UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

DATE

EPA-CASAC-14-XXX

The Honorable Gina McCarthy
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: CASAC Review of the EPA's *Second Draft Policy Assessment for the Review of the Ozone National Ambient Air Quality Standards*

Dear Administrator McCarthy:

The Clean Air Scientific Advisory Committee (CASAC) Ozone Review Panel met on March 25 - 27, 2014, to review the EPA's *Second Draft Policy Assessment for the Review of the Ozone National Ambient Air Quality Standards*, hereafter referred to as the Second Draft PA. This letter highlights the chartered CASAC's consensus advice, followed by consensus responses to the charge questions from the Agency. We have also attached individual review comments from the CASAC Ozone Review Panel.

Overall, the Second Draft PA is an excellent summary of information needed to judge the adequacy of the current National Ambient Air Quality Standards (NAAQS) for ozone and to consider alternative standards. The information on emissions, atmospheric chemistry and common patterns of ozone concentration is presented clearly and is appropriately characterized. The sections on "background" ozone (i.e., ozone that originates from precursors from natural sources or manmade international emissions) are extensive and appropriately characterized, although in our consensus responses to charge questions, we have some technical comments that should be addressed to improve the coverage of this issue. More importantly, the Second Draft PA is not clear as to how background estimates might impact ~~both~~ the primary and secondary standards and whether these impacts ~~how this might~~ may differ regionally. The Second Draft PA cites a 2002 court decision (*American Trucking Associations, Inc. v. EPA, 283 F.3d at 379*) that allows the EPA to consider relative proximity to peak background levels when evaluating alternative standards but it also cites a case where the court said "attainability and technological feasibility are not relevant considerations in the promulgation of the NAAQS.~~"~~" (*American Petroleum Institute v. Costle*, 665 F. 2d at 1185). ~~We were left wondering~~ The Second Draft

Commented [HS1]: Maybe instead say "The Draft PA was silent as to how EPA intends to …."

PA was silent as to how the EPA intends to navigate between these two legal guidelines when considering background ozone in a policy and standard-setting context. This question became an important issue in the CASAC deliberations as we listened to public ~~commenters who presented~~comments that included information ~~on~~regarding high background levels in the intermountain Western United States.

In addressing the adequacy of the primary standard, the Second Draft PA presents scientifically sound information on the health effects evidence for each major effect category: lung function decrements, pulmonary inflammation, respiratory symptoms, respiratory morbidity and respiratory mortality. The CASAC finds scientific justification that current evidence and exposure/risk information call into question the adequacy of the current standard. Furthermore, there is scientific support for the need to revise the standard to achieve additional public health protection. The CASAC supports the scientific rationale presented in the Second Draft PA on these points.

The CASAC concurs with the staff's justifications in the Second Draft PA for retaining the current indicator (ozone), averaging time (maximum daily 8-hour average) and form (4th highest maximum daily 8-hour average, averaged over three years). The indicator of ozone is appropriate based on its causal or likely causal associations with multiple adverse health outcomes and its representation of a class of pollutants known as photochemical oxidants. The current 8-hour averaging time is justified by the combined evidence from epidemiologic and clinical studies referenced in Chapter 4. The CASAC concurs that the ozone standard should be based on the fourth highest, daily maximum 8-hour average value (averaged over three years). This averaging time provides programmatic stability by- allowing for atypical meteorological conditions that can lead to abnormally high ambient ozone concentrations while providing health protection.

[Note to the Reader: At this time, the Panel is still deliberating regarding its advice for a scientifically based upper bound to the range of levels being considered for the primary standard. The Panel will deliberate on the science-based upper level during its upcoming teleconferences~~,~~ and ~~we expect to~~will revise this response to the charge question~~.~~. The Panel will also deliberate on the scientific basis for CASAC's recommendation on the secondary standard so the following text is subject to change.]

With respect to the secondary standard, the CASAC supports the scientific conclusion in the Second Draft PA that the current secondary standard is not adequate to protect against current and anticipated welfare effects of ozone on vegetation. We support retaining the current indicator (ozone) but establishing a revised form of the standard to be the biologically-relevant W126 index accumulated over a 12-hour period (8 am – 8 pm) over the 3-month period ~~of a single year~~ resulting in the maximum value of W126 (henceforth W126). The CASAC recommends a W126 level within the range of 7 ppm-hr to ~~13~~ 15 ppm-hr to protect against current and anticipated welfare effects of ozone on vegetation, taking into account protection of median tree species from annual relative biomass loss of 2% or less and protection of crop species from annual loss of 5% or less. These combinations of indicator, form, averaging time, and level are scientifically justifiable given evidence of current and anticipated welfare effects as captured in the Second Draft PA.

The CASAC recommends that EPA facilitate research needed for the next review of the ozone NAAQS.

Commented [SH2]: Edits from Ed Avol

Commented [FM3]: I do not believe this wording is strong enough. We are willing to say that the secondary standard is not adequate to protect against welfare effects on vegetation. We should clearly state that the current primary standard is not protective of human health.

Commented [R,AG4]: At the meeting, we discussed using a single year or the average of three years, and never came to closure. I am strongly in favor of a three year averaging period, potentially decreasing the level to account for such, for the same reasons that we use a three year average for the primary standard. I would urge everyone to read the reasoning for using a three year average (Page 4-7 of the current PA). Given that this was not agreed upon at the meeting, it should be left out.

Commented [R,AG5]: I recall from the meeting we had discussed and decided upon a range of 7-15. It is probably not appropriate to change at this time. This should be further discussed. We should openly discuss the reasons for changing this. For now, it should reflect our discussions at the meeting.

For the health-based standard, we note that the Second Draft PA outlines key uncertainties and research that needs to be addressed for future reviews of the health-based standards. Specifically, we underscore the need for research to address the characterization of the exposure-response function, the identification of population thresholds, the role of co-pollutants and temperature in modifying or contributing to ozone effects, alternative modeling specifications, population-based information on human exposure for at-risk populations, time-activity data to improve population-based exposure and risk assessment and the characterization of background levels. For the secondary standard, the Second Draft PA also identifies uncertainties and needed research to develop data and better methods for extrapolating results to plant species for which exposure-response functions have not been developed; assessing the effects of ozone on climate (and the effects of climate on ozone); characterizing the effects of ozone on whole ecosystem structure and function; and evaluating how the public judges the adversity of various kinds of ecological effects including foliar injury and estimated reduced tree biomass growth. This policy-relevant research could be conducted in collaboration with other federal and non-governmental organizations to improve our understanding of ozone effects in support of the next review of the ozone NAAQS.

Although CASAC was not asked to comment ~~any questions~~ about international transport of ozone, ~~but~~ we would like to call your attention to this issue as a ~~separate~~ matter separate from ~~not taken into account in~~ our ~~development of~~ advice regarding the standard. North American background ozone is defined by the EPA as the ozone that would be present in U.S. surface air in the absence of North American anthropogenic emissions. It can be estimated using global models by conducting simulations with North American anthropogenic emissions set to zero. Results indicate that background is only partly natural (lightning, biosphere, fires, stratospheric influence) and is enhanced by anthropogenic sources outside North America. Estimates of this external anthropogenic enhancement are fairly consistent across models [Fiore et al., 2009]. Zhang et al. [2011] estimate it to be 9 ppb on average in spring-summer at low-altitude sites and 13 ppb at high altitude sites (>1.500 m elevation), with half ~~from~~ being driven by atmospheric photochemistry involving anthropogenic methane and half from anthropogenic NOx and non-methane Volatile Organic Compounds (VOC) emitted in other continents. There is currently no international legal agreement on ozone or its precursors that would effectively deal with long-range transport, despite the recommendations by the National Academy of Sciences (2009) and the Task Force on Hemispheric Transport of Air Pollution (2010) that such an agreement be sought. Given the significant portion of ozone coming from anthropogenic sources outside North America, CASAC recommends that EPA seek opportunities for international cooperation to reduce long-range transport of ozone.

Commented [DC6]: I recommend retaining the original version because the EPA does not preclude us from considering the issue. So to avoid potential misinterpretation, the original version is less restrictive and thus, better. Holly may be a better judge on this.

Commented [SH7]: Edit from Ed Avol

Overall, we find the Second Draft PA to be adequate for its intended purpose of providing a strong scientific basis for findings regarding the lack of adequacy of current primary and secondary ozone air quality standards, for scientifically justifiable indicators, averaging times, and forms for alternative revised primary and secondary standards, and for scientifically justifiable ranges of levels for each of the primary and secondary standards. The CASAC appreciates the opportunity to provide advice on the Second Draft PA and looks forward to receiving the Agency's response.

Sincerely,

Dr. H. Christopher Frey, Chair
Clean Air Scientific Advisory Committee

Enclosures

NOTICE

This report has been written as part of the activities of the EPA's Clean Air Scientific Advisory Committee (CASAC), a federal advisory committee independently chartered to provide extramural scientific information and advice to the Administrator and other officials of the EPA. The CASAC provides balanced, expert assessment of scientific matters related to issues and problems facing the agency. This report has not been reviewed for approval by the agency and, hence, the contents of this report do not necessarily represent the views and policies of the EPA, nor of other agencies within the Executive Branch of the federal government. In addition, any mention of trade names or commercial products does not constitute a recommendation for use. The CASAC reports are posted on the EPA website at: http://www.epa.gov/casac.

U.S. Environmental Protection Agency
Clean Air Scientific Advisory Committee (CASAC)

CHAIR

Dr. H. Christopher Frey, Distinguished University Professor, Department of Civil, Construction and Environmental Engineering, College of Engineering, North Carolina State University, Raleigh, NC and Visiting Professor, Department of Civil and Environmental Engineering, Adjunct Professor, Division of Environment, Hong Kong University of Science and Technology

MEMBERS

Mr. George A. Allen, Senior Scientist, Northeast States for Coordinated Air Use Management (NESCAUM), Boston, MA

Dr. Ana Diez-Roux, Dean, School of Public Health, Drexel University, Philadelphia, PA

Dr. Jack Harkema, Professor, Department of Pathobiology, College of Veterinary Medicine, Michigan State University, East Lansing, MI

Dr. Helen Suh, Interim Chair, Director of Population Health Doctoral Program, Department of Health Sciences, Northeastern University, Boston, MA

Dr. Kathleen Weathers, Senior Scientist, Cary Institute of Ecosystem Studies, Millbrook, NY

Dr. Ronald Wyzga, Technical Executive, Air Quality Health and Risk, Electric Power Research Institute, Palo Alto, CA

SCIENCE ADVISORY BOARD STAFF

Mr. Aaron Yeow, Designated Federal Officer, U.S. Environmental Protection Agency, Science Advisory Board (1400R), 1200 Pennsylvania Avenue, NW, Washington, DC 20460

U.S. Environmental Protection Agency
Clean Air Scientific Advisory Committee
Ozone Review Panel

CHAIR

Dr. H. Christopher Frey, Distinguished University Professor, Department of Civil, Construction and Environmental Engineering, College of Engineering, North Carolina State University, Raleigh, NC and Visiting Professor, Department of Civil and Environmental Engineering, Adjunct Professor, Division of Environment, Hong Kong University of Science and Technology

MEMBERS

Mr. George A. Allen, Senior Scientist, Northeast States for Coordinated Air Use Management (NESCAUM), Boston, MA

Mr. Ed Avol, Professor, Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA

Dr. Michelle Bell, Professor, School of Forestry and Environmental Studies, Yale University, New Haven, CT

Dr. Joseph D. Brain, Cecil K. and Philip Drinker Professor of Environmental Physiology, Department of Environmental Health, Harvard School of Public Health, Harvard University, Boston, MA

Dr. David Chock, Independent Consultant, Bloomfield Hills, MI

Dr. Ana Diez-Roux, Dean, School of Public Health, Drexel University, Philadelphia, PA

Dr. David A. Grantz, Department of Botany and Plant Sciences, University of California at Riverside, Kearney Agricultural Center, Parlier, CA

Dr. Jack Harkema, Professor, Department of Pathobiology, College of Veterinary Medicine, Michigan State University, East Lansing, MI

Dr. Daniel Jacob, Professor, Atmospheric Sciences, School of Engineering and Applied Sciences, Harvard University, Cambridge, MA

Dr. Steven Kleeberger, Professor and Lab Chief, Laboratory of Respiratory Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC

Dr. Frederick J. Miller, Independent Consultant, Cary, NC

Dr. Howard Neufeld, Professor, Department of Biology, Appalachian State University, Boone, NC

Dr. Armistead (Ted) Russell, Professor, Department of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA

Dr. Helen Suh, Interim Chair, Director of Population Health Doctoral Program, Department of Health Sciences, Northeastern University, Boston, MA

Dr. James Ultman, Professor, Chemical Engineering, Bioengineering Program, Pennsylvania State University, University Park, PA

Dr. Sverre Vedal, Professor, Department of Environmental and Occupational Health Sciences, School of Public Health, University of Washington, Seattle, WA

Dr. Kathleen Weathers, Senior Scientist, Cary Institute of Ecosystem Studies, Millbrook, NY

Dr. Peter Woodbury, Senior Research Associate, Department of Crop and Soil Sciences, College of Agriculture and Life Sciences, Cornell University, Ithaca, NY, U.S.A.

Dr. Ronald Wyzga, Technical Executive, Air Quality Health and Risk, Electric Power Research Institute, Palo Alto, CA

SCIENCE ADVISORY BOARD STAFF

Dr. Holly Stallworth, Designated Federal Officer, U.S. Environmental Protection Agency, Science Advisory Board (1400R), 1200 Pennsylvania Avenue, NW, Washington, DC 20460

CASAC Consensus Responses
Charge Questions
On the Second Draft Policy Assessment
for the Review of the National Ambient Air Quality Standards for Ozone

Introduction (Chapter 1)

1. *Does the Panel find the introductory and background material (sections 1.1 and 1.2) to be appropriately characterized and clearly communicated?*

This chapter is an excellent introduction to the legislative requirements of the Clean Air Act as they apply to ozone. It also provides an excellent history of the standard setting process; Table 1-1 does a good job summarizing the different NAAQS that have been historically set. The historical review of legal challenges to previous decisions and their eventual outcomes is well done. The chapter clearly delineates the scope and approach for the Second Draft PA. This review is especially informative about the ruling by the U.S. Court of Appeals for the District of Columbia concerning the failure of the CASAC to be clear in its recommendations on "whether it was providing scientific or policy recommendations…." which led to retention of the 75 ppb standard instead of a standard between 60 and 70 ppb.

2. *In section 1.3, we describe the general approach for the review. This includes the key aspects of the approach employed in the last review in judging the adequacy of the then-existing standards and in selecting revised standards. Does the Panel find this description of the approach in the previous review adequate and clear? Does the summary of the approach in the current review appropriately describe important considerations in this review?*

Sections 1.3.1 through 1.3.1.2 clearly delineate the Administrator's stated rationale for lowering the NAAQS in 2008 based on the weight of the evidence at that time. The process by which this decision was made is explained in a clear and lucid manner. Figure 1-1 is an excellent flow diagram that establishes the basis for the policy assessment on the primary standard. The description of the five-point hierarchical weight of evidence classifications and how they are used in the Second Draft PA (e.g., greatest weight placed on those effects deemed to be caused, or likely to be caused, by ozone) is clearly explained. The fact that the Second Draft PA focuses on at-risk populations for which the evidence is "adequate" is also made clear (see page 1-21). Stressed throughout is the EPA's concern about the uncertainty and likelihood of ozone effects at lower exposures, which is important for setting a lower end to the recommended range for the NAAQS. How the EPA deals with this type of uncertainty is very well explained throughout this chapter and the rest of the Second Draft PA.

How the EPA deals with controlled human exposure studies and epidemiologic studies is well delineated. There is considerable discussion about the possible existence of threshold effects, with the EPA concluding that there is no evidence for a sharp break point between exposures of 80 and 60 ppb.

There is an extensive discussion of how the welfare risks were evaluated in previous reviews (section 1.3.2.1) and why in the last review the secondary standard was set equal to the primary

standard, even though the CASAC had recommended a different form and level from the primary standard (i.e., a cumulative exposure index, the W126; pg 1-30). The discussion of the weight of the evidence and how it was considered in the Second Draft PA is quite understandable and well written.

In several places (e.g., pages 1-27, lines 7-10; pages 1-36, lines 24-27) the document states that at low W126 exposures, the "magnitude of the response becomes increasingly uncertain." This statement is misleading; the exposure-response functions for seedlings of some tree species definitely show growth losses (decrements) at chronically low exposure levels, whereas other species show decrements only at higher exposures. Crop species show a similar range of responses. For both crop and tree species, there is strong evidence of adverse effects on common, ecologically and economically important species at exposure levels at and below 7 ppm-hr (W126).

Overall, this is a well written introduction which sets the stage for the rest of the Second Draft PA.

Ozone Monitoring and Air Quality (Chapter 2)

1. *To what extent does the Panel agree that the most relevant information on monitoring (section 2.1), emissions and atmospheric chemistry (section 2.2), and common patterns of O_3 concentrations (section 2.3) is presented, and to what extent is the information presented appropriately characterized and clearly communicated?*

Chapter 2 is concise and well written. In sections 2.1 to 2.3, the discussions of monitoring, atmospheric chemistry and ozone patterns are well done. The chapter focuses primarily on the issues involving background ozone, which is an appropriate discussion, but it should further discuss the response of lower level ozone to controls (e.g., as an extension of the section on atmospheric chemistry or as a stand-alone section). As shown in the Second Draft HREA, this is a very important consideration. The Second Draft PA should be clear that controls designed to reduce the peak levels of ozone (e.g., the 4th highest annual MDA8) may not be effective at reducing lower levels of ozone on more typical days and may actually increase ozone levels on days where ozone concentrations are low. The EPA should consider showing the meteorologically-adjusted ozone levels in the figures.

