315	330	345
T8	CELA	180	180	180	160	120	105	90	90	90	90	90	90

Table 2. Example 24 element sequences which are used to create a number of possible features in the cross differencing plots in Appendix A 2.4. Site CELA is simply a label; this data was created to provide idealized instances: (a) speed as hourly averages, in miles per hour and (b) direction as hourly averages, in degrees.

method. In the first case,

$$(S1 \triangle S1) = \begin{matrix} 0.0 & \dots & 0.0 \\ \vdots & 0.0 & \vdots \\ 0.0 & \dots & 0.0 \end{matrix}$$

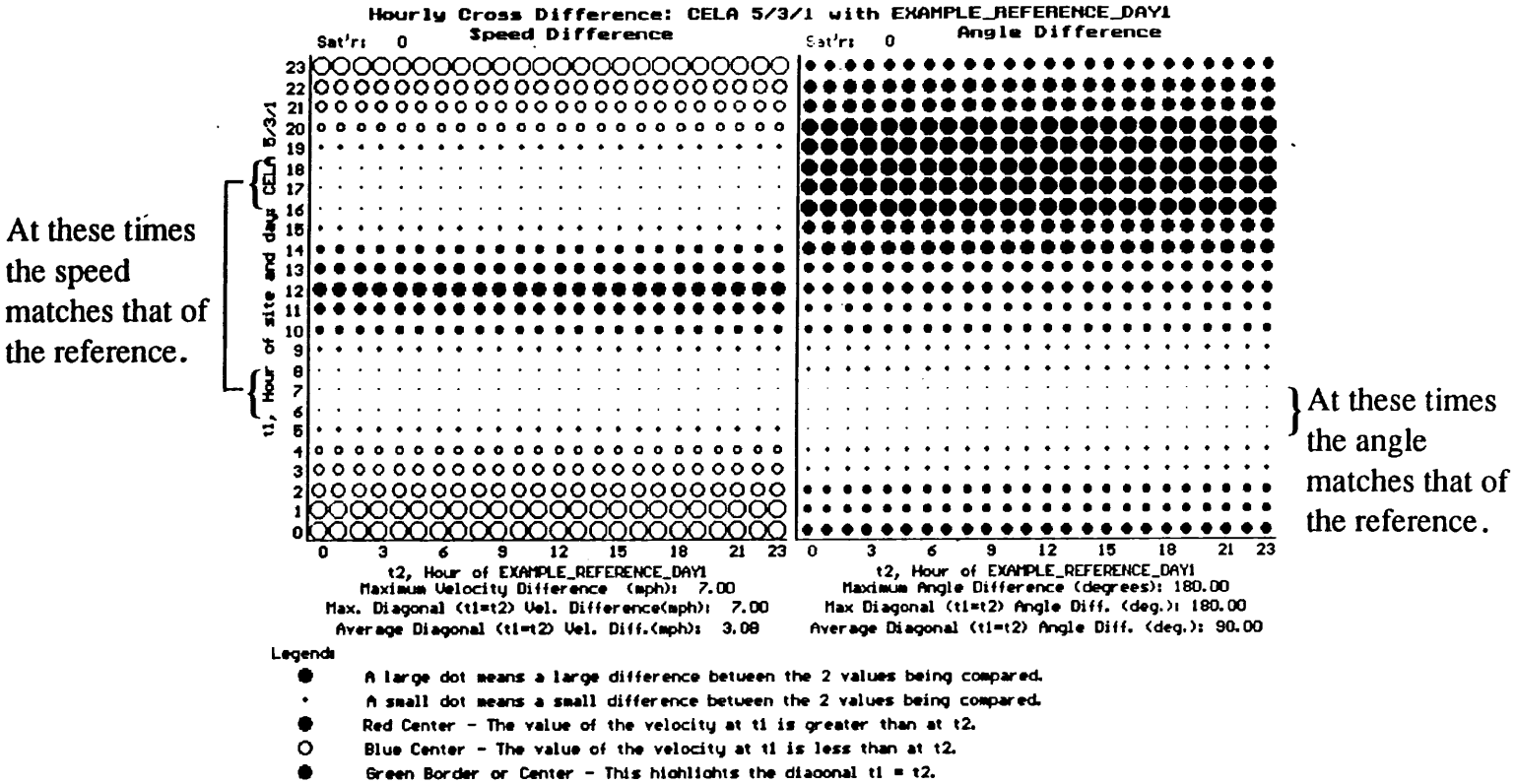
wind speeds are identical at all hours, thus all elements of $(S1 \triangle S1)$ are 0.0. In the second case,

$$(T1 \triangle T2) = \begin{matrix} 1.0 & \dots & 1.0 \\ \vdots & 1.0 & \vdots \\ 1.0 & \dots & 1.0 \end{matrix}$$

wind directions are uniformly different at all hours, thus all elements of $(T1 \triangle T2)$ are 1.0. This says the directions between the two series are at a maximum difference at all hours.

Example 2, Figure 3, is an idealized, but realistic speed profile. (See Table 2a example number S3). Similar profiles are observed in regions with light nighttime winds and strenghtening wind flow observed until midday at which time it slowly diminishes. In this case, the flow ramps from 0 mph to 12 mph between midnight and noon then ramps down to zero from noon to the following midnight. For this example day, the direction changes are a very idealized version of the wind turning through all 360 degrees starting from the north at midnight and returning to the north on the following midnight, T7 of Table 2b. The example task is to show at what times the wind speed was at or near 7 mph, and at what times was the direction near 90 degrees. The natural reference for speed filtering is then the uniform 7 mph sequence, S1, and similarly for direction the uniform 90 degrees sequence, T1 is chosen. Figure 3 shows the results of $(S3 \triangle S1)$ and $(T7 \triangle T1)$. To find the times when series $S3$ was around 7 mph and series $T7$ was around 90 degrees, look for the regions of small circles on the plots. Because of the choice of filters, these are seen as horizontal regions centered at 7 and 17 hours for speed and centered at 6 hours for direction.

Further interpretation of the speed is possible by examining the diagonal. Starting at time 0, the diagonal is made up of opened circles indicating the wind



At these times the speed matches that of the reference.

At these times the angle matches that of the reference.

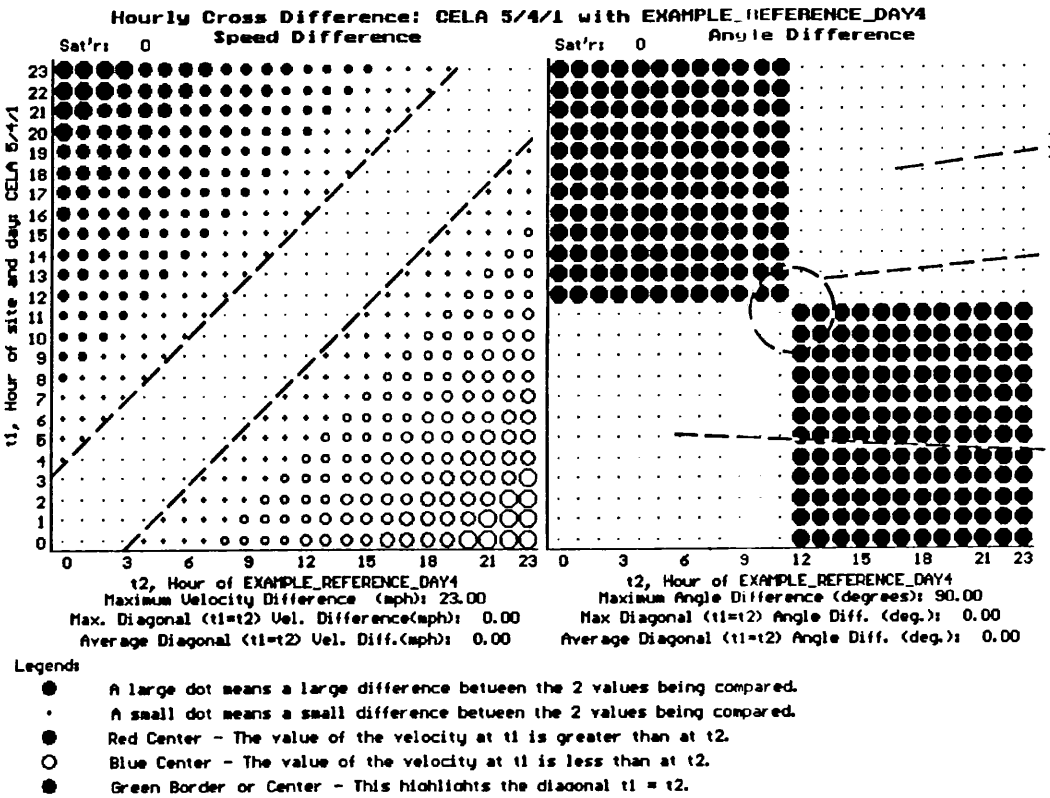
Figure 3. Cross differencing example 2, (left) speed: ramps from 0 mph at hour 0 to 12 mph at hour 12 then back down to 0 mph at hour 23 [S3 table 2] cross differenced with a uniform speed of 7 mph [S1 table2] as a reference, (right) direction: starts at 0 degrees at hour 0 and increases by 15 degrees at each hour to 345 degrees at hour 23 [T7 table 2] cross differenced with a uniform direction of 90 degrees [T1 table 2] at all 24 hours as a reference. (Note: CELA 5/3/1 is the label given to the comparison day consisting of speed S3 and direction T7. EXAMPLE_REFERENCE_DAY1 is speed S1 and direction T1.)

speed of S3 is smaller than the reference $S1$. These opened circles become smaller and pass through a minimum around a time of 7 hours. Between 7 hours and 17 hours, $S3$ is greater than $S1$, passing through a maximum difference at a time of 12 hours, as seen by the filled gray circles. As the wind speed on day $S3$ diminishes, it again passes through a minimum difference and becomes less than $S1$ in the hours from 19 to 23. A similar description of the direction diagonal places the maximum angular difference between $T7$ and $T1$ at around a time of 18 hours.

Time Shifts and Time Dilation and Compression using Diagonal Symmetry and Asymmetry

Until now, only comparisons between a dynamic series and uniform or static series have been made. The real power of this method lies in comparing two dynamic series. In particular, it will be seen that the method works very well when comparing two series which are likely to be similar because they are of similar origins. The similarity of wind flows from day to day (i.e. 24 hour blocks) at a particular site provides an ideal set of series on which to do comparisons. For instance, comparing one day's flow at a site to that of the next day's is a perfectly reasonable endeavor to undertake using this method. Because the days are expected to be similar in their bulk features, the times when they are different is highlighted by the method. The examples presented below all involve comparing two similar series which differ in carefully orchestrated ways. By creating these idealized, but realistic examples, it is hoped that the reader will be able to develop the skills to interpret the real cases presented later in this paper.

Example 3 shown in Figure 4 is a comparison which highlights similarity. It depicts the dynamic series for wind speed and direction compared to themselves, $(S2 \triangle S2)$ and $(T3 \triangle T3)$. The resulting matrices have the distinct property of creating symmetric plots about the 45 degree diagonal line $t_1 = t_2$. When



Dashed lines bound region of minimum difference.

Figure 4. Cross differencing example 3, (left) speed: ramps from 0 mph at hour 0 to 23 mph at hour 23 [S2 table 2] cross differenced with itself [S2 table 2] as a reference, (right) direction: uniform at 90 degrees from hour 0 to hour 11 then abruptly changes to 180 degrees at hour 12 then remains constant until hour 23 [T3 table 2] cross differenced with itself [T3 table 2] as a reference. (Note: CELA 5/4/1 is the label given to the comparison day consisting of speed S2 and direction T3. EXAMPLE_REFERENCE_DAY4 is speed S2 and direction T3.)

reflected across the diagonal, the circle sizes are the same, except the sign of the wind speed circles is different. This is seen as opened circles on one side and closed circles on the other. The diagonal of the wind speed plot is seen as a route of minimum difference, bounded in the plot by dashed lines.

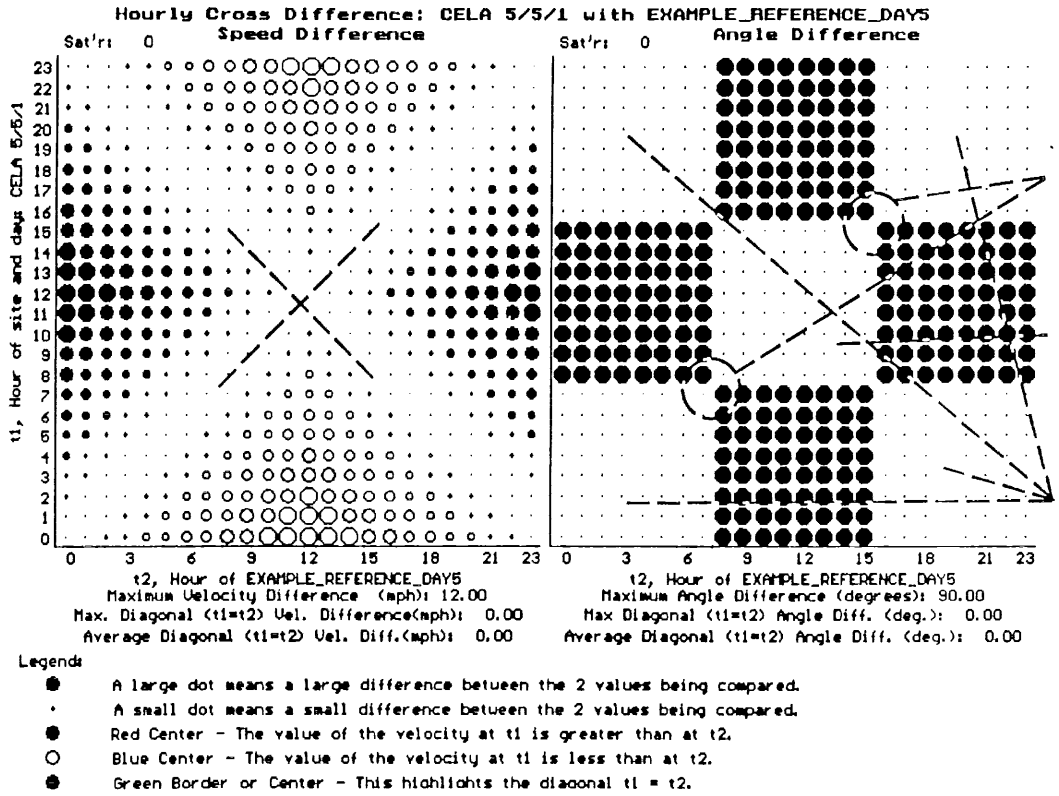
It is constructive to consider the analogy which can be drawn between the cross difference plots of two dynamic series and an elevation contour map. To start the analogy, suppose each circle of the plot corresponds to a mountain peak. The size of each circle is proportional to the height of the mountain peak. Opened cirles are cold snow covered mountains corresponding to the cross difference being negative. Closed cirles are hot arid high desert mountains corresponding to the cross difference being positive. Groups of big circles are mountain ranges, whereas smaller circles at the edge of the mountains are foothills. Lines of small circles are valleys, and patches of small circles are low elevation basins. Edges between small circles and large circles are sharp cliffs. The reader, a potential user of this method, is like a pioneer who wishes to scout out the best path from the lower left corner to the upper right corner of the map. If the diagonal line from these corners is filled with mountainous obstacles, the user knows the two series being compared are dissimilar. However, a valley or basin along the diagonal means the series compare well for that range of times. Beneath each plot is a legend which tells the pioneer the height of the tallest mountain, the height of the tallest pass and the average elevation along the diagonal route in this section of map. Finally, the pioneer may discover there is a route which is not necessarily the diagonal route but has the lowest elevation in going from near the lower left corner to near the upper right corner. This route may cause the pioneer to travel extra distance, but this is offset the reduced travel time involved in taking the easier route. If such a route is found in the cross difference plot, it means the series are similar in elements but offset in time, or similar to one another in expanded or compressed time frames. When the pioneer encounters

a new map, the first step in the process would be to map out the low elevation route across the plot. In each of the remaining example cross difference plots, the reader should endeavor to map out this route of lowest elevation. Once found, the guidelines given below explain how to interpret the route in terms of the differences which exist between the series.

With the above analogy in mind, reexamine example 3. On the speed map, the pioneer is going to go right along the valley running along the diagonal since the minimum travel distance and minimum travel elevation are one and the same. This is equivalent to saying that the two series are exactly alike. The direction map would be traversed along the diagonal also, but it is worth noting the scenery which the pioneer would see. On entering the map at location (0,0), the pioneer would be in region 1, a flat basin with cliffs in the distance both to the right, left and front. Based on taking the shortest path alone, the pioneer would stick to the diagonal. At noon, the pioneer would pass through the very narrow gorge which joins regions 1 and 2 between locations (11,11) and (12,12). This narrow gorge corresponds to a sharp directional change in both series simultaneously. The pioneer would stick to the diagonal in region 2 based only on minimizing travel distance. Future examples will have valleys and narrow gorges off of the diagonal.

Example 4 which is an idealized version of a typical day at one of the Los Angeles Basin sites, $(S3 \triangle S3)$ and $(T4 \triangle T4)$ shown in Figure 5. The diagonal symmetry and a minimum valley along the diagonal are again characteristics of both the wind speed and direction plots. Both of these plots have distinct characteristics which appear in the real plots, though the features have been accentuated to make reading them unambiguous. The distinct fingerprint of the speed pattern for this day is the large X made up of two intersecting lines of small circles (see annotations on the figure). The intersection point of these two lines is either a maximum or minimum. For this day, this point is a maximum,

Dashed "X" marks the time when the velocity is at a maximum.



The 2 circled regions mark the time of the direction transitions.

Region 2

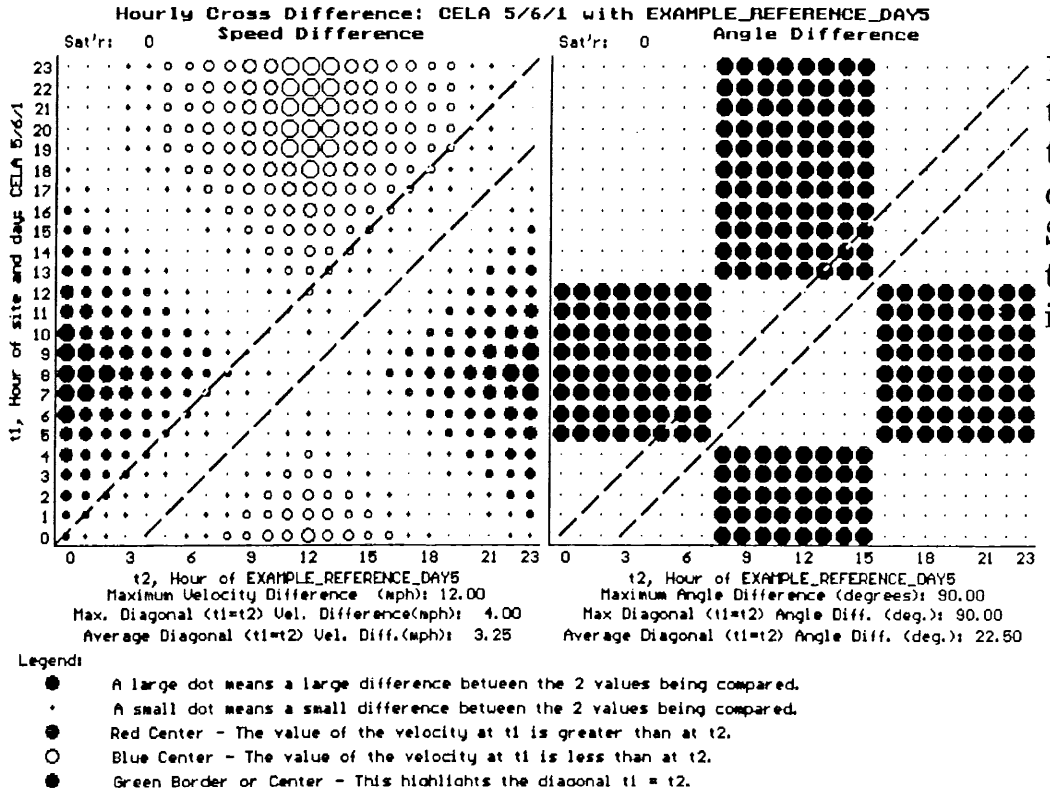
Region 1

Figure 5. Cross differencing example 4, (left) speed: ramps from 0 mph at hour 0 to 12 mph at hour 12 then back down to 0 mph at hour 23 [S3 table 2] cross differenced with itself [S3 table 2] as a reference, (right) direction: uniform at 90 degrees from hour 0 to hour 7 then abruptly changes to 180 degrees from hour 8 to hour 15 at which time the direction returns to 90 degrees from hour 16 to hour 23 [T4 table 2] cross differenced with itself [T4 table 2] as a reference. (Note: CELA 5/5/1 is the label given to the comparison day consisting of speed S3 and direction T4. EXAMPLE_REFERENCE_DAY5 is speed S3 and direction T4.)

which can be reasoned out by noting the relationship between the positive and negative circles elsewhere on the plot relative to the X. A common variation of the X pattern is seen by jumping ahead to example 6, Figure 7. The variation arises from the comparison day's speed either being smaller or exceeding the maximum speed of the reference day, but otherwise the two speed series are identical. This leads to a separated X; the point of separation is seen by the annotations on the wind speed plot of Figure 7.

The direction pattern of example 4 also has a distinct fingerprint which commonly appears. This is the checker-board pattern, characteristic of two sharp direction changes. In this case annotations mark the sharp changes occurring between 7 and 8 (am) and between 15 and 16 (or 3 and 4 p.m.). Pairs of sharp direction changes (especially one pair) are commonly observed in cross difference plots from actual applications. This observation has a simple physical explanation resulting from the persistent cyclic nature of the diurnal wind flows being considered. If the angle changes on average, it must on average change back to its original position. If a strong angle change occurs once, it is likely that some geographical feature at the site or in its surroundings is creating cause for the change. Consequently, a sharp change back to the first direction is often observed to maintain the true cyclic nature of the wind patterns.

