



Western States Petroleum Association
Credible Solutions • Responsive Service • Since 1907

September 1, 2004

Via email to ddrechsl@arb.ca.gov

Dr. Deborah Drechsler, Ph.D.
Research Division, P.O. Box 2815,
California Air Resources Board,
Sacramento, CA 95814

RE: Proposed New California Ozone Standard

Dear Dr. Drechsler:

The Western States Petroleum Association (WSPA) and the American Petroleum Association (API) represent companies that explore, develop, refine, market and distribute petroleum and petroleum products throughout the United States. WSPA is composed of nearly 30 companies that have operations within the 6 Western States including California. API is the national trade association of America's oil and natural gas industry, and represents more than 400 members involved in all aspects of the oil and natural gas industry. Both WSPA and API members are owners and operators of major facilities regulated under the California ozone standard, and producers and marketers of fuels that are often targeted as a means to reduce ozone precursors. As such, we have a direct and substantial stake on the outcome of this proposal. With this letter, we are providing comments on the proposal to adopt a new 8-hour ozone standard for the state of California.

WSPA and API have credible technical and scientific experience

WSPA has been involved in air quality technical and policy issues relating to attainment of federal and state ozone standards since the 1970s. Over this 30 year period, WSPA has worked closely with ARB and other agencies on a host of technical issues on standard setting and implementation and planning including: modeling for ozone and PM impacts, development of the UAM and CB chemical mechanisms, inter-

basin pollutant transport, NOx and ROG impacts on ozone concentrations, fugitive emission factors, emission estimation techniques for criteria and non-criteria pollutants, impacts of I/M on vehicle emissions, SMOG check, and regulatory programs including reformulated gasoline, AB2588, and ORVR.

API has had similar experience working with USEPA on ozone and PM modeling, attainment demonstrations, criteria setting, emission estimation techniques, deposition and regional transport. Together, both organizations, have sought to have the best science used in policy development and in making regulatory decisions.

We have Dr. Al Lefohn of ASL Associates, Dr. Stan Hayes of Environ Corporation and Professor Paul Switzer of Stanford University Department of Statistics to help us review the available documentation and prepare detailed comments. These nationally recognized experts have been deeply involved in many aspects of ozone research, have authored important sections of current and past chapters for the federal Ozone Criteria Documents, and helped develop approaches for assessing ozone risk.

Proposed Ozone Standard is not sufficiently supported

We recognize that the issue of unhealthy ozone concentrations is a concern to all Californians. We also understand that this issue is highly complex and not something that should be proposed without the best scientific and technical data possible coupled with unbiased interpretation of the facts. With this as our basis for making comments, we believe that the proposed document needs to be revised to address our major points of concern. We do not believe, given our review and our understanding of the issues, that the Draft document, as it currently stands, provides sufficient support for adopting a 70 ppb 8-hour average ozone standard. The documentation used to recommend this standard is the same as that upon which EPA arrived at an 80 ppb 8-hour standard. We also note that this recommended standard may be lower than relevant background, and is for all intents and purposes, unattainable.

We understand that California has always pursued standards that are more stringent than those set by EPA for the rest of the country. In this case, as in others, we need to ask whether in doing so, without a clear and strong health basis, such an action is good public policy. While we might be able to recognize this as a "goal", the current law which requires that all feasible controls be adopted to meet a standard which is essentially unattainable raises large concerns with the entire process. As such, it leads to false expectations among the public, and may divert environmental expenditures from other areas which might more effectively protect public health. WSPA and API stress that standard-setting in and of itself does not ensure improved health protection. In fact it is only when the state **achieves** those standards can an improvement in public health be presumed.

WSPA and API Recommendation

We do not believe that the current recommendation is supportable either in terms of form or level. Nor do we believe that it is wise public policy to adopt a standard that is unattainable. Given the impact of this decision, we do not see the need for a rush to judgment based upon questionable evidence and interpretation of the available data. We recommend that CARB review our comments and those of others, and revise the draft document appropriately, including explanations for why comments were not accepted. We do not see why, at this time that ARB needs to adopt a standard different than that already applicable to the rest of the country and suggest that a more thorough dialogue is required on this important decision. We welcome the opportunity to continue discussions with your agency.

After you have had a chance to review these submittals, please feel free to contact me at 310-808-2149, Mr. Kyle Isakower (API) at 202-682-8314, or Dr. Mark Saperstein, (BP, Chair of WSPA Task Force) at 714-228-6716 .

Sincerely.

Cc: Dr. Alan Lloyd
Ms. Catherine Witherspoon
Mr. Mike Schieble
Ms. Catherine Reheis-Boyd

**Comments on the California Ambient Air Quality Standard for Ozone
Document (CAAQSOD)**

Allen S. Lefohn, Ph.D.
A.S.L. & Associates
111 North Last Chance Gulch
Suite 4A
Helena, Montana 59601
406-443-3389
asl@attglobal.net
asl-associates.com

Prepared for

American Petroleum Institute
1220 "L" St., N.W.
Washington, D.C. 20005

and

Western States Petroleum Association
1115 11th Street
Suite 150
Sacramento, CA 95814

August 31, 2004

1. Summary

1.1 Introduction

The staff of the Air Resources Board (ARB) and Office of Environmental Health Hazard Assessment (OEHHA) have reviewed the scientific basis for the California Ambient Air Quality Standard for O₃ to determine its adequacy to protect public health, including the health of infants and children. Based on the results of the staff review and their findings, staff has recommended that the California ambient air quality standard for O₃ be revised. Staff has recommended the following revision be made to the California ambient air quality standard for O₃:

1. Retain O₃ as the indicator for oxidant air pollution.
2. Ozone 1-hour-average Standard – Retain the 1-hour-average standard for O₃ at 0.09 ppm.
3. Ozone 8-hour-average Standard – Establish an 8-hour-average standard for O₃ at 0.070 ppm.
4. For both the 1-hour and 8-hour O₃ standards, the concentrations for the standards noted above are established as “not to be exceeded”.
5. Ozone Monitoring Method – retain the current monitoring method for O₃, which uses the ultraviolet (UV) absorption method for determining compliance with the state Ambient Air Quality Standard for O₃. Incorporate all federally approved UV methods (listed at <http://www.epa.gov/ttn/amtic/criteria.html>) as California Approved Samplers for O₃. This will result in no change in air monitoring practices, but will align state monitoring requirements with federal requirements.

Based on these recommendations, a comprehensive review has been undertaken to evaluate the evidence for supporting the recommendation for an 8-hour O₃ standard at 0.07 ppm. In the pages that follow, O₃ background, *policy-relevant background*, human-health controlled exposure studies, and level of stringency and assumptions of

equivalency between a 1-hour 0.09 ppm and 8-hour 0.07 ppm standard are discussed. In a separate report, Dr. Paul Switzer, Department of Statistics, Stanford University, has reviewed the statistical methods used in the epidemiologic studies of ambient O₃ and mortality as discussed in Chapter 12 and Appendix B of the California Ambient Air Quality Standard for Ozone Document (CAAQSOD). Reference is made to Dr. Switzer's findings in the comments that follow in this section.

1.2 Policy-Relevant Background

The CAAQSOD states that from a regulatory perspective, the important distinction is not between “natural” and “anthropogenic” O₃, but between O₃ produced by controllable emissions and O₃ due to emissions beyond the reach of regulation.

Anthropogenic O₃ produced outside the jurisdiction of an agency and transported into a control region is functionally indistinguishable from that due to natural processes. Within the range of concentrations due to such external or uncontrollable sources, those concentrations that may impact determinations of compliance with air quality standards or limit the potential air quality improvements due to control programs have been defined by the CAAQSOD as *policy-relevant background*.

Estimates of *policy-relevant background* concentrations need to consider the important contribution from stratospheric O₃, as well as other natural sources. There is a large variability among global models on the attribution of the contribution of natural O₃ to the background. Although the CAAQSOD states that ground level impacts from fires are typically in the range of 15-25 ppb, such is not necessarily the case as indicated in this report. It is premature for the CAAQSOD to be concerned that surface O₃ levels in

California are currently influenced by the long-range transport from Asia. The scientific evidence for an Asian influence on surface O₃ concentrations on the United States is weak as documented in these comments and further research efforts are required.

The CAAQSOD emphasizes that the violations associated with the proposed 0.07 ppm 8-hour average standard would occur during the summer, when stratospheric O₃ contributions are thought to be minimal. However, when one characterizes the hourly average concentrations collected in 2003 for 184 monitoring sites in California, one finds that violations of the proposed 8-hour average standard occur during spring, summer, and fall. This implies that *policy-relevant background* concentrations that occur during seasons other than summertime will have to be characterized so that emission control actions result in optimum reductions in hourly average O₃ concentrations. At some monitoring sites in California, when stratospheric O₃ predominates in comparison to anthropogenic sources during the spring, it may *not* be possible for regulators to control hourly average concentrations in the 0.05 – 0.06 ppm range using emission reduction strategies and these concentrations may be associated with some of the 8-hour average violations of the proposed standard.

In fairly pristine locations in the United States, such as Yellowstone National Park in Wyoming, violations of the proposed California 8-hour standard of 0.070 ppm occur during the spring. The ambient concentrations experienced at Yellowstone National Park in the spring represent *policy-relevant background* as defined in Chapter 4 of the CAAQSOD. This implies that the proposed 8-hour standard will be difficult to attain in areas that are affected by stratospheric O₃ during the spring and that perhaps the

methodology used by Staff to propose the form and level of the 8-hour standard provides highly uncertain results.

The empirical data provide a solid indication to CAAQSOD that *policy-relevant background* O₃ is more than likely higher than the 15-35 ppb discussed in the document. Using models that provide highly uncertain concentration estimates provides an overly optimistic message to those who are responsible for implementing control strategies. In some of the modeling efforts to estimate natural background O₃ concentrations within North America, investigators removed all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer, but not biomass burning). Because the State of California does not plan to eliminate all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer), the estimates for the range of *policy-relevant background* concentrations will be greater than the background (i.e., natural background, in North America and anthropogenic and natural background outside of North America) values estimated by these models.

1.3 Human Health Effects – Controlled Exposure Studies

Experimental exposures of human volunteers to air pollutants under controlled laboratory conditions can provide useful pathophysiological information directly relevant to standard setting. Historically, most controlled studies of the effects of O₃ have used exposure protocols of one to two hours in duration, primarily based on a diurnal O₃ concentration profile typical of the Los Angeles air basin, and the notion that peak concentration, rather than cumulative dose, was the most important factor mediating O₃

toxicity. Based on the hypothesis that a broad plateau lasting as long as eight to twelve hours was important, “square-wave” or constant-concentration profiles were used by researchers to investigate the effects of multi-hour exposure to an O₃ concentration lower than the 1-hour ambient air quality standard. Unfortunately, this plateau pattern does not accurately reflect a typical urban O₃ exposure profile and in fact, it is quite rare. As such, health studies based on this exposure pattern are not helpful for estimating health impacts of typical ambient O₃ exposures.

Hazucha *et al.* (1992) were the first to design a protocol utilizing 8-hour exposures with both square and peaked O₃ concentration profiles. The study compared responses to exposure to a constant O₃ concentration (i.e., square-wave pattern) of 0.12 ppm, and to a variable concentration (i.e., triangle-shaped pattern) profile (linear increase from 0 to 0.24 ppm over four hours, followed by linear decrease from 0.24 to 0 ppm over 4 hours). Adams (2003a) also compared responses to square wave and triangular exposure scenarios using the same 8-hour average concentration of 0.08 ppm. In both Hazucha *et al.* (1992) and Adams (2003a), the cumulative exposure in the corresponding square-wave and triangular exposure protocols were essentially the same. Similarly, within each study, the final group mean FEV1 decrements between the two exposure scenarios were very similar. However, significant differences between the square-wave and triangular exposure scenarios were observed when comparing the hour-by-hour results. These findings suggest that relying on pre-post measurement differences across the 8-hour period for developing the proposed standard could miss important health impacts. They also illustrate that the O₃ induced FEV1 response is dependent on the dose rate as well as the cumulative dose of O₃ inhaled, at least when the O₃ concentration is

variable such as in the case of typical urban area diurnal exposure patterns. The important ramification of the results reported by Hazucha *et al.* (1992) and Adams (2003a) is that a non-linear dose-response relationship is evident. The higher hourly average concentrations elicit a greater effect than the lower hourly average values in a non-linear manner. *A major implication of a non-linear dose-response relationship is that the same 8-hour average concentration, with different distributions of hourly average concentrations, will elicit a different adverse effect.*

This points out an important limitation of a form of the proposed 8-hour O₃ standard, which averages hourly concentrations. The proposed level of the 8-hour standard described in the CAAQSOD is designed to protect the public from square-wave exposures. Since few people experience this exposure pattern, the level of the standard should be designed to protect the public from the family of realistic hourly average concentrations, over an 8-hour period, that elicits an adverse effect. Simply using data derived from square-wave exposure regime experiments provides an imprecise estimate of the realistic patterns of hourly average concentrations one is attempting to affect by emission reductions in order to protect the public's health. Given the importance of the non-linear dose-response relationship, it is very important that CAAQSOD reconsider the inappropriate use of an 8-hour standard that *averages* over the period of time. Hazucha's and Adams' research, based on realistic O₃ exposure patterns, which found dosing rate influencing health effects, indicates that in setting the level of the 8-hour standard the higher hourly average concentrations should be provided greater weight than the lower levels. Therefore, the use of a simple averaging scheme over an 8-hour period, which assumes a linear dose-response relationship (which does not occur), is inappropriate.

Hourly average O₃ control concentrations used in the controlled human exposure experiments were near zero. Obviously, these concentrations are lower than (1) the *policy-relevant background* levels discussed in Chapter 4 and (2) those hourly average concentration levels monitored at relatively remote monitoring sites in the United States or in other parts of the world (Lefohn *et al.*, 1990; Lefohn *et al.*, 2001). The use of concentrations below the cleanest monitoring sites in the world raises the concern that results obtained from the 6.6-h experiments may provide *overestimates* of health effects likely to be experienced by the population, as well as overestimates of health benefits that could be achieved with O₃ control strategies when compared to reference concentrations that represent *policy-relevant background* levels. Thus, the experimental design used in the controlled exposure studies precludes interpreting their results as demonstrations of significant difference between the health impact of ambient exposure and *policy-relevant background*.

Lastly, although the definition of an adverse health effect is based on guidelines developed by the American Thoracic Society, the guidelines tend to be qualitative in nature and therefore subjective. It is recommended that CAAQSOD should develop explicit criteria for determining the magnitude of changes in pulmonary function tests and symptoms used in controlled human exposure studies to represent adverse health effects.

In conclusion, with respect to human controlled exposure studies, because (1) the relationship between O₃ dose-rate and subsequent health impacts has not been addressed, (2) studies have not accounted for *policy relevant* O₃ background levels, and (3) there is ambiguity surrounding the definition of adverse health effects, the basis for the level of the proposed 8-hr standard is not well supported.

1.4 Level of stringency and assumptions of equivalency between a 1-hour 0.09 ppm and 8-hour 0.070 ppm standard

The CAAQSOD has looked at the relationships between ambient exposure and 1-hour and 8-hour concentrations. Much of this is discussed, tabulated, and graphically displayed in Chapter 7 and summarized in portions of Chapter 8. The CAAQSOD notes in Section 8.1 that one of the factors considered in assessing health protectiveness of each ambient standard is “the degree of outdoor exposure in California relative to the level of the standard”. Further, the document notes that effects at concentrations at or below the current State standard of 0.09 ppm, averaged over one hour, would provide evidence for the need for a more stringent standard, an averaging time different than the current 1-hour standard, or both. Given this, coupled with the decision to retain the current 1-hour standard, it appears that the CAAQSOD has turned to analyses of ambient air quality to show the relationship between exposure for the 1-hour 0.09 ppm standard and alternative concentrations for an 8-hour standard.

As described in Chapter 10, the CAAQSOD, in order to calculate changes in exposure to O₃ that reflect a hypothetical attainment of the proposed ambient air quality standards, has used a proportional linear rollback procedure. The use of a proportional rollback methodology does not mimic the atmosphere’s response to changes in O₃ precursors, which is a *non-linear* process. This means that the hourly concentrations within an 8-hour average do not respond in a linear manner when emission reductions occur. Second, as a result of the *non-linear* process, the amount of reduction needed to attain an 8-hour 0.070 ppm standard will be *significantly* more than that for a 90 ppb standard. Both of these statements are predicated on our understanding of atmospheric

chemistry of O₃ formation. In addition, there is a serious question of whether the 8-hour standard is set at a concentration (i.e., 0.070 ppm) that really represents at times *policy-relevant background* levels. With a proposed 8-hour 0.070 ppm California standard, the difficulty in attainment increases to the point that it is not clear that serious amounts of reductions of anthropogenic emissions will result in attaining a 0.070 ppm level.

Realistically, the proposed 8-hour California standard of 0.070 ppm may actually be an *Objective* (i.e., goal) instead of a standard.

Using models, several investigators have commented on the difficulty of reducing the mid-level hourly average concentrations, while reducing the fourth highest 8-hour average daily maximum concentration. Winner and Cass (2000) noted that the higher hourly average concentrations were reduced much faster than the mid-level values during simulation modeling for the Los Angeles area. Reynolds *et al.* (2003) analyzed ambient O₃ concentrations used in conjunction with the application of photochemical modeling to determine the technical feasibility of reducing hourly average concentrations in central California, using the 1990 August 3-6 San Joaquin Valley Air Quality Study episode.

Reynolds *et al.* (2003) have commented on possible chemical explanations for the observation that more prominent trends in peak 1-hour O₃ levels occur than for trends in peak 8-hour O₃ concentrations or in occurrences of mid-level (i.e., 0.06 –0.09 ppm) concentrations. The authors noted that when anthropogenic VOC and NO_x emissions are reduced significantly, the primary sources of O₃ precursors are biogenic emissions and CO from anthropogenic sources. Chemical process analysis results indicated that slowly reacting pollutants such as CO could be contributing on the order of 10 – 20% of the O₃ produced. Moreover, the authors noted that process analysis indicated that as NO_x was

reduced, the process for O₃ formation became more efficient, producing more molecules of O₃ for each molecule of NO_x. That is, decreasing emissions were offsetting the increased effectiveness of making O₃.

In conclusion, what has been proposed is a standard that will be extremely difficult to attain and perhaps even unattainable. Different exposure patterns (i.e., hourly frequency distributions) will occur than the predicted distributions based on a simple linear rollback. The proposed 8-hour standard will require much more stringent emission reductions than the current state 1-hour standard and it may not be possible to obtain the realistic emission reductions required to attain an 8-hour average standard at 0.070 ppm.

1.5 A review of the statistical methods used in the epidemiologic studies of ambient O₃ and mortality cited by CAAQSOD

In a separate report, Dr. Paul Switzer has reviewed the statistical methods used in the epidemiologic studies of ambient O₃ and mortality cited by CAAQSOD. As indicated by Dr. Switzer, the CAAQSOD synthesis of epidemiologic studies often contains important caveats regarding modeling issues. While his review deals principally with mortality time-series studies, the issues raised may be applicable to the much larger group of morbidity studies.

As indicated in Dr. Switzer's comments, the epidemiologic studies cited by CAAQSOD point to a string of inconsistent results when variations in ambient O₃ are related to variations in mortality, both for short-term and long-term exposures. Examples of the pattern of inconsistent and inconclusive findings include the following:

- Sharply different mortality effect estimates for summer and winter, which should not exist under the model of additive proportional effects used in the analyses.

- Instability of O₃ mortality effect estimates resulting from different model specifications of weather effects and time trends.
- Instability of O₃ effect estimates resulting from different selections of monitoring sites within cities.
- Heterogeneity of O₃ effect estimates across cities.
- Ozone effect lags that are inconsistent across cities and across studies.
- Exposure-response relations that are inconsistent across cities and across studies.
- Inconsistencies between short-term and long-term effect studies.

In his report, Dr. Switzer discusses several statistical concerns that were not adequately addressed:

1. Confounding of weather and time trends with O₃ effects
2. Heterogeneity of O₃ effects and effect modification
3. Heterogeneity of exposure within study areas
4. The relation between exposure and response
5. Mortality displacement
6. Long-term O₃-mortality studies

Dr. Switzer indicates the following in his comments:

1. **Sensitivity of O₃ effect estimates to model specification.** This issue was brought to light in the HEI reanalysis in the context of time and weather adjustments, and serves as a cautionary tale. The reported effects of O₃ are often difficult to discern and are inconsistent among cities, regions, seasons, and time lags. Such inconsistencies may be suggestive of modeling inadequacies, particularly in regard to unmodeled confounding and unexplained effect modifiers. That O₃ effect estimates are delicate is not surprising given that they are superimposed on much stronger effects due to concomitant weather variations, for example. Without a clear understanding of the reasons for inconsistent effects estimates, one cannot rule out the possibility that O₃ effect estimates are model artifacts.
2. **Enforced additivity in the analysis model.** The analysis models relied on by CAAQSOD assume that O₃ effects are necessarily the same at any temperature, even when restricted to summer data.
3. **Enforced linearity of exposure-response.** Because O₃ health effect estimates are inconsistent across studies, cities, seasons, etc., putative benefits of ambient O₃

mitigation are difficult to know. Enforced model linearity of exposure-response, as in the case of the analysis models relied on by the CAAQSOD relies on, conceals heterogeneity of response. Additionally, clinical studies with controlled exposure have shown a nonlinear relationship between exposure and response [Hazucha *et al.* (1992) and Adams (2003a)]. Pooling of response functions to obtain linearity is not statistically justified and leads to regulatory dilemmas.

4. **Spatial variability of O₃ health effect estimates within cities.** There has been insufficient attention to the issue of spatial variability of effect estimates within cities based on selection or combination of monitors.
5. **Incomplete characterization of the relations between ambient O₃ exposure, individual PM exposure, individual PM susceptibility to health effects, and community level health effect measures.** The models that CAAQSOD uses for the analysis of community health effects of O₃ do not have any link to individual response functions.
6. **No evaluation of the possibility of mortality displacement.** Some studies suggest that acute mortality effects are consistent with mortality displacement in frail populations. This issue is important for public policy and therefore needs to be studied so that regulatory decisions can truly address mitigation.
7. **Unresolved inconsistencies of O₃ effect estimates.** The following inconsistencies are unresolved: seasonal differences, regional grouping, spatial heterogeneity both between cities and within cities, time lag selection, and treatment of gaseous pollutant confounders.

Based on his thorough review, Dr. Switzer believes that the available epidemiologic evidence on O₃ mortality cannot be used to draw robust conclusions regarding the circumstances and magnitudes of ambient O₃ mortality, in particular whether reported O₃ effects are causative. Without a clear understanding of the reasons for inconsistent effects estimates, Dr. Switzer believes that one cannot rule out the possibility that O₃ effect estimates are model artifacts.

1.6 References

Adams, W. C. (2003a) Comparison of chamber and face mask 6.6-hour exposure to 0.08 ppm ozone via square-wave and triangular profiles on pulmonary responses. *Inhal Toxicol* 15:265-81.

Hazucha M. J; Folinsbee L. J.; Seal E. Jr. (1992) Effects of steady-state and variable ozone concentration profiles on pulmonary function. *Am Rev Respir Dis* 146:1487-93.

2. Policy-Relevant Background Ozone

2.1 Introduction

The California Ambient Air Quality Standard for Ozone Document (CAAQSOD) states (page 4-1) that from a regulatory perspective, the important distinction is not between “natural” and “anthropogenic” O₃, but between O₃ produced by controllable emissions and O₃ due to emissions beyond the reach of regulation. In a policy context, anthropogenic O₃ produced outside the jurisdiction of an agency and transported into a control region is functionally indistinguishable from that due to natural processes. As noted on page 4-1, within the range of concentrations due to such external or uncontrollable sources, those concentrations that may impact determinations of compliance with air quality standards or limit the potential air quality improvements due to control programs are defined as the *policy-relevant background*.

Establishing ranges for hourly average O₃ *background* concentrations is important for several reasons. In addition to identifying the range of *policy-relevant background* concentrations to better understand the relationship between emission reductions and resultant O₃ exposure patterns, human health and vegetation researchers need to use the range of *policy-relevant background* concentrations in their controls so that the models that result from their investigations reflect true comparisons with ambient background conditions. If such is not implemented, then it may be possible that overestimations occur in the biological modeling predictions.

In the comments that follow, specific focus is provided on the following issues:

- Estimates of *policy-relevant background* concentrations need to consider the important contribution from stratospheric O₃, as well as other natural sources;

- There is large variability among global models on the attribution of the contribution of natural O₃ to the background;
- The CAAQSOD states that ground level impacts from fires are typically in the range of 15-25 ppb. Such is not necessarily the case;
- Given the limitations discussed in this report with the Lin *et al.* (2000) and Jaffe *et al.* (2003b) trending analyses, the scientific evidence for an Asian influence on surface O₃ concentrations on the United States is weak and further research efforts are required;
- Based on the limitations associated with the methodologies implied by Jaffe *et al.* (2003b) and Lin *et al.* (2000), it is premature for the CAAQSOD to be concerned that surface O₃ levels in California are currently influenced by the long-range transport from Asia.
- The CAAQSOD emphasizes that the violations associated with the proposed 0.07 ppm 8-hour average standard would occur during the summertime, when stratospheric O₃ contributions are thought to be minimal. However, when one characterizes the hourly average concentrations collected in 2003 for 184 monitoring sites in California, one finds that violations of the proposed 8-hour average standard occur during spring, summer, and fall;
- Because violations of the proposed 8-hour average standard occur during spring summer, and fall, *policy-relevant background* concentrations that occur during seasons other than summer will have to be characterized so that emission control actions result in optimum reductions in hourly average O₃ concentrations;
- At some monitoring sites in California, when stratospheric O₃ predominates in comparison to anthropogenic sources during the spring, it may *not* be possible for regulators to control hourly average concentrations in the 0.05 – 0.06 ppm range using emission reduction strategies;
- The empirical data provide a solid indication to CAAQSOD that *policy-relevant background* O₃ hourly average concentrations, as defined on page 4-1, are more than likely higher than the 15-35 ppb discussed in the document. Using models that provide highly uncertain concentration estimates provides an overly optimistic message to those who are responsible for implementing control strategies.
- In some of the modeling efforts to estimate natural background O₃ concentrations within North America, investigators removed all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer, but not biomass burning). Because the State of California does not plan to eliminate all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer), the estimates for the range of hourly average *policy-relevant background* concentrations will be greater than the 4-hour afternoon average background (i.e., natural background, in North America and anthropogenic and natural background outside of North America) values estimated by these models.

- The proposed 8-hour standard of 0.07 ppm is violated in pristine places, such as Yellowstone National Park in Wyoming. The ambient concentrations experienced at Yellowstone National Park in the springtime represent *policy-relevant background* as defined in Chapter 4 of the CAAQSOD. This implies that the proposed 8-hour standard will be difficult to attain in some areas that are affected by stratospheric O₃ during the spring and that perhaps the methodology used by Staff to propose the form and level of the 8-hour standard provides highly uncertain results.

2.2 Historical Measurements

Historical measurements show that O₃ concentrations were highly variable under “pre-industrial” conditions. Although not representing natural O₃ background, attempts have been made, using historical data, to estimate O₃ concentrations in the late 1800s and early 1900s (Bojkov, 1986; Volz and Kley, 1988; Thompson, 1992). Large ranges of concentrations have been reported in the literature. Using the Schönbein method, Bojkov (1986) concluded that the average daily maximum of the surface O₃ partial pressure in the Great Lakes area of North America was approximately 0.019 ppm and the European measurements between the 1850s and 1900 were mostly in the range of approximately 0.017 to 0.023 ppm. Marengo *et al.* (1994) reported average annual O₃ concentrations of 0.010 ppm from 1875 to 1895. However, higher concentrations, for the period 1889-1900, were reported for Zagreb, Croatia by Lisac and Grubišić (1991). The authors reported that the 1893-1900 O₃ concentrations were approximately 0.036 ppm for the multi-hour average daytime period and 0.030 ppm for the average nighttime period.

The earliest measurements of O₃ concentration, using the Schönbein method, should be regarded as approximate rather than absolute (e.g., Pavelin *et al.*, 1999). Many uncertainties exist when attempting to relate data collected by the Schönbein method to absolute O₃ concentrations. Relative humidity variation among different monitoring sites

makes it difficult to compare Schönbein values. During the second half of the nineteenth century, quantitative measurements of surface O₃ were made at the Observatoire de Montsouris, located on the outskirts of Paris (Volz and Kley, 1988). Beginning in 1876 and continuing for 31 years, daily measurements were carried out. The method had a positive interference when H₂O₂ and NO₂ were present and a negative interference when SO₂ was present.

The quality of the O₃ data collected at Montsouris, as well as other locations in the late 1800s and early 1900s, is unclear and any comparison of concentrations inferred from measurements during this period with current concentrations at remote, rural sites should be made with great caution (Lefohn *et al.*, 1992). In addition, it is unknown to what extent the Montsouris data represent O₃ concentrations in Europe or the Northern Hemisphere in the nineteenth century. Because of the uncertainty of the quality of the data and the interference from the large emissions of sulfur dioxide, which would result in an underestimate of O₃ concentrations, it is difficult to quantify the differences between surface O₃ concentrations measured in the last half of the nineteenth century at certain locations in either Europe or North America with those currently monitored at remote locations in the world. There is little reliable observational information and estimates of increases during the first half of the twentieth century rely primarily on modeling results. These models imply that the geochemical background O₃ concentration has increased due to anthropogenic emissions (Yienger *et al.*, 1999).

Under pre-industrial conditions, O₃ at ground level is largely the result of three processes: meteorologically regulated downward mixing of O₃ from the stratosphere, local boundary layer O₃ formation due to photochemical reactions of natural precursors,

and regional to continental scale impacts of large biomass fires producing episodic releases of large volumes of O₃ precursors (VOCs and NO_x). Since there are only a few O₃ measurements available from pre-industrial times, global distributions of “natural” concentrations must be inferred from chemical kinetics and atmospheric modeling, which are subject to great uncertainty.