2. *With regard to information on estimating O_3 concentrations associated with non-anthropogenic sources or "background O_3" (section 2.4), to what extent is this information appropriately characterized and clearly communicated?*

The discussion of background ozone is extensive and generally clear though some pieces are missing. First, the method by which the various backgrounds are calculated or taken into account (e.g., models used) should be further described, and the uncertainties discussed. This section should discuss how the background might impact various standards (both health and welfare-based) and how that might differ regionally~~.~~, and that there are events where the background contribution may lead to levels similar to those being considered as potential standards.~~.~~ The discussion of the source apportionment model estimates is too minimal to really understand what is being done and its importance is not evident; consequently, it should be moved to the appendix. The issues of using monitoring to estimate background should be discussed.

There appears to be confusion in the proper use of the zero-out approach. On p. 2-12, lines 23-26, and again on p. 2A-7, lines 231-241, the Second Draft PA indicates that removing NO_x emissions completely and unrealistically could lead to inflated estimates of background ozone in urban areas where NO_x titration of ozone is significant. The authors consider this a paradoxical result of the zero-out approach. It is not apparent what is meant by paradoxical in this case as the behavior is well known, and is part of the non-linearity in ozone chemistry (see above suggesting a further discussion of the increase in lower ozone levels from controls). Furthermore, the authors ran separate Community Multi-scale Air Quality (CMAQ) model runs for base case (designated as total), natural background (NB), North American Background (NAB) and United States Background (USB), and then took the ratios of one of the background runs to the base case runs as the percent contribution of the background to the base case. This is inappropriate because it assumes ozone additivity; however, ozone chemistry is nonlinear. The zero-out approach is not intended for use in assigning relative contributions, and relative contribution is not a meaningful concept unless the perturbation due to some source emissions is small enough that linearity can be approximated. Based on the above discussion, in the case of the zero-out approach, it is best to change the text in the chapter from "percent contribution" (like in Fig. 2-11 on p. 2-18) to "ratio" when the background ozone is compared to the base case ozone.

Adequacy of the Primary Standard (Chapter 3)

1. *To what extent does section 3.1 (Evidence-based Considerations) capture and appropriately characterize the key aspects of the evidence assessed and integrated in the ISA? To what extent is staff's consideration of the health effects evidence, including the adversity of reported respiratory effects and public health implications technically sound and clearly communicated at an appropriate level of detail? In the Panel's view has the information been appropriately interpreted for the purpose of assessing the adequacy of the current standard?*

Section 3.1 captures the important studies discussed in the ISA and integrates them into a logical narrative to summarize the important findings in each of the major effect categories: lung function decrements, pulmonary inflammation, respiratory symptoms, respiratory morbidity, and respiratory mortality. The adversity of the effects and their implications for public health are discussed in a straightforward and clear manner that leads the reader through the body of data.

The CASAC notes that Figure 3.1, which is intended to illustrate mode-of-action, does not present a logical sequence of steps leading to specific endpoints. The discussion of mode-of-action could be more clearly communicated as it has been in the Appendix. Various specific technical comments are included in individual panel member comments. In some cases, individual panelists suggest changes in the wording or thrust of some of the sentences and paragraphs. The discrepancy between epidemiological studies supporting anti-oxidants being partially protective of ozone-induced lung function decrements and controlled human studies showing no such protection is one such example.

2. *With regard to the presentation of the exposure and risk information for the purpose of assessing the adequacy of the current standard, to what extent is the information, including associated limitations and uncertainties, sufficiently characterized, appropriately interpreted and clearly communicated?*

The major exposure and risk information requisite to assess the adequacy of the current standard is well presented for the adjusted air quality data, exposure-based considerations, and risk-based considerations. Overall, the salient points are clearly presented with enough detail to allow the reader to make judgments on how much weight to assign to any limitations or uncertainties. Individual panelist comments are provided that, if addressed, will strengthen these considerations. Major points are supported by reference to studies or data presented either earlier in the Second Draft PA or in the Second Draft ISA or Second Draft HREA, which leads to the conclusion that the points have been appropriately interpreted.

3. *In the Panel's view, does the discussion in section 3.4 provide an appropriate and sufficient rationale to support staff's preliminary conclusion that the current evidence and exposure/risk information call into question the adequacy of the current standard and that it is appropriate to consider revising the standard to achieve additional public health protection?*

Section 3.4 clearly articulates the findings and points that underpin Staff's preliminary conclusion to call into question the adequacy of the current standard and the appropriateness of revising it. The CASAC finds scientific justification that current evidence and exposure/risk information call into question the adequacy of the current standard. Furthermore, there is scientific support for the need to revise the standard to achieve additional public health protection.

Consideration of Potential Alternative Primary Standards (Chapter 4)

1. *In the Panel's view, has the evidence and exposure/risk information, including associated limitations and uncertainties, been appropriately characterized and interpreted for the purpose of considering potential alternative standards?*

This chapter provides a comprehensive, clear, and carefully documented assessment of potential alternative primary standards for ozone. As discussed in some individual comments, it could be improved to minimize repetition within the chapter and with earlier chapters. The discussion could be more focused on the salient points such as in the discussion of studies conducted in locations that meet possible alternative standards (p. 4-12 to 4-13) and the cut-point analysis (p. 4-14 to 4-15). Figures 4.1 – 4.8 could use some cosmetic improvements.

2. *In the Panel's view, does the discussion in section 4.6 provide an appropriate and sufficient rationale, supported by the discussions in sections 4.1 through 4.4, to support staff's preliminary conclusions regarding alternative primary standards (including the indicator, level, averaging time and form) that it is appropriate to consider?*

Section 4.6 is well written with a clear rationale presented for each of the possible alternative standards. Scientific justification for retaining the current indicator (ozone), averaging time (maximum daily 8-hour average), and form (4th highest maximum daily 8-hour average, averaged over three years) is provided, as well as the scientific basis for consideration of alternative levels.

The CASAC concurs with the Second Draft PA that the current indicator, averaging time, and form be retained. The indicator of ozone is appropriate based on its causal or likely causal

associations with multiple adverse health outcomes and its representation of a class of pollutants known as photochemical oxidants.

The current 8-hour averaging time is justified by the combined evidence from epidemiologic and clinical studies referenced in Chapter 4. Results from clinical studies, for example, show a wide range of respiratory effects in healthy adults following 6.6 hours of exposure to ozone, including pulmonary function decrements, increases in respiratory symptoms, lung inflammation, lung permeability, decreased lung host defense, and airway hyperresponsiveness. These findings are supported by evidence from epidemiological studies that show causal associations between short-term exposures of 1, 8 and 24-hours and respiratory effects and "likely to be causal" associations for cardiovascular effects and premature mortality. The 8-hour averaging window also provides protection against the adverse impacts of long-term ozone exposures, which were found to be "likely causal" for respiratory effects and premature mortality.

Regarding the form of the standard, the CASAC concurs that the ozone standard should be based on the fourth highest, daily maximum 8-hour average value (averaged over three years). This provides health protection while allowing for atypical meteorological conditions that can lead to abnormally high ambient ozone concentrations which, in turn, provides programmatic stability.

[Note to the Reader: At this time, the Panel is still deliberating regarding its advice for a scientifically-based upper bound to the range of levels being considered. The Panel will deliberate on the science-based upper level during its upcoming teleconferences, and we expect to revise this response to the charge question. The text that follows is based on the current status of what the Panel has discussed as of its most recent March 25 - 27, 2014 meeting.]

The Second Draft PA concludes that the scientific evidence and available information support consideration of a new primary ozone standard within the 60 ppb to 70 ppb range based on ozone as the indicator, an 8-hour averaging time, and 4th highest daily form of the standard (averaged over three years). The CASAC concurs that 60 ppb is an appropriate and justifiable scientifically based lower bound for a revised primary standard. This is based upon findings of adverse effects, including clinically significant lung function decrements and airway inflammation, after exposures to 60 ppb ozone in healthy adults with moderate exertion (Adams 2006; Schelegle et al. 2009; Brown, 2008; Kim et al., 2011), with limited evidence of adverse effects below 60 ppb. The CASAC further notes that clinical studies do not address sensitive subgroups, such as children with asthma, and that there is a scientific basis to anticipate that the adverse effects for such subgroups are likely to be more significant at 60 ppb than for healthy adults.

[Note to the Reader: The panel has not yet reached a decision regarding scientific advice for an upper end of a recommended revised standard. Here, we summarize the scientific evidence associated with levels of 70 ppb and 65 ppb~~:~~.]

At a level of 70 ppb for the averaging time and form of the current standard, clinical and epidemiological studies show adverse effects to human health. As discussed in the Second Draft HREA and the Second Draft PA, approximately 1% of children are estimated to experience exposures of 70 ppb or more (daily maximum 8-hour exposures) in an ozone season, with 3% to 10% of children (depending on urban area) estimated to experience one or more exposures of concern at or above 60 ppb. Up to 5% of children are expected to experience ozone-induced lung

function decrements greater than or equal to 15% in a "worst case" year with respect to meteorological conditions. Exposures of these magnitudes have been shown to result in significant adverse effects. For example, controlled human exposure studies show respiratory symptoms combined with clinically significant lung function decrements following ozone exposures to 60 ppb to 70 ppb in healthy individuals. These findings suggest that ozone exposures of 70 ppb pose significant concern, especially for children, asthmatics, the elderly and other at risk populations.

Eight-hour ozone exposures (in the form of the current standard) at levels less than 70 ppb have also been shown to be harmful to human health, although to a lower percent of the population and with overall less severity as compared to what would occur at 70 ppb. An alternative standard level of 65 ppb would reduce the frequency of occurrence of lung function decrements of 15% or higher, as compared to a level of 70 ppb, but does not eliminate such occurrences. Further, an alternative standard level of 65 ppb ~~will~~would lead to lower frequency of short-term and long-term premature mortality than the current standard or a level of 70 ppb. The frequency of lung function decrements and premature mortality decreases even further when the alternative standard is lowered to 60 ppb. As noted earlier, based on results for clinical studies of healthy adults, and scientific considerations of differences in responsiveness of asthmatic children compared to healthy adults, there is scientific support that 60 ppb is an appropriate exposure of concern for asthmatic children.

Other suggestions for improving the section were relatively minor, including better labeling, annotation, and discussion of tables and figures, such as Figure 4-13, and adding sub-headings for the consideration of 70, 65, and 60 ppb standard levels.

3. *Does the Panel have any recommendations regarding additional interpretations and conclusions based on the available information that would be appropriate for consideration beyond those discussed in this chapter?*

The chapter provides a cogent presentation of the rationale behind the Second Draft PA's conclusions and the key uncertainties and areas for future research (Section 4.7). The CASAC recommends future research to address key uncertainties related to ozone health effects. This research, which could be conducted in collaboration with other federal and non-governmental organizations, will help to improve our understanding of ozone health effects in support of the next review of the ozone NAAQS.

Adequacy of the Secondary Standard (Chapter 5)

1. *To what extent does the information in sections 5.1 through 5.5 capture and appropriately characterize the key aspects of the evidence for ozone welfare effects assessed and integrated in the ISA? To what extent does the information in section 5.1 (Nature of Effects and Biologically Relevant Exposure Metric) appropriately summarize the nature of ozone welfare effects and to what extent does it appropriately characterize the evidence with regard to biologically relevant exposures?*

This chapter does a nice job summarizing evidence for welfare effects of ozone exposure and for linking ecological to welfare effects. We support EPA's continued emphasis on Class I and other

protected areas. The discussion of flux based metrics is useful and appropriate, concluding that potential benefit may eventually derive from such metrics but that excessive uncertainty remains at this time. We note that when discussing the W126 metric, the phrase "Exposure-Response" (E-R) should be used in place of "Concentration-Response". The discussion regarding the reduction of uncertainty associated with Open Top Chamber (OTC) derived E-R relationships since the last review (page 5-10; page 5-31) is appropriate and clearly presented. It should be further emphasized in the revised chapter that the National Crop Loss Assessment Network (NCLAN) studies covered multiple locations in the U. S. and multiple crops, along with multiple ozone exposure levels, using consistent methods. All of these factors are important because they enhance the reliability of the results.

2. *To what extent is staff's consideration of the welfare effects evidence, including the implications of reported vegetation effects with regard to adversity to public welfare technically sound and clearly communicated at an appropriate level of detail? In the Panel's view has the information been appropriately interpreted for the purpose of assessing the adequacy of the current standard?*

The linkage of ecological effects of ozone to welfare effects is effective in this draft. The effort to monetize welfare impacts and benefits is appropriate although techniques for this monetization are not yet fully developed. Specifically, calculation of consumer and producer surpluses is a useful contribution to quantification of welfare effects. However, this national approach does not adequately account for negative effects on individual farmers and forest owners in high-ozone areas. Tabulating and discussing the number of counties in which yield loss is predicted to exceed a threshold of 5% for individual sensitive crops for alternative candidate standards should be added to Table 6-4 to help address this issue.

The interaction of agriculture and forestry as modeled by Forest and Agricultural Sector Optimization Model with Greenhouse Gases (FASOM-GHG) is mentioned in both the Second Draft WREA and the Second Draft PA but not adequately explained.

The cottonwood data (Figures 5-1 and 5-3) receive too much emphasis. These results are from a gradient study that did not control for ozone and climatic conditions and show extreme sensitivity to ozone compared to other studies, as already noted in the text (page 5-14, line 5). Although they are important results, they are not as strong as those from other experiments that developed E-R functions based on controlled ozone exposure. ~~; thus, they require further confirmation, especially because they show extreme sensitivity to ozone compared to other studies~~. .

Commented [SH8]: Edit from Kathleen Weathers

Commented [SH9]: Edit from Kathleen Weathers

As discussed above, E-R functions for individual tree seedling species, supported by results from other methods such as FACE and naturally-occurring gradients, demonstrate that some species are very sensitive to ozone and show decreased growth at very low chronic exposure levels, while other species show little response even at much higher levels. A similar result is found for crop species. Thus there is strong evidence of decreased growth and yield of sensitive tree and crop species at very low ozone levels.

The method used with FASOM for forest growth is based on individual species' E-R functions, but this is strictly accurate only for forest stands comprised of single species. For mixed-species

forest stands, competition among species with different sensitivity will reduce overall stand growth losses, but also exacerbate effects on sensitive species. Some panel members were concerned about the difficulty of interpreting "median" response for both ozone-sensitive and relatively insensitive species. It should not be assumed that species of unknown sensitivity are tolerant to ozone. It is more appropriate to assume that the sensitivity of species without E-R functions might be similar to the range of sensitivity for those species with E-R functions.

The spatial extent and degree of impact on sensitive species expected at current ozone exposures, the current standard, and at alternate standards should be better quantified for different regions of the country, such as counties. For example, rather than focusing solely on the median relative biomass loss (RBL), the number of counties containing sensitive tree species that are expected to have growth loss of greater than 2~~%, etc.~~% should be quantified.

In addition, the Second Draft PA should clarify whether the denominator basal area in the calculation includes only the 12 species with available E-R functions, or rather includes all species. If it is all species, then there is a bias introduced by implicitly assuming that species without E-R functions are insensitive to ozone. If it only includes the species with E-R functions, then the interpretation will vary depending on what fraction of the basal area is represented by species without known E-R functions. In either case this requires explicit consideration.

3. *With regard to the presentation of the exposure and risk information for the purpose of assessing the adequacy of the current standard, to what extent is the information, including associated limitations and uncertainties, sufficiently characterized, appropriately interpreted and clearly communicated?*

This chapter makes a strong case that the current secondary standard fails to protect vegetation and ecosystem services from adverse effects. The form of the current standard is inadequate to provide such protection. From correlation analysis based on the Higher Order Direct Decoupled Method (HDDM), the EPA suggests that a W126 level of 15 ppm-hr may in many cases be similar to the current standard. As injury is clearly observed below 15 ppm-hr, the chapter demonstrates that the current standard is inadequate to protect against welfare effects due to ozone. Despite the paucity of data in certain areas, the Second Draft PA makes appropriate statements about causality and risk with which to evaluate the adequacy of the current welfare standard.

The uncertainties are well described and appropriately interpreted. As mentioned previously, there is some confusion that should be clarified regarding increasing uncertainty of responses at lower ozone levels versus smaller magnitudes of response. There is adequate evidence for effects in sensitive species at levels of W126 well below 15 ppm-hr. In Figure 5-5 some information is needed about the referenced sites. Uncertainties do not weaken the case for a more stringent standard. The paucity of data on ozone sensitivity of most U.S. plant species should not prevent the Second Draft PA from discussing "anticipated" effects on a number of unidentified sensitive species. As in the Second Draft WREA, the consideration of uncertainty is repetitive to the point of detracting from the considerable strength of the analyses and the evidence on which they are based.

The discussion of economic losses due to bark beetles and fire perpetuates the confounding of spatial association with causation in the Second Draft PA. As stated, this does not contribute much

to the assessment of risk nor evaluation of the adequacy of the current or alternative standards.

4. *In the Panel's view, does the discussion in section 5.7 provide an appropriate and sufficient rationale to support staff's preliminary conclusion that the current evidence and exposure/risk information call into question the adequacy of the current standard and that it is appropriate to consider revising the standard to achieve additional public welfare protection?*

The CASAC concurs with the justification in this section that the form of the standard should be changed from the current 8-hr form to the cumulative W126 index and finds that the discussion provides an appropriate and sufficient rationale.

This section clearly demonstrates that ozone induced injury may occur in areas that meet the current standard, based on the correlative similarity between the current standard and a level of the W126 index of 15 ppm-hr. Most of the analyses found effects below 15 ppm-hr (many at 10 or even 7 ppm-hr). Based on review of relevant science, the CASAC concludes that the range of values that should be considered for the W126 standard should not exceed 13 ppm-hr. Most definitely, the CASAC does not support a level as high as 15 ppm-hr or 17 ppm-hr. These levels should not be included in the revised PA as options for an alternate secondary standard.