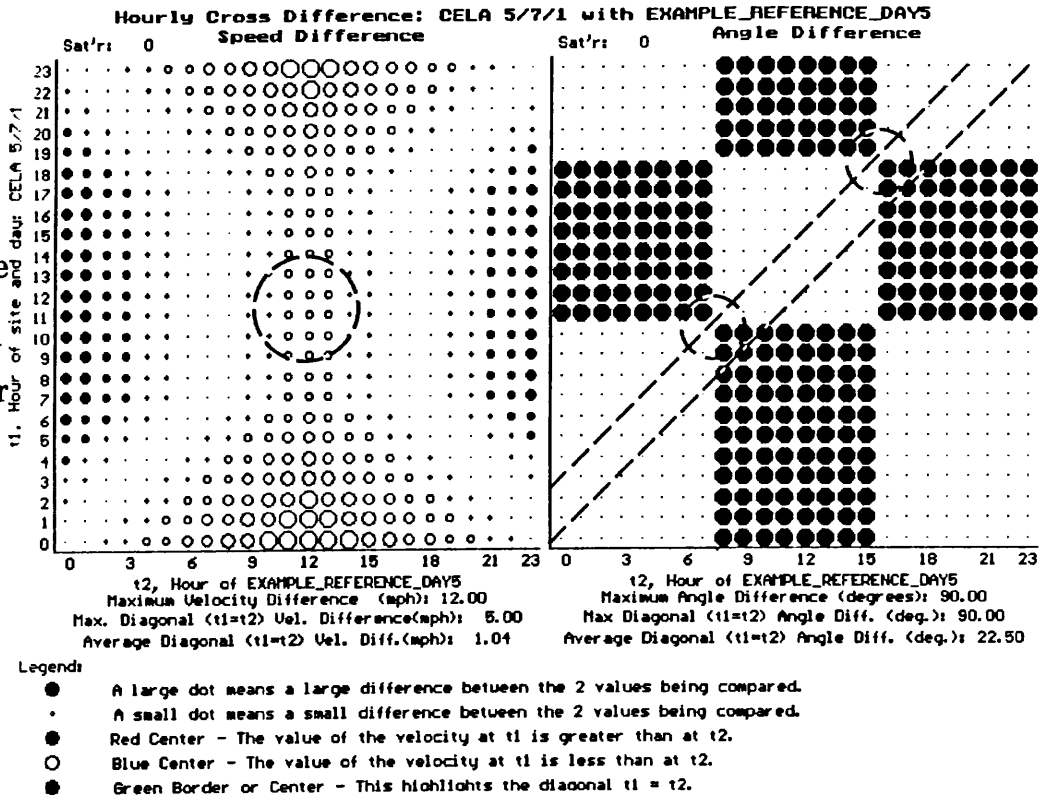
Time series which are similar to one another in element size and order, but are shifted in time, have a different 45 degree symmetry line $t_1 = t_2 + n$ where n is plus (the comparison series lags the reference series in time) or minus (the comparion series leads the reference). To demonstrate a series leading the reference, the speed time series $S3$ is shifted four hours ahead in time to make $S7$ and the direction time series $T4$ is shifted three hours ahead to make $S5$. In example 5, the two new series are cross differenced with the series from which they were derived as $(S7 \triangle S3)$ and $(T5 \triangle T4)$ and are shown in Figure 6. The characteristic X marking the wind speed maximum has moved down four hours,



Dashed lines are the diagonal and the minimum difference. Speed is a 4 hour time shift, direction is a 3 hour shift.

Figure 6. Cross differencing example 5, (left) speed: ramps from 4 mph to 12 mph from hour 0 to hour 8 then back down to 1 mph at hour 19 [S7 table 2] cross differenced with a similar ramp starting at 0 mph and reaching a maximum of 12 mph at hour 12 [S3 table 2] as a reference, (right) direction: 2 transitions occur 3 hours earlier [T5 table 2] than the reference [T4 table 2]. Both the speed and direction examples demonstrate how a time shift looks. (Note: CELA 5/6/1 is the label given to the comparison day consisting of speed S7 and direction T5. EXAMPLE_REFERENCE_DAY5 is speed S3 and direction T4.)

The circled region shows the times when the comparision day has speeds lower than the reference day.



The 2 circled regions mark the time of the delayed direction transitions.

Figure 7. Cross differencing example 6, (left) speed: ramps from 0 mph at hour 0 to 7mph at hour 7 remains at 7 mph until hour 17 and then ramps back down to 1 mph by hour 23 [S6 table 2] cross differenced with with speed ramping from 0 to 12 mph from hour 0 to hour 12 then back to 1 mph by hour 23 [S3 table 2] as a reference, (right) direction: 2 transitions occur 3 hours later [T6 table 2] than the reference [T4 table 2]. This speed example is commonly observed when an average day is compared to a single day. (Note: CELA 5/7/1 is the label given to the comparison day consisting of speed S6 and direction T6. EXAMPLE_REFERENCE DAY5 is speed S3 and direction T4.)

placing it at 8 a.m., while the checker board pattern of the wind direction is shifted three hours the same way. The wind direction difference plot on the right side of Figure 5, example 6, is a case of the comparison series lagging the reference by three hours. A simple rule to follow is this: when a minimum valley in a plot is observed to the right of $t_1 = t_2$, the events of the comparison series happened earlier in the day than in the reference. A similar shift to the left in a minimum valley means the events happened later.

The comparison of a time series which is a time dilation or compression of the reference series shows a symmetry line $t_1 = Nt_2$ where N is a positive fraction less than unity for time dilation and N is greater than unity for a time compression. In example 7, the wind speed $S3$ was exaggeratedly compressed into half its normal period ($N = 2$) to give $S5$ then ($S5 \triangle S3$) was done and shown on left half of Figure 8. Notice the X has been compressed, and the line of symmetry has been changed to an angle of 22.5 degrees. A time dilation would give angles greater than 45 degrees.

Obtaining Time Constants of Change from the Plots

Cross differencing is useful in identifying characteristic times associated with wind speeds and wind directions. The two times which are relevant to the investigations made later in this paper are as follows: (1) the length of time a series remains at a (nearly) constant value t_d and (2) the length of time it takes a series to change from one constant value to another t_t. When analyzing a time series pair of wind speed and direction, two questions can be quickly answered by examining the cross difference plot: (1) what is the duration of time with like wind speeds and/or like directions, (2) what length of time is required for a particular directional change? This information is contained in all of the examples Figures 2 - 8 and will be discussed in terms of the vertical width rule.

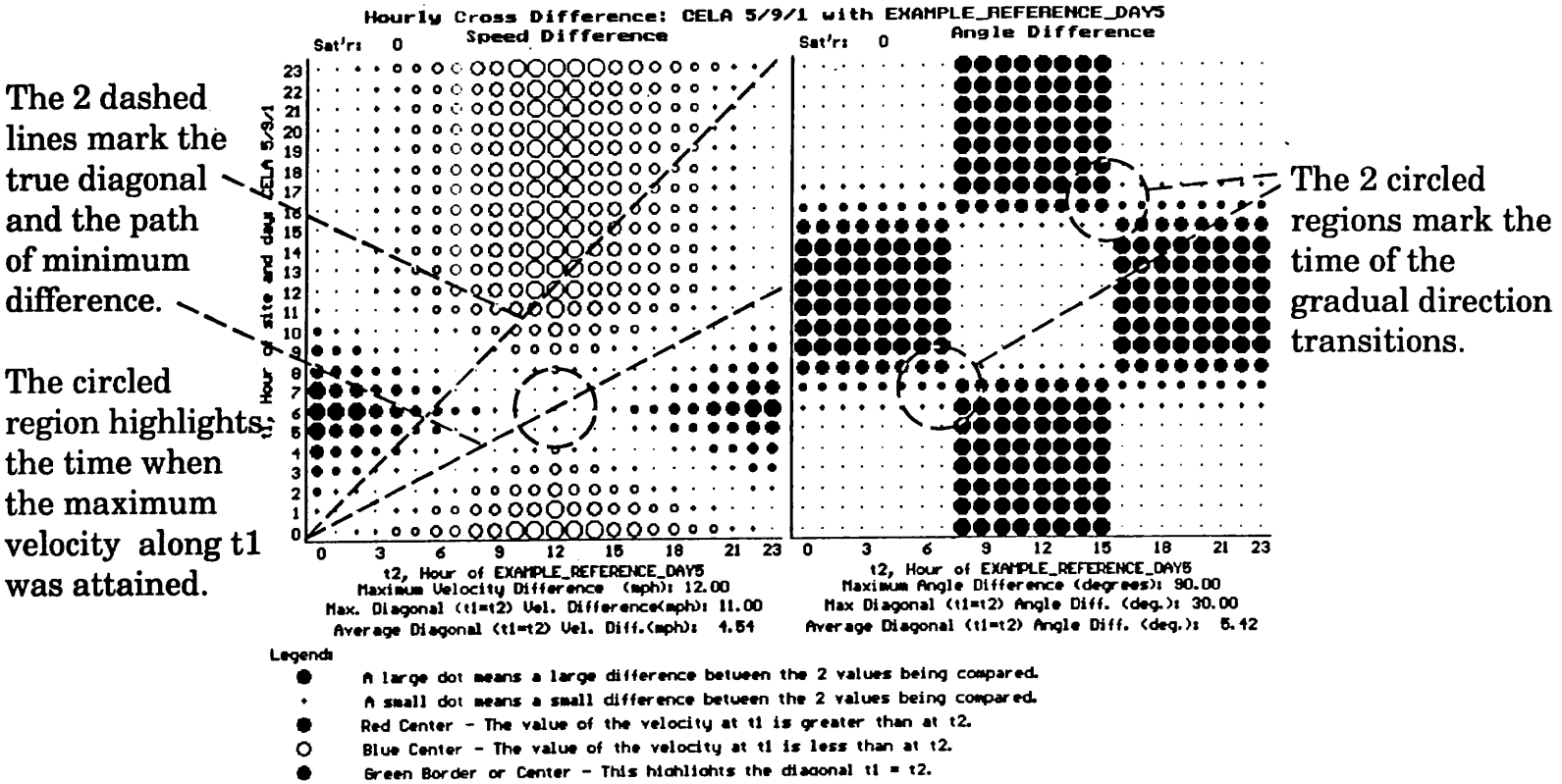


Figure 8. Cross differencing example 7, (left) speed: ramps from 0 mph to 12 mph from hour 0 to hour 6 then ramps back down to 0 mph by hour 12 and stays there until hour 23 [S5 table 2] cross differenced with the same ramp occuring over twice the time [S3 table 2] as a reference, (right) direction: 2 gradual transitions occuring over a period of 3 hours [T8 table 2] cross differenced with the identical transitions occuring over a single hour [T4 table 2] as a reference. (Note: CELA 5/9/1 is the label given to the comparison day consisting of speed S5 and direction T8. EXAMPLE_REFERENCE_DAY5 is speed S3 and direction T4.)

The Vertical Width Rule

The rule is this: to find the duration time constant, t_d, for an element of the comparison series, count the number of like vertical circles which are in a contiguous strand with the element in question. The duration time constant is the number of circles counted (in hours for the series being looked at herein).

A corollary is as follows: to find the t_d for an element of the reference series, count the number of like horizontal circles which are contiguous with the element in question.

The rule also applies to finding the transient time constant, t_t, for an element of the comparison (or reference) series, except that instead of counting like circles around an element, contiguous strands of monotonically increasing or decreasing sized elements are counted. Such strands of changing circles are looked for between strands of uniform circles. Categorically, every element falls into one or the other type of strand.

This rule will be used to answer the above two questions. First, to answer the question of how long the speed remains relatively constant, consult the speed plot of Figure 5 for an example. From any element along the diagonal $t1 = t2$, count the number of the smallest circles up and down from it; in this case there are 3 or 5 depending on which diagonal element. Stop when the circle changes size. At a glance, one knows that the wind speed at any time during the day is within one circle size (one sixth) of the maximum speed difference of 12 mph, or 2 mph of its neighbor's wind speed. Depending on which diagonal element was chosen, this is between 3 and 5 hours. An equivalent procedure is used to determine the duration of constant angle difference. (See the annotations made on Figure 5.) The annotations designate two regions in the plots. Region 1 covers a time period of 16 continuous hours for both the comparison and reference series. It appears as four pieces, but by using the plot itself one can infer that the four regions must all be the same direction. Region 2 covers a time period of eight

hours.

To determine the period of time over which a particular directional change occurs, the second issue in question, examples will be used. Locate the directional change in Figure 5 between 7 and 8 a.m.. Do this by going to the diagonal at hour 8, now step back to hour 7 vertically. Notice that the circle size changes full scale. Remember each change in circle size represents a 16.7 percent increment on the maximum observed difference scale. For the above example, the change is from smallest to largest in one step. This means 100 percent of the maximum observed direction change for this comparison is observed across the hours from 7 to 8. Noted below the plot is Max's value of 90 degrees. Then, for this case, $t_t \leq 1$ hour and a 90 degree change occurs over this time. In the case of a gradual angle change, the circles would grow larger gradually when moving downward in the vertical direction. Under such conditions, the sharp checkerboard pattern of the entire plot has usually been replaced by one with less pronounced edges. In these cases, $t_t \geq 1$ hour. An example of this is shown in the angle difference of Figure 8, $(T8 \triangle T4)$ where the directional change occurring between 5 and 8 a.m. gives $t_t = 4$ hours. To understand this, go to $(t_1 = 0, t_2 = 4)$ and move vertically along the small circles until they begin to get larger. Count the number of time steps where the circles increase. Stop when the circles become constant at the new large size. This is four hour steps.

Developing a Seasonal Average as a Reference

A reference for speed and direction cross differencing was created for each of the 31 wind reporting stations located in the Los Angeles Basin (Appendix A1.9). The reference was created by vector averaging the wind vectors for every day from May through September of 1988, at each hour and at each site. These results are reported in Section 1.9. The composite vector average was then converted back into the 31 speed and direction time series, 24 hours long, one pair of series for each site. The 1988 smog season wind data was chosen as the reference because it best satisfied a number of criteria which were desirable. Recapitulating from Appendix A1.9, the desired criteria are (1) the year was not exceptionally low or high, statistically, in observed levels of ozone, and (2) the year was as close in time to the year 1987 as possible (unfortunately, 1987 ozone levels were too low and therefore that year was rejected).

The persistent diurnal patterns at each site justify using a composite average series as a reference. Appendix A1.9 provides a detailed examination of the statistics of the day to day variation in the wind patterns at each site.

Two topological properties arise when using a seasonal average as the reference in a cross difference. The first relationship is that hour 0 and hour 23 are connected to one another in time. Conceptually, this is the equivalent of taking the plot and connecting columns 0 and 23 so that they are side by side to form a cylinder. The second relationship arises when both the reference and the comparison series are composed of averages. As described above, the sheet is first folded into a cylinder. Then the cylinder is connected by bending one end to meet the other, matching like times, into a torus. Practically, this means regions with similar sized circles are the same across the 23 to 0 hour boundary. This fact may be useful in interpreting the plots done for the section Individual Days Compared to Seasonal Averages below. When comparing a particular day to an average day, the cylinder topography is the rule to use. This applies to the

plots derived from one or two average days.

Results

Average Summertime Dynamics

The cross difference is computed for 31 sites in the Los Angeles Basin from the composite average 1988 smog season day at each site and compiled into Figures 12 - 27 in Appendix A2.4. From these plots, the following data is obtained at each site: (1) the average number of transition periods, (2) the average time the transition periods occur, (3) the average duration of transition periods, (4) the average angle change on transition, (5) the maximum angle difference between any two hours on the average day, and (6) the time of the extremum in the wind speed. Numbers (1), (2), (3), and (5) are reported in Table 3. (It is assumed that the reader having read to this point has enough skill to interpret the plots in Appendix A2.4.) In interpreting the plots, recall regions across the boundary from 23 to 0 hours of both the reference and comparison edges are topologically connected because both are average days.

The wind speed plots of an average day cross differenced with itself are straightforward to read and contain information that could be obtained by a simple time series plot, so only a brief mention of them will be made. On each day, a distinct extremum in the wind speed is seen via the characteristic X, which occurs atop of the time when the extremum is observed. In this respect, all of the sites are identical except that Palm Springs is at a minimum wind speed at midday, while the other 30 sites are approaching a maximum.

The angle difference plots for each site generally have more features than the wind speed plots and are able to provide a considerable amount of information about the average behavior of the summertime winds at that location in a way which is more difficult to convey in simple time series plots. Again, consult Figures 12 - 27 in Appendix A2.4. The cross differencing plots provide an excellent method for determining the length of time a site remains in both the macroscale and mesoscale systems. The method also provides a way to examine the struc-

Station Name	Angle Changes	Hour of Changes	Max. Degree Change	Transition Period in Hrs.
ANAH	1		72.8	3
AZUS	2	4-5,6-7	113.6	2
BANN	0		40.6	0
BURK	2	9-10,16-17	44.5	0
CELA	2	2-4,5-6	178.5	3
COST	2	2-3,6-7	96.2	4
CRES	3	5-6,10-11 20-21	76.3	5
FONT	2	2-3,6-7	179.5	4
GLEN	2	23-0,6-7	179.8	7
HAWT	2	3-4,6-7	87.5	4
HEME	1	6-7	107.6	1
LAHB	2	3-4,7-8	59.9	4
LANC	0		32	0
LGBH	0		156.8	0
LSAL	4	1-2,5-6 7-8,17-18	70.1	2
LYNN	2	4-5,10-11	57.0	6
NEWL	2	4-5,6-7	138.0	2
NORC	2	4-5,5-6	109.7	1
PASA	2	3-4,5-6	177.5	2
PERI	5	1-2,2-3, 4-5 10-11,20-21	167.5	4
PICO	2	2-3,7-8	53.7	4
PLSP	2	8-10,13-14	179	4
POMA	1	7-8	41.9	0
RDLD	2	6-7,22-23	179.9	7
RESE	3	6-7,17-18 20-21	177.3	2
RIVR	2	5-6,10-11	179.9	3
SNBO	2	5-6,7-8	71.0	2
TORO	2	7-10,20-22	142.1	7
UPLA	1	7-8	126.9	1
WHIT	2	3-4,6-7	179.2	3
WSLA	2	1-2,5-6	178.3	4

Table 3. Features of the Average Angle Difference Plots at 31 Sites

ture of the transition periods both on average and on specific days. The cross difference plots alone do not provide unique identification of the specific flow system category to which a site belongs. Thus, this chapter is complemented by the subject matter of Appendix A1.9 which discusses the specific types of flows exhibited by each of the 31 sites and the standard deviation of the wind speed and direction at each site and hour taken over the 151 1988 smog season days. Information presented in Section 2.3 may be helpful in understanding some aspects of the following results.

Qualitatively, each of the sites exhibit the same features. For instance, of the 31 sites 28 show at least two distinct flow regimes on a single day. Additionally, 3 of the 28 show more than two distinct flow regimes. For these 28 sites, 75 to 90 percent of the smog season day is spent in a single flow pattern. This period of a single pattern represents the macroscale flow regime. During the remaining 10 to 25 percent of the time, the site is in a transitional or mesoscale flow which is identified in the plots by a flow direction different from the macroscale. The ratio of the time spent in the macroscale flow versus the time spent in the mesoscale flow is a measure of the relative strengths of the two flows. Though there is no direct evidence to substantiate this hypothesis, the ratio is potentially an indicator of the probability of observing the mesoscale flow on any particular day.

The similarity between the different sites and the common features observed in many of the plots can be used to refine the conceptual model of the transition period. First, notice 20 sites on average have a pair of large angle changes separating a one to six hour period with a uniform flow direction. Second, the timing and duration of each of the transition periods is different for each site, with 19 out of 31 sites having a pair of direction changes between 0 and 12 a.m. Of the four sites with only one transition, the transition occurs between 6 and 8 a.m.

Though 22 out of 31 of the sites examined have an even number of pairs of wind direction changes, it is possible to observe an odd number of changes or no distinct angle changes at all. The sites which exhibited no change had one or both of the following characteristics: (1) it was surrounded by relatively level terrain or, (2) it was surrounded by highly channeling terrain. These environmental features create conditions which minimize mesoscale forcing, making it difficult to observe such effects even under conditions of very mild synoptic scale pressure forcing. Relatively flat terrain does not generate slope wind, and the absence of nearby mountains excludes mountain valley drainage winds. The flat terrain does not confine the wind to any one direction, making gradual directional changes possible. The presence of channeling terrain provides a guide for the wind and under normal conditions restricts the wind flow to the two directions parallel to the channel. Mild, but persistent synoptic scale pressure gradients which exist between the two air basins joined by a channel creates a steady uni-directional wind flow capable of suppressing mesoscale effects. Of the four stations exhibiting gradual or no change, three (ANAH, LANC, and LGBH) were surrounded by flat terrain. The fourth station was Banning (BANN) which is located in the valley channel (San Gorgonio Pass) linking the LA basin to Palm Springs. The six sites which exhibited an odd number of direction changes are accounted for by the explanation that a sharp change was paired with a gradual change which returned the wind to the original direction.

The definition of the term *transition period* is refined for sites in the Los Angeles Basin in light of the above findings. The transition period is actually the time between two generally sharp directional changes which surround the period when mesoscale effects dominate the local wind field. Because the winds during this period are light, often below 1.0 ms^{-1}, their importance to transport has usually been overlooked. However, when one realizes that these winds are light, but directionally not variable, they could be responsible for relocating emissions

on the order of 3 - 15 kilometers from the original release point during the 1 - 5 hours of a typical transition period. Finally, the dispersion which has usually been noted during these transition periods is probably attributable to the two sharp directional changes which bracket both ends, and not necessarily to variable winds occuring all during the transition period itself.

Individual Days Compared to Seasonal Averages: SCAQS 1987 SF_6 Tracer Days

To complement the conclusions drawn in previous chapters, the wind flow dynamics of the two summertime 1987 sulfur hexafluoride tracer studies done as part of the Southern California Air Quality Study are studied using the cross differencing technique. The dynamics of three sites were examined, including one near the release site (Downtown Los Angeles, CELA), a site in the northwestern basin where the tracer was observed (Burbank,BURK), and a site in the eastern desert (Upland, UPLA) where ozone was observed but tracer. These sites represent a source region (CELA) and two potential receptor sites (BURK,UPLA). The goal of this section is to provide additional evidence that the morning dynamics of the release site were largely responsible for the location of the late morning and afternoon observations of tracer in the basin. Six pairs of speed-direction cross differences were computed and put into Figures 9 - 11, the plots from the same site from July 15 and September 3 both being put on the same page. The cross differences were computed comparing the specific day to the average smog season day of 1988.