2.3 Stratospheric Contributions

One source of natural O₃ is the stratosphere. However, O₃ in the stratosphere, isolated far above the ground, is only *policy relevant* if it is transported to lower altitudes, where it may impact terrestrial organisms. Under unperturbed meteorological conditions, the troposphere and stratosphere do not mix, so stratospheric O₃ generally does not affect ground-level O₃. Strong storms in the troposphere can occasionally cause stratospheric air to be drawn downward into the lower atmosphere. Such stratospheric intrusions or “tropopause folding events” can bring high concentrations of O₃ far down into the troposphere. In the midlatitudes of the northern hemisphere, these events tend to occur in spring, accompanying deep pressure “troughs.” When present, stratospheric O₃ will be greatest at high mountain locations; effects at lower elevations do occur, but less frequently (Lefohn *et al.*, 2001).

Researchers who have focused on the evidence for stratospheric intrusions affecting monitoring sites in the eastern United States have reported infrequent occurrences of elevated hourly average O₃ concentrations exceeding 0.12 ppm (e.g., Logan, 1989). Using thresholds substantially lower than the 0.12 ppm level, Lefohn *et al.* (2001) documented evidence for several North American sites where the stratosphere apparently contributed to a frequent number of exceedances of the hourly average concentration

equal to and above 0.05 and 0.06 ppm. The occurrence of elevated O₃ concentrations due to stratospheric sources in parts of Europe is well documented (e.g., Derwent *et al.*, 1976; Stohl *et al.*, 2000), as well as in Japan (Wakamatsu *et al.*, 1989).

Results reported by Singh *et al.*, (1978) for the United States showed a distinct seasonal variation of tropospheric O₃, with a maximum in the spring when 1-hour O₃ concentrations approached or exceeded 0.08 ppm. Lefohn *et al.* (2001) reported the frequent occurrence of hourly average O₃ concentrations \geq 0.05 and 0.06 ppm during the late winter and spring months across North America and northern Europe. Additionally, Lefohn *et al.* (2001) reported that in 1998 there were 385 occurrences of hourly average concentrations \geq 0.05 ppm that occurred during April in Yellowstone National Park (WY). For most sites, the selected thresholds were exceeded most frequently in April and May but were also noted in March and June. By June, and often in May, the potential for pollution-related O₃ production exists even at higher latitude sites. The analysis by Lefohn *et al.* (2001) suggested that even in these months, many of the threshold-exceeding cases were associated with natural O₃ sources.

To explore a possible reason for the high number of hourly occurrences for the threshold concentrations, Lefohn *et al.* (2001) investigated several occurrences of elevated O₃ hourly average concentrations in the United States and Canada. An O₃ episode occurred in Boulder, Colorado (EPA AIRS Site 080130011) on May 6, 1999. At 1700 UTC (1000 LST), an hourly average concentration of 0.060 ppm was recorded and by 2100 UTC (1400 LST), the maximum hourly average O₃ concentration of 0.076 ppm was measured. At 0200 UTC on May 7, 1999 (1900 LST on May 6), the hourly average

concentration declined to 0.059 ppm. Figure 2-1 shows the O₃ vertical profile that was recorded at Boulder, Colorado on May 6, 1999, at 1802 UTC (1102 LST).

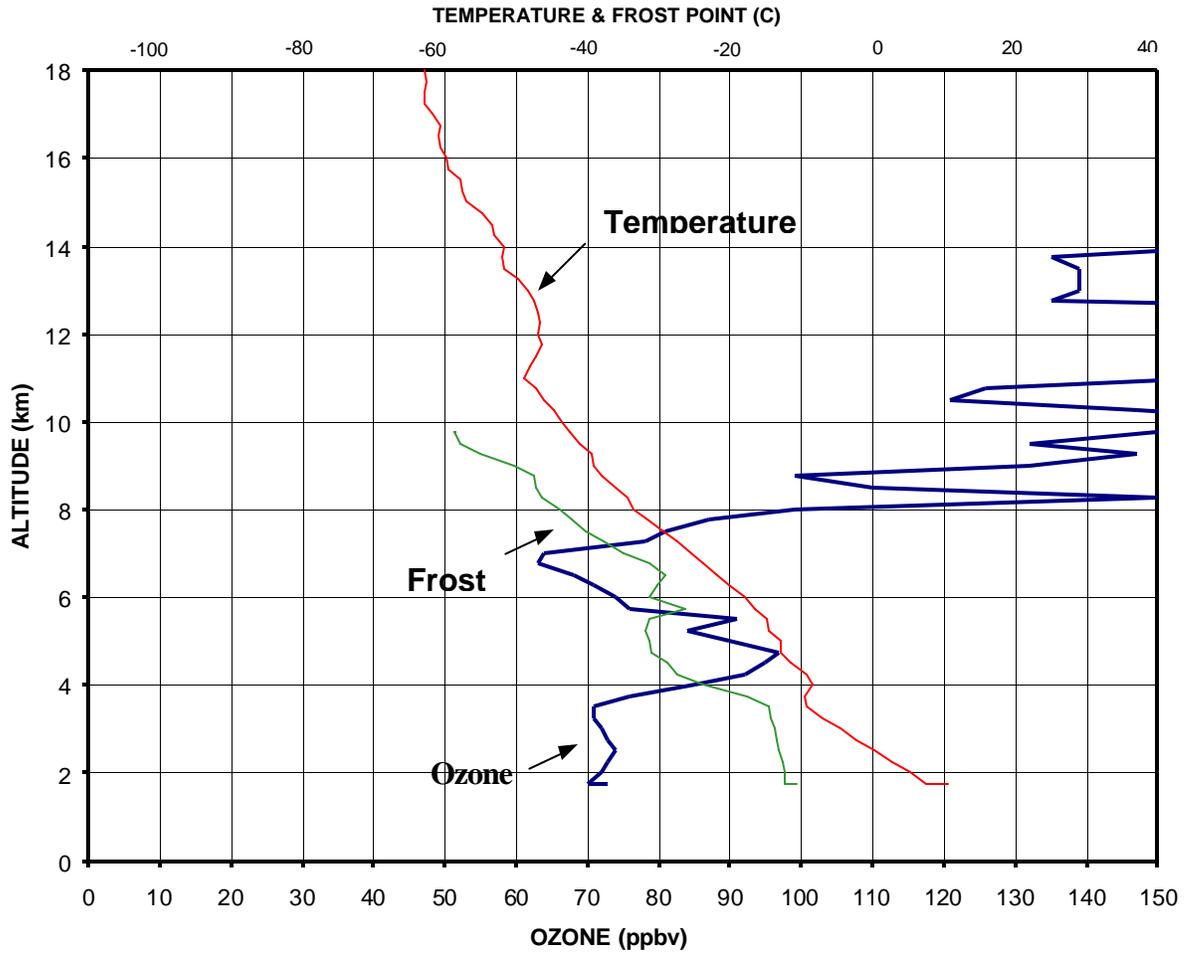


Figure 2-1. Ozone vertical profile at Boulder, Colorado on May 6, 1999, at 1802 UTC. Source: Lefohn *et al.* (2001)

During this intrusion event, the presence of recently injected stratospheric air is seen in the layer at 8-11 km. At 4-5.5 km, air was injected into the troposphere upstream of Boulder and has descended as the upper level low moves overhead at the time of the sounding. Air near the surface that is extremely dry, which is a strong signature for

stratospheric influence, has mixed down from the layer above as the air continues to descend over the Rocky Mountains. Although these high O₃ concentrations (~0.075 ppm) will be diluted as the air is transported eastward, this will raise the level of the unperturbed (no easily discernible sign of an intrusion) O₃ background. It has also contributed a significant high O₃ "event" to the distribution of hourly values measured at Boulder.

Based on continuous LIght Detection And Ranging (LIDAR) measurements of the tropospheric O₃ profile on May 6 from a site near Boulder, and an analysis of the meteorological data, stratospheric intrusion was the cause of the elevated O₃ measurements recorded on this day. Stratospheric intrusion processes, as well as local photochemical processes, affect the Boulder, Colorado, site during May. Prior to May, elevated O₃ concentrations are associated with stratospherically influenced air as illustrated in the example for May 6, 1999. At this site, higher ground-level concentrations of O₃ are present during the summer months. These higher concentrations at ground level are associated with anthropogenic sources.

Although Lefohn *et al.* (2001) could not eliminate the possibility that anthropogenic-derived precursors could be responsible for the O₃ concentrations measured at the sites investigated by Lefohn *et al.* (2001), because of the seasonal timing and episodic occurrences of these concentrations, the authors concluded that there was a strong possibility that natural sources of O₃ were responsible for these exposures. Although the analysis was limited to the more northerly latitudes, the authors felt that it might be expected that a significant contribution from natural sources was also important in producing O₃ levels in the 0.05 – 0.06 ppm or higher range at more southerly locations

where stratospheric intrusions also have an important impact on the tropospheric O₃ distribution (Moody *et al.*, 1995; Oltmans *et al.*, 1996; Cooper *et al.*, 1998).

Oltmans *et al.* (personal communication) have described the collection of ozonesonde measurements beginning at Trinidad, California in August 1997 and have continued on a weekly basis. Twenty-nine O₃ profiles were obtained from April 17 – May 20, 2002 during ITCT 2K2. Surface O₃ measurements began in mid-April 2002 and continue to the present at Trinidad Head. Figure 2-2 illustrates the O₃ mixing ratios from April 17 – May 18, 2002.

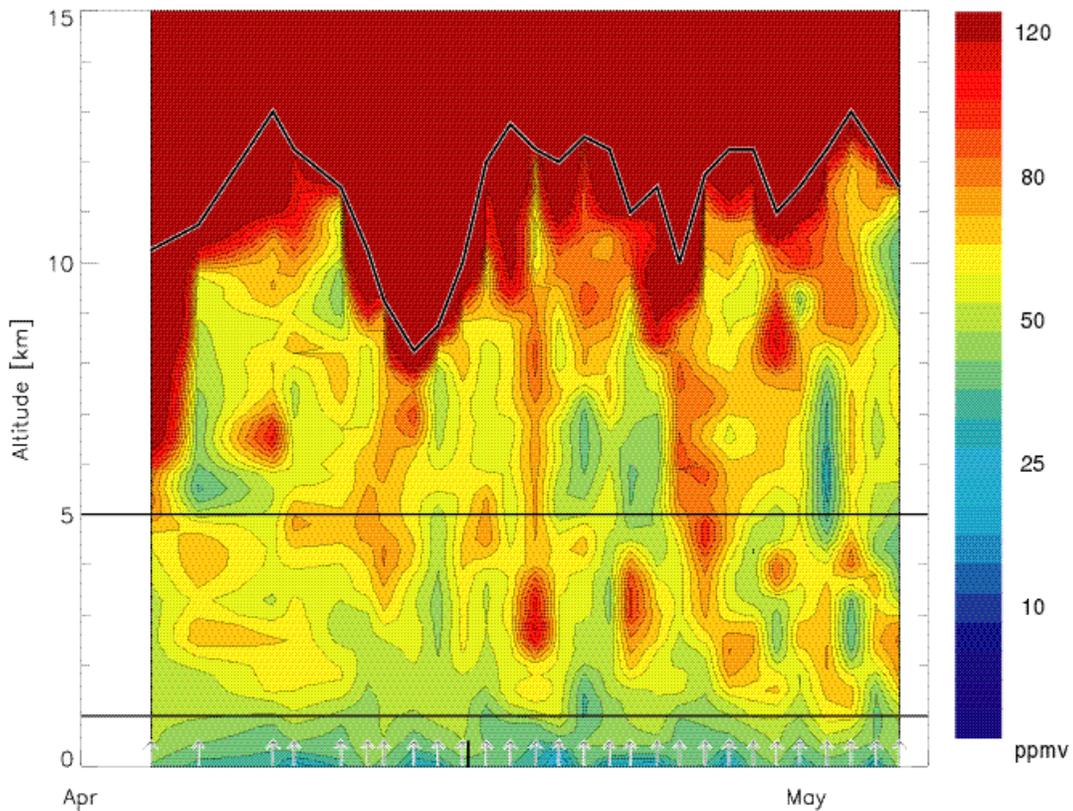


Figure 2.2. Ozone mixing ratio at Trinidad, California (April 17 – May 18, 2002)

Oltmans *et al.* (personal communication) reported that there is a prominent seasonal cycle in O₃ throughout the troposphere at Trinidad Head with a maximum in

April-May. The April-May period of 2002 appears to have average O₃ amounts very similar to the longer-term average. The daily O₃ profiles during ITCT 2K2 provided an opportunity for looking at the details of the spring maximum. During ITCT 2K2, approximately three-fourths of the profiles showed enhanced O₃ layers (exceeded average by >10 ppb) in the 1 – 5 km altitude region. In the 1-5 km altitude region, 22 of 29 profiles had layers with peak values exceeding 70 ppb (this represents an enhancement of ~10 ppb above the average profile). The large majority of these layers appeared to be the result of transport of O₃ from the stratosphere (based on very low humidity, trajectory behavior and potential vorticity). Oltmans *et al.* (personal communication) concluded that the seasonal maximum in tropospheric O₃ at Trinidad receives an important contribution from stratospheric O₃. Oltmans (personal communication) noted that based on the measurements taken at Trinidad Head, mountainous terrain to the east of the area might experience elevated O₃ concentrations in the springtime that were associated with important contributions from the stratosphere.

On page 3-9 of the CAAQSOD, the document states that in continental areas far removed from direct anthropogenic effects, O₃ concentrations are generally 20 - 40 ppb (it is unclear if this is an hourly average value or an 8-hour maximum value). However, one finds higher concentrations than 20 – 40 ppb at isolated monitoring sites in locations such as Montana (Glacier National Park, elevation 963 m) and Wyoming (Yellowstone National Park, elevation 2.47 km). For the period 1987 – 2003, these sites experienced maximum 8-hour average concentrations in the range 0.057 – 0.078 ppm (Yellowstone N.P.) and 0.054 to 0.065 ppm (Glacier N.P.). In 1999 at Yellowstone National Park, all seven of the maximum 8-hour average values ≥ 0.070 ppm occurred during the

springtime, which appears to be associated with stratospheric contributions. At no time were the hourly average concentrations above 0.079 ppm during any of the seven 8-hour average exceedances. During March, April, and May, there were 509, 517, and 458 occurrences of hourly average concentrations ≥ 0.05 ppm, respectively. Thus, Yellowstone National Park experienced more than 50% of its hourly average concentrations during each of the 3 springtime months ≥ 0.05 ppm. The Yellowstone National Park site in Wyoming would violate the proposed California 8-hour standard of 0.07 ppm. Although the maximum 8-hour average concentration of 0.07 ppm was not exceeded in Glacier National Park, it appears that sites at elevations below 1 km are experiencing O₃ concentrations affected by the stratosphere. Because the proposed California 8-hour standard will be exceeded at some locations during the springtime, estimates of *policy-relevant background* concentrations need to consider the important contribution from stratospheric O₃, as well as other natural sources.

2.4 Using Low Spatial Resolution Models to Predict Policy-Relevant Background

A study by Fiore *et al.* (2003) used a global chemical transport model (CTM) with a 2° x 2.5° horizontal resolution (approximately 200 km by 220 km in the northern United States) to determine the origin of hourly average concentration ≥ 0.05 ppm reported by Lefohn *et al.* (2001), and to conduct a more general quantitative analysis of background as a function of season, site location, and local O₃ concentration. Results from modeling studies, where the background is estimated by zeroing anthropogenic emissions in North America, were on the low end of the 25 – 45 ppb range. The practice of zeroing anthropogenic emissions may provide questionable results because the removal of a very large portion of the emissions inventory may introduce additional

uncertainties into the model due to very large perturbations. CARB does not endorse use of zero emissions scenarios when performing sensitivity simulations for regulatory purposes (Ziman, personal communication). Lefohn *et al.* (2001) have indicated that occurrences of O₃ concentrations ≥ 0.05 ppm at remote northern U.S. (e.g., Yellowstone National Park and Glacier National Park), western Canadian, and northern European sites in spring appear to be of stratospheric origin. However, Fiore *et al.* (2003) attribute these events to O₃ produced from North American anthropogenic emissions, with some contributions from hemispheric pollution (i.e., O₃) produced from anthropogenic emissions outside North America, and only a *minor* stratospheric contribution (almost always < 10 ppb).

There are uncertainties in the GEOS-CHEM CTM model calculations as described by Fiore *et al.* (2003) in the temporal variability of O₃ originating from different sources on shorter time scales that must be recognized. These uncertainties relate to the partitioning of O₃ among its sources. The uncertainties stem in part from the lack of seasonal variability in the stratosphere-troposphere exchange of O₃ (Fusco and Logan, 2003), the geographical variability of this exchange, and the variability in the exchange between the free troposphere and the planetary boundary layer in the model. Ideally, the predictions resulting from an ensemble of models should be compared with each other and with observations so the range of uncertainty inherent in the model predictions can be evaluated. Unfortunately, such a comparison has not been performed. Simulation of stratospheric intrusions is difficult in global models, and stratosphere-troposphere exchange of O₃ in these models is generally parameterized. There are uncertainties in both the lifetime of O₃ and the rate of photochemical production of O₃.

These uncertainties are especially important in the spring, when the relative role of stratospheric intrusions may be at its largest compared to photochemistry.

Fusco and Logan (2003) have compared the Harvard-GISS model with the GEOS-CHEM model and have noted that the amplitude of the seasonal cycle of the O₃ flux in the Harvard-GISS model is much larger than that of the modeled flux in the GEOS-CHEM model used by Fiore et al. (2002, 2003) in both hemispheres. In the Northern Hemisphere (NH), the flux of O₃ in the Harvard-GISS model remains low throughout much of the fall before rising sharply in the early winter. The NH flux of O₃ then remains high throughout the winter and spring. In the GEOS-CHEM model, the seasonal cycle of the O₃ flux in the NH is much weaker. As a result, large discrepancies arise between the modeled fluxes in early fall and in late spring. This produces significant differences in modeled tropospheric O₃.

Even if the magnitude of cross-tropopause O₃ fluxes in global CTMs were calculated correctly in an annual mean sense, it should be noted that stratospheric intrusions occur episodically following the passage of cold fronts at mid-latitudes. Of major concern is the ability of global-scale CTMs to simulate individual intrusions and the effects on surface O₃ concentrations that may result during these events. Higher resolution models capable of spatially and temporally resolving stratospheric intrusions and capable of resolving O₃ variability on hourly time scales have not been applied to this problem. Until numerical model estimates of the global tropospheric O₃ budgets can be made with global models that have the capability to resolve stratospheric-tropospheric exchange processes, the modeling results reported by Fiore *et al.* (2002, 2003) should be treated with caution.

Significant differences exist between the predictions estimated by Fiore *et al.* (2002, 2003) and the estimates of background provided with observational data (e.g., Singh *et al.*, 1978; Lefohn *et al.*, 2001). Fiore *et al.* (2002, 2003) have used their global chemical transport model to estimate background levels. Fiore *et al.* (2003) estimate that natural background O₃ levels (4-hour afternoon average concentrations) are in the 10-25 ppb range and *never* exceed 40 ppb. Fiore *et al.* (2003) conclude that based on their model, the stratospheric contribution to surface O₃ is of minor importance, typically well below 20 ppb. The authors conclude that although stratospheric intrusions might occasionally elevate surface O₃ at high-altitude sites, these events are rare. Fiore *et al.* (2002) estimate that in the summer no more than 2 ppb O₃ at the surface can be ascribed to stratospheric origin.

Lefohn *et al.* (2001) observed frequent hourly average concentrations ≥ 0.05 ppm at locations deemed to be clean sites in the northern part of the western U.S, southern Canada, and Northern Europe, during a time of the year when regional photochemical activity was judged to be low (late winter – early spring). These events were attributed to enhancements to the typical seasonal levels of 0.03-0.04 ppm by natural sources, such as mixing down of air that had a stratospheric origin. The global model used by Fiore *et al.* (2002, 2003), on the other hand, finds no significant contribution to such events from air that may have enhanced O₃ from the stratosphere, but rather concludes that such events result from the impact of regional air pollution. This is in keeping with results (Li *et al.*, 2002) from this model, which conclude that earlier studies on the source of high O₃ events seen in the spring at Bermuda result from the transport of pollution from North America, rather than from strongly subsiding air from the upper troposphere that was

probably enhanced with O₃ from the stratosphere (Moody *et al.*, 1996, Oltmans and Levy II, 1992). Although the observed O₃ cannot be decomposed into its contributing sources as can theoretically be done in the model, there is a significant difference in the attribution of the source of O₃ events ≥ 0.05 ppm seen at “clean” sites between the observational studies and the global chemical transport model.

Measurements of O₃ at the surface at remote (or “clean”) sites will always reflect a mixture of air from different sources. In the Fiore *et al.* (2002) model, sources are “shutoff” with an attribution based on the modeled O₃ compared to the actually measured O₃. Using the technique in the Harvard global model (Fiore *et al.*, 2002, 2003), little or no contribution to the background was found from stratospheric sources, particularly in events with O₃ amounts ≥ 0.05 ppm, during late winter or spring. Recently published results of a project carried out in Europe (Stohl *et al.*, 2003) detailed a number of features of stratosphere-troposphere exchange (STE). In addition, a recent observational study (Olsen *et al.*, 2002) found a relatively large flux of O₃ from the stratosphere into the troposphere at extratropical latitudes of the Northern Hemisphere, compared to budgets found in a number of global photochemical models (Collins *et al.*, 2000). It is also noted in Olsen *et al.* (2002) that use of a single direct potential vorticity and O₃ relationship to infer O₃ fluxes may produce unreliable estimates because of seasonal and geographical variations in this relationship. Stohl *et al.* (2003) provide a concise summary of the current understanding of STE. A model intercomparison looking at actual STE events found significant variations in model results that depended significantly on the type and horizontal resolution of the model (Meloan *et al.*, 2003; Cristofanelli *et al.*, 2003). In particular, it was found that the Lagrangian perspective was necessary to characterize the

depths and residence times of individual events (Sprenger and Wernli, 2003; James *et al.*, 2003a; 2003b). As noted in the review by Stohl *et al.* (2003), on a global basis, the stratospheric flux of O₃ to the troposphere only amounts to 10-20% of the photochemical O₃ production. However, with the longer photochemical lifetime of O₃ in winter and spring at extratropical latitudes, and with stratosphere-troposphere exchange maximized during this time of year, the actual contribution to tropospheric O₃ may be significantly larger (Stohl *et al.*, 2003).

At the present stage of development in global models, there is no single model that has been accepted as providing the best estimate of background surface O₃. There is large variability among global models on the attribution of the contribution of natural O₃ to the background. Models, such as the one used by Lelieveld and Dentener (2000), utilize horizontal resolutions that are fairly low (e.g., 3.75° x 5°), which adds great uncertainty to the predicted results. The issue of the determination of the O₃ background is unsettled and controversial. A cautious approach to the determination of background surface O₃ amounts, including information from a suite of models, as well as observational approaches, is warranted.

2.5 The Importance of Biomass Fires

As pointed out in the CAAQSOD on page 4-6, biomass fires accelerate natural O₃ formation by distilling large amounts of VOCs out of plant material and producing CO and NO_x as products of combustion. The large amount of carbonaceous aerosol reduces solar UV flux, slowing photochemical processes near the fire, but the very high concentrations of combustion gases and long lifetimes of very large smoke plumes cause O₃ formation far downwind by oxidation of CO and methane (Crutzen, 1995). Very large

fires observed by satellite have been shown to have high O₃ concentrations accompanying their smoke plumes far downwind (Jenkins and Ryu, 2003). The CAAQSOD notes on page 4-6 that O₃ concentrations associated with fires are highly variable, and that large tropical fire plumes have been observed to commonly have 70 ppb O₃ or more in the middle troposphere. However, on page 4-6 of the CAAQSOD, the authors state that ground level impacts are much less, typically in the range of 15-25 ppb (Jenkins and Ryu, 2003). Such is not necessarily the case. For example, Altshuller and Lefohn (1996) noted that the large fires in Yellowstone National Park in 1988 resulted in hourly average concentrations of 98 ppb at the surface. The CAAQSOD notes on pages 4-7 and 4-8 that air pollution measurements influenced by large fires are generally recognized as exceptional events and excluded from consideration for standards compliance. It should be noted that the regulatory community in California has yet to decide whether or not to include fires in future-year simulations for O₃ attainment planning (Ziman, personal communication). Regulators should be aware of the occurrence of major fires in proximity to monitoring sites and the possibility that such events may affect the hourly average concentrations recorded during these periods.

2.6 The Evidence for Long-Range Transport from Asia Affecting O₃ Concentrations

The CAAQSOD notes on page 4-8 that O₃ due to urban and industrial emissions in east Asian megacities (Beijing, Tokyo, Hong Kong, Taipei, Shanghai, etc.) has been observed to reach western North America in springtime (Jaffe *et al.*, 1999; Jaffe *et al.*, 2003b; Jaffe, McKendry, Anderson, and Price, 2003). The document notes that although modeling exercises aimed at determining the Asian air pollution contribution in western

North America have consistently shown frequent enhancement of carbon monoxide and aerosols, O₃ effects have been harder to discern. The CAAQSOD notes that Jaffe *et al.* (2003b) report statistical analyses of O₃ records from rural sites from northern California to the Olympic Peninsula from the mid-1980s to 2002 that show a broad regional increase in “background” O₃. According to Jaffe *et al.* (2003b), all sites showed a statistically significant increase in springtime O₃ (about 4 ppb/decade) for days selected for oceanic influence, with stronger gradients for higher elevation sites. The CAAQSOD notes on page 4-9 that this is consistent with the vertical gradient pattern of Asian aerosol impacts (VanCuren and Cahill, 2002), ozonesonde observations (Newchurch, et al., 2003), and increased transport exposure above the marine boundary layer shown in transport modeling (Jaegle et al., 2003). However, it is important to carefully review the surface trending evidence provided in Jaffe *et al.* (2003b).

Jaffe *et al.* (2003b) used a 15-year record of O₃ from Lassen Volcanic National Park, a rural elevated site in northern California, data from two aircraft campaigns conducted in 1984 and 2002 over the eastern North Pacific, and observations spanning 18 years from five U.S. west coast, marine boundary layer sites, and reported that O₃ in air arriving from the Eastern Pacific in spring has increased by approximately 10 ppbv, i.e. 30% from the mid 1980s to the present.

The data used by Jaffe *et al.* (2003b) as listed in Table 1 of that paper were:

- CITE-1C (April-May 1984)
- Point Arena, CA (24 April-5 May 1985)
- Lassen Volcanic National Park (1988-2002)
- Redwood National Park (1988-1995)
- Point Reyes N.S. (1988-1992)
- Cheeka Peak, WA (March-April 1997-98); March-May 2001-02)
- Trinidad Head, CA (19 April-18 May, 2002)
- ITCT 2K2(April-May, 2002)

Taken from Jaffe *et al.* (2003b), Figure 2-3 below shows the springtime,

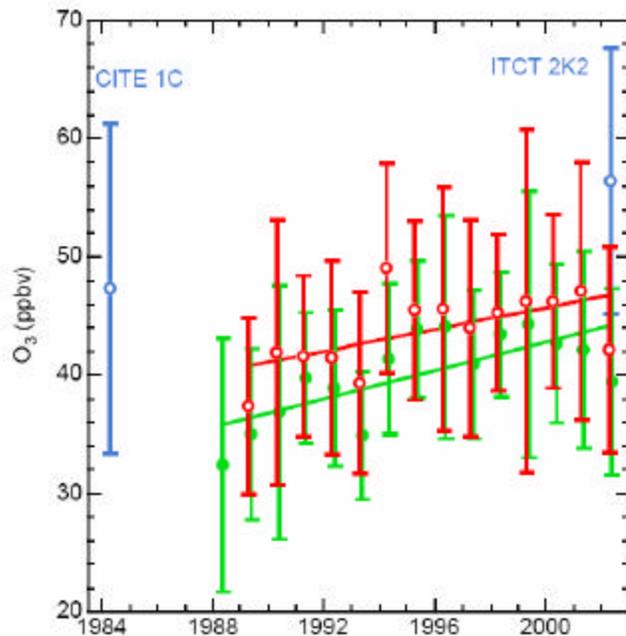


Figure 2-3. Spring mean mixing ratio \pm 1 standard deviation for background O₃ at Lassen Volcanic NP using the marine trajectory dataset (green) and the wind selected dataset (red). Also shown is data from the CITE 1C and ITCT aircraft campaigns. Source: Jaffe *et al.*, 2003b).

background O₃ mixing ratios measured at Lassen from 1988-2002 and the results from two aircraft campaigns (i.e., CITE 1C and ITCT 2K2). The monitor in Lassen Volcanic National Park in Northern California is located in the northwest corner of the park, away from any major emission sources or urban centers. However, as noted by the authors, it does lie approximately 240 km inland from the Pacific coast. The nearest city is Redding, California, 70 km to the west. The authors believed that because of its proximity to urban areas, Lassen required close investigation of the possible effects of North American

influences on observed O₃ levels. For Lassen, the authors used isentropic back-trajectories to segregate the data, since trajectories reflect the air parcel history in a way that local wind speed and direction cannot. They computed 10-day isentropic back-trajectories twice each day (0 and 12 GMT) during all seasons for the entire period of the data record (1988-2002), using the European Center for Medium range Weather Forecasting (ECMWF) data and previously published methods. The authors noted that because of prevailing westerly winds, most trajectories arrive at Lassen from the North Pacific. An air mass that took longer than 24 hours to reach the site from the Pacific coast (defined by 125°W longitude) or crossed over California south of 39° N latitude was classified as “local”. An air mass that arrived at the Lassen site having spent less than 24 hours over land was classified as “marine”. Each air mass type was associated with a 12-hour average O₃ concentration, centered on either 0 or 12 GMT. For spring, summer, fall and winter, the marine dataset contained 39, 28, 31, and 33 % of the full dataset. The authors defined the seasons as the three-month periods beginning on the first of March, June, September and December.

Figure 2-3 from Jaffe *et al.* (2003b) shows the spring (March-May) mean mixing ratio \pm 1 standard deviation for background O₃ at Lassen Volcanic NP, using the marine trajectory dataset (green) and the wind selected dataset (red). Also shown are data from the CITE 1C and ITCT aircraft campaigns. The authors reported that in spring, a consistent, statistically significant, positive trend in background O₃ mixing ratios was found, regardless of which data set was used.

