**UNITED STATES ENVIRONMENTAL PROTECTION AGENCY**
**WASHINGTON D.C. 20460**

**OFFICE OF THE ADMINISTRATOR**
**SCIENCE ADVISORY BOARD**

April 15, 2016

EPA-CASAC-16-002

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

Subject: CASAC Review of the EPA's *Integrated Science Assessment for Sulfur Oxides – Health Criteria (External Review Draft – November 2015)*

Dear Administrator McCarthy:

The Clean Air Scientific Advisory Committee (CASAC) Sulfur Oxides Panel met on January 27-28, 2016, to peer review the EPA's *Integrated Science Assessment for Sulfur Oxides – Health Criteria (External Review Draft – November 2015),* hereafter referred to as the Draft ISA. The CASAC's consensus responses to the agency's charge questions and the individual review comments from the CASAC Sulfur Oxides Panel are enclosed.

Overall, the CASAC finds the draft ISA to be a comprehensive assessment of the available science relevant to understanding the health impacts of exposure to sulfur oxides, but has several recommendations on improving and enhancing the scientific clarity and rigor of the assessment.

The CASAC finds the Executive Summary to be comprehensive and well written. The CASAC suggests that the EPA consider revising the language for a broader, non-technical audience and eliminating technical jargon as much as possible. The correlation between maximum 5-minute $SO_2$ concentrations with corresponding 1-hour concentrations should also be summarized in the ES. Definitions for short- and long-term exposures need to be consistent throughout the ES as well as the entire ISA. It should be clearly stated in the ES and early on in Chapter 1 that the current 1-hour $SO_2$ Primary NAAQS is based heavily on the results of controlled human exposure studies. Chapter 1 adequately summarizes and integrates the key findings of subsequent chapters. However, the CASAC suggests that a brief rationale on the use of $SO_2$ as the indicator be provided in the introductory paragraph of the chapter.

The CASAC finds that the source categories and definitions of major sources are inconsistent throughout Chapter 2 as well as the entire ISA and recommends that these be consistent. The chapter should include locations and emissions for point sources (energy-generating units, integrated steel and iron mills and smelters) near urban centers. The chapter should also elaborate on $SO_x$ transformations and chemical compositions and summarize $SO_2$ measurement methods, interferences, and quantitative

comparisons. Data from the 2013-2014/15 time period with all monitor sites (~400 sites) should be included in the data analysis. A discussion and referencing of power law relationships between peak-to-mean ratio and averaging time should be included as well as additional puff/plume models and performance evaluations using $SO_2$ monitoring data. Additional rationale and data analyses should be provided to illustrate the relationship between $SO_2$ and co-pollutants.

The CASAC supports Chapter 3 as a distinct, standalone chapter on exposure. However, the material should be reorganized to eliminate redundancy within the chapter and with material that belongs in other chapters, to ensure each key topic is covered coherently in one place, to improve the clarity and flow, and to distinguish between properties that are relevant to all criteria air pollutants versus those that are specific to $SO_2$ (e.g., components of exposure and the role of reactivity, exposure modeling, exposure measurement error). Chapter 3 should cover exposure assessment and modeling as needed to support interpretation of epidemiological studies. It should also cover exposure modeling that is needed to apply exposure-response relationships derived from clinical experiments to quantify $SO_2$ exposure and risk assessments. The response to the Chapter 3 charge questions provides one possible way to reorganize the chapter.

Chapter 4 on dosimetry and modes of action is a well-written and generally complete. The section on respiratory absorption would be much improved by adding a succinct, but general discussion of the transport of any inhaled reactive gas as well as the definitions of commonly-used dosimetrics. $SO_2$-metabolizing capabilities should be more completely discussed, including a comparison of the nose with other respiratory tract regions as well as with extra-pulmonary tissues.

Chapter 5 on health effects summarizes and distills a very large literature base, but can be improved by focusing on the key studies relevant to the assessment of the evidence for the causal framework and providing a more rigorous assessment of study quality. The CASAC concurs with the determination of a "causal relationship" between short-term $SO_2$ exposure and respiratory effects and the determination of a "suggestive but not sufficient to infer a causal relationship" between long-term $SO_2$ exposure and respiratory effects. The CASAC also concurs with the determination of an "inadequate to infer the presence or absence of a causal relationship" between cardiovascular effects and long-term $SO_2$ exposure and the determination of a "suggestive but not sufficient to infer as causal relationship" between total mortality and short-term $SO_2$ exposure. The CASAC is not convinced that the current evidence supports a change in the determinations (from "inadequate to infer the presence or absence of a causal relationship" to "suggestive but not sufficient to infer as causal relationship") for long-term $SO_2$ exposure and total mortality, reproductive/developmental effects, and cancer. This is due to potential co-pollutant confounding and lack of toxicologic evidence to support biological plausibility. The CASAC does not concur with the change in the determination (from "inadequate to infer the presence or absence of a causal relationship" to "suggestive but not sufficient to infer as causal relationship") for short-term $SO_2$ exposure and cardiovascular effects due to potential confounding from co-pollutants. The $SO_2$-effects in studies of myocardial infarction hospital admissions were significantly reduced or eliminated in the two-pollutant model, and the one study that had temporally-resolved estimates of $SO_2$ exposures found no association between hourly ambient $SO_2$ concentrations and risk of myocardial infarction.

Chapter 6 presents a good introduction to the identification of populations of greater risk to adverse health consequences from $SO_2$ exposure and does a good job of summarizing existing literature on the extent to which health effects are potentially modified by various factors. However, greater clarification is needed for several components of this chapter including: clearly articulating the three factors that

cause populations to be at increased risk of $SO_2$-related health effects; clarifying terminology and data presented in the tables; expanding on descriptive information on populations at risk; and better integration with Chapter 5 (particularly the discussion of effect modification).

The CASAC appreciates the opportunity to provide advice on the Draft Sulfur Oxides ISA and looks forward to reviewing the Second Draft Sulfur Oxides ISA.

Sincerely,

**/Signed/**

Dr. Ana Diez Roux, 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 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
# Sulfur Oxides Panel

**CHAIR**
**Dr. Ana V. Diez Roux**, Dean, School of Public Health, Drexel University, Philadelphia, PA

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

**Dr. Judith Chow**, Nazir and Mary Ansari Chair in Entrepreneurialism and Science and Research Professor, Division of Atmospheric Sciences, Desert Research Institute, Reno, NV

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

**Dr. Elizabeth A. (Lianne) Sheppard**, Professor of Biostatistics and Professor and Assistant Chair of Environmental & Occupational Health Sciences, School of Public Health, University of Washington, Seattle, WA

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

**CONSULTANTS**
**Dr. John R. Balmes**, Professor, Department of Medicine, Division of Occupational and Environmental Medicine, University of California, San Francisco, San Francisco, CA

**Dr. James Boylan**, Program Manager, Planning & Support Program, Air Protection Branch, Georgia Department of Natural Resources, Atlanta, GA

**Dr. Aaron Cohen**, Principal Scientist, Health Effects Institute, Boston, MA

**Dr. Alison Cullen**, Professor, Evans School of Public Policy and Governance, University of Washington, Seattle, WA

**Dr. Delbert Eatough**, Professor of Chemistry, Department of Chemistry and Biochemistry , Brigham Young University, Provo, UT

**Dr. H. Christopher Frey**, Glenn E. Futrell Distinguished University Professor, Department of Civil, Construction and Environmental Engineering, College of Engineering, North Carolina State University, Raleigh, NC

**Dr. William C. Griffith**, Associate Director, Department of Environmental and Occupational Health Sciences, Institute for Risk Analysis & Risk Communication, School of Public Health, University of Washington, Seattle, WA

**Dr. Steven Hanna**, President, Hanna Consultants, Kennebunkport, ME

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

**Dr. Farla Kaufman**, Epidemiologist, Office of Environmental Health Hazard Assessment, Reproductive and Cancer Hazards Assessment Section, California EPA, Sacramento, CA

**Dr. David Peden**, Distinguished Professor of Pediatrics, Medicine & Microbiology/Immunology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

**Dr. Richard Schlesinger**, Associate Dean, Dyson College of Arts and Sciences, Pace University, New York, NY

**Dr. Frank Speizer**, Edward Kass Distinguished Professor of Medicine, Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA

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

**SCIENCE ADVISORY BOARD STAFF**

**Mr. Aaron Yeow**, Designated Federal Officer, U.S. Environmental Protection Agency, Science Advisory Board, Washington, DC

# U.S. Environmental Protection Agency
# Clean Air Scientific Advisory Committee

**CHAIR**
**Dr. Ana V. Diez Roux**, Dean, School of Public Health, Drexel University, Philadelphia, PA

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

**Dr. Judith Chow**, Nazir and Mary Ansari Chair in Entrepreneurialism and Science and Research Professor, Division of Atmospheric Sciences, Desert Research Institute, Reno, NV

**Dr. Ivan J. Fernandez**, Distinguished Maine Professor, School of Forest Resources and Climate Change Institute, University of Maine, Orono, ME

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

**Dr. Elizabeth A. (Lianne) Sheppard**, Professor of Biostatistics and Professor and Assistant Chair of Environmental & Occupational Health Sciences, School of Public Health, University of Washington, Seattle, WA

**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, Washington, DC

**Consensus Responses to Charge Questions on**
**EPA's Integrated Science Assessment for Sulfur Oxides – Health Criteria**
**(External Review Draft – November 2015)**

## Executive Summary and Chapter 1 – Summary of the Integrated Science Assessment (ISA)

*The Executive Summary is intended to provide a concise synopsis of the key findings and conclusions of the Sulfur Oxides ($SO_X$) ISA for a broad range of audiences. Please comment on the clarity with which the Executive Summary communicates the key information from the $SO_X$ ISA. Please provide recommendations on information that should be added or information that should be left for discussion in the subsequent chapters of the $SO_X$ ISA.*

The CASAC finds the Executive Summary and Chapter 1 to be comprehensive and well organized. The Policy-Relevant Considerations Section is particularly well written and could be used as a model for other sections in the Executive Summary. Table ES-1 is also very useful and appropriate.

However, the CASAC suggests that the language of the Executive Summary be revised for a broader, non-technical audience by eliminating technical jargon. In addition, the same word-for-word sentences and paragraphs found in Chapter 1 should not be used in the Executive Summary. Important points made in Chapter 1 (and in the following chapters) should be reiterated, but reworded, for the more general (lay) audience. Important changes made in subsequent chapters for the revised draft of the ISA should also be reflected in the Executive Summary and Chapter 1.

The CASAC also has the following specific comments and suggestions:

- It should be clearly stated in the Executive Summary and Chapter 1 that the controlled human exposure studies are the principal rationale behind the 2010 1-hour Primary Sulfur Dioxide ($SO_2$) National Ambient Air Qualtiy Standard (NAAQS) and that this standard also provides protection from chronic exposure effects;
- Correlation of maximum 5-minute $SO_2$ concentrations with corresponding 1-hour concentrations should be summarized;
- Definitions for short- and long-term exposures need to be clear and consistent throughout the text of the entire ISA, including the Executive Summary;
- Some of the footnotes used in the first page of the Executive Summary should be elevated to the body of the text; and
- It is important that ambient background concentrations of $SO_2$ be mentioned in the Executive Summary.

*Chapter 1 summarizes key information from the Preamble about the process for developing an ISA. Chapter 1 also presents the integrative summary and conclusions from the subsequent detailed chapters of the $SO_X$ ISA and characterizes available scientific information on policy-relevant issues.*

*Please comment on the usefulness and effectiveness of the summary presentation. Please provide recommendations on approaches that may improve the communication of key findings to varied audiences and the synthesis of available information across subject areas. What information should be added or is more appropriate to leave for discussion in the subsequent detailed chapters?*

Overall Chapter 1 adequately summarizes and integrates the key findings of subsequent chapters. The summary table and references are well done. In addition, summary statements used in the subsections of Chapter 1 are well-crafted and could be reiterated in subsequent ISA chapters. The CASAC suggests that a brief rationale for why $SO_2$ is used as the indicator for gaseous SOx be provided in the introductory paragraph of this chapter. In the Conclusion section (1.8), the authors correctly emphasize that the current ambient air quality standard is heavily based on the results of the controlled human exposure studies, but this needs to be more strongly reflected in the earlier sections (e.g., 1.6.1).

**Chapter 2 – Atmospheric Chemistry and Ambient Concentrations of Sulfur Oxides**

*Chapter 2 describes scientific information on sources, atmospheric chemistry, and measurement and modeling of ambient concentrations of gaseous sulfur oxides.*

*To what extent is the information presented regarding sources, chemistry, and measurement and modeling of ambient concentrations accurate, complete, and relevant to the review of the $SO_2$ NAAQS?*

Sources of Sulfur Dioxide

Figures where source emissions are summarized are inconsistent with respect to emission ranges and definitions of major sources (e.g., > 100, 1000, or 2000 tons/year). It would be more informative to include a table containing $SO_2$ emission rates and co-pollutants (e.g., hydrogen sulfide, dimethyl sulfide, oxides of nitrogen, carbon monoxide, and particulate matter) for major sources (including smelters, steel mills, on- and off-road diesel vehicles, and marine vessels). Estimated emission rates from volcanoes and their potential impacts also should be given.

Section 2.2 should include a discussion of the Data Requirements Rule for the 1-hr $SO_2$ NAAQS as it pertains to identifying major sources. A summary of the types of sources included and a map showing the location of each source would be helpful in understanding the existing large $SO_2$ sources. Changes in this pattern should also be noted. In particular, a map with major $SO_2$ sources (> 2000 tons/year) grouped by source category can be used to illustrate areas needed for modeling and/or additional monitoring.

Atmospheric Chemistry and Fate

The use of terms $SO_2$ and $SO_x$ needs to be clarified up front. $SO_2$ chemistry, aqueous phase formation of sulfate ($SO_4^=$) and other compounds, as well as their linkages to exposure and deposition in the respiratory system, need to be emphasized.

The importance of pollution sources and formation of non-sulfate compounds such as inorganic particulate S(IV) species, organic S(IV) species (e.g., bis-hydroxy dimethyl sulfone) and organic S(VI) species (e.g., alkyl sulfates) requires additional discussion. Studies such as Alarie et al. (1973) and Amdur (1971) demonstrated the relationship between exposure to inorganic S(IV) compounds and exacerbation of $SO_2$ inhalation responses in animals. These compounds are potential confounders or moderators of $SO_2$ health effects in epidemiological studies where copper smelter or integrated steel mill emissions are abundant and the possible influence of these compounds should be discussed.

Measurement Methods

Personal Monitoring Techniques and emerging technologies in small, low-cost microsensors in Section 3.2.1 are part of the measurement methods. This section should be combined with and discussed in Sections 2.4.1 – 2.4.3. The data completeness requirements for calculation of 5-min and/or 1-hr averages should be clarified. Details on how values below lower detection limits (LDLs) are handled in calculating averages should also be provided.

Available Federal Reference Methods (FRMs) or Federal Equivalence Methods (FEMs) for $SO_2$ measurements and their specifications should be described. A summary table of $SO_2$ measurement comparison studies that specifies operating principles, averaging times, accuracy, precision, LDLs, interferences, etc. would be informative. Potential reasons for discrepancies between FRM/FEM and other measurement methods should be discussed. This would be important for determining the equivalence between standard monitors (1-hr averages with LDL < 2 ppb) and trace monitors (5-min averages with LDL < 0.2 ppb). It would also be helpful to show a time series of these comparisons and verify their similarities and differences by location, season, and time of day.

*Please comment on the extent to which available information on the spatial and temporal trends of ambient $SO_2$ concentrations at various scales has been adequately and accurately described. In particular, what is the extent to which the analyses of recently available 5-min $SO_2$ concentration data are informative in considering relationships between 5-min and 1-hr $SO_2$ concentrations?*

Environmental Concentrations

The data analysis from the current 2010-2012 data at 42 sites in six core-based statistical areas (CBSA)/metropolitan areas should be extended to include 2013-2014 $SO_2$ data at all sites (2015 data will be available on May 1, 2016). Methods to classify "high" versus "moderate" concentration sites should be clarified. The five- to six-year time period (2010-2014/15) will better represent the equivalence, comparability, and/or predictability between maximum 5-min and 1-hr data. The most recent 3-year design value should be used to demonstrate attainment status.

As the ultraviolet fluorescence (UVF) FRM instruments are capable of producing short-duration averages, consistent reporting of each 5-min average by the states is recommended. This allows for the examination of consecutive elevated 5-min $SO_2$ concentrations to evaluate the duration of plume touchdowns and downwind mixing. These data will facilitate analyses of exposure durations and patterns, especially when $SO_2$ concentrations exceed the 200 ppb health benchmark. Making each 5-min average available would also provide a useful database to evaluate a 5-min $SO_2$ NAAQS indicator or to allow future studies that might require a full continuous database.

Comparisons and scatter plots should clarify the time periods and averaging times. The 5-minute maximum concentrations for any hour should correspond with the hourly average concentrations for that same hour (i.e., no data points should be below the 1:1 line on Figure 2-28). No statistical analysis of the results is included, and there is not much of a conclusion on the peak-to-mean ratios (PMRs) besides the statement, "These results emphasize that 1-hr average concentrations at or below 75 ppb, for the most part, represent hourly 5-minute maximum values below 200 ppb." Additional statistical analysis of the data should be performed with the extended data set (2010-2014/15 for all sites) to build a more compelling conclusion on the PMRs.

The utility of 5-min data can be demonstrated based on the "0.2 power law" relationships (Slade 1968; Turner 1970). Figure 2-28 shows that hourly maximum 5-min values are 1-3 times higher than the corresponding hourly values with an average PMR of 1.6 for 5-min averages. This is consistent with "0.2 power law" relating PMR to averaging time. As the modeling community uses PMR as guidance for averaging time to the 0.2 power, a literature review examining the effect of PMR by time of day (stability), wind speed, stack height, distance from sources, and site locations should be performed.

The zone of influence in Figures 2-14 to 2-19 can be better presented to include $SO_2$ emissions from each National Emissions Inventory (NEI) facility to describe spatial variability and distributions of ambient $SO_2$ concentrations. This comparison should include facilities (e.g., coal- and oil-fired power plants, smelters and integrated steel mills) expected to produce high local concentrations under appropriate meteorological conditions. Major copollutants associated with these sources should also be indicated. Figure 2-17 includes a site with an integrated steel mill (A) but does not include urban $SO_2$ monitoring data from the lead smelter south of St. Louis, which is now closed. A comparison of ambient concentrations before and after the closure with the expanded data set at this site would be valuable. The incorrect identification of the location of the Hayden copper smelter (Figure 2-19) should be corrected. The source zone of influence (e.g., 5, 10, 15 km) centered on major NEI facilities should be hypothesized beforehand (rather than after the statistical output) in order to guide the analysis. Maps of spatial variations of a 3-year average of daily maximum 1-hr $SO_2$ concentrations at the annual $99^{th}$ percentile based on existing Air Quality System (AQS) monitors for the United States and for each selected core-based statistical area (CBSA)/metropolitan area would provide insight on spatial variations of $SO_2$ concentrations.

The temporal variation between summer and winter or cold and warm seasons should be illustrated. The plots presenting time series distribution for $99^{th}$ percentile daily maximum 1-hr $SO_2$ (> 75 ppb) should be added to illustrate potential non-attainment areas.

Dispersion Modeling

Lagrangian puff models (e.g., CALPUFF and SCICHEM) should be considered in addition to steady-state plume models (e.g., AERMOD). Other models may better simulate situations with light and variable winds and convergence zones, as the Gaussian plume model is limited to the assumption of steady-state, and straight-line plumes over a 1-hr modeling period.

Model performance evaluations (MPEs) should be conducted with monitoring data for each model application, the type of MPE should be specified, and acceptance criteria should be determined (Hanna and Chang 2012). The model performance statistics should be considered when weighing the credibility of the results (e.g., less weight for areas without monitors) and the resulting exposure impacts. For comparison with the NAAQS, regulatory models usually predict the peak of the concentration distribution, not necessarily corresponding to the same sampling site and time period (e.g., using Q-Q plots of highest ranked modeled versus highest ranked observed concentrations). However, models used for risk and exposure assessments require the prediction of concentration distributions paired in time and space using scatter plots and statistics to draw proper conclusions.

Side-by-side maps comparing modeled and measured $SO_2$ concentrations for a given region (e.g., Huang et al. 2011) would shed light on the representativeness of air quality modeling to detect plume touchdowns and the adequacy of the number of monitor sites within 15 km of major $SO_2$ sources.