It is immediately evident from the plots for CELA, Figure 9, that the wind speed near the release site was below the seasonal average in the afternoon on both of the tracer days. At CELA, the maximum wind speed difference on July 15 was 5.9 mph below average at 4 p.m. The maximum wind speed on September 3 is 4 mph below average at 4 p.m. The standard deviation, σ, in the speed for this time at CELA based on the seasonal average for 1988 is 3.2 mph. Both days have values greater than the standard deviation. The angle difference on July 15 indicates the transition period began later and ended earlier than average. On September 3, there was no transition period at all. On both days, this indicates that there was little or no time between the surface flow pattern from the previous day and the onset of the flow on the tracer day. On the same days, the other

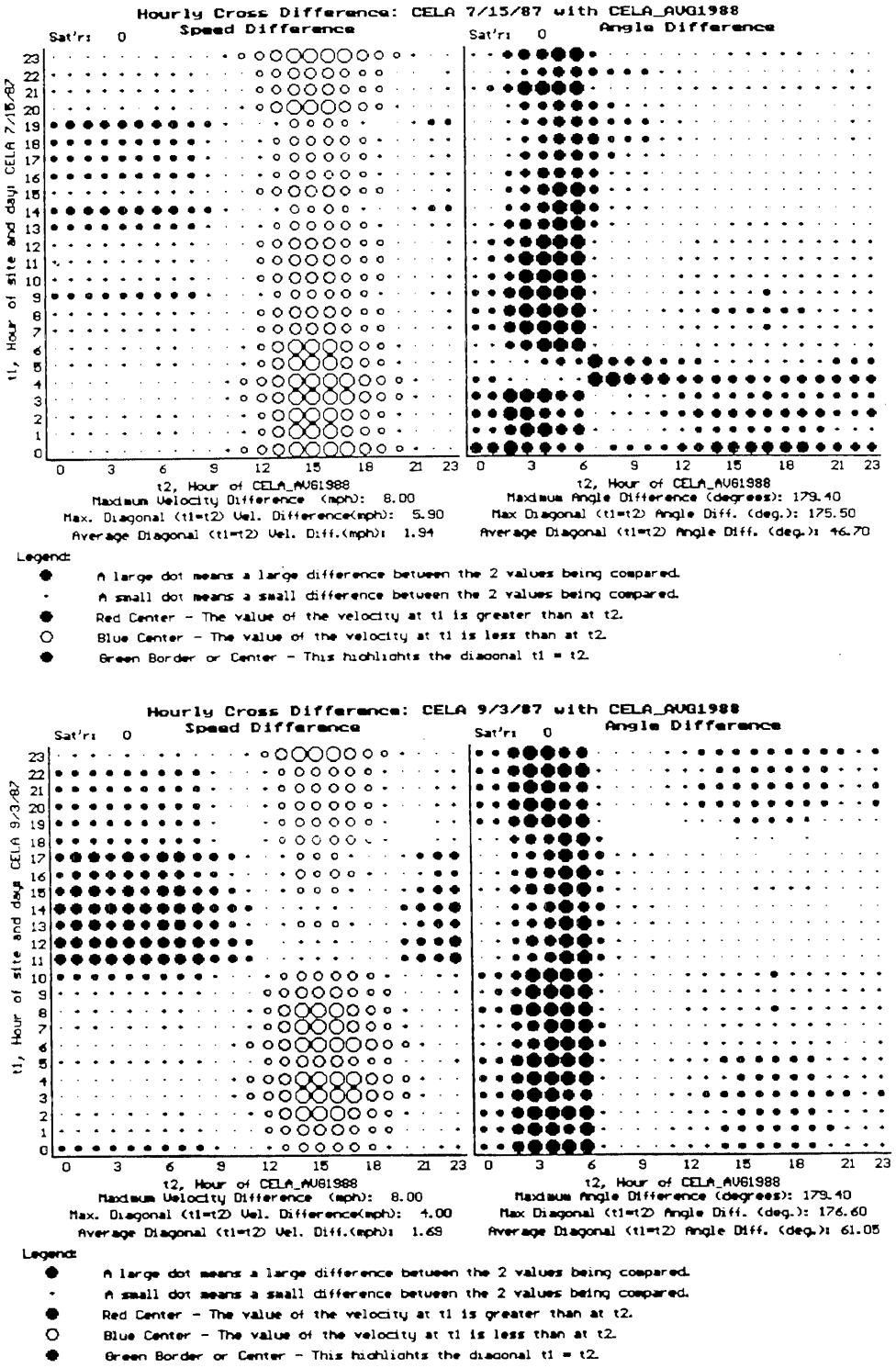


Figure 9. Speed and angle cross difference at Downtown LA (CELA) (top) July 15, 1987,(bottom) September 3, 1987. In both cases the reference for comparison is the average smog season day 1988 at CELA. Note the similarity between both days.

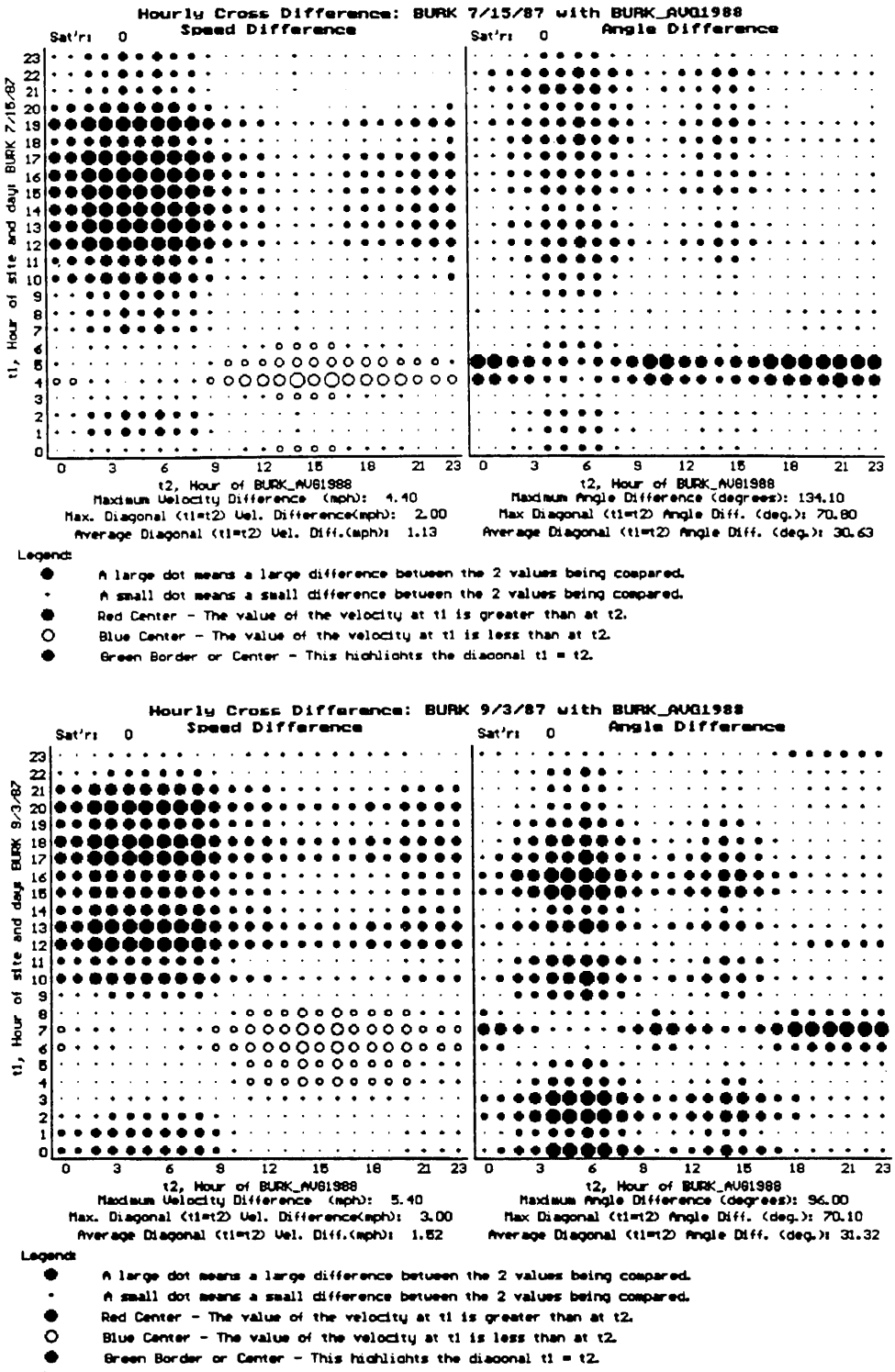


Figure 10. Speed and angle cross difference at Burbank (BURK) (top) July 15, 1987, (bottom) September 3, 1987. In both cases the reference for comparison is the average smog season day 1988 at BURK. Note the similarity between both days.

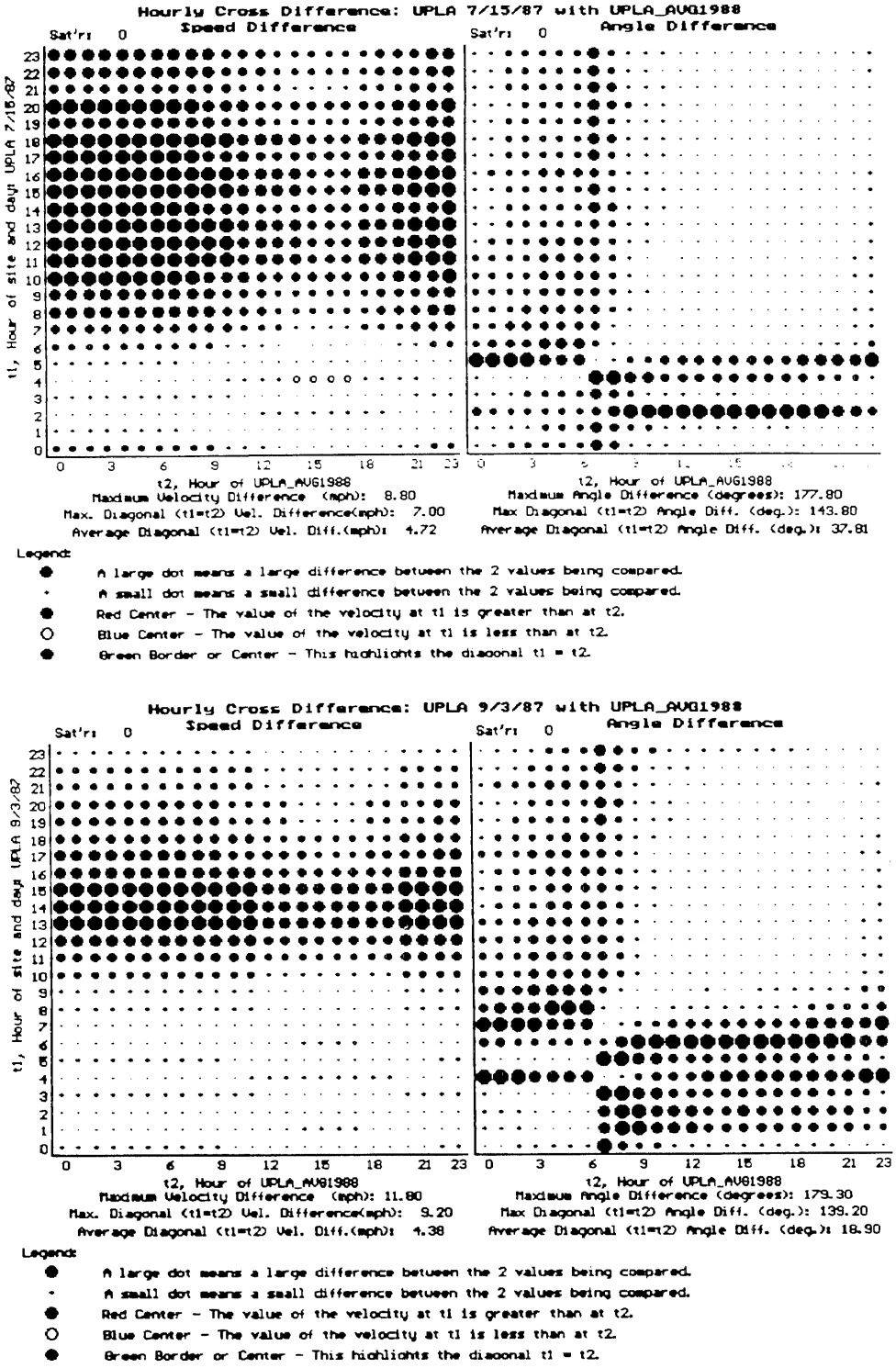


Figure 11. Speed and angle cross difference at Upland (UPLA) (top) July 15, 1987, (bottom) September 3, 1987. In both cases the reference for comparison is the average smog season day 1988 at UPLA. Note the similarity between both days.

two sites, Burbank (BURK, Figure 10) and Upland (UPLA, Figure 11), have wind speeds which are higher than the average. For BURK, maximum wind speed difference on July 15 was 2 mph ($\sigma = 3.2$) above average at 5 p.m. On September 3, the wind speed is 3 mph above average ($\sigma = 2.6$) at 6 p.m. For UPLA, maximum wind speed difference on July 15 was 7 mph ($\sigma = 1.2$) above average at 6 p.m. On September 3, the wind speed is 9.2 mph above average ($\sigma = 1.3$) at 3 p.m. The morning transition at Burbank is weak on July 15, and occurs two hours later than average on September 3. On the afternoon of both days, the angle plot does not exhibit the typical strong direction change. The directional change is seen in Figure 10 as the black region between hours 12 and 15 of the reference day, and extending roughly from hour 10 to hour 20 of the comparison days. The angle difference plot at Upland has a simple interpretation: on July 15, the morning transition period comes to an end two hours early; on September 3 the angle profile is a close match to the average day (with the exception of an abrupt 139 degree angle difference recorded at hour 4).

Taken collectively, the results from these three stations describe the following circumstances. The two major outlets from the basin are experiencing higher than normal outflow rates. Downtown Los Angeles, the branch point between the two outlets, is experiencing a strongly suppressed wind flow speed because of the sharper than average divergence which resides in the air space above this location. The consequence of this sharper than average divergence region is the creation of a sharp division between which emissions from the downtown area are routed out of each outlet. For the two tracer experiments, all of the tracer was routed via the route which passes by Burbank to the northeast of the release point. Less than one percent of the tracer was observed to head towards Upland or any other location in the eastern basin.

Conclusions

The Cross Differencing Method

1. The cross differencing method is easy to implement, requires commonly available computing hardware, and is fast to use once reading the plots is mastered.
2. The method has been successfully demonstrated using wind fields from both seasonal average days and individual days.
3. A natural application of the method, but one not examined in this paper, is the comparison of two wind fields for the same site. This approach is likely to be very useful in visually identifying times when numerical models succeed and fail at a particular location. Furthermore, the comparison of the results of different numerical models can be accomplished by this method.

Transition Periods in the Los Angeles Basin

1. The transition periods primarily have an effect on the morning rush hour. On average, strong direction change occurs within the morning rush hour at 27 of the 31 sites, while only 2 of the 31 sites (Burbank, Reseda) have direction changes during the evening rush hour.
2. On average, the transitions are characterized by two sharp directional changes; one occurring at the end of the previous day's wind flow and the other at the start of the new day's flow. The wind direction often changes on the order of 100 or more degrees on transition.
3. Transport during a one to five hour transition period may relocate emissions from between 3 - 15 kilometers from the source because of the directional nature of the flows during these periods.

SCAQS SF_6 Days

1. The two days examined show similar dynamics both near the downtown Los Angeles release site and at two sites situated near basin air flow exits. Both days had strongly suppressed wind speeds near the downtown area and higher than average speeds at the exits.

Accountability Modeling needs to accurately account for the timing of the transitions, and their flow direction and speed. Mesoscale effects are observed primarily after dark, and most commonly in the early morning hours.

A2.4 Appendix

Cross Difference Plots for Average Smog Season Days at 31 Sites in Los Angeles

The following 16 pages contain cross difference plots for the 31 sites listed in Table 1. Each plot is generated by performing the cross differencing method on the seasonal average for that site with itself. Certain observations which can be made from the plots are summarized in Table 2 and discussed within the text.

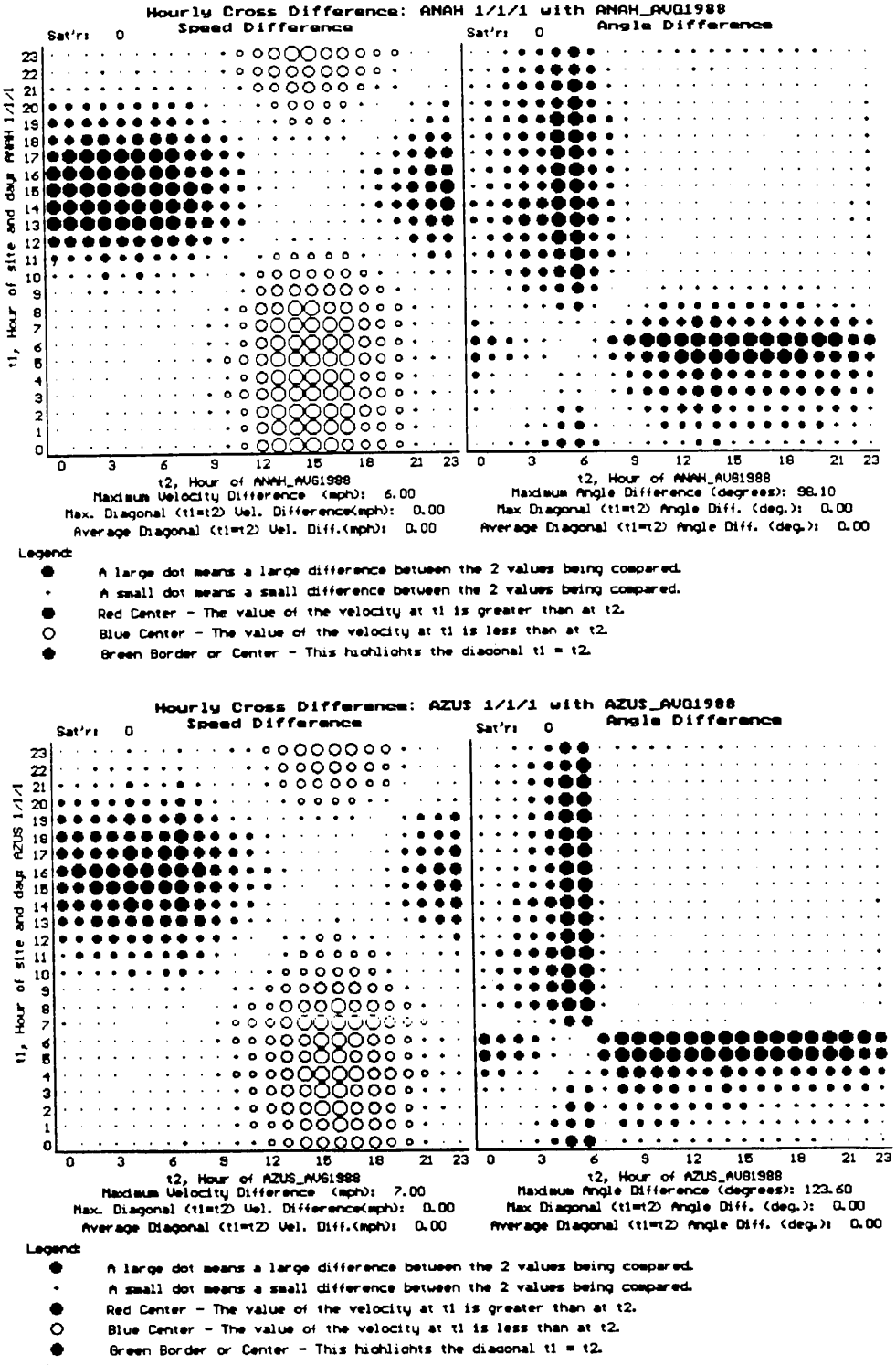


Figure 12. Speed and angle cross difference (top) Anaheim (ANAH) average smog season day 1988 compared to itself, (bottom) Azusa (AZUS) average smog season day 1988 compared to itself.

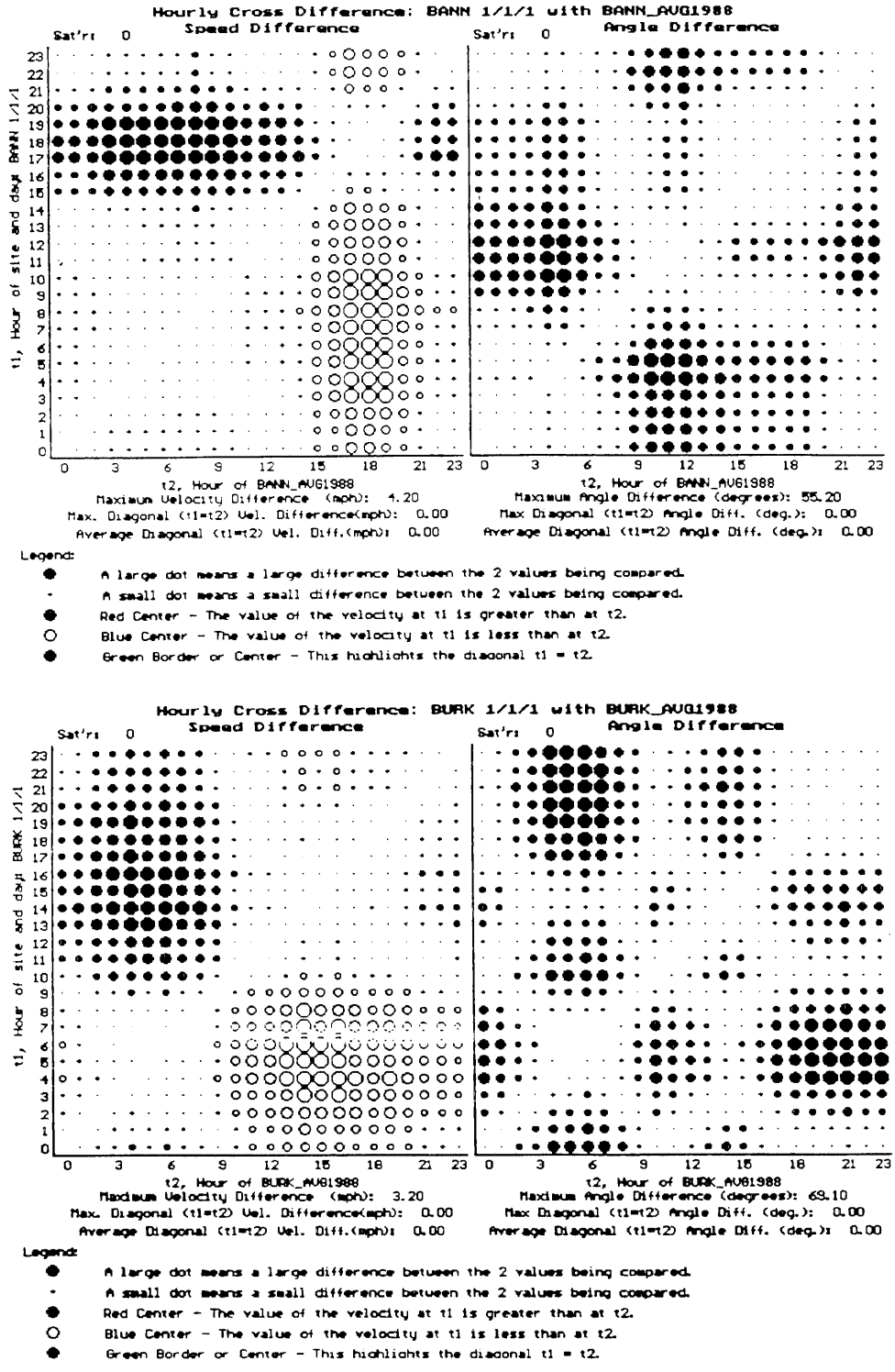


Figure 13. Speed and angle cross difference (top) Banning (BANN) average smog season day 1988 compared to itself, (bottom) Burbank (BURK) average smog season day 1988 compared to itself.