Jaffe *et al.* (2003b) reported that positive trends in O₃ mixing ratios were found in *all* seasons. Using the full data set, the trends were statistically significant in all seasons,

and they were statistically significant in summer regardless of which data set was used. The authors were not certain whether the trends in other seasons reflected changes in background O₃ or changes in regional photochemistry. The authors noted that, in particular, O₃ levels at Lassen in the summer were clearly affected by local in-situ production to a much greater extent than in spring.

Jaffe *et al.* (2003b) presented several arguments that indicated it was unlikely that local photochemical production of O₃ has made a significant contribution to the positive temporal trend in springtime background O₃. First, the authors believed that they had selected measurements from sites and meteorological conditions that minimized influence from North American emissions. Second, the authors hypothesized that photochemical production of O₃ was slow in Northern California in spring (i.e., March-May). Third, the trend is uniform at all sites including the aircraft data that probed well out over the Pacific Ocean. Fourth, tabulated inventories suggested that emissions of O₃ precursors have decreased significantly in the state of California over the past 15 years. Based on this analysis, the authors concluded that background O₃ in the eastern North Pacific during spring had increased. Although the Lassen Volcanic National Park site is not close to any major emission sources or urban centers, the site experiences maximum hourly average O₃ concentrations above 0.08 ppm during April – May and above 0.10 ppm during the summertime, suggesting local or regional photochemical production, at least during summer and possibly during the springtime.

Table 2-1 is a summary of the maximum hourly average concentrations experienced at Lassen during the 1988 – 2003 period. Elevated hourly average concentrations occur during April and May at the Lassen National Park monitoring

Table 2-1. Maximum hourly average concentrations measured at Lassen Volcanic National Park. (Source: EPA AQS Database)

Year	March	April	May
1988	.056	.067	.089
1989	.050	.072	.067
1990	.065	.071	.069
1991	.064	.062	.060
1992	.056	.056	.071
1993	.051	.058	.084
1994	.062	.085	.082
1995	.062	.066	.078
1996	.062	.066	.073
1997	.058	.072	.072
1998	.066	.078	.071
1999	.050	.085	.082
2000	.063	.066	.071
2001	.056	.068	.078
2002	.063	.058	.070
2003	.061	.063	.069

site. To investigate whether the source of these elevated levels may be associated with anthropogenic emissions in the regional area to the west of the park site, a similar table was prepared as above to describe the hourly maximum concentrations of O₃ that occur to the west of the Lassen site. Table 2-2 is a summary of the maximum hourly average concentrations experienced at a monitoring site during the 1999 – 2003 period at Redding (CA), which is west of the Lassen monitoring site. It appears that although the authors believed that photochemical production of O₃ is slow in northern California in spring, local photochemical production may be very important in the Redding, CA area during the springtime. Elevated hourly average concentration levels of O₃ occur in the areas to the west (i.e., Redding, CA and Anderson, CA) and to the south (Chico, CA; Tuscan, CA;

Table 2-2. Maximum hourly average concentrations measured at Redding, CA (060890004). (Source: EPA AQS Database)

Year	March	April	May
1999	.050	.101	.096
2000	.072	.073	.094
2001	.046	.054	.077
2002	.044	.060	.079
2003	.038	.043	.065

Red Bluff, CA) of the park site. As indicated from the hourly average data measured in Redding, levels in the 0.09 – 0.101 ppm range occur during the April-May periods at these locations and well above 0.10 ppm during the summertime. Although Jaffe *et al.* (2003b) believed that the Lassen Volcanic National Park site was not influenced in the spring (March-May) by anthropogenic sources, it appears that the site is influenced by the transport of O₃ resulting from anthropogenic emissions within the regional area.

If the most important source of O₃ at Lassen Volcanic National Park were associated with Asian emissions, one would have expected to observe an increase only during the March-May period (based on the authors' hypothesis). However, as noted in Jaffe *et al.* (2003b), the authors also observed a trend also during the summertime. Clearly, photochemistry is important in both spring and summer in the regional area near Lassen. Based on the monitored elevated levels of hourly average concentrations to the west of the Lassen monitoring site during the spring and the summer, one should quantify the importance of the regional photochemical activity. Although emissions of O₃ precursors may have decreased in California as a whole over the 1988 – 2003 monitoring

period, there still may be regional increases in areas that could affect air quality in Lassen. Although data appear to show reductions in emissions in the Redding area, it appears that these reductions have not reduced O₃ levels in the region. The CAAQSOD (page 7-79) notes that the Northeast Plateau Air Basin (NEPAB), which is in the northeast corner of California and includes Lassen, Modoc, and Siskiyou counties, has experienced a slight upswing during the last several years in the maximum concentrations for both the 1-hour and the 8-hour time periods.

The five other surface sites (Point Arena, CA; Redwood National Park; Point Reyes, CA; Cheeka Peak, WA; and Trinidad Head, CA) reported in Jaffe *et al.* (2003b) all lie within a few km of the coast. Because none of these monitoring sites provided a continuous record of more than 8 years, the authors combined the measurements from these sites and therefore obtained a *discontinuous*, 18-year record of springtime, background O₃ in the coastal marine boundary layer (MBL).

Figure 2-4 (from Jaffe *et al.*, 2003b) shows the trending of these sites. The authors reported a spring mean mixing ratio ± 1 standard deviation for background O₃ at 5 MBL sites with linear regression lines. The data have been selected by local wind direction and speed to minimize local influences. The linear fit to the data from the 4 sea level sites (solid line), yields a slope and year 2000 O₃ mixing ratio (with 95% confidence intervals) of 0.50 ± 0.36 ppbv/year, 39.9 ± 3.3 ppbv, and an r^2 of 0.44. If the higher altitude and latitude Cheeka Peak (WA) site is included (dashed line), the calculated slope and year 2000 mixing ratio become 0.78 ± 0.28 ppbv/year, 42.9 ± 2.4 ppbv, and the r^2 is 0.68.

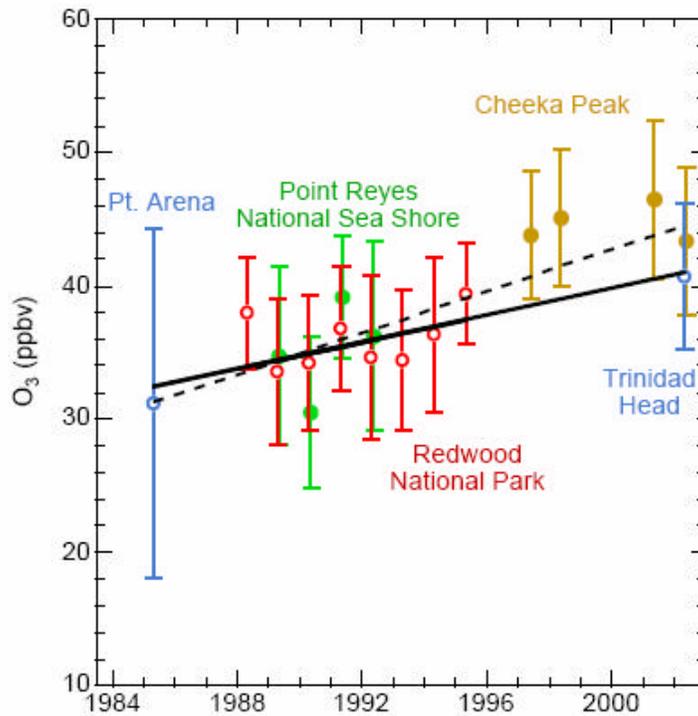


Figure 2-4. Spring mean mixing ratio ± 1 standard deviation for background O₃ at 5 MBL sites with linear regression lines. The data have been selected by local wind direction and speed to minimize local influences. Source: Jaffe *et al.* (2003b).

Unfortunately, there was no statistical rationale provided to explain why one could combine the data from the 5 monitoring sites and use them in the above figure to test for a trend. The only thing that the monitoring sites apparently have in common is that they are in the coastal marine boundary layer. Unless a convincing statistical argument is made using the hourly average concentration monitoring values to show that the data should be combined in the figure, using different monitoring sites, spread over a large geographic area over different periods of time, can only lead to suspect results. For example, if one eliminates the Point Arena, Trinidad Head, and Cheeka Peak data and investigates a trend for Point Reyes National Seashore and Redwood National Park, it is

not evident that a trend exists. The Point Arena data were collected over the period April 24 – May 5, 1985. This is a very short time period and may not reflect the springtime levels occurring during March - May 1985. Ozone concentrations vary from year to year because of meteorological variability. Jaffe *et al.* (2003b) presented no evidence that the data collected at Point Arena over the short monitoring period in 1985 was representative of the entire March-May period for that year (no data were collected) for all 5 monitoring sites. Therefore, no statistical rationale was made to justify the comparison of the Point Arena data with other data collected during other years at other monitoring sites along the coast. The data at Point Arena are only representative of the data collected at that time at that location. No rationale is provided by Jaffe *et al.* (2003b) as to why one can use Point Arena, Trinidad Head, or Cheeka Peak data to compare with Point Reyes or Redwood National Park data. The assumption is that because these are considered MBL sites, one should be able to compare the data with one another. Each monitoring site is unique in its distribution of O₃ hourly average concentrations. Unless a characterization is made that provides a statistical rationale as to why the sites can be compared, one should not place the data in a scatter diagram and try to fit a trending line between the data points.

One approach for evaluating the MBL hypothesis by the authors is to compare the *simultaneous* measurements of the hourly O₃ values. Point Reyes and Redwood NP did have simultaneous measurements during 1989 - 1992. Although the monitor at Redwood National Park collected data in 1988, the Point Reyes National Sea Shore monitor collected data in 1988 only for the months of October, November, and December.

Table 2-3 illustrates the median (50th percentile) hourly average concentration by month (using all of the hourly average concentration data).

Table 2- 3. 50th percentile hourly average concentration measured at Point Reyes N.S. and Redwood N.P. (Source: EPA AQS Database)

Year	March		April		May	
	Point Reyes	Redwood	Point Reyes	Redwood	Point Reyes	Redwood
1989	.032	.028	.034	.028	.034	.028
1990	.032	.030	.031	.028	.029	.030
1991	.039	.030	.039	.035	.036	.032
1992	.032	.028	.034	.030	.036	.026

Based on the review of data, it appears that the means for the Point Reyes National Sea Shore monitoring site are almost always greater than the means for the Redwood National Park monitoring site. Thus, the two monitoring sites exhibit different hourly average concentration distributions and there is no statistical rationale to combine the data for the two sites on the same graph. Although the two sites are characterized as MBL, this criterion alone does not provide a justification for including them in Figure 2-4 above to test for a trend. Although it may be possible that a trend exists, there is concern that the methodology utilized by the authors may result in an apparent trend that is not statistically justified. Again, because at least two of the monitoring sites (the only ones on the graph with an extended period of data collection) originate from different hourly average concentration distributions, summary data from these monitoring sites cannot simply be placed in a diagram and regressed.

The authors have apparently assumed that to determine marine O₃ levels, it is sufficient to select periods of relatively high onshore wind. The California Bay Area Quality Management District (Umeda, 1993) has observed that modeled particles released near the surface in the late afternoon around San Francisco Bay are occasionally carried offshore and southward at night, and some particles are brought near to the

Pinnacles during the following afternoon by the sea breeze. Recent modeling efforts in the South Coast Air Basin in California have indicated that offshore O₃ is apparently underestimated in models. Thus, the assumption that air from out of the west has originated in the west may be incorrect. Instead, as noted above, air from out of the west may have originated inland, been transported offshore, and returned onshore. The work by Jaffe *et al.* (2003b) does not present strong evidence for the long-range transport of O₃ from Asia.

The work by Lin *et al.* (2000) has been cited as additional evidence of the long-range transport of O₃ from Asia. Statistical analyses of U.S. O₃ data (Lin *et al.*, 2000) appeared to indicate that “background” effects were being driven by an increase in minimum O₃ concentrations at rural locations in the U.S. at the same time as peak O₃ concentrations are declining due to U.S. emission controls. They estimate a mean increase in U.S. “background” O₃ of 3-5 ppb between 1980 and 1998. The authors examined the long-term trend of background O₃ in surface air over the United States from 1980 to 1998, using monthly probability distributions of daily maximum 8-hour average concentrations at a large collection of rural sites in the AIRS database as shown in Table 2-4, and reported that O₃ concentrations decreased at the high end of the frequency distribution but increased at the lower end of the distribution. The increase was statistically significant at the 5% level in spring and fall. The authors hypothesized that the increase was due to the long-range transport of pollutants from Asia. The largest increase observed by Lin *et al.* (2000) was in the *northeastern United States*, which, as they noted, was inconsistent with the hypothesis that the increase was attributable to transport from Asia. It is interesting to note that many of the rural monitoring sites used

Table 2-4. AIRS monitoring sites used in the Lin *et al.* (2000) analysis.

AIRS ID	Site Name	County	Land Use	8-Hour Avg. (1998-2000)
010735002	JEFFERSON CO	JEFFERSON CO	RURAL – RESIDENTIAL	0.092
040132001	GLENDALE	MARICOPA CO	RURAL – RESIDENTIAL	0.075
040191018	TUCSON	PIMA CO	RURAL – DESERT	0.072
051191002	NORTH LITTLE ROCK	PULASKI CO	RURAL – FOREST	0.087
060012001	HAYWARD	ALAMEDA CO	RURAL – RESIDENTIAL	
060652002	INDIO	RIVERSIDE CO	RURAL -	0.090
060690002	HOLLISTER	SAN BENITO CO	RURAL – RESIDENTIAL	0.073
060710005	CRESTLINE	SAN BERNARDINO	RURAL – RESIDENTIAL	0.146
060731006	SAN DIEGO CO	SAN DIEGO CO	RURAL – RESIDENTIAL	0.100
060830008	CAPITAN	SANTA BARBARA CO	RURAL -	0.063
060833001	SANTA BARBARA CO	SANTA BARBARA C	RURAL – AGRICULTURAL	0.066
061113001	EL RIO	VENTURA CO	RURAL – AGRICULTURAL	0.068
080013001	WELBY	ADAMS CO	RURAL – AGRICULTURAL	0.072
090131001	STAFFORD	TOLLAND CO	RURAL – FOREST	0.089
121035002	TARPON SPRINGS	PINELLAS CO	RURAL – RESIDENTIAL	0.082
132151003	MUSCOGEE CO	MUSCOGEE CO	RURAL – AGRICULTURAL	0.093
132470001	ROCKDALE CO	ROCKDALE CO	RURAL – AGRICULTURAL	0.111
171192007	MADISON CO	MADISON CO	RURAL – AGRICULTURAL	0.086
180970042	INDIANAPOLIS	MARION CO	RURAL – AGRICULTURAL	0.088
201730001	SEDGWICK CO	SEDGWICK CO	RURAL – AGRICULTURAL	0.077
210150003	BOONE CO	BOONE CO	RURAL – AGRICULTURAL	0.086
210670001	FAYETTE CO	FAYETTE CO	RURAL – AGRICULTURAL	0.076
210910012	HANCOCK CO	HANCOCK CO	RURAL – RESIDENTIAL	0.089
220331001	E. BATON ROUGE PAR	E. BATON RGE PA	RURAL – AGRICULTURAL	0.094
220470002	IBERVILLE PAR	IBERVILLE PAR	RURAL – RESIDENTIAL	0.088
220950002	ST JOHN BAP. PAR	ST JOHN BAP. PA	RURAL – INDUSTRIAL	0.087
230052003	CAPE ELIZABETH	CUMBERLAND CO	RURAL – RESIDENTIAL	0.077
240030014	ANNE ARUNDEL CO	ANNE ARUNDEL CO	RURAL – AGRICULTURAL	0.107
240251001	EDGEWOOD	HARFORD CO	RURAL – COMMERCIAL	0.100
240313001	ROCKVILLE	MONTGOMERY CO	RURAL – RESIDENTIAL	0.090
240330002	GREENBELT	PRINCE GEORGES	RURAL – AGRICULTURAL	0.099
250171801	SUDBURY	MIDDLESEX CO	RURAL – AGRICULTURAL	0.085
260370001	CLINTON CO	CLINTON CO	RURAL – AGRICULTURAL	0.079
260492001	GENESEE CO	GENESEE CO	RURAL -	0.086
260812001	KENT CO	KENT CO	RURAL – AGRICULTURAL	0.087
290470003	CLAY CO	CLAY CO	RURAL – RESIDENTIAL	0.086
290470005	CLAY CO	CLAY CO	RURAL – AGRICULTURAL	0.089
291831002	ST CHARLES CO	ST CHARLES CO	RURAL – AGRICULTURAL	0.094
291890006	ST LOUIS CO	ST LOUIS CO	RURAL – RESIDENTIAL	0.090
310550032	OMAHA	DOUGLAS CO	RURAL – AGRICULTURAL	0.071
340010005	ATLANTIC CO	ATLANTIC CO	RURAL – RESIDENTIAL	0.090
340071001	CAMDEN CO	CAMDEN CO	RURAL – COMMERCIAL	0.101
340190001	FLEMINGTON	HUNTERDON CO	RURAL – AGRICULTURAL	0.098
340273001	MORRIS CO	MORRIS CO	RURAL – AGRICULTURAL	0.096
350011012	BERNALILLO CO	BERNALILLO CO	RURAL – DESERT	0.075
350130008	DONA ANA CO	DONA ANA CO	RURAL – AGRICULTURAL	0.073
360310002	ESSEX CO	ESSEX CO	RURAL -	0.080
360631006	NIAGARA CO	NIAGARA CO	RURAL – AGRICULTURAL	0.085
360650004	ONEIDA CO	ONEIDA CO	RURAL – FOREST	0.073
361173001	WAYNE CO	WAYNE CO	RURAL – AGRICULTURAL	0.081

Table 2-4. (continued).

AIRS ID	Site Name	County	Land Use	8-Hour Avg. (1998-2000)
370810011	GUILFORD CO	GUILFORD CO	RURAL – RESIDENTIAL	0.094
371191005	MECKLENBURG CO	MECKLENBERG CO	RURAL – INDUSTRIAL	0.098
371191009	MECKLENBURG CO	MECKLENBERG CO	RURAL – AGRICULTURAL	0.104390230001
	CLARK CO	CLARK CO	RURAL – AGRICULTURAL	0.093
390610010	HAMILTON CO	HAMILTON CO	RURAL – INDUSTRIAL	0.085
391331001	PORTAGE CO	PORTAGE CO	RURAL – AGRICULTURAL?	0.093
391351001	PREBLE CO	PREBLE CO	RURAL – AGRICULTURAL	0.080
401430174	GLENPOOL	TULSA CO	RURAL – AGRICULTURAL	0.081
410050004	CLACKAMAS CO	CLACKAMAS CO	RURAL – AGRICULTURAL	0.072
410090004	COLUMBIA CO	COLUMBIA CO	RURAL – AGRICULTURAL	0.056
420430401	HARRISBURG	DAUPHIN CO	RURAL – COMMERCIAL	0.090
420990301	PERRY CO	PERRY CO	RURAL -	0.085
450150002	BERKELEY CO	BERKELEY CO	RURAL – INDUSTRIAL	0.081
450230002	CHESTER CO	CHESTER CO	RURAL – COMMERCIAL	0.088
450370001	EDGEFIELD CO	EDGEFIELD CO	RURAL – AGRICULTURAL	0.085
450791002	RICHLAND CO	RICHLAND CO	RURAL – AGRICULTURAL	0.095
470370026	NASHVILLE-DAVIDSON	DAVIDSON CO	RURAL – FOREST	0.091
470650028	CHATTANOOGA	HAMILTON CO	RURAL – FOREST	0.097
470651011	HAMILTON CO	HAMILTON CO	RURAL – AGRICULTURAL	0.097
471571004	SHELBY CO	SHELBY CO	RURAL – AGRICULTURAL	0.097
471632002	SULLIVAN CO	SULLIVAN CO	RURAL – RESIDENTIAL	0.091
471650007	SUMNER CO	SUMNER CO	RURAL – INDUSTRIAL	0.100
510410004	CHESTERFIELD CO	CHESTERFIELD CO	RURAL – RESIDENTIAL	0.087
510850001	HANOVER CO	HANOVER CO	RURAL – AGRICULTURAL	0.100
530330010	KING CO	KING CO	RURAL – FOREST	0.063
551390011	OSHKOSH	WINNEBAGO CO	RURAL – AGRICULTURAL	0.076

from the AIRS database included in the Lin *et al.* (2000) analysis (e.g., Crestline [CA], Rockdale [GA], Edgewood [MD], Camden [NJ], Guilford [NC], and Sumner [TN]) are heavily influenced by transport from polluted areas: 66 percent of the O₃ monitoring sites used in the Lin *et al.* (2000) analysis exceeded the national 8-hour O₃ standard for the 3-year period 1998-2000. On page 3-4 of the CAAQSOD, the authors point out that, for the South Coast Air Basin, the highest O₃ concentrations are observed in the San Bernardino Mountains, where the Crestline monitoring site is located. It should be noted that in the AIRS database, the land use designation of *rural* does not mean that the site is mostly isolated from the long-range transport of episodic occurrences of O₃ concentrations that occur in or near urban areas. A land use characterization of “rural” does not imply that a specific location is

isolated from anthropogenic influences. Based on the methodology used by Lin *et al.* (2000), the reported results should not be used as evidence that long-range transport from Asia is influencing surface O₃ concentrations in the United States.

The CAAQSOD on page 4-11 indicates that long-range transport of anthropogenic O₃ may grow as Asian energy consumption increases the continent's NO_x emissions. Modelling studies, which have large uncertainties, indicate that the Asian O₃ increment in North America could double over the next few decades. The authors of the CAAQSOD believe that, assuming the temporal pattern of transport remains unchanged, such an *estimated* impact might increase mean O₃ concentrations by 2-6 ppb, with the potential effect on peak transport events unknown. However, given the limitations discussed earlier with the Lin *et al.* (2000) and Jaffe *et al.* (2003) trending analyses, the scientific evidence for an Asian influence on surface O₃ concentrations on the United States is weak and further research efforts are required.

2.7 The Importance of Exogenous O₃ Sources

The CAAQSOD stresses that the exogenous O₃ sources discussed in Chapter 4 are generally not major contributors to observed peak O₃ concentrations in California. This is thought to be mainly due to a seasonal mismatch between peak transport and local O₃ maxima. The CAAQSOD emphasizes that the violations associated with the proposed 0.07 ppm 8-hour average standard would occur during the summertime, when stratospheric O₃ contributions are thought to be minimal. However, when one characterizes the hourly average concentrations collected in 2003 for 184 monitoring sites in California, one finds that violations of the proposed 8-hour average standard occur during the springtime, summer, and

fall. Table 2-5, while not describing all violations that occurred in California during 2003, provides an indication that violations occurred outside of the summertime.

Table 2-5. Period of time when violations occur.

AIRS ID (m)	Year	Location	Month	Day	Start Hour	8-Hour Average	Elev.
060170012	2003	ECHO SUMMIT	5	15	22	0.079	2250
060170012	2003	ECHO SUMMIT	5	16	1	0.076	2250
060250006	2003	CALEXICO - EAST	5	24	12	0.078	0
060250006	2003	CALEXICO - EAST	3	12	10	0.075	0
060270101	2003	DEATH VLY NM	5	23	10	0.084	125
060270101	2003	DEATH VLY NM	5	24	10	0.084	125
060270101	2003	DEATH VLY NM	4	10	12	0.080	125
060730006	2003	SAN DIEGO	10	19	11	0.083	135
060730006	2003	SAN DIEGO	5	20	11	0.078	135
060730006	2003	SAN DIEGO	3	9	11	0.074	135
060834003	2003	VANDENBERG AFB	10	24	14	0.077	104
060834003	2003	VANDENBERG AFB	10	27	16	0.074	104
060834003	2003	VANDENBERG AFB	3	30	14	0.071	104

Violations of the 8-hour standard occurred (outside of the summer months) during the months of March, April, and October. Thus, it appears that one will have to characterize *policy-relevant background* concentrations that occur outside of the summertime so that emission control actions result in optimum reductions in hourly average O₃ concentrations. Besides regulatory control considerations, the distribution of hourly average *policy-relevant background* concentrations is important for epidemiological analyses, risk assessments, and clinical human health and vegetation experiments. On page 3-6, the CAAQSOD states that California is particularly prone to springtime stratospheric O₃ intrusions and thus, risk

assessments involving seasons other than summertime will have to use realistic *policy-relevant background* values that reflect the contribution from stratospheric O₃ as well as other natural sources. In addition, researchers implementing clinical human health and vegetation experiments will require knowledge of the variability of the *policy-relevant background* hourly average concentrations over seasons when violations of the proposed 8-hour average occur. Those involved in epidemiological analysis should be advised of the range of *policy-relevant background* concentrations so that statistical models provide realistic predictions. Thus, it is important for the CAAQSOD to provide estimates of the variability of *policy-relevant background* hourly average concentrations during seasons when elevated levels of O₃ occur in California.

The document mentions that stratospheric O₃ intrusions generally occur as a result of large-scale atmospheric disturbances, conditions that are inimical to the stable, stagnant conditions necessary to support buildup of pollutants in an urban area. The document (page 4-9) states that the probability of stratospheric O₃ adding to a high O₃ concentration due to anthropogenic emissions is low. In addition, stratospheric O₃ events can be recognized by unique atmospheric chemistry [i.e., very dry air, low aerosol concentrations, a general lack of anthropogenic precursor gases (VOCs and NO_x), and very low carbon monoxide (CO) concentrations compared to typical urban plumes].

The CAAQSOD concludes that the above chemical signature should make it possible to recognize stratospheric intrusion and to classify associated O₃ concentrations as “exceptional events”. However, there are times when there is downward mixing of stratospheric O₃ with the troposphere and the resulting mixture is later transported down under optimum meteorological conditions to the surface, with the result that hourly average

concentrations in the 0.05 to 0.06 ppm range occur. As indicated by Lefohn *et al.* (2001), these occurrences are more frequent than the natural *episodic* events (i.e., high hourly average concentrations experienced over short time periods) originating from the stratosphere. Thus, it may not be possible for regulators to classify these occurrences as “exceptional events” because the hourly average concentrations in the 0.05 – 0.06 ppm range would not necessarily be distinguishable from hourly average concentrations in the same range that were associated with anthropogenic emissions.

As indicated by Lefohn *et al.* (2001) and the CAAQSOD, one would anticipate that the higher-elevation sites would be expected to be influenced more by stratospheric O₃ than the lower-elevation sites. As pointed out earlier, Oltmans (personal communication) indicated that based on the ozonesonde measurements taken at Trinidad Head, mountainous terrain to the east of the area might experience elevated O₃ concentrations in the springtime that were associated with important contributions from the stratosphere. Unfortunately, monitoring data were not available in the mountainous terrain to the east of Trinidad Head. However, monitoring data for Redding (060890004), Anderson (060890007), and Lassen Volcanic National Park (060893003) offer an opportunity to investigate the O₃ exposures that are experienced at different elevations. The Lassen monitoring site appears to be influenced by anthropogenic sources to the west during both springtime and summertime. Elevated hourly average concentration levels of O₃ in Redding have been recorded in the 0.09 – 0.101 ppm range during the April-May periods. However, if it were possible to identify a year when minimum occurrences of hourly average concentrations greater than or equal to 0.05 ppm occur in the Redding area during the springtime, it might be possible to identify the possible influence of stratospheric O₃ on the exposure patterns at the Lassen monitoring site.

It appears that the O₃ exposures in the Redding area in the spring of 2003 were fairly low and provided an opportunity to investigate the springtime exposures at Lassen National Park.

Hourly averaged data were characterized for the Redding (149 m) and Anderson (498 m) monitoring sites for 2003 and compared to the Lassen site (1788 m) in an effort to identify whether the lower elevation monitoring sites experienced fewer hourly average concentrations greater than or equal to 0.05 ppm during the springtime months. Figure 2-5 illustrates the number of hourly average concentrations greater than or equal to 0.05 ppm for the three monitoring sites by month. In 2003, for the months March – June, Lassen Volcanic National Park experienced many more hourly average concentrations greater than or equal to 0.05 ppm than either of the two lower elevation monitoring sites to the west. However, during the July – September period, the monitoring site in Anderson experienced more frequent occurrences of hourly average concentrations greater than or equal to 0.05 ppm. This implies that the Anderson site (0.096 8-hour average) is more influenced by anthropogenic sources than either the Lassen (0.073 ppm 8-hour average) or Redding (0.076 ppm 8-hour average) sites.

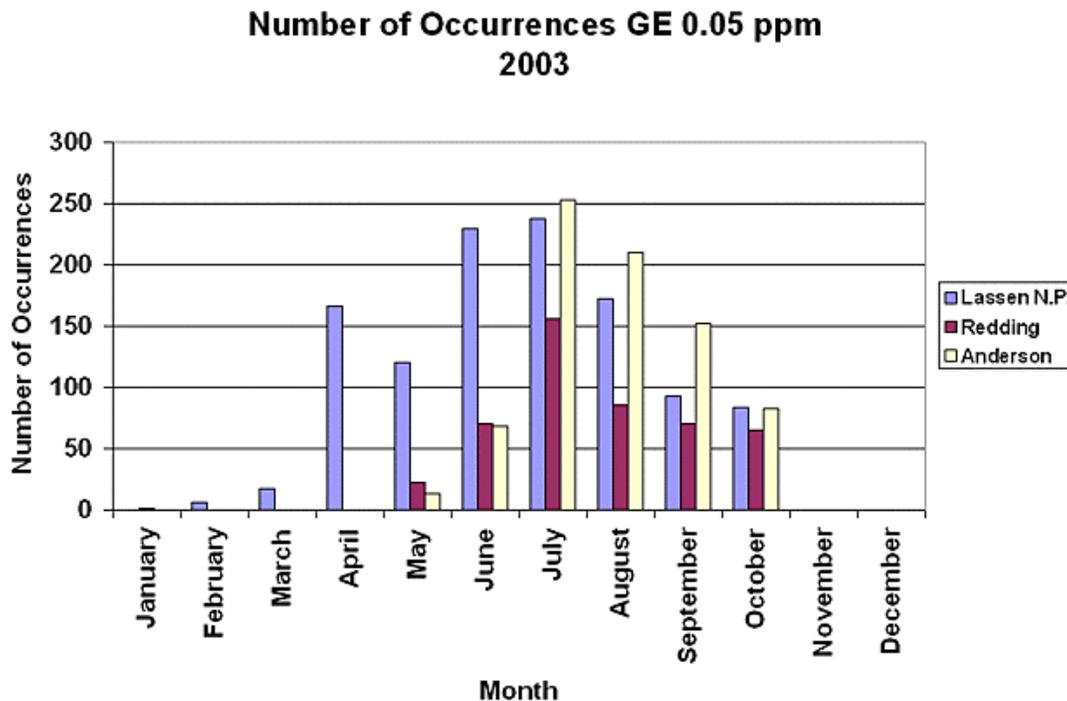


Figure 2-5. Comparison of the number of hourly average concentrations ≥ 0.05 ppm for monitoring sites located at Lassen National Park, Redding, and Anderson.