The ~~CASAC does not recommend the use of a three-year averaging period. We favor a single-year averaging period, which will provide more protection for annual crops and for the anticipated cumulative effects on perennial species. The scientific analyses considered in this review, and the evidence upon which they are based, are from single-year results. If a 3-year averaging period is established, then the upper limit will need to be reduced to protect against one-year ozone peaks. We consider this further in the response to charge questions for Chapter~~ 6.

Commented [R,AG10]: As noted above, I strongly disagree with this statement. There are strong reasons for using a three year average. I would much prefer using a three year average that may be set somewhat lower to protect against high summers.

In summary, the CASAC concurs with the Second Draft PA that the current secondary standard is inadequate with respect to form, level and averaging time. Only the current indicator (ozone) should be retained. The potential revised standard is considered further in the response to the charge questions for Chapter 6 below.

Consideration of Potential Alternative Secondary Standards (Chapter 6)

1. *In the Panel's view, has the evidence and exposure/risk information, including associated limitations and uncertainties, been appropriately characterized and interpreted for the purpose of considering levels of protection and potential alternative standards?*

2. *In the Panel's view, does the discussion in section 6.5 provide an appropriate and sufficient rationale, supported by the discussions in sections 6.1 through 6.4, to support staff's preliminary conclusions regarding alternative secondary standards (including the indicator, level, averaging time and form) that it is appropriate to consider?*

3. *Does the Panel have any recommendations regarding additional interpretations and*

conclusions based on the available information that would be appropriate for consideration beyond those discussed in this chapter?

Charge Questions 1, 2, and 3 are all closely related and are being addressed together in this summary response.

The Second Draft PA makes a very strong case, consistent with previous CASAC judgment, for changing the secondary metric to the W126 averaged over the highest three-month interval. Accumulation over the 08:00 a.m. – 08:00 p. m. daytime 12-hour period is an acceptable means of generalizing across latitudes and seasons.

The suggestion in Section 6.2 to use a 3-year averaging period is not supported by the available data. We have not supported it in the past nor do we support it here. The primary justification for a 3-year averaging period is to improve the program stability of the classification of regions as being in or out of compliance. The proposed form includes a 3-month period, so it is not nearly as sensitive to extreme events as an hourly or 8-hour averaging period. The case has not been made that welfare benefits from the stability of a 3-year average are greater than those from using the biologically-relevant 1-year value. If a 3-year averaging period is implemented, it should be at a lower level than a single-year standard to protect against single unusually damaging years that will be obscured in the average. ~~Previous EPA analyses (Mintz, 2011) suggested that a reduction of approximately 2 ppm-hr to 4 ppm-hr would provide the requisite margin, but updated calculations are needed.~~

The key issue in the Second Draft PA with respect to welfare effects is the level of the standard. If protection of the most sensitive members of the community is extended to components of ecosystems that impact public welfare, then several potential levels of the standard should be considered. A 2% biomass loss is an appropriate scientifically-based value to consider as a benchmark of adverse impact for long-lived perennial species such as trees, because effects are cumulative over multiple years. For example, a 2% annual RBL fully explains the observed ~~the~~ biomass loss of 21% over 7 years in the study of pure stands of aspen at the Rhinelander, WI FACE site (Wittig et al. 2009). The CASAC considers it significant that a similar value of 1% to 2% for tree seedling biomass loss was recommended previously by a consensus meeting of experts on ecological effects of ozone (Heck and Cowling 1997). In our scientific judgment, it is appropriate to identify a range of levels of alternative W126-based standards that aim for not greater than 2% RBL for the median tree species. With a sample size of only 12 species, and assuming that the available sample is representative, the confidence interval for the population median is inclusive of a range of the 5th to 7th ranked among species for which E-R functions are available. Thus, we take into a consideration a possible range of W126-based levels that account for statistical uncertainty in the median estimate.

The CASAC considered carefully the data presented in Figure 5-2 and Table 6-1 of the Second Draft PA. The Monte Carlo analysis (red dots, Fig. 5-2) should not be used in evaluating the effect of ozone on RBL of tree seedlings. This analysis overemphasizes the species for which relatively few E-R functions are available, is biased toward the few less sensitive response functions available for some individual species, makes unsupported assumptions regarding the representativeness of available response functions, and confounds intra- and inter-species variability in unquantifiable ways. We favor using a measure of central tendency of the data, specifically the median across species (the green line in Fig. 6-4). This analysis provides the

Commented [SH11]: The Mintz analysis has been superseded by new technical information submitted by EPA on 5-8-14. See "Agency Follow up" posting listed as "Additional Technical Information on W126 Averaging Time."

Commented [R,AG12]: As from above. I will make the point that the three year averaging period for health was to be protective, not because humans happen to react to the third year of the 4th highest level. It is because of the points made in lines 14-22 on page 4-7.

Commented [R,AG13]: We should have discussed this at the meeting. Too much additional analysis and discussion is coming up now that was not brought up then.

median of best available estimates within each species, and the median across species with all species treated equally. Table 6-1 presents the RBL results for individual species for different levels of W126. This table demonstrates that a range of 7 ppm-hr to 13 ppm-hr will protect against RBL of 2% for at least 5 of the 12 species, consistent with statistical uncertainty regarding the median specie given the observed sample of only 12 species. With compounding over the harvest cycle or life span of these species, this will result in considerably greater cumulative RBL as discussed above. For the more sensitive tree seedlings, a value closer to the lower end of the range of 7 ppm-hrs would be more appropriate. The range of 7 ppm-hr to 13 ppm-hr is protective of median crop loss of 5% or less, with lower values within the range offering more protection than higher values.

Visible foliar injury is even more sensitive than RBL of 2%, with W126 values below 10 ppm-hr required to reduce the number of sites showing visible foliar symptoms. Crop loss appears to be less sensitive than these other indicators, largely because of the CASAC judgment that a 5% yield loss represents an adverse impact, and in part due to more opportunities to alter management of annual crops. A level of W126 that is protective of 2% RBL for the median tree seedling will thus likely be protective of crop yields at 5% yield loss. This discussion, like that in the Second Draft WREA and Second Draft PA, focuses on annual values for W126, and lower values should be considered for a multi-year average.

The spatial region for which a standard is intended to be protective should be specified. Some analyses are done for large regions of the country, some for counties etc. A county scale is appropriate for assessing crop yield loss. Calculating producer and consumer surpluses at national or large region scales does not provide adequate protection. Farmers growing sensitive crops in high ozone locations can be considered a "sensitive population" for welfare impacts, and crop yields under these conditions should be protected.

Care should be taken not to overstate uncertainties. For example, there is quite a lot of certainty in estimates of biomass loss for forest tree seedling species and crop species for which exposure-response (E-R) functions have been developed. Because several dominant crop species have E-R functions, there is a quite a lot of certainty about impacts of ozone on crop yield across most annual cropland in the USA.

There is considerable uncertainty in extrapolating from the 12 forest tree species to all forest tree species in the US. It should be anticipated that there are species of vegetation that are highly sensitive to ozone that do not have E-R functions, and others that are insensitive. It is scientifically justifiable to extrapolate from the known E-R curves, assuming that they are representative of the un-sampled population.

Based on scientific considerations, the CASAC recommends that a revised standard be set using the W126 cumulative index, with accumulation over the 08:00 a.m. – 08:00 p.m. 12-hour period of the highest 3-months over a single year. The CASAC further recommends that a scientifically justifiable level for the revised Secondary Standard be considered within the range of 7 ppm-hr to 13 ppm-hr. If three-year averaging is chosen, the range would have to be calculated to provide equivalent protection given the possibility of one-year peaks, as previously noted.

References:

Adams, W.C. 2006. Comparison of chamber 6.6 hour exposures to 0.04 – 0.08 ppm ozone via square wave triangular profiles on pulmonary responses. *Inhalation Toxicology*. 18:127-136. - http://dx.doi.org/10.1080/08958370500306107

Brown, J.S.; T.F. Bateson, and W.F. McDonnell, W.F. (2008). Effects of exposure to 0.06 ppm ozone on FEV1 in humans: A secondary analysis of existing data. *Environ Health Perspect* 116: 1023-1026. http://dx.doi.org/10.1289/ehp.11396

Fiore, A.M., and 47 others, Multi-model Estimates of Intercontinental Source-Receptor Relationships for Ozone Pollution, *J. Geophys. Res*., 114, D04301, 2009.

Heck, W. W. and E. B. Cowling. 1997. The need for a long term cumulative secondary ozone standard - an ecological perspective. *Environmental Manager*. (1~~)::~~): 23-33.

Kim, C.S., N.E. Alexis, A.G. Rappold, H. Kehrl, M.J. Hazucha, J.C. Lay, M.T. Schmitt, M. Case, R. B. Devlin, D.B. Peden, D. Diaz-Sanchez. 2011. Lung function and inflammatory responses in healthy young adults exposed to 0.06 ppm ozone for 6.6 hours. *Am J Respir Crit Care Med.* 183: 1215-1221. http://dx.doi.org/10.1164/rccm.201011-1813OC

~~Mintz, David. 2011. Relationship between W126 Annual Values and 3-Year Average. Memo to Docket, EPA-HQ-OAR-2005-13726.~~

National Research Council. 2009. Global Sources of Local Pollution: An Assessment of Long-Range Transport of Key Air Pollutants to and from the United States. Washington, DC: The National Academies Press.

Schelegle, ~~ES;~~E.S.; Morales, ~~CA;~~C.A.; Walby, ~~WF;~~W.F.; Marion, S; Allen, ~~RP~~R.P. 2009. 6.6-hour inhalation of ozone concentrations from 60 to 87 parts per billion in healthy humans. *Am J Respir Crit Care Med.* 180: 265-272. http://dx.doi.org/10.1164/rccm.200809-1484OC

Task Force on Hemispheric Transport of air Pollution. 2010. Hemispheric Transport of Air Pollution, Part D: Answers to Policy-Relevant Science Questions, Air Pollution Studies No. 20. Prepared by the Task Force on Hemispheric Transport of Air Pollution acting within the framework of the Convention on Long-range Transboundary Air Pollution. United Nations, New York and Geneva. http://www.htap.org/publications/2010_report/2010_Final_Report/HTAP%202010%20Part%20D%20110407.pdf

Wittig, V. E., E. A. Ainsworth, S. L. Naidu, D. F. Karnosky, and S. P. Long. 2009. Quantifying the impact of current and future tropospheric ozone on tree biomass, growth, physiology and biochemistry: a quantitative meta-analysis. *Global Change Biology*. 15:396-424.

Zhang, L., D.J. Jacob, N.V. Smith-Downey, D.A. Wood, D. Blewitt, C.C. Carouge, A. van Donkelaar, D.B.A. Jones, L.T. Murray, and Y. Wang, Improved estimate of the policy-relevant background ozone in the United States using the GEOS-Chem global model with 1/2ox2/3o

horizontal resolution over North America, *Atmos. Environ.*, 45, 6769-6776, 2011.

Appendix A

Formatted: Font: 12 pt

Individual Comments on the Policy Assessment
For the Review of the Ozone NAAQS (Second Draft)
CASAC Ozone Review Panel

Contents

George Allen A-2
Joe Brain A-10
David Chock A-12
Ana Diez Roux A-14
David Grantz A-18
Daniel J. Jacob A-21
Steven Kleeberger A-22
Fred Miller A-23
Howard S. Neufeld A-26
Armistead (Ted) Russell A-29
Helen Suh A-31
James Ultman A-34
Sverre Vedal A-36
Peter Woodbury A-38
Ronald E. Wyzga A-47

George Allen

~~Revised Comments from George Allen on EPA's February 2014 Second Draft Policy Assessment Document for Ozone~~

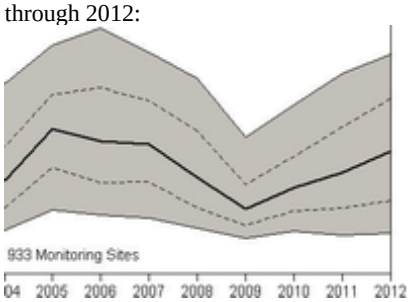
Charge Questions on O3 Monitoring and Air Quality (Chapter 2).

Ch. 2, Q #1. To what extent does the Panel agree that the most relevant information on monitoring (section 2.1), emissions and atmospheric chemistry (section 2.2), and common patterns of O3 concentrations (section 2.3) is presented, and to what extent is the information presented appropriately characterized and clearly communicated?

Formatted: No underline

Formatted: Space After: 0 pt

Section 2.1, Monitoring. National trends of annual 4th highest max 8-hour values are plotted in Figure 2-2, page 2-4. Section 2.1 appropriately notes the distinct drop between 2002 and 2004, consistent with the drop in summer NOx emissions due to the "NOx SIP call", and notes the decreasing trend between 2000 and 2009. While that time period does appear to have a distinct downward trend, it is also informative to look at this same plot constrained to the period from 2004 through 2012:



This 9-year time period starting after the NOx SIP call drop in O3 shows no indication of a trend, and leads to a very different conclusion. While both this time window and EPA's interpretation of O3 trends over 2000 through 2009 are valid, the latter is totally driven by a one-time intervention; since then there does not appear to be any progress in reducing O3 concentrations. Year to year variations in summer meteorological conditions can play a large role in "raw" trends of less than 15 to 20 years duration. In the past EPA has presented trend estimates that include adjustments for meteorology; see: http://www.epa.gov/airtrends/weather.html and references therein. It may be helpful to present such "adjusted trend" data to better assess progress since the NOx SIP call; these adjustments tend to pull "high" years down and "low" years up in concentration. 2013 O3-season data is now in AQS; even though those data will not be "certified" for another 3 months, it would be informative to add that data for 2013 to this trend plot.

Section 2.2, Emissions and Atmospheric Chemistry. This section is a clearly written and concise summary of this topic. It makes the interesting point regarding the co-benefit from NOx emission reductions for O3, NO2, and PM2.5.

Section 2.3, Air Quality Concentrations. This section is a clearly written and concise summary of this topic. Footnote 7, page 2-10, makes the sometimes overlooked point that since O3 is measured seasonally in most areas, the 4th highest day's value is similar to the 98th percentile annual metric

form used for some other criteria pollutants.

Section 2.4, Background O3.

Ch. 2, Q #2. With regard to information on estimating O3 concentrations associated with nonanthropogenic sources or "background O3" (section 2.4), to what extent is this information appropriately characterized and clearly communicated?

This section is the core of this chapter. Estimates of background O3 now play a minimal role in the REA document given the new approach to estimating risk under various emission reduction scenarios using the HDDM rollback method and total O3 concentrations.

Background O3 is still a factor in the Policy Assessment however, since a 2002 court decision allows EPA to consider background levels when evaluating risk for alternative (lower) standards (section 1.3.1, page 1-26, lines 17-19). But case law also states that "that attainability and technical feasibility are not relevant considerations in the setting of a NAAQS" (section 1.2.1, page 1-4 lines 19-21, API v. Costle, 1981). It is unclear how EPA might navigate between these two legal guidelines in terms of how background O3 would be used in a policy and standard-setting context.

EPA performed new 2007 base "year" (7-months) zero-out and CAMx source apportionment modeling that is presented in this section. Section 2.4.2 (page 2-16, lines 6-14) introduces the concept of "apportionment-based US background" O3 (AB-USB?) as the most relevant metric for estimation of a "fractional background" metric. Overall this chapter is difficult to follow. It seems the new modeling's utility is to confirm earlier outcomes using these improved modeling approaches. It would be helpful if the chapter could better focus on these specific issues.

Section 2.4.3: The discussion on page 2-17, lines 1-12, is helpful in understanding the fractional contribution of background O3 on days with elevated O3.

Figures 2-13 and 2-14 on page 2-19 could benefit from the addition of "N" (# of site-days) to each bin. Figure 2-13 is difficult to interpret without this information.

Section 2.4.4 presents background O3 in the context of the 12 urban case study areas. Table 2-2 (page 2-21) seems to be the most relevant presentation, but is only for site-days > 60 ppb; it would be helpful to also include this information for site-days >65 and > 70.

Section 2.4.5 presents background O3 in the context of a W126 secondary standard form for four locations (2 are large urban areas). Page 2-22 lines 10-12 refer to figure 2-7 (page 2-7) as showing high observed 2010-2012 W126 values for these four sites. These sites can not readily be identified on this map, so these values also need to be provided in a text form.

Other Comments.

Chapter 1, Section 1.2.2. (History of O3 NAAQS reviews)

This section is a useful summary of recent actions and court rulings regarding the O3 NAAQS review process. There are two topics regarding the 2008 NAAQS revision and subsequent reconsideration process that I would like to expand on that may be relevant to the current review.

First is the "Reconsideration" of the 2008 O3 primary NAAQS, discussed on page 1-9. It is not widely known that in the summer of 2011, EPA sent a final rule to OMB/ORIA with 65 ppb as the standard. This and the role that ORIA played in the process was detailed in a New York Times article on Nov. 17, 2011:

www.nytimes.com/2011/11/17/science/earth/policy-and-politics-collide-as-obama-enters-campaign-mode.html

It is worth noting that EPA proposed this standard (in the middle of the CASAC range of 60 to 70 ppb) based on the scientific literature used for the 2008 rule – e.g., new studies since 2006 could not be considered in this reconsidered rulemaking. EPA subsequently posted the full preamble of what was to be the final Summer 2011 rule, as well as the draft impact analysis at~~:~~:http://www.epa.gov/glo/actions.html

http://www.epa.gov/glo/actions.html

While this background material is not directly relevant to the current review process, it provides some useful context for the decisions the agency made in 2011 during the reconsideration process, and may be of interest to the current O3 panel.

