2.2 TRANSITION PROBABILITIES TO CHARACTERIZE SYNOPTIC SCALE DYNAMICS AFFECTING LOCAL OZONE COMPOSITION

Scott Beaver and Ahmet Palazoglu*
University of California, Davis, California

Saffet Tanrikulu
Bay Area Air Quality Management District, San Francisco, California

1. INTRODUCTION

In a previous study (Beaver and Palazoglu, 2006), a novel form of cluster analysis was applied to hourly, ground-level wind measurements from the San Francisco, CA Bay Area (Figure 1) to identify surface flow patterns affecting local ozone composition. Days exhibiting similar diurnal cycles for the wind field were grouped using the Extended Empirical Orthogonal Functions model (Weare, 1982). The analysis indicated 4 distinct, recurring, mesoscale wind patterns affecting regional ozone buildup processes differently. The result of this cluster analysis is summarized in Figure 2, showing the cluster assignments for each day from 1 June through 30 September of the years 1996 to 2003. Upon forming composite 500 hPa weather maps for each cluster, it was clear that synoptic conditions largely influence mesoscale flow in the Bay Area. Thus, the clusters represent a set of synoptic atmospheric states, and the sequence of daily cluster labels (Figure 2) indicates the evolution of the synoptic meteorology through a progression of states.

The clusters capture either predominately cyclonic or anti-cyclonic influences of the 500 hPa pressure level. Two of the clusters are associated with a trough positioned along the Pacific coastline and encompassing the Bay Area study region. These clusters correspond to ventilated conditions, denoted V1 and V2, in which the marine layer penetrates the Bay Area, channels through a gap in the Coastal Range at the delta of the San Francisco Bay, and flows into the Central Valley. Both cyclonic clusters have relatively low ozone levels, however V2 is a deeper trough, providing increased ventilation and reduced pollutant compositions relative cluster V1. The remaining clusters, H1 and H2, exhibit anti-cyclonic motion at the 500 hPa pressure level, affecting ground level transport and dispersion such that ozone compositions are elevated and highly variable relative to the ventilated clusters. Cluster H2 corresponds to a persistent cell of high pressure forming over the western United States. This cluster exhibits weakened marine layer flow which enters through the mouth of the Bay but does not penetrate through the Bay Area and into the Central Valley, as for the ventilated regimes V1 and V2. Cluster H1 is associated with a passing, offshore ridge of high pressure which causes a northerly shift in wind direction and reduced marine flow entering through the mouth of the Bay.

Breaking the entire ozone season into only 4 synoptic regimes is quite coarse, however the cluster analysis is useful because it reveals the major weather patterns defining the Bay Area's synoptic ozone climatology. Composite spatial fields for upper atmospheric data (500 hPa geopotential height) distinguish synoptic features associated with the clusters. Composite maps of ground level observations indicate the mesoscale flow responses affecting regional pollutant transport and dispersion.

Clearly, however, Bay Area air quality cannot be characterized fully by considering only the presence of a small number of highly generalized meteorological features. One approach by which the synoptic ozone climatology can be more finely resolved is to force the clustering algorithm to identify a larger number of meteorological states. It must be recognized, however, that not all ozone variability can be explained in terms of *static* (single day) meteorological patterns. Additional *dynamic* (multiday) events affecting regional ozone levels exist as well, and such transitions of the synoptic state can be identified by examining the sequence of cluster labels describing the time evolution of the synoptic meteorology. This distinction between static pat-

* *Corresponding author address*: Ahmet Palazoglu, Dept. Chemical Engineering & Materials Science, University of California, One Shields Avenue, Davis, CA 95616. E-mail: anpalazoglu@ucdavis.edu

terns, indicating the presence of a particular synoptic feature on a given day, and dynamic patterns, describing the evolution of such synoptic features over time, is further discussed by Comrie (1992).

In our current study, several statistical tools are introduced for dynamic analysis of the sequence of cluster labels observed in Figure 2. Transitions between the clusters indicate dynamic events that are not captured by the static cluster patterns themselves. We provide a battery of statistical tests to determine which transitions are likely or unlikely to result from a given state, allowing inference of physical mechanisms driving the evolution from one synoptic regime to the next. Energetically favored transitions, i.e. those that occur at a relatively high frequency, indicate the presence of naturally preferred progressions of synoptic states. Improbable sequences of cluster labels suggest types of atmospheric activity not represented in the static cluster patterns—relatively infrequent synoptic states appearing as inconsistencies in the sequence of cluster labels. Such dynamic relations allow the development of a conceptual model explaining historical Bay Area air quality as well as allowing inference of future ozone levels.

2. BINOMIAL TRANSITION PROBABILITIES

First, the *realization* of a cluster is defined. A single realization of a given cluster occurs for each longest possible, continuous set of days bearing the same cluster label, as determined by inspection of the results of Figure 2. The day immediately preceding some realization of cluster r either bears another cluster label $s \neq r$, is unlabeled (as a small fraction of the days cannot be assigned to any cluster with reasonable confidence), or falls outside the study period (in the event that the realization includes the date 1 June for any year). Similarly, the day immediately following the realization of some cluster bears a different (or no) cluster label, unless the realization includes the edge date 30 September.

A *transition* is a dynamic event lasting either 2 or 3 days. Moving forward in time (from left to right on Figure 2), a transition occurs when the cluster label changes from r to s, with no intermediate days bearing a different (or no) label. A 2-day transition occurs when one day is assigned wholly to r and the next day wholly to s, whereas a 3-day transition occurs when there is an intermediate, transitional day doubly assigned to both clusters r and s. The difference between the 2-day and 3-day transitions is highlighted in Figure 3. Regardless of their durations, transitions from cluster r to cluster s occur in some observable proportion, and it is of interest to characterize the relative frequencies of all such possible transitions. To ensure that the transition probabilities from a given cluster sum to unity, all of the days bearing no cluster label are lumped to form a $(k+1)^{\text{th}}$ "cluster" which is appended to the original set of k clusters.

Probabilities for transitions occurring *from* cluster r *to* cluster s, with $r \neq s$, are assumed to follow a binomial distribution, as each transition can be viewed as the outcome of a binary decision: given that a transition occurs from cluster r, the transition is to cluster s with probability p_{rs} and to some other cluster with probability $(1 - p_{rs})$.

The Wald method (Brown et al., 2001) is commonly used to compute estimates $\hat{p}_{rs}$ and confidence limits C_{rs} for the true transition probabilities p_{rs}, where n_{rs} is the number of transitions from state r to s, with diagonal elements n_{rr} taken as 0 because they imply no transition. Also, $N_r = \sum_s n_{rs}$, the total number of transitions occurring from state r to any other state.

$$\hat{p}_{rs} = \frac{n_{rs}}{N_r} \tag{1}$$

$$C_{rs} = z_{\alpha/2}\sqrt{\frac{\hat{p}_{rs}(1-\hat{p}_{rs})}{N_r}} \tag{2}$$

This definition of the transition probability estimates ensures that $\sum_s \hat{p}_{rs} = 1$, however no such claim can be made for $\sum_r \hat{p}_{rs}$ because a different N_r is used to normalize each row; when viewed as a matrix, one should only consider the rows of $\hat{p}_{rs}$ but not the columns. The Wald statistics are flawed because of the possibility of producing $C_{rs} > \hat{p}_{rs}$, indicating a negative lower confidence limit for a probability bounded on [0, 1] by definition. This problem becomes common for small $\hat{p}_{rs}$ and/or small sample size N_r.

The Wilson statistics (Brown et al., 2001) provide an alternative approach to estimating p_{rs}. Estimate $\hat{p}^W_{rs}$ and confidence limit C^W_{rs} are intended for use with smaller sample size N_r and are not as prone to producing unrealistic confidence bounds as the Wald statistics.

$$\hat{p}^W_{rs} = \frac{n_{rs} + \frac{z_{\alpha/2}}{2}}{N_r + z^2_{\alpha/2}} \tag{3}$$

$$C^W_{rs} = \frac{z_{\alpha/2}\sqrt{N_r}}{N_r + z^2_{\alpha/2}}\sqrt{\hat{p}^W_{rs}(1-\hat{p}^W_{rs}) + \frac{z^2_{\alpha/2}}{4N_r}} \tag{4}$$

One disadvantage of the Wilson statistics is that $\sum_s \hat{p}^W_{rs}$ is not guaranteed to equal unity.