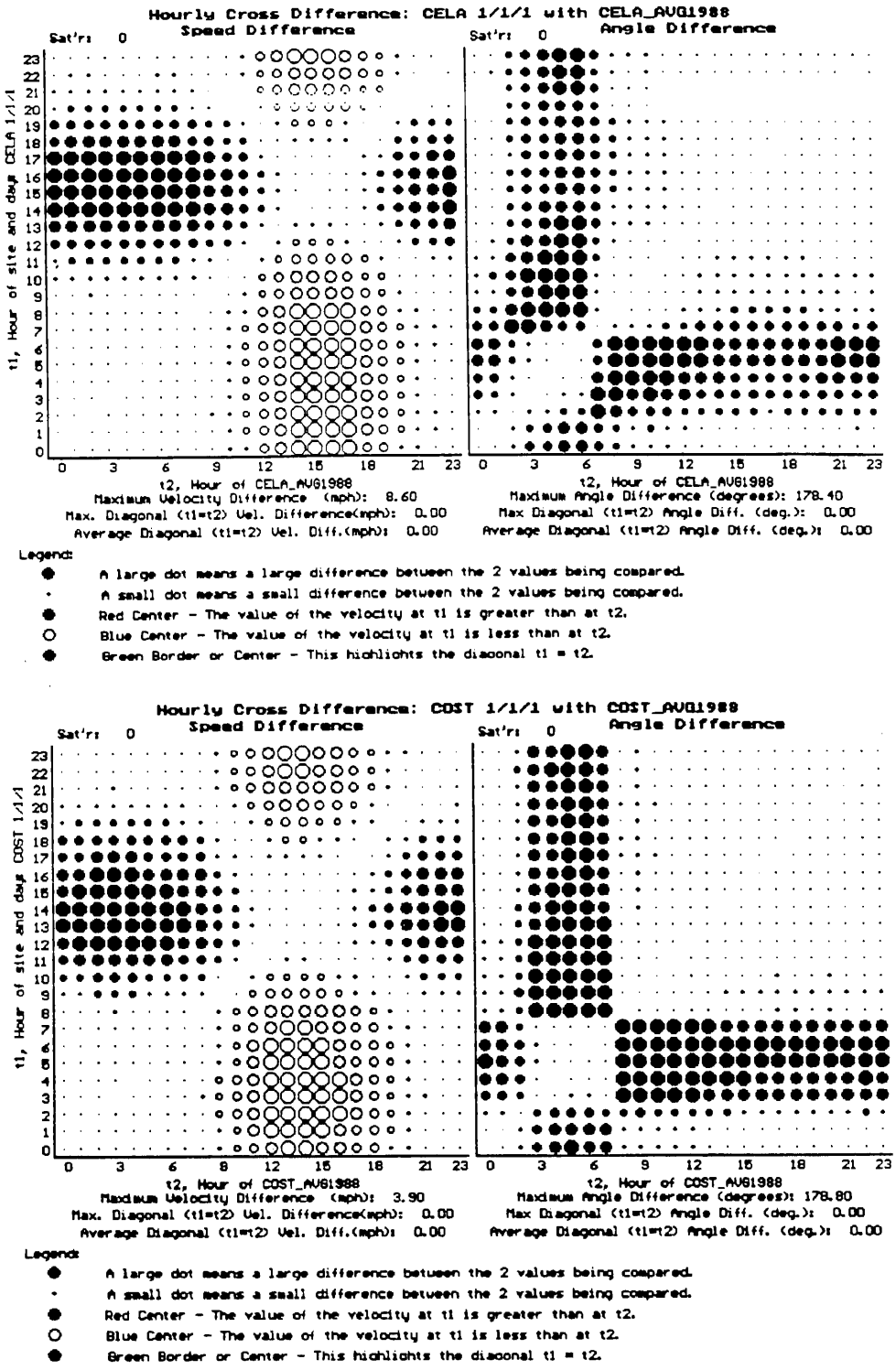


Figure 14. Speed and angle cross difference (top) Central Los Angeles (CELA) average smog season day 1988 compared to itself, (bottom) Costa Mesa (COST) average smog season day 1988 compared to itself.

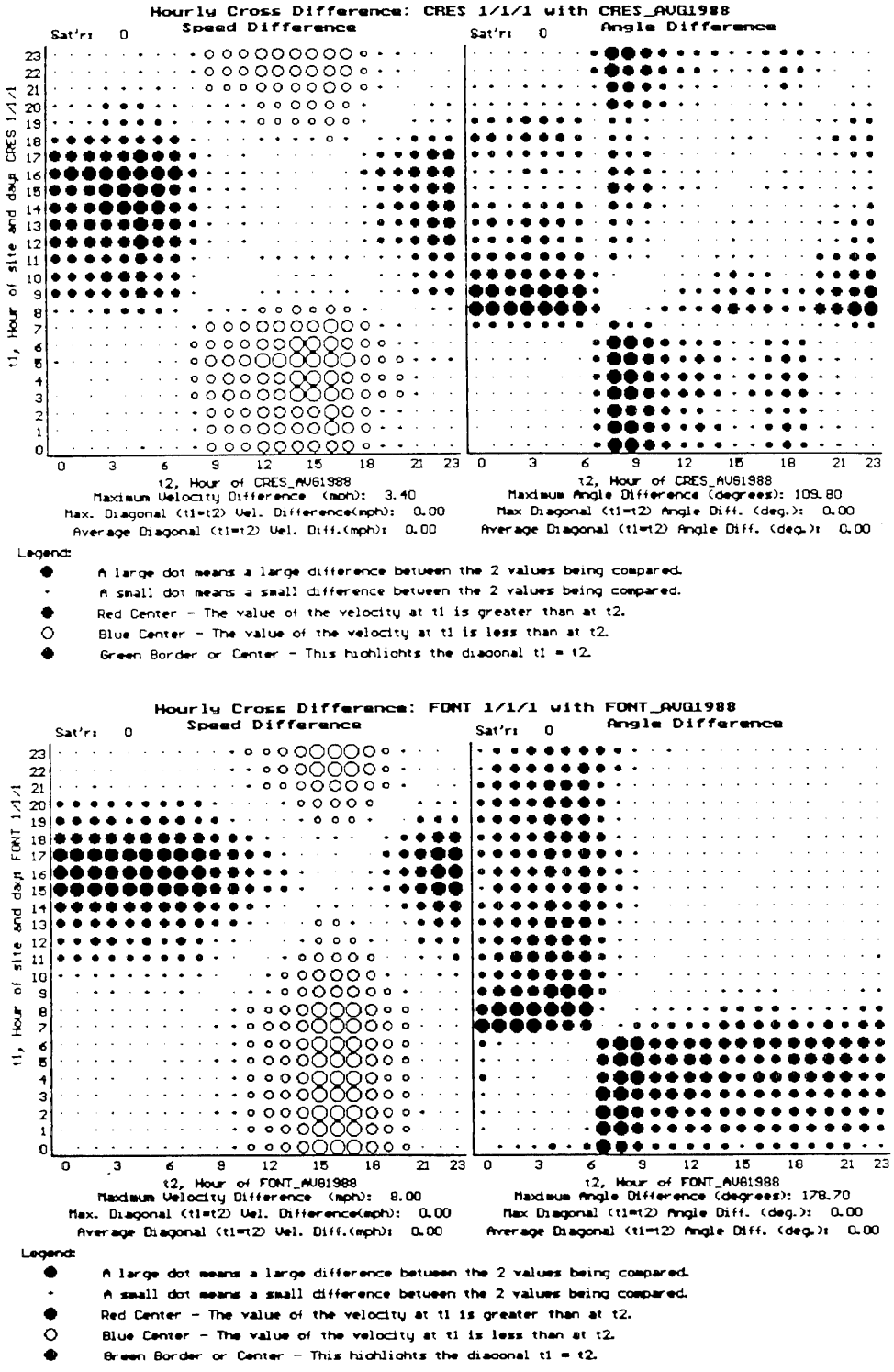


Figure 15. Speed and angle cross difference (top) Crestline (CRES) average smog season day 1988 compared to itself, (bottom) Fontana (FONT) average smog season day 1988 compared to itself.

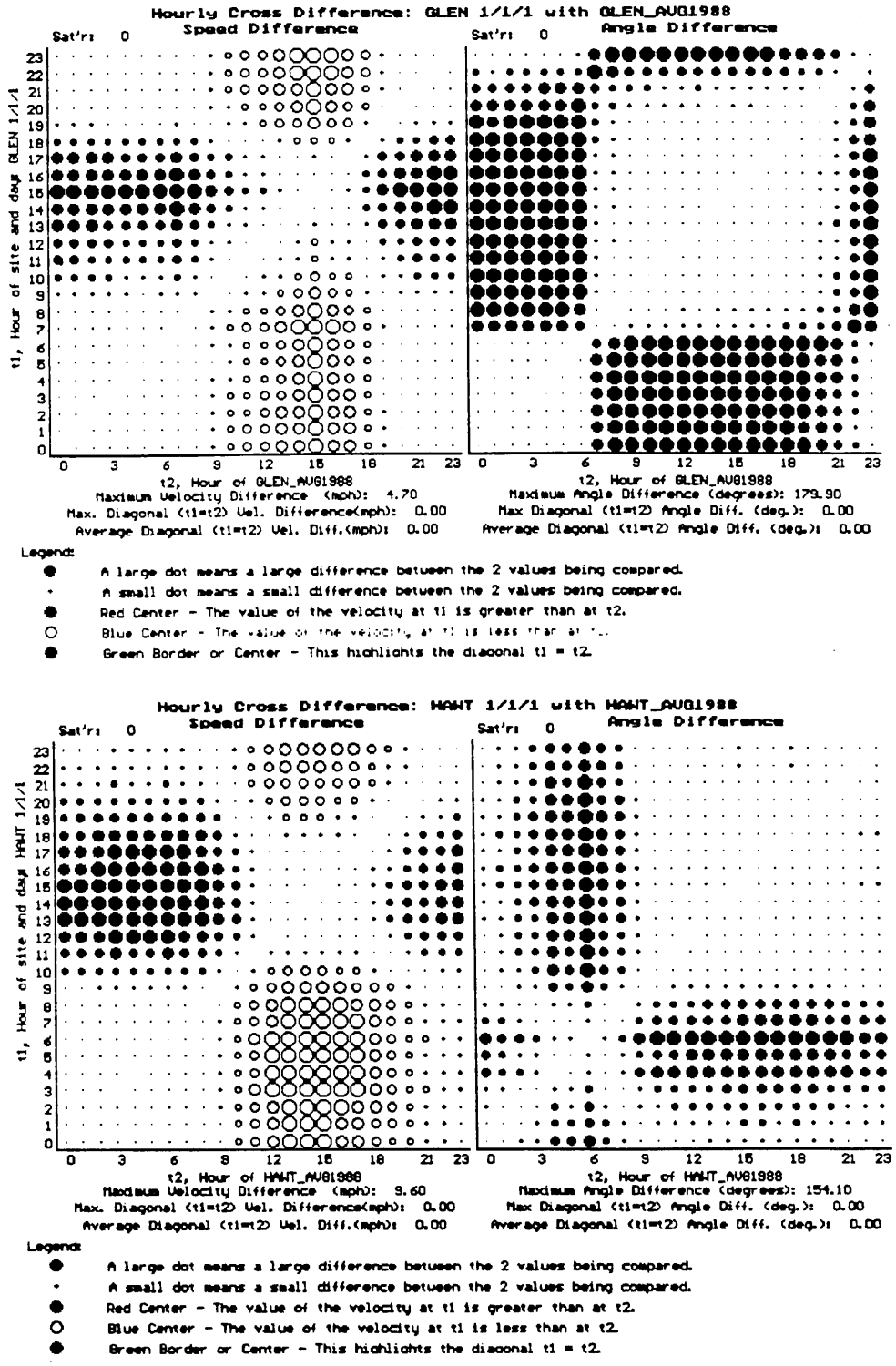


Figure 16. Speed and angle cross difference (top) Glendora (GLEN) average smog season day 1988 compared to itself, (bottom) Hawthorne (HAWT) average smog season day 1988 compared to itself.

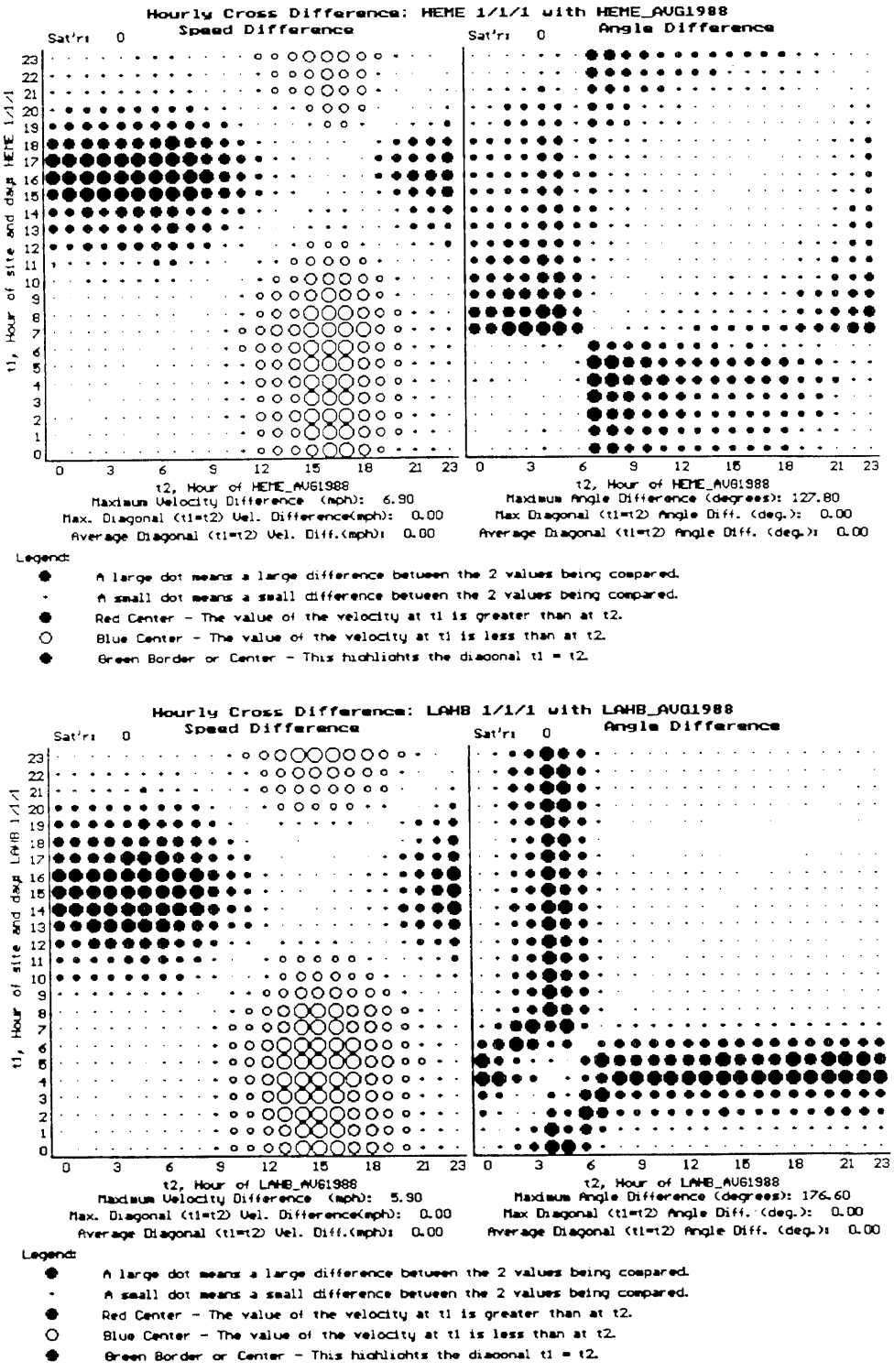


Figure 17. Speed and angle cross difference (top) Hemet (HEME) average smog season day 1988 compared to itself, (bottom) La Habra (LAHB) average smog season day 1988 compared to itself.

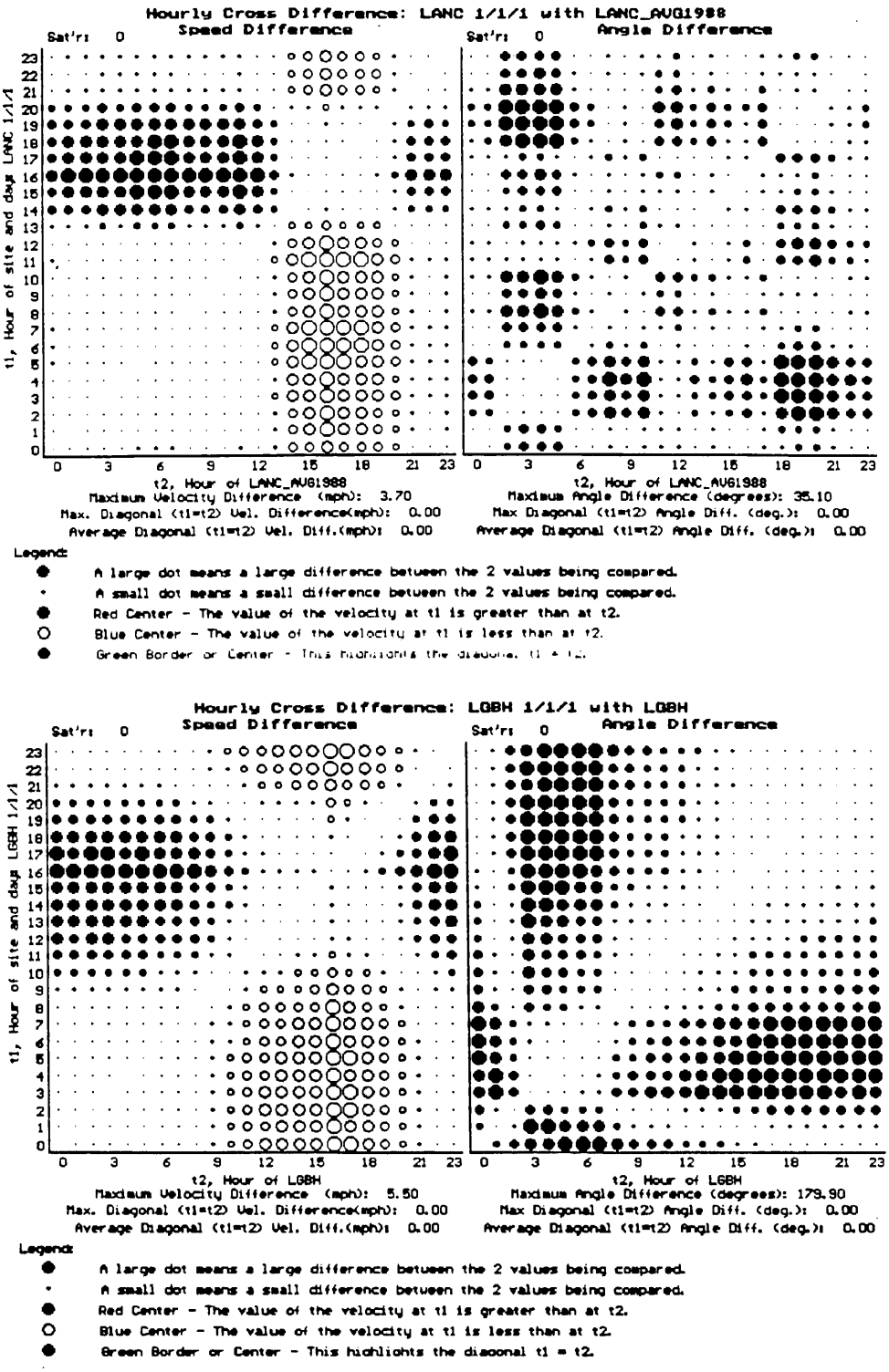


Figure 18. Speed and angle cross difference (top) Lancaster (LANC) average smog season day 1988 compared to itself, (bottom) Long Beach (LGBH) average smog season day 1988 compared to itself.

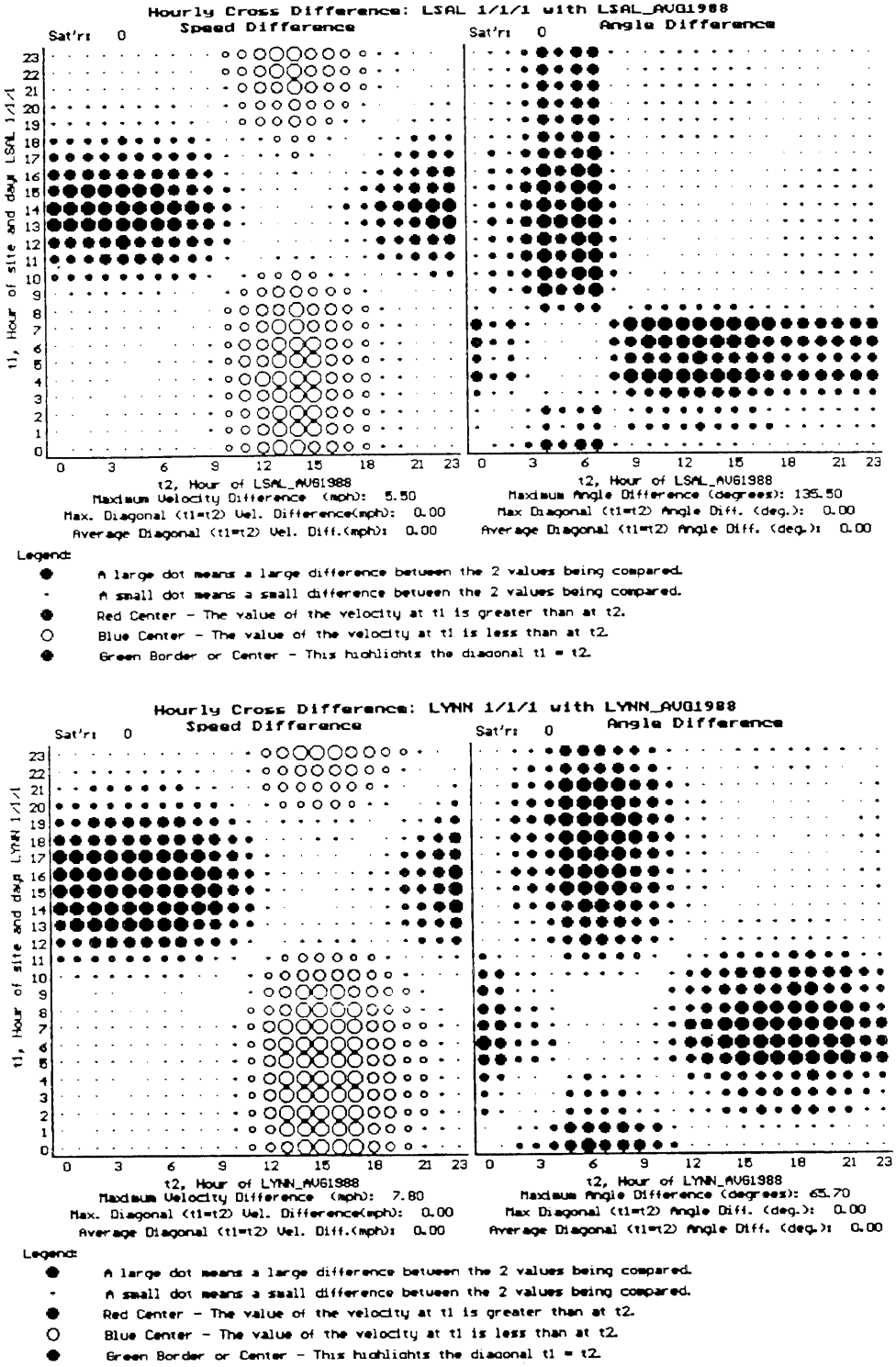


Figure 19. Speed and angle cross difference (top) Los Alamitos (LSAL) average smog season day 1988 compared to itself, (bottom) Lynwood (LYNN) average smog season day 1988 compared to itself.