*How informative is the analysis of correlations between $SO_2$ and co-occurring pollutant concentrations for interpretation of epidemiologic studies?*

Co-pollutants

Only Pearson correlations were given to address the relationship between $SO_2$ and other co-pollutants. Since many of the scenarios that lead to high co-pollutant correlations are known, the discussions should begin with some suggested hypotheses followed by analysis. Individual sites and dates and/or times that show high correlation coefficients (e.g., $R^2 > 0.8$ or 0.9) may warrant additional analysis and discussion (e.g., are they related to plume touchdown?). The possibility that $SO_2$ is more correlated with CO and average $PM_{2.5}$ (rather than $PM_{2.5}$ sulfur) needs to be explained.

Speculation on low wintertime ozone ($O_3$) and $SO_2$ correlations needs to be supported with additional data analysis. It should address the regional nature of $O_3$ formation versus local $SO_2$ plume touchdown. There is higher $SO_2$ and less $O_3$ due to the less oxidation of $SO_2$ and decreased photochemistry during wintertime.

## Chapter 3 – Exposure to Ambient Sulfur Dioxide

*Chapter 3 describes scientific information on exposure to ambient $SO_2$ and implications for epidemiologic studies. To what extent is the discussion on methodological considerations for exposure measurement and modeling clearly and accurately conveyed, appropriately characterized, and relevant to the review of the $SO_2$ NAAQS?*
*Please comment on the accuracy, level of detail, and clarity of the discussion regarding exposure assessment and the influence of exposure error on effect estimates in epidemiologic studies of the health effects of $SO_2$.*

The CASAC supports having exposure as a distinct, standalone chapter. The primary recommendation is that the material in Chapter 3 should be reorganized. This reorganization should eliminate redundancy within the chapter and with material that belongs in Chapters 2 or 4, ensure each key topic is covered coherently in one place, improve the clarity and flow, and distinguish between properties that are relevant to all criteria air pollutants versus those that are specific to $SO_2$ (e.g., components of exposure and the role of reactivity, exposure modeling, exposure measurement error). Consideration should be given to adding a glossary of exposure-related terms, possibly by expanding upon the list of acronyms and definitions starting on p. xii. The important terms in this chapter should be defined (e.g., exposure, exposure concentration, exposure error and bias) and consistently applied. Most papers reviewed in the previous Sulfur Oxides ISA document (U.S. EPA, 2008) do not need full review in this document. The discussion of ventilation should be moved to Chapter 4 and all in-depth discussions of deterministic air quality modeling (e.g., CMAQ, AERMOD) should be consolidated into Chapter 2 with appropriate cross-referencing in Chapter 3. Chapter 3 should cover exposure assessment and modeling as needed to support interpretation of epidemiological studies. It should also cover exposure modeling that is needed to apply exposure-response relationships derived from clinical experiments to quantify $SO_2$ exposure and risk assessments. The discussion of probabilistic exposure modeling should focus on providing the necessary background to relate estimated exposure-response relationship results to populations in the Risk and Exposure Assessment (REA).

The CASAC outlines below one possible way of reorganizing Chapter 3. The first section could describe the context and scope of the chapter, clearly conveying the purpose for discussing exposure in the ISA. This could be followed by a conceptual overview of human exposure that incorporates specific discussion of the unique features of $SO_2$ as a pollutant and how this informs the understanding and modeling of exposure to $SO_2$. The discussion of exposure should not be conflated with dose. The introduction to the next section could cover applications of exposure assessment relevant to this review, addressing the distinction between exposure assessment for epidemiologic inference versus for risk assessment. This would be followed by specific discussion of exposure metrics directly relevant to the health studies reviewed in the document; presumably this section will cover exposure modeling in the context of these exposure metrics. Finally, it is important to discuss exposure measurement error and its implications for the interpretation of epidemiologic studies.

As part of clearly defining exposure error, the ISA should distinguish between error associated with how well the given metric is measured or modeled and error associated with the use of a given metric as a surrogate for the preferred metric. Specifically, exposure error includes: (1) the exposure measurement error that comes from the uncertainty of the exposure metric being used relative to an uncertain version of that metric; and (2) the exposure error that arises when a surrogate target parameter of interest in the epidemiologic study is used rather than a target parameter, perhaps unobservable, that one might prefer. None of the observational studies reviewed in the ISA are corrected for exposure measurement error. Exposure error should be considered when evaluating the conclusions to be drawn from these studies. The discussion should recognize that papers on the implications of exposure measurement error have made their case differently depending upon study design: through simulation studies alone (e.g., time series studies) or by theoretical developments that are then demonstrated using simulation studies (e.g., cohort studies). Thus the conclusions that can be drawn about measurement error properties in cohort studies are more generalizable than for time series studies. The concepts of Berkson and classical measurement error apply to both short- and long-term exposure metrics, and often both types of measurement error are present. When investigators use statistical models to predict exposures (e.g., land use regression), the exposure measurement error is no longer purely Berkson or classical but rather Berkson-like and classical-like. These distinctions should be made. The exposure measurement error discussion should conclude with the current (limited) state of knowledge about the impacts of measurement error on inference from epidemiologic studies.

There is an uneven level of detail across tables (i.e., the information presented in the tables versus the discussion of the tables in the text) that should be addressed. Models have uncertainty and model results are not true values, which should be acknowledged. More complex models do not necessarily reduce or eliminate uncertainty, which should also be acknowledged. Finally, correlations derived from data with many values below level of detection (LOD) should not be reported. This is misleading and should not be summarized in the document, even if it is reported in the peer-reviewed literature. It would be appropriate for the document to indicate that the literature exists and then to explain why the misleading correlations are not being reported.

The discussion of air exchange rates is too narrow and not well motivated or organized. There is no mention of databases of measured air exchange rates. The text could be more systematic in reviewing air exchange rates not just for “buildings” but for other enclosed microenvironments, such as vehicles. The issue of indoor reaction rate also merits a clearer and more accurate coverage and discussion. The state of science of stochastic population-based exposure modeling is not accurately portrayed and should be updated to account for recent improvements and applications of models such as the Stochastic Human

Exposure and Dose Simulation (SHEDS) and Air Pollutant Exposure (APEX) models, and improvements to their key input data. Similarly, discussion of the Consolidated Human Activity Database (CHAD) should be appropriately updated.

The uncertainty in the understanding of the appropriate exposure time scales (i.e., appropriate amount of exposure time-averaging) and how they correspond to various health outcomes should be clearer and the implications should be discussed. The controlled human exposure studies show health effects at the 5-minute time scale. Epidemiologic studies typically focus on different outcome measures and use exposures on the 24-hour average or 1-hour daily maximum time scales. In epidemiologic studies the critical time windows for effects are often not known; thus, it is difficult to tell whether exposure metrics are aligned with the optimum time scales of effects. These and other factors contribute to exposure measurement error, although one could argue that they fall outside the classical and Berkson classifications. Nonetheless, the implications of different averaging times and of possible misalignment of epidemiologic time-averaging with effect time scales should be introduced and discussed.
The discussion of co-pollutants should address some additional topics as co-pollutants can have implications for measurement error and its impact on inference from epidemiologic studies. Due to the changing nature of the pollutant mixture over time, the correlation of co-pollutants with $SO_2$ at various time scales could be quite different historically than currently. Historical exposures are more relevant to epidemiologic studies. Furthermore, correlations between co-pollutants can be higher when modeled exposures are used in the analysis, in part because the same input data are relevant for multiple pollutants. The revised ISA should more clearly explain and discuss the significance of confounders versus exposure modifiers as they pertain to the development and interpretation of exposure metrics used in epidemiological-based analysis or for scenario-based exposure modeling.

**Chapters 4 - Dosimetry and Modes of Action**

*Chapter 4 characterizes scientific evidence on the dosimetry and modes of action (MOA) for $SO_2$. Dosimetry and modes of action are bridged by the absorption and reaction of $SO_2$ in the epithelial lining fluid to form $SO_2$-derived products (e.g., sulfite and/or S-sulfonates) that are widely distributed throughout the body.*

*Are there topics that should be added or receive additional discussion? Similarly, are there topics for which discussion should be shortened or removed?*

This is a well-written chapter that generally meets the objectives of: (1) being a complete and accurate description that provides a good context for Chapter 5 on the health effects; (2) containing a good discussion and comparison of inhaled, ingested and endogenous pathways for $SO_2$/sulfite; and (3) providing modes of action for short-term and long-term respiratory effects as well as for extra-pulmonary effects.

Although the first paragraph of Section 4.2 notes that few studies have been done since the previous ISA, there are detailed discussions of studies that were done prior to this. It would be best to summarize results of these earlier studies to make the chapter more concise. Although this chapter is not intended to cover health effects, it contains some rather extensive discussions of this. The portions of these discussions that have a direct bearing on dosimetry or MOA should be retained. Other portions could be omitted.

*To what extent is the discussion of the chemistry of inhaled $SO_2$ and the processes of absorption, distribution, metabolism, and elimination accurate, complete, and relevant to the review of the $SO_2$ NAAQS?*

Section 4.2.2 on respiratory absorption would be much improved by a succinct but general discussion of the transport of any inhaled reactive gas. A conceptual description of how overall uptake and regional distribution of tissue dose depend on inhaled concentration, minute ventilation and gas properties is particularly needed. A definition of commonly-used terminology is also needed.

$SO_2$ metabolizing capabilities should be more completely discussed, including a comparison of the nose with other respiratory tract regions as well as with extra-pulmonary tissues. For example, it has been reported in studies with dogs that the activity of sulfite oxidase is comparable in the nose, trachea, and proximal and medium bronchi, but lower in lung parenchyma. As a start, see the paper by Maier et al. (1999) on xenobiotic-metabolizing enzymes in the canine respiratory tract for more specifics.

*Please comment on the discussion comparing endogenously generated and ingested sulfite with that derived from ambient inhalation.*

This portion of the chapter was well done, and no major changes are needed.

*To what extent are the discussion and integration of the potential modes of action underlying the health effects of exposure to sulfur oxides presented accurately and in sufficient detail? Are there additional modes of action that should be included in order to fully characterize the underlying mechanisms of sulfur oxides?*

Although a pathway for extra-pulmonary effects has been established for particles and ozone, there are insufficient data to conclude that there are such pathways for $SO_2$. Particularly because the $SO_2$ response is so rapid, within minutes, it is not reasonable to expect extra-pulmonary effects will have time to occur. Thus, an extra-pulmonary MOA could be deemphasized.

Extrapolating the robust literature on obesity and obstructive sleep apnea, it is very likely that functional nasal obstruction with increased oropharyngeal breathing will occur in obese and overweight individuals of all ages. This will likely be true for breathing at rest. A discussion of how this might affect pulmonary dose distribution and the MOA of $SO_2$ would be desirable.

**Chapter 5 - Integrated Health Effects of Exposure to Sulfur Oxides**

*Chapter 5 presents assessments of the health effects associated with short-term and long-term exposure to sulfur oxides. The discussion is organized by health effect category, exposure duration, outcome, and scientific discipline.*

*To what extent does this chapter accurately reflect the body of evidence from previous and recent epidemiologic, controlled human exposure and toxicological studies? What are the views of the panel on the integration of this evidence and the relative emphasis placed on each source of evidence?*

The authors are to be commended for a chapter that summarizes and distills a very large literature. That said, some of the epidemiological evidence presented in the chapter could be more accurately characterized. The precision of effect estimates needs to be made clear in a consistent manner. Currently, different sections of the chapter (e.g., respiratory and cardiovascular) treat precision of effect estimates differently.

Overall, the integration of the evidence from controlled human exposure, epidemiological, and toxicological studies is good and the relative emphasis placed on each source of evidence is reasonable. However, the chapter is lengthy. A suggestion about how to potentially shorten the chapter is to focus on the key studies that are relevant to the assessment of the evidence using the causal framework. Both the text and the tables could be revised in such an approach. If detailed descriptions of less relevant studies can be eliminated then more emphasis could be placed on a careful evaluation, synthesis and integration of the studies that provide the strongest evidence, including details that would allow a more rigorous assessment of study quality, especially for epidemiological studies.

*Considering the discussion of the strengths and limitations of the evidence in the text and tables within Chapter 5, to what extent is the causal framework appropriately applied to evidence for each of the health effect categories to form causal determinations?*

The CASAC concurs with the determination of a "causal relationship" between short-term $SO_2$ exposure and respiratory effects and the determination of a "suggestive but not sufficient to infer a causal relationship" between long-term $SO_2$ exposure and respiratory effects. The CASAC also concurs with the determination of "inadequate to infer the presence or absence of a causal relationship" between cardiovascular effects and long-term SO2 exposure and the determination of "suggestive but not sufficient to infer a causal relationship" between total mortality and short-term $SO_2$ exposure. The CASAC is not convinced the evidence has sufficiently changed since the 2008 ISA to warrant a change in the determinations (from "inadequate to infer the presence or absence of a causal relationship" to "suggestive but not sufficient to infer as causal relationship") for long-term $SO_2$ exposure and total mortality, reproductive/developmental effects, and cancer. The new epidemiological studies do not adequately address the potential confounding of the long-term exposure to the $SO_2$-outcome association (total mortality and cancer) and there is a lack of toxicological evidence to support biological plausibility (reproductive/developmental effects). Because of the potential for co-pollutant (especially $PM_{2.5}$) confounding of observed effects of both short-term and long-term exposures to $SO_2$ in epidemiological studies, there is still considerable uncertainty about the role of $SO_2$. Unless there are new studies that address this uncertainty, either experimental or epidemiological with adjustment for co-pollutant confounding, the CASAC recommends retaining the determination of an "inadequate to infer the presence or absence of a causal relationship" for these health effects and long-term $SO_2$ exposure. The CASAC does not concur with the change in the determination (from "inadequate to infer the presence or absence of a causal relationship" to "suggestive but not sufficient to infer as causal relationship") for short-term $SO_2$ exposure and cardiovascular effects due to potential confounding from co-pollutants. The $SO_2$-effects in studies of myocardial infarction hospital admissions were significantly reduced or eliminated in the two-pollutant model, and the one study that had temporally-resolved estimates of $SO_2$ exposures found no association between hourly ambient $SO_2$ concentrations and risk of myocardial infarction.

Measurement error in exposure assessment for epidemiological studies is treated superficially. The text leads readers to conclude that measurement error always leads to a bias toward the null, but bias away from the null can also occur, particularly in studies of long-term exposure. This should be stated clearly in the relevant sections. The revised treatment of measurement error suggested for Chapter 3 can be referenced in Chapter 5.

The chapter is inconsistent in how it discusses co-pollutant correlations. For example, in the respiratory section, $SO_2$ correlations with other criteria pollutants are repeatedly described as "low to moderate," yet in the cardiovascular section, the correlations are described as "moderate to high" with other pollutants. In general, the chapter needs more depth regarding the effects of multi-pollutant exposures. Co-exposure to other pollutants is currently treated as adding to uncertainty. Although this is an important issue, exposure to mixtures of pollutants in the "real world" also may have greater health impact than exposure to single pollutants in controlled human exposure studies.

*The conclusions in the draft $SO_X$ ISA regarding the respiratory effects of $SO_2$ exposure rely heavily on controlled human exposure evidence demonstrating effects of short-term peak exposures. Interpretation of the epidemiologic studies is more complicated due to the longer averaging time used in these studies. Please comment on the extent to which the evidence pertaining to the lowest concentrations associated with effects is appropriately characterized.*

The CASAC concurs with the reliance on the experimental human evidence where possible. The interpretation of the epidemiological evidence is more complicated. Many of the studies that found associations between $SO_2$ and asthma outcomes report mean concentrations and ranges of concentrations that are very low. It is a bit hard to reconcile the associations reported in these studies with the experimental evidence that indicates lack of bronchoconstriction in most subjects below 200 ppb and respiratory symptoms below 400 ppb. That said, the strongest effect of $SO_2$ in epidemiological studies of asthma exacerbations is for short lags and this finding is somewhat consistent with the controlled human exposure study evidence of bronchoconstriction with short duration exposures.

It is important to note that only mild-moderate asthmatic subjects who were both relatively young and healthy have been studied in controlled human exposure studies and thus the results of these studies may not be representative of the responses of more at-risk individuals. Obesity has not been considered as an effect modifier regarding short-term exposures and asthma exacerbations. Discussion of exposure-response function shapes (e.g., linearity) is overly confident given the very limited empirical exploration of this issue for most outcomes. There is also a need to define the magnitude of changes necessary for an outcome examined in human clinical studies to be considered an adverse outcome.

**Chapter 6 - Populations and Lifestages Potentially at Risk for Health Effects Related to Sulfur Dioxide Exposure**

*Chapter 6 evaluates scientific information and presents conclusions on factors that may modify exposure to $SO_2$, physiological responses to $SO_2$ exposure, or risk of health effects associated with $SO_2$ exposure. Consistent with previous ISAs for ozone, lead, and oxides of nitrogen, conclusions on these at-risk factors inform at-risk lifestages and populations.*

*To what extent has the available scientific evidence from epidemiologic, controlled human exposure, and toxicological studies been integrated to inform conclusions on at-risk populations and/or lifestages? Is there information available on other key at-risk factors that is not included in the draft $SO_X$ ISA and should be added?*

Chapter 6 presents a good introduction to the identification of populations at greater risk to adverse health consequences of $SO_2$ exposure. Overall the chapter does a good job of summarizing existing literature on the extent to which health effects are potentially modified by various factors. The EPA has made substantial progress in improving the content and presentation of the chapter on populations at risk. However, greater clarification is needed for several components of this chapter.

A few suggestions to further improve the clarity and informativeness of the chapter are below:

1. Populations can be "at risk" of higher $SO_2$-related effects because:
    a. They spend more time in areas with ambient concentrations that are harmful to health (e.g., based on where they live or time spent outdoors);
    b. They experience a greater internal dose when exposed to a given ambient concentration (e.g., because of ventilation rate or exercise activities) and hence suffer a greater adverse effect; and/or
    c. They have other factors that act synergistically with the air pollution exposure to enhance its adverse effect (e.g., age, sex, socioeconomic status, race, genetic predisposition, existing disease, behaviors like smoking or diet, weight).

    These three types of factors may also themselves cluster and interact. These categories are sometimes implicit in the discussion but are not always clearly articulated in the chapter. The EPA could consider whether these categories would be helpful in framing and structuring the chapter.

2. The tables are quite comprehensive and helpful. However further clarification of the terminology and the data presented in them is necessary for readers to fully understand the summaries. The EPA could further specify the criteria used to determine presence of effect modification (e.g., statistical significance of the interaction, magnitude of difference in effect across strata). The symbols in tables also need to be further clarified. For example, it is not clear what the "—" symbol means. Since the results shown are complex, additional clarification in the text may be helpful.

3. Consider expanding on descriptive information on populations at risk (e.g., expand Table 6-2 to include children, include descriptives of prevalence of outdoor activities for various subgroups). The numbers of "at risk" individuals in the tables at the beginning of the chapter would be much more informative if material from Chapters 3 and 5 were considered, as well as information on activity patterns from the CHAD and/or APEX databases.

4. There should be better integration with Chapter 5, particularly the discussion of effect modification.

**References**

Alarie, Y., Wakisaka, I., Oka, S. (1973). Sensory irritation by sulfate aerosols. *Environ. Physiol. Biochem.*, 3:182-184.

Amdur, M.O. (1971). Aerosols Formed by Oxidation of Sulfur Oxide. *Arch Environ Health*, 23:459-468.

Hanna, S.; Chang, J. (2012). Acceptance criteria for urban dispersion model evaluation. *Meterol and Atmos Physics*, 116(3):133-146.

Huang, M,; Carmichael, G.R.; Spak, S.N.; Adhikary, B.; Kulkarni, S.; Cheng, Y.; Wei, C.; Tang, Y.; D'Allura, A.; Wennberg, P.O.; Huey, G.L.; Dibb, J.E.; Jimenez, J.L.; Cubison, M.J.; Weinheimer, A.J.; Kaduwela, A.; Cai, C.; Wong, M.; Bradley Pierce, R; Al-Saadi, J.A.; Streets, D.G.; Zhang, Q. (2011). Multi-scale modeling study of the source contributions to near-surface ozone and sulfur oxide levels over California during the ARCTAS-CARB period. *Atmos. Chem. Phys.*, 11:3173-3194.

Maier, K. L., Wippermann, U., Leuschel, L., Josten, M., Pflugmacher, S., Schröder, P., Sandermann, H., Takenaka, S., Ziesenis, A., and Heyder, J. (1999). Xenobiotic-metabolizing enzymes in the canine respiratory tract. *Inhalation toxicology*, 11(1):19-35.

Slade, D.H. (1968). Meteorology and Atomic Energy, U.S. Atomic Energy Commission (No. TID-24190). *U.S. Atomic Energy Commission*, 445.