Hypothesis testing is used to determine transitions that are either *favored* or *disfavored*— those that occur more frequently or less frequently than would be expected by chance, respectively. As the null hypothesis, it is assumed that transitions occur independently

of the originating state r. Therefore, the null hypotheses transition probability p^0_{rs} should be proportional to the total number of realizations of cluster s. Because of the forward nature of the transition statistics, the number of state realizations for each cluster cannot be computed from n_{rs} but must be counted to avoid edge effects when computing the value of p^0_{rs}.

$$p^0_{rs} = \frac{\#\text{ realizations of cluster } s}{\#\text{ realizations of all clusters } r \neq s} \quad (5)$$

The confidence intervals for p_{rs} calculated using both the Wald and Wilson statistics are compared to p^0_{rs} to determine if the null hypothesis is rejected, indicating that the number of transitions n_{rs} is significantly different than would be expected by chance. Pairs of clusters for which p^0_{rs} lies outside of $\hat{p}_{rs} \pm C_{rs}$ or $\hat{p}^W_{rs} \pm C^W_{rs}$ experience forward transitions that are statistically significant, as identified using the Wald or Wilson statistics, respectively. Pairs of clusters for which p^0_{rs} is below the lower confidence bound for either confidence interval are favored, or occur more frequently than by chance, while pairs of clusters for which p^0_{rs} exceeds either of the upper confidence bounds indicate disfavored transitions. The Wilson statistics tend to be more conservative than the Wald statistics, less often indicating trends of borderline significance.

Assuming that the null hypothesis is true, that state transition probabilities are given by p^0_{rs}, the likelihood of observing n_{rs} can be computed. $P(n_{rs}|p^0_{rs}, N_r)$ is the binomial probability of observing n_{rs} given N_r total transitions from state r and that the null hypothesis is true.

$$P(n_{rs}) = \binom{N_r}{n_{rs}} (p^0_{rs})^{n_{rs}} (1 - p^0_{rs})^{N_r - n_{rs}} \quad (6)$$

$$\binom{N_r}{n_{rs}} = \frac{N_r!}{n_{rs}!(N_r - n_{rs})!} \quad (7)$$

Relatively small values for $P(n_{rs}|p^0_{rs}, N_r)$ indicate that null hypothesis is not likely, and that the transition from cluster r to s occurs more or less frequently than would be expected by chance. This third measure is used in conjunction with the formal hypothesis testing using the Wald and Wilson statistics to form a battery of statistical methods which can determine cluster transitions that are favored or disfavored.

3. APPLICATION OF METHODS

The sequence of daily cluster labels shown in Figure 2 is first examined to determine all realizations for each cluster. There are 21, 65, 57, and 66 realizations appearing as multiday runs for clusters H1, H2, V1, and V2 respectively. Additionally, 22 days cannot be assigned to any cluster with reasonable confidence and remain unlabeled. This set of unlabeled days is considered cluster U for the transition probability calculations— there are 16 lone days and 3 pairs of consecutive days, for a total of 19 realizations. Thus, the original sequence of cluster labels for 976 days (122 days for 8 years) is represented by 228 individual cluster realizations.

Using the above numbers of state realizations, null hypothesis transition probabilities are calculated. The results are shown in Table 1. For example, there are $57+66+21+19 = 163$ realizations of states other than H2. Therefore, if the null hypothesis is true, we should expect a fraction 57/163 (or 0.35) of the transitions from state H2 to occur to V1.

Based on the observed number of forward transitions between each pair of states (Table 2), transition probabilities and corresponding confidence bounds are estimated using $\alpha = 0.05$. Confidence intervals using Wald statistics ($\hat{p}_{rs} \pm C_{rs}$) and Wilson statistics ($\hat{p}^W_{rs} \pm C^W_{rs}$) are given in Tables 3 and 4, respectively. Statistically significant transitions are boldfaced.

Both the Wald and Wilson statistics are in agreement that transitions H1→H2 and H2→V2 are favored, whereas H1→V1 and H1→V2 are disfavored. The Wald statistics additionally indicate that transitions H2→H1, V1→U, and V2→H1 are disfavored, whereas the Wilson statistics do not. Note that the Wald statistics indicate a negative lower confidence bound for the V1→U transition which is indicated as disfavored.

Likelihoods of observing the proportions of state transitions given the null hypothesis is true are calculated using Equation 6. Results in Table 5 generally corroborate the Wald and Wilson statistics. Relative frequencies for transitions H1→H2, H2→V2, and H1→V1 are likely to occur by chance at lower than the 0.001 level, while the H1→V2 transition likelihood is near the 0.01 level— observed proportions for these state transitions are highly unlikely to occur by chance, in agreement with both statistics. Additionally, H2→H1 is likely at the 0.02 level, suggesting this transition is disfavored as suggested by the Wald but not Wilson statistics. The transitions V1→U and V2→H1, which are indicated as disfavored by the Wald statistics only, are moderately likely (0.03 and 0.04, respectively), suggesting that these transitions may be randomly driven and are neither favored nor disfavored.

4. DISCUSSION

4.1. H1→H2 TRANSITION

The transition from H1 to H2 is heavily favored, indicating some physical mechanism driving this progression of states. While this transition occurs only 14 times in

the 8-yr study period, it has important implications for Bay Area air quality. Days near the H1→H2 transition points have some of the most severe ozone levels in the study period and account for over one third of the 20 multiday exceedances.

Each occurrence of this transition is of the 3-day type, with an intermediate day assigned to both clusters H1 and H2. Ozone levels at most locations are usually higher on these doubly assigned, transitional days than their immediately preceding (assigned solely to H1) and following (assigned solely to H2) days. This effect is illustrated in Figure 4, showing the changes in daily maximum ozone level at each monitoring station between pairs of consecutive days during the 3-day H1→H2 transition periods. Between the last day assigned to H1 and the transitional day between H1 and H2, ozone levels increase significantly at nearly all monitoring stations, especially at the monitors typically downwind of the urban source areas. On the next day, ozone levels tend to decrease at most stations (but may still remain above the exceedance threshold) as the synoptic state transitions into conditions typical of H2— a cell of high pressure forming over northern California and other western states. Note, however, that ozone levels at Livermore decrease less than for the other stations (or sometimes even increase) as H2 is fully realized and produces blocking anti-cyclonic flow in the Central Valley which diverts a polluted airmass to the Livermore Valley. Los Gatos also is likely to have increasing ozone levels into the realization of H2 due to carryover effects as transport from the South Bay into the Santa Clara Valley weakens.

In general, the H1→H2 transition can be described as producing relatively high ozone levels, typically lasting for several days but peaking on the transitional day, with the location of daily maximum ozone level shifting from the Santa Clara Valley to the Livermore Valley. Not all of the 14 H1→H2 transitions follow this generalization, however, and a notable outlier is the transition occurring on 6 September 1996. Between 5 September (assigned solely to H1) and 6 September (doubly assigned), ozone levels increase as expected, with especially large increases in the Santa Clara Valley. Between 6 September and 7 September (assigned solely to H2), ozone levels exhibit a substantial increase at all stations (shown in bottom of Figure 4), instead of decreasing as would be expected. Despite these large increases, however, ozone levels remain just below the exceedance threshold. Ozone levels peak on the next day, with exceedances occurring at Livermore and Los Gatos on 8 September. The atypical synoptic evolution associated with this outlier H1→H2 transition will be discussed later in this section.

The 14 occurrences of the H1→H2 transition share similar mesoscale air flows, typically associated with a shift in the surface wind direction from northerly to westerly over the 3-day transition period. Realizations of cluster H1 are typically preceded by a high pressure cell forming over the Pacific Ocean, far west of the continent, during conditions in which a trough is present along the Pacific coastline (i.e. one of the ventilated clusters V1 or V2). The offshore high pressure center migrates toward the continent, displacing the trough inland and possibly pinching off a cell of low pressure over the western United States— H1 is realized when the Bay Area is caught between these cells and experiences northerly upper atmospheric flow. Cluster H2 represents conditions with weak, westerly marine flow into the Bay Area. The transitional days experience ground level wind directions intermediate that of H1 and H2 and low wind speeds.