Second, this section is a useful summary of the court's 7/23/2013 decision (reissued 12/11/13) regarding the 2008 O3 NAAQS rule. While this ruling upheld the 2008 primary NAAQS of 75 ppb, it did so based on an unusual interpretation of CASAC's intent in saying a range of 60 to 70 ppb was appropriate to consider. This may have implications on how future CASAC advisory reviews are written. Although this issue is not part of the PA review, it would be helpful if EPA staff could provide some guidance to CASAC on how to avoid future perceived ambiguity in its recommendations.

On page 1-10 lines 10-14, the court's rationale for the 0.075 ppm not being inconsistent with CASAC advice is correctly explained. However it is informative to look more closely at the detailed wording of this argument, especially since it may require the CASAC to be more explicit in future letters regarding this issue.

The court decision is summarized at:

http://www.lawandenvironment.com/2013/07/mississippi-v-epa-support-of-the-clean-air-science-advisory-committee-is-not-necessary-to-affirm-epas-naaqs/

The decision itself:

http://www.lawandenvironment.com/wp-content/uploads/2013/07/08-1200-1447980.pdf

The essence of the decision, on page 41:

"...in order for EPA to explain adequately its reasons for disagreeing with CASAC, CASAC itself must be precise about the basis for its recommendations. Because in this case CASAC failed to specify whether the 0.070 ppm level it recommended as a maximum rested on a scientific conclusion about the existence of adverse health effects at that level, EPA's invocation of scientific uncertainty and more general public health policy considerations satisfies its obligations under the statute."

Additional excerpts from the decision relevant to CASAC follow.

Pages 38-39:

EPA did not make such a specific scientific determination about the 0.070 ppm level that served as the ceiling of CASAC's recommendation; instead, EPA referred generally to declining certainty below 0.075 ppm. Had CASAC reached a scientific conclusion that adverse health effects were likely to occur at the 0.070 ppm level, EPA's failure to justify its uncertainty regarding the existence of adverse health effects at this level would be unacceptable. Indeed, it is a familiar principle that agencies may not "merely recite the terms 'substantial uncertainty' as a justification for [their] actions"; instead, they "must explain the evidence which is available, and must offer a rational connection between the facts found and the choice made." State Farm, 463 U.S. at 52 (internal quotation marks omitted). In other words, EPA must explain why the evidence on which CASAC relied cannot support the degree of confidence CASAC placed in it. This is especially true given the added layer of stringency imposed by EPA's obligations under section 307(d)(6).

But we are unable to determine whether CASAC reached any such scientific conclusion. Although CASAC stated that "overwhelming scientific evidence" supported its recommendation that the standard be set no higher than 0.070 ppm, Mar. 2007 CASAC Letter, at 2, it never explained whether this proposal was based on its scientific judgment that adverse health effects would occur at that level or instead based on its more qualitative judgment that the range it proposed would be appropriately protective of human health with an adequate margin of safety. Indeed, although CASAC concluded that "there is no longer significant scientific uncertainty regarding [its] conclusion that the current 8-hr primary NAAQS must be lowered," given the "large body of data clearly demonstrat[ing] adverse human health effects at the current level," CASAC recognized that "[s]cientific uncertainty does exist with regard to the lower level of ozone exposure that would be fully-protective of human health." Oct. 2006 CASAC Letter, at 5.

To be sure, EPA's statutory obligation to respond to CASAC does not evaporate whenever CASAC exercises judgment amidst scientific uncertainty. Quite to the contrary, had CASAC acknowledged uncertainty in the scientific evidence but explained that, based on its expert scientific judgment, it nonetheless believed adverse health effects were likely to occur at the 0.070 ppm level, then section 307(d)(6) would have required EPA to explain why it disagreed with this scientific conclusion. Put differently, to the extent that CASAC has exercised scientific judgment, EPA must respond in kind. But because CASAC never made clear the precise basis for its recommendation, all we know for certain is this: both CASAC and EPA believed the existence of adverse health effects to be certain at the 0.08 ppm level and reached differing conclusions about what level below 0.08 ppm was requisite to protect the public health with an adequate margin of safety. [end quote]
~~[end quote]~~

Formatted: Space After: 0 pt

Page 41:

Formatted: Space After: 0 pt

Absent a definitive scientific conclusion from CASAC that adverse health effects would occur at the 0.070 ppm level, we must assume that it too took these same considerations into account and simply exercised its judgment to recommend a standard set at a lower level. Although both CASAC and EPA must exercise public health policy judgment when confronted with scientific evidence that does not direct it to a specific outcome, it is to EPA's judgment that we must defer. And (as noted earlier):

But in order for EPA to explain adequately its reasons for disagreeing with CASAC, CASAC itself must be precise about the basis for its recommendations. Because in this case CASAC failed to specify whether the 0.070 ppm level it recommended as a maximum rested on a scientific conclusion about the existence of adverse health effects at that level, EPA's invocation of scientific uncertainty and more general public health policy considerations satisfies its obligations under the statute.

Consideration of Potential Alternative Primary Standards (Chapter 4). These are general comments regarding this chapter.

Section 4 of the draft Policy Assessment is a concise summary of the science presented in the ISA and the health REA with regard to evaluating the health impacts of the current and alternative ozone NAAQS. I agree that all aspects of the form of the current ozone NAAQS are appropriate to retain. It is worth noting that in much of the country, ozone is monitored for 6 or 7 months of the year; this makes the 4th-highest 8-h daily mean form similar to the 98th percentile form used in other NAAQS with a daily form.

When choosing ozone concentrations to consider for alternative (lower) values of a standard, there is a continuum of decreased risk along with increased uncertainty of the quantitative assessment of that risk as ozone exposures decrease. Thus there are no "bright risk lines" at any one ozone concentration when considering values for a revised standard. However, the risk and exposure assessment process creates bright lines at intervals of 5 ppb, ranging from 80 to 60 ppb, focusing on health risks at standards of 70, 65, and 60 ppb in this review. This approach is appropriate since there is no meaningful value in attempting to characterize differences in risk between standards of 70 and 69 ppb for example; the supporting science simply isn't that precise, and we are working from a starting point of 75 ppb (the current ozone NAAQS).

Section 4.6 of the draft Policy Assessment summarizes the reduction in risk for various health endpoints associated with alternative standards of 70, 65, and 60 ppb. 70 ppb is the upper end of the range recommended in this assessment. I agree with the many panel members who expressed concern during the March 27, 2014 CASAC meeting that 70 ppb was not sufficiently protective of vulnerable populations. In that context, EPA's discussion of risks and the science supporting those risks for 65 ppb (starting on page 4-52) is informative to this process. Lines 3-13 on that page provide a framework for this value, noting 65 ppb is "somewhat below" the 70 ppb level where health effects "judged adverse" by the ATS were reported, and higher than the level where lung function decrements "that could be adverse in individuals with lung disease." 65 ppb is also above the 60 ppb level where "pulmonary inflammation has been reported."

Risk is not eliminated at 65 ppb, but this is a concentration where the science behind the reported health effects of concern can be considered to be sufficiently robust for the NAAQS process. Thus I recommend that EPA consider setting a revised ozone standard that is no higher than 65 ppb. This is consistent with the CASAC letter of March 30, 2011 regarding the ozone "reconsideration" review process, which says that 70 ppb "would provide little margin of safety". This review was based on the science available to the panel in the 2008 NAAQS revision, and thus does not reflect new research since 2006. It is also consistent with the "first version" of EPA's final rule for the reconsideration process, which went to OMB in 2011 with a standard of 65 ppb (see my comments on Chapter 1 of this document).

Finally, although not directly relevant to this process, it is worth noting that a new Canadian Ambient Air Quality Standard (CAAQS) for ozone was approved in 2012; this replaces the existing Canada Wide Standard of 65 ppb. The ozone CAAQS of 63 ppb (same form as the current ozone NAAQS) goes into effect in 2015, followed by a standard of 62 ppb in 2020. More information on the ozone CAAQS can be found at:

http://gazette.gc.ca/rp-pr/p1/2013/2013-05-25/html/notice-avis-eng.html#d106

Ed Avol

CHAPTER 1 INTRODUCTION

1. P1-2, line 16 and Footnote 5: not clear that this footnote or caveat are needed; there could be other reasons, including substantive technical as well as basic grammatical or textual, reasons that might lead to a request for a 2nd draft document.
2. P1-14, line 4: With all due respect, there was considerable controversy over the Administrator's recommendation of a 0.075 ppm ozone standard. One cannot "know" what the Administrator "relied" on to make his determination. It would be more appropriate (and accurate) to say that "…the Administrator *was provided* the available scientific evidence…*He subsequently* revised the level of the 8-hour primary O3 standard…"
3. P1-14, line 21: We cannot know what "The Administrator *believed*…" This should be changed to something like "In the opinion of the Administrator…"
4. P1-19, Figure 1-1: A more accurate presentation of the review of the primary standard would suggest that the **Averaging Time** box in the figure be re-worded to allow for the possibility of either a longer *or shorter* -term averaging time.
5. P1-21, lines 8-12: This is a pretty convoluted and wordy sentence, almost obscuring the message. The point here is that based on available research, there is no evidence of a biological threshold for ozone health effects (…and that's a lot shorter and more understandable sentence!).
6. P1-27, lines 21 forward: Why is so much time and effort spent in this document discussing what a former Administrator did or did not do with regard to the last review? While it's appropriate to refer to this, it seems like there is an inordinate amount of text devoted to revisiting (or arguably, attempting to re-formulate) the historical record with regard to what a former Administrator did or did not rely on to make his prior determination. The issue before the current CASAC is NOT what the prior review did or did not decide; the issue before the current CASAC is the strength of the cumulative current evidence regarding ozone exposure and protection of the public's health, and the clarity and appropriateness of proposed staff recommendations to the current Administrator.
7. P1-27, line29: We do not and cannot know to what extent the Administrator *carefully considered the public comments*…or to what *degree he further recognized that several additional lines of evidence had progressed sufficiently since the 1997 review to provide a more complete and coherent picture*… (p1-28, lines 1-3); this phrasing should be removed. The comments and information were provided, the Administrator had access to them, and he made a determination. (This PA is not the place to justify or re-interpret previous Administrator's actions, but rather to present current data and judgments).
8. (typographical error; p1-35, line 35, "evidence")
9. P1-40, lines 25 to 34: this is a useful paragraph describing the organization of the overall document, but it is completely out of place at the end of Chapter 1. This should be moved to the first few pages of Chapter 1, prior to the Background section, and an overview sentence could be added describing the presentation in Chapter 1.

CHAPTER 2 OZONE MONITORING AND AIR QUALITY

10. P2-8, line 13: This is an important point but is awkwardly phrased. Recommend re-wording to say, "…contributes to subsequent O_3 formation further downwind."
11. P2-9, line 7: replace "high" with "elevated", so that sentence reads "…result in a higher frequency and duration of days with elevated O_3 conditions."
12. P2-9, line16: replace "high" with "higher" and "low" with "lower"
13. P2-10, lines 26 to 27: The lower rates of chemical scavenging in downwind rural areas can often result in characteristically broader, more gradual hourly ambient O_3 concentrations, rather than the more"peaky" O_3 spikes often seen in urban areas with substantive rush-hour-type combustion NO contributions. This observation helps to clarify why locations such as Lake Arrowhead downwind of Los Angeles can exhibit elevated eight-hour levels compared to surrounding reporting locations.
14. P2-11 forward, Section 2.4Background O_3: Although I appreciate the figures, depth, and discussion regarding background ozone levels, this presentation seems to conflict with the prior determination that background levels are less important, since total ozone concentrations are being used in the current document to assess risk. This section might be valuable as a supplement or Appendix, but if total ozone is the perspective and approach being used, this whole section seems overly detailed, out of scope, and a diversion from keeping the presentation focused and crisp. If staff were to consider every possible question that might be asked of them and seek to develop and provide as complete an answer to every other possible question, the resulting document would be several times larger than it already is and considerably less accessible to the Reader. In my opinion, this document should focus on the approach, implications, and process put into place by the related ISA and HREA, and not delve off into addressing every possible corollary or potential derivative consideration.

CHAPTER 4 CONSIDERATION OF ALTERNATIVE PRIMARY STANDARDS

(No specific comments on specific sections)

This chapter provides a useful and carefully documented assessment of alternative primary standards for ozone. Although I personally found it wordy and a bit indirect, it does provide a comprehensive evaluation of the anticipated health implications of alternative ozone standards in the 60 to 70ppb range. The figures and tabular summaries regarding assorted health indices should be of particular utility to the Administrator in her assessment.

Joe Brain

O3 Monitoring and Air Quality (Chapter 2): This chapter provides a description of the current O3 monitoring network and recent concentrations, information on emissions and atmospheric chemistry, common patterns and variability in O3 concentrations, as well as, discussion of current information on estimating O3 concentrations associated with non-anthropogenic sources.

1. To what extent does the Panel agree that the most relevant information on monitoring (section 2.1), emissions and atmospheric chemistry (section 2.2), and common patterns of O3 concentrations (section 2.3) is presented, and to what extent is the information presented appropriately characterized and clearly communicated?

I believe that to a great extent, the most relevant information on monitoring, emissions, atmospheric chemistry, and resulting patterns of ozone concentration are presented in Chapter 2. A great deal of information is presented concisely. Especially the figures are helpful in understanding the geographic distribution of the network and the ozone concentrations that have been measured.

A key question is not answered. How does the information collected – the geographic distribution of monitoring stations – correlate with the distribution of humans throughout the United States. It's clear that information is more densely collected in the northeast and in coastal California. We know that that makes sense. But overall, if one simply designed a system of monitoring sites that was solely based on accurately measuring exposure of as many inhabitants as possible, would it be the same or would it be different?

For example, one can imagine that the monitoring sites were selected decades ago. To what extent does the current network operating today represent recent changes in population, such as dramatic declines in Detroit and growth in Las Vegas. There should be a paragraph or so defending the network or at least explaining the extent to which it resembles an ideal, unbiased network. This same discussion should be made not only in relation to all people, but in relation to more responsive populations. What's the distribution of children? What's the distribution of people with preexisting disease and/or the elderly? To what extent do these distributions resemble the current ozone monitoring network?

Ozone, as we all know, is not a primary emission from cars, trucks, factories, or other anthropogenic sources. As we know, ozone comes from two primary precursors, oxides of nitrogen and volatile organic compounds. When they combine in the presence of sunlight, ozone is produced. Carbon monoxide and methane may modulate these chemical reactions.

Shouldn't we also have a map showing the distribution of sites that monitor these precursors or perhaps a map showing the relative quantitative importance of these precursor emissions? Should we then comment on the extent to which ozone creation is produced from local sources to ones that contribute because of long distance transport mechanisms? Of course, these issues are perhaps more related to control strategies than to the focused mandate of this report.

2. With regard to information on estimating O3 concentrations associated with

nonanthropogenic sources or "background O3" (section 2.4), to what extent is this information appropriately characterized and clearly communicated?

I am delighted at the well-developed Section 2.4, which deals with background levels of ozone and their contributions from natural sources, as well as from anthropogenic emissions outside our borders. The PRB or the NAB (North American background) is important. (There is also the USB, US background.) The text and the figures are extremely useful, and clearly demonstrate that these background levels are appreciable.

Figures 2-8 through 2-10 include the phrase "zero-out modeling." The text, especially 2-12, has a thorough discussion of the limitations of this strategy. But what approaches to dealing with this background are rational and how have they been incorporated into this document? This clearly affects the magnitude of attributable outcomes to ozone and to the nature of the dose-response curves at low and realistic levels. It also calls into question pivotal clinical studies. For example, when examining the responses of exercising human chamber subjects to 60, 70, 80, or 100 ppb ozone, should their measurements be compared to 0 ppb or is it more relevant and appropriate to compare that to background ozone levels? That has rarely been done, but would clearly diminish the magnitude of the changes reported. Again, I would be interested in variability in ozone levels in these geographic areas. The hypothesis that greater variability is associated with heightened responses, even if the average values are the same, is a hypothesis worth exploring. I also think there should be more discussion of co-exposures such as ozone and PM or ozone and heat stress.

David Chock

CHAPTER 2: OZONE MONITORING AND AIR QUALITY

Charge questions:
1. To what extent does the Panel agree that the most relevant information on monitoring (section 2.1), emissions and atmospheric chemistry (section 2.2), and common patterns of ozone concentrations (section 2.3) is presented, and to what extent is the information presented appropriately characterized and clearly communicated?

2. With regard to information on estimating ozone concentrations associated with non-anthropogenic sources or "background ozone" (section 2.4), to what extent is this information appropriately characterized and clearly communicated?

This Chapter provides a very clear and concise picture of the monitoring network, the recent observed ozone concentration trends, both spatially and temporally, and a very brief but relevant and credible description of the emissions and atmospheric chemistry. It also provides a clear description of the different definitions of background ozone concentrations.

There appears to be a confusion in the proper use of the zero-out approach. On p. 2-12, lines 23-24, and again on p. 2A-7, lines 231-241, the authors indicate that removing NOx emissions completely and unrealistically could lead to inflated estimates of background ozone in urban areas where NOx titration of ozone is significant. The authors consider this result as a paradoxical result of the zero-out approach. The implication is that background ozone should not be higher than when anthropogenic precursor emissions are involved. This is erroneous, and in fact contradictory to the acknowledged notion that ozone chemistry is nonlinear. Furthermore, the authors ran separate CMAQ runs for base case (designated as total), NB, NAB and USB, and then took the ratios of one of the background runs to the base case runs as the percent contribution of the background to the base case. This is again erroneous, because it assumes ozone additivity, which violates chemistry's nonlinearity. The zero out approach is not intended for use to assign relative contributions, and relative contribution is not a meaningful concept unless the perturbation due to some source emissions is so small that linearity can be approximated. The source-apportionment method can be used to assign relative background contributions when base case (total) emissions are run, but the resulting relative background contributions are valid only for these particular base case conditions, and the contributions cannot be converted to absolute concentrations to represent background ozone for other emission scenarios including the background emissions alone. Note also that the source apportionment methodology used in CAMx is an approximate tool taking some ozone nonlinearity (local identifications of NOx-limited and VOC-limited environments) into consideration. Its general applicability to a wide range of conditions remains to be established.