Turner, D.B. (1970). Workbook of Atmospheric Dispersion Estimates (No. 999-AP-26). *U.S. Environmental Protection Agency.*

U.S. EPA (2008). Integrated Science Assessment for Sulfur Oxides: Health Criteria (EPA/600/R-08/047F). Research Triangle Park, NC: U.S. Environmental Protection Agency, National Center for Environmental Assessment.

# Appendix A

**Individual Comments by CASAC Sulfur Oxides Panel Members on EPA's Integrated Science Assessment for Sulfur Oxides – Health Criteria (External Review Draft – November 2015)**

# Mr. George A. Allen

## General Comments

Over the last several NAAQS review cycles, the structure of the ISA has become more and more complex, at least in part in response to CASAC review requests. We now have a Preamble (generic, for all ISAs), a Preface, an Executive Summary, and a Summary chapter – all before the actual core ISA chapters. Since to some extent (?) they all need to stand on their own, reading them in sequence is painfully redundant. While I don't have explicit suggestions as to how simplify or combine these documents, this may be worth considering.

HERO continues to be a very useful tool for reviewing these NAAQS process documents, and EPA should be encouraged to maintain it.

## Executive Summary

*The Executive Summary is intended to provide a concise synopsis of the key findings and conclusions of the SOX ISA for a broad range of audiences. Please comment on the clarity with which the Executive Summary communicates the key information from the SOX ISA. Please provide recommendations on information that should be added or information that should be left for discussion in the subsequent chapters of the SOX ISA.*

Overall, the Executive Summary clearly communicates the key findings of this ISA. Given the presumably non-technical audience of executives for this section, some of the terms used tend to be fairly technical for a general audience. Sometimes these terms are defined, but sometimes not. One example is on page xli, lines 23-24: "Endogenous sulfite from the catabolism of ingested sulfur-containing amino acids far exceeds exogenous sulfite..." – this may be too technical for this section, and more appropriate for Chapter 1?

Table ES-1 is a concise summary of SO2 exposure and health effects and changes in the causal determination status since the last ISA. Footnote B notes changes to the phrasing of causal determinations. Five of these categories have been moved up one level; are any of these changes in part because of changes in the phrasing?

The "Policy-Relevant Considerations for Health Effects Associated with Sulfur Dioxide Exposure" at the end of the Executive Summary is very well written. This is the level and style that may serve as a model for other sections of this summary.

## Specific Comments on the Executive Summary

Page xxxvii: The first page of the Executive Summary is heavily footnoted. Footnote 2, and maybe some others important to understanding the context of this ISA, could be promoted into the text, since it explains what this ISA is actually covering – essentially only SO2 (…"only SO2 is present in the atmosphere at relevant concentrations."). It would help if this could be more clearly explained in the first

paragraph – defining the backstory of how we get from SOX (a wide range of oxidized gaseous and particle sulfur species) to SO2 (the NAAQS indicator); some of this is done in the first paragraph of Chapter 1. While the criteria pollutants Pb and CO are unambiguous, others are less so. Where does the term SOX (and NOX) come from in a regulatory sense – the CAA?

As noted previously, the ISA needs to be very consistent throughout in using terms like SOX and SO2, using SOX only where essential – presumably only when setting the ISA stage.

Footnote 4, about blue hyperlinks, seems out of place in this paragraph which has no blue hyperlinks.

Page xxxviii, line 20: "joint exposure to other pollutants". Elsewhere in this section, the term "copollutants" is used for this – it would help to define that here. Do these terms refer only to NAAQS indicators?

Page xxxix, lines 21-23: "Multiple evaluations of AERMOD's performance against field study databases over averaging times from 1 hour to 1 year have indicated that the model is relatively unbiased in estimating upper-percentile 1-hour concentration values." Modeling the 1-h form of the current SO2 NAAQS is much more challenging than modeling the old form of daily or annual SO2 concentrations. From a NAAQS compliance point of view, AERMOD is generally considered to provide a conservative result – e.g., generally over-predicting measured SO2 at the surface, even if this outcome is primarily due to constraints on actual monitoring locations.

Page xxxix, lines 31-33: central site monitors and "…the relatively high spatial variability in SO2 across an urban area." This statement needs clarification here and on page 3-2. It is correct for urban areas that have large sources of SO2 in or near them. However, as SO2 point sources are cleaned up, many urban areas have very low and relatively uniform levels of SO2.

**Chapter 1**

*Chapter 1 summarizes key information from the Preamble about the process for developing an ISA. Chapter 1 also presents the integrative summary and conclusions from the subsequent detailed chapters of the SOX ISA and characterizes available scientific information on policy-relevant issues. Please comment on the usefulness and effectiveness of the summary presentation. Please provide recommendations on approaches that may improve the communication of key findings to varied audiences and the synthesis of available information across subject areas. What information should be added or is more appropriate to leave for discussion in the subsequent detailed chapters?*
This chapter effectively summarizes the key findings of subsequent chapters, and reads better than the Executive Summary. A few specific comments follow.

Page 1-1, first paragraph: SO2 as the NAAQS indicator for gaseous SOX is just dropped into the narrative. A brief rational and/or background for using SO2 would be helpful.

Page 1-10, lines 27-29: "…SO2 correlations with NO2 and CO were observed to be high, it is possible the data may have been collected before recent rulemaking to reduce sulfur content in diesel fuel…". This fuel cleanup is an important point, and is part of why (along with very few large SO2 point sources) urban SO2 is now usually very low and effectively very homogeneous from an exposure / health effects

perspective. Some of these fuel S reductions are still ongoing; New York City started to prohibit 3000 PPM S #6 heating oil only a year ago (#4 is still allowed), and it will be several more years before all NYC heating oil will be less than 30 ppm S.

Section 1.4.2 (Assessment of Human Exposure), pages 1-8 to 1-11: This section covers exposure issues in great detail. But even more so than for NO2, the key drivers behind the NAAQS form and level come from controlled human exposure studies as noted in Section 1.8, conclusions (page 1-28). Health outcomes from longer or chronic SO2 exposures come from epidemiological studies, but it is the need to control acute exposures (minutes to 1 hour) that also reduces the more chronic exposures. It may be useful to have a brief discussion of this elsewhere in Chapter 1 (1.6.1 perhaps) since it is prominent in the Chapter 1 conclusion paragraph (1.8).

Table 1-1 (pages 1-21 to 1-23) is an appropriately detailed summary of causal determinations. The different sections of the column on the right listing SO2 concentrations associated with effects can be confusing, and would benefit from additional formatting to separate subsections; see the first part of page 1-21 for an example.

Number of monitors reporting 5-minute data. This is inconsistent across the document. Page 1-7 and 2-91 say “more than 400”. Page 2-29 has 195 monitors. Table 2-6 and page 2-62 text has 309 monitors “to date”. Some of these differences may be due to the increasing number over time because of the monitoring requirements in the June 2010 revision to the SO2 NAAQS.

**Chapter 2**

Section 2.6.1, Dispersion Modeling.

AERMOD performance is generally considered acceptable if within a factor of two of actual measured SO2 metrics such as 99$^{th}$ percentile design values. This is certainly an issue in a compliance context, and possible an issue if modeled SO2 exposures are used in acute effect health studies. A robust assessment of AERMOD performance is in Frost (2014), “AERMOD performance evaluation for three coal-fired electrical generating units in Southwest Indiana”
http://www.tandfonline.com/doi/abs/10.1080/10962247.2013.858651
This reference is not in the current draft of the ISA.

There is minimal mention and no discussion of “puff” lagrangian models in this section, and the potential for improvement in modeling performance relative to AERMOD. The CASAC review of the SO2 IRP clearly recommended that this be addressed (page 2), along with a discussion of the SO2 “Data Requirements Rule” that allows modeling to be used for NAAQS compliance purposes. Under this rule, modeling effectively becomes part of the SO2 compliance monitoring network, and thus it is important to include it in any discussion of the network since modeling could replace existing or newly required monitoring sites.

**Chapter 3**

3.2.1.1 Section 2.4.1 discussed the sensitivity of the commonly used UVF SO2 measurement method, with detection limits of 2 to 0.2 ppb. This section says: “Ultraviolet fluorescence (UVF) detection of SO2 has a high detection limit relative to ambient levels” (page 3-2, lines 19-20). While this may be true for ambient SO2 levels often observed in areas without large SO2 sources, it is not relevant for concentrations of SO2 that are of concern for health effects.

3.2.1.2 The “personal HEADS” sampler runs at 4 lpm, not 10 (10 is the flow for the full-size fixed-site version of HEADS). The discussion of personal monitoring / Citizen Science “sensors” includes some citations that are not relevant, including one with a low-end detection limit of 15 PPM (not PPB).

# Dr. John R. Balmes

**Comments on Chapter 5**

*Chapter 5 presents assessments of the health effects associated with short-term and long-term exposure to sulfur oxides. The discussion is organized by health effect category, exposure duration, outcome, and scientific discipline.*

*a. To what extent does this chapter accurately reflect the body of evidence from previous and recent epidemiologic, controlled human exposure and toxicological studies? What are the views of the panel on the integration of this evidence and the relative emphasis placed on each source of evidence?*

In general, the chapter characterizes the body of evidence on SO2 health effects in a fairly accurate way. That said, some of the epidemiological evidence presented in the chapter could be more accurately characterized. I do not think it is appropriate to present associations between SO2 exposure estimates and outcomes per 5 or 10 ppb increase in SO2 when the association is not statistically significant without noting the lack of significance. The sections on cardiovascular outcomes are better at representing nonsignificant associations with wording such as follows: "the 95% CI for the association was wide, indicating an imprecise association." I would prefer that the chapter state something to the effect that "while the analysis between SO2 and asthma incidence suggested an association, the strength of the association did not reach statistical significance. That said, there was an apparent increase in the OR for incident asthma of X% for a 5 ppb increase in SO2."

Overall, the integration of the evidence from controlled human exposure, epidemiological, and toxicological studies is good and the relative evidence placed on each source of evidence is reasonable.

*b. Considering the discussion of the strengths and limitations of the evidence in the text and tables within Chapter 5, to what extent is the causal framework appropriately applied to evidence for each of the health effect categories to form causal determinations?*

I agree with the chapter's assessment of the strength of the evidence regarding causal determination for both respiratory effects and both short and long-term exposure. I also agree that the evidence for cardiovascular effects of long-term exposure is inadequate. Furthermore, I agree that the evidence for the association of total mortality with long-term exposure is suggestive. In do not think that the evidence for cardiovascular effects of short-term exposure, total mortality and short-term exposure, and cancer and long-term exposure has changed substantively since the 2008 ISA was published.

*c. The conclusions in the draft SOX ISA regarding the respiratory effects of SO2 exposure mrely heavily on controlled human exposure evidence demonstrating effects of short-term peak exposures. Interpretation of the epidemiologic studies is more complicated due to the longer averaging time used in these studies. Please comment on the extent to which the evidence pertaining to the lowest concentrations associated with effects is appropriately characterized.*
I agree with draft ISA's reliance on the experimental human evidence where possible. I also agree that interpretation of the epidemiological evidence is more complicated. Many of the studies that found associations between SO2 and asthma outcomes report mean concentrations and ranges of concentrations that are very low. I find it hard to reconcile the associations reported in these studies with

the experimental evidence that indicates lack of bronchoconstriction in most subjects below 200 ppb and respiratory symptoms mostly above 400 ppb. That the strongest effect of SO2 in epidemiological studies of asthma exacerbations is for short lags is somewhat consistent with the controlled human exposure study evidence of bronchoconstriction with short duration exposures.

**Specific Comments**

p. 5-6, line 6 AHR should be spelled out with first use.

p. 5-6, lines 17-18 It should be noted here that this lack of studies in children is due to ethical considerations.

p. 5-7, line 14 Should be "…that is…" Subset is the subject and is singular.

p. 5-13, line 10 The Balmes et al. study used eucapneic hypernea rather than exercise.

p. 5-20, line 33 The specific lung changes need to be specified. The percent changes are not interpretable without this specification.

p. 5-21, lines 5 and 8 Should be "45-minute" and there is an important comma to be inserted to make the sentence clear: "… air followed by SO2, and O3 followed by O3…".

p. 5-21, line 25 I find the characterization of recent studies as consistent with the 2008 ISA, "finding some positive associations" as a bit disingenuous. The data as a whole are inconsistent as is stated in the next sentence so yes there are some positive associations, but there are also some studies that found no associations. I think it is better to characterize the data as inconsistent.

p. 5-51, line 22 Should be "…effects… were not assessed…".

p. 5-70, lines 31-32 Nasal discharge/congestion and upper respiratory symptoms are NOT clinical symptoms "demonstrating exacerbations of COPD".

p. 5-71, line 5 Should add "or lung function" to the end of this sentence.

p. 5-71, lines 24-25 Because nasal discharge/congestion and upper respiratory symptoms are not symptoms of COPD, I would delete "COPD" from line 24.

p. 5-78, lines 7-10 1.6% vs. 0.8% is hardly a difference. The earlier Wong et al. study had a wider CI.

p. 5-80, line 21 Should be "A limited number of studies has examined…".

p. 102, lines 33-34 Changes in FEF25-75 and FEV1 are not indicative of "a constriction in the upper airways". The sentence should simply state that the SO2 exposures "induced decreases in FEF25-75 and to a lesser extent FEV1."

p. 5-142, lines 2-3 This sentence somewhat mischaracterizes the results reported by this study, of which I am one of the senior authors. The combined association between SO2 exposure during the first 3 years

of life and asthma incidence pooling data from all four sites was not statistically significant (OR 1.03 [0.94-1.13). I think it is somewhat misleading to suggest that we found an association for SO2 exposure during the first 3 years of life and I don't understand the values given in the ISA text [OR = .16 (0.73−1.84) per 5 ppb SO2]. The only statistically significant association for SO2 that we reported was for Houston, but given the lack of association in the pooled data, I don't think a concentration-response function should be included here. That said, it is remarkable that the pooled effect estimate (OR=1.03) is the same as that of the Clark et al. study in Vancouver.

p. 5-239, lines 12-13 Stating here that "Primary pollutants such as NO2 and CO typically show moderate to high correlations with SO2" is in contrast to multiple other place in the ISA text where these correlation with other criteria pollutants are described as "low to moderate".

p. 5-274, line 15 Again, it is stated here that "SO2 is low to moderately correlated with other pollutants".

# Dr. James Boylan

**Comments on Chapter 2**

*Chapter 2 describes scientific information on sources, atmospheric chemistry, and measurement and modeling of ambient concentrations of gaseous sulfur oxides.*

*a. To what extent is the information presented regarding sources, chemistry, and measurement and modeling of ambient concentrations accurate, complete, and relevant to the review of the $SO_2$ NAAQS?*

Page 2-3: Figure 2-1 needs to add the label for the y-axis. Also, many of the "Industrial" source category names have been cut-off which makes it difficult to distinguish them from each other.

Figures 2-2, 2-3, 2-4: The text on page 2-3 states "Counties that have not fully achieved the current air quality standard for $SO_2$, being partially or entirely out of attainment, are outlined in yellow in Figures 2-2, 2-3 and 2-4." However, the legend in the figures describes these areas as "Counties with ambient $SO_2$ 1-hr concentrations that exceed 75 ppb" or "Counties with $SO_2$ concentrations above the NAAQS". The two are not identical because areas can exceed the NAAQS, but not be designated nonattainment. Also, the author should add which 3-year design value was used to determine if the county was nonattainment. Next, the author should explain why (A) industrial cement production and (B) industrial chemical and allied products manufacturing are included in Figure 2-3(A) and Figure 2-3(B) when it looks there are many other source categories with higher $SO_2$ emissions (see Figure 2-1).

Page 2-9: Figure 2-5 should include EPA $SO_2$ projections to 2017 or 2018. These are readily available as part of the modeling EPA did for the recently proposed Cross-State Air Pollution Rule (CSAPR).

Section 2.2: The author should include a discussion on the Data Requirements Rule for the 1-hour $SO_2$ NAAQS since this rule will identify the sources that need to characterize air quality through ambient monitoring or through air quality modeling. States were required to submit a list to EPA by January 15, 2016, that identifies all sources within its jurisdiction that have $SO_2$ emissions that exceeded the 2,000 tons per year (tpy) annual threshold during 2014. A summary of the types of sources included in the list and a map showing the location of each source would be very helpful in understanding the large sources of $SO_2$ that still remain in the U.S.

*b. Please comment on the extent to which available information on the spatial and temporal trends of ambient $SO_2$ concentrations at various scales has been adequately and accurately described. In particular, what is the extent to which the analyses of recently available 5-min $SO_2$ concentration data are informative in considering relationships between 5-min and 1-hr $SO_2$ concentrations?*

Pages 2-24 to 2-84: Section 2.5 "Environmental Concentrations" is based on 2010-2012 data. However, 2013 has been available since May 1, 2014; 2014 data has been available since May 1, 2015; and 2015 will be available on May 1, 2016 (some of the 2015 data is already available). By including this additional data, the number of data points will nearly double adding to the robustness of the analysis.

Pages 2-41 to 2-46: Figures 2-14, 2-15, 2-16, 2-17, 2-18, and 2-19 need to include the $SO_2$ emissions for each NEI Facility. This will help explain the spatial variability in the $SO_2$ concentrations at the monitors.

Page 2-68: In Figure 2-28, many sites with high $SO_2$ concentrations are classified as "Moderate Conc Sites" (e.g., Cleveland, Pittsburgh, and Payson/Phoenix) while other sites with much lower $SO_2$ concentrations are classified as "High Conc Sites" (e.g., New York City and St. Louis). It seems that the method for identifying the classification needs to be updated. In addition to only looking at $SO_2$ data from these six cities, the rest of the $SO_2$ data pairs should be analyses to support the author's conclusion on page 2-69 that "These results emphasize that 1-hour average concentrations at or below 75 ppb, for the most part, represent hourly 5-minute maximum values below 200 ppb." Specifically, the author should create individual scatter plots for every monitor with data and summarize the results with a single scatter plot with all the $SO_2$ data pairs from all the $SO_2$ monitors in the country. Next, there are a number of data points that are below the 1:1 line in Cleveland and St. Louis. By definition, the 5-minute maximum value can never be lower than the 1-hour average value if the 5-minute maximum value is extracted from the same hour of data. In general, it is inappropriate to compare a 5-minute maximum value to a 1-hour average value if the 5-minute maximum value is extracted from a different hour. However, since the 2010 1-hour $SO_2$ NAAQS uses the daily maximum 1-hour concentration to represent to the daily maximum 5-min concentration, it would also be useful to see the comparison of the daily maximum 1-hour concentration vs. the daily maximum 5-min concentration in addition to the comparison for the same hour of data. Finally, the author does not include any statistical analysis of the results and does not include much of a conclusions on the peak to mean ratios besides "These results emphasize that 1-hour average concentrations at or below 75 ppb, for the most part, represent hourly 5-minute maximum values below 200 ppb." Additional statistical analysis of the data must be performed with the full data set (all monitors and all years) to build a more compelling conclusion on the peak to mean ratios.

Page 2-71 to 2-73: The time series plots in Figure 2-29, 2-30, and 2-31 should replace the red circles with a time series line so that the lower values can be distinguished from each other.

*c. How informative is the analysis of correlations between $SO_2$ and co-occurring pollutant concentrations for interpretation of epidemiologic studies?*

Page 2-82 to 2-83: The author states "The lower wintertime $SO_2$-$O_3$ correlation could be directly linked to relatively low $O_3$ concentrations during this time of year due to less photochemical $O_3$ production. At such low ambient levels that are presumably near the instrument detection limit, $O_3$ measurements may be subject to substantial measurement error, which may lead to poor correlations between O3 and other pollutants, including $SO_2$." I agree with the first sentence. However, I don't agree with the second sentence since no evidence has been presented to support this conclusion. In many areas of the country, winter ozone levels are well above instrument detection limit and hover around 30-40 ppb. A reason for the low $SO_2$-$O_3$ correlation in all seasons is likely due to the regional nature of ozone formation vs. the local nature of $SO_2$ plume hits. In the wintertime, there is higher $SO_2$ and less ozone due to the less oxidation of $SO_2$ and less formation of ozone due to decreased photochemistry.

Page 2-84: The author states "However, given that a small subset of monitors report relatively strong copollutant correlations, confounding may need to be considered on a study-by-study basis." The author should examine the handful of data point that have a high correlation coefficient to help identify the reason (e.g., power plant plume hits). Then, these situations can be accounted for in exposure studies.