This shift in wind direction is evidenced in Figure 5, showing time averaged wind fields for the hours 1200–1600 LT (local time) averaged among the last days fully assigned to H1 (Figure 5a), the doubly assigned transitional days (Figure 5b), and the following days fully assigned to H2 (Figure 5c). The precursor response to the shift in the flow direction is illustrated in Figure 6, showing hourly time series for NO level at 2 South Bay stations. Increased overnight carryover of NO is observed leading into the transitional day at the South Bay monitors in Fremont and San Jose; severe ozone levels are observed in the downwind Santa Clara Valley during the afternoon hours of the following day. While H1 has elevated ozone levels relative to the ventilated states V1 or V2, ozone levels do not increase beyond the exceedance threshold until near the transition to cluster H2. Thus, days assigned to H1 but preceding the transition to H2 generally have moderate air quality, however the northerly shift in Bay Area winds gives warning that a prolonged period of relatively high ozone levels is likely to ensue within several days.

The evolution of the synoptic meteorology associated with the H1→H2 transition is demonstrated in Figure 7 for the transition occurring on 9 August 2002. (Weather maps are obtained from NCEP Reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, CO, USA web site at http://www.cdc.noaa.gov/.) On 5 August, cluster V1 is realized and a trough is present over the Pacific coast. A large pocket of high pressure is present north of the Hawaiian Islands, however at this point the trough buffers the Bay Area from its effects. Cluster H1 is first realized on 6 August (not shown), when it has migrated sufficiently close to the Pacific coast to affect mesoscale conditions in the Bay Area. By 8 August the trough has been displaced inland and a low pressure cell has been pinched off— both the offshore high pressure and onshore low pressure contribute to northerly upper atmospheric flow over the Bay Area

associated with H1. The transition to H2 occurs on 9 August (not shown), and by 10 August a typical H2-type high pressure center over the southwestern United States is present, in addition to residual offshore high pressure. The wide, east-west band of high pressure positioned over the Bay Area is typical of H1→H2 conditions resulting in multiday exceedance periods. By the next day, the offshore high pressure cell has dissipated and onshore high pressure (i.e. cluster H2) prevails.

There is of course significant variability among the 14 H1→H2 transitions occurring in the study period, however the recurring pattern of offshore high pressure displacing a coastal low pressure system holds for all cases. The H1→H2 transitions resulting in multiday exceedances exhibit a wide, east-west band of high pressure linking the onshore and offshore high pressure centers (as with 10 August 2002, shown in Figure 7). We note that for the cases not resulting in exceedances, the offshore high pressure cell dissipates rapidly and is no longer present by the time that the onshore high pressure center has developed.

A slightly different trajectory than the above example is realized for the previously discussed outlier H1→H2 transition occurring on 6 September 1996, as depicted in Figure 8. On 3 September, the day before H1 is realized, no cluster label is assigned— a low pressure cell exists along the coast well north of the Bay Area. Instead of being displaced eastward, as typical of H1→H2 transitions, the low pressure is instead displaced to the north by tropical high pressure expanding northward, as observed on 6 September, the transitional day. By 8 September the tropical high pressure has formed a slight ridge along the Pacific coast, triggering an exceedance in Los Gatos in addition to the usual Santa Clara Valley and Livermore Valley locations. The H1→H2 transition occurring on 21 September 2003 is similar to the 6 September 1996 outlier— the synoptic evolution involves the northward displacement of a coastal trough, in this case by high pressure originating over Mexico, resulting in exceedances in Los Gatos in addition to Livermore and San Martin. These 2 outlier H1→H2 transitions are distinguished by the fact that H1 is not preceded by a ventilated state— for these 2 outliers, H1 is immediately followed by an unlabeled day which is itself preceded by anti-cyclonic conditions. Thus, these unusual sequences of cluster labels indicate modified H1→H2 transitions in which a different type of synoptic activity is triggering exceedances at locations including Los Gatos.

4.2. H1→V1/V2 TRANSITIONS

Because the H1→H2 transition is so heavily favored, it is logical that the transition probability statistics indicate H1→V1 and H1→V2 as being disfavored. Cluster H1 is realized as offshore high pressure migrates toward California, producing northerly flow which allows temperature and pressure to build over land due to the lack of westerly marine layer intrusion. Thus, H1-type weather systems produce conditions conducive to the formation of H2-type patterns, and it is uncommon for a low pressure system (i.e. V1 or V2) of sufficient strength to appear during the short duration in which H1 is present to displace the offshore anti-cyclonic system and prevent H2 from being realized.

Nonetheless, there are 3 transitions occurring from H1 to a ventilated state during the study period (3 July 1997, 20 September 2000, and 11 June 2002). In each case, a deep polar low expands southward, preventing the H1-type offshore high pressure cell from taking its usual trajectory toward the continent, and it is instead forced to migrate north. Figure 9 provides an example using the 20 September 2000 occurrence. On 19 September (assigned to H1), a typical H1-type offshore high pressure system dominates Bay Area conditions. A deep polar low is present over the Great Lakes in Canada but does not yet affect the Bay Area. By September 20 (assigned to V2), the polar low has moved south, preventing the usual shoreward trajectory of the offshore high. On September 21 (assigned to V2), the offshore high has been forced north toward the Gulf of Alaska as the polar low expands to the south and west. For the other 2 cases (3 July 1997 and 11 June 2002), the H1 offshore highs are displaced by lows originating from the Gulf of Alaska and moving south along the Pacific coast.

4.3. TRANSITIONS FROM H2

The statistics also suggest that the H2→H1 transition may be disfavored, though they are in disagreement for this transition. Taking the conceptual model of H2 being onshore high pressure and H1 being offshore high pressure, the H2→H1 transition violates the natural west to east progression of Pacific weather systems— this sequence would seem to imply a high pressure cell forming over land and moving offshore in an anomalous westward trajectory. The H2→H1 transition in fact occurs 3 times (26 September 1997, 25 September 1999, and 12 September 2000), however the above conceptual model which is essentially the reverse of the H1→H2 transition breaks down. Instead, on each occurrence of the H2→H1 transition an offshore low pressure cell is present to the south of the study domain, as shown in Figure 10 for the days preceding the transitions. Thus, the H2→H1 sequence of cluster labels actually captures a new type of weather pattern not accounted for in the set of 4 static cluster patterns describing various ridge

and trough conditions which dominate Bay Area summers.

A final significant transition is the favored H2→V2 sequence of cluster labels. Because the H2→H1 transition is rare, the anti-cyclonic H2 is generally observed to transition into one of the ventilated states, V1 or V2. Cluster V2 represents a deeper trough pattern than V1, suggesting that a low pressure system of sufficient strength is required to displace the H2-type high pressure cell once it forms over California's Central Valley. Large decreases in ozone levels at all inland regions occur immediately upon the transition from H2 to V2. This transition often provides overnight relief from intensifying episodes of poor air quality, such as the 3-day episode occurring from 10–12 July 1999, in which regional, daily maximum 8-hr ozone levels (obtained at Livermore and/or Concord, with similar levels at each monitor) reach 92, 116, and 122 ppb, respectively, for these 3 days assigned to H2. The next day, 13 July, is fully assigned to V2 and maximum regional ozone drops sharply to 59 ppb as a polar airmass pushes south to form a coastal trough.

4.4. TRANSITIONS FROM V1/V2

The Wald statistics alone indicate the V2→H1 transition as disfavored, however the other tests in the statistical battery fail to confirm this and it is thus noted that no significant transitions are realized from either clusters V1 or V2. Thus, these states favor the development of no other states, and little predictive capability may be possible from either of these ventilated regimes. Given one of the ventilated states is present, the trough will remain until it is displaced by high pressure introduced by developing synoptic conditions in proximal regions of the globe.