Based on the characterized monitoring data for 2003, it appears that the Lassen Volcanic National Park may have been influenced during the March – June period by stratospheric O_3 . However, an alternative explanation might be that the meteorological situation in Redding and Anderson might not have been conducive to photochemical O_3 production. Transport of precursors from these sites may have allowed the production of O_3 in route to Lassen National Park. One would have to perform additional analyses to confirm this possibility. As noted on page 4-9, the CAAQSOD mentions the possibility of long-range transport of anthropogenic O_3 from Asia. The evidence for the long-range transport from Asia affecting O_3 hourly average concentrations is not very strong at this time. However, if one were to hypothesize that the long-range transport from Asia were influencing the frequency of 0.05 ppm and above hourly average concentrations occurring at the Lassen site

during the spring of 2003, one would have to conclude that the Asian influence appeared to be minimal at the Redding and Anderson monitoring sites. In other words, the O₃ associated with the long-range transport from Asia remained aloft until intercepted by the Lassen monitoring site. However, Jaffe *et al.* (2003b) included the 5 marine boundary layer sites in their trends analysis, which implied that the lower-elevation sites were thought to be influenced by the long-range transport from Asia. Thus, based on the assumptions of Jaffe *et al.* (2003b), one would have anticipated that the monitoring sites at Redding and Anderson would have been influenced by the long-range transport from Asia. Because only the monitoring site at Lassen in 2003 experienced the enhancement in hourly average concentrations greater than or equal to 0.05 ppm, it is assumed that the enhancement is associated with stratospheric O₃. Therefore, it may be concluded that at some monitoring sites in California, when stratospheric O₃ predominates in comparison to anthropogenic sources, it may *not* be possible for regulators to control hourly average concentrations in the 0.05 – 0.06 ppm range, using emission reduction strategies.

On page 4-9, the CAAQSOD mentions the possibility of long-range transport of anthropogenic O₃ from Asia. The document points out that downward mixing of Asian O₃ to the surface is not possible when a strong surface inversion is present, thus locally generated high O₃ concentrations are unlikely to be enhanced by long-range transport. The document notes that there is also a seasonal mismatch between the peak of long-range transport and California's O₃ seasons. The CAAQSOD notes that periods of effective long-range transport are generally restricted to late winter and spring (Berntsen, Karlsdottir and Jaffe, 1999), while high O₃ due to local sources in California tend to occur in late summer and fall. Based on the limitations associated with the methodologies implied by Jaffe *et al.* (2003b) and Lin

et al. (2000), it is premature for the CAAQSOD to be concerned that surface O₃ levels are currently influenced by long-range transport from Asia.

2.8 Estimating Policy-Relevant Background Concentrations

In estimating the definition of *policy-relevant background* on page 4-1 of the CAAQSOD, one will have to define a range of concentrations due to *external* or *uncontrollable* sources that may impact determinations of compliance with air quality standards or limit the potential air quality improvements due to control programs. In the Fiore *et al.* (2003) analysis, the investigators removed all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer, but not biomass burning) within North America from their model before estimating natural background. Because the State of California does not plan to eliminate all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer), the estimates for the range of *policy-relevant background* concentrations will be greater than the background (i.e., natural background, in North America and anthropogenic and natural background outside of North America) values hypothesized by Fiore *et al.* (2003).

The CAAQSOD concludes on page 4-11 that “background” O₃ in California is dominated by natural tropospheric and stratospheric processes. The effects of occasional very large biomass fires and anthropogenic emissions are secondary factors. CAAQSOD appears to be willing to accept the modeling predictions of Fiore *et al.* (2002, 2003). However, the predicted “natural background” O₃ range of 15-35 ppb near sea-level, with a maximum of about 40 ppb, are highly uncertain. The CAAQSOD on page 4-11 states that exogenous

enhancements to “natural” levels generally are small (about 5 ppb), and are unlikely to alter peak concentrations. The document states that at altitudes above 2 km, stratospheric intrusions can elevate peak ambient concentrations to 45-50 ppb. It is unclear if these are 4-hour afternoon average concentrations as estimated by Fiore *et al.* (2002, 2003) or hourly average values. If these values are 4-hour afternoon average concentrations, then the hourly average concentrations as indicated above will be higher.

The authors of the document believe that the timing, spatial extent, and chemical characteristics of stratospheric air mass intrusions make these events recognizable in air quality records, providing that the affected region has a fairly extensive monitoring network and that multiple air quality parameters (CO, VOC, PM, RH) were being measured as well. *However, such is not the case when stratospheric injections into the troposphere have the opportunity to mix with background air prior to being transported down to the surface at a later time.* Although the authors believe that elevations below 2 km may not be affected by stratospheric O₃, it appears that Lassen Volcanic National Park (1788 m) is affected in the springtime by stratospheric O₃. Although a paucity of data exist in California, it may be possible that monitoring sites at elevations below 1788 m are also affected by stratospheric O₃.

The Fiore *et al.* (2002, 2003) model predicted 4-hour afternoon average concentrations. Given the (1) different definitions used for background by CAAQSOD and Fiore *et al.* (2002, 2003), (2) low spatial resolution of the model resulting in large uncertainties, and (3) multi-hour averaging (4-hour afternoon averages) of the hourly average concentrations, the hourly average concentrations of *policy-relevant background* near sea level and at higher elevations are more than likely higher than provided by CAAQSOD.

Fiore *et al.* (2003) stated that natural background levels never exceed 0.04 ppm. Given the empirical observations during springtime in which the influence of stratospheric O₃ on surface exposures at various locations in the world has been documented, it appears that the modeling results published by Fiore *et al.* (2002, 2003) should to be used with extreme caution.

From a regulatory perspective, it is important to identify the variability of O₃ hourly average concentrations associated emissions that are beyond the reach of regulation. These concentrations may impact determinations of compliance with air quality standards or limit the potential air quality improvements due to control programs. In addition, human health and vegetation researchers should be applying these concentrations in their controls so that the effects models that result from their investigations reflect true comparison ambient background conditions. As indicated earlier, if such realistic exposures are not used, it may be possible that overestimations in effects will occur in the modeling predictions.

As indicated earlier, empirical data collected at monitoring sites that are far removed from urban sources indicate that hourly average concentrations are much higher than the values reported by Fiore *et al.* (2003) for background O₃. However, the empirical data provide a solid indication to CAAQSOD that *policy-relevant background O₃* as defined on page 4-1 is more than likely higher than the 15-35 ppb. Using models that provide highly uncertain concentration estimates provides an overly optimistic message to those who are responsible for implementing control strategies.

In fairly pristine locations in the United States, such as Yellowstone National Park in Wyoming, violations of the proposed California 8-hour standard of 0.070 ppm occur during the springtime. The ambient concentrations experienced at Yellowstone National Park in the

springtime represent *policy-relevant background* as defined in Chapter 4 of the CAAQSOD. This implies that the proposed 8-hour standard will be difficult to attain in areas that are affected by stratospheric O₃ during the springtime and that perhaps the methodology used by Staff to propose the form and level of the 8-hour standard provides highly uncertain results.

2.9 Conclusion

The CAAQSOD states that from a regulatory perspective, the important distinction is not between “natural” and “anthropogenic” O₃, but between O₃ produced by controllable emissions and O₃ due to emissions beyond the reach of regulation. Anthropogenic O₃ produced outside the jurisdiction of an agency and transported into a control region is functionally indistinguishable from that due to natural processes. Within the range of concentrations due to such external or uncontrollable sources, those concentrations that may impact determinations of compliance with air quality standards or limit the potential air quality improvements due to control programs have been defined by the CAAQSOD as *policy-relevant background*.

Establishing ranges for hourly average O₃ *background* concentrations is important for several reasons. In addition to identifying the range of *policy-relevant background* concentrations to better understand the relationship between emission reductions and resultant O₃ exposure patterns, human health and vegetation researchers need to use the range of *policy-relevant background* concentrations in their controls so that the models that result from their investigations reflect true comparisons with ambient background conditions. If such is not implemented, then it may be possible that overestimations will occur in the biological modeling predictions.

Estimates of *policy-relevant background* concentrations need to consider the important contribution from stratospheric O₃, as well as other natural sources. There is a large variability among global models on the attribution of the contribution of natural O₃ to the background. Although the CAAQSOD states that ground level impacts from fires are typically in the range of 15-25 ppb, such is not necessarily the case. It is premature for the CAAQSOD to be concerned that surface O₃ levels in California are currently influenced by the long-range transport from Asia. The scientific evidence for an Asian influence on surface O₃ concentrations on the United States is weak and further research efforts are required.

The CAAQSOD emphasizes that the violations associated with the proposed 0.07 ppm 8-hour average standard would occur during the summertime, when stratospheric O₃ contributions are thought to be minimal. However, when one characterizes the hourly average concentrations collected in 2003 for 184 monitoring sites in California, one finds that violations of the proposed 8-hour average standard occur during spring, summer, and fall. This implies that *policy-relevant background* concentrations that occur during seasons other than summertime will have to be characterized so that emission control actions result in optimum reductions in hourly average O₃ concentrations. At some monitoring sites in California, when stratospheric O₃ predominates in comparison to anthropogenic sources during the springtime, it may *not* be possible for regulators to control hourly average concentrations in the 0.05 – 0.06 ppm range using emission reduction strategies.

The empirical data provide a solid indication to CAAQSOD that *policy-relevant background* O₃ is more than likely higher than the 15-35 ppb discussed in the document. Using models that provide highly uncertain concentration estimates provides an overly optimistic message to those who are responsible for implementing control strategies. In some

of the modeling efforts to estimate natural background O₃ concentrations within North America, investigators removed all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer, but not biomass burning). Because the State of California does not plan to eliminate all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer), the estimates for the range of *policy-relevant background* concentrations will be greater than the background (i.e., natural background, in North America and anthropogenic and natural background outside of North America) values estimated by these models.

2.10 References

- Altshuller, A. P.; Lefohn, A. S. (1996) Background ozone in the planetary boundary layer over the United States, *J. Air Waste Manage. Assoc.*, 46: 134-141.
- Berntsen, T.; Karlsdottir S.; Jaffe D. (1999) Influence of Asian emissions on the composition of air reaching the North Western United States. *Geophys Res Lett* 26(14):2,171-2,174.
- Bojkov, R. (1986) Surface ozone during the second half of the nineteenth century. *Mon Weather Rev* 25:353-352.
- Collins, W.J.; Derwent, R.G.; Johnson, C.E.; Stevenson, D.S. (2000) The impact of human activities on the photochemical production and destruction of tropospheric ozone, *Q. J. R. Meteorol. Soc.* 126: 1925-1951.
- Cooper, O. R.; Moody, J. L.; Davenport, J. C.; Oltmans, S. J.; Johnson, B. J.; Chen, X.; Shepson, P. B.; Merrill, J. T. (1998) Influence of springtime weather systems on vertical ozone distributions over three North American sites. *J. Geophys. Res.* 103: 22,000-22,013.
- Cristofanelli, P. et al. (2003) Stratosphere-troposphere exchange: A model and method evaluation, *J. Geophys. Res.* 108(D12), 8525, doi:10.1029/2002JD002600.
- Crutzen, P. (1995) My life with O₃, NO_x and other YZOXs. Nobel Lecture, December 8, 1995.

- Derwent, R. G.; McInnes, G.; Steward, J. N. M.; Williams, M. L. (1976) The occurrence and significance of air pollution by photochemically produced oxidant in the British Isles: 1972-1975, Report LR 227 (AP), Warren Spring Lab., Stevenage, U.K.
- Fiore, A. M.; Jacob, D. J.; Bey, I.; Yantosca, R. M.; Field, B. D.; Fusco, A. C.; Wilkinson, J. G. (2002). Background ozone over the United States in summer: Origin, trend, and contribution to pollution episodes. *J. Geophys. Res.* 107(D15): ACH 11-1 – 25.
- Fiore, A.; Jacob, D. J.; Liu, H.; Yantosca, R. M.; Fairlie, T. D.; Li, Q. (2003) Variability in surface ozone background over the United States: implications for air quality policy. *J. Geophys. Res. (Atmos.)*: 108: D24, 4787,ACH 19-1 – 19-10.
- Fusco, A. C.; Logan, J. A. (2003) Analysis of 1970-1995 trends in tropospheric ozone at Northern Hemisphere midlatitudes with the GEOS-CHEM model. *J. Geophys. Res. (Atmos.)*: 108, NO. D15, 4449, doi:10.1029/2002JD002742, ACH 4-1 – 4-25.
- Jaegle, L.; Jaffe D.; Price H.; Weiss-Penzias P.; Palmer P.; Evans M.; Jacob D.; Bey, I. (2003) Sources and budgets for CO and O₃ in the northeastern Pacific during the spring of 2001: Results from the PHOBEA-II Experiment. *J Geophys Res* 108(D20).
- Jaffe, D. et al. (1999). Transport of Asian air pollution to North America. *Geophys Res Lett* 26(6):711-714.
- Jaffe, D.; Price H.; Parrish, D.; Goldstein, A; Harris, J. (2003b) Increasing background ozone during spring on the west coast of North America. *Geophys Res Lett* 30(12).
- Jaffe, D.; McKendry, I.; Anderson, T.; Price, H. (2003) Six 'new' episodes of trans-Pacific transport of air pollutants. *Atmos Env* 37m 391-404.
- Jaffe, D.; Price, H.; Parrish, D.; Goldstein, A.; Harris, J. (2003b) Increasing background ozone during spring on the west coast of North America. *Geophys Res Lett* 30(12).
- James, P.; Stohl, A.; Forster, C.; Eckhardt, S.; Seibert, P.; Frank, A. (2003a) A 15-year climatology of stratosphere-troposphere exchange with a Lagrangian particle dispersion model: 1. Methodology and validation. *J. Geophys. Res.* 108(D12), 8519, doi:10.1029/2002JD002637.
- James, P.; Stohl, A.; Forster, C.; Eckhardt, S.; Seibert, P.; Frank, A. (2003b) A 15-year climatology of stratosphere-troposphere exchange with a Lagrangian particle dispersion model: 2. Methodology and validation. *J. Geophys. Res.* 108(D12), 8522, doi:10.1029/2002JD002639.
- Jenkins, G.; Ryu, J-H. (2003) Linking horizontal and vertical transport of biomass fire emissions to the Tropical Atlantic Ozone Paradox during the Northern Hemisphere winter season: climatology. *Atmos Chem Phys Discuss* 3, 5,061-5,098.

- Lefohn, A. S.; Shadwick, D. S.; Feister, U.; Mohnen, V. (1992) Surface-level ozone: Climate change and evidence for trends. *J. Air Waste Manage. Assoc.* 42: 136-144.
- Lefohn, A. S., Oltmans, S. J.; Dann, T.; Singh, H. B. (2001) Present-day variability of background ozone in the lower troposphere. *J. Geophys. Res.* 106(D9): 9945-9958.
- Li, Q.; Jacob, D.J.; Fairlie, T.D.; Liu, H.; Yantosca, R.M.; Martin, R.V. (2002) Stratospheric versus pollution influences on ozone at Bermuda: Reconciling past analyses, *J. Geophys. Res.*, 107, 10.1029/2002JD00213.
- Lelieveld, J.; Dentener, F. (2000). What controls tropospheric ozone?. *J Geophys Res* 105(D3).
- Lin, C-Y.; Jacob, D.; Munger, J; Fiore, A. (2000) Increasing background ozone in surface air over the United States. *Geophys Res Lett* 27(21):3,465-68.
- Lisac, I; Grubiši, V. (1991) An analysis of surface ozone data measured at the end of the 19th century in Zagreb, Yugoslavia. *Atmos. Environ.* 25:481-486 (1991).
- Logan, J. (1989) Ozone in rural areas of the United States, *J. Geophys. Res.* 94: 8511-8532.
- Marenco, A.; Gouget, H.; Nédélec, P.; Pagés, J.-P.; Karcher, F. (1994) Evidence of a long-term increase in tropospheric ozone from Pic du Midi data series: Consequences: Positive radiative forcing. *J. Geophys. Res.* 99: 16617-16632.
- Meloan, J., et al. (2003) Stratosphere-troposphere exchange: A model and method intercomparison, *J. Geophys. Res.* 108(D12), 8526, doi:10.1029/2002JD002274.
- Moody, J. L.; Oltmans, S. J.; Levy II, H.; Merrill, J. T. (1995) Transport climatology of tropospheric ozone: Bermuda, 1988-1991. *J. Geophys. Res.* 100: 7179-7194.
- Newchurch, M. J., et al. (2003) Vertical distribution of ozone at four sites in the United States. *J Geophys Res* 108(D1).
- Olsen, M.A.; Douglass, A.R.; Schoeberl, M.R. (2002) Estimating downward cross-tropopause flux using column ozone and potential vorticity. *J. Geophys. Res.* 107(D22), 4636, doi:10.1029/2001JD002041.
- Oltmans, S. J. (personal communication). Mr. Oltmans, NOAA/CMDL, Boulder, Colorado provided the data to A.S. Lefohn.
- Oltmans, S.J.; Levy II, H. (1992) Seasonal cycle of surface ozone over the North Atlantic, *Nature*, 358, 392-394, 1992.
- Oltmans, S. J.; Levy II, H.; Harris, J. M.; Merrill, J. T.; Moody, J. L.; Lathrop, J. A.; Cuevas, E.; Trainer, M.; O'Neill, M. S.; Prospero, J. M.; Vömel, H.; Johnson, B. J. (1996)

- Summer and spring ozone profiles over the North Atlantic from ozonesonde observations. *J. Geophys. Res.* 101: 29,179-29,200.
- Oltmans, S. J.; Johnson, B.; Harris, J.; Ives, M.; Snibel, W. (personal communication). Mr. Oltmans, NOAA/CMDL, Boulder, Colorado provided the data to A.S. Lefohn.
- Pavelin, E. G.; Johnson, C. E.; Rughooputh, S.; Toumi, R. (1999). Evaluation of pre-industrial surface ozone measurements made using Schönbein's method, *Atmos. Environ* 33: 919-929.
- Singh, H. B.; Ludwig, F. L.; Johnson, W. B. (1978) Tropospheric ozone: Concentrations and variabilities in clean remote atmospheres. *Atmos. Environ.* 12: 2185-2196.
- Springer, M.; Wernli, H. (2003) A northern hemispheric climatology of cross-tropopause exchange for the ERA 15 time period (1979-1993), *J. Geophys. Res.* 108(D12), 8521, doi:10.1029/2002JD002636.
- Stohl, A.; Spichtinger-Rakowsky, N.; Bonasoni, P.; Feldmann, H.; Memmesheimer, M.; Scheel, H. E.; Trickl, T.; Hubener, S.; Ringer, W.; Mandl, M. (2000) The influence of stratospheric intrusions on alpine ozone concentrations. *Atmos. Environ* 34: 1323-1354.
- Stohl, A. et al. (2003) Stratosphere-troposphere exchange: A review and what we have learned from STACCATO, *J. Geophys. Res.* 108(D12), 8516, doi:10.1029/2002JD002490.
- Thompson, A.M. (1992) The oxidizing capacity of the Earth's atmosphere: Probable past and future changes. *Science* 256:1157-1165.
- Umeda, T. (1993) Monterey Bay Area Transport Study Final Report. Vol 2. Trajectory Analysis. Bay Area Air Quality Management District. San Francisco, CA.
- VanCuren, R.; Cahill, T. (2002) Asian aerosols in North America: Frequency and concentration of fine dust. *J Geophys Res* 107(D24).
- Volz, A.; Kley, D. (1988) Evaluation of the Montsouris series of ozone measurements made in the nineteenth century. *Nature.* 332: 240-242.
- Wakamatsu, S.; Uno, I.; Ueda, H.; Uehara, K.; Tateishi, H. (1989) Observational study of stratospheric ozone intrusions into the lower troposphere. *Atmos. Environ.* 23: 1815-1826.
- Yienger, J. J.; Klonecki, A. A.; Levy II, H.; Moxim, W. J.; Carmichael, G. R. (1999) An evaluation of chemistry's role in winter-spring ozone maximum found in the northern midlatitude free troposphere. *J. Geophys. Res.* 104: 3655-3667.

Ziman, S. (personal communication). ChevronTexaco Energy Technology Company.
Richmond, CA.

3. Human Health Effects – Controlled Exposure Studies

3.1 Introduction

Experimental exposures of human volunteers to air pollutants under controlled laboratory conditions can provide useful pathophysiological information directly relevant to standard setting. The primary outcome assessed in human-controlled exposure studies is lung function with the amount of air that can be exhaled during the first second of a forced expiratory maneuver (forced expiratory volume in one second or FEV1), a metric that can reflect lung obstruction, being key. Other lung function metrics and respondent self-report of symptoms are other common outcomes. Typically, outcomes are assessed before and after an exposure regime to the pollutant(s) of interest has occurred. The change in function is then compared with that obtained before and after exposures to control conditions, i.e., filtered air (FA).

Controlled human exposure studies are superior to epidemiological studies for purposes of establishing exposure-response relationships. This is because exposures to the pollutant(s) of interest can be precisely controlled and measured, and therefore exposure accuracy and health effects closely reflecting ambient exposure scenarios can be obtained. Because epidemiology involves observational rather than experimental studies, estimating personal exposure levels and controlling potential confounding factors are extremely difficult. The uncertainty inherent in estimating these study factors translates into uncertainty about the causal nature of any statistical associations between exposure and health outcome that might be observed. As such, the current set of epidemiological studies regarding O₃ is not appropriate for use in identifying causal effects or estimating risks.

Despite the advantages of controlled human exposure studies over epidemiologic ones, there are some concerns about them, particularly with respect to the multi-hour O₃ exposure chamber studies as they relate to the proposed 8-hour standard. The authors of the CAAQSOD have missed important features of studies that have addressed some of the limitations noted above.

Specifically, there are concerns about the controlled exposure studies which include (1) using exposure scenarios that do not reflect *ambient* conditions, (2) the analytic scheme of determining health impacts by comparing the experimental pre-post (FEV1) results, and (3) using a reference of filtered air (i.e., 0 ppm concentration) as the control condition. In addition, the CAAQSOD needs to quantify the definition of an adverse health effect so that the standard-setting process is based on biologically meaningful differences and not just statistically significant numerical test result differences.

3.2 Exposure Scenario and Analytic Strategy

On page 11-75, the CAAQSOD notes that historically, most controlled studies of the effects of O₃ have used exposure protocols of one to two hours duration, primarily based on a diurnal O₃ concentration profile typical of the Los Angeles air basin, and the assumption that peak concentration, rather than cumulative dose, was the most important factor mediating O₃ toxicity. The document states that analysis of air quality data from other parts of the US, however, has shown another widespread pattern characterized by a six-to-eight hour period with a relatively constant O₃ concentration near but below that of the existing federal one-hour ambient standard (0.12 ppm) (US EPA, 1996). Based on the hypothesis that a broad plateau lasting as long as eight to twelve hours was important, “square-wave” or constant-

concentration profiles were used by researchers to investigate the effects of multi-hour exposure to an O₃ concentration lower than the 1-hour ambient air quality standard. Unfortunately, this plateau pattern does not accurately reflect a typical urban O₃ exposure profile and in fact, it is quite rare.

Lefohn and Foley (1993) analyzed 925 exposure regimes across the United States, identified from 166 site-years of data for the period 1987-1989, that met the following criteria: (1) the site never experienced an exceedance of an hourly average concentration equal to or greater than 0.12 ppm and (2) the site experienced 8-hour daily maximum average concentrations greater than 0.08 ppm. For those monitoring sites that met the above two criteria, the authors identified the number of times the 8-hour daily maximum average concentration exceeded 0.08 ppm during the monitoring year; the data were then organized into seven categories. In *no case* could Lefohn and Foley (1993) identify a monitoring site that experienced a “square-wave” or constant O₃ concentration type of exposure (i.e., the occurrence of 8-hour daily maximum averages greater than 0.08 ppm but less than or equal to 0.082 ppm, which contained only hourly average concentrations greater than 0.08 ppm but less than or equal to 0.082 ppm). Recently, Lefohn and Shadwick (personal communication), using all O₃ monitoring sites in the 2002 AIRS database with 8-hour average concentrations greater than or equal to 0.07 ppm, characterized the hour-by-hour changes in the 8-hour windows associated with the daily maximum 8-hour average O₃ value. Based on the results of the analysis that identified the characterization of the sequential patterns associated with the 8-hour exposures, Lefohn and Shadwick concluded that only 1.51% of the 28,184 sequential patterns could be classified as representing a “square wave” profile (arbitrarily defined as experiencing a range of 4 ppb or less). Thus, the occurrences of “square-wave

profiles” are rare. As such, health studies based on this exposure pattern are not helpful for estimating health impacts of typical ambient O₃ exposures.

Although the CAAQSOD on page 11-75 appears to accept without verification that there exists a broad plateau lasting as long as eight to twelve hours, in contrast to the sharp one- to two-hour peak identified earlier in the Los Angeles, CA area, it is important for the authors to characterize and present the hourly average monitoring data to support the hypothesis that “square-wave” profiles are more frequent in California than reported by Lefohn and Foley (1993) and Lefohn and Shadwick (personal communication) for the entire United States.

Hazucha *et al.* (1992) were the first to design a protocol utilizing 8-hour exposures with square and peaked O₃ concentration profiles. The study compared responses to exposure of a constant O₃ concentration (i.e., square-wave pattern) of 0.12 ppm, and to a variable concentration (i.e., triangle-shaped pattern) profile (linear increase from 0 to 0.24 ppm over four hours, followed by linear decrease from 0.24 to 0 ppm over 4 hours). The results are illustrated in Figure 3-1 (Figure 11-3 in the CAAQSOD), below as described on page 11-80.

The total inhaled effective dose of O₃ was equivalent for the two exposures (difference < 1%). Exposure to the constant O₃ concentration induced a group mean decrement in FEV1 of approximately 5% by the fifth hour of exposure, which did not change over the remainder of the exposure. In contrast, with the variable concentration protocol, the response over the first three hours was minimal, followed by a mean decrease in FEV1 over hours 4 through 6 that peaked at approximately 10%. There appears to be a lag in

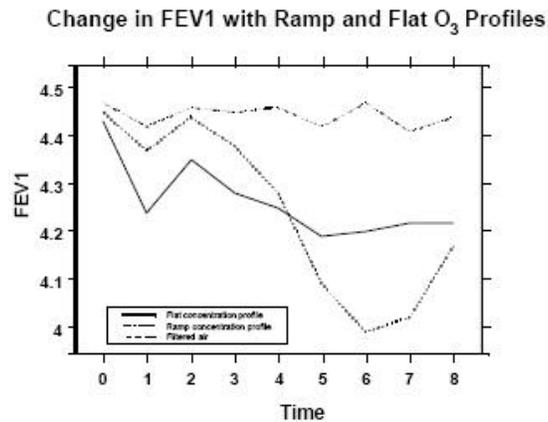


Figure 11-3 Comparison of the change in FEV1 with ramp and flat O₃ concentration profiles. Derived from Hazucha, 1993.

Figure 3-1. Change in FEV1 comparing square-wave exposure with the varying hourly-by-hour exposure. Source: Hazucha et al., 1992.

development of the maximal response, since the maximal O₃ concentration occurred at hour 4, yet the maximal FEV1 response occurred at hour 6. FEV1 improved during the last two hours of the exposure, and by the end of the exposure the FEV1 decrement was nearly identical to that following the constant concentration exposure.

Adams (2003a) also compared responses to square wave and triangular exposure scenarios using the same 8-hour average concentration of 0.08 ppm. He studied healthy adults who were exposed to O₃ in an environmental chamber and through a facemask system that allowed natural breathing while they completed a 6.6-hour protocol. The subjects inhaled FA and 0.08 ppm O₃ in the chamber and the face mask system. Two types of O₃ exposure patterns were used: (1) the usual square-wave profile, and (2) an acute triangular profile, in which O₃ concentration was changed each hour from 0.03 to 0.07 to 0.10 to 0.15 to 0.08 to 0.05 ppm. Thirty young adults (15 of each gender) served as subjects. The two exposure

methods, face mask and chamber, yielded near identical total inhaled O₃ doses, in both the square-wave and triangular exposure profiles, and produced very similar post-exposure pulmonary function, symptoms, and exercise ventilatory pattern responses. Figure 3-2 summarizes the results of the experiments. The protocol key is as follows: 1 -- chamber, FA; 2 -- face mask, FA; 3 -- chamber 0.08 ppm O₃, square wave; 4 -- face mask 0.08 ppm O₃, square wave; 5 -- chamber diurnal (peaked) O₃ with 8-hr average of 0.08 ppm; 6 -- face mask diurnal (peaked) O₃ with 8-hr average of 0.08 ppm.