Comments on Section 2.6 – Atmospheric Modeling

Page 2-87: The author states "One limitation of the Gaussian approach is the assumption of steady conditions over a 1-hour modeling period and over the plume transport distance to the receptors." This is especially true for low winds which is when the $SO_2$ concentrations are the highest. For this reason, a Lagrangian model such as CALPUFF or SCICHEM is preferred.

Page 2-89: The author presents AERMOD model performance for multiple field-study databases. In these examples, the emissions characteristics and background concentrations were well known; meteorological data were available on site; and tracer concentrations were measured at multiple locations where high plume impacts were expected. In these cases, model was generally shown to be acceptable for comparison to air quality standards. However, there are many application of AERMOD where the emissions are not know, the background concentrations are not known, and the meteorological data is not on site (typically from the nearest airport that can a considerable distance away from the emission source). In these cases, the model performance is usually very poor or not evaluated at all which can lead to large modeled uncertainties. Therefore, model performance evaluations (MPEs) should be performed with monitoring data for each model application. The model performance statistics should be considered when weighing the credibility of the results and the resulting exposure impacts. Where no MPE was performed or no monitors exist, the model results may be significantly biased and should be given less weight in risk and exposure assessments. Finally, the type of model performance evaluation that was performed should be understood. Regulatory models need to predict the peak of the concentration distribution, unpaired in time and space, for comparison to air quality standards. However, models used for risk and exposure assessments, require skill at predicting concentration distributions paired in time and space. Typically, EPA's dispersion modeling results are summarized in Q-Q plots of highest ranked modeled vs. highest ranked observed concentrations, unpaired in time and space. This approach can make modeling results that are not skilled at predicting concentration distributions paired in time and space look good. Below is an example (Brode, 2008) of model vs. observed concentrations using a Q-Q (unpaired) plot (Figure 1) and a scatter (paired) plot (Figure 2). Model applications used for risk and exposure assessments should not use Q-Q plots to evaluate model performance. Rather, they must use scatter plots and statistics that are paired in time and space to draw proper conclusions on risk and exposure.

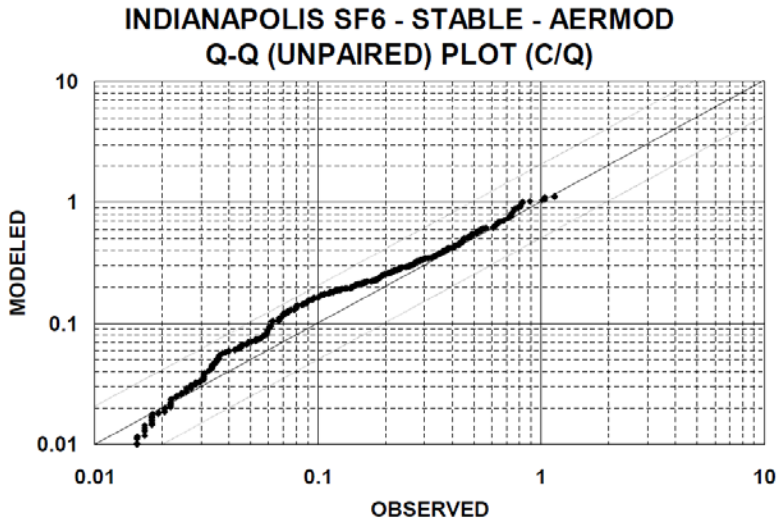



**Figure 1.** Typical AERMOD evaluation using a Q-Q (unpaired) plot.

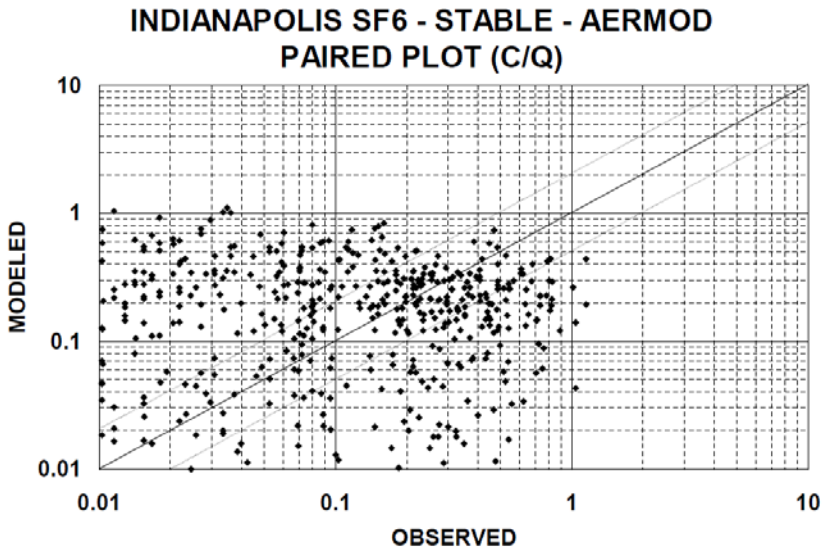



**Figure 2.** Previous example showing AERMOD results using a scatter (paired in time and space) plot.

Reference

Brode, R. W. (2008), AERMOD Evaluation for Non-Guideline Applications, 9th Conference on Air Quality Modeling, Research Triangle Park, NC; October 9, 2008.

# Dr. Judith Chow

Overall, Chapter 2 is informative and well-written. The responses to the three charge questions below point to areas that need additional verification with references provided at the end.

a. *To what extent is the information presented regarding sources, chemistry, and measurement and modeling of ambient concentrations accurate, complete, and relevant to the review of the $SO_2$ NAAQS?*

Section 2.2 Sources of Sulfur Dioxide (Page 2-1)

Gaseous $SO_x$ needs to be better defined. In addition to $S_2O$, $SO_3$, and $H_2SO_4$ (Page 2-1, Lines 7-8), sulfur monoxide (SO; later acknowledged on Page 2-18, Line 1), sulfur dioxide ($SO_2$), thiosulfate ($S_2O_3$), and heptoxide ($S_2O_7$) should be included. As atmospheric sulfur budgets are mainly composed of $SO_2$ and sulfate ($SO_4^=$), both gas and particle phases should be acknowledged. Even though most of the adverse environmental and health effects of $SO_4^=$ should be addressed in the PM ISA, information on the $SO_2/SO_4^=$ relationship can be used to examine the influences and correlations with co-pollutants.

Source categories in Figure 2-1 (Page 2-3) need to be clarified (labels should be added to both axes). Several source subcategories are not shown on the x-axis. What is the difference between the "Industrial Other" and "Industrial" subcategories? Does "Boilers-Oil" represent fuel-oil or residual oil combustion (one would expect higher $SO_2$ emissions from residual oil combustion)? What source category covers smelter emissions? Figure 2-1 emphasizes the importance of coal combustion, especially the electrical generating unit (EGU) fraction, as the major source of U.S. $SO_2$ emissions. Since many of these source subcategory emissions are low compared to EGU, it may be more informative to replace this figure with a table containing emission rates of $SO_2$ and co-pollutants (e.g., $NO_x$, CO, and PM) from major sources based on either the 2011 or 2014 National Emission Inventory (NEI; U.S.EPA, 2013b).

The EPA handout shows three Hilo, Hawaii, sites with 99th percentile daily maximum 1-hr $SO_2$ concentrations > 600 ppb for the period of 2010-2012. However, volcanoes are only acknowledged in Section 2.2.5.2 without an emission rate estimate.

Do the yellow-outlined areas in Figures 2-2 and 2-3 (Pages 2-4 to 2-6) represent counties with a single 1-hr $SO_2$ concentration exceeding 75 ppb or the $99^{th}$ percentile daily maximum 1-hr $SO_2$ concentrations and for what time periods? Emphasis is given to cement production as well as industrial chemical and allied products manufacturing, without acknowledging the smelter sources. Emissions from these source subcategories should be isolated in Figure 2-1 that corresponds to Figure 2-3. A map showing locations and sizes of commercial marine vessel emissions, which is the third largest emission category in Figure 2-1, would be helpful (e.g., Figure 2 of Wang et al., 2007).

The definition of "major emitting facility" needs to be consistent. It is defined as >100 tons/yr on Page 2-7 (Lines 4-5) but as facilities with >1,000 tons/yr $SO_2$ emissions in Figure 2-4 (Page 2-8). Janoe and Craft (2015) note a 2,000 tons/year threshold for the 2013 attainment designations. For the consent decree, the limits are 2,600 tons/yr $SO_2$ with an annual average emission rate of at least 0.45 lb per million BTUs or >16,000 tons/yr for a single plant. Emission rate categories bracketing these ranges in Figure 2-4 would be informative.

In Figure 2-5 (Page 2-9) it would also be useful to add years when the regulations for on-road (2006) and off-road (2007) diesel sulfur fuel standards were in place and when the North American Emission Control Areas for shipping were implemented (DieselNet, 2014).

Section 2.3 Atmospheric Chemistry and Fate (Page 2-18)

Figure 2-10 (Page 2-23) should reference Figure 7.19 of Seinfeld and Pandis (2006), as the reference cited (Seinfeld and Pandis, 1998) is out of print. What are losses due to dry and wet deposition? Dry deposition of $SO_2$ is much larger than that of $SO_4^=$ (Clarke et al., 1997; Shadwick and Sickles, 2004; Sickles and Shadwick, 2007), while aqueous $SO_4^=$ may be removed through occult deposition of large fog or cloud droplets (Dollard et al., 1983; Lillis et al., 1999; Pandis and Seinfeld, 1989).

Section 2.4 Measurement Methods (Page 2-24)

A list of available Federal Reference Methods (FRMs), Federal Equivalence Methods (FEMs), and related citations (e.g., Blacker et al., 1973; Kok et al., 1989; Luke, 1997; Stelson and Bao, 1988; Trieff et al., 1968; West and Gaeke, 1956; Wright et al., 1989) should be given. A summary Table of $SO_2$ measurement comparison studies (e.g., Ferek et al., 1997; Kim and Kim, 2001; Leppanen et al., 2005; Luke, 1997; Medina et al., 2011; Terraglio and Manganelli, 1962; Vadjic et al., 1992) should be added that specifies operating principles, averaging time, accuracy, precision, lower detection limits (LDL), interferences, etc. Discuss potential reasons for discrepancies between FRM/FEM and other measurement methods. This would be important for determining the equivalence between standard monitors (1-hr averages with LDL <2 ppb) and trace monitors (5-min averages with LDL <0.2 ppb) as well as within a given instrument type. It would also be helpful to show a time series of these comparisons and verify similarities and differences by location, season, and time of day.

The number of sites with a collocated standard monitor (1-hr) and trace monitor (5-min) should also be specified. Different Air Quality System (AQS) monitor IDs are given in Table 2-8 (Pages 2-47 and 2-48), but not much information is provided regarding the precision and comparability between collocated 1-hr (standard monitor) and 5-min (trace monitor) $SO_2$ concentrations. Only six out of the 42 sites in the six core-based statistical area (CBSA)/metropolitan areas reported 5-min $SO_2$ measurements.

The discussion of instrument biases and uncertainties regarding central site monitoring with the ultraviolet fluorescence (UVF) method in Section 3.2.1.1 (Page 3-2, Lines 17-27) needs to be combined with the discussion in Section 2.4.2 (Page 2-26) regarding positive and negative interferences.

The performance specifications in Table 2-4 (Page 2-25) list an LDL of 2 ppb, zero drift of 4 ppb, and interference of 5 ppb for FRM $SO_2$. These levels seem high given the low ambient 1-hr average $SO_2$ concentrations of 4 ppb (Page 2-35, Lines 8-14) or 1-3 ppb for the $50^{th}$ percentile national statistics during 2010-2012 (Table 2-7, Page 2-37). A comparison between continuous $SO_2$ measurements and weekly averaged Clean Air Status and Trends Network (CASTNET) filter pack or other commercially integrated samplers should be provided to give some perspective on performance for low (<5 ppb) ambient $SO_2$ concentrations and their comparability for annual averages (Bennett et al., 1994; Bytnerowicz et al., 2002; Chow et al., 1993; Sickles et al., 1999; Sickles and Shadwick, 2002). As there are ~90 CASTNET sites that measure $SO_2$, it would be good to add the CASTNET monitors to the AQS monitors shown in Figure 2-11 (Page 2-30).

Other $SO_2$ measurement methods (Section 2.4.3) might include Fourier transform infrared spectroscopy (FTIR) (Horrocks et al., 2001) and microsensors (Do and Chen, 2007; Dye et al., 2014; Ohira and Toda, 2005; Roberts et al., 2014; Toda et al., 1998; U.S.EPA, 2013a). Personal Monitoring Techniques in Section 3.2.1.2 (Page 3-2) belong in "Other Sulfur Dioxide Measurements" under Section 2.4.3 (Page 2-27).

Less costly and more portable instruments for measuring criteria pollutants are of growing interest for community involvement and exposure assessment (Dye et al., 2014; Snyder et al., 2013; Wang and Brauer, 2014). The discussion on these small, low-cost sensors in Section 3.2.1.2 (Personal Monitoring Techniques) should be treated as a stand-alone method in Section 2.4.3 (Other Sulfur Dioxide Measurements) since it has been used for source characterization and ambient monitoring.

As there are a variety of solid-state options now available to replace the photomultiplier tube (PMT), the "PMT" (Page 2-24, Line 28) might be changed to "detector." Section 2.4.4.2 (starting on Page 2-32) might mention the importance of using the proper material for seasoning the sampling probes (Mamane and Donagi, 1976; Wohlers et al., 1977).

Section 2.5 Environmental Concentrations (2-34)

The Executive Summary (Page xxxvii, Lines 15-18) says that EPA requires states to report either the highest 5-min centration for each hour of the day or all twelve 5-min concentrations for each hour of the day. This is inconsistent with Chapter 1(Page 1-7, Lines 18-19), which says that "new $SO_2$ monitoring guidelines require states to report 5-min data." As the UVF FRM instruments are capable of producing short-duration averages, consistent reporting of each 5-min average by the states is preferred. This allows for the examination of consecutive elevated 5-min $SO_2$ concentrations in order to evaluate the duration of plume touchdown/downwind mixing and to better understand the exposure durations and patterns, especially when $SO_2$ concentrations exceed the 200 ppb health benchmark. Requiring states to report each 5-min average also provides a database to evaluate a future 5-min $SO_2$ NAAQS indicator.

Table 2-7 should specify how the mean (arithmetic average?) was calculated when many measurements were below LDLs. The method for treating values below LDL in annual average calculations needs to be specified (Kushner, 1976). Does Figure 2-13 (Page 2-39), the map of the $99^{th}$ percentile of daily 24-hr average $SO_2$ concentrations reported at AQS monitoring sites during 2010−2012, include both 5-min and 1-hr $SO_2$? Figures 2-12 and 2-13 (Page 2-38 and 2-39) might include a category of 75-140 ppb, since 140 ppb was the 1971 24-hr $SO_2$ NAAQS limit. Why are <13 ppb and >108 ppb used to limit the categories in Figure 2-13? It seems like rounder numbers such as 15 and 110 ppb would be more logical.

Section 2.6 Atmospheric Modeling (Page 2-85)

The purpose of applying dispersion modeling to estimate $SO_2$ concentrations needs to be clarified. The Executive Summary (Page xxxix, Lines 11-12) notes that "dispersion models can be used to estimate $SO_2$ concentrations in locations where monitoring is not practical or sufficient and that "modeling is critical to the assessment of the impact of future sources or proposed modifications where monitoring cannot inform, and for the design and implementation of mitigation techniques" (Page xxxix, Lines 16-18). Will modeling efforts focus on: 1) location where monitoring is not practical or sufficient; 2) areas with large point sources (e.g., areas >2,000 tons/yr or with major coal-fuel power plant with >16,000 tons/year, Hamel, 2015); or 3) areas where $SO_2$ levels exceed the NAAQS? Including examples of

CMAQ or AERMOD derived 99$^{th}$ percentile daily maximum 1-hr $SO_2$ concentrations compared to measured ambient $SO_2$ concentrations would illustrate how well modeling results can be used to determine attainment.

Can a dispersion model represent the zones defined in Figures 2-14 to 2-19 (Pages 2-41 to 2-46)? Side-by-side maps denoting the comparison between modeled and measured ambient $SO_2$ concentrations for selected regions (e.g., Huang et al., 2011) would shed light on the representativeness of dispersion modeling to catch plume touchdown and the adequacy of the number of monitor sites within 15 km of major $SO_2$ sources.

A discussion of some of the AERMOD options and comparison with $SO_2$ and inert tracer data would be informative about how well the modeled distributions represent the measured concentrations (Frost, 2014; Guerra et al., 2014; Hanna et al., 2001; Irwin, 2014; Isakov et al., 2007; Paine et al., 2015; Rehbein et al., 2014; Rood, 2014).

*b. Please comment on the extent to which available information on the spatial and temporal trends of ambient $SO_2$ concentrations at various scales has been adequately and accurately described. In particular, what is the extent to which the analyses of recent available 5-min $SO_2$ concentration data are informative in considering relationships between 5-min and 1-hr $SO_2$ concentrations?*

Section 2.5.2 Spatial Variability (Page 2-35)

The number of $SO_2$ sites should be clarified. Chapter 1 states that there are more than 400 monitoring sites across the U.S. reporting 5-min data (Page 1-7, Lines 21-22); whereas the introduction in Section 2.4.4 (Page 2-27, Line 7) shows that there are approximately 400 monitors reporting 1-hr $SO_2$ concentrations to EPA's AQS. This includes 195 monitors in the SLAMS (State and Local Air Monitoring Stations) network that report 5-min average $SO_2$ concentrations in 2012 (Page 2-29, Lines 12-13). These numbers differ from those shown in Table 2-6 (Page 2-36) with 337 monitors reporting 1-hr data and 309 reporting 5-min data for the period of 2010-2012.
Much can be learned from the temporal and spatial analyses of past studies of $SO_2$ and $SO_4^=$ (e.g., Blanchard et al., 2013a, b; Hand et al., 2012; Hidy et al., 1978; Husain et al., 1998; Husain et al., 2004; Malm et al., 2002). The definition of "near the source of origin" (Page 2-31, Line 8) should be clarified. This implies a micro- (<100 m) to middle-scale (100-500 m) zone of influence (Chow et al., 2002). Plume touchdowns from tall stacks may occur several kilometers downwind of the emission point. Is there a reference to the population weighted emissions index (PWEI) estimates described (Lines 12-15, Page 2-31)? If this combines population densities with emissions in the same CBSA, it may give a biased estimate of exposure, especially in the eastern U.S. where county areas are small. A better approach (and possibly the one used here) would combine modeled and measured $SO_2$ concentrations with population densities (e.g., Zou et al., 2009).

The explanation and purpose of Figures 2-14 through 2-19 (Pages 2-41 to 2-46) are unclear. What do the "buffer zones" represent? Define what the letters and circles mean. With a $SO_2$ reaction rate of a few percent per hour (Calvert and Stockwell, 1983), and wind speeds aloft that are often equal to or exceed 15 km/hr at the exit of a tall stack, there is probably ample $SO_2$ left in the plume to touch down far beyond 15 km. If $SO_2$ monitors within 15 km of large point sources are expected to experience the greatest emission impact, the source zones of influence within 5, 10, and 15 km circled in the selected

six CBSA/metropolitan areas (Figures 2-14 to 2-19) should be plotted based on major NEI facilities rather than centered on ambient $SO_2$ monitors. Maps of spatial variations of a 3-year average of daily maximum 1-hr $SO_2$ concentrations at the annual 99$^{th}$ percentile based on existing AQS monitors for the U.S. and for each selected CBSA/metropolitan area would provide insight on spatial variations of $SO_2$ concentrations.

One would not expect high correlations among monitors that detect plume touchdown. A high level at one location means the relatively narrow plume is not being measured at another location. This is largely confirmed in Figures 2-20 and 2-22 (Pages 2-50 and 2-52). Higher correlations would imply some accumulation of $SO_2$ within the mixed layer, which would be evident in a diurnal time series plot. Comparison of frequency distributions in Figures 2-21 and 2-23 (Pages 2-51 and 2-53) provide more realistic estimates of exposure likelihood. Figures 2-26 and 2-27 (Pages 2-59 and 2-60) show evidence of touchdown as surface layers couple to layers aloft a few hours after sunrise.

Note that Payson and Phoenix, AZ, represent two different airsheds separated by terrain and should be treated separately. The Payson measurements in Table 2-8 (Pages 2-47 and 2-48) serve as a good example for a case study, with 50$^{th}$ percentile $SO_2$ of 50 ppb that is one to two orders of magnitude lower than the 99$^{th}$ percentile (295 ppb) and the maximum 1-hr (1,501 ppb). The distance between Site D in Payson/Phoenix and the copper smelter should be given (Page 2-46, Lines 8-9). More spatial/temporal analysis on wind direction/wind speed along with dispersion modeling estimates that associate ambient $SO_2$ concentrations with the location of plume touchdown should be illustrated to differentiate maximum 5-min concentrations from the daily maximum 1-hr $SO_2$ concentrations at the 99$^{th}$ percentile.