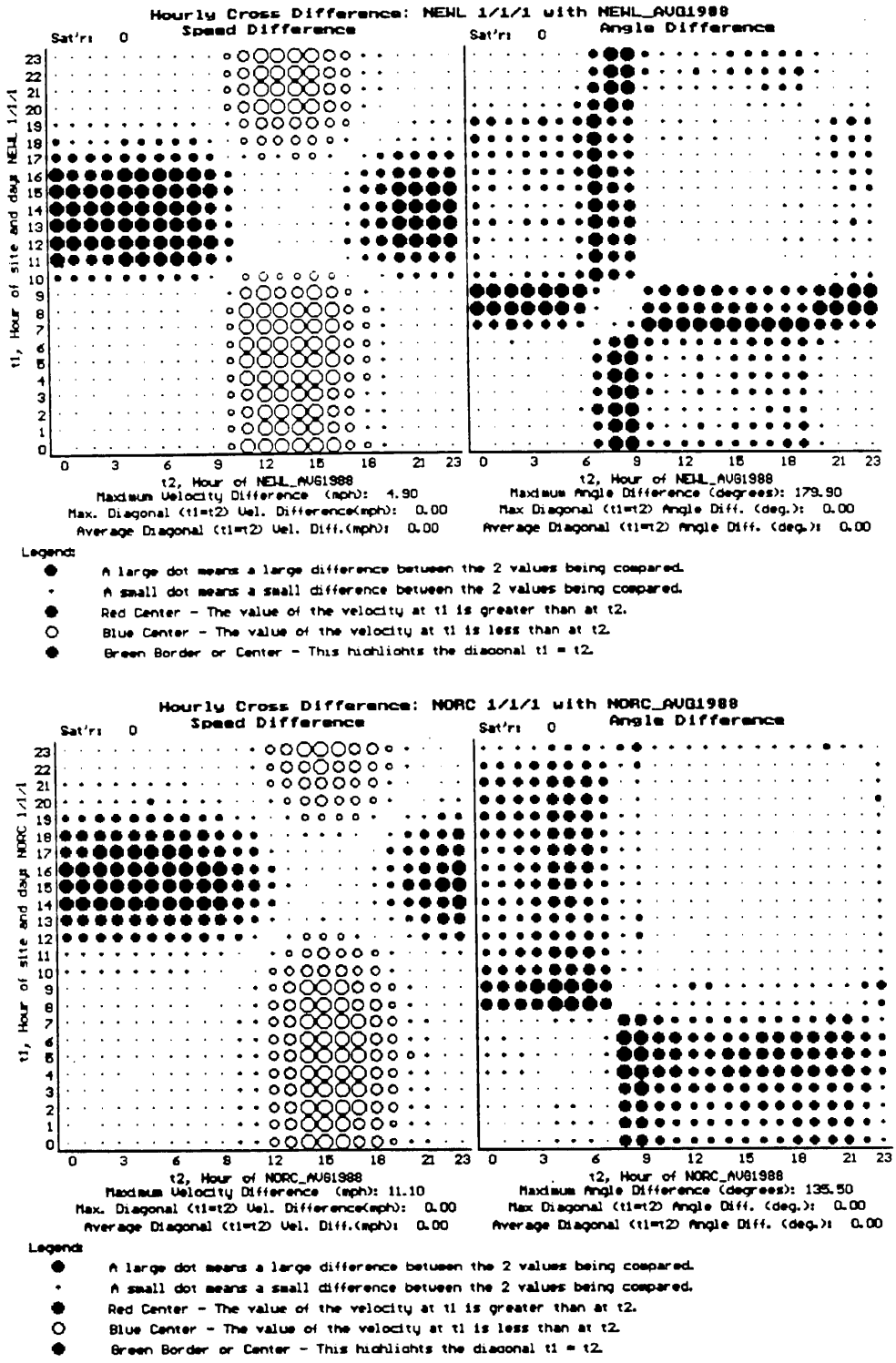


Figure 20. Speed and angle cross difference (top) Newhall (NEWL) average smog season day 1988 compared to itself, (bottom) Norco (NORC) average smog season day 1988 compared to itself.

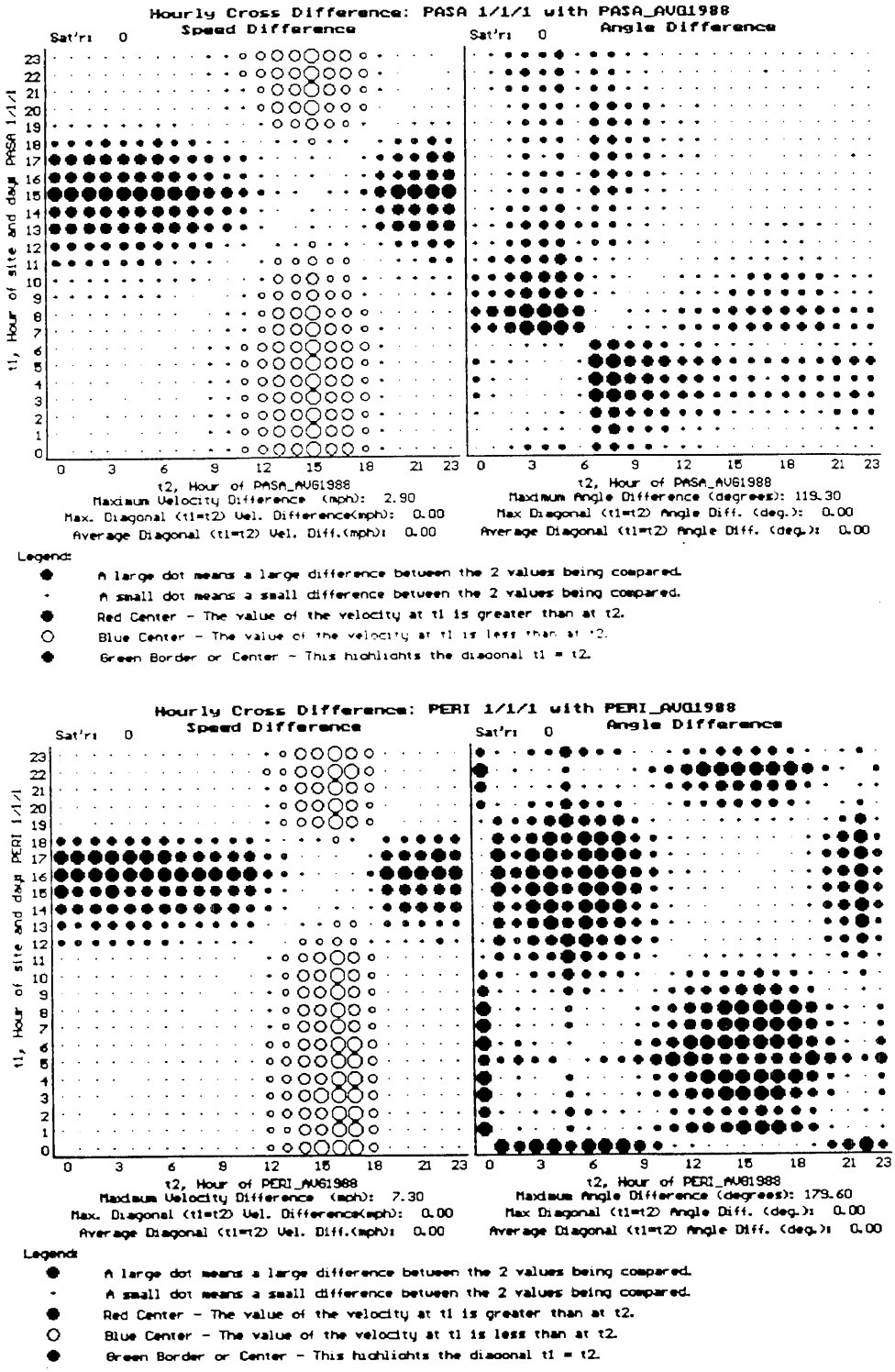


Figure 21. Speed and angle cross difference (top) Pasadena (PASA) average smog season day 1988 compared to itself, (bottom) Perris (PERI) average smog season day 1988 compared to itself.

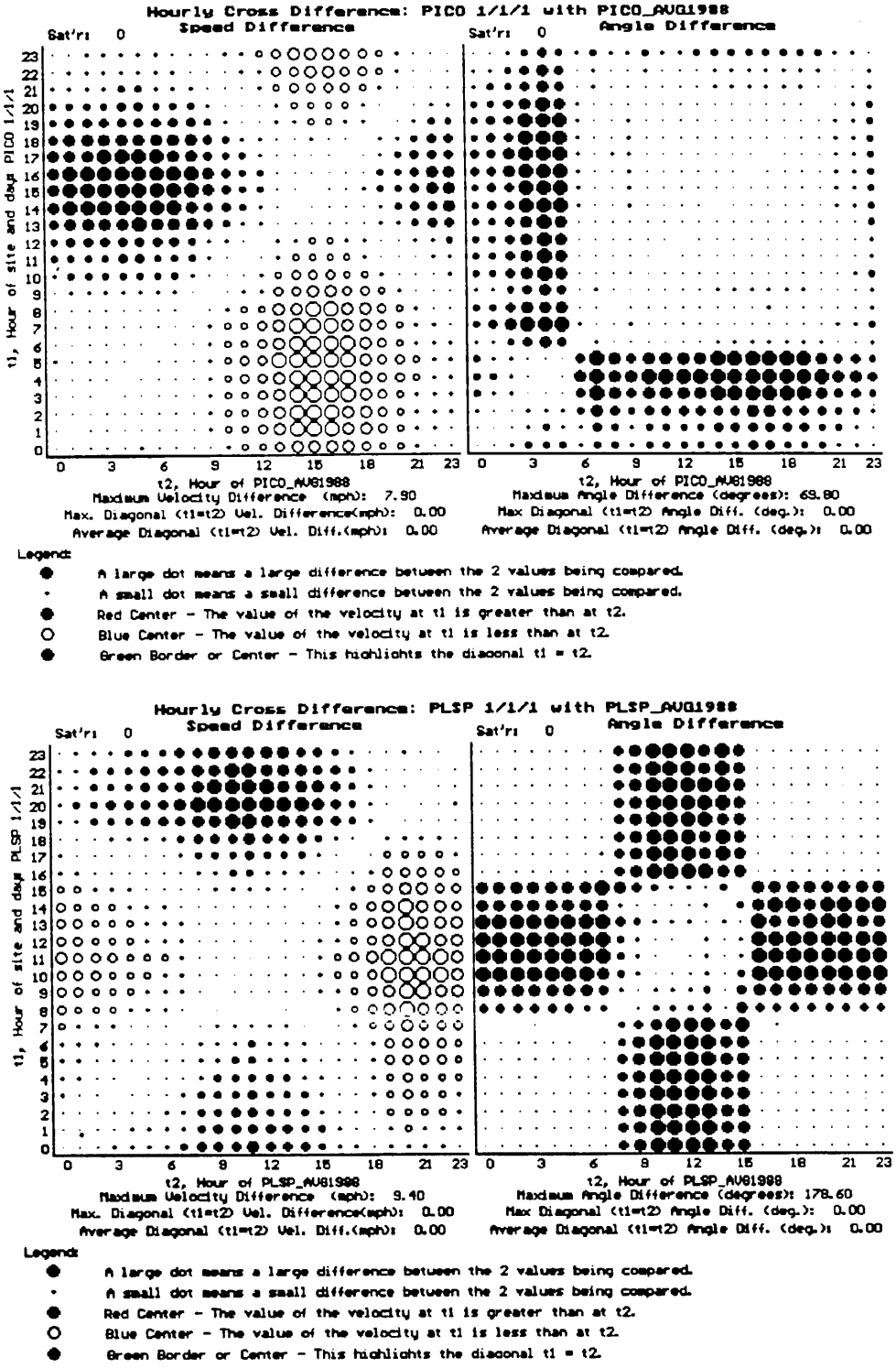


Figure 22. Speed and angle cross difference (top) Pico Rivera (PICO) average smog season day 1988 compared to itself, (bottom) Palm Springs (PLSP) average smog season day 1988 compared to itself.

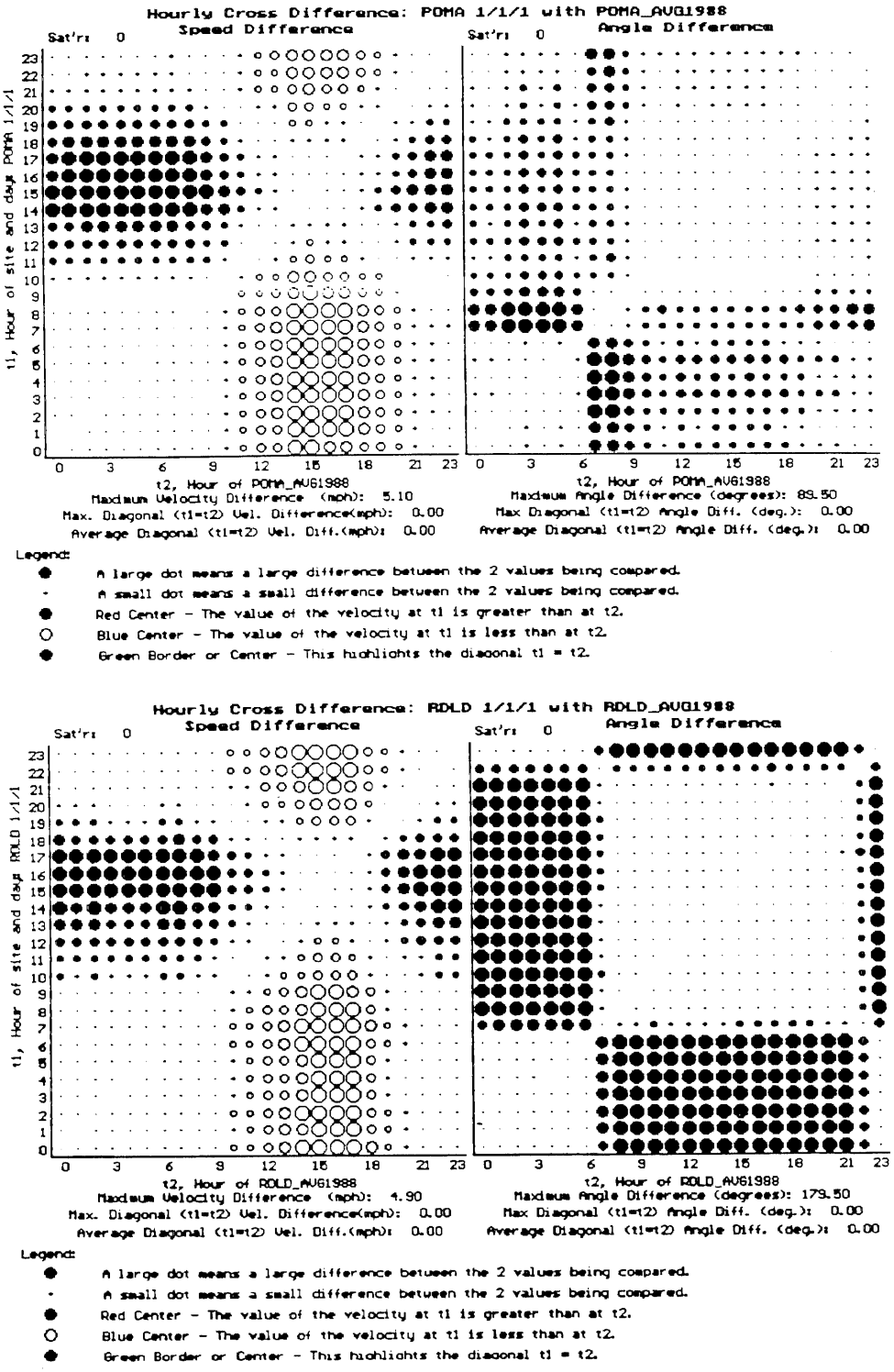


Figure 23. Speed and angle cross difference (top) Pomona (POMA) average smog season day 1988 compared to itself, (bottom) Redlands (RDLD) average smog season day 1988 compared to itself.

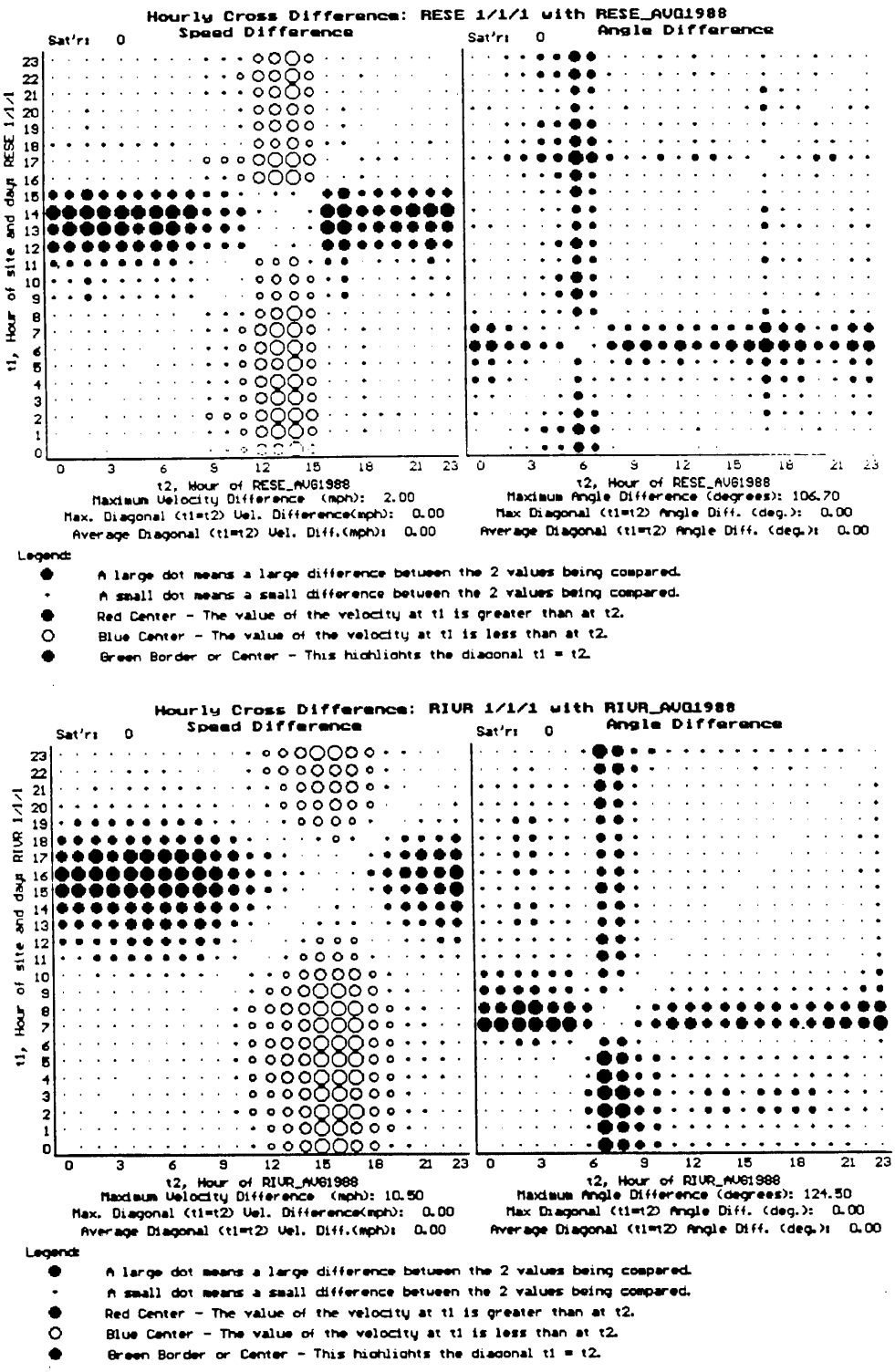


Figure 24. Speed and angle cross difference (top) Reseda (RESE) average smog season day 1988 compared to itself, (bottom) Riverside (RIVR) average smog season day 1988 compared to itself.