Based on the above discussion, in the case of the zero out approach, it is best to change the text in the Chapter from "percent contribution" (like in Fig. 2-11 on p. 2-18) to "ratio" when the background ozone is compared to the base case ozone. Here two model results are compared, and the ratios need not be less than 100%. In the case of source apportionment, "percent contribution" can be used, but is applicable only to the base case run. A separate model run still has to be made to determine the background-only ozone concentrations. So if the authors are willing to replace "percent contribution" by "ratio", the zero-out approach is actually more straightforward and does

not involve the additional assumptions in the source-apportionment methodology.

In the case of W126, the EPA uses a counterfactual assumption by determining the ozone concentrations by separately zeroing out different background emissions based on the definitions of NB, NAB and USB. The procedures are more complicated but are presented in the Appendix. The resulting estimates of background fractions for different definitions of background are generally consistent with but slightly lower than those for the MDA8. This is partly due to the definition of W126 which has a higher weighting for higher ozone concentrations.

One very minor comment: Since seasonal means have been used quite extensively, it is better to incorporate the definition of season in the text rather than leaving it in the footnote (footnote 8 on p. 2-13).

Overall, other than the misstatements about the zero out approach and the erroneous implications that background ozone cannot become greater than when anthropogenic emissions are involved, this is an outstanding chapter in terms of clarity, conciseness, and scientific credibility. The authors should be congratulated for a job well done.

Ana Diez Roux

Chapter 3. Adequacy of the Primary Standard

1. To what extent does section 3.1 (Evidence-based Considerations) capture and appropriately characterize the key aspects of the evidence assessed and integrated in the ISA? To what extent is staff's consideration of the health effects evidence, including the adversity of reported respiratory effects and public health implications technically sound and clearly communicated at an appropriate level of detail? In the Panel's view has the information been appropriately interpreted for the purpose of assessing the adequacy of the current standard?

Overall section 3.1 appropriately describes key aspects of the evidence. I found the consideration if the evidence to be technically sound, and the information appropriately interpreted for the purpose of assessing the adequacy of the current standard. My main comment is that the section would benefit from synthesis and emphasis of the most important facts relevant to assessing the adequacy of the current standard. There is also some repetition within sections (for example the section on at risks populations repeats the key message several times).

2. With regard to the presentation of the exposure and risk information for the purpose of assessing the adequacy of the current standard, to what extent is the information, including associated limitations and uncertainties, sufficiently characterized, appropriately interpreted and clearly communicated?

Section 3.2 also contains abundant repetitions from the REA and could be synthesized.

The section refers to two important issues in estimating the health impact of alternative standards:

1. It is noted that the simulations used to estimate Ozone levels under alternative standards result in spatial patterns different than those observed in the epidemiologic studies on which the health effects measures are based. This would result in different health impacts than those predicted from the epidemiologic studies if one or both of the following conditions are met (a) factors associated with space modify the effects of ozone on heath or (b) spatial mobility of persons within the area is a key driver of individual-level exposures. If we are confident that the impact of these two conditions is absent or negligible then we can be confident in the expected health benefits as predicted despite the change in the spatial pattern.

2. It is noted that based on the approach used to model ozone reductions under alternative standards, ozone levels may actually rise in some areas when meeting lower overall standards. This is because of the dynamics used to model ozone reductions. It should be noted that as a consequence the estimates of the health effects are not precisely the health impacts of reducing ozone to a certain level, but rather the health impact of meeting an alternative standard *through a postulated set of changes to precursors* (some of which results in reductions and some of which result in increases in ozone). This is a subtle but important difference I think. It may be useful to at least note this. Also, is the approach used to model meeting alternative standards (which results in increases in some locations but decreases in others) realistic?

Pg 3-106 lines 23-28 suggest that since approximately 30-60% of the average daytime O3 is attributable to US anthropogenic sources, then 30-60% of total 03- associated health risks in the urban case studies is attributable to US anthropogenic sources. I don't think this statement is accurate: if the reductions in ozone exposure necessary to eliminate or sharply reduce ozone associated health effects can be achieved through redctions in US anthropogenic sources alone, then much more than 30-60% in health effects can be attributed to anthropogenic sources.

3. In the Panel's view, does the discussion in section 3.4 provide an appropriate and sufficient rationale to support staff's preliminary conclusion that the current evidence and exposure/risk information call into question the adequacy of the current standard and that it is appropriate to consider revising the standard to achieve additional public health protection?

Overall I though this section was adequate but could benefit from synthesis and emphasis.

Chapter 4. Consideration of Potential Alternative Primary Standards

1. *In the Panel's view, has the evidence and exposure/risk information, including associated limitations and uncertainties, been appropriately characterized and interpreted for the purpose of considering potential alternative standards?*

   Overall I found the chapter to be very well written and to the point. The point regarding ozone serving as an indicator for a standard meant to provide protection against photochemical oxidants is well taken. The discussion regarding averaging times is focused and supported by appropriate evidence. The discussion regarding the form was also very well written. The points supporting he use of an nth highest daily maximum (as opposed to an expected excedance or percentile-based form) were well stated, however I found the justification of the 4th highest daily max (as opposed to the nth highest) incomplete.

   The section on controlled human exposures studies is an excellent summary although it loses focus in the latter part (pg 4-10 line 19 through pg 4-11 line 6). For example it is not clear why panel studies are discussed here as they are not controlled human exposures studies. The section on page 4-11 lines 6-20 should be consistent with and avid repetitions with pg 4-10 lines 1-18.

   The approach of summarizing associations in cities meeting various alternative standards may be informative but the point of this analysis is not stated clearly and the overall conclusion is not well stated (pg 4-13). What can we conclude then from table 4-1? If studies conducted in areas that have met lower standards do not show an effect do we conclude then that the standard produces appropriate health protection? But if they do does this suggest that an even lower standard is necessary? The logic of this analysis needs to be clarified.

   The subsequent section (on associations below various cutpoints) is clearer but the conclusion could also be summarized more clearly. Is the key point that a standard of 60 ppb is protective whereas a standard of 65 or 70 is not because studies for which all

exposures were below 65-70 still reported associations whereas those at levels below 60 did not? I also found Table 4-2 confusing. The main point needs to be summarized. The reference to the table in the text was confusing.

The section on protection from long term exposures is well done and convincing.

Section 4.4.2.1 would benefit from a final statement of the key conclusions derived from figure 4-1 to 4-4. The same applies to section 4.4.2.2. the bullets are useful but an overall summary statement of what we can conclude from these bullets taken together would be very helpful.

Section 4.4.2.3. The reason for the large difference in the % reduction in mortality associated with meeting a standard of 70 ppb for areas with area wide concentrations > 40ppb and >60ppb is not clear (the footnote does not help clarify). This also applies to other health outcomes. Also the rationale for reporting these two particular estimates is not presented. These types of estimates are repeated later in the chapter so their meaning needs to be clarified.

I'm not sure I would characterize a 9% reduction in ozone associated mortality as a "small " change (pg 4-41 , top of page). In any case it is larger than the effect observed with a standard f 70 ppb so it is not clear why it is considered small.

The chapter also does a reasonable job of grappling and acknowledging the complex issue of uncertainties.

Minor comment: avid using the word mortality (a rate) when you mean total number of deaths (as in Figure 4-10).

2. *In the Panel's view, does the discussion in section 4.6 provide an appropriate and sufficient rationale, supported by the discussions in sections 4.1 through 4.4, to support staff's preliminary conclusions regarding alternative primary standards (including the indicator, level, averaging time and form) that it is appropriate to consider?*

The section provides an appropriate and sufficient rationale. Overall he section is very well organized and the arguments are laid out n a clear and compelling way. A few clarifications, particularly of the data presented, would make this an outstanding chapter.

Tables 4-4 and 4-5 are clear but I found figure 4-13 cryptic. It is not clear exactly what is shown on the Y axis. Is it the ratio of deaths attributable to ozone for alternative standards compared to the 75 ppb standard ? Maybe label the x axis : total ozone attributable deaths, ozone attributable deaths at ozone levels > 20, >40 and >60.

It is difficult to follow the calculations reported on page 4-51 lines 14-21. "For days with area wide concentrations at or above 20ppb a standard with a level of 70….". Is this derived from figure 4-13? But if so isn't this the reduction in deaths attributable to ozone above 20ppb (**not** on days with area wide concentrations at or above 20 ppb??)

3. *Does the Panel have any recommendations regarding additional interpretations and conclusions based on the available information that would be appropriate for consideration beyond those discussed in this chapter?*

   No additional recommendations. Overall this is an excellent chapter. The final section in particular is very well done.

David Grantz

Chapter 5. Adequacy of the Secondary Standard.

1. To what extent does the information in sections 5.1 through 5.5 capture and appropriately characterize the key aspects of the evidence for ozone welfare effects assessed and integrated in the ISA? To what extent does the information in section 5.1 (Nature of Effects and Biologically Relevant Exposure Metric) appropriately summarize the nature of ozone welfare effects and to what extent does it appropriately characterize the evidence with regard to biologically relevant exposures?

This chapter does a nice job off summarizing evidence for welfare effects of ozone exposure. The paragraph (page 1-37, lines 4-14) provides an excellent rationale for considering the range of available studies, and how they might be assessed. Similarly, lines 30-36 on page 1-37 provide an excellent rationale for emphasis on Class I and other protected areas.

The discussion of flux based metrics is useful and appropriate, concluding that potential benefit may eventually derive from such metrics but that excessive uncertainty remains at this time. The discussion regarding reduced uncertainty associated with OTC derived C-R relationships (page 5-10; page 5-31) is appropriate and clearly presented.

The reference to EPA 2013 (page 5-13, line 5) is curious. Meta-analysis has demonstrated that reduced carbon transport to roots is a generally observed phenomenon, though exceptions exist. This sentence needs to be reviewed, for accuracy and for syntax.

2. To what extent is staff's consideration of the welfare effects evidence, including the implications of reported vegetation effects with regard to adversity to public welfare technically sound and clearly communicated at an appropriate level of detail? In the Panel's view has the information been appropriately interpreted for the purpose of assessing the adequacy of the current standard?

The linkage of ecological effects to welfare effects is clearly and appropriately explained, and in sufficient detail.

The cottonwood data (Figures 5-1 and 5-3) receive too much emphasis. These data are clearly outliers, as already noted in the text (page 5-14, line 5) and require further confirmation.

The effort to monetize welfare impacts and benefits is appropriate, though techniques for this are not yet fully developed. Specifically, calculation of consumer and producer surplus data is a useful contribution to quantification of welfare effects. The interaction of agriculture and forestry as modeled by the Forestry and Agricultural Sector Optimization Model with Greenhouse Gases (FASOM-GHG) is mentioned in both the WREA and the PA but not adequately explained.

3. With regard to the presentation of the exposure and risk information for the purpose of assessing the adequacy of the current standard, to what extent is the information, including associated limitations and uncertainties, sufficiently characterized, appropriately interpreted and clearly communicated?

Given that a W126 of 15 ppm hr is in many cases approximately equivalent to the current standard, and that injury is clearly observed below 15 ppm hr, the chapter demonstrates that the current standard is inadequate to protect against welfare effects due to ozone. The consideration of uncertainty (page 5-22, line 23 on) represents an appropriate level of concern and level of interpretation. A similar tone should be adopted in the WREA. However, as in the WREA, the

periodic consideration of uncertainty is again accumulated and repeated later (e.g. Page 5-27). Consolidation near the end of the chapter would enhance the impact of the evidence, reduce redundancy, and substantially shorten the document.
It is unclear why the information on foliar injury in four National Wildlife Refuges (Table 5-6) did not appear in the Risk Assessment.
The discussion at the top of page 5-55, related to economic losses due to bark beetles and to fire, perpetuates the questionable discussion in the WREA, confounding association with any evidence of causation. As stated these effects do not contribute much to the assessment of risk due to ozone nor to the adequacy of the current or alternative standards.
In Table 5-8 the ranges for Normal and Dry Palmer Z indices overlap. Is this intentional?

4. In the Panel's view, does the discussion in section 5.7 provide an appropriate and sufficient rationale to support staff's preliminary conclusion that the current evidence and exposure/risk information call into question the adequacy of the current standard and that it is appropriate to consider revising the standard to achieve additional public welfare protection?

This section can be made more concise and impactful. It clearly demonstrates that ozone induced injury may occur in areas that meet the current standard. It systematically justifies the conclusion that the form of the standard should be reconsidered, and begins to build the case for levels that should be considered.
There is a wrong word/typographical error at page 5-62, line 2, which alters the meaning.

Chapter 6. Consideration of Potential Alternative Secondary Standards.

1. In the Panel's view, has the evidence and exposure/risk information, including associated limitations and uncertainties, been appropriately characterized and interpreted for the purpose of considering levels of protection and potential alternative standards?

It is well justified and consistent with previous CASAC judgment, that a cumulative, non-threshold metric, such as W126 is an appropriate form for the standard. Accumulation over the 08:00 – 20:00 daytime 12 hour period is an acceptable means of generalizing across latitudes and seasons. The compromises involved in this selection are clearly explained.
It is not convincing to argue that a multi-year averaging period is superior to a single year standard. Many cultivated and un-managed species are annuals, or cultivated as annuals. Perennials including trees may compound annual effects, but the effects are inherently single year impacts. Potential differences in soil moisture do not really bear on this question, as there are many factors that vary between years. The only significant reason provided in the chapter to consider a three year average is for increased statistical stability (page 6-11, lines 14-27). This may be sufficient reason, but this should be clearly stated. If a three year averaging period is implemented, it should be at a lower level than a single year standard to protect against single unusually damaging years that will be obscured in the average.
The key issue in the PA with respect to Welfare Effects is the level of the standard. It is surprising to find the statement at page 2-20, line 17-18, that "we are not able to identify a range of appropriate W126 index values". If protection of the most sensitive members of the community is extended to components of ecosystems that impact public welfare, then several potential levels of the standard suggest themselves. Both visibility and seedling biomass exhibit identifiable thresholds at around 10 ppm hr (though visibility is more of a slope change than a decline to low

injury level) and for relative biomass loss it is closer to 7 ppm hr. Crop loss appears to be less sensitive than these other indicators from the current analysis.
There is substantial redundancy in this chapter that could be consolidated for clarity and brevity.

2. In the Panel's view, does the discussion in section 6.5 provide an appropriate and sufficient rationale, supported by the discussions in sections 6.1 through 6.4, to support staff's preliminary conclusions regarding alternative secondary standards (including the indicator, level, averaging time and form) that it is appropriate to consider?

Given the apparent break points around 7-10 ppm hr for various endpoints, it is difficult to understand how staff comes to suggest a standard as high as 17 ppm hr. In many cases, 15 ppm hr is nearly equivalent to the current standard, which is considered here to be inadequate to protect against ozone induced welfare effects. If uncertainty is invoked to prevent over-protection, then 15 ppm hr is the highest justifiable level for the revised standard based on the internal logic of the WREA and PA. If, as seems more likely given patchy species and spatial data coverage, the uncertainty results in potential underestimation of risk, then a lower level should be considered.

3. Does the Panel have any recommendations regarding additional interpretations and conclusions based on the available information that would be appropriate for consideration beyond those discussed in this chapter?

None.

Daniel J. Jacob

***Introduction (Chapter 1)*:**
1. Does the Panel find the introductory and background material (sections 1.1 and 1.2) to be appropriately characterized and clearly communicated?

Yes. My only confusion is in the use of "We" at various points in the text. It isn't clear to me who "We" refers to.

On page 1-9 lines 3-4, "a causal relationship between O3 and 8-hour exposures" doesn't seem to make sense. Probably a typo.

2. In section 1.3, we describe the general approach for the review. This includes the key aspects of the approach employed in the last review in judging the adequacy of the then-existing standards and in selecting revised standards. Does the Panel find this description of the approach in the previous review adequate and clear? Does the summary of the approach in the current review appropriately describe important considerations in this review?

It is generally fine. I only have a few concerns:

1. Page 1-38: The choice of the W126 index as metric for the secondary standard is not revisited as part of the current review. Why is it not?
2. Page 1-40: It's not clear if the analysis of the background contribution to ozone is to have solely a scientific purpose or also a policy purpose.
3. Page 1-40, lines 22-24: the notion that background ozone does not significantly impact the W126 index has been challenged by recent work, particularly for the Intermountain West where the background is high.

Steven Kleeberger

Chapter 4: Consideration of Potential Alternative Primary Standards

1. In Panel's view, has the evidence and exposure/risk information, including associated limitations and uncertainties, been appropriately characterized and interpreted for the purpose of considering potential alternative standards?

I believe the information has been appropriated characterized and interpreted.

2. In the Panel's view, does the discussion in section 4.6 provide an appropriate and sufficient rationale, supported by the discussions in sections 4.1 through 4.4, to support staff's preliminary conclusions regarding alternative primary standards (including the indicator, level, averaging time and form) that it is appropriate to consider?

In my opinion, the rationale and discussion that lead to preliminary conclusions are appropriate. For more clear presentation, staff may wish to consider sub-headings for the consideration of 70, 65, and 60 ppb standard levels. Conclusions (p 4-57) are clearly stated, and the rationale for not considering levels below 60 ppb is appropriate. However, in general, the table and figure legends should provide more detail to enable the reader to understand better what is presented. Specifically, legends for Figures 4.1 – 4.4 and 4.5 – 4.8 are good examples where more detail in the figure legends would help the reader understand all of the included information.

3. Does the Panel have any recommendations regarding additional interpretations and conclusions based on the available information that would be appropriate for consideration beyond those discussed in this chapter?