5. CONCLUSIONS

Aside from the elucidation of several mechanisms affecting Bay Area ozone levels, the transition probabilities allow certain inferences about the predictability of future ozone levels given the current synoptic state. An important example is the northerly shift in winds caused by offshore anti-cyclonic motion. Though the resulting ground level dispersion pattern rarely allows ozone levels to increase to the NAAQS exceedance level, the northerly shift provides imminent warning that a multi-day episode of severe ozone levels will likely ensue within several days unless a deep polar low pushes south to displace the offshore high. The analysis also reveals that onshore high pressure systems conducive to ozone buildup remain intact until a low pressure system of sufficient strength arrives to provide relief. The trough patterns, on the other hand, seem to offer little predictive capability. No significant transitions occur from either ventilated state, and such transitions are essentially random, being driven by developing conditions elsewhere in proximal regions of the globe.

Characterization of the dynamics implied by a sequence of daily cluster labels is useful for developing a conceptual model of synoptic ozone climatology. The transition probabilities identify several atmospheric events affecting regional ozone levels, however, they do not form a complete statistical characterization of the time evolution of Bay Area meteorology. In our future work, the concepts of persistency and intermittency will be used to compliment the transition probabilities in further describing synoptic evolution as represented by a sequence of daily cluster labels.

6. Acknowledgements

Funding provided by the California Air Resources Board and the Central California Ozone Study is gratefully acknowledged.

REFERENCES

Beaver, S. and Palazoglu, A., 2006: Cluster Analysis of Hourly Wind Measurements to Reveal Synoptic Regimes Affecting Air Quality. *J. Appl. Meteor. Climatol.*, in press.

Brown, L. D., Cai, T. T., and DasGupta, A., 2001: Interval Estimation for a Binomial Population. *Statistical Science*, **16**, 101–117.

Comrie, A. C., 1992: An Enhanced Synoptic Climatology of Ozone Using a Sequencing Technique. *Physical Geography*, **13**, 53–65.

Lorenz, E. N., 1956: Empirical Orthogonal Functions and Statistical Weather Prediction. Scientific Report 1, Statistical Forecasting Project. Massachusetts Institute of Technology Defence Doc. Center No. 110268, 49 pp.

Weare, B. C. and Nasstrom, J. S., 1982: Examples of Extended Empirical Orthogonal Function Analyses. *Mon. Wea. Rev.*, **110**, 481–485.

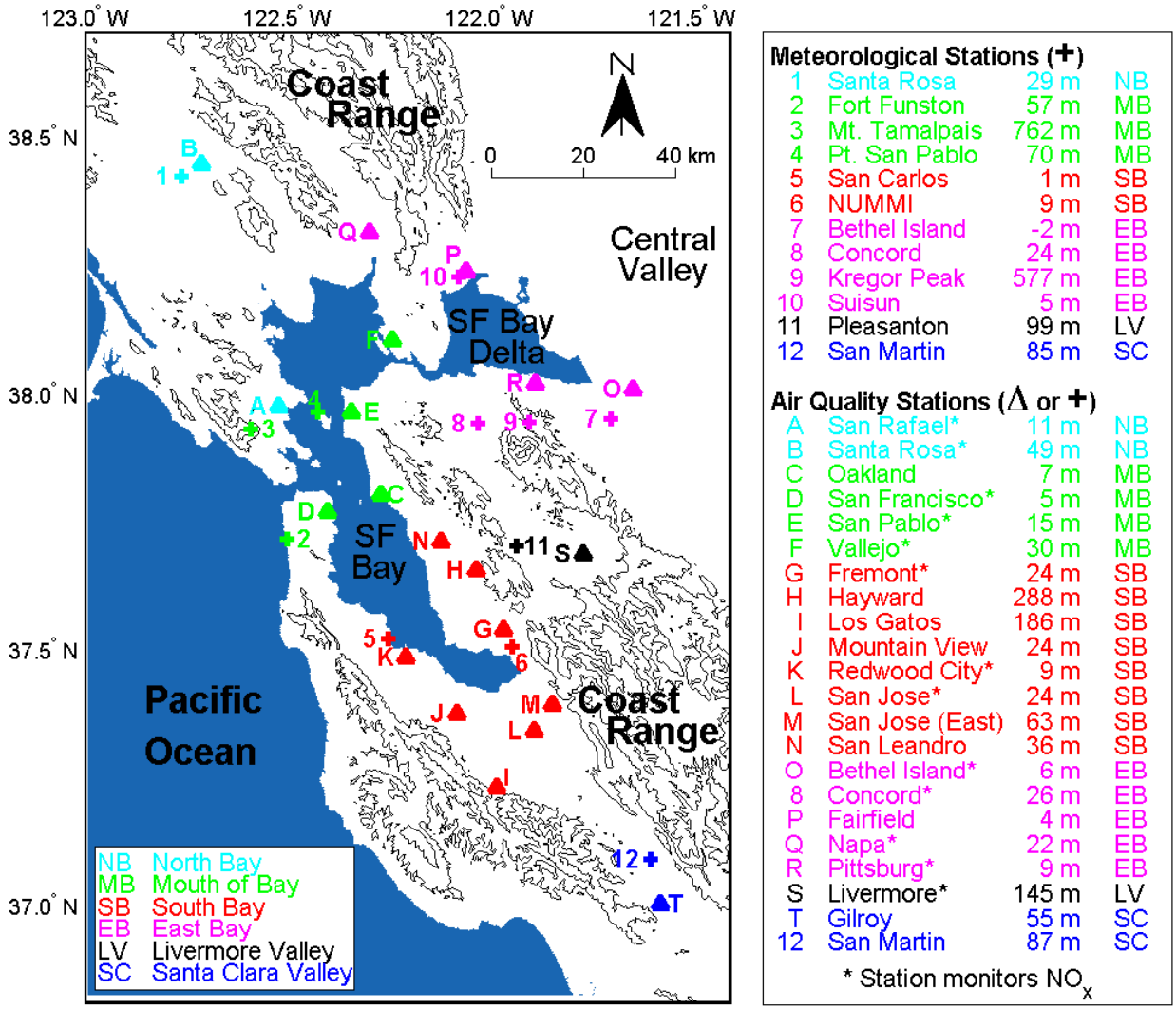

Figure 1: Study region with locations of meteorological and air quality monitors. Contour lines are at 300, 600, and 900 m. Names, elevations, and subregion classifications for the stations are given in the legend. Note that the air quality monitors at Concord and San Martin are located in close proximity to meteorological monitors. These 2 stations are not shown explicitly but should be referenced under the meteorological stations bearing the same name.

Table 1: Null hypothesis transition probabilities p^0_{rs}.

	To Cluster V1	To Cluster V2	To Cluster H1	To Cluster H2	To Unlabeled
From Cluster V1	–	.39	.12	.38	.11
From Cluster V2	.35	–	.13	.40	.12
From Cluster H1	.28	.32	–	.31	.09
From Cluster H2	.35	.41	.13	–	.12

Table 2: Number of forward transitions between each pair of states.

	To Cluster V1	To Cluster V2	To Cluster H1	To Cluster H2	To Unlabeled
From Cluster V1	–	22	11	22	2
From Cluster V2	29	–	4	26	7
From Cluster H1	1	2	–	14	4
From Cluster H2	19	39	3	–	4

Table 3: Estimates and $\alpha = 0.05$ confidence limits (in parentheses) for transition probabilities using Wald statistics. Statistically significant transitions, as determined by comparison with Table 1, are shown in boldface.

	To Cluster V1	To Cluster V2	To Cluster H1	To Cluster H2	To Unlabeled
From Cluster V1	–	.39 (.13)	.19 (.10)	.39 (.13)	**.04 (.05)**
From Cluster V2	.44 (.12)	–	**.06 (.06)**	.39 (.12)	.11 (.07)
From Cluster H1	**.05 (.09)**	**.10 (.13)**	–	**.67 (.20)**	.19 (.17)
From Cluster H2	.29 (.11)	**.60 (.12)**	**.05 (.05)**	–	.06 (.06)

Table 4: Estimates and $\alpha = 0.05$ confidence limits (in parentheses) for transition probabilities using Wilson statistics. Statistically significant transitions, as determined by comparison with Table 1, are shown in boldface.