It is difficult to interpret the spatial concentrations exhibited in Figures 2-20 to 2-23 (Pages 2-50 to 2-53). Were the corresponding 24-hr average $SO_2$ concentrations between monitors used for Pearson correlations? One would expect a lack of correlation as plume touchdown at one site precludes a peak concentration at another site.

Section 2.5.3 Temporal Variability (Page 2-55)

The time period of ambient concentrations (1990-2012) in Figure 2-24 (Page 2-55) should match the emissions trend (1990-2013) in Figure 2-5 (Page 2-9). If data are available, Figure 2-24 needs to be updated for $SO_2$ to include the longer period from 1980-2014/15 to demonstrate the trend of decreasing $SO_2$ over the past 35 years.

As emission reductions can be related to the decline of both $SO_2$ and $SO_4^=$ concentrations, a map showing sites with collocated $SO_2$ and $SO_4^=$ measurements (e.g., from IMPROVE and CSN sites) could be used to illustrate temporal and spatial variations. Hand et al (2012) demonstrate the decline of $SO_2$ and $SO_4^=$ concentrations (in percent per year) in the U.S. for the period of 2001-2010. Diurnal variations of $SO_2$ and $SO_4^=$ with temperature can be used to illustrate atmospheric oxidation and/or local cold pools. Diel variabilities shown in Figures 2-26 and 2-27 (Pages 2-59 and 2-60) may be more descriptive if summer versus winter or warm versus cold seasons are shown separately. Plume touchdown in early morning hours with higher ambient $SO_2$ concentration is expected during summer, as vertical mixing is enhanced under warmer temperatures.

Time series distributions are a better way to characterize exposure than correlations in Figures 2-29 and 2-31 (Pages 2-71 and 2-73). The same plots should be generated for 99$^{th}$ percentile daily maximum 1-hr $SO_2$ concentrations that correspond to 75 ppb to illustrate the frequency of exceeded 1-hr $SO_2$ NAAQS.

What emissions were used in Figure 2-32 (Page 2-77) to estimate policy-relevant background (PRB) $SO_2$ using the MOZART model? What is the representativeness of the maximum PRB of 0.03 ppb $SO_2$ based on modeling with 2001 meteorological data? This background level is one to two orders of magnitude lower than the LDL of 2 or 0.2 ppb for $SO_2$ measurements.

c. *How informative is the analysis of correlations between $SO_2$ and co-occurring pollutant concentrations for interpretation of epidemiologic studies?*

Section 2.5.5 Co-pollutant- Correlations

Chapter 2 focuses on a single pollutant with only five out of 99 pages designated to show Pearson correlations with other co-pollutants. It is not clear how many sites are included in Figures 2-35 to 2-38 (Pages 2-81 to 2-84) to address correlations between $SO_2$ and other co-occurring pollutants. Individual sites and dates and/or times that show a correlation coefficient higher than 0.8 or 0.9 may warrant additional analysis. The possibility that $SO_2$ is more correlated with CO and average $PM_{2.5}$ (rather than $PM_{2.5}$ sulfur) needs to be explained.

**References**


Bennett, R.L., Stockburger, L., Barnes, H.M., (1994). Comparison of sulfur measurements from a regional fine particle network with concurrent ACID MODES network results. Atmospheric Environment 28, 409-419.

Blacker, J.H., Confer, R.G., Brief, R.S., (1973). Evaluation of the reference method for determination of sulfur dioxide in the atmosphere (Pararosaniline method). JAPCA 23, 525-527.

Blanchard, C.L., Hidy, G.M., Tanenbaum, S., Edgerton, E.S., Hartsell, B.E., (2013a). The Southeastern Aerosol Research and Characterization (SEARCH) study: Spatial variations and chemical climatology, 1999-2010. Journal of the Air & Waste Management Association 63, 260-275. http://www.tandfonline.com/doi/pdf/10.1080/10962247.2012.749816

Blanchard, C.L., Hidy, G.M., Tanenbaum, S., Edgerton, E.S., Hartsell, B.E., (2013b). The Southeastern Aerosol Research and Characterization (SEARCH) study: Temporal trends in gas and PM concentrations and composition, 1999-2010. Journal of the Air & Waste Management Association 63, 247-259. http://www.tandfonline.com/doi/pdf/10.1080/10962247.2012.748523

Bytnerowicz, A., Tausz, M., Alonso, R., Jones, D., Johnson, R., Grulke, N., (2002). Summer-time distribution of air pollutants in Sequoia National Park, California. Environ. Poll 118, 187-203.

Calvert, J.G., Stockwell, W.R., (1983). Acid generation in the troposphere by gas-phase chemistry. Environ. Sci. Technol 17, 428A-443A.

Chow, J.C., Engelbrecht, J.P., Watson, J.G., Wilson, W.E., Frank, N.H., Zhu, T., (2002). Designing monitoring networks to represent outdoor human exposure. Chemosphere 49, 961-978. https://www.researchgate.net/publication/10982533_Designing_monitoring_networks_to_represent_outdoor_human_exposure?ev=prf_pub

Chow, J.C., Watson, J.G., Bowen, J.L., Frazier, C.A., Gertler, A.W., Fung, K.K., Landis, D., Ashbaugh, L.L., (1993). A sampling system for reactive species in the western United States, in: Winegar, E.D., Keith, L.H. (Eds.), Sampling and Analysis of Airborne Pollutants. Lewis Publishers, Ann Arbor, MI, pp. 209-228.

Clarke, J.F., Edgerton, E.S., Martin, B.E., (1997). Dry deposition calculations for the Clean Air Status and Trends Network. Atmospheric Environment 31, 3667-3678.

DieselNet, (2014). Emission standards: Summary of worldwide diesel emission standards, Ecopoint, I., Mississauga, ON, Canada. http://www.dieselnet.com/standards/#na

Do, J.S., Chen, P., (2007). Amperometric sensor array for NOx, CO, O2 and SO2 detection. Sensors and Actuators B-Chemical 122, 165-173.

Dollard, G.J., Unsworth, M.H., Harve, M.J., (1983). Pollutant transfer in upland regions by occult precipitation. Nature 302, 241-243.

Dye, T.S., Gordon, B.J., Roberts, P.T., (2014). Air sensors: Quality data for the right application. EM, 22-27.

Ferek, R.J., Covert, P.A., Luke, W., (1997). Intercomparison of measurements of sulfur dioxide in ambient air by carbonate-impregnated filters and Teco pulsed-fluorescence analyzers. J. Geophys. Res 102, 16267-16272.

Frost, K.D., (2014). AERMOD performance evaluation for three coal-fired electrical generating units in Southwest Indiana. Journal of the Air & Waste Management Association 64, 280-290.

Guerra, S.A., Olsen, S.R., Anderson, J.J., (2014). Evaluation of the SO2 and NOX offset ratio method to account for secondary PM2.5 formation. Journal of the Air & Waste Management Association 64, 265-271.

Hand, J.L., Schichtel, B.A., Malm, W.C., Pitchford, M.L., (2012). Particulate sulfate ion concentration and SO2 emission trends in the United States from the early 1990s through 2010. Atmos. Chem. Phys 12, 10353-10365.

Hanna, S.R., Egan, B.A., Purdum, J., (2001). Evaluation of the ADMS, AERMOD, and ISC3 dispersion models with the OPTEX, Duke Forest, Kincaid, Indianapolis and Lovett field datasets. International Journal of Environment and Pollution 16, 301-314.

Hidy, G.M., Mueller, P.K., Tong, E.Y., (1978). Spatial and temporal distributions of airborne sulfate in parts of the United States. Atmospheric Environment 12, 735-752.

Horrocks, L.A., Oppenheimer, C., Burton, M.R., Duffell, H.J., Davies, N.M., Martin, N.A., Bell, W., (2001). Open-path Fourier transform infrared spectroscopy of $SO_2$: An empirical error budget analysis, with implications for volcano monitoring. J. Geophys. Res 106, 27647-27659.

Huang, M., Carmichael, G.R., Spak, S.N., Adhikary, B., Kulkarni, S., Cheng, Y., Wei, C., Tang, Y., D'Allura, A., Wennberg, P.O., Huey, G.L., Dibb, J.E., Jimenez, J.L., Cubison, M.J., Weinheimer, A.J., Kaduwela, A.P., Cai, C., Wong, M., Pierce, R.B., Al-Saadi, J.A., Streets, D.G., Zhang, Q., (2011). Multi-scale modeling study of the source contributions to near-surface ozone and sulfur oxides levels over California during the ARCTAS-CARB period. Atmos. Chem. Phys 11, 3173-3194.

Husain, L., Dutkiewicz, V.A., Das, M., (1998). Evidence for decrease in atmospheric sulfur burden in the eastern United States caused by reduction in $SO_2$ emissions. Geophys. Res. Lett 25, 967-970.

Husain, L., Parekh, P.P., Dutkiewicz, V.A., Khan, A.R., Yang, K., Swami, K., (2004). Long-term trends in atmospheric concentrations of sulfate, total sulfur, and trace elements in the northeastern United States. Journal of Geophysical Research-Atmospheres 109. 10.1029/2004jd004877.

Irwin, J.S., (2014). A suggested method for dispersion model evaluation. Journal of the Air & Waste Management Association 64, 255-264.

Isakov, V., Venkatram, A., Touma, J.S., Koracin, D., Otte, T.L., (2007). Evaluating the use of outputs from comprehensive meteorological models in air quality modeling applications. Atmospheric Environment 41, 1689-1705.

Janoe, S., Craft, Z., (2015). 1-hr SO2 NAAQS: Designations and implementation. EM 21, 4-8.

Kim, K.H., Kim, M.Y., (2001). Comparison of an open path differential optical absorption spectroscopy system and a conventional in situ monitoring system on the basis of long-term measurements of $SO_2$, $NO_2$, and $O_3$. Atmospheric Environment 35, 4059-4072.

Kok, G.L., Dasgupta, P.K., Adams, D.F., Appel, B.R., Farwell, S.O., Knapp, K.T., Pierson, W.R., Reiszner, K.D., Tanner, R.L., (1989). Method 704B: Determination of sulfur dioxide content of the atmosphere (Formaldehyde Absorber/ Pararosaniline Method), in: Lodge, J.P. (Ed.), Methods of Air Sampling and Analysis, 3 ed. Lewis Publishers, Chelsea, MI, pp. 499-502.

Kushner, E.J., (1976). On determining the statistical parameters for pollution concentration from a truncated data set. Atmospheric Environment 10, 975-979.

Leppanen, S., Anttila, P., Lattila, H., Makkonen, U., (2005). Long-term comparison of filter method and sensitive analyser in monitoring of sulphur dioxide. Atmospheric Environment 39, 2683-2693.

Lillis, D., Cruz, C.N., Collett, J.L., Jr., Richards, L.W., Pandis, S.N., (1999). Production and removal of aerosol in a polluted fog layer: model evaluation and fog effect on PM. Atmospheric Environment 33, 4797-4816.

Luke, W.T., (1997). Evaluation of a commercial pulsed fluorescence detector for the measurement of low-level $SO_2$ concentrations during the Gas- Phase Sulfur Intercomparison Experiment (Paper 96JD03347). J. Geophys. Res 102, 16255-16266.

Malm, W.C., Schichtel, B.A., Ames, R.B., Gebhart, K.A., (2002). A ten-year spatial and temporal trend of sulfate across the United States. J. Geophys. Res. Atmos 107, ACH 11-11-ACH 11-20. doi:10.1029/2002JD002107.

Mamane, Y., Donagi, A.E., (1976). The effects of sampling line on the measurement of sulfur dioxide: A case study. JAPCA 26, 991-992.

Medina, D.S., Liu, Y.D., Wang, L.M., Zhang, J.S., (2011). Detection of sulfur dioxide by cavity ring-down spectroscopy. Environ. Sci. Technol 45, 1926-1931.

Ohira, S.I., Toda, K., (2005). Micro gas analysis system for measurement of atmospheric hydrogen sulfide and sulfur dioxide. Lab on a Chip 5, 1374-1379.

Paine, R., Samani, O., Kaplan, M., Knipping, E., Kumar, N., (2015). Evaluation of low wind modeling approaches for two tall-stack databases. Journal of the Air & Waste Management Association 65, 1341-1353. 10.1080/10962247.2015.1085924.

Pandis, S.N., Seinfeld, J.H., (1989). Mathematical modeling of acid deposition due to radiation fog. J. Geophys. Res 94, 12911-12911.

Rehbein, P.J.G., Kennedy, M.G., Cotsman, D.J., Campeau, M.A., Greenfield, M.M., Annett, M.A., Lepage, M.F., (2014). Combined analysis of modeled and monitored SO2 concentrations at a complex smelting facility. Journal of the Air & Waste Management Association 64, 272-279.

Roberts, T.J., Saffell, J.R., Oppenheimer, C., Lurton, T., (2014). Electrochemical sensors applied to pollution monitoring: Measurement error and gas ratio bias - A volcano plume case study. Journal of Volcanology and Geothermal Research 281, 85-96.

Rood, A.S., (2014). Performance evaluation of AERMOD, CALPUFF, and legacy air dispersion models using the Winter Validation Tracer Study dataset. Atmospheric Environment 89, 707-720.

Seinfeld, J.H., Pandis, S.N., (1998). Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. John Wiley & Sons, New York, NY.

Seinfeld, J.H., Pandis, S.N., (2006). Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. John Wiley & Sons, New York, NY.

Shadwick, D.S., Sickles, J.E., II, (2004). Sample size for seasonal mean concentration, deposition velocity and deposition: A resampling study. Atmospheric Environment 38, 477-489.
Sickles, J.E., II, Shadwick, D.S., (2002). Biases in Clean Air Status and Trends Network filter pack results associated with sampling protocol. Atmospheric Environment 36, 4687-4698.
Sickles, J.E., II, Shadwick, D.S., (2007). Seasonal and regional air quality and atmospheric deposition in the eastern United States. J. Geophys. Res. Atmos 112.
Sickles, J.E., II, Hodson, L.L., Vorburger, L.M., (1999). Evaluation of the filter pack for long-duration sampling of ambient air. Atmospheric Environment 33, 2187-2202.
Snyder, E.G., Watkins, T.H., Solomon, P.A., Thoma, E.D., Williams, R.W., Hagler, G.S.W., Shelow, D., Hindin, D.A., Kilaru, V.J., Preuss, P.W., (2013). The changing paradigm of air pollution monitoring. Environ. Sci. Technol 47, 11369-11377.
Stelson, A.W., Bao, X., (1988). Water interference of pulsed fluorescence $SO_2$ monitor. JAPCA 38, 420-420. http://www.tandfonline.com/doi/pdf/10.1080/08940630.1988.10466393
Terraglio, F.P., Manganelli, R.M., (1962). Laboratory evaluation of sulfur dioxide methods and the influence of ozone oxides of nitrogen mixtures. Anal. Chem 34, 675-677.
Toda, K., Inoue, H., Sanemasa, I., (1998). Miniaturized detector of sulfur dioxide based on the flow conductometry of an absorbing solution. Bunseki Kagaku 47, 727-734.
Trieff, N.M., Wohlers, H.C., O'Malley, J.A., Newstein, H., (1968). Appraisal and modification of West-Gaeke method for sulfur dioxide determination. Journal of the Air Pollution Control Association 18, 329-331. http://www.tandfonline.com/doi/pdf/10.1080/00022470.1968.10469136
U.S.EPA, (2013a). DRAFT Roadmap for Next Generation Air Monitoring. http://www.epa.gov/research/airscience/docs/roadmap-20130308.pdf
U.S.EPA, (2013b). 2011 National Emissions Inventory Data, Agency, U.S.E.P., Research Triangle Park, NC. http://www.epa.gov/ttn/chief/net/2011inventory.html
Vadjic, V., Gentilizza, M., Cackovic, M., Eskinja, I., (1992). Comparison of mass concentrations of SO2 determined in air by two different methods. Environ. Mon. Assess 21, 19-26.
Wang, A., Brauer, M., (2014). Review of next generation air monitors for air pollution, Canada, E., Gatineau, QC. http://elk.library.ubc.ca/handle/2429/46628
Wang, C.F., Corbett, J.J., Firestone, J., (2007). Modeling energy use and emissions from North American shipping: Application of the ship traffic, energy, and environment model. Environ. Sci. Technol 41, 3226-3232.
West, P.W., Gaeke, G.C., (1956). Fixation of sulfur dioxide as disulfitomercurate (II) and subsequent colorimetric estimation. Anal. Chem.
Wohlers, H.C., Trieff, N.M., Newstein, H., Stevens, W., (1977). Sulfur dioxide adsorption on and desorption from Teflon, tygon, glass, stainless steel and aluminum tubings. Atmospheric Environment 1, 121-121.
Wright, R.S., Rickman, E.E., Jr., Green, A.H., Sickles, J.E., II, Decker, C.E., McElroy, F.F., (1989). Laboratory and field evaluations of extrasensitive $NO_2$ and $SO_2$ analyzers, Proceedings, 82nd Annual Meeting of the Air and Waste Management Association. Air & Waste Management Association, Pittsburgh, PA.
Zou, B., Wilson, J.G., Zhan, F.B., Zeng, Y.N., (2009). An emission-weighted proximity model for air pollution exposure assessment. Science of the Total Environment 407, 4939-4945.

# Dr. Aaron Cohen

**Comments on Chapter 5: Integrated Health Effects of Exposure to Sulfur Oxides**

**General Comments**
The authors are to be commended for a chapter that, for the most part, effectively summarizes and distills a very large literature. Causal determinations have been updated for eight classes of health outcomes, and for five of these the determination has changed since the 2008 ISA, in all cases having been strengthened. For the most part, the rationale for the determinations has been well described.

I have reviewed each section of Chapter 5 below, in each case focusing on the extent to which the authors' have adequately justified the proposed determinations presented in the Executive Summary (Page xliv, Table ES-1).

I have several general comments with regard to the following issues:

- Exposure measurement error deserves additional attention. The reader is, for the most part asked to assume bias to the null, though bias away from the null may be an issue in long-term exposure studies. In addition, it should be noted that measurement error correction has, for the most part, not been done.

- Discussion of E-R function shapes is overly confident with regard to linearity given the very limited empirical exploration of this issue for most outcomes.

- With regard to co-pollutant confounding, the authors should note which pollutants are most important in this regard. Arguably, that would be $PM_{2.5}$ and sulfates, though relatively few studies have this information. Low correlations in the EPA air pollution data base do not necessarily provide reassurance in this regard; the correlations that matter are those in the evidentiary studies.

**Section 5.2.1 Respiratory Morbidity Short-Term Exposure**
This section is well-written and presents clearly and logically the evolution of the evidence since the 2008 ISA. The determination that the current evidence, based largely on the evidence for exacerbation of asthma from both observational and experimental studies, is consistent with a causal relationship is well-justified.

Epidemiologic evidence for asthma exacerbation is broadly consistent with experimental evidence in that the strongest effects appear to occur at the shortest lags.
Page 5-17, lines 18-30: The justification for down-weighting Goodman et al. 2015 seems reasonable.
Page 5-23, lines 14-16: Any idea why effects were seen in just one treatment group? Ages in this small study ranged from 12-65 years. Were treatment groups balanced on age or correlates of disease severity?

Page 5-29, lines 4-8 and Page 5-33, lines 9-11: The exposure-response relationships could be presented more clearly and explicitly. In addition to Table 5.2 a figure(s) might help.

Page 5-35, Table 5-8: In each of the three studies the investigators failed to report key information on either $SO_2$ exposure (O'Connor et al.; Spira-Cohen et al.) or effect estimates (Dales et al.) did EPA ask the authors to provide these data? If not, could they?

Page 5-36, line 9-10: O'Connor et al. actually reported increased RR for each symptom. The authors have mis-interpreted the evidence because they have used the 95% CI to perform a statistical significance test.

Page -86, line 5: rather than "there was uncertainty," better to say the estimate was very imprecise.

Page 5-108, line 8: better to say "supportive of an inverse association…"

Page 5-134, lines 12-13: the evidence for linearity of the mortality exposure-response function is based on very limited exploration of alternatives and this should be noted.

**Section 5.2.2 Respiratory Morbidity Long-Term Exposure**

The determination that the evidence is now "suggestive" of a causal relationship rather than "inadequate," is based on two new studies of asthma incidence in children and one new experiment in rats. This is justifiable given how "suggestive" is defined (Preamble, Table II). However, the results of the longitudinal epidemiologic studies cannot confidently be interpreted as due to $SO_2$ exposure as opposed to other pollutants and, perhaps, other risk factors.