I do not have additional interpretations or conclusions to add to the chapter. However, I would like to commend the staff for including the key uncertainties and areas for future research (4.7). This is very helpful for the reader to understand that the 'ozone field' still requires much additional effort/studies to fully understand all of the consequences of exposure to this pollutant.

Fred Miller

Introduction (Chapter 1):

1. Does the Panel find the introductory and background material (sections 1.1 and 1.2) to be appropriately characterized and clearly communicated?
2. In section 1.3, we describe the general approach for the review. This includes the key aspects of the approach employed in the last review in judging the adequacy of the then-existing standards and in selecting revised standards. Does the Panel find this description of the approach in the previous review adequate and clear? Does the summary of the approach in the current review appropriately describe important considerations in this review?

Response: The introductory and background material provides the reader with a synopsis of the evolution of the NAAQS review process for photochemical oxidants together with insights into major litigation actions that have occurred over the last 30 years. The description of the approach in the previous review is clearly presented, and the summary of the current review approach captures the salient features of the review process.

O3 Monitoring and Air Quality (Chapter 2):

1. To what extent does the Panel agree that the most relevant information on monitoring (section 2.1), emissions and atmospheric chemistry (section 2.2), and common patterns of O3 concentrations (section 2.3) is presented, and to what extent is the information presented appropriately characterized and clearly communicated?
2. With regard to information on estimating O3 concentrations associated with nonanthropogenic sources or "background O3" (section 2.4), to what extent is this information appropriately characterized and clearly communicated?

Response: My comments address the layout of the figures and tables in this chapter and do not address the charge questions. The authors should consider a different color scheme for some of the figures -- it is hard for the reader to sometimes discern differences (e. g., Figure 2-1). Since figures and tables should stand alone, the legends sometimes need to be expanded so the reader does not have to refer to the text to interpret the table or figure (e. g., Figure 2-2).

Adequacy of the Primary Standard (Chapter 3):

1. To what extent does section 3.1 (Evidence-based Considerations) capture and appropriately characterize the key aspects of the evidence assessed and integrated in the ISA? To what extent is staff's consideration of the health effects evidence, including the adversity of reported respiratory effects and public health implications technically sound and clearly communicated at an appropriate level of detail? In the Panel's view has the information been appropriately interpreted for the purpose of assessing the adequacy of the current standard?

Response: This section captures the important studies discussed in the ISA and integrates them into a logical narrative of what the important findings were in each category of effect such as lung function decrements, pulmonary inflammation, respiratory symptoms, and respiratory mortality. The adversity of the effects and their implications for public health is discussed in a straightforward

and clear manner that leads the reader through the body of data that has been amassed for this ubiquitous pollutant. The level of detail used is appropriate for the task at hand. Most importantly, the studies have been appropriately interpreted and discussed in Section 3.1.3 in the context of assessing the adequacy of the current standard. Section 3.1.3 presents the discussion of adversity in a manner that one might find in a legal briefing document, which speaks to the clarity of the points raised and supported by facts brought forward from the ISA and the HREA.

On page 3-15, the statement is made that the group mean decrements in various controlled human exposure studies at 60 ppb O_3 are not consistently statistically significant. While a correct statement, the authors should add that this is due to a lack of consistency in statistical power among the studies reflecting an inadequate number of subjects in some of the studies. There are places where paragraphs are duplicated almost verbatim on the same page, such as on page 72 for the first 2 paragraphs. On page 3-81, there is a reference to Table 3-12 but no such table appears in the chapter.

2. With regard to the presentation of the exposure and risk information for the purpose of assessing the adequacy of the current standard, to what extent is the information, including associated limitations and uncertainties, sufficiently characterized, appropriately interpreted and clearly communicated?

Response: The exposure and risk information pertinent to assessing the adequacy of the current standard is presented for 3 main categories: the adjusted air quality data, exposure-based considerations, and risk-based considerations. The salient points are clearly presented, and enough detail is provided so the reader can ascertain how much weight to assign to any limitations or uncertainties. Major points being made are always supported by reference to studies or data presented either earlier in this draft of the PA or in the ISA or HREA, which leads one to conclude that the points have been appropriately interpreted.

Consideration might be given to having the Y-axis on each of Figures 3-7 to 3-10 all be from 0 to 30 percent. This would allow the reader to better understand the importance of the different points that are being made in this body of figures.

3. In the Panel's view, does the discussion in section 3.4 provide an appropriate and sufficient rationale to support staff's preliminary conclusion that the current evidence and exposure/risk information call into question the adequacy of the current standard and that it is appropriate to consider revising the standard to achieve additional public health protection?

Response: Absolutely! – Section 3.4 is a "slam dunk". This section, which is just slightly over 5 pages in length, clearly articulates the findings and points that underpin Staff's preliminary conclusion to call into question the adequacy of the current standard and the appropriateness of revising it.

Consideration of Potential Alternative Primary Standards (Chapter 4):

1. In the Panel's view, has the evidence and exposure/risk information, including associated limitations and uncertainties, been appropriately characterized and interpreted for the purpose of considering potential alternative standards?

2. In the Panel's view, does the discussion in section 4.6 provide an appropriate and sufficient rationale, supported by the discussions in sections 4.1 through 4.4, to support staff's preliminary conclusions regarding alternative primary standards (including the indicator, level, averaging time and form) that it is appropriate to consider?
3. Does the Panel have any recommendations regarding additional interpretations and conclusions based on the available information that would be appropriate for consideration beyond those discussed in this chapter?

Response: The evidence and exposure/risk information together with an acknowledgement of the inherent limitations and uncertainties is presented in a logical manner and the data have been appropriately interpreted. The discussion in section 4.6 is well constructed and the case for considering alternative standards is articulated clearly. However, the CASAC O_3 Panel should discussion meeting the summary paragraphs for the 3 alternative levels considered by staff as the arguments/discussions presented in the PA have this reviewer inclined to want to revisit the upper level of 70 ppb based upon scientific body of evidence considerations and not on any policy recommendations.

Howard S. Neufeld

Chapter 1: Introduction

1. This chapter is an excellent introduction to the history of the standard setting process and the legal challenges to previous decisions and their outcomes. It also clearly outlines the goals of the PA. This was a well-written and well organized chapter.

2. Yes and yes. I have no other substantive comments.

Chapter 2: Monitoring and Air Quality

*First item here deleted as it was incorrect – the 8 hr standard should be expressed in units of concentration. I had previously confused it with an exposure index that has hours of ppb*hrs.*

1. The discussion on background ozone is satisfactory and clearly explains how staff is dealing with this concept.

Chapter 3: Adequacy of the Primary Standard

1. The analysis of the ISA conclusions is well done. The particular emphasis placed on studies on the low end of ozone concentrations is commendable. Staff has done a careful and thorough analysis of respiratory effects purportedly caused by ozone. The writing is clear and understandable to the general public.

2. I thought staff did an excellent job of translating the ISA and Risk Assessment results into a statement of adequacy with regards to the current standard. Their analysis of the uncertainty and variability is both thorough and complete and serves to strengthen their conclusions reached about the adequacy of the current standard.

3. As noted above, staff has properly interpreted the new scientific studies and reached a logical conclusion that the current standard is not adequately protecting human health. This chapter was one of the strongest of all in the document: well synthesized, well justified, and with appropriate conclusions drawn from the analyses. I feel, therefore, that they are justified in proposing that the current standard should be revised.

Chapter 4: Consideration of Potential Alternative Primary Standards

1. The analyses of the various standard attributes (indicator, averaging time, form, and level) are well done. The conclusions that follow are appropriately stated and thoroughly justified.

2. I agree in all respects with the conclusions in section 4.6. I think the analyses preceding this section clearly justify the proposals in this section. I found nothing with which to disagree. In particular, staff's analysis of the difficulties in setting the standard too low are much appreciated and their detailed consideration of alternative metrics are to be commended. They clearly justify why the current indicator, averaging time and form of the standard should be retained, but that the level should be lowered.

3. I have no additional recommendations here.

Chapter 5: Adequacy of the Secondary Standard

1. Staff make a strong case (perhaps stronger than in any previous reviews) both that the current secondary standard fails to protect vegetation and ecosystem services from adverse effects, and that the form of the standard is inadequate to provide such protection. Sections 5.1 through 5.5 clearly lay out the argument for the impacts of ozone on ecological processes and ecosystem services. Their descriptions and analyses of the uncertainties are refreshingly clear and unbiased. Despite the paucity of data in certain areas, staff is still able to make statements of causality and risk with which to evaluate the adequacy of the current welfare standard.

2. Staff did an excellent job regarding technical soundness and providing clarity to the public. I might note that during the risk analysis, and also here, staff compares changes in seedling growth under ozone exposure to adult tree changes in circumference, expressed as percent losses. However, circumference as a surrogate for growth is somewhat questionable since growth in diameter would imply an increase in area, which is related to changes in circumference by the square (and total growth, i.e. volume, would increase by the cube). A 10% reduction in circumference would translate into a 19% reduction in trunk area. How this might affect the conclusions drawn I don't know, but it suggests that using circumference underestimates the effects.

   With regard to adequately interpreting the data for the purpose of assessing the adequacy of the current standard, I am satisfied with the analyses presented and agree with staff's interpretations.

3. The uncertainties are well laid out and explained and place the interpretations into proper perspective. It appears both in this section, and in the previous chapter, that staff has gone to great lengths to justify the conclusions reached from their analyses of the ISA and REA documents as well as making sure to follow the requirements of the Clean Air Act. I was very impressed with this aspect of the current PA.

4. I agree with most of the conclusions in section 5.7. I support moving from the 8-hr standard form to the cumulative W126 index, and I agree with using the maximum 3 month interval. I think a strong case can still be made that the timing should be for an individual year, rather than averaged over three years. I am concerned about the level and at what upper limit it should be set. Most of the analyses seemed to find effects below 15 ppm*hrs (many at 10 or even 7), so it would seem reasonable to set it lower than 15 ppm*hrs, and not as high as 17 ppm*hrs. Also, in many cases 15 ppm*hrs is approximately equivalent to the current standard, so keeping it that high would engender little benefit.

Chapter 6: Consideration of Alternative Secondary Standards

1. Staff make a very strong case for switching the secondary form to the W126 and for averaging it over three months and using the highest three month interval as the metric that is evaluated. The lack of an arbitrary threshold and the higher weightings for higher ozone concentrations are the appealing aspects of this ozone metric. I also support the use of daylight hours for this metric, from 8 am to 8 pm. There could be some quibbling about whether this time interval is shifted too late in the day, and misses some uptake early in the morning, but as staff notes, ozone at low elevations rarely rises to high values before 8 am. However, in mountain regions, ozone is high over a 24 hour period, and as such, starting

the metric at 8 am instead of 7 am or even 6 am, may miss some aspects of the exposure impacts. However, given that only one time interval can be selected, and that most areas across the country are not at high elevations, I can support using the 8 am to 8 pm interval.

2. I am less convinced that the three year average form is better than a single year form. I fully understand staff's arguments about stability but remain to be convinced that this would provide superior protection (requisite protection) given the comments from CASAC on this topic from earlier reviews. And if the increased statistical stability is that important, then the level should be reduced in order to protect the most sensitive vegetation.

   There seems little justification for allowing the increased flexibility of setting the level of the W126 above 15 ppm*hrs. See comments above in the Chapter 5 section. Most of the tree data show 2% annual growth losses at up to 14 ppm*hrs. It would seem prudent to take out the statement that the Administrator could consider exposure values above 15 ppm*hrs.

3. I have no further recommendations except one: the PA document contains many redundancies, which take away from its ability to clearly transmit its message to the public. This is especially so for Chapters 5 and 6. If there is any prudent way to shorten these chapters some, I think that would improve this document.

In conclusion, I would like to commend EPA Staff for their due diligence and hard work in putting these documents together. I was much impressed by the analyses and conclusions. They represent a real step forward from the first drafts.

Armistead (Ted) Russell

Overall, I found the PA informative and providing much of the information needed to inform the Administrator in regards to potentially modifying the ozone NAAQS. It generally has a good discussion on the adequacy of the current health and welfare standards, and potential revisions. Further, the preliminary revisions are in line with the evidence provided in the ISA and the analyses in the two REAs. The current presentation, for the most part, picks a reasonable balance between the desire to make the PA readable, concise and to the point, and providing sufficient information. The greatest need has to deal with issues involving the increases in lower levels of ozone in response to controls designed to reduce higher ozone levels. This issue impacts the risk and exposure assessments and the form of the standard.

Chapter 2: O3 Monitoring and Air Quality

1. To what extent is the most relevant information on monitoring, emissions and chemistry, and common patters of ozone concentrations is presented, etc.:

The current Chapter 2 is very streamlined; too much so. At present, it focuses primarily on the issues involving background ozone, which is an appropriate discussion, but there is at least one larger issue that needs to be addressed here in some detail, that being the response of lower level ozone levels to controls. As shown in the Health REA this is a very important consideration, and should be discussed in more detail in Chapter 2 as it is an important consideration in the potential form of a standard, and the possible limitations of the current form. As such, the PA should provide a discussion of how the observed ozone levels at various percentiles have been found to be evolving, e.g., the decreases in higher levels and increases in lower levels. This should be augmented with results from the modeling. The PA should provide the Administrator with a firm understanding that controls oriented at reducing the peak levels of ozone (e.g., the 4th highest annual MDA8) may not be that effective at reducing more typical levels and may actually increase ozone levels on lower ozone levels and also increase 24 hour levels on a broader range of days. Chapter 2 needs an overall summary. What are the major take-home points for the Administrator (and others) from Chapter 2?

2. Is the discussion of Background Ozone appropriately characterized and adequately communicated?

The discussion of background ozone is much more extensive than any other part of the air quality characterization. None-the-less, there are some missing pieces. First, the method by which the NB is calculated (e.g., models used) should be further described. The discussion of the source apportionment model estimates is much too minimal to really understand what is being done. Provide an extra sentence or two. The sentences beginning on 2-16, line 9 going to line 14 are not clear. The potential use of monitoring to estimate background should be discussed. I would also include a bit more on the range of controversy surrounding this issue. This should be recapped later on as well. In general, however, the discussion does provide a good and reasonably thorough assessment of the "background ozone" issue on setting a standard.
In the PA, I would also bring forward more of the results on the levels of controls needed to just meet various levels of the standards being assessed, both for the health and welfare standards. One could use the two tables from the Appendices for the H-REA (Table 2 of Appendix 4) and the W-REA (Table 4A-2). This could be part of a synthesis as well.

Page 2-16, line 16: It is not apparent that the difference between 66 and 59 is due to the definitional approach versus the difference in model .
2-17, l21-23: What exactly does this mean? In particular, does "but for" mean if background sources were not present, there would not be an exceedance, or does it mean, if only background sources were present, there would be an exceedance?
2-9, l34: When discussing specific metrics, be very careful as to what is being said as to not be ambiguous.
2-9 l 14. I think you mean "intrusions" not "inversions"

Chapter 4: Form of the health based standard. Given the potential for controls to increase lower levels of ozone, one might consider a different form of the standard that would be protective at lowering high levels of ozone and also decreasing mid and lower levels. This should at least be discussed at a level that could lead the Administrator/reader to confidently say that the current form is appropriate. Indeed, one might come to the conclusion that another form would be better.

Chapter 6. One of the conclusions of the W-REA was that just meeting the current standard (75 ppb) leads to very similar ozone levels as a W126 of 15. This should be brought out more. Indeed, I am not sure that the difference is beyond the uncertainties in the approaches used. Given that a W126 of 15 roughly corresponds to a 75 ppb standard, the rationale for considering a W126 of 17 should be further discussed.

The PA could use a synthesis as well. The Synthesis should include how the health and welfare standards might work together. Further, it could identify the critical findings that would likely drive the decision to keep or revise the standards, including the characterization of the likely benefits of various choices of the standards and the uncertainties that are key. Maybe this will be included as part of the Executive Summary in the next draft. At present, the Executive Summary is a bit too short and weak. For example, it currently uses the phrase "call in to question", which might be stated as finds that the current standard is inadequate to protect health, and provide the specific evidence to suggest so. One thing that the synthesis could provide is the similarities in ozone fields at different levels of the health and welfare standards under consideration. It could also include tables showing the levels of control to meet those levels.

Executive Summary:
As noted above, the ES is rather short, and appears to be a work in progress. Compared to the ES's for the two REAs, this one does not stand up. It really does not do the PA justice. An important point from the W-REA should be made here, that being that a W126 of 15 ppm-hrs is very similar to just meeting the current standard. The ES could also use a synthesis.

Helen Suh

Adequacy of the Primary Standard (Chapter 3)

1. *To what extent does section 3.1 (Evidence-based Considerations) capture and appropriately characterize the key aspects of the evidence assessed and integrated in the ISA? To what extent is staff's consideration of the health effects evidence, including the adversity of reported respiratory effects and public health implications technically sound and clearly communicated at an appropriate level of detail? In the Panel's view has the information been appropriately interpreted for the purpose of assessing the adequacy of the current standard?*

The section is comprehensive. While this comprehensive summary is accurate and clear, it does appear in other documents. Further, its volume is overwhelming. As a result, the Chapter is somewhat unfocused, and evidence supporting its conclusions regarding the adequacy of the primary standard are not sufficiently targeted. As one suggestion to increase the clarity of the Chapter, Section 3.4 could be moved to the beginning of the Chapter, with the summaries of findings and other sections used to support the conclusions of Section 3.4. Other comments and suggestions are discussed briefly below:

- Figures need additional annotation, better titles and a better color scheme (for example Figures 3-7 to 3-10).
- The discussion of the policy-relevant background risks should be removed or moved to the discussion of the history.
- Subsection 3.1.2
    - The subsection 3.1.2, which addresses the nature and strength of conclusions, could be more targeted, focused less on a summary of all findings but instead on findings contributing to the adequacy determination. In this way, it can provide an introduction to the topics discussed in later subsections, including adversity of effects and concentrations associated with observed impacts, with these later subsections being at the heart of Section 3.1.
    - While causal determinations were stated at relevant places in the text, it would be helpful to include at the beginning of Section 3.1.2 a table that lists the causal determinations made in the 2006 and current ISA. This table should be accompanied by a brief discussion that explains how health outcomes were selected for further discussion in Section 3.1.2.
    - Summary figures and/or tables should be added to Subsection 3.1.2 in place of text for the major outcomes. While summary figures of epidemiological studies were included (such as Figure 3-2), a preferable model for summarizing findings may be to include tables, such as Tables 4-23 and 5-19 from the NOx ISA, which summarize information succinctly across evidence and study types to support causality determinations. It may also be helpful to note in these summary figures and tables studies or findings that show ozone impacts at standard-relevant ozone levels.
    - The summary of ozone-attributed health outcomes in epidemiological studies should include a brief discussion of the potential for confounding of the observed effects, since is it is relevant to causal determinations.
    - Page 3-53, line 29-33 The statement beginning "Generally, the epidemiologic studies used nearest air monitors to assess ozone concentrations…explained by the different exposure assignment methods used…" relates not only to this subsection but also to previous

subsections. It should be mentioned in earlier subsections, if only to say that significant associations were found despite the presence of exposure error resulting from the use of ambient ozone concentrations to assess exposures.