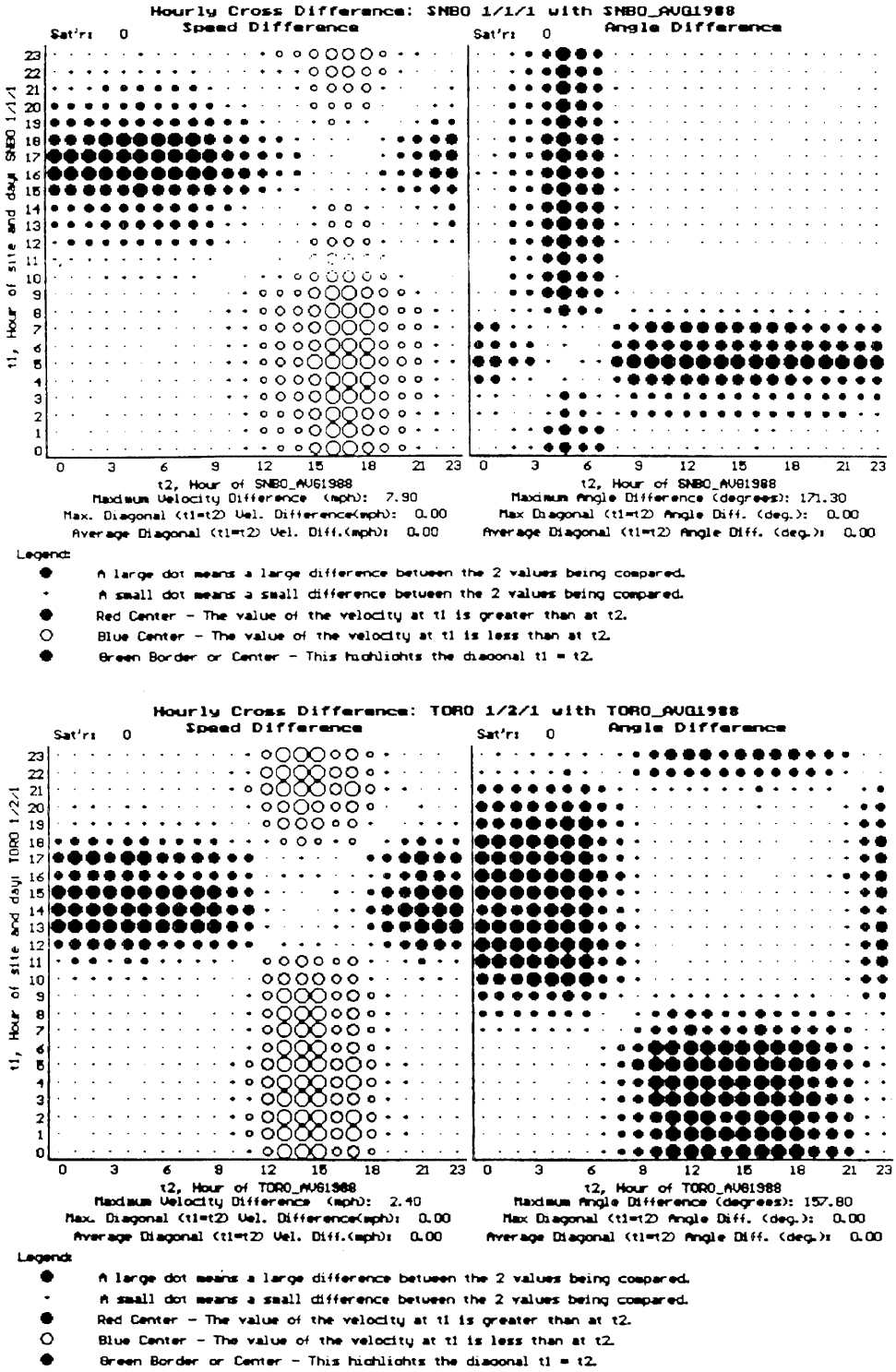


Figure 25. Speed and angle cross difference (top) San Bernardino (SNBO) average smog season day 1988 compared to itself, (bottom) El Toro (TORO) average smog season day 1988 compared to itself.

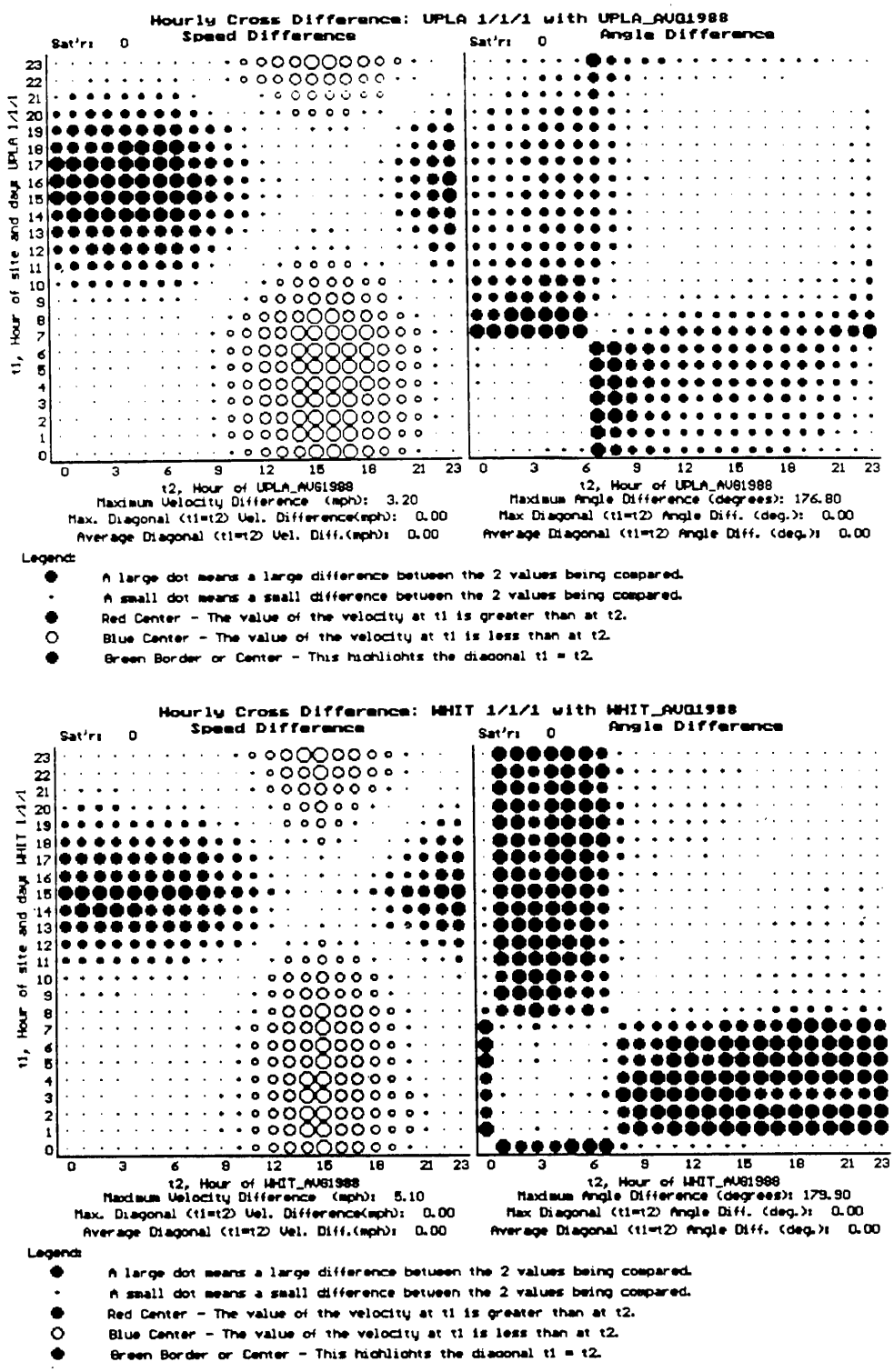


Figure 26. Speed and angle cross difference (top) Upland (UPLA) average smog season day 1988 compared to itself, (bottom) Whittier (WHIT) average smog season day 1988 compared to itself.

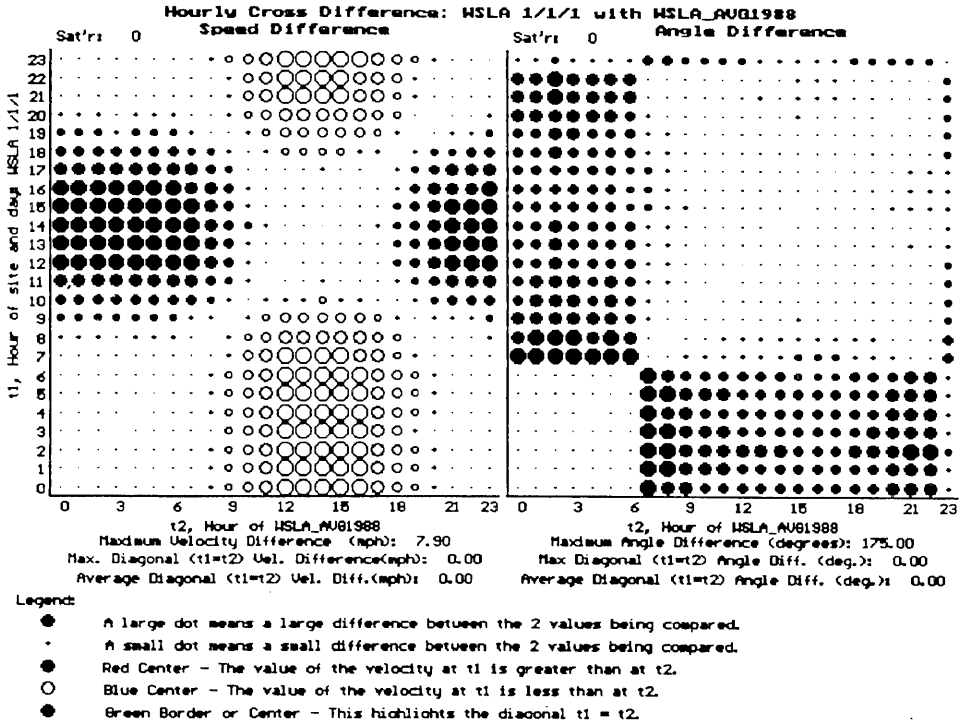


Figure 27. Speed and angle cross difference West Los Angeles (WSLA) average smog season day 1988 compared to itself.

References

1. Banta R. and Cotton W. R. (1981), "An Analysis of the Structure of Local Wind Systems in a Broad Mountain Basin," Journal of Applied Meteorology, **20** (11), pp. 1255-1266.
2. Blumen W. (1984), "An Observational Study of Instability and Turbulence in Nighttime Drainage Winds," Boundary-Layer Meteorology, **28**, pp. 245-269.
3. Defant F. (1951), "Local Winds," Compendium of Meteorology, Am. Met. Soc., Boston, MA, pp. 655-663.
4. Doran J.C. and Horst T. W. (1983), "Observations and Models of Simple Nocturnal Slope Flows," Journal of the Atmospheric Sciences, **40** (3), pp. 708-717.
5. Edinger J. G. (1959), "Changes in the Depth of the Marine Layer over Los Angeles Basin," Env. Sci. and Tech., **16** (3), p. 219.
6. Horrell R. S. (1987), "Atmospheric Buoyancy Driven Downslope Flows with Entrainment from a Ground Source," Masters Thesis, Department of Chemical Engineering, California Institute of Technology, Pasadena, CA 91125.
7. Horrell R. S., Deem M., Wyckoff P., Shair F. and Crawford N. (1989), "Ground Release SF_6 Tracer Experiments Used to Characterize Transport and Dispersion of Atmospheric Pollutants During the Southern California Air Quality Study of 1987," Air and Waste Association 82nd Annual Meeting, paper 89-138.2.
8. Kondo J. and Takeshi S. (1988), "A Simple Model of Drainage Flow on a Slope," Boundary-Layer Meteorology, **43**, pp. 103-123.
9. Scorer R.S. (1978), "Environmental Aerodynamics," Ellis Horwod, Chichester, U.K., p. 488.
10. Schroeder M.J. (1961), "Down-Canyon Afternoon Winds," Bulletin American Meteorological Society, **42** (8), pp. 527-542.

11. Shair et al. (1982), "Transport and Dispersion of Airborne Pollutants Associated with the Land Breeze-Sea Breeze System," Atmospheric Environment, **16** (9), pp. 2043-2053.

Chapter Three

Neural Network Computing Techniques to Process Air Quality Data

3.1 Introduction

The recent rapid growth in the field of collective computational systems frequently referred to as neuro-computing or neural networks is attributable to the recent work of Hopfield (1982) on associative memories, which showed how a collection of neuron-like processing units could be used to perform computationally intensive tasks. Further work by Rumelhart (1986) showed how networks which are strictly feedforward could be taught to map inputs to outputs even when the problem is not linearly separable. These authors and others have demonstrated how the parallel processing strategies utilized by biological systems can be emulated, and used in significant engineering and scientific applications.

Neural networks are attractive for several reasons. First, they are composed of many processing elements connected in a very parallel fashion. Therefore, a physical embodiment of such a system (Holler, 1989) capable of parallel data processing, will scale in time with the number of layers, rather than the number of weights as is the case with a sequential processor. Second, such systems are adaptive. If the circumstances in which the network performs its computations change, the network can be retrained to accommodate. Additionally, if trained properly, these networks can generalize. That is, they can interpolate and extrapolate over the input set, not suffering from the same deficiencies that expert systems suffer from when data is incomplete or not part of the training set (Fozzard, 1988).

Lapedes and Farber (1987) show that neural networks are able to accurately approximate a large class of functions. Additionally, they are capable of predicting the future for chaotic functions as well or better than known methods.

Recent developments have expanded the field to systems which are capable of handling time-dependent functions by means of information feedback within the network (Fang, 1990). Applications such as stock-market price pattern recognition, Kamijo (1990), and word recognition from a stream of characters (Benson,

1989) are among topics commonly being explored with recurrent networks.

An obscure application of neuro-computing exists which is relevant to the field of meteorology (Hu, 1964). Hu describes the use of an adaptive linear element (Adaline) system to do weather forecasting, specifically the occurrence of rain in the San Francisco Bay Area. He was able to conclude that the technique compared favorably with the U. S. Weather Bureau forecasts up to 24 hours into the future and was better from 24 to 36 hours into the future. More recently, good preliminary results have been obtained using a neuro-computing technique applied to the meteorological classification of satellite data (Smotroff, 1990). In both cases networks were proven capable of making classifications or forecasts which were traditionally done by experts.

The purpose of this paper is to demonstrate with two examples that additional applications in the science of meteorology and air quality data processing exist for neural networks. The first example demonstrates how automated Air Quality (AQ) data reconstruction can be accomplished using a network. The second example demonstrates the forecasting ability which can be embodied in a network. In both cases the examples were chosen to be representative of a class of commonly confronted AQ problems, and to produce results which were in and of themselves worthy of note. A second purpose was to explore an alternative approach for forecasting to the traditional approaches of numerical computing which have come under scrutiny (Zannetti, 1989).

3.2 Neuro – computing Strategy

The following section summarizes a general neuro-computing strategy. Persons unfamiliar with neuro-computing may find these introductory sections informative, and a useful preparation for the material that follows. Neuro-computing consists of 6 major steps:

1. Deciding whether a neuro-computing tactic is appropriate for a given computing task;
2. collecting and preprocessing the training data;
3. selecting a network architecture;
4. choosing the appropriate learning algorithm;
5. training the network;
6. computing with the trained network.

Often, when a physical system is being modeled with a neural network the state variables for the system are the inputs to the neural network, and the next or future states of the system are the outputs. In the discussion below the word "system" will be understood to mean physical system, and the word "task" will be understood to mean the task of modeling a physical system.

On the Appropriateness of a Neuro-computing Strategy

The decision to use a neuro-computing strategy to perform a given computing task is based on any combination of the following reasons:

1. The computing job would benefit from the speed of parallel processing or is inherently a parallel process. To actually obtain a speedup due to parallelism, it is necessary to utilize a physical system, i.e. an embodiment in hardware of a neural network, which takes advantage of the parallelism.
2. The functional mapping from input to output is not known, or understood, or practically obtained by known analytical methods.
3. The computing task to be accomplished would benefit from the ability to be updated as new or changing data are acquired.

4. An associative memory of past functional relationships is desired. Recurrent networks can perform "pattern completion," as well as, "pattern recognition" of inputs which are presented temporally.
5. The set of state variables for the system is incomplete. In this case, the network learns to generalize.

Collection of Training Data

The collection of data for training a neural network involves gathering data from a representative set of input variables and the associated output variables for the system being studied. Networks typically have multiple inputs and outputs, and are therefore referred to as input and output vectors. The notation convention I_i refers to the ith input element, and O_j refers to the output from the jth neuron. The outputs which are part of the training set are typically referred to as *target outputs*, *target vectors*, or just *targets*. The notation for targets is the same as for outputs except the letter T is used. The set of input and target vectors used to train the network is referred to as the *training set*. When the number of input variables is large or thought to contain duplicate or spurious inputs a principal component analysis (PCA) may be used to reduce the number of inputs down to the most significant subset of inputs (Baldi, 1989).

Training of a Network

Training a neural network is defined, for the purpose of this manuscript, as the process of presenting input vectors from the training set to the network (in its current state), and comparing the associated output vectors the network computes to the target output. The difference between the output and target vectors determines the error for that presentation. A vector of errors is computed for each input in the training set, one error for each output which is to be compared with a target. An algorithm such as error back propagation (Rumelhart, 1986), Madaline III, Real-Time Recurrent Learning Algorithm (Williams, 1988), etc. is used to adjust the internal parameters of the network to minimize

the error at each output. This definition applies to a mainstream class of neurocomputing strategies and is generally referred to as supervised training. This is a general definition, but it does not apply to all types of networks. Specific training algorithms will be discussed below.

Computation with the Network

The neural networks discussed herein achieve their computational power by performing the following calculation for a given input. The output of neuron j, O_j, is computed using Equations 3.2.1 - 3.2.4.

$$O_j = f_j[\sigma(I,O,W)] \tag{3.2.1}$$

where I and O are the input and output vectors, and W is the weight matrix which connects the inputs to the outputs and which also connects each of the outputs to all of the output units. The function, $f_j[\sigma]$, is the transfer function, and is typically a sigmoidal function like,

$$f(\sigma) = a + b \cdot tanh(\sigma), \tag{3.2.2}$$

or the so-called sigmoidal squashing function

$$f(\sigma) = a + \frac{b}{1 + e^{-\beta\sigma}}. \tag{3.2.3}$$

The function σ_j is computed by,

$$\sigma_j(I,O,W) = \sum_{k=1}^{m} I_k W_{jk} + \sum_{k=m}^{m+n} O_k W_{jk} - \gamma_j \tag{3.2.4}$$

where there are m inputs and n outputs. The symbol γ_j represents the adjustable bias weight associated with each neuron. Each neuron sums the contributions of all the inputs (the first sum in Equation 3.2.4) with the contributions from all of the neurons (the second sum in Equation 3.2.4), and subtracts off a fixed bias (γ in Equation 3.2.4). This sum is then passed through a nonlinear function (eg Equations 3.2.2 or 3.2.3) to produce the output of each of the neurons.

The above formulation describes a general, fully interconnected network which includes feedback. This definition can be reduced to a feedforward system by holding all weights which feed information backward in the system to zero. Something quite noteworthy is that a fully connected recurrent network can actually be trained to zero out any connections which feedback if the problem is inherently feedforward. The network learns that the recurrent weights are superfluous. From an engineering and application point of view this is of considerable value since the architecture of the network need not be assumed, only the number of neurons.

One measure of a network's performance and an indicator of the degree to which a network has been trained is the sum squared error defined in Equation 3.2.5 for input-output pattern q.

$$\epsilon_q = \frac{\sum_{j=1}^{n}[T_j - O_j]^2}{n} \tag{3.2.5}$$

where n is the number of output units. The sum squared error is computed for each input-output pair. A means of comparing network performance between two separate input-ouput pattern files is the *average epoch sum squared error* defined by Equation 3.2.6. It is the average of the sum squared errors in a single training set,

$$E = \frac{\sum_{j=1}^{p}\epsilon_j}{p} \tag{3.2.6}$$

where p is the number of pattern pairs presented in an epoch. Epoch is defined as the corpus of input and output pairs presented to the network for training.

Training Algorithms

Two network training algorithms are employed in the simulations to be described below: the feedforward form of the Madaline III algorithm by Andes (1989) is used in Section 3.3, and the recurrent algorithm of Williams and Zipser (1988) in Section 3.4.

Training Feedforward Networks

There are several algorithms to train feedforward networks of the architecture discussed in Section 3.2. Among them are backpropagation (Rumelhart, 1986) which has been discussed and used extensively in neuro-computing literature, and more recently Madaline III (Andes, 1989). The Madaline III algorithm is the most recent form of the Adaline (Widrow, 1988) family of adaptive element training algorithms. This algorithm has the particular distinction of being capable of learning on analog devices where the transfer functions are not as precisely described as in Equations 3.2.1 to 3.2.4. To obtain the increased speed benefit which neural networks offer, the use of hardware implementations is essential. The feasibility of such a use was sought, and therefore Madaline III was used. Madaline III is a very straight forward algorithm which performs gradient descent on the sum squared error in weight space. The reason it is useful for hardware implementations of a neural network is that it accounts for the nonidealities of the processing elements. Simply put, Madaline III perturbs the inputs of each neuron at every layer and determines whether the mean squared error is reduced or not. With this information and the input to each weight, a weight update rule is obtained.

Training Recurrent Networks

A network which has feedback among neurons is termed a recurrent network. Algorithms to train this sort or network have appeared only recently and have been authored by Williams and Zipser (1988), Pineda (1989), and Pearlmutter (1989). Additional computational overhead is acquired when training recurrent

networks compared to feedforward only networks. From an engineering point of view the additional effort may be worthwhile since this form of network is capable of organizing itself into its own architecture. The algorithm used by Williams and Zipser begins training with each neuron connected to every other neuron, and every input from the outside world. As training proceeds the elements of the weight matrix which connects some neurons to the external inputs, and to other neurons disappear. This process of zeroing weights isolates neurons into hidden layers. In applications where the required number of neurons, and layering of the neurons is not known *a priori*, training a recurrent network can provide the answers.

Among other things, recurrent connections confer upon a network the ability to identify patterns in a time series, and to predict future patterns. For this reason recurrent networks are well suited to forecasting applications.

3.3 A Feedforward Network for Ozone Data Reconstruction

A two-layer feedforward network was used to reconstruct air quality (AQ) data from incomplete data sets. The task of reconstructing missing AQ data is commonly done by hand by an air quality expert. For small data sets this is practical, but for the larger data sets which are generated in an urban area, i.e. Los Angeles, this may require a considerable amount of man-hours. Consider for example that roughly 100 air quality variables are being measured and recorded from 30 stations for each of the 24 hours in a single day by the South Coast Air Quality Management District in the Los Angeles Basin. Somewhere between four to eight percent of this data is routinely missing or upwards of 5760 data points daily, or approximately two million data points yearly. (This estimate is made based on the ozone data reported by the SCAQMD for 1987).

Neuro-computing offers a viable means of handling missing AQ data. The restoration technique can be used to minimize the time devoted to data reconstruction by AQ experts. This can be done as a pre-processing step prior to data archival or as an imbedded data filter in an existing application which accepts the raw data.