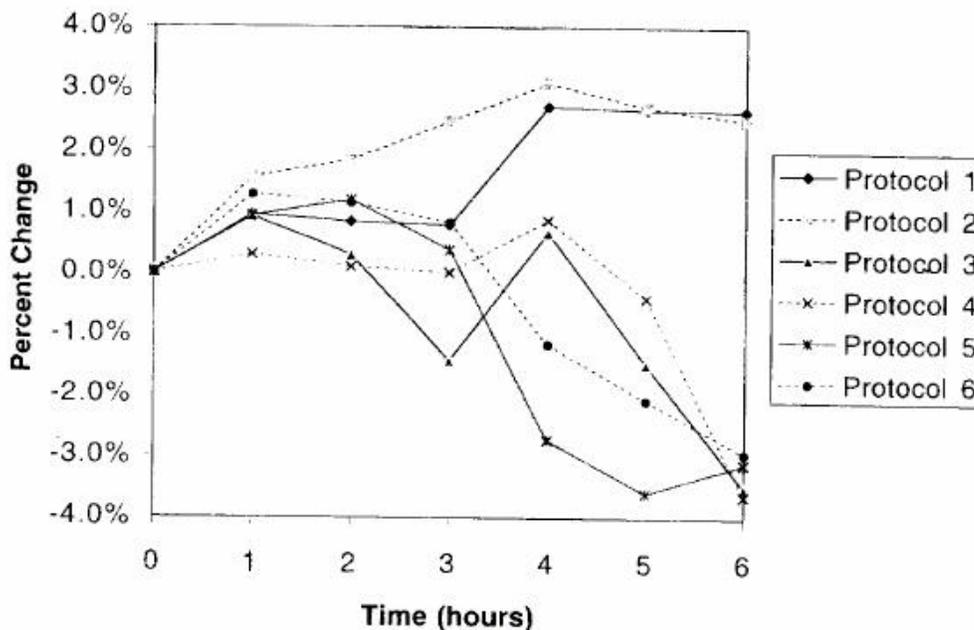


FIGURE 1. Hour-by-hour percent change in FEV_{1.0}.

Figure 3-2. Change in FEV1 comparing square-wave exposure with the varying hourly-by-hour exposure. Source: Adams et al. (2003a).

At the end of the 6.6-hour experiment, no statistical significant difference was observed for group mean changes in FEV1 between the square-wave (Protocols 3 and 4) or triangular exposures (Protocols 5 and 6). However, significant differences were observed when evaluating the hour-by-hour results. In reviewing Figure 3-2 above, it is clear that both

triangular exposure scenarios (Protocols 5 and 6) had greater percent change decrements in FEV 1.0 than the square-wave exposures. In addition, the triangular exposure used in the chamber (Protocol 5) is very similar in shape to the results published by Hazucha *et al.* (1992).

In both Hazucha *et al.* (1992) and Adams (2003a) the cumulative exposure in the corresponding square wave-triangular exposure protocols were essentially the same. Similarly, within each study the final group mean FEV1 decrements between the two exposure scenarios were very similar. However, significant differences between the square and triangular exposure scenarios were observed when comparing the hour-by-hour results. For both Hazucha and Adams, the triangular exposure scenario resulted in earlier onset of lung function decrements and possibly a greater number of people experiencing deficits than the square wave. In other words, the pre-post analytic strategy may miss adverse health effects, and the more realistic diurnal exposure scenario appears to be more sensitive to health impacts than the square wave profile used in most controlled exposure studies. These findings suggest that relying on pre-post measurement differences across the 8-hour period for developing the proposed standard could miss important health impacts. They also illustrate that the O₃ induced FEV1 response is dependent on the dose rate as well as the cumulative dose of O₃ inhaled, at least when the O₃ concentration is variable such as in the case of typical urban area diurnal exposure patterns (Adams, 2003a, Lefohn and Foley, 1993, Lefohn and Shadwick (personal communication)).

The important ramification of the results reported by Hazucha *et al.* (1992) and Adams (2003a) is that a non-linear dose-response relationship is evident. The higher hourly average concentrations elicit a greater effect than the lower hourly average values in a non-

linear manner. *A major implication of a non-linear dose-response relationship is that the same 8-hour average concentration, with different distributions of hourly average concentrations, will elicit a different adverse effect.*

This highlights an important limitation of a form of the proposed 8-hour O₃ standard, which averages hourly concentrations. Because the hourly average concentrations experienced over an 8-hour period are rarely constant (i.e., square wave), it is important that CAAQSOD recognize that the works by Hazucha *et al.* (1992) and Adams (2003a) provide information that will allow one to begin to characterize the realistic patterns of hourly average concentrations that are relevant for standard setting. The proposed level of the 8-hour standard described in the CAAQSOD is designed to protect the public from square-wave exposures. Since few people experience this exposure pattern, the level of the standard should be designed to protect the public from the family of realistic hourly average concentrations, over an 8-hour period, that elicits an adverse effect. Simply using data derived from square-wave exposure regime experiments provides an imprecise estimate of the realistic patterns of hourly average concentrations one is attempting to affect by emission reductions in order to protect the public's health. Given the importance of the non-linear dose-response relationship, it is very important that CAAQSOD reconsider the inappropriate use of an 8-hour standard that *averages* over the period of time. Research by Hazucha and Adams, based on realistic O₃ exposure patterns which found dosing rate influencing health effects, indicates that the higher hourly average concentrations should be provided greater weight than the lower levels in setting the level of the 8-hour standard. Therefore, the use of a simple averaging scheme over an 8-hour period, which assumes a linear dose-response relationship (which does not occur), is inappropriate.

This limitation of the controlled exposure studies also applies to the studies where other health outcomes were self-reported symptoms such as respiratory irritation, cough, wheeze, and pain upon deep inspiration. Since square wave exposure regimes do not mimic those experienced under ambient conditions, the authors of the CAAQSOD should reconsider the relevance of these health outcome results to impacts caused by actual ambient O₃ exposure.

3.3 Definition of an Adverse Health Effect

The CAAQSOD appropriately uses the guidelines published by the American Thoracic Society (ATS, 2000) in defining an adverse health effect. However, the guidelines are qualitative, in nature, making their application to a given study subjective. In particular, CAAQSOD should develop sets of criteria for pulmonary function tests (e.g., FEV1) and subject-reported symptoms.

The symptom criteria should address which symptoms are relevant, the magnitude of difference between “exposed” and “control” that is biologically meaningful, and the severity, frequency, and/or duration of occurrence of the symptom that render it an adverse health effect. Thus, for example, Adams (2002) uses the self-reported severity of symptoms ranging from throat tickle to pain on deep inspiration to create a total symptom score as the analytic metric. However, it is not clear that the ATS intended symptoms to be combined or that any perceptible symptom be considered as an adverse effect. In addition, ATS (2000) notes in its guidelines for symptoms that where symptoms result in a “diminished quality of life,” the negative effect(s) must be measurable. ATS (2000) acknowledges that at the individual level the degree of symptomology associated with diminished quality of life not requiring clinical

care or change in medication has not yet been agreed upon. This type of symptom response is typical of the controlled studies CAAQSOD references in setting the 8-hour standard.

With respect to group effects, while ATS indicates “any detectable” increment in symptom frequency should be considered an adverse effect, this is somewhat vague (e.g., a change of any magnitude or a statistically significant change) and contradicts the CAAQSOD qualitative principle that the incremental increase should be biologically meaningful.

Similarly, where FEV1 is the endpoint of concern, CAAQSOD provides no guidelines or criteria regarding the changes in FEV1 that would be considered biologically important. This is particularly important where the health assessment is based on population data rather than individual risk. In this case, statistical significance may be achieved based on the number of subjects in the study where the response differences noted between exposure and control outcomes are small and perhaps medically unremarkable. Therefore, it is recommended that the authors of CAAQSOD make explicit their quantitative levels of outcomes that represent an adverse health effect and reconsider the level of the proposed 8-hour standard in light of the ambiguity surrounding the use of symptom and FEV1 endpoints from available controlled human studies.

3.4 Background Ozone Not Used in Controlled Studies

Hourly average O₃ control concentrations that are used in the controlled human exposure experiments are near zero. Obviously, these concentrations are lower than (1) the *policy-relevant background* levels discussed in Chapter 4 and commented on in an earlier section of this report and (2) those hourly average concentration levels monitored at relatively remote monitoring sites in the United States or in other parts of the world (Lefohn

et al., 1990; Lefohn *et al.*, 2001). The use of concentrations below those of the cleanest monitoring sites in the world raises the concern that the results obtained from the 6.6-h experiments may provide *overestimates* of health effects the population would likely experience and overestimates of health benefits that could be achieved with O₃ control strategies when compared to reference concentrations that represent *policy-relevant background* levels. Thus, the experimental design used in the controlled exposure studies precludes interpreting their results as demonstrating any significant difference between the health impact of ambient exposure and *policy-relevant background*. Therefore, the basis for the level of the proposed 8-hr standard is not well supported.

3.5 Conclusion

As discussed earlier, Hazucha *et al.* (1992) and Adams (2003a) illustrated that the FEV1 response is dependent on the *dose rate* as well as the *cumulative dose* of O₃ inhaled, at least when the O₃ concentration is variable and that O₃ concentration is the most important factor in determining responses to O₃ exposure. This implies that one should not use the 8-hour *average* concentration as the form of the standard to protect human health. Rather, one should use the accumulation of the hourly average concentrations over an 8-hour period, with appropriate weighting being provided to the individual hourly average concentrations (Lefohn and Foley, 1993). In addition, health impacts should be assessed hourly during multi-hour chamber experiments rather than just at the onset and post-exposure period. Lastly, the use of exposures at 0 ppb in the multi-hour chambers experiments may result in an overestimate of effects. Thus, it is not only the proposed level of the 8-hour standard that is of questionable validity because of the design and analysis of the controlled human exposure studies, but the form of the 8-hour standard as well.

3.6 References

- Adams, W. C. (2002) Comparison of chamber and face-mask 6.6-hour exposures to ozone on pulmonary function and symptoms responses. *Inhal Toxicol* 14:745-64.
- Adams, W. C. (2003a) Comparison of chamber and face mask 6.6-hour exposure to 0.08 ppm ozone via square-wave and triangular profiles on pulmonary responses. *Inhal Toxicol* 15:265-81.
- Adams, W. C. (2003b) Relation of pulmonary responses induced by 6.6-hr exposures to 0.08 ppm ozone and 2-hr exposures to 0.30 ppm ozone via chamber and face-mask inhalation. *Inhal Toxicol* 15:745-59.
- American Thoracic Society. (2000) What constitutes an adverse health effect of air pollution? 161. *Inhal Toxicol* 161:665-73.
- Hazucha M. J; Folinsbee L. J.; Seal E. Jr. (1992) Effects of steady-state and variable ozone concentration profiles on pulmonary function. *Am Rev Respir Dis* 146:1487-93.
- Lefohn, A. S.; Foley, J. K. (1993) Establishing Ozone Standards to Protect Human Health and Vegetation: Exposure/Dose-Response Considerations. *J. Air Waste Manag. Assoc.* 43(2):106-112.
- Lefohn, A.S.; Shadwick, D. S. (personal communication).
- Lefohn, A. S.; Krupa, S. V.; Winstanley, D. (1990) Surface Ozone Exposures Measured at Clean Locations Around the World. *Environmental Pollution.* 63(3):189-224.
- Lefohn, A. S.; Oltmans, S. J.; Dann, T.; Singh, H. B. (2001) Present-day variability of background ozone in the lower troposphere. *J. Geophys. Res.* 106(D9):9945-9958.

4. Level of stringency and assumptions of equivalency between a 1-hour 90 ppb and 8-hour 70-ppb standard

4.1 Introduction

The California Ambient Air Quality Standard for Ozone Document (CAAQSOD) has recommended in Chapter 8 that the current 1-hour 90 ppb never to be exceeded standard be retained and that a new 8-hour 70 ppb never to be exceeded standard be adopted as well. Both of these recommendations are based on available chamber data and epidemiological studies. But, at the same time, it is quite clear that CAAQSOD has also looked at relationships between ambient exposure and 1-hour and 8-hour concentrations. Much of this is discussed, tabulated, and graphically displayed in Chapter 7 and summarized in portions of Chapter 8.

The CAAQSOD notes in Section 8.1 that one of the considerations in assessing health protectiveness of each ambient standard is “the degree of outdoor exposure in California relative to the level of the standard”. Further, the document notes that effects at concentrations at or below the current State standard of 90 ppb, averaged over one hour, would provide such evidence for the need for a more stringent standard, an averaging time different than the current 1-hour standard, or both. Given this, coupled with the decision to retain the current 1-hour standard, it appears that the CAAQSOD has turned to analyses of ambient air quality to show the relationship between exposure for the 1-hour 90 ppb standard and alternative concentrations for an 8-hour standard. Some of this information is displayed in the tables in Chapter 7 and is summarized in Table 8-2. In particular, Table 8-2 predicts that a majority of 8-hour design values would fall within the 70 to 80 ppb based on a rollback to a 1-hour maximum of 90 ppb at the design monitors for each air district within the state.

4.2 Rollback Methodology

As described in Chapter 10, the CAAQSOD, in order to calculate changes in exposure to O₃ that reflect a hypothetical attainment of the proposed ambient air quality standards, has used a proportional linear rollback procedure. The use of a proportional rollback methodology does not mimic the atmosphere's response to changes in O₃ precursors, which is a *non-linear* process. This means that the hourly concentrations within an 8-hour average do not respond in a linear manner when emission reductions occur. Second, as a result of the *non-linear* process, the amount of reduction needed to attain an 8-hour 70 ppb standard will be *significantly* more than that for a 90 ppb standard. Both of these statements are predicated on our understanding of atmospheric chemistry of O₃ formation. Work that supports these conclusions is described below. In addition, there is a serious question of whether the 8-hour standard is set at a concentration (i.e., 70 ppb) that really represents at times *policy-relevant background* levels, which are different than the levels that result from the modeling efforts cited in Chapter 4. If so, this has serious ramifications for overestimating benefits assumed in a risk assessment.

It is recognized that the primary concern of the CAAQSOD is to propose a standard whose form and level are protective of the public's health. In a previous section of this report, the uncertainty of the database associated with the development of the form and level of the proposed 8-hour standard was described. There is concern that the form and level are not adequate for protecting the public's health. In addition, there is concern that if the 8-hour standard of 70 ppb were adopted by California, it might not be attainable. Using the EPA's Aerometric Information Retrieval System (AIRS), Lefohn *et al.* (1998) reported that for the period 1993-1995, approximately 50% of the areas that violated the 8-hour standard were

influenced by 4 or more occurrences of mid-level hourly average concentrations (i.e., 60 - 90 ppb). In addition, the authors identified those sites that demonstrated a significant reduction in O₃ levels for the period 1980-1995. Using the data from the sites that experienced reduced O₃ levels over the period of time, Lefohn *et al.* (1998) investigated whether the rate of reduction of the mid-level hourly average concentrations was similar to the rate experienced by the high hourly average concentrations. The analysis indicated that the hourly average concentrations in the mid range were reduced slower than the hourly average concentrations above 90 ppb. The figure below is an example that shows that the higher hourly average concentrations (i.e., above 90 ppb) decreased at a faster rate (greater negative rate per year) than the hourly average concentrations in the mid-level range. The numbers of hourly average concentrations in the low end of the distribution also decreased. Apparently, both the high and low ends of the distribution were moving toward the center of the distribution.

As control strategies are implemented, the rate of change to reduce the higher hourly average concentrations will be greater than the rate of change of the mid-level concentrations. Figure 4-1 from Lefohn *et al.* (1998) illustrates the disproportionate reduction of the hourly average concentrations. Note that for a monitoring site in Ventura County that actual reduction in emissions resulted in both the frequency of the higher hourly and lower hourly average concentrations being reduced. The frequency of occurrence of the mid-level hourly average concentrations (i.e., the 50 – 70 ppb range) increased. This meant that both ends of the hourly average concentration distribution were “squeezed” towards the middle of the distribution. The result is that as serious emission reductions occur, at some monitoring sites in California, there will be an increase in the mid-level hourly average concentrations. Clearly a proportional rollback approach is meaningless for predicting actual changes in

hourly average concentrations. Lefohn *et al.* (1998) identified a similar response in hourly average concentrations as a result of emission reduction for other California sites.

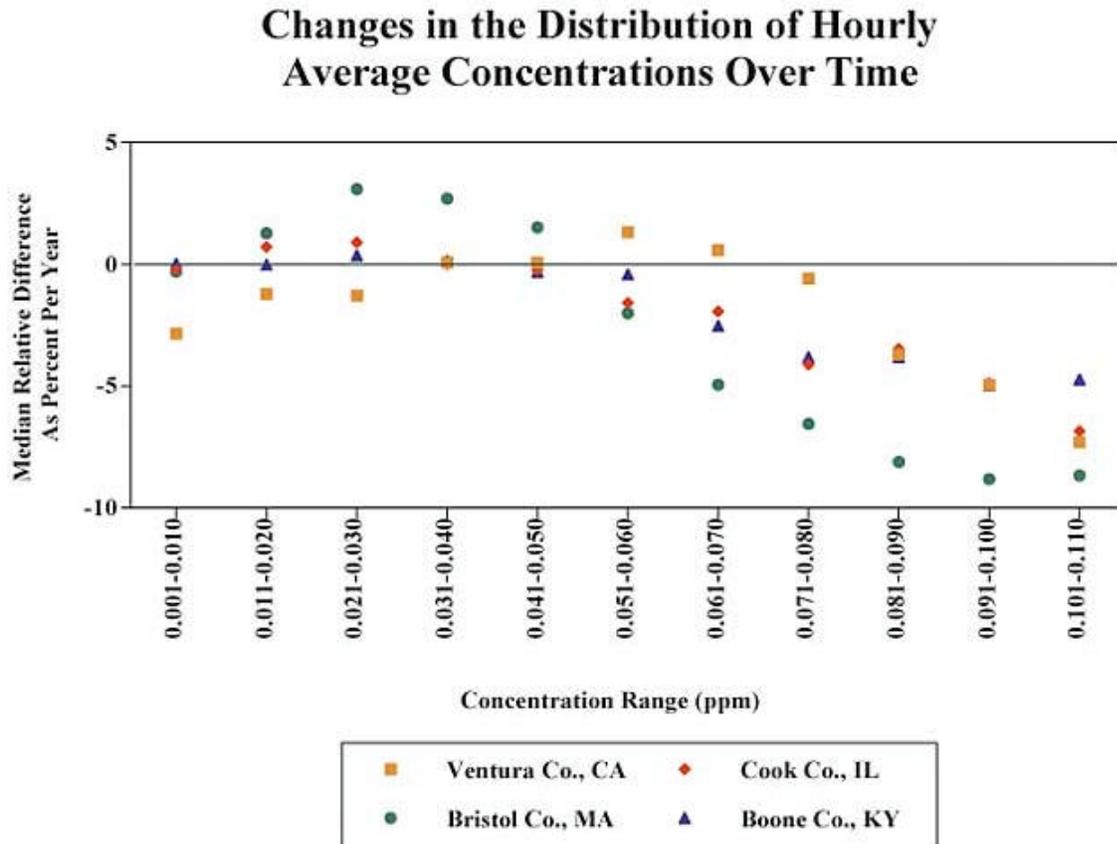


Figure 4-1. Changes in distribution of hourly average concentrations over time. Source: Lefohn *et al.* (1998).

Based on the results published by Lefohn *et al.* (1998), which used empirical data, it appears that when control strategies are implemented, the higher hourly average concentrations will be reduced faster than the mid-level values. Similar to the results obtained for the hourly values, the rate of decline for the 8-hour daily maximum values in the mid range will be much slower than the higher 8-hour values. Figure 4-2 illustrates the

slowing down process for Fairfield County, Connecticut. Note the rapid decrease in the early years and then a "flattening" of the curve in the later years.

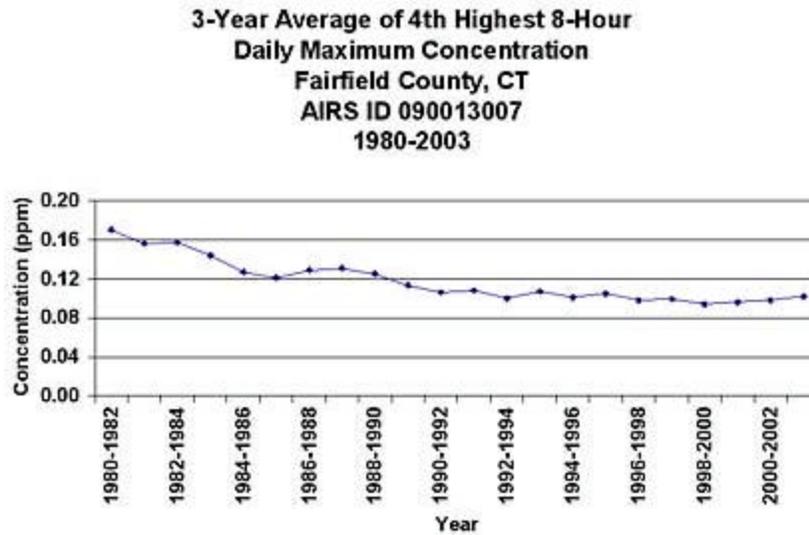


Figure 4-2. Trends of the 8-hour standard for the period 1980-2003.

This flattening has been observed nationwide, including California. The EPA has recently published the report, *The Ozone Report - Measuring Progress Through 2003* (EPA, 2003). The Agency notes that there has been a slowing down in the reduction of both the 1-hour and 8-hour average concentrations. Figures 4-3 and 4-4, taken from that report, illustrate the changes over the period 1980-2003.

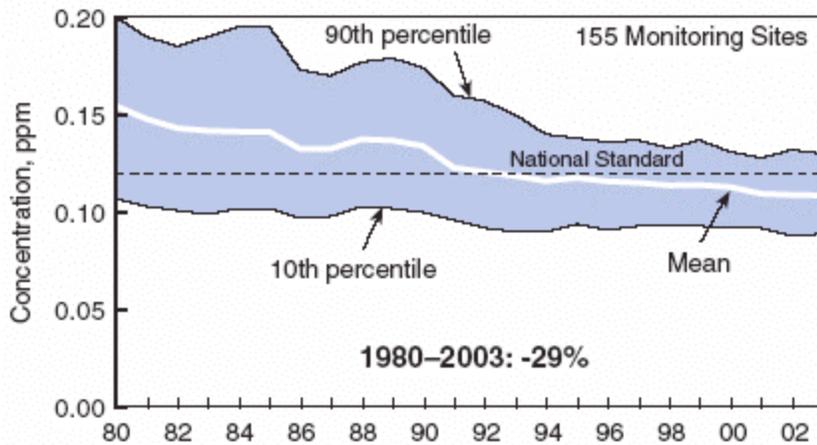


Figure 6. One-Hour Ozone Air Quality Trend, 1980–2003, Based on Running Fourth Highest Daily Maximum 1-Hour Ozone Value over 3 Years.

Figure 4-3. One-hour ozone air quality trend, 1980-2003, based on running 4th highest daily maximum 1-hour ozone value over 3 years. Source: *The Ozone Report - Measuring Progress Through 2003*.

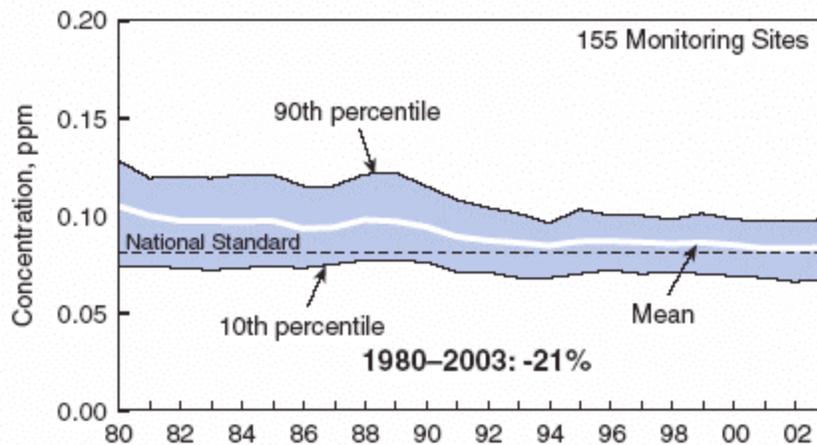


Figure 8. 8-Hour Ozone Air Quality Trend, 1980–2003, Based on 3-Year Rolling Averages of Annual Fourth Highest Daily Maximum 8-hour Ozone Concentrations.

Figure 4-4. 8-hour ozone air quality trend, 1980-2003, based on running 4th highest daily 3-year rolling averages of annual 4th highest daily maximum 8-hour ozone concentrations. Source: *The Ozone Report - Measuring Progress Through 2003*.

4.3 Attaining the Proposed California 8-Hour Standard

For most sites that violate the 8-hour 80 ppb federal standard, attainment may be extremely difficult. With a proposed 8-hour 70 ppb California standard, the difficulty inherent in mid-level reductions suggests the possibility that severe reductions of anthropogenic emissions will fail to result in attaining a 70 ppb level. Realistically, the proposed 8-hour California standard of 70 ppb may actually be an *Objective* (i.e., goal) instead of a standard.

Using models, several investigators have commented on the difficulty in reducing the mid-level hourly average concentrations, while reducing the fourth highest 8-hour average daily maximum concentration. Winner and Cass (2000) noted that the higher hourly average concentrations were reduced much faster than the mid-level values during simulation modeling for the Los Angeles area. Reynolds *et al.* (2003) analyzed ambient O₃ concentrations used in conjunction with the application of photochemical modeling to determine the technical feasibility of reducing hourly average concentrations in central California, using the 1990 August 3-6 San Joaquin Valley Air Quality Study episode. The following four isopleths show how O₃ responds to precursor reduction. Each isopleth was created by using the modeling outputs for a specific combination of precursor reduction done in 10% increments (for example 10% VOC reduction and 20% NO_x reduction would be one of 121 possible combinations), followed by plotting the results. The axes represent the amount of reduction from the 1999 baseline emissions, with 100% of each precursor being in the top right-hand corner. Ozone concentrations are represented by the isopleths. These isopleths can be calculated for either the 1-hour average or the 8-hour average, and illustrate how O₃ would respond to reduction for the area specified. As can be seen in the Fresno isopleths (Figures 4-5 and 4-6), in order to achieve a 90 ppb O₃ concentration for the peak

hour, 85% reduction of NO_x from the baseline is needed. However, for the same day, a 70 ppb concentration for the peak 8-hour requires about 92% reduction of NO_x.

Similarly, for the Bakersfield isopleths (Figures 4-7 and 4-8), the 90 ppb peak 1- hour O₃ concentration requires an 82% reduction of NO_x from the baseline. Yet, for the 70 ppb peak 8-hour O₃ concentration, a 86% reduction will be required. Similar findings were observed for other inland Central California locations Reynolds *et. al.*, 2003), as well as major areas in the eastern half of the United States (Reynolds *et. al.*, 2004).

Reynolds *et al.* (2003) have commented on possible chemical explanations for the observation that more prominent trends in peak 1-hour O₃ levels occur than for trends in peak 8-hour O₃ concentrations or in occurrences of mid-level (i.e., 60 – 90 ppb) concentrations. The authors noted that when anthropogenic VOC and NO_x emissions are reduced significantly, the primary sources of O₃ precursors are biogenic emissions and CO from anthropogenic sources. Chemical process analysis results indicated that slowly reacting pollutants such as CO could be contributing on the order of 10 – 20% of the O₃ produced. Moreover, the authors noted that process analysis indicated that as NO_x was reduced, the process for O₃ formation became more efficient, producing more molecules of O₃ for each molecule of NO_x. That is, decreasing emissions were offsetting to some extent the increased effectiveness of making O₃.

Peak 1-hour Ozone Isopleths (ppb) for Fresno Subregion - 6 August 1999

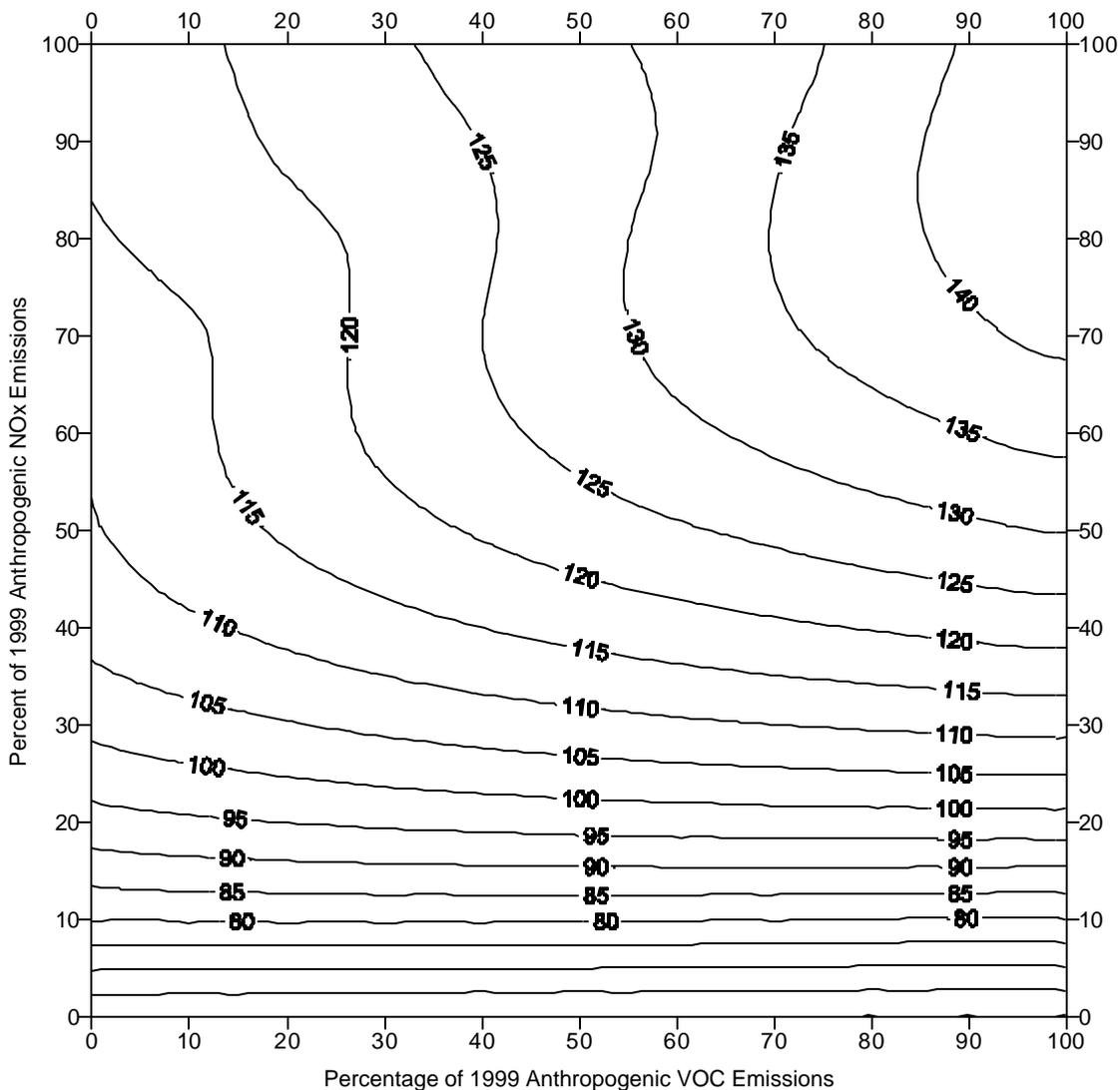


Figure 4-5. Peak 1-hour ozone isopleths (ppb) for Fresno Subregion August 6, 1999.