- Section 3.1.3: It seems that much of this section is a repeat of earlier discussions, although with a different focus. To reduce repetition, it may make sense for this discussion to be incorporated into section 3.1.2, possibly using the following outline
    - Health outcome 1
        - Summary of findings (including causality determination and whether it has changed)
        - Adversity of impacts
        - Ozone concentrations associated with health effects
    - Health outcome 2 (etc.)

- Section 3.1.4: Important discussion that summarizes information across health effects. The summary tables 3-1 and 3-2 were very helpful. This subsection would be improved with further organization and integration of the findings across health effects, as this would help to demonstrate coherence with regard to the lowest ozone concentration at which key health impacts have been observed. The discussion of coherence doesn't begin for many pages (page 3-71) and is consequently buried. It would be better to summarize the earlier discussion more concisely to highlight the key points of the section.

2. With regard to the presentation of the exposure and risk information for the purpose of assessing the adequacy of the current standard, to what extent is the information, including associated limitations and uncertainties, sufficiently characterized, appropriately interpreted and clearly communicated?

Again, this section is comprehensive and accurate. The discussion on page 3-94 was clear, to the point, and clearly related to standard-setting considerations. Similarly, the discussion of uncertainties was well written and concise. However, in other areas, the section was overly repetitive and should be summarized more concisely to highlight discussions that focus more specifically on standard-relevant issues.

3. In the Panel's view, does the discussion in section 3.4 provide an appropriate and sufficient rationale to support staff's preliminary conclusion that the current evidence and exposure/risk information call into question the adequacy of the current standard and that it is appropriate to consider revising the standard to achieve additional public health protection.

Section 3.4 was a terrific synthesis of the current health and exposure/risk information. In some ways, the chapter should lead with this section, as it captures the key points, summarizes the evidence clearly and concisely in a manner that keeps the focus on standard-setting considerations.

Consideration of Potential Alternative Primary Standards (Chapter 4)

1. In the Panel's view, has the evidence and exposure/risk information, including associated limitations and uncertainties, been appropriately characterized and interpreted for the purpose of considering potential alternative standards?

This Chapter was well written, comprehensive, clear, and focused. It was a pleasure to read. Suggestions to improve the chapter are rather minor. They include:

- In general, the figures and tables need further descriptions, better titles and better annotation. For example, several figures, such as Figures 4-1 and 4-2 and Figures 4-3 and 4-4, seem extremely similar (with the same title) and it is not clear what the graphs are depicting.
- The question on page 4-8 (and following text) should be reworded to focus on lung function decrements and airway hyperresponsiveness rather than controlled human studies given the inclusion of panel study results in this section. Correspondingly, the title of the subsection "epidemiological evidence" to hospital admissions and mortality should be reworded to focus on health outcomes rather than study type.
- The discussion on page 4-14 of the cut-point analysis is a repeat of that in Chapter 3. Rather than repeat the discussion, instead reference it and draw conclusions relative to the level of the standard.
- The interpretation of Figure 4-10 is unclear from the text and the figure.
- Figure 4-9 (and other figures), the y axis should be labeled as "deaths" and not "mortality". In the text, it may be helpful to explain reasons for the observed between-city variability in figures such as Figure 4-9. Is this between-city variability important?

2. *In the Panel's view, does the discussion in section 4.6 provide an appropriate and sufficient rationale, supported by the discussions in sections 4.1 through 4.4, to support staff's preliminary conclusions regarding alternative primary standards (including the indicator, level, averaging time and form) that it is appropriate to consider?*

Yes. The section was clear, well-written, and effective.

3. *Does the Panel have any recommendations regarding additional interpretations and conclusions based on the available information that would be appropriate for consideration beyond those discussed in this chapter*?

The chapter was a cogent presentation of the rationale behind the report's conclusions.

James Ultman

To what extent does section 3.1 (Evidence-based Considerations) capture and appropriately characterize the key aspects of the evidence assessed and integrated in the ISA? To what extent is staff's consideration of the health effects evidence, including the adversity of reported respiratory effects and public health implications technically sound and clearly communicated at an appropriate level of detail? In the Panel's view has the information been appropriately interpreted for the purpose of assessing the adequacy of the current standard?

The chapter does an excellent job of summarizing the current evidence-based considerations of short-term and long-term O_3 health effects and placing them in the contexts of adversity to the individual as well as public health implications.

Section 3.1.1

I believe that appendix 3A describes specific modes of action much better than current text in section 3.1.3. I would go so far as to replace the latter by the former.

Regarding figure 3-1 (copied from figure 5-8 in the ISA), the mode-of-action pathways for specific adverse effects do not necessarily follow from the hierarchy in the figure. For example, inflammation and cell remodeling are both placed on the second level while epithelial metaplasia appears on the third level. I would argue that a more accurate sequence for the metaplastic changes that have been observed in toxicological studies might be: inflammation (second level)→cell repair and remodeling (third level)→ metaplasia (fourth level not currently on figure).

The mode-of-action discussion implies that secondary oxidation products rather than ozone itself drive the respiratory responses. This point of view is based on well-accepted dosimetry analyses indicating that ozone is almost completely reacted during its transit through the mucous layer in conducting airways. However, these analyses are not directly supported by experimental observations. Furthermore, it is still quite possible that ozone quantitatively penetrates the thin surfactant layer in alveoli distal to the conducting airways.

The first step in the respiratory response is described as oxidant stress, but this catch-all term does not tell the whole story. While ozone and some of its reaction products are oxidants that can be detoxified by endogenous antioxidants, other reaction products (e.g., aldehydes) cause cell damage by non-oxidative processes.

Page 3-13 (line 6) gives a literal interpretation of the McDonnell and the Schelegle models in terms of a specific mechanism (i.e., oxidant stress). Although this is a reasonable hypothesis, one should recognize that both models are simply two-stage mathematical constructs that include a build-up of inhaled *ozone* dose by continuous inhalation in competition of a reduction in biologically-effective dose by some clearance or metabolic process. It is also possible that such a reduction is the result of absorption of ozone in airways regions (such as the nose and mouth) that are proximal to the airway sites where dysfunction actually occurs.

Section 3.1.2.2

Entry is missing from the table of contents.

With regard to the presentation of the exposure and risk information for the purpose of assessing the adequacy of the current standard, to what extent is the information, including associated limitations and uncertainties, sufficiently characterized, appropriately interpreted and clearly communicated?

I think that the document is quite effective in presenting risk reduction information between recent conditions and just meeting the current standard.

In the Panel's view, does the discussion in section 3.4 provide an appropriate and sufficient rationale to support staff's preliminary conclusion that the current evidence and exposure/risk information call into question the adequacy of the current standard and that it is appropriate to consider revising the standard to achieve additional public health protection?

This section provides a sufficient rationale for staff's conclusion the current standard does not provide adequate health protection.

Sverre Vedal

Ch. 3. Adequacy of the current standard

Adequacy of the Primary Standard (Chapter 3)
1. To what extent does section 3.1 (Evidence-based Considerations) capture and appropriately characterize the key aspects of the evidence assessed and integrated in the ISA? To what extent is staff's consideration of the health effects evidence, including the adversity of reported respiratory effects and public health implications technically sound and clearly communicated at an appropriate level of detail? In the Panel's view has the information been appropriately interpreted for the purpose of assessing the adequacy of the current standard?

- Although there is some evidence from epi studies that anti-oxidants partially protect from lung function declines due to ozone exposure (3-15), the more direct evidence from human clinical studies does not support this.
- Toxicologic morphologic changes do not provide evidence of bronchial hyperresponsiveness – these changes occur at the level of the respiratory bronchiole and alveoli, which probably does not influence larger airways effects such as those in asthma. Perhaps there is some problem with the wording or intent here (3-40, lines 28-32).
- Observations on three endpoints (and conclusions on 3-119):
    - Long-term exposure effects on lung-function: valiant attempts are made to provide evidence for this using relatively weaker studies, whereas the best study (CHS) showed no effects.
    - Short-term exposure effects on symptoms in asthmatic children: arguments to justify not considering the two multi-city (multi-site, actually) studies that find no evidence of associations. I wonder if this same effort would have been made if these were the only two studies in which evidence for associations was found?
    - New onset asthma: argument is presented as to why the lack of a confirming main effect of long-term ozone exposure in the CHS is ok in light of the interesting gene polymorphism interaction analysis.
- In "Pulmonary Structure and Function" (3-44) there is no mention of the CHS study that provides the strongest evidence against long-term exposure effects of ozone on lung function.
- Does Section 3.1.3, Adversity of Effects, really require so much exposition? It's very repetitious of earlier material.
- Is compression of the ozone distribution due to model-based air quality adjustments realistic, i.e., decreases in high ozone concentrations and increases in low concentrations? (3-86)
- It is difficult to accept, if I'm understanding this correctly, that the percentage of children experiencing lung function declines is approximately the same as the percentage of children exposed (Figures 3-7 through 3-14). For example, in Atlanta, 14% of children are estimated to experience at least one exposure of concern at or above 60 ppb (Fig. 3-7) and 17% of children are estimated to experience at least one day of >10% decline in FEV1 (Fig. 3-11).
- In addition to the sensitivity of epidemiologic-based risk on C-R functions due to choice of study region, e.g., large vs. small (3-114), there was also seeming substantial sensitivity to choice of regional vs. national C-R functions.
- Missing references detected: Rojas-Martinez 2007 (3-44, line 11), Joad 2006 (3-45, line 16), NRC 2008 (3-85), US EPA 2007 (3-98).
- [Refs to check: Kim 2011 AJRCCM re: inflamm at 60 ppb, Lin EHP 2008 on first asthma

admission; Fanucchi 2006 infant rhesus]

2. With regard to the presentation of the exposure and risk information for the purpose of assessing the adequacy of the current standard, to what extent is the information, including associated limitations and uncertainties, sufficiently characterized, appropriately interpreted and clearly communicated?

I found this section to be among the best sections of Chapter 3.

3. In the Panel's view, does the discussion in section 3.4 provide an appropriate and sufficient rationale to support staff's preliminary conclusion that the current evidence and exposure/risk information call into question the adequacy of the current standard and that it is appropriate to consider revising the standard to achieve additional public health protection?

Most definitely. It's well done.

Peter Woodbury

Executive Summary

Page ES-2, line 26. I don't think suggesting that values "somewhat above" 15 ppm-hrs is appropriate. If the EPA staff judge that a value above 15 ppm-hrs for W126 should be considered, then the analysis throughout the WREA and the PA should include a specific value above 15 ppm-hrs so that this suggestion can be reviewed by CASAC and others.

***Introduction (Chapter 1)*:** This chapter provides context for the review, including the background of past reviews, as well as the scope and approach for the current review. This includes discussion of the basis for the current standard.

1. Does the Panel find the introductory and background material (sections 1.1 and 1.2) to be appropriately characterized and clearly communicated?

Yes, this material is important and the coverage is appropriate.

2. In section 1.3, we describe the general approach for the review. This includes the key aspects of the approach employed in the last review in judging the adequacy of the then-existing standards and in selecting revised standards.

Does the Panel find this description of the approach in the previous review adequate and clear?

Yes, this material is important and the coverage is appropriate.

Does the summary of the approach in the current review appropriately describe important considerations in this review?

Overall, this summary is cogent and useful. However, regarding the secondary standard, certain important conclusions are misleading and require revision. For example, on page 1-27, lines 7-10 states that the "magnitude of the response becomes increasingly uncertain". A similar statement is made on page 1-36, lines 24-27. This is somewhat misleading. Data such as the concentration-response functions for individual tree seedling species, supported by results from other methods such as FACE and naturally occurring gradients demonstrate that some species are very sensitive to ozone and show decreased growth at very low chronic exposure levels, while other species show little response to much higher levels. A similar result is found for crop species. Thus there is strong evidence of decreased growth and yield of some common tree and crop species at very low ozone levels. The more important source of uncertainty at these low levels is determining what degree of growth decrement should be considered unacceptable to protect public welfare. This issue extends throughout the PA, and the PA could be strengthened by more specifically quantifying the spatial extent and degree of impact expected at current ozone exposures, the current standard, and at the alternate standards. For example, rather than focusing on the "median RBL", quantify the number of counties containing sensitive tree species that are expected to have growth loss of greater than 1%, 2%, etc.

***O3 Monitoring and Air Quality (Chapter 2)*:** This chapter provides a description of the current

O3 monitoring network and recent concentrations, information on emissions and atmospheric chemistry, common patterns and variability in O3 concentrations, as well as, discussion of current information on estimating O3 concentrations associated with non-anthropogenic sources.

1. To what extent does the Panel agree that the most relevant information on monitoring (section 2.1), emissions and atmospheric chemistry (section 2.2), and common patterns of O3 concentrations (section 2.3) is presented, and to what extent is the information presented appropriately characterized and clearly communicated?

2. With regard to information on estimating O3 concentrations associated with non-anthropogenic sources or "background O3" (section 2.4), to what extent is this information appropriately characterized and clearly communicated?

The large difference in "counterfactual" vs "source apportionment" methods for estimating backgrounds, is important for W126 (p. 2-23 and elsewhere). This issue is challenging to describe, but is done reasonably well. In particular, the summary on Page 2-26 and 2-27 is helpful..

"Anthropogenic" emissions of VOCs are distinguished from "natural" sources. However, as mentioned on Page 2-7 (lines 19-20) and 2-8 (lines 1-4), "natural" emissions can include human-influenced emissions from fire, agriculture, forestry, and other land management practices. This issue of definitions should be further clarified (see my further comment below). Even more importantly, the extent to which such human-influenced emissions are included in "natural background" or other "background" ozone scenarios, should be clarified, and preferably quantified since it could affect interpretation of what portion of ozone is potentially controllable.

On a more minor but related point, the terms "anthropogenic", "man-made" and "manmade" all seem to be used to mean the same thing, with "anthropogenic" being the most commonly used term. Perhaps the term "man-made" should be used throughout for emissions directly from human activity, with the term "anthropogenic" used more broadly to include indirect human-influenced emissions such as from agriculture, forestry, and other land management practices. Or if the term "anthropogenic" is used throughout then, make it clear that this narrower and I believe non-standard definition is being used.

Adequacy of the Secondary Standard (Chapter 5): This chapter discusses key aspects of the welfare effects evidence and exposure/risk information, particularly relevant to consideration of adequacy of the current secondary standard and specifically describes staff's consideration of this information in reaching preliminary conclusions about the adequacy of the current standard.

1. To what extent does the information in sections 5.1 through 5.5 capture and appropriately characterize the key aspects of the evidence for ozone welfare effects assessed and integrated in the ISA?

In general, the information from the ISI is presented appropriately. The brief quotes are particularly helpful, as are the references to specific portions of the ISI.
To what extent does the information in section 5.1 (Nature of Effects and Biologically Relevant Exposure Metric) appropriately summarize the nature of ozone welfare effects and to what extent does it appropriately characterize the evidence with regard to biologically relevant exposures?

In general, Section 5.1 is useful and appropriate, and the use of questions and answers is a good format. However, I suggest rearranging the order of material within each subsection such that the question is answered at the end of the section discussing each question. For example, the current section on page 5-1, lines 21-33 should be placed just before Page 5-3, line 9. Furthermore, each question should be clearly answered. For example, Page 5-7, before Line 1, there should be a short paragraph that directly answers the question about appropriate paradigm posed on Page 5-5, lines 29-30.

Page 5-2, line 25. Replace "vegetative species" with "many species of vegetation".

Page 5-13, line 5. Delete "although"

2. To what extent is staff's consideration of the welfare effects evidence, including the implications of reported vegetation effects with regard to adversity to public welfare technically sound and clearly communicated at an appropriate level of detail?

In general, the draft is appropriate and substantially improved from the previous draft, with some caveats. First, it is important to appropriately address the fact that the sensitivity of most tree species and many crop species has not been quantified in terms of a C-R function. It should not be assumed that species of unknown sensitivity are not sensitive to ozone. For example, on page 5-18, lines 28-29, discusses "if present in these specially protected areas", referring to 7 of the 12 tree species for which C-R functions are available. More appropriately, it should state that "if ozone-sensitive species are present". This may sound like a minor point about language, but I believe it is actually an important point about how to apply the available scientific data to ecosystems, and it has large implications. For example, stating that "Half (6/12) of species with known C-R functions would have growth decreases greater than 5%, and of these species are representative of responses of unmeasured species, this degree of impact would occur in [state percentage of studied locations with this level of response]. In brief, it is important not to assume that unmeasured species are not sensitive to ozone, it is much more appropriate to assume that the sensitivity of species without C-R functions might be similar to the range of sensitivity for those species with C-R functions.