*Section 5.2.2.1 Development of Asthma*

It would help to discuss briefly which risk factors in addition to air pollution are most important to assess re. confounding and effect measure modification. When information on a specific covariate (e.g., housing characteristics or smoking) is unavailable to what extent can more readily available data (e.g., SES) serve as a surrogate and how much residual confounding would be expected. This should then be discussed re. studies of Clark et al. and Nishimura et al. which studies include different covariates, as do the cross-sectional studies in Table 5-28.

Is loss to follow-up an issue in either Clark et al. or Nishimura et al.? It is not mentioned.

**Section 5.3.1 Cardiovascular Effects Short-Term Exposure**

The determination that the evidence is now "suggestive" of a causal relationship rather than "inadequate," is based on new epidemiologic studies, there having been little or no new toxicologic studies. The epidemiologic evidence on MI triggering is given particular weight (Page 5-225, lines 6-10 and Page 5-223, lines 12-14). I find the new determination harder to defend in this case because: 1) the new studies share the main weakness noted for the older studies in the ISA 2008, namely that any observed effects my actually be due to other pollutants; 2) $SO_2$ effects in studies of MI hospital admissions were substantially reduced or eliminated in two-pollutant models (Page 5-175; lines 18-37; and 3) the only study that precisely characterized exposure at the time of MI onset found no effect (Bhaskaran et al. 2011; Page 5-175; lines 8-10).

Page 5-186, lines 15-29: Are the null studies stronger or weaker methodologically? To say that results are "inconsistent" just states the obvious. Bell et al. (2008) appear to have used considerably more extensive monitoring data than other studies and to have conducted an assessment of alternative exposure models none of which resulted in increased risk of stroke admission associated with $SO_2$.

Page 5-194, lines 29-31: this clearly worded description of the evidence for Heart Failure applies with equal force to Venous Thromboembolism (Page 5-194, lines 12-14)

Page 5-200, lines 5-6: what is meant by "clinically relevant?" Is clinical relevance an issue of general concern? If so, there are other places in the chapter, e.g., effects on sub-clinical markers, where this issue should be addressed.

Page 5-209, line 4: typo: should read "…consistent with…"

Page 5-208, lines 9-11 and Page 5-209, lines 7-8: Does this imply that the exposure-response relationship may be nonlinear?

Page 5-213, lines 1-2: the evidence for linearity of the mortality exposure-response function (Chen et al. 2013; Page 5-211, lines 1-6) is based on very limited exploration of alternatives and this should be noted.

**Section 5.3.2 Cardiovascular Effects Long-Term Exposure**

The determination that the regarding cardiovascular effects of long-term exposure to $SO_2$ is "inadequate" to make causal determination is well-supported by the evidence reviewed in this section. The authors correctly note the inability to attribute increased rates of cardiovascular disease to $SO_2$ given high correlations with co-pollutants including $PM_{2.5}$ (Page 5-234, lines 8-11) and the absence of supporting experimental and clinical evidence (Page 5-237, lines 11-14).

**Section 5.4 Reproductive and Developmental Effects**

The determination that the evidence is now "suggestive" of a causal relationship rather than "inadequate," is based on a growing, largely epidemiologic, literature on adverse birth outcomes that is generally well-summarized in the Chapter. The reasons why the evidence was considered "inadequate" in the 2008 ISA are summarized on Page 5-241, lines 2-6. Although there are now more studies these same reasons: lack of toxicologic support, potential confounding by co-pollutants, and uncertainty regarding exposure, still seem to apply to the current literature as noted by the authors (e.g., Page 5-256, lines 14-17).

**Section 5.5 Mortality**

**Section 5.5 Total Mortality Short-term exposure**

The determination that the evidence remains "suggestive" of a causal relationship is well-supported by the current review, given that the shortcomings of the literature identified in the 2008 ISA (Page 5-261, lines 7-10) have largely not been addressed in more recent studies.

Specifically, the potential for confounding of observed $SO_2$ effects by co-pollutants remains largely unresolved and, to the extent that confounding by co-pollutants has been assessed, neither $PM_{2.5}$ nor sulfate PM has been included.

Page 5-284, lines 12-16: the evidence for linearity of the mortality exposure-response function is based on very limited exploration of alternatives and this should be noted. That said, to my eye neither Figure 5-22 nor Figure 5-23 strongly support a such a definitive statement on linearity, especially given the uncertainty regarding the role of $SO_2$ vs. other pollutants that $SO_2$ may well represent.

**Section 5.5 Total Mortality Long-term exposure**

The determination that the evidence is now "suggestive" of a causal relationship rather than "inadequate," is based on epidemiologic studies reported since the 2008 ISA. However, although the number of studies has increased the authors offer little evidence that the uncertainties identified in the 2008 ISA have been reduced (Page 5-289, lines 1-9), which they appear to acknowledge, at least for European (Page 5-298, lines 10-11) and Asian (Page 5-299, lines 7-9) studies.

The possibility that "sulfate or other particulate $SO_x$…" might be responsible for increased mortality associated with $SO_2$ was noted in the 2008 ISA and in the current chapter (Page 5-306; lines 6-9). Absent experimental evidence to the contrary, this remains a tenable explanation for the observed mortality effects of $SO_2$.

The authors' conclusion re. the epidemiologic evidence (Page 5-306, lines 9-18) seems inconsistent with the change in determination from "inadequate" to "suggestive."

**Section 5.5 Cancer - Long-term exposure**

It is not entirely clear to me why the determination has changed from the 2008 ISA wherein the evidence was considered "inconclusive." Given the evidence reviewed in the chapter that still seems the best description. Notably the two of the three largest studies reviewed (Krewski et al. 2009 and Brunekreef et al. 2009) report null results for lung cancer. Studies of lung and other cancers are more numerous but largely share the same limitations as earlier ones with regard to whether other pollutants are responsible for observed effects.

A summary Table and/or forest plot would be helpful in presenting the epidemiologic studies.

## Dr. Alison Cullen

In light of the charge, the most important overall need for Chapter 3 of the Sulfur Oxides Draft ISA is to improve clarity, reduce redundancy and emphasize new information, as well as, to clearly identify and distinguish the science that drives policy. Substantial rewriting is warranted.

The 2015 Draft ISA serves as a review of the policy relevant science and an update to the 2008 ISA, with a particular emphasis on new information available between January 2008 and April 2015. The most relevant studies from previous assessments are also carried forward. In the 2015 Draft, studies from previous assessments and those available since 2008 are presented together in tables (e.g., Tables 3-3 and 3-4) using alphabetical ordering by author. A temporal ordering would be useful as updates in the language regarding causal determination for relationships between sulfur dioxide and health effects has been strengthened in the current Draft ISA (e.g., from "inadequate to infer" to "suggestive") presumably based on the newly available studies.

Confusion in the use of basic terms leads to a lack of clarify in places. For example exposure and exposure concentration are sometimes used interchangeably. The definition of exposure as related to conditions at the human boundary, in contrast to the movement of pollutants across the human boundary (as implied by the inclusion of Table 3-6 with ventilation rates), remains an important distinction. Specific passages and segment edit suggestions will be included in the next set of comments.

Section 3.2.2.1 lacks clarity in the discussion of comparisons of measurements and model results, as well as inter-comparison of models. Clarify whether the SPM and EWPM were both compared to AERMOD, and also one or several sets of model results were compared to measurements from 3 monitoring sites. Or if other comparisons are implied these should be laid out. A tabular format could work well for making these comparisons evident.

In the discussion of land use regression models the distinction between spatial variability and inter-individual variability is sometimes lost, whereas particularly for exposure assessment the different impacts of these different types of variability are important.

The discussion of "exposure error" and "bias" is quite difficult to follow as written and would benefit from a more systematic layout. Reference to spatial and temporal variability, and its interaction with the relationship between site monitoring versus personal monitoring, in particular needs more clarity beyond the blanket reference to wider confidence intervals around exposure and health effect estimates. The layout in the figures of Section 3.3.3.2 is valuable so perhaps this could be used as a model for other sections.

Eliminate and tighten up repetitive language. For example, language introducing Section 3.3.3.4 appears in exact duplicate twice.

Edit presentations related to correlation (e.g., personal-ambient relationships). These currently incorporate measures that fall below the detection limit, as well as measures above detection limits, and then report the inevitable impact of the proportion of below detection limit measures on both correlation and $R^2$. This information appears in multiple sections and could be consolidated. The representation of

measures below the detection limit needs clarification. It may help to include a parallel presentation focused solely on the measures above detection limit for comparison. However, overall, the conclusion from page 3-48,

> "Low correlations in situations with a high proportion of samples below the detection limit should not be interpreted as evidence for the lack of a relationship between personal exposure and ambient SO2 concentrations."

is correct, thus there is a legitimate question regarding how much analysis and discussion is warranted for results that include a high proportion of BDL data.

The document should focus on information (specific to sulfur oxides) which forms the basis for the conclusions presented. Within Section 3.2.1 there is a tremendous amount of excellent background material on monitoring, the instruments and methods that are used and the development of the field over time. Much of this is generic. There is also a substantial discussion of emerging technology for sulfur dioxide monitoring, that is currently in its early stages and not yet commercially available. Some of this text could be removed and the content simply referenced from other sources and documents, to increase clarity of the presentation, and to sharpen the focus.

Similarly, the background information related to modeling (e.g., time activity, micro-environmental exposure, total personal exposure, etc.) does not always focus on considerations related to sulfur dioxide. Background is important, but the balance of strictly background text, versus that which is necessary for a discussion of sulfur dioxide, is skewed to the former. This document could rely more extensively on the 2008 ISA and referencing content that was reported there. This would help to improve focus of the current Draft ISA.

The use of tabular formats for bias information (e.g., in Table 3-2) is an example of a very efficient use of space which reduces the need for more extensive discussion in the individual sections with no loss of clarity.

Table 3-9 includes substantial detail associated with studies of reproductive timing and effects however this content is not clearly explained. For example, the duration is sometimes expressed in terms of trimesters of pregnancy or months prior to conception; however the interpretation of this information for the ISA is missing. In short, the level of detail in Table 3-9 is inconsistent with the discussion provided.

In Table 3-10, and the text directly above it, there is confusion between the terms "true value" and modeled estimate. Model results are necessarily estimates and carry a full range of uncertainty and variability which must be taken into account when they are applied to real world applications.

# Dr. Delbert Eatough

**Comments on Chapter 2: Atmospheric Chemistry and Ambient Concentrations of Sulfur Oxides**

*Chapter 2 describes scientific information on sources, atmospheric chemistry, and measurements and modeling of ambient concentrations of gaseous sulfur oxides.*

a. *To what extent is the information presented on sources, atmospheric chemistry, and measurements and modeling of ambient concentrations complete, and relevant to the review of the $SO_2$ NAAAQS?*

Major Points:

While EPA has done a good job of reviewing the concentrations and chemistry of the many gaseous species which can form $SO_2$, they have been incomplete in their review of the chemistry of $SO_2$, touching only on the formation of sulfate. They have not mentioned the formation of both inorganic and organic particulate S(IV) and organic S(VI) compounds which can form from the chemistry of $SO_2$ in plumes and in urban environments. The organic compounds include dimethyl sulfate, monomethyl sulfuric acid and bis-hydroxymethyl sulfone (a S(IV) compound). The inorganic S(IV) compounds can be significant, especially in smelter, integrated steel mill and coal-fired power plant plumes. The alkyl sulfates and sulfone are often a significant fraction of urban particulate sulfur oxides. These oxysulfur compounds will probably have potential health effects quite different from sulfate and the S(IV) compounds will be analyzed as sulfate by most techniques presently used for analysis. The fraction of the aerosol S(IV) species which is probably most relevant to the ISA is the particulate inorganic S(IV) compounds which have been shown to exacerbate $SO_2$ inhalation responses in animals (Alarie 1973, Amdur 1971). I had commented on the potential importance of these species in my April 29, 2014 comments to EPA in connection with the previous $SO_2$ review meeting. I would be happy to provide additional information as requested. I will make additional comments in this area in my response to parts b. and c. of this charge questions. I have also included and amplified on the points made in my 2014 comments in this area (modified to emphasize this ISA) in an Appendix at the end of my comments.

Other Points:

I believe that EPA has not been careful in the use of the terms $SO_2$ and $SO_X$. I would recommend the total document be named Sulfur Dioxide" and not "Sulfur Oxides" as they correctly point out this is the focus of the ISA. Likewise, the same change should be made to the heading of each of the chapters. While they point out there are many compounds present in the group "Sulfur Oxides", they point out the only relevant compound for the NAQQS review is $SO_2$. They also refer to $SO_X$ as only gaseous compounds when the more common usage is for both gas and particulate species, including sulfate, e.g. definition of $SO_X$ on page 2.2 line 7.

The relationships among the various sectors given in Figure 2.1, Figure 2-5, and Table 2-1 are not clear to me. Where are oil refineries, smelters and integrated steel mills included in these sectors (which appear to vary in name among the two figures and table)? Having a Figure (such as Figure 2-3) for these three sources would be of interest to me. I focus on these sources because they can frequently emit low

lying emission, are often high emitters, are often close to urban centers and will all form, in addition to $SO_2$ and sulfate, the non-sulfate particulate oxysulfur species I have outline above.

The counties which exceed 75 ppb (or higher) one hour $SO_2$ average are quite clear in Figures 2-2 and 2-3, they cannot be distinguished in Figure 2-4. Being able to so do would be useful.

The ratio of the red and green lines in Figure 2-5 looks to be about 4-5 which is quite different than suggested by Figure 2-1 where EGU = Coal exceeds all other sectors by a factor of nine. This conflict needs to be reconciled..

EPA has been careless in the identification of free radicals in Section 2.3. I would recommend the use of the • symbol for clarity, e.g. OH•. For example, on line10, page 2-10, $HSO_3$ is called the bisulfate ion, however, what is meant is $HSO_3$•, the radical. Equation 2-1 would then read:

$$SO_2 + OH\bullet + M \rightarrow HSO_3\bullet + M$$

with similar changes throughout this section.

Page 2-21, line 13. Add that sulfuric acid is usually quickly neutralized by ammonia and give the reactions occurring.

Page 2-12 line 16. “particle” should be "aqueous”.

Page 2-12 line 22. I would add the Pitts, Finlayson-Pitts book as it is better on chemistry than the ones listed.

Sections 2.3.2 and 2.3.3. We have written two papers which you might want to incorporate into this section (they are attached in the covering e-mail but not included in my comments because of copy right restriction). I can provide copies upon request,

> Eatough D.J., Caka F.M. and Farber R.J. (1994) "The Conversion of $SO_2$ to Sulfate in the Atmosphere," *Israel J. Chemistry*, **34**, 301–314.

is a review of all plume studies on the conversion of $SO_2$ to sulfate conducted to the time of the article, which covers the great majority of plume studies. It shows that the non-aqueous phase, clear sky conversion daytime conversion rate is constant when adjusted for T and relative humidity and is dominated by the hydroxyl radical chemistry (the second order chemistry not being important). It also shows that the effect of forming an aerosol aqueous phase on the conversion rate is dramatic. The figure in this manuscript which summarizes these points follows:

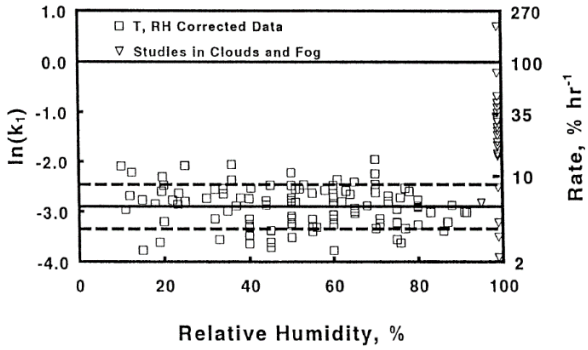



Fig. 3. Variation of $\ln(k_1)$ for the homogeneous conversion of $SO_2(g)$ to sulfate as a function of relative humidity. Solid line represents mean, dashed line, 1 standard deviation from mean. The $k_1$ values have been corrected to 50% RH using eq 12 and to 25 °C using eq 13. Also shown as the inverted triangles are uncorrected $\ln(k_1)$ values for the conversion of $SO_2$ to sulfate in fog and clouds.

Eatough D.J., Arthur R.J., Eatough N.L., Hill M.W., Mangelson N.F., Richter B.E., Hansen L.D. and Cooper J.A. (1984) "Rapid Conversion of $SO_2$(g) to Sulfate in a Fog Bank," *Environ. Sci. Technol.*, **18**, 855–859.

is a report on the formation of sulfate in an oil fired power plant plume and demonstrates both the rapid aqueous phase conversion of $SO_2$ to sulfate in a fog bank and the immediate shift to gas phase hydroxyl radical chemistry when the plume exits the fog bank. This paper also demonstrates the formation of bis-hydroxymethyl sulfone in the plume, but only in the absence of the fog bank. The Figure from the paper illustrating these points follows:

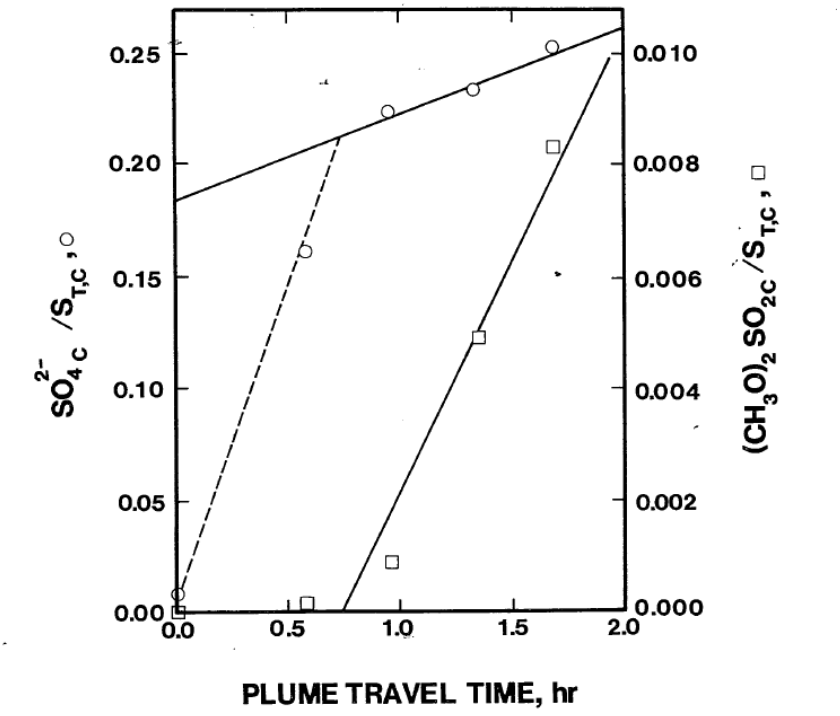



**Figure 3.** Variation of $SO_4^{2-}{}_C/S_{T,C}$ and $(CH_3O)_2SO_2/S_{T,C}$ vs. plume travel time for the aerosol samples collected Sept 26.

The material I have included in the Appendix further amplifies on the chemistry discussed in these two articles.

b. *Please comment on the extent to which available information on the spatial and temporal trends of ambient $SO_2$ concentrations at various scales has been adequately and accurately described. In particular, what is the extent to which the analysis or recently available 5-min $SO_2$ concentration data are informative in considering relationships between 5-min and 1-hr $SO_2$ concentrations?*

Major Points:

I found Section 2.5.2.2 on Urban Spatial Variability and Section 2.5.3 on Temporal Variability hard to correlate until I found they were working with a different data base in the two sections. Section 2.5.2.2 contains Table 2.8 which has 1-hour max sulfur dioxide concentration distribution for 42 monitors in six core-based statistical area/metropolitan focus areas. The nine highest monitors based on the 99 percentile data in this data set can be identified. Section 2.5.3 discusses the total national data set up to Section 2.5.3.3 and then discusses the 5 minute average data for the same 42 monitors as discussed in Section 2.5.2.2. However, the nine highest monitors based on the 99 percentile data for the 5 minute data set are different for this data set and some of the 1 hour 99 percentile data are higher than the 5 minute 99 percentile data, which cannot be. Comparisons between the two are therefore suspect, e.g. Figure 2-28. The section on comparison of hourly average and 5 minute data would be more convincing if only hourly averaged data with complete 5 minute averaged data were used. The consensus report on Chapter 2 recommend expanding the data set through 2014.