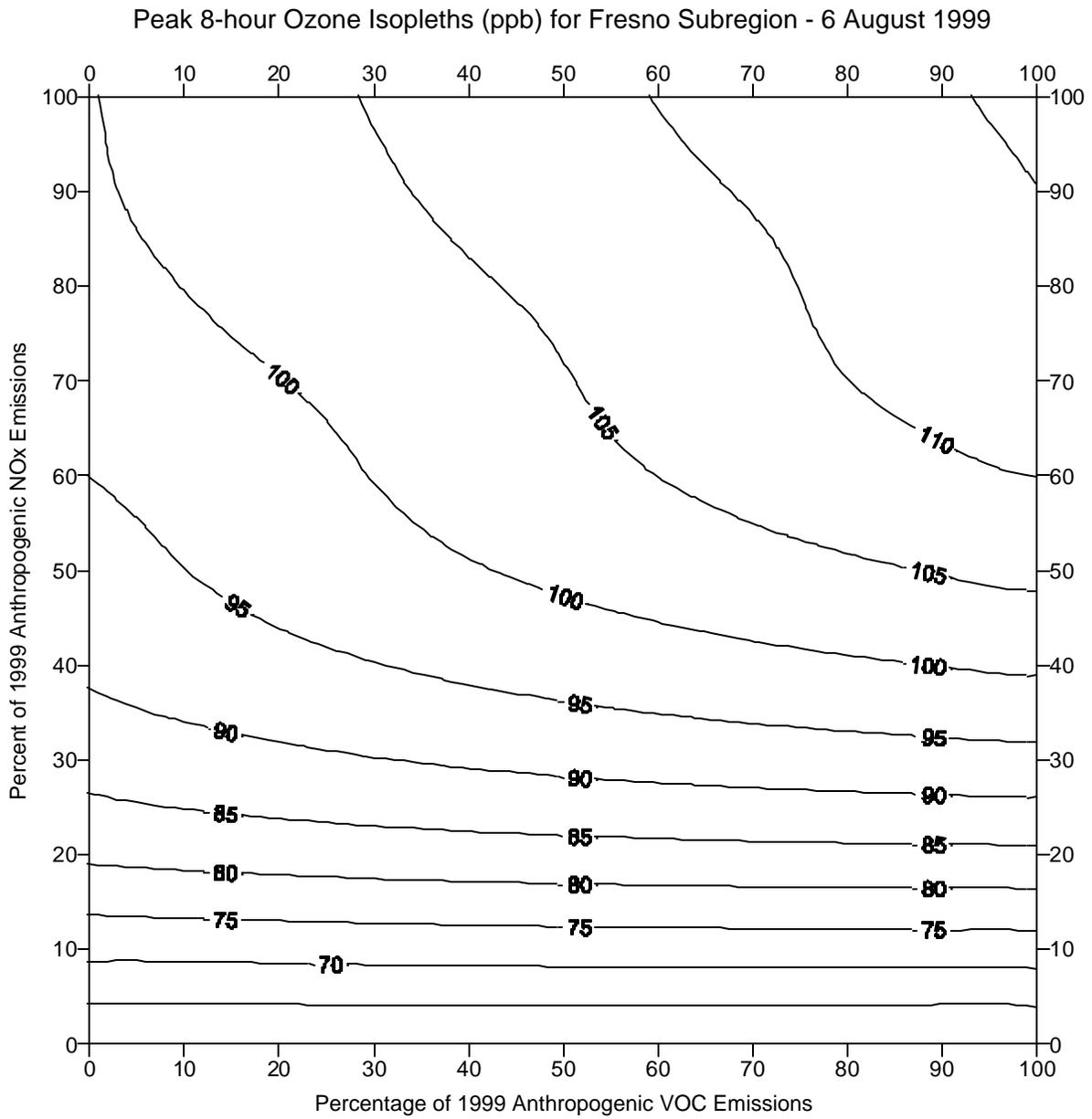


Figure 4-6. Peak 8-hour ozone isopleths (ppb) for Fresno Subregion August 6, 1999.

Peak 1-hour Ozone Isopleths (ppb) for Bakersfield Subregion - 6 August 1999

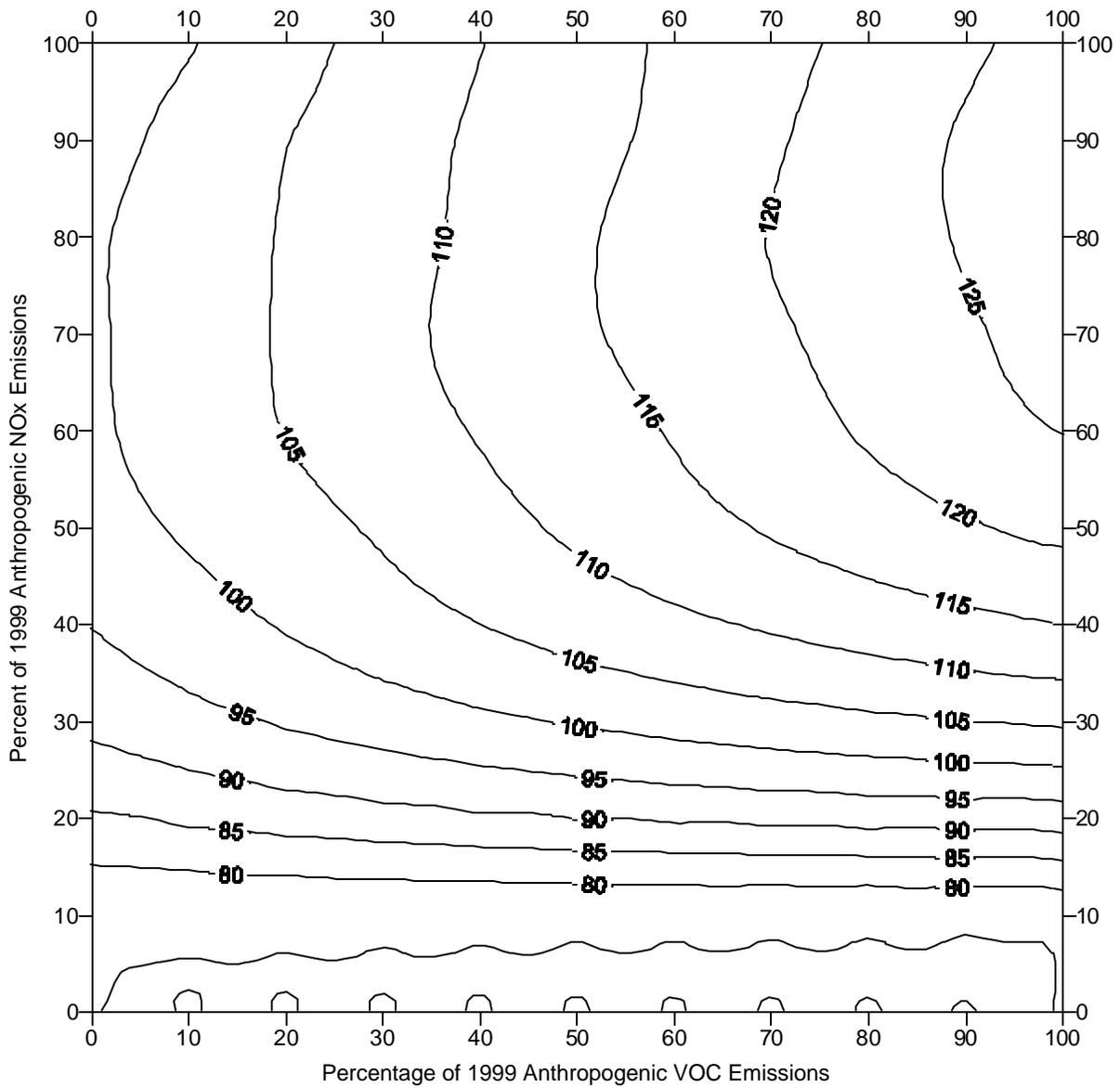


Figure 4-7. Peak 1-hour ozone isopleths (ppb) for Bakersfield Subregion August 6, 1999.

Peak 8-hour Ozone Isopleths (ppb) for Bakersfield Subregion - 6 August 1999

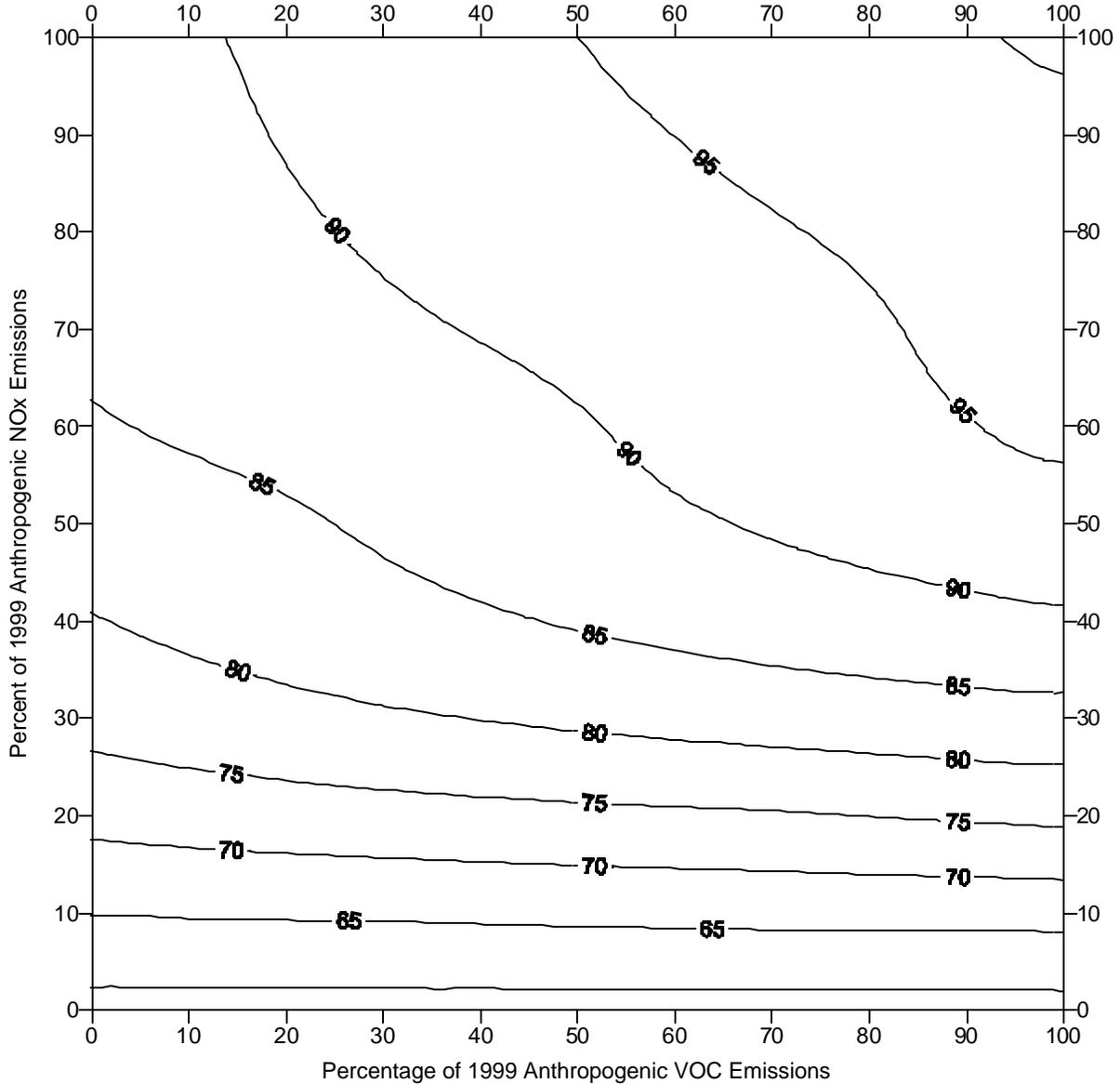


Figure 4-8. Peak 8-hour ozone isopleths (ppb) for Bakersfield Subregion August 6, 1999.

This was also seen in analysis of the ambient air quality data when looking at the change in the O₃ to NO_x ratio between 1990 and 2000, as shown in Figure 4-9. The ratio of O₃ to NO_x represents the efficiency of the atmosphere to produce O₃. There were significant NO_x emissions reduction in the Bakersfield area between 1990 and 1999, both for the

weekday (WD) and weekend (WE) days. Yet, with these large reductions, the ambient ratio shows an increase, and provides an independent correlation with the model results showing that the reduction of NO_x is offset by the efficiency of the atmosphere to produce more O_3 per molecule of NO_x , thus requiring more reduction of NO_x to meet the applicable O_3 standard. This was observed for other locations in Central California as well.

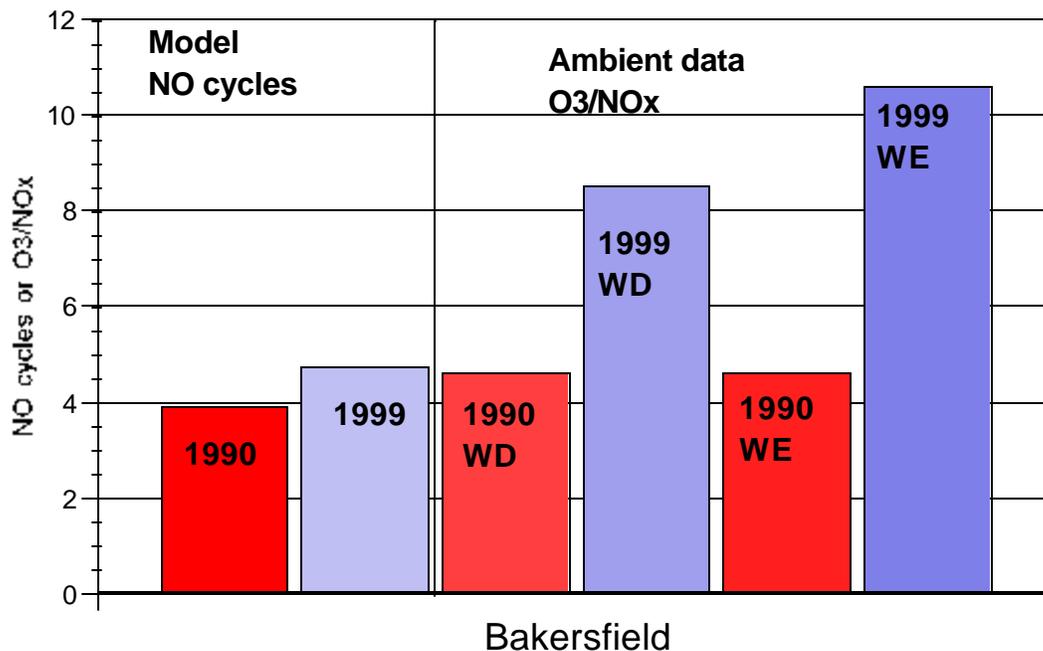


Figure 4-9. Change in the O_3 to NO_x ratio between 1990 and 2000.

4.4 Conclusion

What this means relative to the suppositions used by the CAAQSOD in selecting an 8-hour standard, in part based on looking at equivalency of exposure between the 1-hour 90 ppb standard and the proposed 8-hour 70 ppb standard, is that the document has failed to take into account that the atmosphere is not a linear system. In reality, what has been proposed is a standard that will be extremely difficult to attain and perhaps even unattainable. In addition,

different exposure patterns (i.e., hourly frequency distributions) will occur than the predicted distributions based on a simple linear rollback. The proposed 8-hour standard will require much more stringent emission reductions than the current state 1-hour standard and it may not be possible to obtain the actual emission reductions required to attain an 8-hour average standard at 70 ppb.

References

- Lefohn, A. S.; Shadwick, D. S.; Ziman, S. D. (1998) The Difficult Challenge of Attaining EPA's New Ozone Standard. *Environmental Science & Technology.*, 32(11):276A-282A
- Reynolds, S. D.; Blanchard, C. L.; Ziman, S. D. (2003) Understanding the effectiveness of precursor reductions in lowering 8-hr ozone concentrations. *J. Air & Waste Manage. Assoc.* 53: 195-205.
- Reynolds, S. D.; Blanchard, C. L.; Ziman, S. D. (2004) Understanding the effectiveness of precursor reductions in lowering 8-hr ozone concentrations-Part II The eastern United States. *J Air & Waste Manage. Assoc.* in press.
- Winner, D. A.; Cass, G. R. (2000) Effect of emissions control on the long-term frequency distribution of regional ozone concentrations. *Environmental Science & Technology.* 34: 2612-2617.

ENVIRON

August 31, 2004

MEMORANDUM

To: Kyle Isakower, American Petroleum Institute

From: Stan R. Hayes, ENVIRON

Subject: **Initial Comments on California's Draft Ozone Staff Report**

As a part of their review of the California ambient air quality standard for ozone (CAAQS), the Air Resources Board (ARB) and the Office of Environmental Health Hazard Assessment (OE-HHA) of the California Environmental Protection Agency (Cal/EPA) are reviewing the scientific literature on public exposure, atmospheric chemistry, and the health effects of exposure to ozone. Results of that review are contained in a Draft Ozone Staff Report (Cal/EPA 2004) entitled "Review of the California Ambient Air Quality Standard for Ozone," which was released for public review on June 21, 2004.

At the request of the American Petroleum Institute (API), we have reviewed portions of the Draft Ozone Staff Report relating to health effects and epidemiology (primarily in Chapters 11 and 12). This memorandum summarizes our initial comments.

We note that the volume of material considered in the Draft Ozone Staff Report and cited by it is extensive and complex. Given limited time and resource constraints, it is not possible at this time for us to review all of the material and topics discussed in that report. Our review, therefore, is limited in focus. We ask that our comments be viewed as preliminary. We reserve the right to expand and supplement our comments later, if appropriate and should additional opportunity, time, and resources become available.

The Draft Ozone Staff Report, as released for review, did not yet contain Chapter 10, "Quantifying the Health Benefits of Reducing Ozone Exposure." The draft of that chapter was released by ARB on August 24. Because of the timing of that release, coming after completion of our review of the earlier report version, we do not include comments here on Chapter 10.

SUMMARY

Specific comments are summarized as follows:

1. Serious concerns exist about the use of PM epidemiological studies to assess ozone acute mortality effects.
2. Further analyses are needed before drawing conclusions from epidemiological studies about the effects of ozone exposure on mortality and other serious health endpoints.
3. A more precise and quantitative definition of adverse effects is needed.

4. Further discussion of the basis for the staff report's ozone CAAQS recommendations is needed, particularly relative to the justification for a more stringent standard than the federal NAAQS.
5. As called for in the staff report, additional research to investigate the response of human subjects to multi-hour exposures to ozone in the range 0.04 to 0.08 ppm is warranted and needed.
6. Additional quantification of the nature and standard setting implications of uncertainties, individually and in combination, is warranted and needed.

COMMENTS

1. Serious concerns exist about the use of PM epidemiological studies to assess ozone acute mortality effects. Over the past decade, numerous epidemiological studies have been conducted, such as those that examine the statistical association between acute mortality and particulate matter (PM). Seeking to avoid potential confounding between PM and other pollutants, many studies also examined the relationship between daily mortality and ozone. The staff report (Section 12.4.4) extracts the ozone mortality associations reported in those studies (see, for example, the NMMAPS study as re-analyzed by Dominici et al. 2003) and concludes that:

“...there is now sufficient evidence to reach the preliminary conclusion that summer season O₃ is likely to be an independent risk factor for premature mortality.” (p. 12-67)

“The median risk from current studies appears to be about 3% pr [sic] 40 ppb change in 1 hr O₃ concentration.” (p. 12-77)

While a full analysis of this issue is beyond the scope of this review, we caution that new, ozone-specific analyses to verify the applicability of PM mortality studies to ozone are warranted and necessary before drawing such conclusions as those above (see also Comment 2). Absent new ozone-specific analyses, there are reasons for serious concern about the reliability of using PM epidemiological studies to characterize the effects of ozone exposure on acute mortality, including the following:

- Not only is the validity of assuming a linear exposure-response relationship for ozone not known, but data for other acute responses show contrary evidence of nonlinearity.
- Ozone presents greater confounding problems than PM, due to its strong seasonal cycle and temperature dependence.
- The most appropriate metric for ozone exposure is not known, and may need to consider multiple factors such as averaging time (1-hour, 8-hour, or seasonal), different measures of maximum concentration, cumulative dose (concentration x time), inter-episode duration, and frequency of repeated elevated concentrations.
- Greater uncertainty exists for ozone than for PM about the degree to which fixed-site measurements accurately characterize personal exposure.
- The role of the more pronounced diurnal profile for ozone is not known.

Among these reasons for concern, the first may be particularly important, especially for ozone at or below the levels of federal and ozone state standards, since by definition those are the concentrations that would occur upon attainment. PM epidemiological studies (as well as epidemiological studies that assess other acute ozone effects, for that matter) assume that the underlying exposure-response relationship is linear. However, the basis for assuming linearity for ozone is not stated in the Draft Ozone Staff Report, which does not present its evidence and reasoning in support of this important assumption.

Adequate data are not available to evaluate directly the validity of this linearity assumption for mortality. However, if linearity were to hold for ozone acute mortality (and likely for other acute effects as well) as assumed in the Draft Ozone Staff Report, one might expect to see signs of that linearity in the data reported for other measures of acute ozone response, such as lung function impairment and lower respiratory symptoms.

While it is not possible within the scope of this limited review to examine all data sets and acute health endpoints, consider several examples for illustration purposes. The linearity hypothesis for lung function impairment and lower respiratory symptoms can be tested using human clinical data from controlled chamber studies conducted by Avol et al. (1984), Kulle et al. (1985), and McDonnell et al. (1983). These studies, which are also cited in the Draft Ozone Staff Report, measured the response of heavily exercising, healthy young adults exposed to a range of ozone concentrations for one hour (Avol) or two hours (Kulle and McDonnell). Response to ozone exposure was measured in terms of lung function impairment (e.g., FEV1¹ decrement) and lower respiratory symptom incidence (e.g., cough and chest discomfort).

During an earlier review of the national ambient air quality standards (NAAQS) for ozone, Hayes et al. (1987; 1989) used data from these three studies to develop exposure-response relationships that were cited by the U.S. Environmental Protection Agency (USEPA) in their Ozone Staff Paper (USEPA 1988) and used in USEPA's Acute Ozone Health Risk Assessment (Hayes et al. 1989; Winkler et al. 1990; Whitfield et al. 1994).

Exposure-response relationships from Hayes et al. (1989) are shown in Figures 1 through 6. Ozone-induced responses, expressed as the fraction of the healthy adult population fraction affected, are shown for two different levels of FEV1 decrement in Figure 1 (10%) and Figure 2 (20%), two different level of cough severity (or lower respiratory symptoms for Avol) in Figure 3 (any including mild) and Figure 4 (moderate/severe), and two different levels of chest discomfort (again, lower respiratory symptoms for Avol) in Figure 5 (any) and Figure 6 (moderate/severe).

Few of the exposure-response relationships show evidence of the hypothesized linearity. In fact, while there is a range of shapes that best fit the different data sets, a distinct sigmoidal shape is evident for many of the dataset-endpoint combinations.

Nonlinearities are also evident in Figure 7 from McDonnell and Smith (1994) (see USEPA's Ozone Staff Paper 1996, p. 34), in which FEV1 response data from 1-, 2-, and 6.6-hour exposure are presented. Also, data from Folinsbee et al. (1991), as presented in the Draft Ozone Staff Report (Figure 11-2, p. 11-55), are used to derive the relationships in Figure 8, which are for three levels of FEV1 decrement (>10%, >15%, and >20%) under 6.6-hour exposures at exercise levels

¹ FEV1 is the volume of air that can be expelled in the first second of a maximal expiration.

intended to represent heavy or strenuous work or play over the period. The narrow range of ozone exposure concentrations considered (0.08, 0.10, and 0.12 ppm) makes it difficult to assess the degree of linearity of the exposure-response relationship at lower concentrations. Those concentrations are important because they are levels that would occur on most days under conditions of standard attainment (see also a later comment supporting the Draft Ozone Staff Report's call for additional research in this area). Nonetheless, the data in Figure 8 are also suggestive of possible response nonlinearities.

We agree with the following statements in the Draft Ozone Staff Report about this issue:

“Responses to duration of exposure are not linear” (Section 11.3.1, p. 11-13)

“The length of the interval between exposures appears to be more critical in determining the long-term impact of repeated exposures than the total duration of the exposure episode” (Section 11.3, p. 11.46)

“The model results suggest that the O₃-pulmonary function response relationship may have a sigmoid rather than a linear shape, suggesting a response plateau” (Section 11.4.3, p. 11-77)

While the discussion here is not definitive, it is suggestive. New, ozone-specific research is needed to better assess the plausibility, likelihood of occurrence, and potential magnitude of ozone mortality effects at concentration levels currently found in ambient air, particularly at ozone levels characteristic of standard attainment.

2. Further analyses are needed before drawing conclusions from epidemiological studies about the effects of ozone exposure on mortality and other serious health endpoints. Conclusions in the Draft Ozone Staff Report regarding the effect of ozone exposure on acute mortality and other serious acute health endpoints, and particularly the quantification of those effects in ozone benefit analyses in Chapter 10, should be deferred until further assessment of the appropriateness of PM acute mortality and other epidemiological studies for ozone standard setting purposes. If warranted based on the findings of that assessment, new ozone-specific research should be conducted to more reliably investigate the relationship between ozone exposure and mortality and other serious health endpoints. That research should include at a minimum the following:

- Use of a nonlinear ozone exposure-response model, one that at least allows for the possibility of a nonlinear response.
- Inclusion of a range of alternative ozone exposure metrics, considering, in addition to measures of the peak concentration level (e.g., 1-, 8-, 24-hr, seasonal averages), the following other potentially important factors: (a) frequency of peaks, (b) accumulated dosage (C x T), and (c) inter-episode duration.
- Use of improved techniques for filtering out such confounding influences as ozone's strong seasonal cycle and temperature dependence.
- Use of methods to ensure accurate quantification of important uncertainties and their standard setting implications.

3. A more precise and quantitative definition of adverse effects is needed. In Section 8.2 of the Draft Ozone Staff Report, the definition of an “adverse effect” is discussed. As the staff report notes:

“It is important to keep in mind the differences between statistical significance and medical or biological significance when considering what constitutes an adverse health effect.” (p. 8-3)

Having stated this appropriate caution, however, the staff report appears at times to equate “detectable” with “adverse.” For example, the staff report states that “any detectable increase in symptom frequency should be considered adverse” (p. 8-4).

Development of a more precise and quantitative definition of adverse effects is warranted and needed, and would allow Cal/EPA to be more precise in stating when and under what circumstances a statistically detectable effect becomes sufficiently large to be adverse. In developing such a definition, Cal/EPA may wish to consider such factors as the following:

- For symptoms, in defining effect adversity, it may be judged appropriate to consider symptom severity, which is typically rated in human chamber studies as “any” (which includes mild), “moderate,” or “severe.” Rather than regarding “any detectable increase” in symptom frequency as adverse, it might be more appropriate to require that, to be considered adverse, a symptom would have to be moderate or severe. This would seem more consistent with 2000 ATS guidelines that “symptoms associated with reduced quality of life or with a change in clinical status (i.e., requiring medical care or a change in medications) should be considered adverse” (Chapter 8, p. 8-4).
- For pulmonary function, the staff report appears to agree that “small, transient reductions in pulmonary function should not necessarily be regarded as adverse” (p. 8-4), but that “reversible loss of lung function in conjunction with symptoms should be considered adverse” (p. 8-4). In applying this definition of effect adversity, Cal/EPA should take into account that lung function decrement and symptoms may not be well correlated in individuals. As a result, when defining effect adversity, Cal/EPA should incorporate the joint probability that both lung function decrements of sufficient magnitude and symptoms of sufficient severity occur in the same individual.
- In any event, specificity is needed in defining an adverse effect. The definition should specify: (a) the degree of lung functional decrement that is judged to be adverse and why (e.g., FEV1 decrements of 10%, 15%, or 20%), (b) whether individual-level responses or group-mean responses are to be used and why, and (c) which degree of symptom severity is to be judged adverse and why (“any” or “moderate/severe”).

4. Further discussion of the basis for the staff report’s ozone CAAQS recommendations is needed, particularly relative to the justification for a more stringent standard than the federal NAAQS. The Draft Ozone Staff Report recommends the form, averaging time, and concentration for the CAAQS. However, the basis for the staff report’s recommendations needs further discussion, clarification, and justification. Of particular relevance are the reasons for adopting a more stringent standard than the federal NAAQS, with an emphasis on the need for such action and the incremental benefits expected to be realized (perhaps to be addressed in a revised version of Chapter 10).

5. As called for in the staff report, additional research to investigate the response of human subjects to multi-hour exposures to ozone in the range 0.04 to 0.08 ppm is needed. Section 1.2 of the Draft Ozone Staff Report, among other things, calls for the funding of “additional research investigating the responses of human subjects to multi-hour exposures to O₃ concentrations between 0.04 and 0.08 ppm.” Such research is warranted and should be conducted. In addition (see earlier comment), new ozone-specific research should be conducted before drawing conclusions from epidemiological studies regarding ozone acute mortality or serious morbidity effects, and especially before using them to quantify such effects in ozone benefits analyses in Chapter 10 of the staff report.