In the Panel's view has the information been appropriately interpreted for the purpose of assessing the adequacy of the current standard?

Please see comments above.

3. With regard to the presentation of the exposure and risk information for the purpose of assessing the adequacy of the current standard, to what extent is the information, including associated limitations and uncertainties, sufficiently characterized, appropriately interpreted and clearly communicated?

The choice of the word "paradigm" seems odd in question on p. 5-5, but I don't have a suggestion of a better term.

Make sure not to define "adverse" effects too narrowly (p. 5-6), loss of biomass growth could be important even if the species is not harvested for timber or fiber. This topic is mentioned

elsewhere, and on Page 5-12 this issue is appropriately broadened, but perhaps on p. 5-6 some mention of other effects could be made, or a reference to other locations that address these broader impacts.

p. 5-13, line 5 delete "although"

Figure 5-1 should be improved by moving the legend to the right of the main figure panel and arranging the legend species in the same order (top to bottom) as in the main figure panel.

p. 5-14, line 16 and elsewhere. As I mentioned in comments on the previous drafts, it is difficult to interpret a "median" response for both ozone-sensitive and relatively insensitive species. Instead, it makes sense to characterize the expected impacts on the sensitive species, and quantify the spatial extent and effect on biomass growth for (1) known sensitive species (i.e., those that are shown be particularly sensitive with their C-R functions, and (2) the same result assuming that the 12 species with known C-R functions represent all tree species. A more complex scheme could be developed to try to extrapolate known species to unknown species based on physiological characteristics, as was done for crops in the FASOM analysis, but there would be substantial uncertainty in such extrapolation.

p. 5-21. I think that the "modeling regions" in Table 5-4, are the 9 large US climate regions shown in Fig. 4-6, but this should be made explicit (the term "modeling region" doesn't seem to be defined in the text currently).

p. 5-24. As in my comments on the first draft WREA and PA, and second draft WREA (Section 6.8) I still have a question about the RBL values weighted by basal area. Does the denominator basal area in the calculation include only the 12 species with C-R functions or does it include all species? If the latter, it is biased. If the former, the interpretation will vary depending on what fraction of the basal area is for species without C-R functions. Furthermore, if the goal is to assess ozone effects on total biomass growth of a mixed-species forest, then this value is not very informative because it will overestimate impacts in mixed species forests because of not including competition between sensitive and insensitive species (see previous comments on competition). If the purpose is to assess ozone impacts on sensitive species, this value is also not informative because it underestimates impacts on sensitive species for the same reason. A comparatively small growth decline in a sensitive species (e.g. 2%) based on a seedling study may translate into a larger effect at the stand scale.

Page 5-25. The method used with FASOM for forest growth is based on individual species C-R functions, but that is only appropriate for mono-specific stands. For mixed-species stands, overall forest growth will not be affected as much as would be implied by a weighted average of the growth rates (or yield losses) from individual C-R functions. This is because of competition among species with different sensitivity to ozone. This is a serious limitation in the approach for mixed-species forests that are common in many parts of the USA.

I still don't agree about ignoring impacts on farmers and forest owners in high ozone areas just because national assessments include winners and losers. An example of such a calculation is presented on page 5-32, line 9-12 for soybean for 2 counties in Kansas. The number of counties in which yield loss is predicted to exceed 1, 2% or 5% could be tabulated for alternative standards.

See comment for Chapter 6 of the WREA related to this topic. Summaries of county-scale information could be added to Table 6-4.

p. 5-30 line 31 etc. Clarify that NCLAN covered multiple locations in the USA and multiple crops, with multiple O3 exposure levels using consistent methods – all of these factors are very important because they mean that the results are highly valuable for national risk assessments.

Check for occurrences of "PSDI", should be "PDSI" throughout.

p. 5-41, line 6. Change "by of" to "by".

Figure 5-5 (page 5-48). In figure legend, provide some information about the sites.

EPA should assure that uncertainties are not suggested or implied to always weaken the case for a more stringent standard. For example, the paucity of data on ozone sensitivity of most US plant species should be considered as "anticipated" that there are a large number of unidentified sensitive species, as well as of course many less sensitive species.

Page 5-62, replace "commiserate" with "commensurate".

Page 5-63, beginning line 31. Rephrase the sentence on line 31 to be more definite (replace "might be" with "are").

Page 5-65, lines 20-23. Replace "likelihood and magnitude of a response become increasingly uncertain" with "magnitude of effects become smaller". As discussed above, the evidence is very strong for tree biomass loss and crop yield loss for sensitive species at W126 values of 5-10. It is the magnitude of the effect that is smaller. And for less sensitive species, there will be little or no biomass loss at low ozone exposure values. Again the important uncertainty is determining what magnitude of an effect is important for welfare, not whether there are any effects at lower ozone exposure levels. This is an important distinction.

4. In the Panel's view, does the discussion in section 5.7 provide an appropriate and sufficient rationale to support staff's preliminary conclusion that the current evidence and exposure/risk information call into question the adequacy of the current standard and that it is appropriate to consider revising the standard to achieve additional public welfare protection?

In general, this section is appropriate. However, because it is a summary section, many comments on other sections of this chapter should also be applied to this section. Also, note my specific comments above for this section (i.e. Pages 5-59 through 5-65).

Consideration of Potential Alternative Secondary Standards (Chapter 6): This chapter discusses key aspects of the welfare effects evidence and exposure/risk information particularly relevant to consideration of potential alternative secondary standards and specifically describes staff's consideration of this information in reaching preliminary conclusions on alternative standards appropriate to consider.

1. In the Panel's view, has the evidence and exposure/risk information, including associated limitations and uncertainties, been appropriately characterized and interpreted for the purpose of considering levels of protection and potential alternative standards?

In general, yes, except for a few points. First, I don't understand the suggestion of values "somewhat above" 15 ppm-hrs. If the EPA staff judge that a value above 15 ppm-hrs for W126 should be considered, then the analysis throughout the WREA and the PA should include a specific value above 15 ppm-hrs so that this suggestion can be reviewed by CASAC and others. Second, there needs to be greater attention paid to sensitive species. While there is analysis of both sensitive crop and tree species, much of the focus is on a median response, for example for tree species. And even when sensitive species are discussed, it is with statements such as 7 of 12 species had a relative yield loss of X below an ozone exposure value of y. However, with trees, the two most sensitive species had a substantial yield loss well below this value. Thus more attention should be paid to including information about commonly occurring sensitive species. Third, more attention should be paid to crop yield loss and tree biomass growth losses at smaller spatial scales such as counties. If yield of a sensitive crop such as soybean is greater than 5% in a county, it affects the farmers in that county, even if at larger regional or national scales there are smaller impacts on yield or on producer surpluses. Fourth, the suggestion to use a 3 year averaging period is not supported by the available data, nor has it been supported by CASAC. The only justification for a 3-year averaging period is to improve the stability of the classification of regions as being in or out of compliance. Greater attention should be paid to accounting for cumulative impacts of a 1% or 2% loss in growth of tree species and other impacts if a 3-year averaging period is to be used. Specifically, a lower value of the standard would be appropriate for a longer averaging period than 1 year. See also detailed comments below.

Additionally it seems to me that there is likely to be a strong bias effect of using across-the-board NOx reductions (see comments for Chapter 4 of WREA). For the large climate regions used, there is a wide variation in ozone exposure values throughout the region. If "across-the-board" cuts are used to reduce ozone exposure in the highest locations, then ozone exposure in locations where it is lower to start with will tend to become very low. But an actual targeted control strategy would likely only reduce NOx in the locations necessary to bring the high ozone-exposure locations into compliance, thus having much less effect on other locations than would across-the-board cuts. If this bias is large, then it strongly affects all of the risk analyses based on the alternate standards, because impacts of ozone under the alternate standards would be underestimated because ozone exposure values are underestimated for much of each region. If this bias is substantial it could mean that the ozone exposure is underestimated and thus the benefit of the alternate standards might all be underestimated. This could mean that the benefits of any of the alternate standards compared to "just meeting" the current standard and are underestimated as well.

Page 6-8, line 4-5. Change to "extremely highly correlated metrics, with Pearson correlation coefficients of 0.99".

Page 6-9, line 16-17. I don't find that this conclusion is warranted – the data support an annual time frame. There should be some compelling reason to use a multi-year time frame. Note that the proposed form already includes a 3-month period, so it is not as sensitive as an hourly or 8-hour period to extreme events.

Page 6-9, lines 18-36. As in my comments in the previous 1st draft PA, it is not appropriate to assume that the only welfare effect of crop yield loss is total producer and consumer surpluses. I think that a goal of avoiding yield losses of sensitive crops of 5% or greater for each county would be appropriate to protect welfare. Farmers growing sensitive crops in high ozone locations can be considered a "sensitive population" for welfare impacts, and their crop yields should be protected. Furthermore, I do not find any support herein for the idea that the data do not support an annual time period for yield losses in annual crops.

Page 6-12. The argument that stability of compliance is of value is stated on this page. I agree with the previous CASAC statement (lines 26-29) that if a multi-year period is chosen for stability purposes, the level of the standard should be lowered to prevent exceedences of a threshold of impact. This is a very important point as it affects the choice of a level depending on the averaging time (1 versus 3 years).

Page 6-17 and 6-18. This discussion of individual species responses to different W126 levels is very helpful and informative, as is Table 6-1. However, I think Table 6-1 could be reformatted to make it easier to read by moving much of the text into column headings, and increasing the number of columns. For example "ppm-hrs" should appear at the top of the first column, under the column heading rather than be repeated on every row. And "median species" could be a column heading for both tree seedlings and crops, as could "loss". These changes would make it easier to see the actual values in the table. Also, it is not clear to me what "varying lower" means in the table.

Page 6-24, lines 10-24. See previous comments herein and for the first draft PA regarding a question with the weighted RBL scheme.

Page 6-25, line 14. Change "great" to "greater".

Page 6-30, lines 14-18. It would be helpful to include here and elsewhere the fraction of US forests represented by the species with C-R functions (probably using basal area) as well as the fraction of US crop area covered by the crop species with C-R functions. This helps quantify the uncertainty, which is quite different for crops and forest tree species. This information is available in the REA and could be summarized here as well.

Beginning page 6-30 line 6 and onward. While it is useful to list the various uncertainties, it would be more helpful to give some idea of how the uncertainty might affect the interpretation. I realize that this is challenging, but my concern that it is important to communicate what is known with reasonable certainty versus what is really unknown. I commend the staff for accomplishing this challenging task well in the REA in Table 7-23. There is quite a lot of certainty in estimates of biomass loss for forest tree seedling species and crop species for which C-R functions have been developed. Because several dominant crop species have C-R functions, there is a quite a lot of certainty about impacts of ozone on crop yield across most annual cropland in the USA. But it is much more uncertain to extrapolate from the 12 forest tree species to all forest tree species in the US. For uncertainty in ozone exposure, while it is true that the sparseness of rural monitors means that in many regions there is uncertainty, there are large portions of the US where monitors are dense enough, and where there are not large mountains or other features that make interpolation

more difficult, such that regional estimates of ozone exposure are pretty certain, even if there are somewhat larger uncertainties for individual locations.

Page 6-39, lines 29-31. This is a non-sequitur. The difficulty in determining the degree of impact that is important for welfare is not related to the question of averaging the standard across 1 or 3 years. Also, as quoted in the PA (page 6-36, line 19), the CASAC said "averaging across years in not recommended".

Page 6-43, line 13-14. As for the tree species, crop results should focus on both a group (such as median) response and also individual species response. This is important because sensitive species such as soybean are very widespread and important crops. Also, there is little attention given in the PA and WREA to non-crop annual species. The results for annual crops can be considered as also indicators for a very large number of annual non-crop species that may have many welfare values. While I appreciate the much greater attention to effects on crop yield in this second draft compared to the first draft, I still think a bit more attention to crops is warranted given the strength of the database on crop yield response to ozone.

Page 6-43, line 35 and elsewhere. While the EPA chose to focus on 2% biomass loss for forest species, greater acknowledgement should be made that CASAC recommended 1 to 2%. This is particularly important in conjunction with the decision by EPA to focus on a 3-year rather than annual averaging time as recommended by CASAC.

Page 6-44, line 21. I think it is worth discussing the implications of focusing on sensitive tree species in addition to the existing focus on median and majority species. I think it is misleading for example on line 14-15 to say "less than 9 or 10" without including that the 2 remaining species have much higher predicted biomass losses. I am not suggesting that these sensitive species should be the primary focus, but rather including them for consideration rather than not including them, as happens with the current emphasis. As for sensitive crop species, these sensitive tree species are ecologically important and widespread. Furthermore, they may also serve as indicators for the high likelihood that there are other sensitive tree species for which C-R functions have not been developed. As an example of how information on sensitive species might be included, the human health section of this PA includes estimates of the number of not just all children, but also asthmatics for example in Table 4-45, to better represent sensitive populations.

Page 6-45, lines 21-23. I don't understand why these lines are here. The EPA has selected a range of appropriate alternate values for the standard, and it is in accord with many previous CASAC recommendations. But this sentence says that the Administrator can reasonably choose a value beyond this range. This is a very open ended statement, and I don't understand what it is based on. If the EPA staff judge that a value above 15 ppm-hrs for W126 should be considered, then the analysis throughout the WREA and the PA should include a specific value above 15 ppm-hrs so that this suggestion can be reviewed by CASAC and others.

Page 6-45 lines 24-35. This is very helpful, pointing out the implications of choosing different values for the standard among the values put forward by EPA and by CASAC.

Page 6-46, lines 3-5 and 13-14. As mentioned above for the previous page, I don't understand the suggestion of values "somewhat above". If the EPA staff judge that a value above

15 ppm-hrs for W126 should be considered, then the analysis throughout the WREA and the PA should include a specific value above 15 ppm-hrs so that this suggestion can be reviewed by CASAC and others.

Page 6-48, lines 12-13. The only support that I find in the PA for a 3-consecutive-year time frame is stability of compliance among years. As discussed in the PA and in my prior comments, CASAC specifically recommends NOT averaging across years, so I think this recommendation should be reconsidered or at least further qualified.

Page 6-48, line 18, Again I find the usage of "somewhat above" to be vague and misleading and impossible to evaluate, see comments above.

2. In the Panel's view, does the discussion in section 6.5 provide an appropriate and sufficient rationale, supported by the discussions in sections 6.1 through 6.4, to support staff's preliminary conclusions regarding alternative secondary standards (including the indicator, level, averaging time and form) that it is appropriate to consider?

For the most part yes, except for the issue of mentioning values "somewhat above 15 ppm-hrs (see comments above). Also, I think more attention needs to be paid to suggesting a lower value for the standard if using a 3-year rather than a 1-year averaging time.

3. Does the Panel have any recommendations regarding additional interpretations and conclusions based on the available information that would be appropriate for consideration beyond those discussed in this chapter?

Please see comments above.

Ronald E. Wyzga

Introduction (Chapter 1):

Charge questions:

1. Does the Panel find the introductiry and background material (sections 1.1 and 1.2) to be appropriately characterized and clearly communicated?

I believe that these sections are extremely well-written. I especially commend the Agency's intent to provide a document "written to be understandable to a broad audience". With respect to the latter, there are a few places where jargon and/or technical terms have crept in that could be clarified for a more naïve audience. For example, p. 1-7, l. 19: "certoriori", p. 1-21: "Controlled Human Exposure Studies"; p. 1-36: "OTC", which I believe refers to open-topped chambers.

In other places legal references could be placed in a footnote as they are distracting from the text; e..g, p. 1-3, ll. 24-28.

2. In section 1.3, we describe the general approach for the review. This includes the key aspects of the approach employed in the last review in judging the adequacy of the then-existing standards and in selecting revised standards. Does the Panel find this description of the approach in the previous review adequate and clear? Does the summary of the approach in the current review appropriately describe important considerations in this review?

This is well-written and provides an important introduction.

A minor comment is that some of the material in section 1.3.1.2.3 is repetitious of material in section 1.2.2.

Some comments on Chapters 3 and 4:

Does the currently available scientific evidence and exposure/risk information …support or call into question the adequacy of the current O_3 primary standard?

The studies of Schelegle et al. (2009) clearly demonstrate that there are adverse health effects among a subpopulation of healthy young subjects after a 6.6 hour expsoure of 72ppb, a level below the current standard; hence there is clear scientific evidence that exposures at the current standard could lead to adverse effects and alternative standards are appropriate for consideration. It should be noted that lung function results from clinical studies at exposure levels at or below 72ppb also demonstrate a clear response. Whether these are "adverse" depends upon the definition of "adverse"; the Schelegle et al study clearly statisfies the definition of "adverse" as defined by the American Thoracic Society, which defines adversity as a combination of lung functions changes accompanied by symptoms. If an alternative definition of "adverse" is used, it should be clearly defined. There is also epidemiological support that finds adverse health responses at contemporary ozone levels in the US.

What is the range of potential alternative standards that are supported by the currently available scientific evidence….?

Clearly, from the above, the upper end must be at or below 72ppb. The Schelegle et al. (2009) study was conducted using healthy volunteers 18-25 years of age for an exposure of 6.6 hours. At issue is how to compensate for the shorter exposure time and the possible impacts among a more sensitive or less healthy population. Figure 2 from the Schelegle et al. (2009) study shows that at higher exposure levels lung function and symptom scores tend to increase with longer exposures. This increase is less apparent for the 70 and 80 ppb protocol exposures; nevertheless prudence suggests that some compensation could be made to adjust for the 6.6 hour exposure period. Panel studies also suggest that more sensitive individuals can respond to levels at or below 70ppb. There is no easy way to extrapolate from exposure levels at which healthy adults respond to comparable levels among sensitive populations, such as asthmatics and children; hence there is uncertainty associated with the lower end of the proposed range.