	To Cluster V1	To Cluster V2	To Cluster H1	To Cluster H2	To Unlabeled
From Cluster V1	–	.39 (.12)	.21 (.10)	.39 (.12)	.06 (.07)
From Cluster V2	.44 (.12)	–	.09 (.07)	.40 (.12)	.13 (.08)
From Cluster H1	**.12 (.14)**	**.16 (.15)**	–	**.64 (.19)**	.24 (.17)
From Cluster H2	.30 (.11)	**.60 (.12)**	.07 (.07)	–	.09 (.07)

Table 5: Likelihood of observing proportions of state transitions given the null hypothesis transition probabilities of Table 1 are true.

	To Cluster V1	To Cluster V2	To Cluster H1	To Cluster H2	To Unlabeled
From Cluster V1	–	.11	.04	.11	.03
From Cluster V2	.03	–	.04	.10	.15
From Cluster H1	**.001**	**.01**	–	**.001**	.08
From Cluster H2	.07	**.001**	.02	–	.07

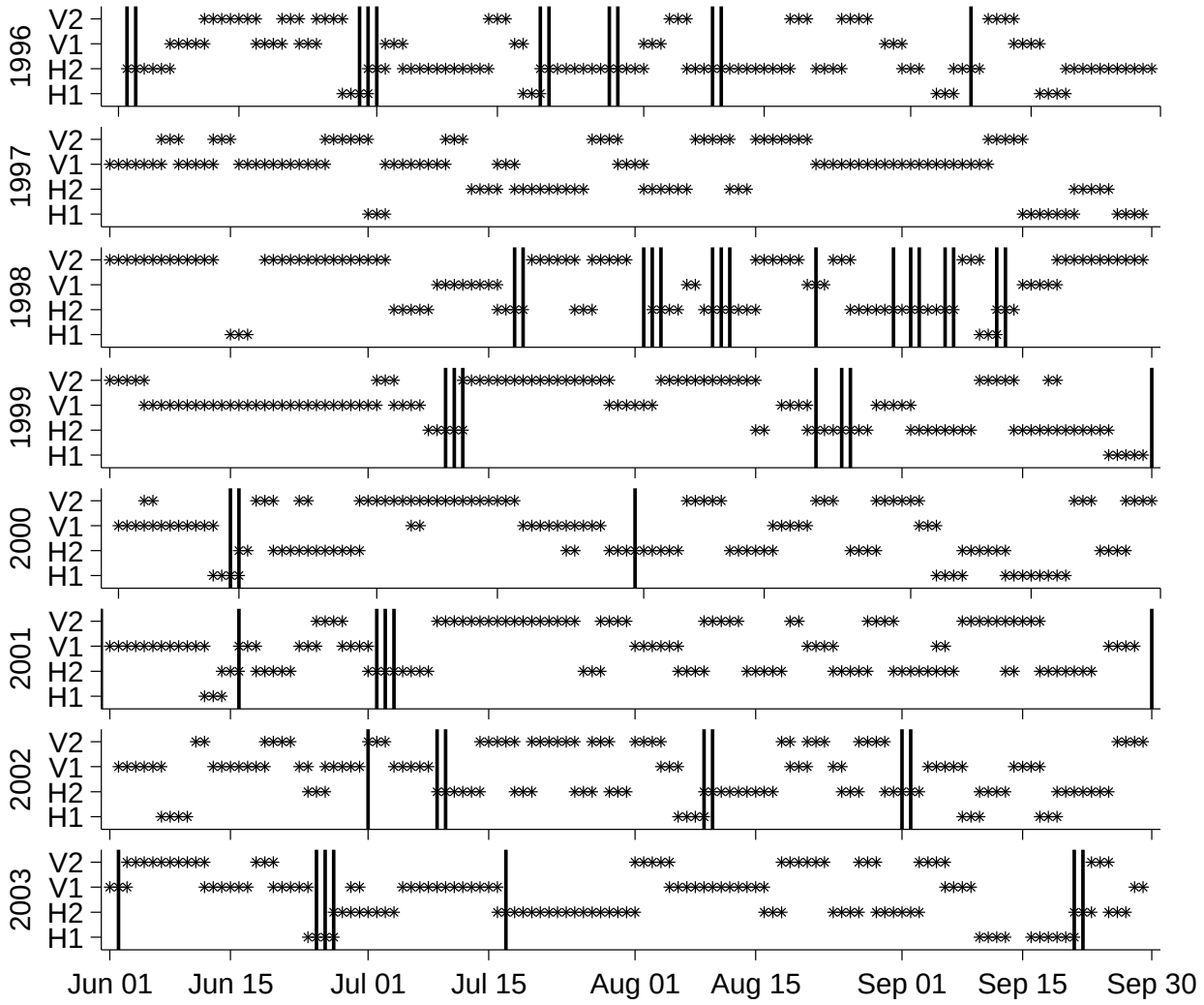

Figure 2: Cluster membership for 1 June through 30 September 1996–2003 observation period. Y-axis position of asterisk indicates cluster membership for each day. Vertical lines indicate days exceeding the NAAQS for 8-hr ozone in the Bay Area.

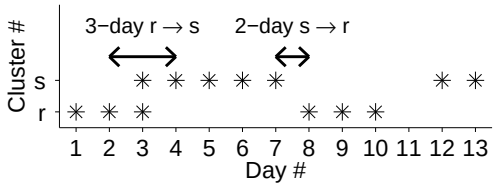

Figure 3: Diagram depicting 2-day and 3-day transitions for a hypothetical 2-cluster solution on an observation period containing 13 days. Note that day #11 is unlabeled and no transition from r to s occurs between days #10 and #12.