6. A more quantitative analysis of the nature and standard setting implications of uncertainties, individually and in combination, is warranted and needed. Significant uncertainties exist in the available health and epidemiological data, particularly with respect to the latter. A more quantitative and systematic approach to the characterization of uncertainty would assist in judging the standard setting implications of that uncertainty and identifying areas where additional research would be most important.

REFERENCES

- Avol, E.L., W.S. Linn, T.G. Venet, D.A. Shamoo, and J.D. Hackney. 1984. Comparative respiratory effects of ozone and ambient oxidant pollution exposure during heavy exercise. *J. Air Pollut. Control Assoc.*, 34:804-809.
- Cal/EPA. 2004. “Review of the California Ambient Air Quality Standard For Ozone, Volumes I and 2.” Public Review Draft. California Environmental Protection Agency, Air Resources Board and Office of Environmental Health and Hazard Assessment, Sacramento, California. June 21.
- Dominici, F., A. McDermott, M. Daniels, S.L. Zeger, and J.M. Samet. 2003. “Mortality Among Residents of 90 Cities. Revised Analyses of Time-series Studies of Air Pollution and Health.” Health Effects Institute, Boston, Massachusetts.
- Folinsbee, L. J., D.H. Horstman, H.R. Kehrl, W.F. McDonnell, T.R. Gerrity, E. Seal, R. Larson, M.J. Hazucha, S. Abdul-Salaam, B. Faucette, and P.J. Ives. 1991. Effects of single and repeated prolonged low-level ozone exposure in man. Presented at: annual meeting of the Society for Occupational and Environmental Health; March; Washington, DC.
- Hayes, S.R., M. Moezzi, T.S. Wallsten, and R.L. Winkler. 1987. “An Analysis of Symptom and Lung Function Data from Several Controlled Ozone Exposure Studies.” SYSAPP-86/120. Systems Applications, Inc., San Rafael, California, for the U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.
- Hayes, S.R., A.S. Rosenbaum, T.S. Wallsten, R.G. Whitfield, R.G. Whitfield, and R.L. Winkler. 1989. Acute ozone exposure-response relationships for use in health risk assessment. Paper No. MPM-L4. Annual Meeting, Society for Risk Analysis, San Francisco, California. October 29 – November 1.
- Hayes, S.R., A.S. Rosenbaum, T.S. Wallsten, R.G. Whitfield, R.L. Winkler, and H. Richmond. 1989. A health risk assessment for use in setting the U.S. primary ozone standard. In:

Atmospheric Ozone Research and Its Policy Implications, T. Schneider and S.D. Lee (Editors), Elsevier Science Publishers, Amsterdam, The Netherlands.

Kulle, T.J., L.R. Sauder, J.R. Hebel, and M.D. Chatham. 1985. Ozone response relationships in healthy non-smokers. *Am. Rev. Respir. Dis.*, 132:36-41.

McDonnell, W.F., D.H. Horstman, M.J. Hazucha, E. Seal, Jr., E.D. Haak, S. Salaam, and D.E. House. 1983. Pulmonary effects of ozone exposure during exercise: dose-response characteristics. *J. Appl. Physiol.: Respir. Environ. Exercise Physiol.*, 54:1345-1352.

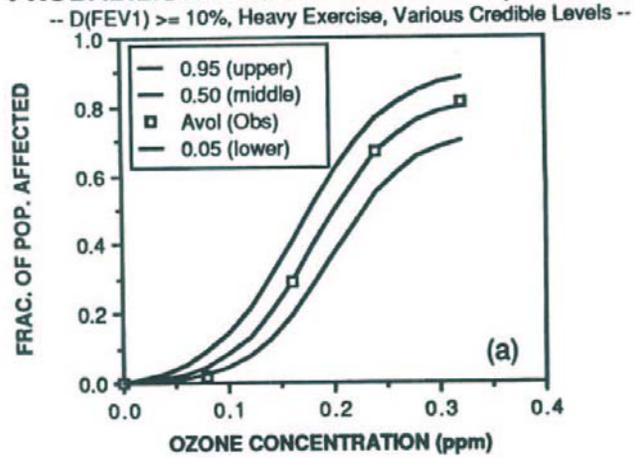
McDonnell, W.F., and M.V. Smith. 1994. Prediction of acute ozone response as a function of exposure rate and total inhaled dose. *J. Appl. Physiol.* 76: 2776-2784.

U.S. EPA. 1988. "Review of the National Ambient Air Quality Standards for Ozone: Preliminary Assessment of Scientific and Technical Information, Staff Paper." U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

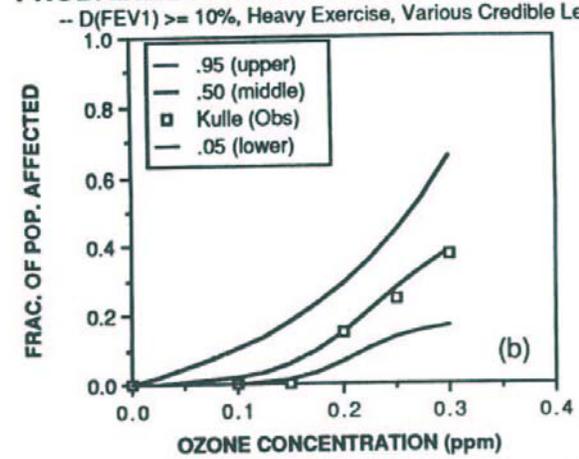
U.S. EPA. 1996. "Review of the National Ambient Air Quality Standards for Ozone, Assessment of Scientific and Technical Information, OAQPS Staff Paper." EPA-452/R-96-007. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina. June.

Whitfield, R.G., H.M. Richmond, S.R. Hayes, A.S. Rosenbaum, T.S. Wallsten, R.L. Winkler, M.L.G. Absil, and P. Narducci. 1994. *Chapter 5, Health Risk Assessment of Ozone*. In: *Ozone: Effects on Public Health and Ecological Impacts, Why Society Must Control Tropospheric Ozone*, ed., D.J. McKee. Chelsea, Michigan: Lewis Publishers.

PROBABILISTIC EXP.-RESP. RELAT. (Avoi Data)



PROBABILISTIC EXP.-RESP. RELAT. (Kulle Data)



PROBABILISTIC EXP.-RESP. RELAT. (McDonnell Data)

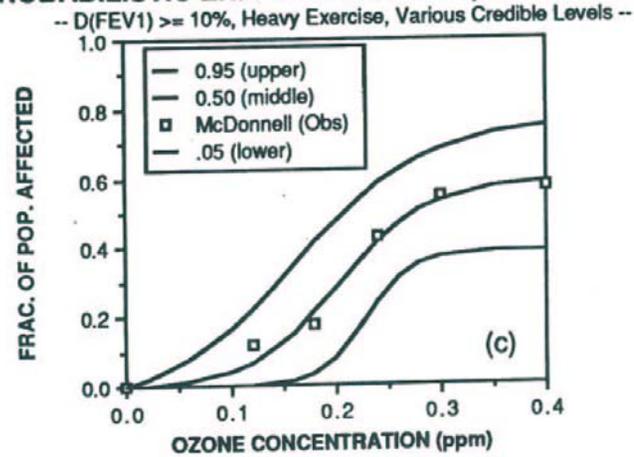


Fig 1. Ozone probabilistic exposure-response relationships for FEV1 decrement \geq 10 percent (heavy exercise).

Source: Hayes et al. (1989)

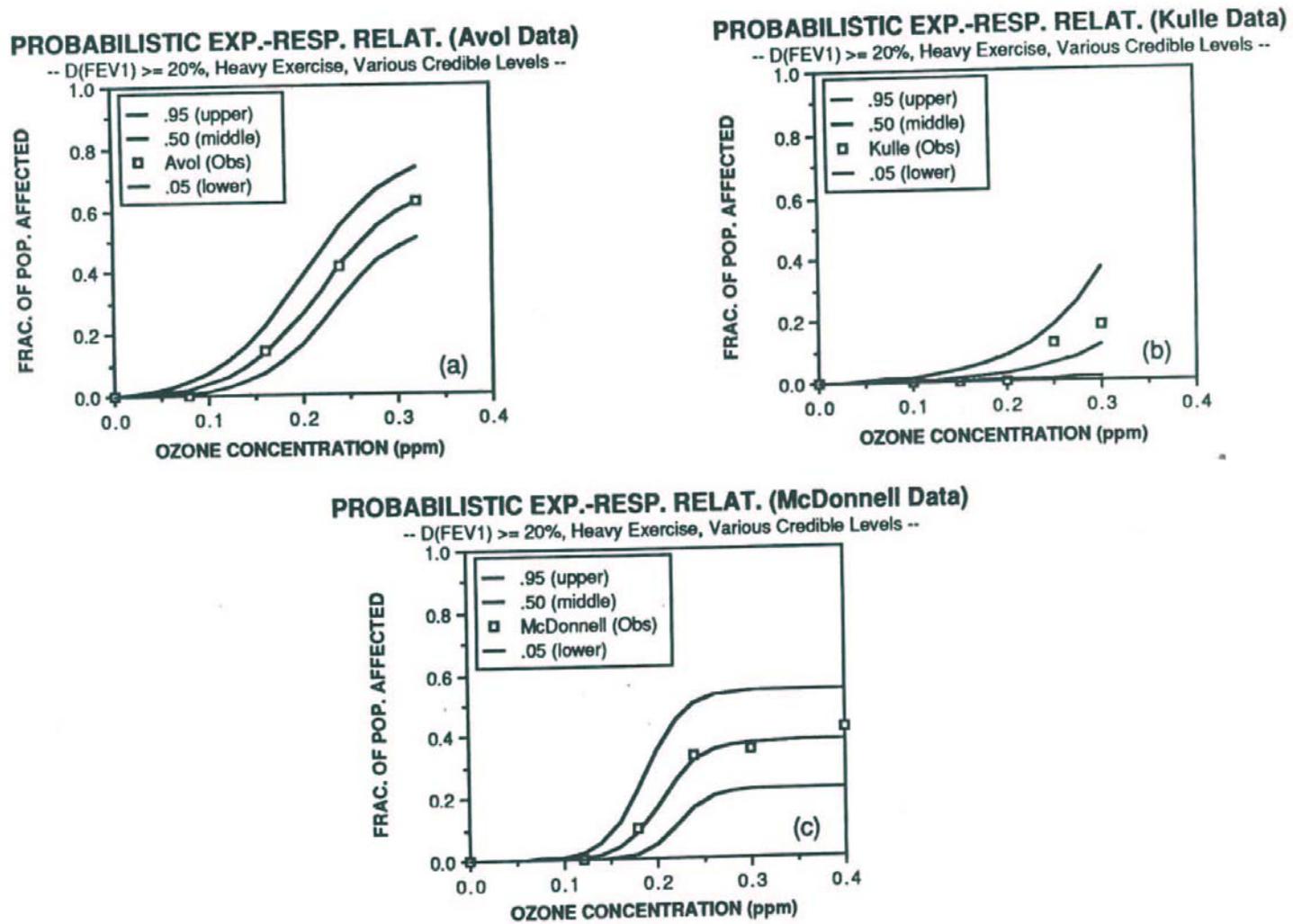
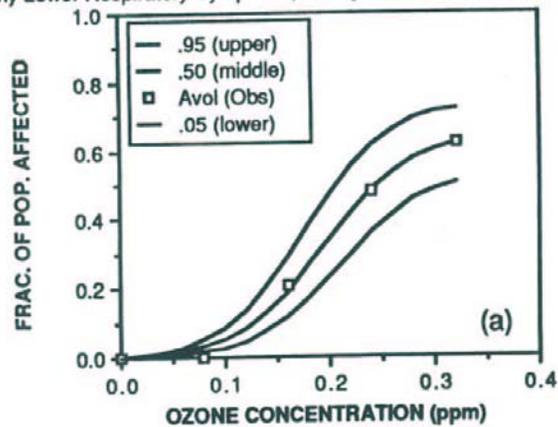


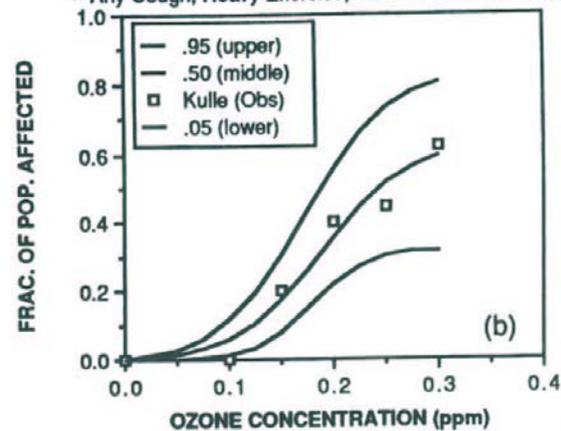
Fig 2. Ozone probabilistic exposure-response relationships for FEV1 decrement ≥ 20 percent (heavy exercise).

Source: Hayes et al. (1989)

PROBABILISTIC EXP.-RESP. RELAT. (Avol Data)
 -- Any Lower Respiratory Symptoms, Heavy Exercise, Various Credible Levels --



PROBABILISTIC EXP.-RESP. RELAT. (Kulle Data)
 -- Any Cough, Heavy Exercise, Various Credible Levels --



PROBABILISTIC EXP.-RESP. RELAT. (McDonnell Data)
 -- Any Cough, Heavy Exercise, Various Credible Levels --

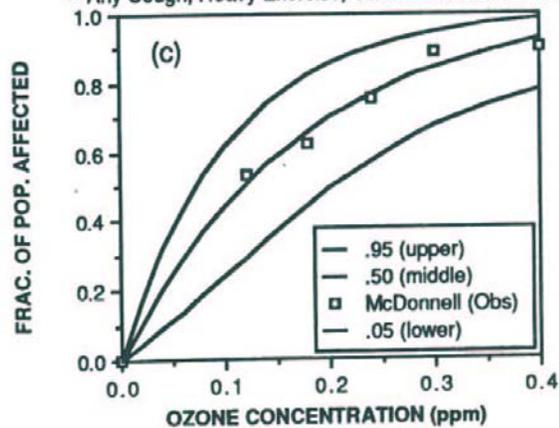
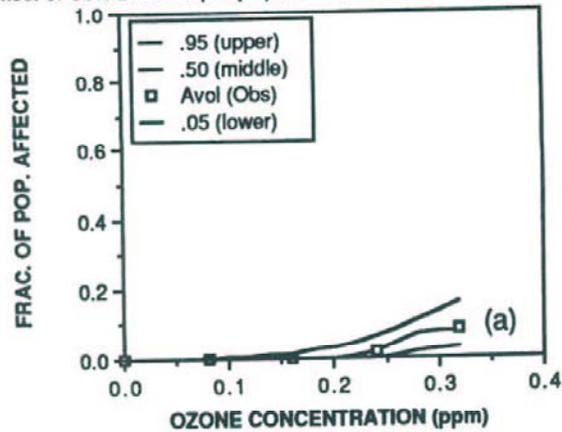


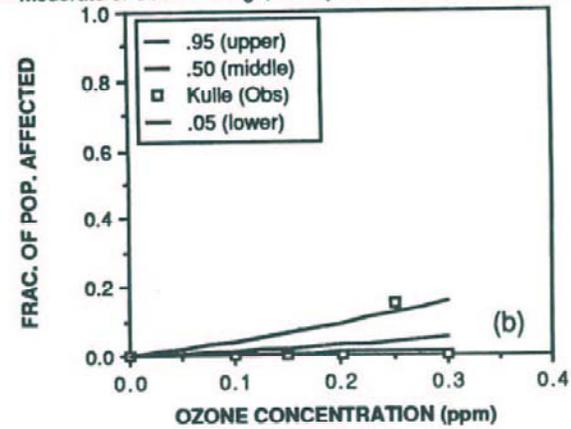
Fig 3. Ozone probabilistic exposure-response relationships for any cough symptoms (mild, moderate or severe; heavy exercise).

Source: Hayes et al. (1989)

PROBABILISTIC EXP.-RESP. RELAT. (Avol Data)
 -- Mod. or Sev. Low. Resp. Symp., Heavy Exercise, Various Credible Levels --



PROBABILISTIC EXP.-RESP. RELAT. (Kulle Data)
 -- Moderate or Severe Cough, Heavy Exercise, Various Credible Levels --



PROBABILISTIC EXP.-RESP. RELAT. (McDonnell Data)
 -- Moderate or Severe Cough, Heavy Exercise, Various Credible Levels --

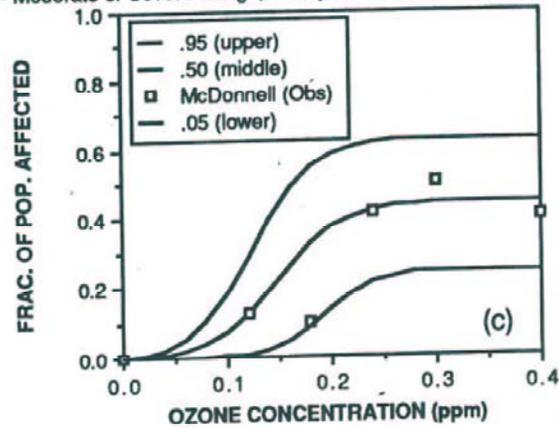
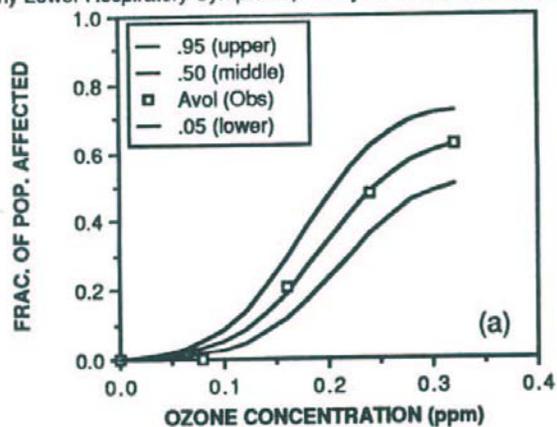


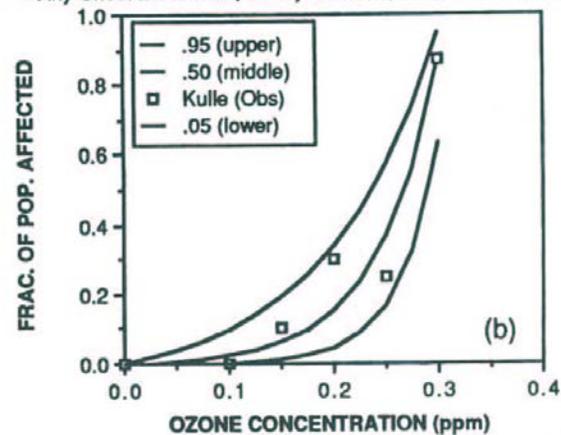
Fig 4. Ozone probabilistic exposure-response relationships for moderate or severe cough symptoms (heavy exercise).

Source: Hayes et al. (1989)

PROBABILISTIC EXP.-RESP. RELAT. (Avol Data)
 -- Any Lower Respiratory Symptoms, Heavy Exercise, Various Credible Levels --



PROBABILISTIC EXP.-RESP. RELAT. (Kulle Data)
 -- Any Chest Discomfort, Heavy Exercise, Various Credible Levels --



PROBABILISTIC EXP.-RESP. RELAT. (McDonnell Data)

-- Any Chest Discomfort, Heavy Exercise, Various Credible Levels --

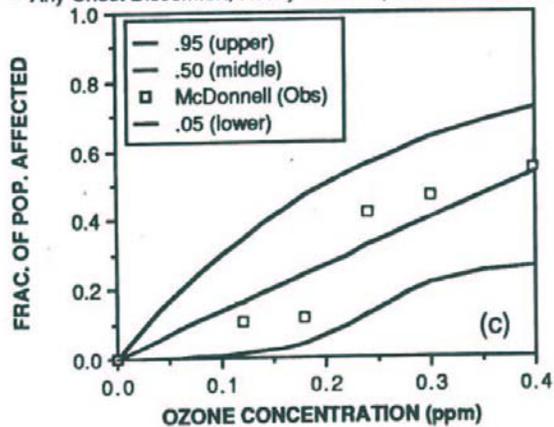


Fig 5. Ozone probabilistic exposure-response relationships for any chest discomfort (mild, moderate, or severe; heavy exercise).

Source: Hayes et al. (1989)

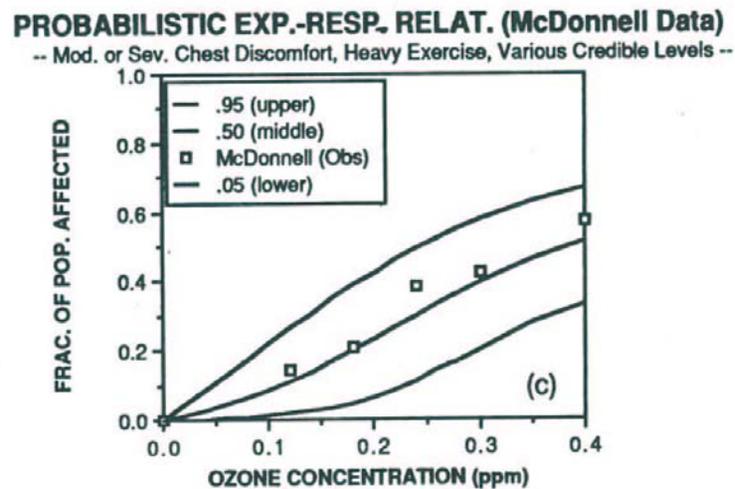
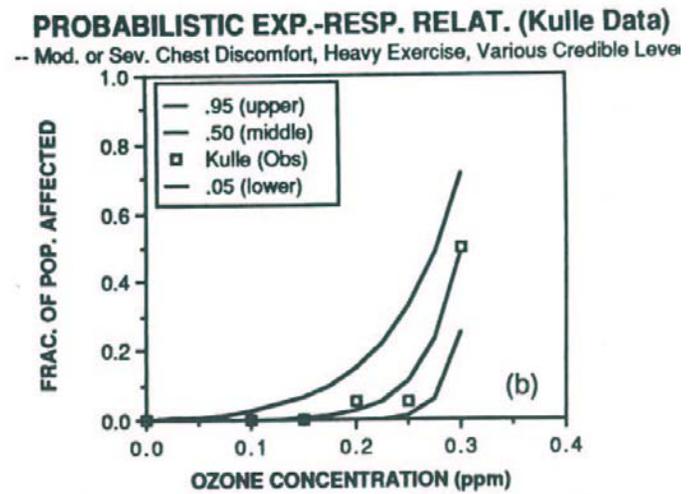
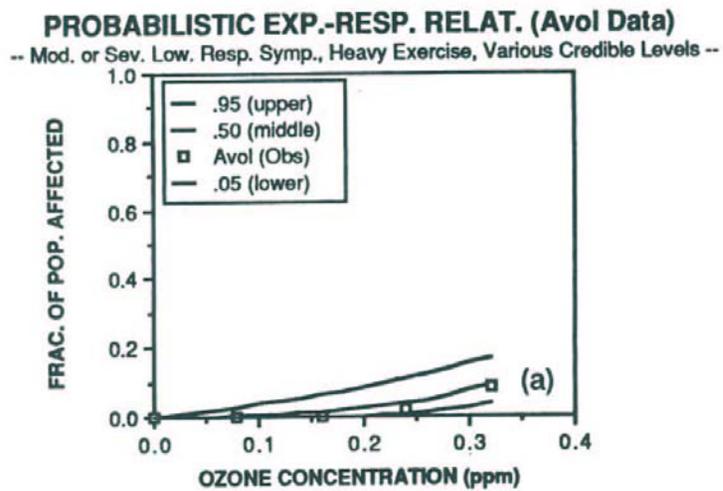
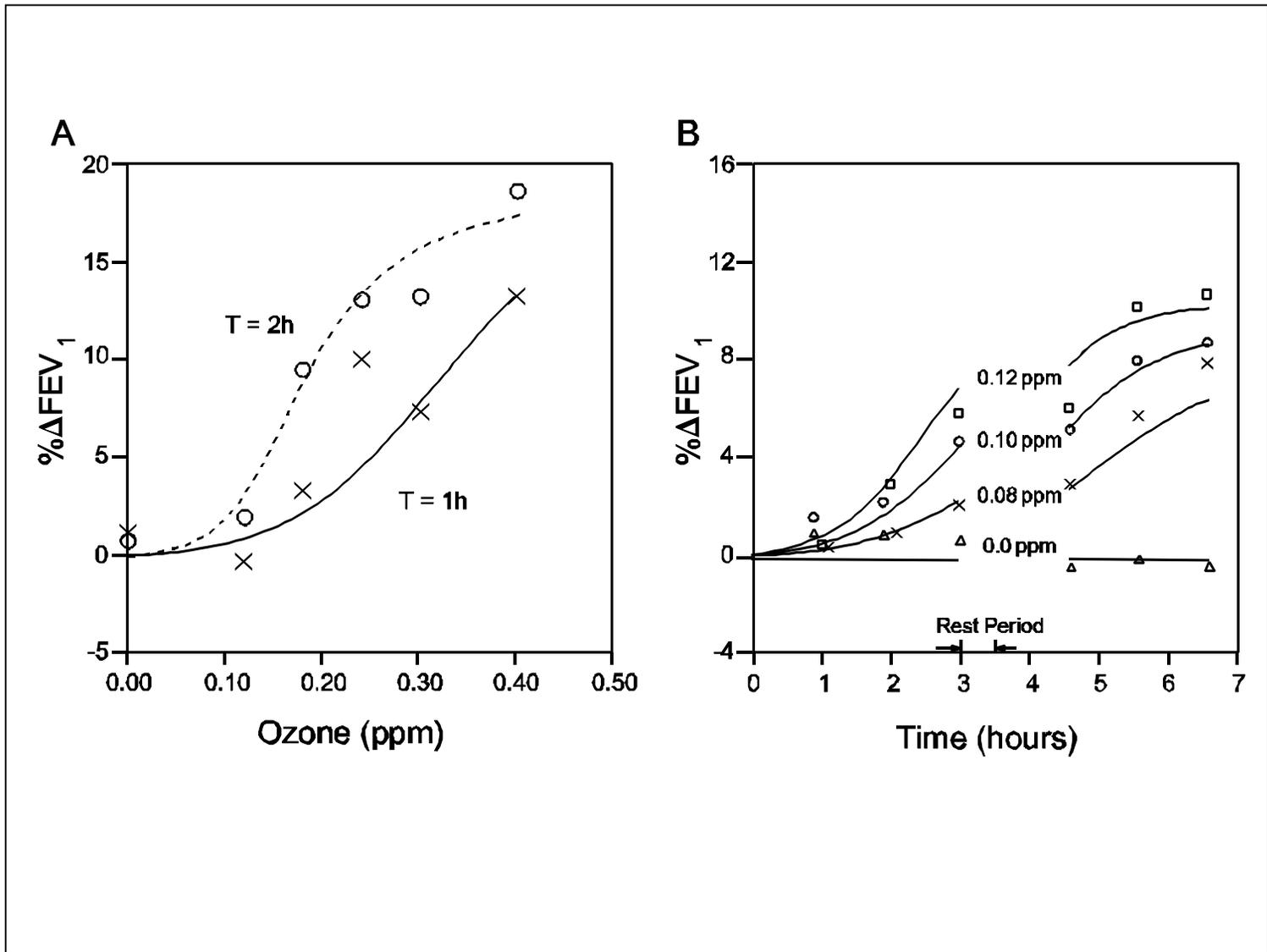


Fig 6. Ozone probabilistic exposure-response relationships for moderate or severe chest discomfort (heavy exercise).

Source: Hayes et al. (1989)



Source: McDonnell and Smith (1994); from USEPA O₃ Staff Paper (1996), p. 34

Fig. 7. Predicted mean decrements in FEV₁ for 1- and 2-hour exposures to ozone with intermittent heavy exercise (A) and 6.6-hour exposures with moderate prolonged exercise (B).

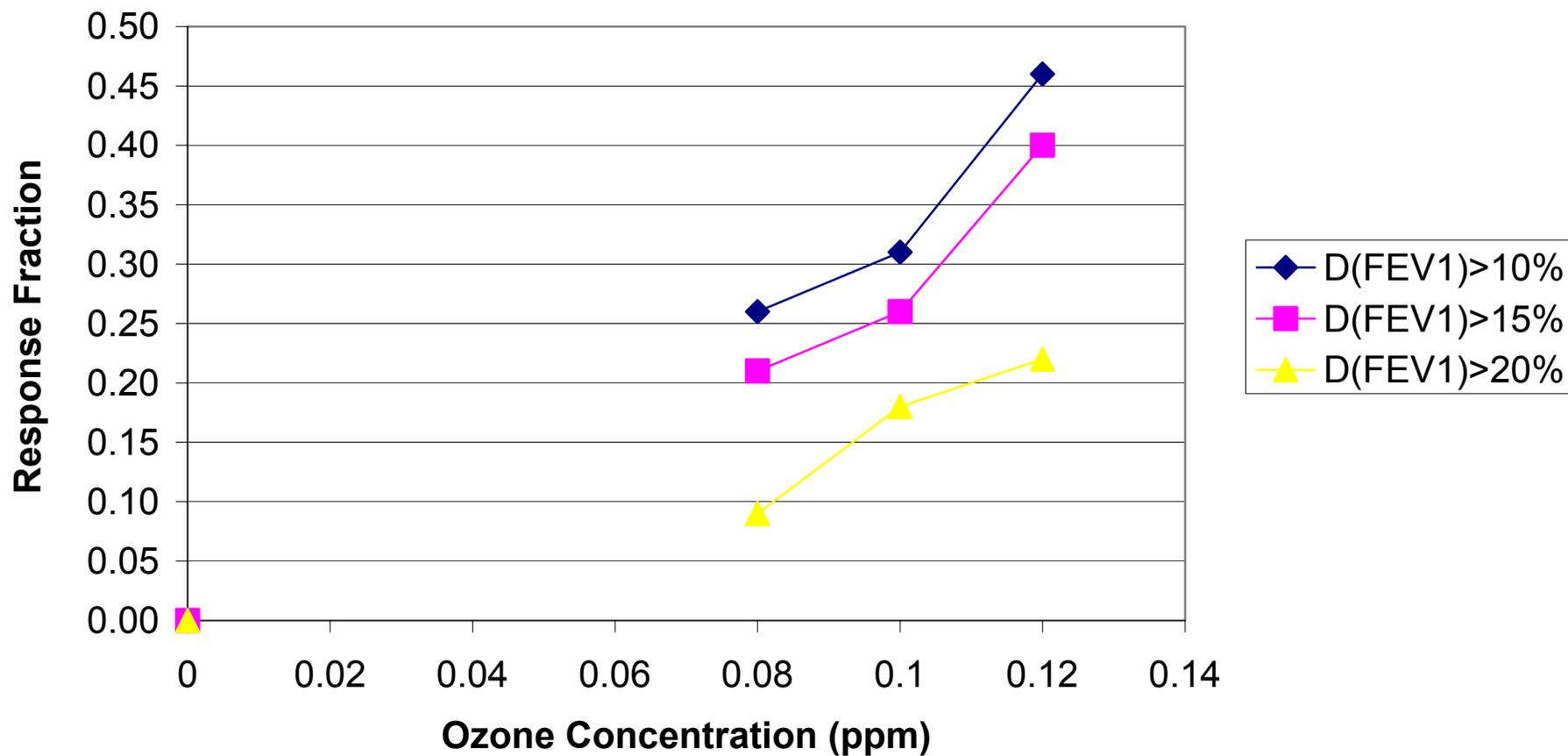


Fig. 8. Decrement in FEV1 for 6.6-hour exposures to filtered air and ozone at 0.08, 0.10, and 0.12 ppm under moderate exercise (derived from Follinsbee et al. 1991, as reproduced in Draft Ozone Staff Report, Fig. 11-12, p. 11-55).