The neuro-computing solution offers an alternative approach for handling AQ data records with missing entries. It uses historical data as a basis for the data reconstruction, and can be updated continuously as new data is collected. Seasonal variations can easily be incorporated into the neural network solution by training individual networks on single seasons. It can reconstruct data across missing "peaks," a deficiency in the data which often leads to the data being discarded, and at minimum requires an expert's judgement to reconstruct the missing data. It can generate a flag which is a measure of the likely performance of the network's restoration based on previous restoration performance.

The simulations performed were done to: (1) validate the above claims, and (2) qualitatively assess the performance of the network in performing data

reconstruction.

Method

A 24 input 24 output, two-layer feedforward network was used to perform ozone data reconstruction. The network was presented with 24 one-hour-averaged values of ozone measured at Newhall, CA, one hour at each input. The network was then trained to output the identical pattern, filling in data at any missing hours. Figure 1 (top) shows the input, output, target, and error of the trained network for July 14, 1989 with a full input set. Figure 1 (bottom) shows the same networks output when an incomplete input is applied to the network.

The network was trained on a single month's worth of data, the 30 days of June 1988. In addition, "missing data" days were created by introducing defects into the training data set in the following way. A day in the data set was duplicated, then one, two, or three inputs were zeroed indicating that the data was missing. The corresponding outputs were kept intact. At least one duplicate day was made for each day of June.

Because the peak in the ozone occurs at approximately the same time of the day during this month it was clear from preliminary training simulations that the network would be incapable of generalizing to be able to reconstruct peaks when the peak occurred more than a couple hours from the average. To alleviate this problem ten additional days were created from the June days by shifting the data forwards or backwards in time up to five hours. (Actual days where the peak was shifted were too infrequent to justify collecting them). "Missing data" days were created for the time shifted days also. In total, there were 70 I/O pairs in the training set.

The sigmoid transfer function was adjusted to give it an output range from 0 to 30, which corresponds to the part per hundred million range of nearly all ozone observations in Los Angeles. To achieve this, the following characteristic parameters were used: $a = 0, b = 30$, and the gain, β was 0.001. The number

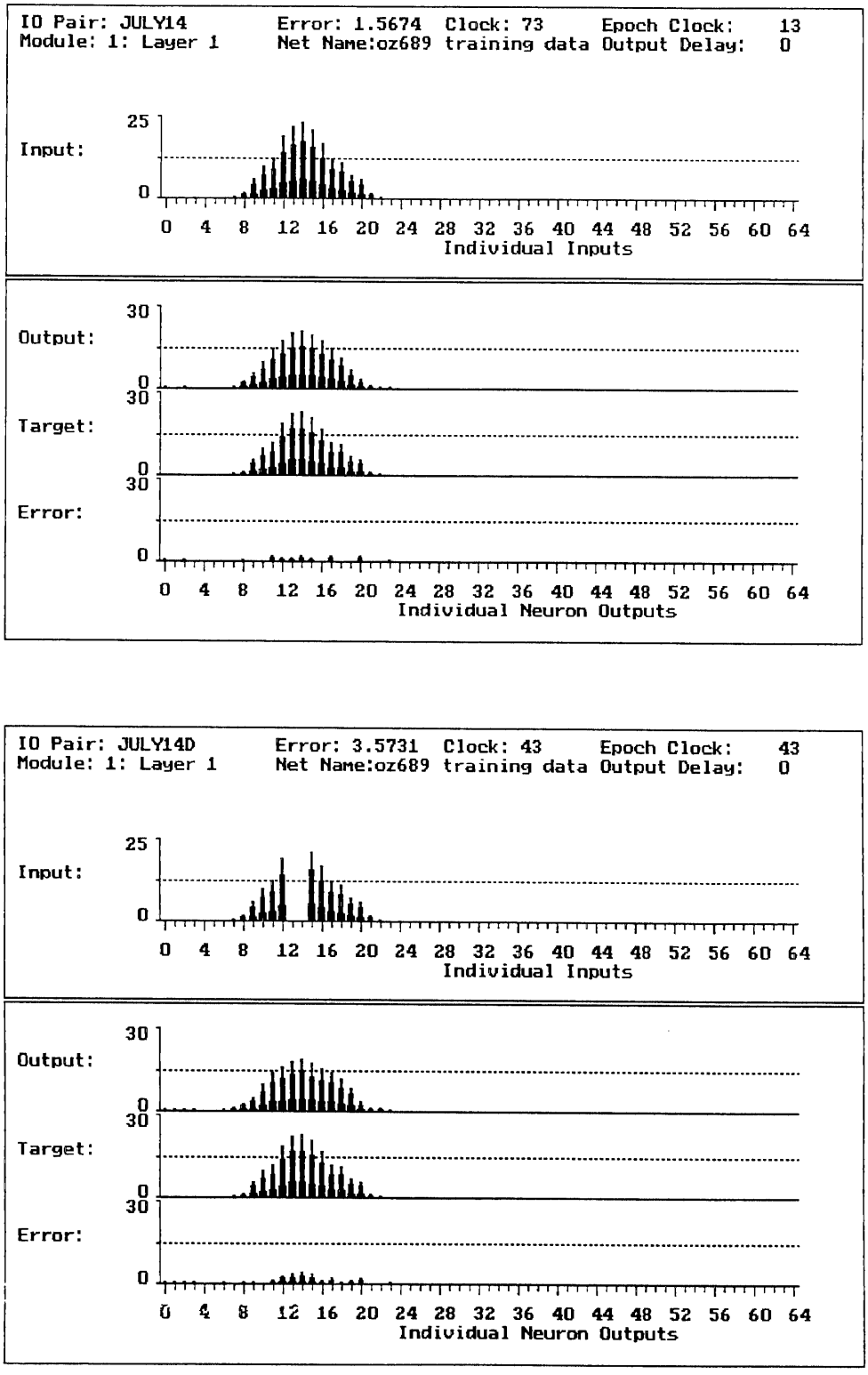


Figure 1. Input, output, target and error for the sample input July 14, 1989: (top) complete input, (bottom) input with two deletions at hours 13 and 14.

of neurons in the second (or hidden) layer was fixed at 24.

The Madaline III algorithm was used to train the network. The algorithm was implemented via DynaMind version 0.8 (NeuroDynamX 1990) PC-based simulator. Training was done on a Zenith Z-386/20 with 80387 math coprocessor. Approximately 500 epochs were required to train the network. Following training the network was tested on two month's worth of data from July 1989 and August 1989. Each simulation required approximately 2 hours. The learning rate was set to 0.05, and was found easily by trial and error.

Results

The average epoch sum squared error, Equation 3.2.6, for the June training data was 1.43 and for the July and August test set it was 2.04. A network with randomized weights has a value of 227 for the June training data. The similarity between these two values indicates that the network was able to process the new data nearly as well as the training data. In physical terms it means the filtering process imposed by the network alters each ozone value by approximately one pphm. For small ozone values, around one pphm, this is of course significant, however it is the large values which are of the greatest concern. For the larger ozone readings, 15 pphm, this means less than 7 percent error.

The effectiveness of the network for data reconstruction can be summarized qualitatively as successful with the following caveats. An increasing number of missing data points decreases the capacity for the network to fill in the missing points. Contiguous missing data points are reconstructed with greater accuracy than missing data which is scattered, especially when the missing points lie within several hours of one another. Figure 2 shows data reconstruction where deletions are not contiguous. Compare this to Figure 1, a similar day, noting the larger error which occurs in reconstruction when the missing hours of data are separated. In cases where the reconstruction of the data was poor, the ozone values surrounding the missing data points tends to be altered, generally in the

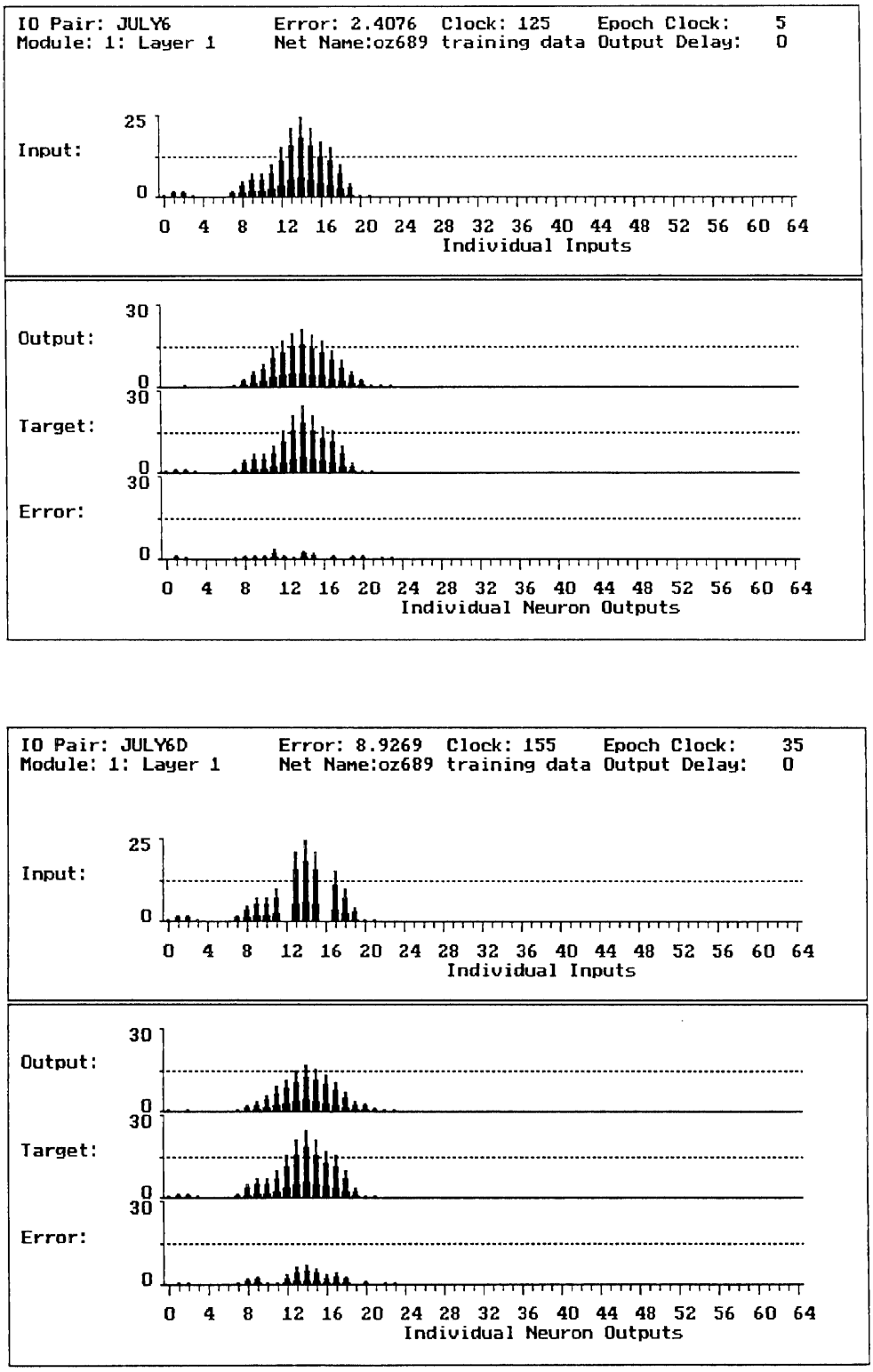


Figure 2. Input, output, target and error for the sample input July 6, 1989: (top) complete input, (bottom) input with two deletions at hours 12 and 16.

direction of being reduced.

Quantitatively the network is capable of reconstructing most missing points, on average, to better than $\pm$ 10 percent of their true value.

3.4 A Recurrent Network for Short – term Ozone Forecasting

Method

A fully interconnected recurrent network was used to continuously forecast short-term ozone levels at a single location based on the previous sequence of hourly averaged ozone levels observed at that location. The network is presented with a single ozone level in pphm from a single site in the Los Angeles Basin and trained to continuously make three predictions of the ozone levels at either: one, two and three hours, or at four, five, and six hours following the hour of the presented value. Through the use of the recurrent connections it is expected that the network will learn how many time steps from the past series of ozone presentations it requires to make its prediction.

The William and Zipser algorithm was used to do the training. The protocol was to present a single input, the ozone from a single hour, to the network and use the ozone levels at three future hours as targets for training. Each day (starting at hour 0) was treated as an independent series, and series were shown to the network in random order. This limited the network to making its prediction based only on the portion of the current day's series which it had seen as input. This decouples the prediction from the time series of the previous day. 24 one hour ozone levels for a given day. The training data for the network was the 30 days' of June 1989 taken at Newhall, CA.

The sigmoidal transfer function used by the network was identical to the one described above for the data reconstructing application. The learning rate was 0.3, which was determined by past experience using this algorithm. Approximately 500 epochs were required to train the network. Following training, the network's performance was tested on one month's worth of data from June 1988.

Since the optimum number of neurons which was obtained for the simulation was not known *a priori*, this was determined by simulation experiments. Three neurons is the lower limit, since three outputs are desired for each input (as

described above). Since simulation time with this algorithm scales with the number of neurons to the fourth power it is desirable to find the fewest number of neurons, above three, after which the epoch average sum squared error does not decrease. Simulations were done using from three to nine neurons. Each simulation was allowed to converge, but was trained on no less than 500 epochs in order to assure comparability between runs. Two simulations were done for each number of neurons. Figure 1 shows a plot of the epoch average sum squared error versus the number of neurons. From this figure it was determined that six neurons was the optimum choice for the rest of the simulations. Two networks were then trained using six neurons each, the first to make predictions for hours one, two and three, the second to make predictions for hours four, five, and six into the future. Under these conditions the first three neurons were arbitrarily assigned the job of making the prediction and neurons four, five, and six were allowed to "float" without being trained. To allow these neurons to float gives them the opportunity to arrange themselves into a hidden layer by the mechanisms discussed previously.

Results

Figures 2 and 3 display the ability of the two networks to predict from one to six hours into the future, from each hour of the day. Each error point is the average error for the forecasts based on 30 days' worth of data. Figure 2 shows the errors obtained from the test data set, and for comparison Figure 3 shows the errors associated with the training data. Note how the test data set has nearly identical errors to the training data indicating that the performance of the network is as good on the test set as it is on the training data. The peak centered around hour 9, seen in the error curves of all the predictions longer than two hours into the future has a useful interpretaion. It results from the increased uncertainty of predicting the peak ozone value which occurs diurnally around hour 14 at Newhall. Logically, each hour further into the future has increased

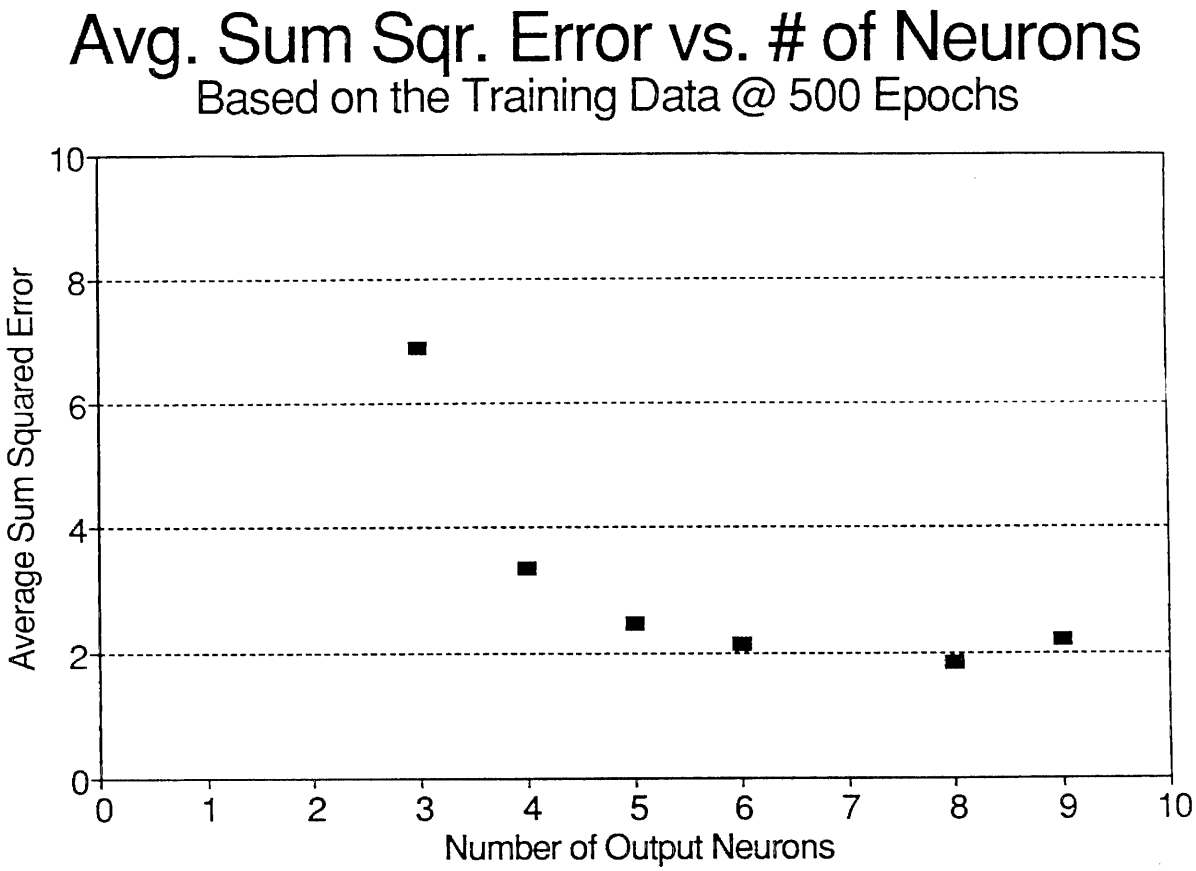


Figure 1. Average sum square error versus the number of neurons.

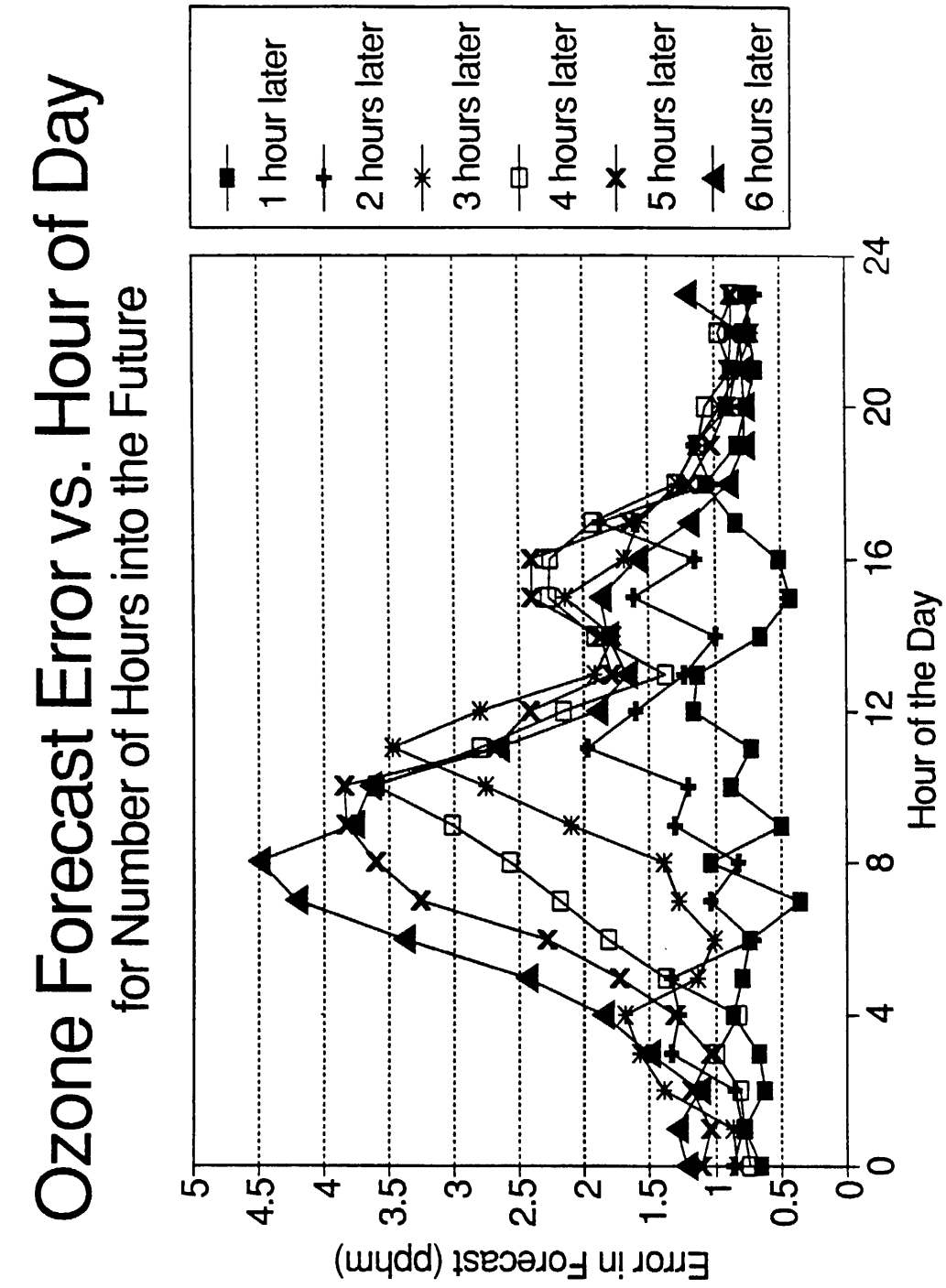


Figure 2. Test data: ozone forecast error versus the hour of the day.