**On the nature of sources responsible for high 99$^{th}$ percentile 1 hour $SO_2$ Data at sampling sites included in Figure 2-12:**

I have correlated the data for the sampling sites given in Figure 2-12 with the possible presence of high concentrations of non-sulfate metal S(IV) aerosol species know to be present in plumes of metal smelters and integrated steel processing plants (the sampling site 99$^{th}$ percentile values is given om the following discussion where it was present in the data supplied by EPA). This comparison is only partially complete because not all sampling sites identified with a color other than blue were included in the data set provided to the committee by EPA.

There are only three currently operating smelters of importance in this regard in the U.S.

The Asarco Cu Smelter at Hayden, AZ (295). Figure 2.12 identifies this site as one of the 6 sites with 99$^{th}$ percentile daily 1-h max $SO_2$ concentration of 201-400 ppb. The summary proved by EPA correctly identifies the city as Hayden and gives the correct coordinates for the site. However, the CBSA name is incorrectly listed as Payson, AZ and the data are incorrectly included in the discussion of 2-19. (This error in Figure 2-19 and Table 2-8 needs to be corrected, it is not clear what the data for D in Figure 2-19 should be).

The Kennecott Cu Smelter in Magna, UT. The Salt Lake City area had no sites in where the 1-h data exceed the 99$^{th}$ percentile 75 ppb limit. However, prior to the installation of new acid plants and the construction of a tall stack, impacts in the Salt Lake Valley were frequent and significant. It has been suggested (see Appendix material to my comments) that metal S(IV) species played a role in the

exacerbation of asthma as seen the early EPA CHESS study conducted when $SO_2$ concentrations in the Salt Lake Valley were high..

The Doe Run Herculaneum lead smelter in Missouri south of St. Louis on the Mississippi River. The previous ISA highlighted emissions from this facility. However, the source site in Herculaneum is given in Figure 2-17, but data from the site are not included in Figure 2-12. It can be anticipated that more local impact was still high (if monitored at Herculaneum during the time frame covered) and some impact from this smelter might have been seen in St. Louis. The smelter was discontinued in 2014 and will no longer be an $SO_2$ contributor to the area.
Integrated steel miles can also be expected to be a significant source of metal S(IV) species. The following operating integrated steel mills can be expected to have influenced the 99$^{th}$ percentile 1 hour $SO_2$ Data for sites given in Figure 2-12 which fell in the 75-200 concentration range (Because of incompleteness in the data made available from EPA for the comparison summarized here, not all sources are clearly identified in this summary. The consensus report recommends making this comparison complete):

The 99$^{th}$ percentile 1 hour $SO_2$ concentrations which I could positively identify are given in ( ).

Weirton Steel Corporation, Weirton, PA

- Steubenville OH (4 miles down the Ohio River)
- Follansbee WV (12 miles down the Ohio River)
- Weirton-Steubenville WV-OH (close to Weirton facility)

USS Edgar Thomson Plant, Pittsburgh (140): Liberty, Pittsburgh (5 miles from facility)

US Steel Corporation, AK Steel Holding and Great Lakes Steel Works, Detroit: Detroit

US Steel Gary Works and ArcelorMittal East Chicago Works, Indiana: Indiana

ArcelorMittal Cleveland Works, Cleveland (80): Cleveland

The sites in Detroit and Indiana are more generic because I could not identify the site from the list provided by EPA but a high site was clearly present in Figure 2-12.

These comparisons indicate high $SO_2$ concentrations (and presumable high inorganic particulate S(IV)) are frequently associated with monitors near integrated steel mills.

<u>Other Points:</u>

Section 2.5.3.1 discusses long-term trends and the material at the top of page 2-56 implies the decrease is due to, first, the reduction in power plant emissions due to the Acid Rain Program and, second, regulations on sulfur content in diesel fuel. However, the diesel fuel content reduction has been much smaller than, for example, industrial fuel combustion emission from other than power plants (Figure 2-5). This material should reflect Figure 2-5 when giving credit.

Page 257, line 7. Should skewed right be skewed higher?

I found the discussion on seasonal trends to be rather conjectural. Will modeling support the causes for the trend listed? Is the frequency of plume touch down known? Have the effects of emission heights and meteorology been considered? I find it hard to accept the statements on highest concentrations being due to increased mixing heights.

Figure 2-11. It is quite difficult to distinguish the NCORE sites (as opposed to SLAMS sites) in this Figure. A change of symbol identification should be considered.

Table 2-6. It would be useful to refer to the "Except during 2010" statement to the footnote in the Table for clarity. I also assume that the completeness criteria is the reason for including only 309 monitors for the 5-min data and 337 monitors for the 1-h data in the subsequent large number of figures in this section. A dropout rate of near 25% seems rather large. Can you add a section on why the percent is so high and whether it is anticipated to improve in the future?

Throughout this portion of the document the Payson/Phoenix area and Gila County are both used to refer to the same geographical area. It would be helpful if only one term were used.

Section 2.5.4 discusses background concentrations using the 2001 data given in the 2008 ISA. There are a few points EPA should consider on this section as given in the next paragraph.

Figure 2-32. It should be made clear in the figure caption that these are 2001 data. I assume the second "Background" figure refers to PRB as defined in the text. This might be made clear. The "Total" top figure is also based on 2001 data. However, as suggested by Figure 5-2, a figure based on current data would have a slightly different distribution and would average less than half the values given in Figure 2-32. If we assume background is about the same (do you have any data to indicate this is not about right?) then the current "Percent Background Contribution" would average about twice the values given in the Figure. What would this imply for the percent due to background in the western states? Some thought and comment should be given to these considerations, particularly the statements at the end of the last paragraph on page 2-76.

Page 276 line 23. Is "absolute" that same as "total"?

Page 276, line26. The apparent impact of volcanic emissions on background extends quite east of the west coast.

Figure 2-33 is missing.

*How informative is the analysis of correlations between $SO_2$ and co-occurring pollutant concentrations for interpretation of epidemiological studies?*

I am not certain why EPA included both Max and 24-hour average co-occurring pollutant comparisons in Figure 2-35. I assume Max means the 1-hour maximum concentrations of the co-pollutant during the 24 hour period for which $SO_2$ daily average concentrations were obtained. It could also mean the co-pollutant concentration during the hour for which the highest $SO_2$ concentrations were seen. In either case, the Max comparison only makes sense if you expect the sources and emission diurnal patterns of

$SO_2$ and the co-pollutant to be the same. EPA needs to better explain what Max means, and how you justify including the comparison, or just drop it.

A similar statement applies to Max in Figure 2-36. There I am even less sure of what Max refers to.

In addition, I am not sure what is meant by $O_3$ Max, and why no average is included, even in Figure 2-36 which includes hourly $SO_2$ data. A correlation here, of course would be quite surprising.

The same comment I have made on Figure 2-35 applies to the seasonal data in Figure 2-37.

Page 2-83, line 2-5. I have frequently used ozone hourly average ozone data in source apportionment analysis for both summer and winter studies and find the expectation of substantial measurement error in these data in the winter to not be justified.

Perhaps the wisest statement in this section is we do not expect co-pollutant correlation confounding in $SO_2$ focused epidemiological studies, but you better look for it anyway.

I have made no comments on Section 2.6.1 Dispersion Modeling, not because I do not think it important, but because I am not knowledgeable on the subject.

**Comments on Executive Summary**

Sources and Human Exposure to Sulfur Dioxide

**Page xxxvii, line 20.** If EPA agrees with and adds material related to my discussion of the potential importance of metal S(IV) species, the confounding effects of particulate S(IV) compounds should be added to this sentence.

**Page xxxvii, line 29.** Likewise the non-photochemical formation of particulate S(IV) species including metal S(IV) species and bis-hydroxymethyl sulfone should be added here.

**Page xxxvii, line 4.** Likewise, smelters and integrated steel mills should be added to the anthropogenic sources.

**Comments on Chapter 1 - Summary of the Integrated Science Assessment**

1.4.1 Emission Sources and Distribution of Ambient Concentrations

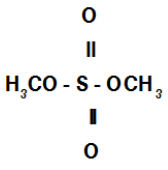
**Page 1-7, line 13.** The removal of $SO_2$ by reaction pathways other than the formation of sulfate, as summarized in the comments on Chapter 2, should be acknowledged and summarized.

**Page 1-7, line 34.** Smelters and integrated steel mills should be added to the anthropogenic sources.

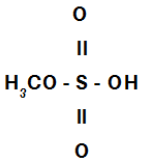
**APPENDIX:**

In this Appendix to my comments I summarize material on the conversion of $SO_2$ to compounds other than sulfate given in my previous April 29, 2014 comments. I also expand on the material given there to discuss the relative importance of the S(IV) and organic S(VI) species to sulfate in plumes from smelters, steel mills and power plants and comment on the chemistry which seems to be involved in the formation of these compounds in the atmosphere.

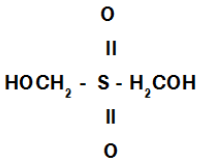
We have published a chapter which summarizes the information available through 1983 on Particulate Inorganic S(IV) and on Bis-Hydroxymethyl Sulfone, an organic S(IV) compound which are formed in emissions from $SO_2$ sources (Eatough 1983). Much of the information in the next two sections is found in that chapter. I have appended that chapter to my covering e-mail and can make a copy available on request. The other two oxysulfer compounds we have identified, dimethyl sulfate and mono methyl sulfuric acid are S(VI) compounds and they are briefly discussed later..




**Dimethyl Sulfate**




**Monomethyl Sulfuric Acid**




**Bis-Hydroxymethyl Sulfone**

**Particulate Inorganic S(IV).**

The early laboratory studies of Amdur (1971) and Alarie (1973) indicated that exposure to both $SO_2$ and metal oxides present in smelter emissions resulted in an enhanced animal response and that the resulting aerosols were irritating. Postulating that this work might reflect the presence of stable metal sulfite species in the aerosols studied, and that the formation of such aerosol species, rather than aerosol sulfate, might explain the results of the EPA CHESS study in the Salt Lake City environment with substantial impact from Cu smelter emissions (EPA 1974), we conducted studies on S(IV) associated with ambient aerosols. This work demonstrated that stable transition metal ion - sulfite species existed in aerosols associated with smelter emissions (Smith 1976, Eatough 1979, Eatough 1980) and formed in aging smelter plumes (Eatough 1981a, Eatough 1982). Studies in several smelter and coal-fired power plant plumes indicated the inorganic sulfite species were present in both primary emission and were also formed in the plume (Eatough 1983). The inorganic S(IV) was most clearly associated with the sum of Fe and Cu in the aerosols (Eatough 1983, 1981c) and was inversely related to aerosol acidity (Eatough 1983). The inorganic S(IV) species were present at from 10 to 30 mol % of the sulfate species in these smelter plume associated aerosols. The sulfite species were less important in urban or coal-fired power plant plumes in the east being from 5 to 10 mol % of the sulfate (Eatough 1978) but varied from 40 mol % in the fresh plume and present at concentrations comparable to sulfate in the aged plume of a coal-fired power plant in Utah (Eatough 1081a, 1983) Concentrations were generally low in oil-fired power plant plumes and varied from 8 mol % of sulfate in New York to 1 mol % of sulfate in Los Angeles (Eatough 1983). We also demonstrated that stable Fe(III) –S(IV) aerosols could be routinely generated in the laboratory (Hilton 1979).

In summary, the measurement, stability and formation of these inorganic S(IV) species in aerosols has been reviewed (Eatough 1983). These S(IV) species were present in emissions from all smelters studied, were also present in emissions from coal-fired power plants and additional material was formed during plume transport in both coal-fired power plant and smelter plumes. Based on the chemistry identified, we would expect to find these metal S(IV) compounds in emission from integrated steel mills. The amount of the S(IV) species, relative to sulfate average 0.1 mol S(IV)/mol sulfate in the coal-fired power plant plumes and 0.5 mol S(IV)/mole sulfate in aged smelter plumes. The formation of S(IV) in smelter plumes increased with decreasing acidity of the aerosol.

***Bis-Hydroxymethyl Sulfone.***

Several different methods of analysis of particulate samples collected from the plumes of coal-fired power plants or from areas heavily impacted by coal-fired boilers indicated that a S(IV) compound distinctly different from inorganic S(IV) was present in the samples (Eatough 1978, Eatough 1981, Richter 1981). This compound was subsequently identified as bis-hydroxymethyl sulfone (Eatough 1984). The sulfone was usually present in emissions from coal- or oil-fired power plants at mol ratios of about 0.5 (range of 0.1 to 1.0) compared to sulfate (Eatough 1983). First order formation of the sulfone was observed in plumes from six different power plants at rates of from 0.4 to 3.0 % $SO_2$/hr. with the observed rate being inversely proportional to atmospheric water partial pressure (Eatough 1983). The sulfone was found in highest concentrations in the Los Angeles Basin in inland samples (mol fraction comparable to sulfate), but was not seen in coastal samples impacted by fog or clouds (Farber 1982). No information is available on the toxicology of bis-hydroxymethyl sulfone.

In the previous April 29, 2014 comments I made the following recommendation:

"A potentially enlightening exercise might be to examine if any relationship exists between the results of epidemiological studies and the source of $SO_2$ for a given epidemiological study. This suggestion is based on the early laboratory studies of Amdur (1971) and Alarie (1973) which indicated that exposure to both $SO_2$ and metal oxides present in smelter emissions resulted in an enhanced animal response and that the resulting aerosols were irritating. I suggest this because it might enlighten whether particulate S(IV) (e.g. absorbed $SO_2$) might be important in exacerbation of asthma. If these S(IV) containing aerosols identified in the above reviewed research account for the enhanced effect of $SO_2$ in the presence of transition metal containing aerosols in animal exposure studies, then this class of compounds may be important in the interpretation of the morbidity effects associated with exposure to pollution from smelter or integrated steel mill sources. A careful review of pertinent epidemiological literature may inform this postulate.

The current set of counties which are nonattainment with respect to the current $SO_2$ NAAQS will probably not provide the needed information. A review of nonattainment counties with populations near or over 100,000 show that with two exceptions, the $SO_2$ exposures are dominated by emissions from coal-fired power plants, where aerosol S(IV) species may be less important. The two exceptions are Jefferson County, MO where about 30% of the $SO_2$ emissions currently are from the Herculaneum Lead Smelter, with the remainder being from coal-fired power plants, and the Steubenville, Weirton region in eastern Ohio and western Pennsylvania, where emission from the Weirton Steel are likewise, a minor portion of the $SO_2$ emissions in the immediate area, with coal fired power plants being more important.

These two locations would only stand out from the other nonattainment areas if the morbidity influence of aerosol S(IV) species was much greater that that associated with $SO_2$ itself. In addition to being a nonattainment area, Jefferson County, MO was also highlighted in the September 2008 Integrated Science Assessment for Sulfur Oxides, but with no epidemiological discussion associated with this nonattainment area (part of the St. Louis MO MSA).

Probably a more fruitful set of data to evaluate the relative importance of aerosol inorganic S(IV) species associated with smelter emissions would involve past epidemiological studies from about two to three decades ago when smelter emission were much more significant, for example from the TX smelters in El Paseo (ASARCO Cu smelter, closed in 1999), and Corpus Christi (ASARCO Pb smelter, closed in 1985), Az smelters (ASARCO Cu smelter in Hayden, currently operating and Phelps Dodge Cu smelter in Douglas, closed in 1987), from the Kennecott Cu smelter in Magna, UT prior to construction of the tall stack, from the Tacoma WA smelter (American Smelting and Refining, a Cu smelter specializing in high As ore refining, closed in 1985), or the smelters in Montana (ASARCO Pb smelter in East Helena, closed in 2001, Anaconda Cu smelter in Anaconda, closed in 1981) and Idaho (Bunker Hill Pb smelter in Kellogg, closed in 1982). I know that several epidemiological studies were conducted at these locations, but I am not familiar with the results of these studies with respect to asthma exacerbation. I recommend that EPA look at this older data to see if an estimate of the relative potency of $SO_2$ and smelter associated aerosol S(IV) species can be determined. There will not be data on the concentrations of inorganic S(IV) in the aerosols emitted from these sources, so total particulate exposure or $SO_2$ would need to be used as a surrogate. The importance of elucidating the effect of these exposures is correctly alluded to in the ISA on Page 4-12, Line 11."

This recommendation was included in the final comments (albeit much more briefly) from the committee to EPA, but has not been acted upon are near as I can tell. In light of the importance of emissions from smelters (infrequent, but not nonexistent) and integrated steel mills, I believe the exercise would still be valuable in determining if aerosol inorganic S(IV) is a confounder to the $SO_2$ asthma response seen in epidemiological studies.

**Organic Oxysulfur Compounds in the Atmosphere**

This section was added to my preliminary comments to provide a written response to the question raised in the preliminary comments by Dr. Daniel Jacob. Compounds identified by BYU and discussed below include:

***Alkyl Sulfates.*** We have previously identified monomethyl sulfuric acid and dimethyl sulfate in power plant plumes (Lee 1980, Eatough 1981b, Hansen 1987) and in the Los Angeles Basin (Eatough 1986, Hansen 1986). The alkyl sulfates have been shown to be present in emissions from both coal and oil-fired power plants (Eatough 1981b). In addition, formation of dimethyl sulfate during plume transport has been seen in the plumes of both an oil-fired and a coal-fired power plant (Hansen 1987). Dimethyl sulfate did not form in the plume of the oil-fired power plant studied while it resided in a fog bank, but formation was seen after the plume exited the fog bank. The rate of conversion of $SO_2$ to dimethyl sulfate was about 0.4 mole %/hr in the oil fired power plant plume and about 0.05 mole %/hr in the coal fired power plant plume. Particulate phase dimethyl sulfate dominated in these two studies. In the Los Angeles Basin studies (Eatough 1986) dimethyl sulfate was only seen in air masses not imbedded in a fog bank, i.e. generally in the inland area. Gas phase dimethyl sulfate was the dominate species in these

studies, was present at highest concentrations in transported plumes in the Inland Empire and was seen to exceed 10 mole % of the total sulfur oxides present. It should be pointed out that at the time of these studies, substantial $SO_2$ emission from power plants were present in the Basin..

Dimethyl sulfate is a mutagen and suspected human carcinogen, so its presence may be important with respect to toxic species, but I am not aware of any data indicating that inhalation will exacerbate asthma.
.

***Other Organic Oxysulfur Compounds.*** Aerosol phase methane sulfonic acid (Panter 1980) and gas phase ethylene sulfite (Jones 1974) have been identified in atmospheres impacted by emission from coal fired power plants, but only at concentrations much less than the above described species.

**References**


Alarie Y., Wakisaka I. Oka S. (1973) "Sensory Irritation by Sulfite Aerosols," Environ. Physiol. Biochem., **3**, 182-184.

Amdur M.O. (1971) "Aerosols Formed by Oxidation of Sulfur Oxide." Arch. Environ. Health, **23**, 459-468.

Eatough D.J., White V.F., Hansen L.D., Eatough N.L. and Cheney J.L. (1986) "Identification of Gas-Phase Dimethyl Sulfate and Monomethyl Hydrogen Sulfate in the Los Angeles Atmosphere," Environ. Sci. Technol., **20(9)**, 867-872.

Eatough D.J. and Hansen L.D. (1984) "Bis-Hydroxymethyl Sulfone: A Major Aerosol Product of Atmospheric Reactions of $SO_2$ (g)," Sci. Total Environ., **36**, 319-328.

Eatough D.J. and Hansen L.D. (1983) "Organic and Inorganic S(IV) Compounds in Airborne Particulate Matter," in Trace Atmospheric Constituents: Properties, Transformations and Fates, **12**, S.E. Schwartz, ed., John Wiley and Sons, New York, NY 221-268

Eatough D.J., Christensen J.J., Eatough N.L., Hill M.W., Major T.D., Mangelson N.F., Post M.E., Ryder J.F., Hansen L.D., Meisenheimer R.G. and Fisher J.W. (1982) "Sulfur Chemistry in a Copper Smelter Plume," Atmos. Environ., **16**, 1001-1015.

Eatough D.J., Richter B.E., Eatough N.L., and Hansen L.D. (1981a) "Sulfur Chemistry in Smelter and Power Plant Plumes in the Western U.S.," Atmos. Environ., **15**, 2241-2253.

Eatough D.J., Lee M.L., Later D.W., Richter B.E., Eatough N.L. and Hansen L.D. (1981b) "Dimethyl Sulfate in Particulate Matter from Coal- and Oil-Fired Power Plants," Environ. Sci. Tech., **15**, 1502-1506.

Eatough D.J. and Hansen L.D. (1981c) "S(IV) Chemistry in Smelter Produced Particulate Matter," Amer. J. Indus. Med., **1**, 435-448.