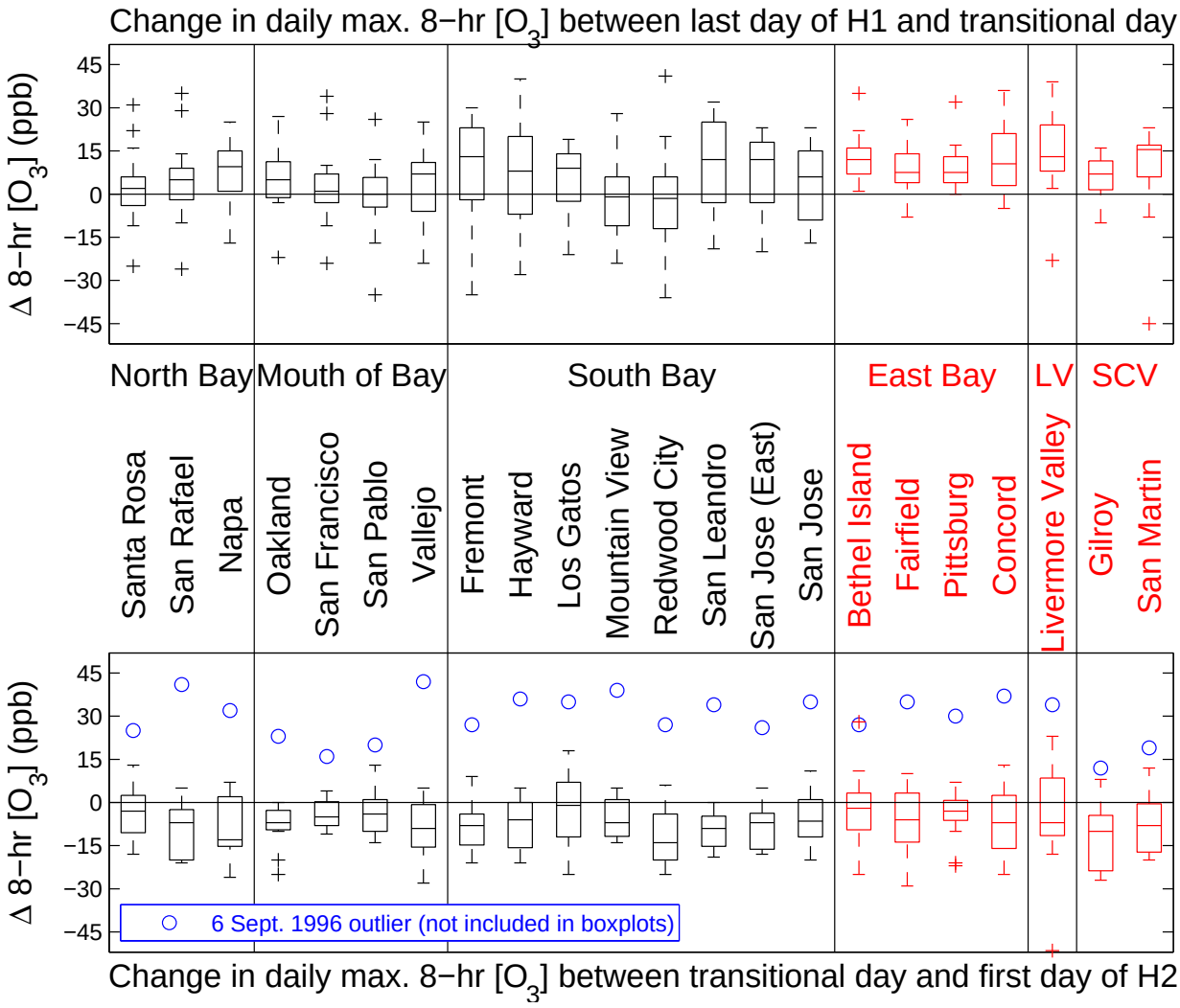

Figure 4: Boxplots showing change in daily maximum ozone composition at 22 monitoring locations. Top shows change in ozone level between last day of H1 and the following transitional day into H2, while the bottom plot shows the change between the transitional day and the next day assigned to H2. Horizontal lines on boxes indicate lower quartile, median, and upper quartile, while the whiskers contain the remaining data except for several extreme values plotted individually using plus signs or circles. Sites shown in red are downwind of the major urban sources and typically experience the highest ozone levels.

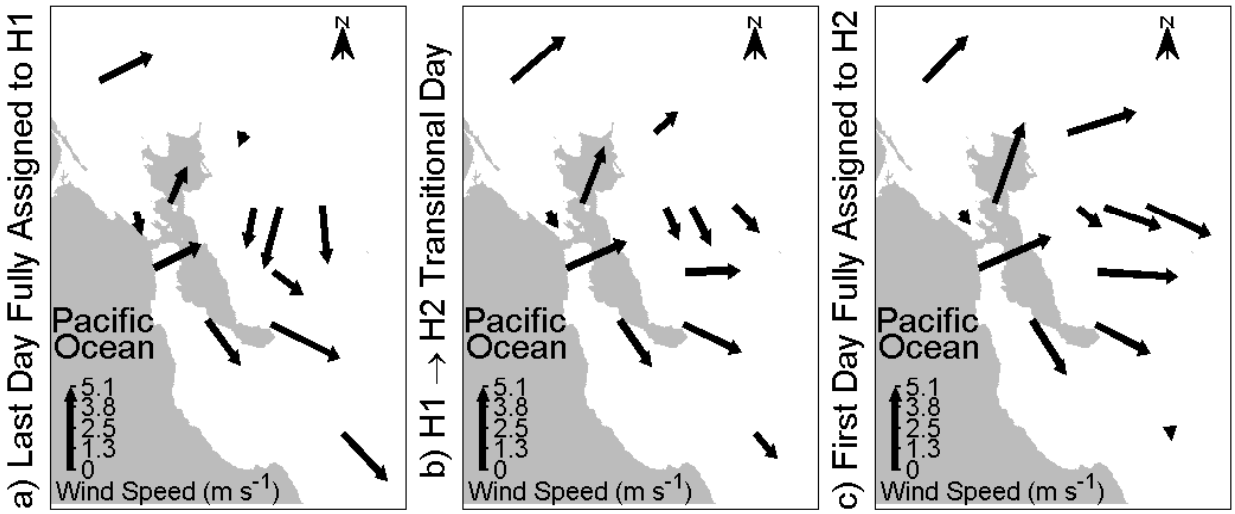

Figure 5: 1200–1600 LT time averaged wind field averaged among groups of 14 days from the H1→H2 transition: a) last days fully assigned to H1, b) doubly assigned transitional days, c) first days fully assigned to H2.

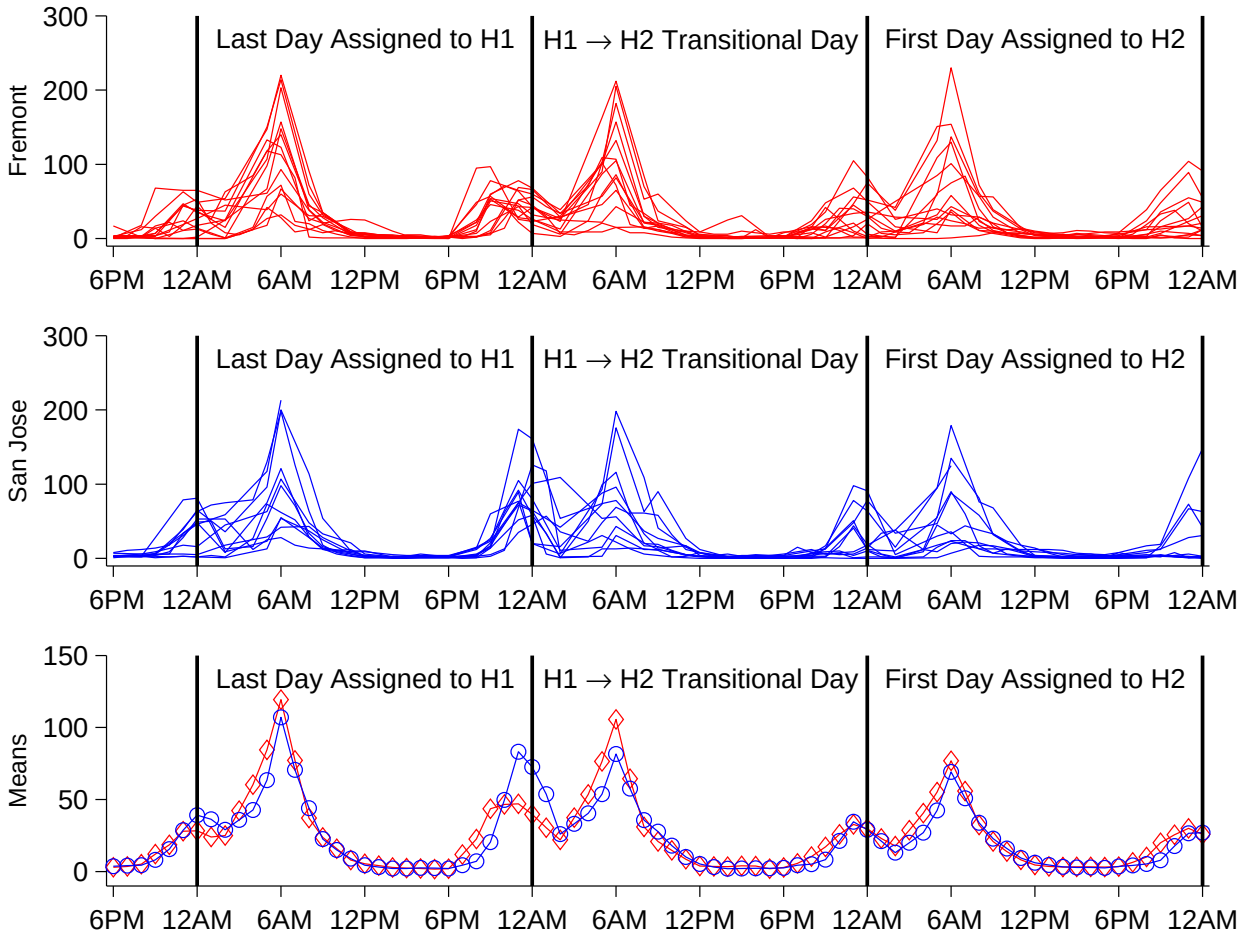

Figure 6: Top two images show hourly time series for [NO] (ppb) at Fremont and San Jose, respectively, for 14 H1→H2 3-day transition periods. Bottom image shows mean trajectories for [NO] at Fremont (red diamonds) and San Jose (blue circles).