A Review of Statistical Methods Used in Epidemiologic Studies of Ambient
Ozone and Mortality
Chapter 12 and Appendix B

Paul Switzer, Ph.D.
Department of Statistics
Stanford University

August 30, 2004

On June 21, 2004 the California Environmental Protection Agency [CEPA] issued its Review of the California Air Quality Standard for Ozone [O3C]. This review addresses Chapter 12 of O3C, which presents a CEPA synthesis of epidemiologic studies, and portions of Appendix B titled Recommendations for an Ambient Air Quality Standard for Ozone [O3R]. While this review deals principally with mortality time series studies, the issues raised here may be applicable to the much larger corpus of morbidity studies. It would be prudent on the part of CEPA to examine the morbidity studies in light of the fundamental statistical issues that are raised below.

The epidemiologic studies cited by CEPA point to a string of inconsistent results when variations in ambient ozone are related to variations in mortality, both for short term and long term exposures. Examples of the pattern of inconsistent and inconclusive findings include the following:

- Sharply different mortality effect estimates for summer and winter, which should not exist under the model of additive proportional effects that is used in the analyses.
- Instability of ozone mortality effect estimates resulting from different model specifications of weather effects and time trends.
- Instability of ozone effect estimates resulting from different selections of monitoring sites within cities.

- Heterogeneity of ozone effect estimates across cities.
- Ozone effect lags that are inconsistent across cities and across studies.
- Exposure-response relations that are inconsistent across cities and across studies.
- Inconsistencies between short-term and long-term effect studies

This review addresses several statistical concerns that were not adequately addressed by CEPA. The issues addressed here are grouped into the following categories.

1. Confounding of weather and time trends with ozone effects
2. Heterogeneity of ozone effects and effect modification
3. Heterogeneity of exposure within study areas
4. The relation between exposure and response
5. Mortality displacement
6. Long-term ozone-mortality studies

1. CONFOUNDING OF WEATHER AND TIME TRENDS WITH OZONE EFFECTS

Estimates of ozone mortality effects vary from city to city, are often non-significant, and frequently negative. This is readily seen, for example, in the reanalyzed multi-city study by HEI [HEI 2003]. Estimates rarely exceed a fraction of a percent of daily mortality for the range of controllable ozone concentrations. Because the effect is so small compared with the range of daily Poisson mortality variation, great care must be taken to assure that estimates are not sensitive to modeling choices.

Ozone variation is substantially correlated with weather variations. Therefore, special care is needed in separating ozone effects from the much larger effects of weather. The HEI reanalysis studies point to the sensitivity of ozone pollutant effect estimates to the precise way in which weather effects are modeled and this is noted as

well in O3C [12-68, 12-70]. Greater flexibility in modeling weather effects is shown to substantially reduce the apparent pollutant effect estimates. Indeed, it is quite possible that the recommended further investigation of weather effects would show that remaining ozone effects are substituting for heretofore unmodeled weather effects

While CEPA acknowledges ozone effect sensitivity to alternative modeling of weather, it does not adequately address a critical modeling assumption -- *additivity* of weather and ozone effects, an assumption that is built into all the ozone effect estimates cited by CEPA. The additivity assumption is very strong and it presumes that the incremental effects of ozone would be the same at any level of temperature and humidity. Thus, for example, the presumption is that incremental ozone effects are the same at moderate temperature and humidity as they are at extreme temperature and humidity. If this assumption should fail, then additive modeling of ozone effects, as relied on by CEPA, can lead to uninterpretable estimates of ozone effects. This is especially true when effects are not proportionally related to ozone, an issue discussed later in this review.

CEPA has argued strongly throughout Section 12.4 of O3C that only season-specific ozone effects are meaningful. Underlying the CEPA argument is the unstated supposition that ozone and weather effects are indeed not additive and/or that the community response to ozone is not proportional to concentration. Additionally, the sharp disagreement between summer and winter finding for ozone effects argues that the effects of weather may not have been adequately addressed.

While it is quite reasonable to consider season-specific modeling of ozone effects, there has still been no investigation into these additivity and proportionality assumptions that are still very much a part of most of the season-specific studies. Nor has there been

an investigation into the effects of alternative models for weather effects for the season-specific studies. Shortcomings of existing modeling strategies are made evident by the troubling finding that higher ozone levels appear to be beneficial on winter days [12-71].

Without a systematic exploration of weather and ozone interaction, we cannot conclude that some part of the weather effects is mistakenly attributed to ozone, even for studies that are season specific. There are several strategies for incorporating non-additivity, i.e., allowing for differential ozone effects at different levels of the confounding variables. One strategy is to allow for the estimation of a joint response surface that includes both ozone and weather. Roberts [2004a] explored the mortality effects of air pollution [particulate matter PM10] and temperature, in combination, using a nonparametric response surface. His analyses indicate that the usual additivity presumption may not be plausible. A second strategy for incorporating non-additivity of ozone effects is to stratify, effectively to estimate different ozone effects for different temperature strata; see Morris and Naumova (1998), who examined mortality and morbidity effects of CO for different temperature strata.

CEPA relies heavily on time-series studies that adjust for seasonal and time trend co-evolution of ozone and mortality. Co-evolution is an important source of potential confounding in time-series studies, which is only partially mitigated by season-specific analysis. The way in which models adjust for co-evolution could substantially affect conclusions regarding health effects on short time scales. The recent HEI (2003) report includes a number of examples that illustrate the sensitivity of pollutant effects estimates to specification of the time trend adjustment. As expected, HEI concluded time trends with richer parameterization capture shorter scale fluctuations in mortality and reduce the

apparent pollutant effects. There is no obvious level at which to cut off time trend effects in favor of pollutant effects and this conundrum is not resolved and is acknowledged by CEPA [12-70].

Although there have been a number of studies that attempt to relate ozone with various health endpoints, CEPA recognizes that the epidemiologic findings are inconsistent with regard to choice of ozone measure and choice of study area, and have unresolved potential for covariate confounding [B-21].

2. HETEROGENEITY OF OZONE EFFECTS AND EFFECT MODIFICATION

CEPA correctly emphasizes the importance of the 90-city study here in that the same modeling strategy was used for all cities in the study. The reanalysis of this study in HEI (2003) reduced pollutant effect estimates and increased their associated standard error estimates. The real issue is: when are separate city analyses combinable [as in Figure 12-2 of O3C] and what is the interpretation of a combined estimate when there are genuine inter-city differences among effect estimates?

The approach used in the NMMAPS multi-city study by Samet. *et al.* (2000) and in their reanalysis HEI (2003) is to allow pollutant effects to be different in different cities and to model this variation as random. Adopting the random effects approach introduces the notion of an overall *population* mean effect. However, this population mean is a model construct which does not reflect the inter-city differences that are part of the model. Formal statistical tests to detect overall heterogeneity among cities will not be informative because of the low power of these tests, as clearly pointed out in the HEI Special Panel review.

It is important to understand the sources of inter-city differences among ozone effect estimates. Without a clear understanding, we cannot rule out the possibility that effect estimates are model artifacts. Therefore, there has been a determined but incomplete effort to relate inter-city effect differences to characteristic differences among cities such as demographics, climate, etc. This is called “effect modification” and is a potentially useful approach. Some putative effect modifiers for which data are available are variable demographic characteristics, climate statistics, proximity to pollutant sources, or statistical summaries of pollutant concentrations akin to climate statistics, and the statistical precision of the effect estimates. However, Samet *et al.* (2000) could not identify any statistically significant pollutant effect modifiers among those that they examined in their 90-city study. Disparities among cities and different studies could arise, for example, through incorrect treatment of confounding variables or an incorrect characterization of the exposure-response relationship.

Finally, it is important to draw attention to regulatory implications of unresolved differences among ozone effect estimates for different cities. The regulatory question concerns the implied reduction in health effects that could be expected from a specific regulatory standard. For example, based on results from the multi-city studies, it is reasonable to suppose that a reduction of ambient ozone will produce no health benefit in some cities, even based on random-effects models.

3. HETEROGENEITY OF EXPOSURE WITHIN STUDY AREAS

Exposures to ambient ozone will vary across a community on any given day. CEPA seems not to have taken account of exposure variability in its assessments,

although inconsistent findings across different age groups and across seasons [Anderson *et al.*, 1998] suggests that exposure variability can impact estimates of ozone effects that are based on assumptions that ignore heterogeneity of exposure. There are two sources of population exposure variability for a given ambient concentration. One obvious source is the heterogeneity among individual microenvironmental trajectories, such as variations in time spent outdoors, variations in residential and workplace penetration and air exchange factors. A second source of exposure variation is the spatial heterogeneity of ozone concentrations, which induces different exposures relative to the monitoring site(s) used to measure ambient ozone.

What is really relevant is whether or not reductions in ambient monitor-site ozone would produce the reductions in community mortality or other health effects that are implied by models that do not consider how exposure is related to measured ambient ozone. Where multiple ozone monitoring sites are available, it would be a salutary exercise to compare effects estimated using different monitoring sites. If one monitor records proportionally lower ambient ozone than a second monitor, the first monitor will show a correspondingly larger unit ozone effect because both monitors are used to explain the *same* community-wide time series of health effects

If the concentration-response relationship were exactly linear, and if the population average exposure to ambient ozone was in constant proportion to the reported ambient ozone, then it could be argued that the estimated effect per unit increase in ambient ozone is not affected by population variability in exposure. But it is important to distinguish between the unit effects of ambient ozone and the unit effects of ozone exposure. The proportionality factor relating population exposures to reported ambient

ozone is likely to be different in different cities. The model-estimated unit effects of ambient ozone would not then be comparable across cities without an understanding of city-specific relationships between exposures and reported ozone. Combined ozone effect estimates across cities, such as those reported by NMMAPS, implicitly and implausibly assume that the relation between monitored ambient ozone and ambient ozone exposure is the same across cities. .

If cross-correlations among monitors are indeed high, and average population exposure is indeed approximately proportional to the monitored ozone values, and the exposure-response relation is indeed linear, then ozone effect estimates should be about the same using any standardized combination of monitors to represent exposure. If this is empirically contradicted, then exposure-response may not be linear, monitored ozone poorly represents population exposure, and the effects of ozone reductions would be hard to anticipate based on the kinds of models and assumptions that are relied on by CEPA. Studies that try to relate ambient ozone with personal exposure show inconsistent results as noted by CEPA [B-14].

4. THE RELATION BETWEEN EXPOSURE AND RESPONSE

A linear [proportional] exposure-response relationship is key to many of the inferences and conclusions that CEPA draws from the studies that it has reviewed. For example, the combined analyses in the multi-site studies rely on a proportionality that relates ozone to mortality that is common to all cities. Studies of exposure error, such as the one by Zeger *et al.* (2000) which concludes that exposure error may bias effect estimates downward, also rely strongly on the proportional effect hypothesis.

Furthermore, the rationale for combining multiple monitors within a city also relies on the assumed proportionality.

Some studies of ozone health effects have indeed discerned a community-level non-linear exposure-response relationship [Burnett *et al.* 1997, Burnett *et al.* 2001, Steib *et al.* 1996, and Ponce de Leon *et al.* 1996]. Additionally, clinical studies with controlled exposure have shown nonlinear relations between exposure and response [Hazucha *et al.* (1992) and Adams (2003), for example]. Exposure measurement error will tend to flatten a non-linear exposure-response curve [Cakmak *et al.* (1999)] making it harder to distinguish between linear and non-linear associations. Although there are simulation study reports that specific threshold exposure-response models for a population could be distinguished even in the presence of exposure measurement error, it is not clear to what extent their findings could be generalized.

When non-proportional effects are allowed in the effect estimation model, then estimated pollutant effects have been seen to depart from proportionality, as was seen in many cities in the multi-city analysis by Daniels *et al.* (2000) and Dominici *et al.* (2002). In these studies, the response is modeled as a low-order parametric spline function of ambient pollution. Application of the spline response model to different cities yielded a variety of response shapes, often with inadequate precision, suggesting that there are statistical difficulties distinguishing between linearity and non-linear spline models. Formal tests for response-function linearity will typically have low statistical power against plausible non-linear alternatives although the implications on non-linearity for regulatory purposes are important. Opting for a linear model because of low statistical power can result in regulatory decisions that will not produce the desired mitigation of

health effects. Unfortunately, CEPA has not addressed the critical issue of exposure-response modeling.

In the multi-city studies, non-linear exposure-response functions for different cities were pooled across cities, as in Schwartz and Zanobetti (2000) and Daniels *et al.* (2000), even though city-to-city differences among pollutant-effect response functions are not obviously within the range of sampling variability. However, a pooled response function, even if it is linear, is not interpretable unless exactly the same exposure-response applies to every city. The putative benefits of ambient pollutant reductions in any particular city cannot be deduced from the pooled response function. However, the conclusions of CEPA rely strongly on questionable commonality and linearity of the pollutant-effect response function, even to the extent that proportionality constants in linear models would need to be the same for every city.

Better insights into the relationship between monitored ambient pollutant concentrations and anticipated community-level ozone health effects could be obtained by explicitly modeling the relationship between monitored ozone and *individual* ozone exposure, such as Dominici *et al.* (2000). However, individual-level *exposure* modeling should go hand-in-hand with the individual-level modeling of pollutant *response* in order to build a model for community-level response to ambient pollution. One can readily construct examples to illustrate that ignoring individual heterogeneity of exposure-response can lead to misleading community-level exposure-response functions.

The assumed linearity [proportionality] between ozone exposure and response has far-reaching regulatory implications. For example, one could double the health effect improvement by doubling the ozone reduction, so there is no obvious regulatory

threshold based on health effects under this assumption. Furthermore, the same reduction of effects could then be achieved by the same reductions in either a high-ozone or a low-ozone city.

A specific regulatory standard implies that society either accepts the pollutant effects at that regulatory threshold or that this threshold is a background that is not affected by regulated sources. So, in the linear exposure-response context, the choice of a regulatory limit boils down to a choice of acceptable risk or a statement about background ozone levels. The regulatory limit cannot be regarded as a threshold for health effects since there is no robust evidence linking the proposed regulatory limits to an epidemiologically demonstrated effects threshold. CEPA has claimed that an exposure-response relationship has been demonstrated in many studies [B-15] while elsewhere it notes that most studies did not explicitly test for thresholds [B-24].

It appears that CEPA has fallen into the trap of basing a standard on the notion of statistically significant associations [B-25]. As remarked earlier, many or most epidemiologic studies do not allow for thresholds in their analysis models. A statistically significant association at a particular concentration level is merely the result of a particular modeling exercise and is not a robust indicator of an effect threshold. Indeed, most analyses presume that there is no effect threshold, which would argue for a zero-level regulation.

There are also fundamental issues related to the compatibility of community linear exposure-response relationships in the context of individual variations in susceptibility, apart from issues related to individual heterogeneity of exposure. For example, suppose that there are individual-specific response thresholds to 1-hour ozone

exposures. Then the assumed linearity in the community response to the ozone index implies that the population proportion that responds to pollutant levels that are between 0 and x is the same as the population proportion that responds to levels between x and $2x$. This is incompatible with plausible assumptions about the population distribution of susceptibility thresholds. Thus, one might question whether the assumed linearity of community exposure-response is even a plausible working hypothesis.

When more than a single index is contemplated, such as a 1-hour and 8-hour daily maximum, then the anticipated benefit of pollutant reductions will depend on the *joint* population distribution of thresholds for the two indices. For example, some individuals may be sensitive to 8-hour ozone but relatively insensitive to 1-hour peaks or vice versa. Therefore, the community response to two different ozone indices is tied to the joint distribution of individual responses. The response of the same individuals to the two different ozone indices needs to be examined. The study of bivariate community response to a pair of ozone indices needs considerable care.

Finally, CEPA raised the issue of model selection in connection with choosing time-series lags and ozone indices that maximize estimates of ozone health effects [12-74, B-16]. The choice of an index is central to the interpretation of exposure-response and the expected mitigation that would be obtained from ambient ozone reductions. Choosing time lags and ozone indices for largest *apparent* effect will definitely introduce statistical bias that exaggerates effect estimates. Indeed, individual studies have used widely varying, but specifically selected, ozone and/or weather indicators. The basis for selection of such indices does not appear to be physiological. Until there is agreement on common indices, there will remain an unassessed statistical selection bias. Simulation

studies by Lumley and Sheppard (2000) have shown that lag selection bias can be of the same order as the estimated effect itself. The evident potential selection bias is disregarded by the conclusions that CEPA has drawn from its synthesis of epidemiologic studies of ozone health effects.

5. MORTALITY DISPLACEMENT

When pollutant mortality effects are largely confined to a frail population with short longevity the effect is referred to as mortality displacement or mortality harvesting. If this were the case, then there would be substantial public policy and regulatory implications. For example, it might be more effective to mitigate pollutant exposure of frail individuals through nursing home and hospital indoor air requirements, as opposed to regulation of ambient pollutant levels. Because of these policy implications it is necessary to address the possibility of mortality displacement if epidemiologic studies are to be used as a basis for air quality standards. CEPA has not addressed mortality displacement.

Several studies have examined the possibility of mortality displacement, and those studies have reached divergent conclusions. Some authors have used a heuristic that mortality displacement in frail populations is characterized by an initial excess of deaths following a high pollution day, followed by a deficit of deaths on later days. For example, Schwartz (2000) analyzed time-series of pollutant and mortality data using distributed-lag models to see if the sequence of lag coefficients exhibited the heuristic pattern. However, the models under which these analyses were performed were not mortality displacement models. The interpretations and conclusions of these studies,

however plausible they sound to the authors, remain purely heuristic because the estimation models are not constructed on the basis of explicit parameters in a mortality displacement model.

Roberts (2004b) examined the behavior of distributed lag coefficients, such as those calculated by Schwartz (2000), using an explicit frail population model and concluded that neither the serial pattern of these coefficients nor their sum is necessarily indicative of mortality displacement. Thus, further study is needed to understand what information is provided by distributed-lag models relevant to the question of mortality displacement.

The study reported by Zeger *et al.* (1999) is based on a decomposition of the mortality time series into shorter and longer time scales, as is a similar study by Zanobetti *et al.* (2000). But these studies did not use an explicit frail population model in their analyses and rely on heuristic arguments. A more promising approach is that of Smith *et al.* (1999), Murray and Nelson (2000), and Roberts (2003), who study pollutant effects explicitly in the context of frail population models. In the examples that they have studied, it appears that any excess pollutant mortality is indeed consistent with mortality displacement in frail populations with mean lifetime on the order of weeks. Their conclusions are in sharp contrast to those of Zanobetti *et al.* (2000).

6. LONG-TERM MORTALITY STUDIES

CEPA refers to several long-term ecological cohort studies of ozone health effects, of which Pope *et al.* (2002) is the latest and most comprehensive. In these long-term studies, both the pollutant concentration and mortality for each city are represented

by single average numbers that do not vary over time. The ecological studies cited by CEPA are cohort studies that are limited to enrolled individuals for whom individual covariate information is available, such as demographic information and smoking habits. The individual covariate information is used to adjust crude mortality rates for the enrolled cohort so as to even out the mortality comparisons between cities. Ozone health effects are inferred by relating time-averaged adjusted mortality to time-averaged monitored ozone across cities.

Both the Pope *et al.* (2002) cohort study of long-term pollutant effects and the Dominici *et al.* (2002) time-series study of short-term pollutant effects involve a comparable number of U.S. cities. However, geographic variation in the cohort studies takes the place of time variation in the time-series studies. City-specific effect modifiers in time-series studies, as discussed earlier, become confounding variables in the cohort studies. A putative confounding variable in a cohort study is one that shows geographic covariability with PM. Thus, demographic adjustments in the cohort studies are a way of accounting for potential confounding of ozone effects by demographic variables. Similarly, between-city variations of co-pollutants and climate variables could be related to between-city variations of ozone and thereby contribute to confounding of ozone effects.

In some important ways, however, a multi-city cohort study suffers from disadvantages vis-à-vis a single-city time-series study. For example, in a time-series study the population at risk is the same each day while in the cohort study the population at risk in each city is different and models are needed to bring the separate at-risk populations into alignment. Also, the assigned ozone concentration for a city needs to be

related not to the average city-wide exposure but rather to the average exposure of the cohort assigned to that city. Further, it is reasonable to suppose that this exposure measurement error will be different for different cities.

In any event, the large study by Pope *et al.* (2002) did not discern an ozone effect on total mortality even when restricted to summer ozone concentrations and to specific causes of death. The negative findings are noted by CEPA [12-51, B-18].

8. CONCLUSIONS

There are a variety of unresolved statistical issues in O3C and O3R. CEPA's synthesis of epidemiologic studies often contains important caveats regarding modeling issues, since effect estimates will be strongly dependent on modeling. However, its own caveats are put aside in drawing conclusions. The available epidemiologic evidence on ozone mortality cannot be used to draw robust conclusions regarding the circumstances and magnitudes of ambient ozone mortality, in particular whether reported ozone effects are causative. The inconsistencies and model dependency of effect estimates seem to have been brushed aside by CEPA, while at the same time it concedes that it is difficult to use the epidemiologic studies for purposes of setting standards [B-15]. Without a clear understanding of the reasons for inconsistent effects estimates, we cannot rule out the possibility that ozone effect estimates are model artifacts. Indeed, CEPA does itself regard the estimates of ozone mortality effects skeptically [B-25]. Below I briefly summarize a few of the points made in this review. The body of the review should be consulted for details and a fuller discussion.

1. **Sensitivity of ozone effect estimates to model specification.** This issue was brought to light in the HEI reanalysis in the context of time and weather adjustments, and serves as a cautionary tale. The reported effects of ozone are often difficult to discern and are inconsistent among cities, regions, seasons, and time lags. Such inconsistencies may be suggestive of modeling inadequacies, particularly in regard to unmodeled confounding and unexplained effect modifiers.

That ozone effect estimates are delicate is not surprising given that they are superimposed on much stronger effects due to concomitant weather variations, for example. Without a clear understanding of the reasons for inconsistent effects estimates, we cannot rule out the possibility that ozone effect estimates are model artifacts.

2. **Enforced additivity in the analysis model.** The analysis models relied on by CEPA assume that ozone effects are necessarily the same at any temperature, even when restricted to summer data. Approaches to mitigate the problem, depending on availability of data, include joint response surface modeling of ozone and its confounders or stratification of the analyses based on confounder categories.
3. **Enforced linearity of exposure-response.** Because ozone health effect estimates are inconsistent across studies, cities, seasons, etc., putative benefits of ambient ozone mitigation are difficult to know. Enforced model linearity of exposure-response, as in the case of the analysis models the CEPA relies on, conceals heterogeneity of response. Pooling of response functions to obtain linearity is not statistically justified and leads to regulatory dilemmas.
4. **Spatial variability of ozone health effect estimates within cities.** There has been insufficient attention to the issue of spatial variability of effect estimates within cities based on selection or combination of monitors.
5. **Incomplete characterization of the relations between ambient ozone exposure, individual PM exposure, individual PM susceptibility to health effects, and community level health effect measures.** The models that CEPA uses for the analysis of community health effects of ozone do not have any link to individual response functions.
6. **No evaluation of the possibility of mortality displacement.** Some studies suggest that acute mortality effects are consistent with mortality displacement in frail populations. This issue is important for public policy and therefore needs to be studied so that regulatory decisions can truly address mitigation.
7. **Unresolved inconsistencies of ozone effect estimates.** The following inconsistencies are unresolved: seasonal differences, regional grouping, spatial

heterogeneity both between cities and within cities, time lag selection, and treatment of gaseous pollutant confounders.

REFERENCES

Adams WC. (2003). Comparison of chamber and face mask 6.6-hour exposure to 0.08 ppm ozone via square-wave and triangular profiles on pulmonary responses. *Inhal Toxicol* 15:265-81.

Anderson HR, Ponce de Leon A, Bland JM, Bower JS, Emberlin J, Strachan DP. 1998. Air pollution, pollens, and daily admissions for asthma in London 1987-92. *Thorax* 53:842-8.

Burnett RT, Cakmak S, Brook JR, Krewski D. 1997. The role of particulate size and chemistry in the association between summertime ambient air pollution and hospitalization for cardiorespiratory diseases. *Environ Health Perspect* 105:614-20.

Burnett RT, Smith-Doiron M, Stieb D, Raizenne ME, Brook JR, Dales RE, Leech JA, Cakmak S, Krewski D. 2001. Association between ozone and hospitalization for acute respiratory diseases in children less than 2 years of age. *Am J Epidemiol* 153:444-52.

Cakmak, S.; Burnett, R.T.; Krewski, D. (1999) Methods for detecting and estimating population threshold concentrations for air pollution-related mortality with exposure measurement error. *Risk Anal.* 19: 487-496.

Daniels, M.; Dominici, F.; Samet, J. M.; Zeger, S. L. (2000) Estimating particulate matter-mortality dose-response curves and threshold levels: an analysis of daily time-series for the 20 largest US cities. *Am. J. Epidemiol.* 152: 397-406.

Dominici, F.; Zeger, S. L.; Samet, J. (2000) A measurement error model for time-series studies of air pollution and mortality. *Biostatistics* 1: 157-175.

Dominici, F.; Daniels, M.; Zeger, S.L.; Samet, J.M. (2002) Air pollution and mortality: estimating regional and national dose-response relationships. *J. Am. Statist. Assn.* 97: 100-111.

Hazucha MJ, Folinsbee LJ, Seal E Jr. 1992. Effects of steady-state and variable ozone concentration profiles on pulmonary function. *Am Rev Respir Dis* 146:1487-93.

Health Effects Institute (2003) Revised analyses of the National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Special report. Boston, MA.

Lumley, T.; Sheppard L. (2000) Assessing seasonal confounding and model selection bias in air pollution epidemiology using positive and negative control analyses.

Environmetrics 11: 705-717.

Morris, R. D.; Naumova, E. N.; (1998). Carbon monoxide and hospital admissions for congestive heart failure: evidence of increased effect at low temperatures. *Environ. Health Perspect.* 106: 649-653.

Murray, C. J.; Nelson, C. R. (2000) State-space modeling of the relationship between air quality and mortality. *J. Air Waste Manage. Assoc.* 50: 1075-1080. Pope, C.A.III; Burnett, R.T.; Thun, M.J.; Calle, *et al.* (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *J. Am. Medical Assn.* 287: 1132-1141.

Ponce de Leon A, Anderson HR, Bland JM, Strachan DP, Bower J. 1996. Effects of air pollution on daily hospital admissions for respiratory disease in London between 1987-88 and 1991-92. *J Epidemiol Community Health* 50 Suppl 1:s63-70.

Pope, C.A.III; Burnett, R.T.; Thun, M.J.; Calle, *et al.* (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *J. Am. Medical Assn.* 287: 1132-1141.

Roberts, S. (2004a) Interactions between particulate air pollution and temperature in air pollution mortality time series studies. Environmental Research, In Press.

Roberts, S. (2004b) An investigation of distributed lag models in the context of air pollution and mortality time series analysis. Journal of Air and Waste Management Association. In Press.

Roberts, S. (2003) Personal communications. Stanford University Department of Statistics.

Samet, J. M.; Dominici, F.; Zeger, S. L.; Schwartz, J.; Dockery, D. W (2000) National morbidity, mortality, and air pollution study. Cambridge, MA: Health Effects Institute; research report no. 94.

Schwartz, J. (2000) Harvesting and long term exposure effects in the relation between air pollution and mortality. *Am. J. Epidemiol.* 151: 440-448.

Schwartz, J.; Zanobetti, A. (2000) Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. *Epidemiology* 11: 666-672.

Smith, R. L.; Davis, J. M.; Speckman, P. (1999) Assessing the human health risk of atmospheric particles. In: Environmental Statistics: Analysing Data for Environmental Policy: proceedings of a Novartis Foundation Symposium; May 1998; London, England. New York, NY: Wiley. [Novartis Foundation Symposium 220, pp. 59-79].

Stieb DM, Burnett RT, Beveridge RC, Brook JR. 1996. Association between ozone and asthma emergency department visits in Saint John, New Brunswick, Canada. *Environ Health Perspect* 104:1354-60.

Zanobetti, A.; Wand, M. P.; Schwartz, J.; Ryan, L. M. (2000) Generalized additive distributed-lag models: quantifying mortality displacement. *Biostatistics* 1: 279-292.

Zeger, S. L.; Dominici, F.; Samet, J. (1999) Harvesting-resistant estimates of air pollution effects on mortality. *Epidemiology* 10: 171-175.

Zeger, S. L.; Thomas, D.; Dominici, F.; Samet, J. M.; Schwartz, J.; Dockery, D.; Cohen, A. (2000) Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ. Health Perspect.* 108: 419-426.