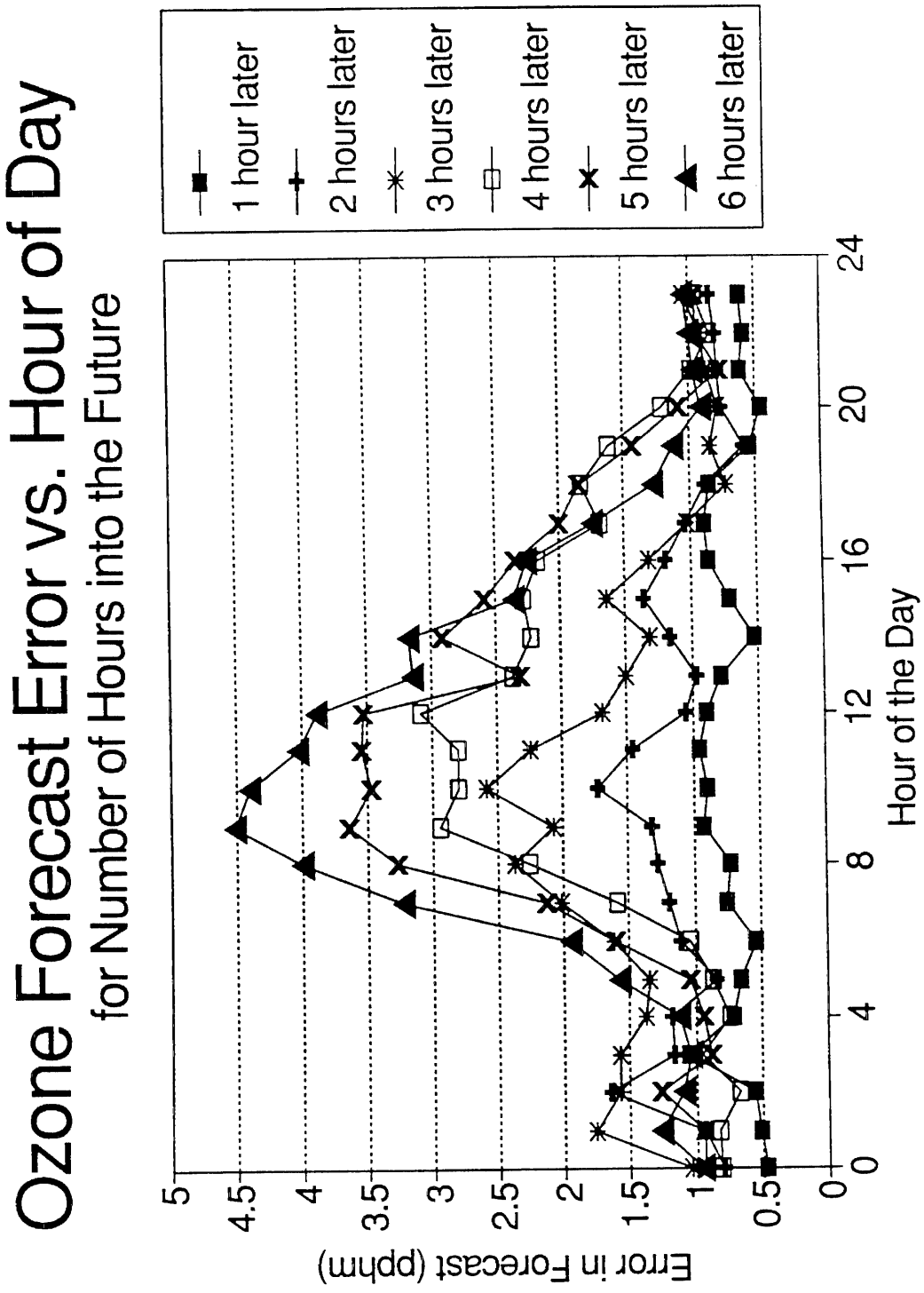


Figure 3. Training data: ozone forecast error versus the hour of the day.

uncertainty in the prediction, and that is the trend borne out in Figures 2 and 3.

Figure 4 shows the comparison of the actual versus the predicted value for a forecast three hours into the future. The network is clearly capable of following the true trajectory of the ozone in time very accurately. The correlation coefficient of the forecast to the observed value is 0.82 with a standard deviation of 2.2 for Figure 4. The performance of the network decays steadily to a correlation coefficient of 0.46 and a standard deviation of 6.1 for a six hour prediction. Figure 5 shows the output from the network for the longer predictions, four, five, and six hours into the future. Besides showing a time series of the networks longer term predictions it demonstrates the oscillating behavior which the network may be subject to on start-up. After approximately one day's worth of data the network settles into the desired forecasting mode. The possibility of a start-up instability is present when the network contains recurrent connections. In this case, predictions are not valid for approximately one day's worth of data.

The outputs of neurons four, five, and six (ie. the hidden units) as a function of time are shown in Figure 6. This figure provides some insight as to how the recurrent network is storing state information and using this information to make predictions. Notice that each hidden unit's output is an oscillator, each with a different amplitude, and phase. Though the mechanism is not completely understood, it appears that the input affects the amplitude of each of these outputs, and then collectively these three hidden outputs are combined to make the prediction. Thus it appears that the state of the network is maintained by three oscillators. The strong feedback coupling is observed during the start-up of the network which makes predictions farther into the future. This is reasonable since such a network may be predicting the future decline in the ozone level while it is experiencing an increase in the level at its input (see Figure 5).

The novelty of this approach makes it hard to compare to other methods

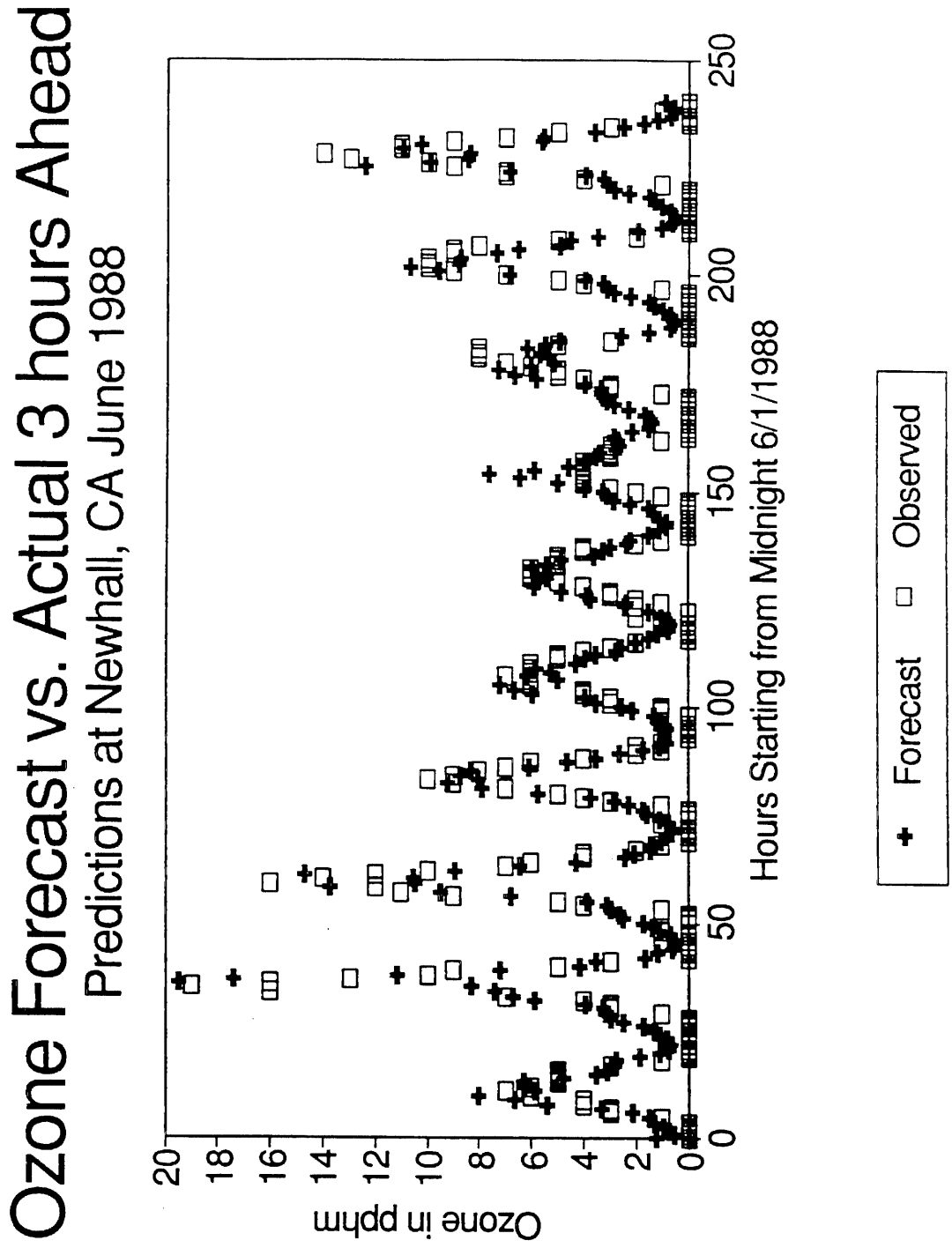


Figure 4. Ozone forecast versus the actual observed level three hours ahead.

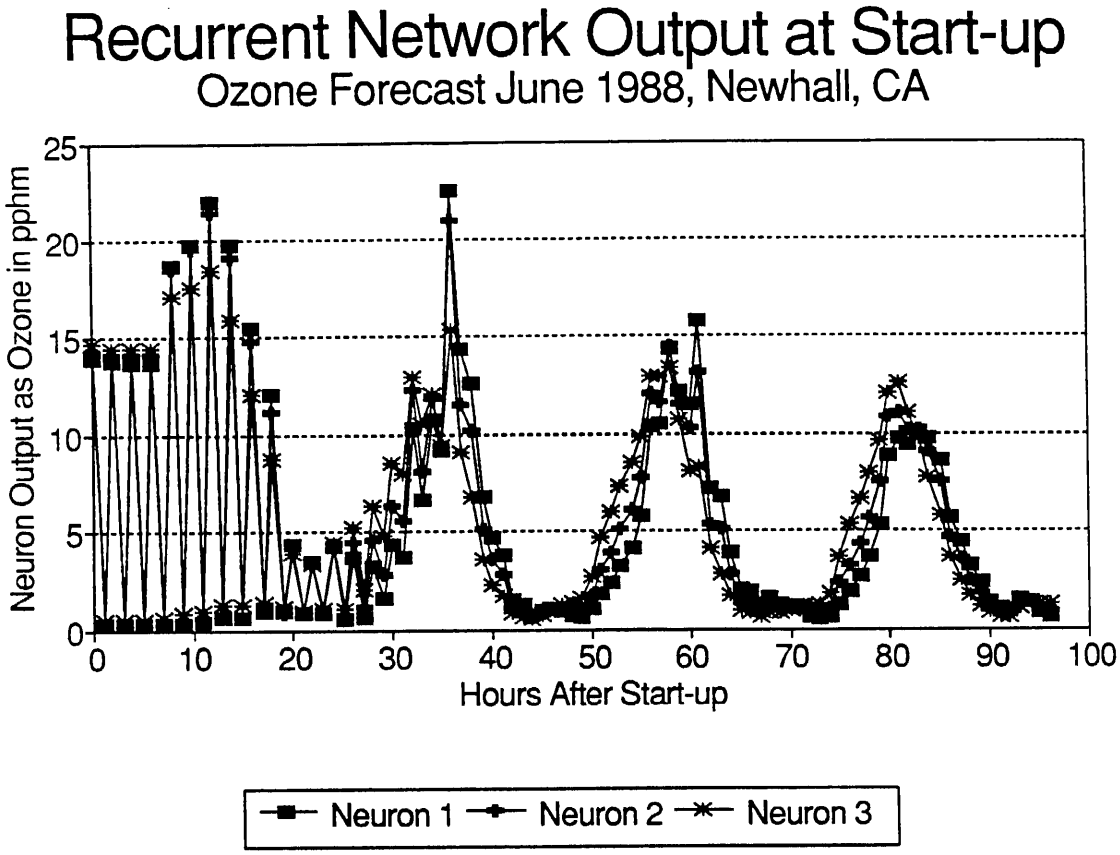


Figure 5. Recurrent network output at start-up for predictions four, five, and six hours into the future.

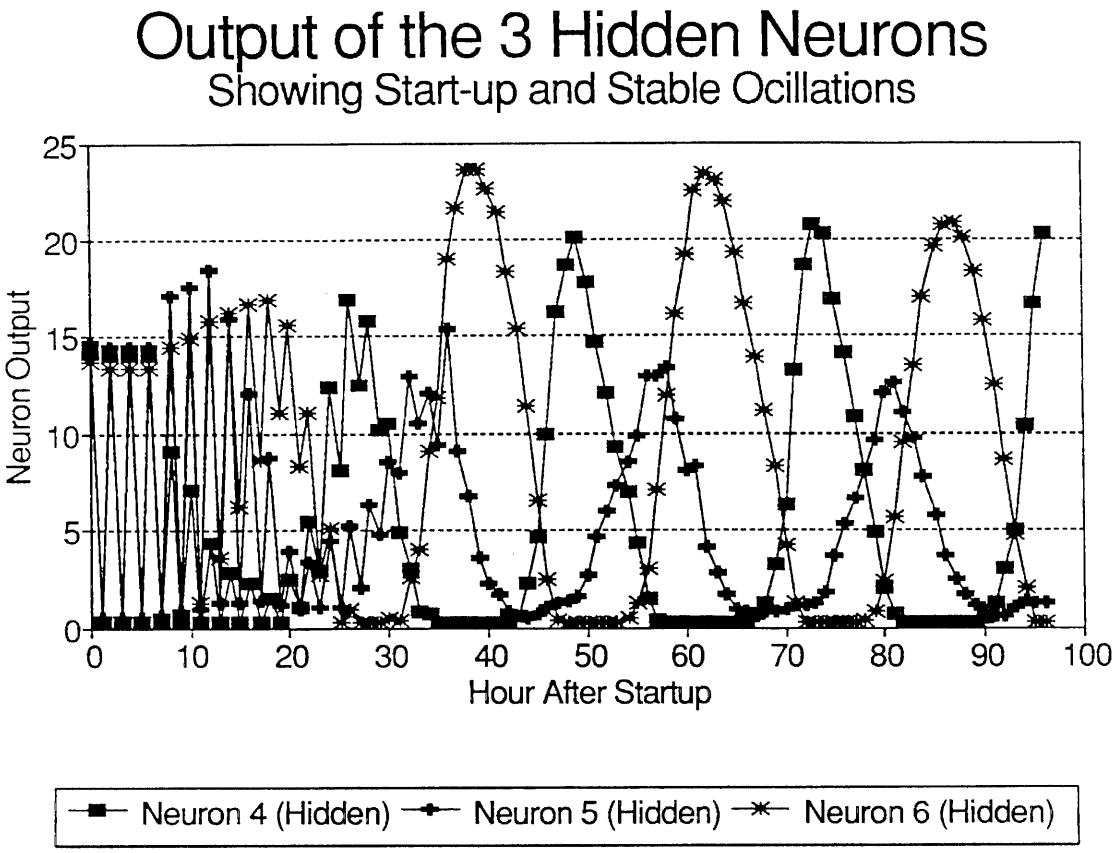


Figure 6. Output of the three hidden neurons as a function of time when predicting four, five, and six hours into the future.

such as the statistical forecast models for daily maximum ozone concentration (Robeson, 1990), though it is clear that this method can be generalized to do this type of forecasting. In lieu of a direct comparison two comparisons were constructed. (1) A linear regression was done using the value of ozone at hour nine to forecast from one to five hours into the future. (2) The 30 day average value was computed from the training set and used as a guess. Hour nine was chosen because the network produced the largest errors a this time when forecasting into the future. The relatively large error suggests that this is the most difficult time from which to forecast, and might be a weak point for the network. The resulting comparison of the methods is made in Figure 7. The neural network systematically had a lower error than either of the other two approaches. It should be noted that the standard deviation of all three methods was sufficiently large so as to prevent the three methods from being statistically distinguished from one another.

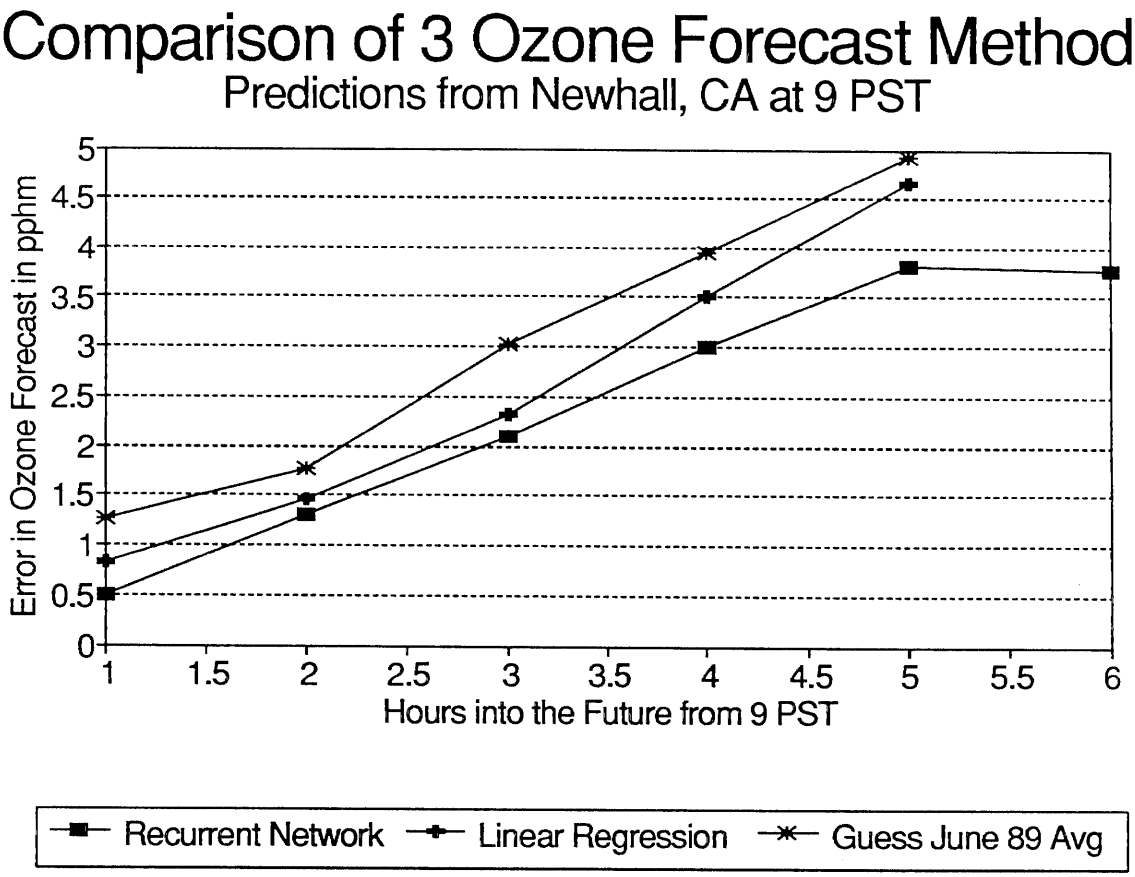


Figure 7. Comparison of three ozone forecast methods: (1) Recurrent network, (2)Linear regression, and (3) Seasonal Average.

Conclusions

Neural networks were capable of being configured and trained for the two classes of problems discussed. In both cases the network's performance was adequate for the task being performed, though direct comparisons with other methods were lacking as a definitive yardstick for comparison. Though it was not elaborated on, neuro-computing is the only method known to this author where the prediction or reconstruction made by the network can also include a measure of the networks confidence in its result.

The techniques employed can be extended to more complex cases with more diverse input sets. For instance, the temperature dependence, wind speed, atmospheric mixing depth, and ozone levels at other locations can all serve as inputs to either the data reconstruction or forecasting network. From the previous discussion it is also clear how networks might be trained to generate wind field forecasts at least several hours into the future. Such a tool might be useful to meteorologists and AQ specialists.

Acknowledgements

I would like to thank Bernard Widrow of the Department of Electrical Engineering at Stanford University for time spent in several discussions on adaptive learning as it pertains to weather forecasting related to the work of his former student Michael Hu. I would also like to thank Dr. John Hopfield, and Ron Benson for helpful suggestions and criticisms in conducting this research and preparing this document. I would also like to thank Rod Brittner of NeuroDynamX Inc. for the unrestricted use of the DynaMind 1.0 simulator for the purpose of conducting research, and Mark Holler of Intel Corporation for discussions on hardware implementation.

I would also like to thank the South Coast Air Quality Management District for access to ozone data, and the helpful suggestions of their senior meteorologist Joe Cassmassi in conducting this research.

References

1. Andes D. (1989), MRIII: A Robust Algorithm for Training Analog Neural Networks, pp. 533-536.
2. Baldi P. and Hornik K. (1989), Neural Networks and Principal Component Analysis: Learning From Examples without Local Minima, Neural Networks **II**, pp. 53-58.
3. Benson R. G. Horrell R. S. and Brittner R. (1989), "Word Recognition from a Stream of Characters," Intel Technical Project Summary.
4. Fang Y. and Sejnowski T. J. (1990), Faster Learning for Dynamic Recurrent Backpropagation, Neural Computation **2**, pp. 270-273.
5. Fozzard R. Bradshaw G. and Ceci L. (1988), A Connectionist Expert System that Actually Works, Neural Information Processing Systems **I**, pp. 248-255.
6. Lapedes A. and Farber R. (1987), How Neural Nets Work, Proceedings of IEEE, Denver Conference on Neural Nets.
7. Hopfield(1982), Neural Networks and Physical Systems with Emergent Collective Computational Abilities, Proceedings of the National Academy of Science **79**, pp. 2554-2558.
8. Holler M. Tam S. Castro H. and Benson R. (1989), An Electrically Trainable Artificial Neural Network (ETANN) with 10240 "Floating Gate" Synapses, IJCNN **II**, pp.191-196.
9. Hu M. J. C. (1964), Application of the Adaline System to Weather Forecasting, Technical Report No. 6775-1.
10. Kamijo K. and Tanigawa T. (1990), Stock Price Pattern Recognition - A Recurrent Neural Network Approach, IJCNN **I.**, pp. 215-221.
11. Pearlmutter B. A. (1989), Learning State Space Trajectories in Recurrent Neural Networks, Neural Computation **1**, pp. 263-269.
12. Pineda F. J. (1989), Recurrent Backpropogation and the Dynamical Approach to Adaptive Neural Computation, Neural Computation **1**, pp 161-172.
13. Robeson S. M. and Steyn D. G. (1990), Evaluation and Comparison of Statistical Forecast Models for Daily Maximum Ozone Concentrations, Atmos. Env. **24b** 2, pp. 303-312.
14. Rumelhart D. E. McClelland J. L. and the PPDP Research Group (1986), Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol 1 Foundations.
15. Smotroff I. G. Howells T. P. and Lehar S. (1990), Meteorological Classification of Satellite Imagery Using Neural Network Data Fusion, IJCNN **II**, pp. 23-28.

16. Widrow B. Winter R. G. and Baxter R. A. (1988), Layered Neural Nets for Pattern Recognition, IEEE Transactions on Acoustics, Speech, and Signal Processing, **36** 7, pp. 1109-1118.

17. Williams R. J. and Zipser D. (1989), A Learning Algorithm for Continually Running Fully Recurrent Neural Networks, Neural Computation **1**, pp. 270-280.

18. Zannetti P. (1989), Can We Continue to Apply Dispersion Models Without a Proper Linkage with Meteorlogical Models?, Air and Waste Management Association Annual Meeting paper 89-43.1.