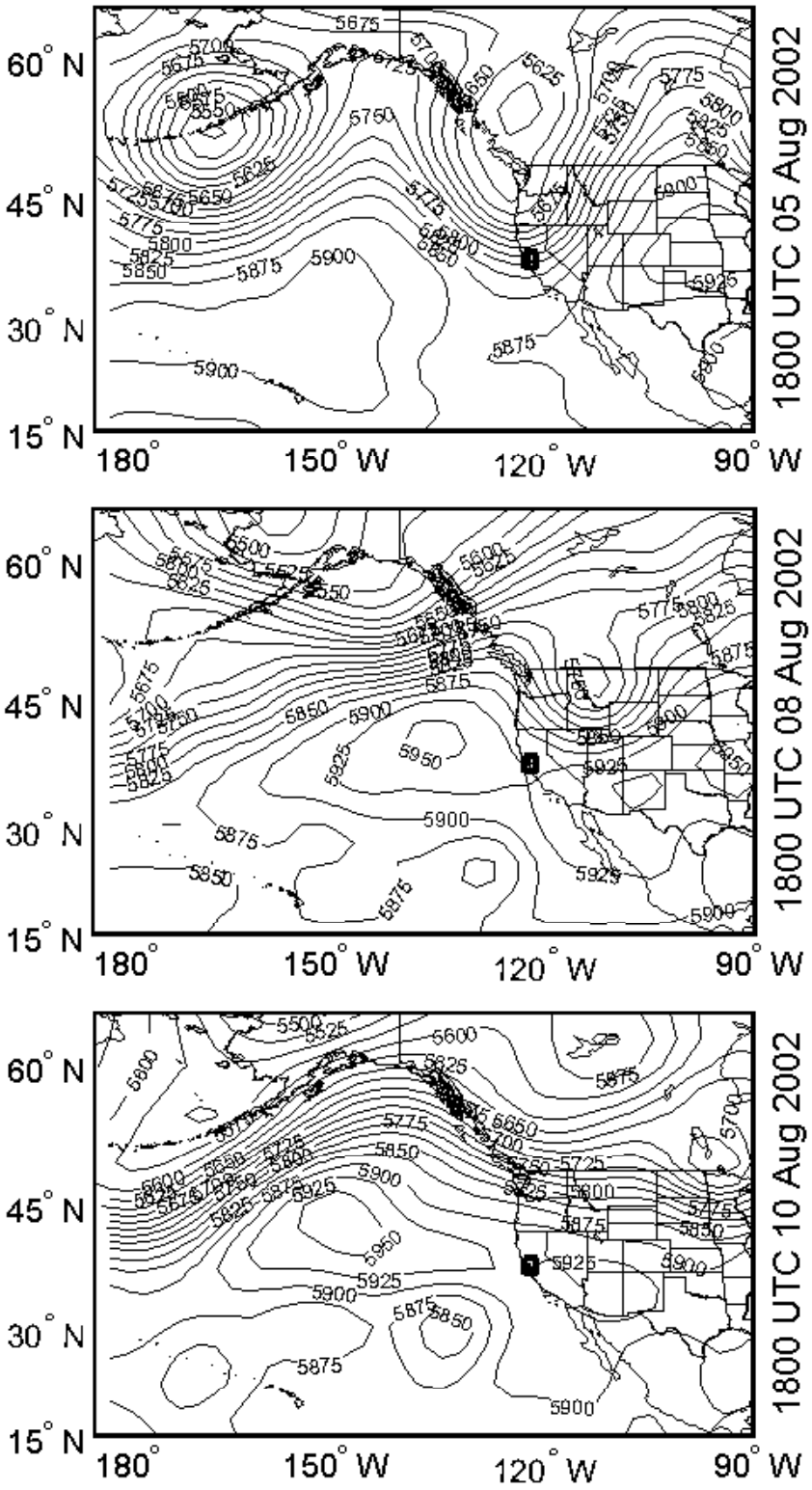

Figure 7: 500 hPa weather maps depicting a typical H1→H2 transition occurring on 9 August, 2002. Weather maps are shown for 5, 8, and 10 August at 1800 UTC, from top to bottom. The Bay Area study region of Figure 1 is highlighted.

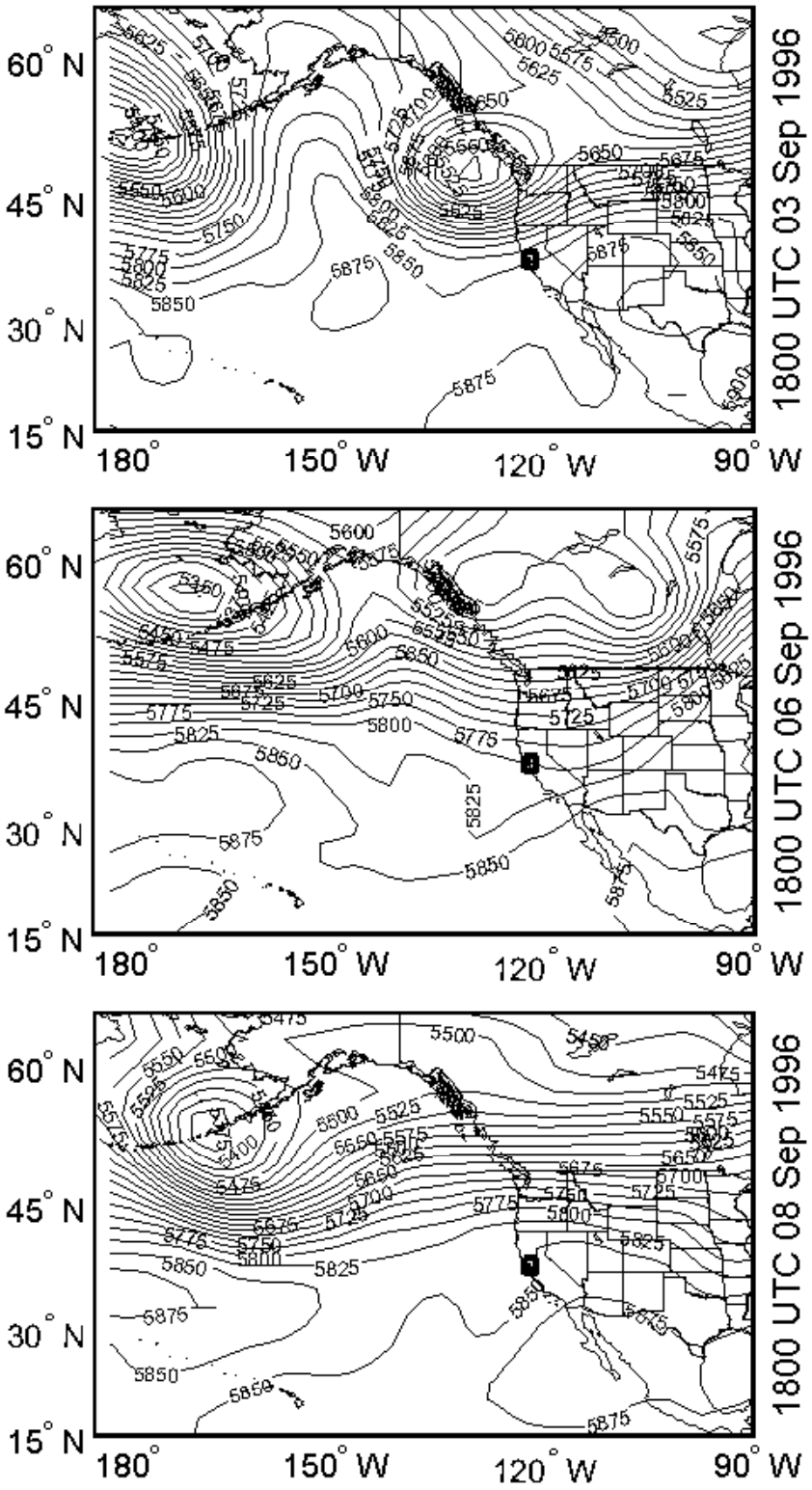

Figure 8: 500 hPa weather maps depicting an atypical H1→H2 transition occurring on 6 September, 1996. Weather maps are shown for 3, 6, and 8 September at 1800 UTC, from top to bottom. The Bay Area study region of Figure 1 is highlighted.

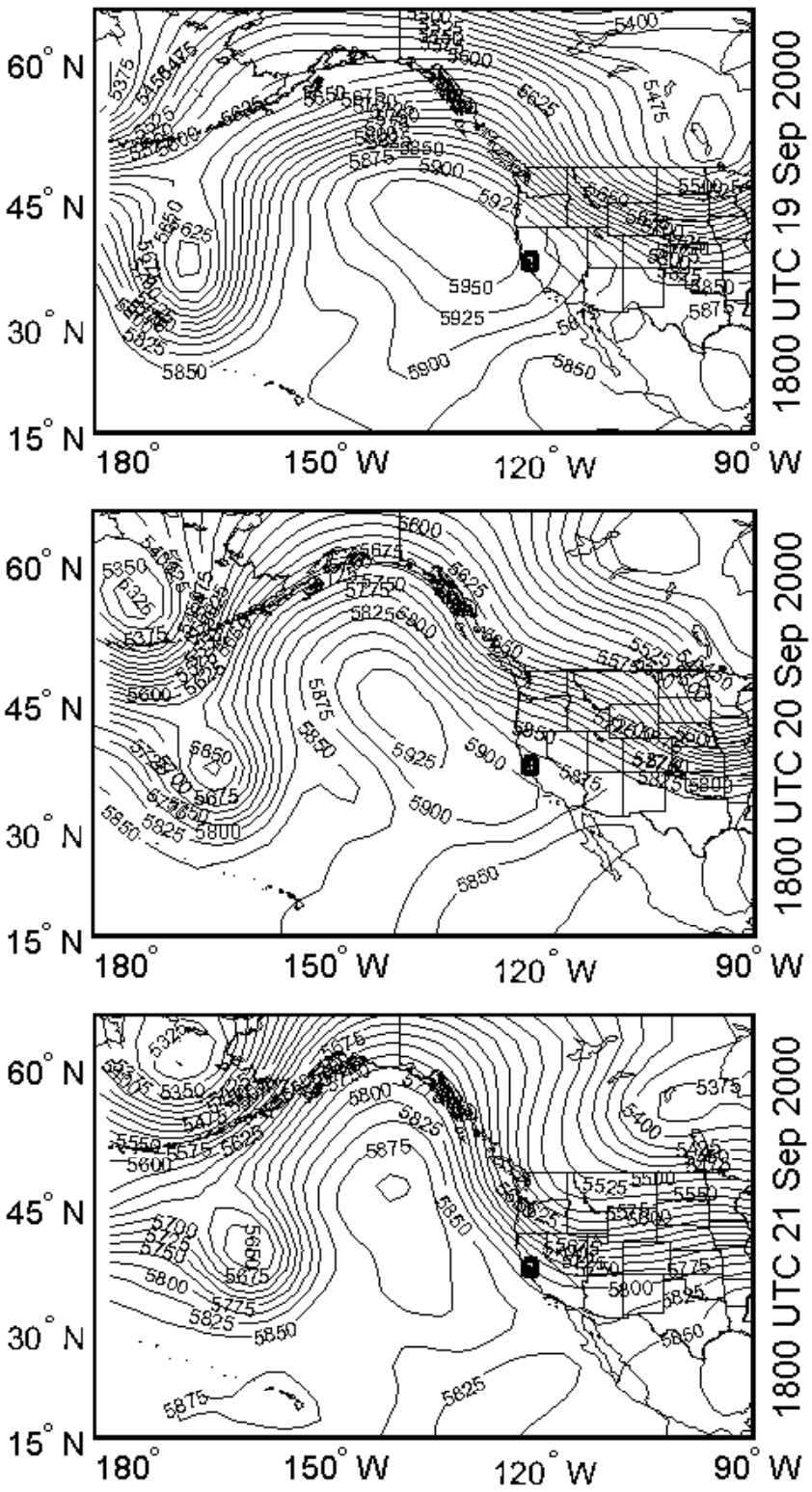

Figure 9: 500 hPa weather maps depicting a typical transition from H1 to a ventilated state (V1 or V2) occurring on 20 September, 2000. Weather maps are shown for 19, 20, and 21 September at 1800 UTC, from top to bottom. The Bay Area study region of Figure 1 is highlighted.

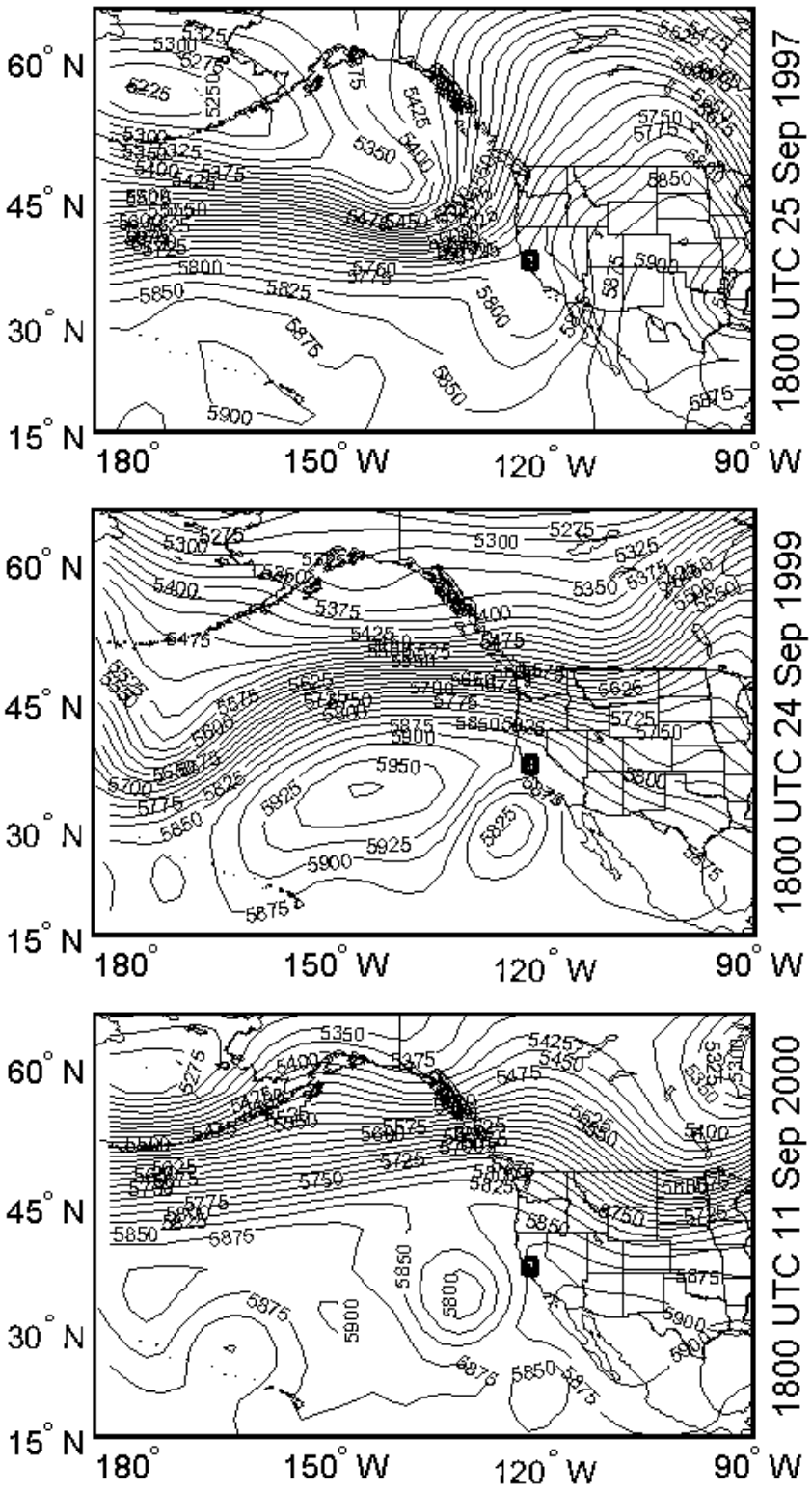

Figure 10: 500 hPa weather maps for days preceding the 3 occurrences of the H2→H1 transition. The Bay Area study region of Figure 1 is highlighted.