Eatough D.J., Eatough N.L., Hill M.W., Mangelson N.F., Ryder J., Hansen L.D., Meisenheimer R.G. and Fischer J.W. (1979) "The Chemical Composition of Smelter Flue Dusts," Atmos. Environ., **13**, 489-506.

Eatough D.J., Major T., Ryder J., Hill M., Mangelson N.F., Eatough N.L., Hansen L.D., Meisenheimer R.G. and Fischer J.W. (1978) "The Formation and Stability of Sulfite Species in Aerosols," Atmos. Environ., **12**, 263-271.

EPA (1974) “Health Consequences of Sulfur Oxides: A Report from CHESS, 1970-1971” EPA-650.1-74-004, May 1974

Farber R.J., Huang A.A., Bregman L.D., Mahoney R.L., Eatough D.J., Hansen L.D., Blumenthal D.L., Keifer W.S. and Allard D.W. (1982) "The Third Dimension in the Los Angeles Basin," Sci. Total Environ., **23**, 345-360.

Hansen L.D., Eatough D.J., Cheney J.L. and Eatough N.L. (1987) "The Formation of Dimethyl Sulfate in Power Plant Plumes," 80th Annual Meeting of the Air Pollution Control Association, Paper No. 87-96.3, New York, New York, 21-26 June.

Hansen L.D., White V.F. and Eatough D.J. (1986) "Determination of Gas-Phase Dimethyl Sulfate and Monomethyl Hydrogen Sulfate," Environ. Sci. Tech., **20(9)**, 872-878. **[362**

Hilton C.M., Christensen J.J., Eatough D.J. and Hansen L.D. (1979) "Fe(III)-S(IV) Aerosol Generation and Characterization," Atmos. Environ., **13**, 601-605.

Jones P.W. (1974) Proceedings of the 1974 APCA Meeting, Paper No. 74-25.

Lee M.L., Later D.W., Rollins D.K., Eatough D.J. and Hansen L.D. (1980) "Dimethyl and Monomethyl Sulfate: Presence in Coal Fly Ash and Airborne Particulate Matter," Science, **207**, 186-188.

Panter R. and Penzhorn R.D. (1980) “Alkyl Sulfonic Acids in the Atmosphere”, Atmos. Environ., **14**, 149-151.

Richter B.E. (1981) “S(IV) and Alkylating Agents in Airborne Particulate Matter,” Ph.D. dissertation, Brigham Young University, Provo, UT.

Smith T.J., Eatough D.J., Hansen L.D. and Mangelson N.F. (1976) "The Chemistry of Sulfur and Arsenic in Airborne Copper Smelter Particulates," Bull. Env. Contamination and Toxicology, **15**, 651-659.

# Dr. H. Christopher Frey

**Comments on Chapter 3**

a. *To what extent is the discussion on methodological considerations for exposure measurement and modeling clearly and accurately conveyed, appropriately characterized, and relevant to the review of the $SO_2$ NAAQS?*

EPA should consider developing a glossary of terms that are used not just in this ISA, but in multiple ISAs, to help enforce consistent and correct use of terminology. For example, the term "exposure" is used correctly in many places, and is correctly defined in the context of inhalation exposure in Equation 3-1. However, "exposure" is later used incorrectly to imply that somehow it accounts for inhalation rate related to activity levels. This implied definition of exposure is not consistent with earlier text. Estimation of the product of exposure concentration and inhalation rate would lead to a 'potential dose,' rather than an "exposure."

A glossary would be useful in keep track of various terms that are complex yet used often in the text. For example, the term "exposure error" is used, but it was difficult to find a good definition of this term in Chapter 3. The term was eventually explained to the reader, but it would help the reader if there was a "go to" place to find the definition, given its repeated usage. Other terms that merit definitions include but are not limited to "exposure bias," "classical error," "Berkson error," "minimum detected level," "ambient concentration," "exposure concentration," "microenvironment," "indoor," and others.

The draft Chapter 3 is somewhat difficult to read. It is not clear in many places. It is repetitive. It is not well organized not only because of repetition but because there is not a clear linear flow of information.

The introduction to the chapter should define more clearly the scope in terms of the exposure scenario that is addressed throughout. In general, exposure can be via multiple routes (inhalation, ingestion, and dermal contact) and for each route there can be multiple pathways. It would help the general reader if this was acknowledged and if a rationale was given as to why the focus of this document is only on the inhalation exposure route.

The scenarios related to exposure pathways are mentioned in passing throughout the chapter, which implies that it would help to have an upfront section that clearly specifies the exposure pathways of interest. These include point sources with elevated stacks (e.g., power plants). The text occasionally refers to near-ground level releases from vehicle sources. Perhaps there are others.
Some context regarding the relationship of the selected exposure routes and pathways for the purpose of health effects assessment pertaining to ambient $SO_x$, with respect to secondary pollutants such as sulfate, would be useful. It would be helpful to the reader to understand, for example, that the formation of secondary sulfate and the direct and indirect human and ecological exposures related to sulfate are not in the scope of this document or review but are handled in other reviews, such as for the primary and secondary PM standard and the secondary $NO_x$ and $SO_x$ review.

The introduction to the chapter should also lay out the relationship between exposure modeling and the overall risk assessment methodology. For example, to the extent that dose-response relationships are

based on controlled clinical studies, the exposure assessment should focus on estimating exposure concentrations. To the extent that the dose-response relationships are based on epidemiological studies that are predicated on fixed site monitoring data, then the exposure assessment would typically focus on ambient concentration as surrogate. This context is not clearly stated and would help the reader understand the scope of this chapter.

A general comment is that this draft chapter often contains assertions that should be supported by reference to literature, but no reference is given. See detailed comments for examples.

The text frequently refers to "$SO_2$" when it should instead be more clear and specific: e.g., $SO_2$ ambient concentration, $SO_2$ exposure concentration.

When referring to a concentration (e.g., ambient, exposure), it is critical to clearly communicate the averaging time regarding the concentration. Specification of the averaging time should be done consistently throughout.

For Section 3.2.1.2 the reader would be grateful for a summary table of the detection methods, their precision, accuracy, averaging time(s), detection limits, interferents, and so on. Also, an indication of their state of development (e.g., well established, demonstrated, in development) would be helpful.

Section 3.2.2 on modeling should be more clear regarding the distinction between modeling ambient concentration versus exposure concentration versus exposure.

Table 3-1 should be revised to more clearly distinguish models for ambient concentration versus models for exposure concentration versus models for exposure. Models for ambient concentration are not typically considered to be "exposure models" but rather they are used in situation for which ambient concentration is used as a surrogate for exposure.

Section 3.2.2.1 should, with better clarity, distinguish between estimated ambient concentration versus estimated exposure concentration.

The discussion of source proximity models begs the question about quantification of emissions. The document should be more clear about the source categories of interest, the elevation at which emissions are released (e.g., tall stacks for power plants, near ground level for vehicles, and so on), and provide context regarding the apportionment of SO2 emissions among source categories. For example, there should be a table that summarizes the National Emission Inventory for $SO_2$ at a national level to help the reader understand the average relative contribution of key source categories. Additional example(s) of localized inventories, such as in proximity to a particular central site monitor, would help provide insight regarding the significance of elevated sources, ground level sources, etc. For example, on page 3-7, lines 21-35, the reader wonders what are the key sources? What is the role of averaging time, spatial resolution, source classification, and completeness of the emission inventory?

The document should always be clear as to whether a reported concentration is measured or estimated using a model (and regarding the averaging time and spatial resolution of the reported concentration).

The comparison of models is a good start but needs further development. For example, in the context of page 3-9 as an example, there should be a more extensive and structured framework for comparing modeling approaches that takes into account factors such as what sources are considered, what receptors are considered, spatial scale, averaging time, goodness of fit metrics, and so on.

Gaussian plume models (e.g., on page 3-10) and Chemical Transport Models (page 3-11) require input data regarding emission sources. As noted above, the document currently lacks adequate information regarding $SO_2$ emissions inventories and regarding source characteristics that are relevant to either the input or interpretation of various modeling approaches. For example, even though LUR models may not require input data regarding emissions, they are based on emission surrogates and interpretation of their results requires knowledge of emissions.

The discussion of CTMs should be more structured. CTMs typically have four main components: (1) emissions, (2) chemistry, (3) meteorology, and (4) initial and boundary conditions.

The text should be clearer to distinguish between models, measurements, and reality. Sometimes statements are made as if ground truth was obtained from a model or measurement, when this may not be the case. As an example, on page 3-12, lines 27-28, a statement is made about cloud-based reactions as if these are known with absolute certainty, but instead it needs to be mentioned that the results are based on a model, not absolute truth.
The text should be careful when referring to "any" – e.g., "so that any inherent biases are understood." It may not always be possible to exhaustively identify, much less understand, "any" (and presumably, all) sources of bias.

With regard to CTMs, the text should be more clear that grid-based models generally underestimate localized peak concentrations because of spatial averaging within a grid cell. One approach to deal with this, plume-in-grid, is mentioned. However, another approach is the used of nested grids, which is not mentioned but should be. Furthermore, the role of the height at which emissions are released should be treated more clearly in the document. On page 3-13, first paragraph, the text implies (perhaps unintentionally) that emissions released in the ground layer of a CTM grid are perhaps of more significance to exposure than those released in higher layers of the grid, such as from stacks of power plants. However, the text does not mention that coal power plants typically release far greater SO2 emissions than other sources (e.g., diesel vehicles) and that the higher point of release also leads to longer range transport and dispersion. Thus, the plume from a power plant can affect a larger downwind population than a plume from a vehicle. These types of distinctions are important when interpreting data from air quality monitors and the results from models, and thus merit discussion.

The text could state or infer implications of the review of modeling approaches with regard to model selection.

CHAD is an extremely important resource and merits a more focused and a bit more detailed treatment than what exists in this draft. CHAD is mentioned in several places, but could instead be reviewed in one part of the document to avoid repetition. Instead of repetition, the treatment of CHAD should be more specific. It would help to have a table that lists the key studies that comprise CHAD and their relevant characteristics – e.g., dates, number of individuals, number of diary days, demographic characteristics of the individuals, etc.

The use of newer methods for obtaining diary data is an important topic and should not be part of a run-on paragraph that includes other topics. However, it would be useful to report whether there are data now available from these newer techniques that would be useful in the current review of the SO2 NAAQS. It is not clear whether data are sufficiently established from these newer techniques to be relevant (e.g., for the REA).

The discussion of air exchange rates is too narrow and not well motivated or organized. Entirely missing is mention of databases of measured air exchange rates. The text could be more systematic in reviewing air exchange rates not just for "buildings" but for other enclosed microenvironments, such as vehicles. The text on page 3-14, lines 25-32 seems to describe a very complex model that arguably does not actually exist. No references are cited to support this text. Some statements in this paragraph are dubious. Making a model more complex does not always, in practice, "reduce the uncertainty" and to state this as if it is a truth is misleading. Adding more variables increases input data requirements, and each variable may have associated uncertainty. The more complex model may also have model uncertainty.

The role of "microenvironmental exposure models" in the last review of the $SO_2$ NAAQS is mentioned only in passing. To the extent that models such as APEX and SHEDS are likely to be relevant to the current review, they merit more treatment than given in this draft. These are also referred to as "stochastic population-based models."

A statement is made that models such as SHEDS or APEX "are not used for exposure assessment in epidemiologic studies." While this is typically the case, the statement as given is false. There are some examples in which models such as SHEDS and APEX have been used to evaluate the utility of estimated exposure estimates, rather than fixed site monitoring or CTM estimates, in explaining variability in adverse effect endpoints. Much of this work has been done for PM rather than $SO_2$, but methodologically these PM examples are relevant. Two examples are Chang *et al*. (2012) and Mannshardt *et al.* (2013). There are other examples.

On page 3-15, the text states that "the state of science for stochastic population exposure models has not changed substantially since the 2008 ISA for Oxides of Nitrogen". This statement is very puzzling. APEX was used extensively in the most recent review of the $O_3$ NAAQS, and various updates to the input data and the model were incorporated during this more recent review process.

Similar to the comments on CTMs, the discussion of stochastic population based exposure models can be more clearly structured. The key inputs to these models include ambient concentration, population demographic data (i.e. location of population and distribution of population by census tract), activity diary data, and data regarding infiltration to enclosed microenvironments. The latter can depend also on housing stock. Each of these merits discussion, but to the extent a topic is already treated in detail earlier in the document, that material can be cross-referenced without repetition.

The term "indoor" is a bit confusing in that it does not connote in-vehicle microenvironments but apparently is defined to include these.

There is no transition from

The overall organization of this chapter is hard to follow.

*b. Please comment on the accuracy, level of detail, and clarity of the discussion regarding exposure assessment and the influence of exposure error on effect estimates in epidemiologic studies of the health effects of $SO_2$.*

This charge questions appears to be aimed especially at Section 3.3 "Exposure Assessment and Epidemiologic Inference" and its constituent parts, including 3.3.5 "Implications for Epidemiological Studies of Different Designs". The title of this section is awkward. This section is itself awkward and hard to follow at times.

A glossary would help the reader with this section of the document, since many key terms are used frequently.

The content was generally reasonable, but the writing could be improved.
Section 3.3.1 defines exposure. As such, this material belongs earlier in the document.

Section 3.3.2 focuses on infiltration and as such could be combined with an earlier section that brings up this topic. The treatment of the infiltration factor would be improved by systematically dealing with evidence for Finf and with evidence separately and individually (to the extent such evidence exists) for P, a, and k. In particular, k is barely treated in this chapter but is quite significant for $SO_2$.

Section 3.3.2.2 the term "Personal-Ambient Relationships" is unclear.

Section 3.3.3.1 on "Activity Patterns" co-mingles two different issues. One is the amount of time spent in each microenvironment. The other is the activity level of an individual at a given time or location. These should be treated separately. Furthermore, the "level of activity" is not given proper context. Level of activity affects inhalation rate. The product of exposure concentration and inhalation rate is a potential dose. Up to this point, no modeling approaches have been discussed that attempt to estimate potential dose. Thus, the context of introducing "level of activity" is not very clear. It is also very misleading to imply that "activity pattern" includes both time in microenvironment and "level of activity" – i.e. it is not correct to state that Equation 3-3 includes both of these.

To the extent that the ISA should address level of activity to help provide insight regarding relationships between exposure concentration and adverse effect, then the ISA should treat this topic separately and more clearly. Inhalation rate is related to metabolic rate, and there are some modeling approaches that are predicated on metabolic rate (e.g., see work of Tom McCurdy and others) that could be mentioned here. The concept of "potential dose" should be introduced. Potential dose may be a more accurate surrogate of adverse effect than exposure concentration. However, there can be errors in estimation of "level of activity" and inhalation rate such that an attempt to model the sensitivity of inhalation rate to level of activity can potentially introduce both random and systematic errors to an analysis.

Figure 3-1 is a very important figure, but the reader wonders whether this can be updated with the most recent version of CHAD and whether there are diary data in CHAD more recent than the 2001 publication date of the figure. My understanding is that CHAD does include more recent diary data.

There have been some questions, for example, in prior CASAC reviews of other criteria pollutants as to whether activity patterns may be changing as a result of changes in lifestyle.

On Page 3-39, it is stated that "exertion level is an important determinant of exposure." This is not correct: exertion level is an important determinant of potential dose, but would not affect exposure concentration.

Page 3-41: CHAD is mentioned again. Instead, have one place in the chapter that focuses on CHAD and either cross-reference to that or augment it only as needed in other places. The information given here is quite out of date. For example, the number of diaries now in CHAD may be larger than mentioned here. Exposure error is mentioned in many places. The reader would appreciate one clearly indicated section in which exposure error is addressed thoroughly. Other sections of the chapter could cross-reference to this. In general, repetition should be avoided. Exposure error is mentioned many times but a deeper attempt to define it does not appear until late in the document.

The summary was not very helpful. Less repetition and more integration/synthesis would be helpful.

Other Comments

Page 3-3, lines 21-38: some context regarding these measurement methods is needed. Which have actually been used in practice? Which are methods that are in development? The paragraph reads a bit like a laundry list and would benefit from more discussion of each technology raised and its context (e.g., microelectrodes using nanomaterials – are these actually deployed in use, are they in development but not yet demonstrated, etc.).

Page 3-8, line 7: "it" is unclear… LUR?
Page 3-8, first paragraph – please be more clear about averaging times.
Page 3-8, line 23, please specify the key source categories that are "industrial" sources.
Page 3-8, 34-35: please indicate the mean concentration so that the reader can put these numbers into context.
Page 3-8, line 37: when reporting that a distribution is skewed, it would be helpful to give supporting information such as the coefficient of variation and the skewness coefficient.

Page 3-9, lines 18-19: This statement is unclear with regard to epistemic foundation. Is the finding that SO concentration "was not influenced by" the sources mentioned based on a model or based on statistical analysis of measurements, or other?

Page 3-9, line 24: "when diesel fuel still contained sulfur" is a very odd statement. Diesel fuel continues to contain sulfur now. However, the allowable maximum sulfur content has changed. The ISA should be much more clear regarding what the allowable sulfur content was and is for highway diesel and nonroad diesel, and the implications of this. As noted in an earlier comment, there should be a section of this chapter that deals with the emission inventory.

Page 3-10, line 16: bicubic spline functions *for what*?

Page 3-10, line 19: IQR *of what*?

Page 3-10, line 33: “industrial/vehicle” is unclear.

Page 3-11, lines 3-4: the reported correlation with vehicle emissions requires some more explanation and interpretation. For example, were these onroad diesel or nonroad diesel vehicles? Where there some location-specific features that lead to the estimated correlation. It is also not very clear as to exactly what was correlated here and the averaging time that was the basis for the correlation.

Page 3-11, line 36: what averaging time?

Page 3-12, line 35: “vertical emissions” is an awkward term and not one that I have seen before. Perhaps this refers to emissions released at an elevated height from a stack. The meaning should be more clear.

Page 3-13, lines 21-22: This may be true, but what is the basis for this claim? Also, define “microenvironment” and provide support from the literature as to why time spent in microenvironments is important. For $SO_2$, what is known from prior work regarding the relative importance of microenvironments? If there are not good examples for $SO_2$, then give examples for other criterial pollutant(s) to at least motivate the importance of this topic. For example, explain how does microenvironmental concentration differ by microenvironment? An example or two would help.
Page 3-14, line 15: “could be”? is this actually being done?

Page 3-16, first paragraph: no references are cited here, but should be. This text cites lengthy tables, but lacks discussion or interpretation of these tables. Such discussion and interpretation should be included.

Table 3-2: consistently identify the averaging time and spatial scale for each approach. There is room to spell out the model names that are abbreviated in the first column. The term “exposure bias” is used in some of the commentary but this term is not clearly defined. How it is different from “exposure error”? The table should also cite references for examples of these approaches. The table should also be more clear as to whether the approach is aimed at ambient concentration (environmental concentration), versus exposure concentration, versus exposure.

Microenvironmental models are not applicable only to panel studies – see comment above about their role in epidemiological studies.

Table 3-3: please define/explain the difference between “local,” “urban,” and “regional.” The column on exposure model type seems rather complicated but why not more clearly define the model types (i.e. in Table 3-2) and then use a more terse notation here that is also more clear.

Page 3-26, line 5: “this framework” – there is no prior framework given, so the antecedent here is undefined.

Page 3-26, line 17: be clear as to whether “indoor” includes all enclosed microenvironments such as vehicles.
Equation 3-3: check if the last term is $C_{o,i}$ or should it be $C_o$?

Equation 3-4 – the discussion here should be clearer that in this case $C_o$ is the outdoor concentration in close proximity to the enclosed microenvironment or to the point of outdoor exposure.

Page 3-28, last paragraph. This paragraph is awkwardly worded and hard to follow. It is also seems to be inaccurate. Line 22 is particularly confusing. A negative concentration is not physically possible and cannot be explained with respect to a detection limit.

Page 3-29, line 13: how do personal activities affect air exchange rate? This may be possible but is not explained and, therefore, is unclear.

Page 3-36, first paragraph. The studies by Sarnat et al. that are cited here are not listed in Table 3-4. It is not clear as to why they do not appear in the table. What is the implication of the information in this paragraph for the ratio of indoor to outdoor concentration? Line 15: "noise" (random error?) in measurements would reduce the magnitude of the correlation between measured concentration and ambient concentration, but this point is not clear.

Page 3-36, second paragraph – is this about the ratio of indoor to outdoor concentration?

Page 3-37, line 2: "large" relative to the mean measured concentration? Or "large" in some absolute sense?

Page 37, Section 3.3.3, first paragraph, there are a lot of issues with this paragraph in terms of lack of clarity. For example, lines 8-9, it is not clear what is really meant here… does this intend to refer to the idea that variability in concentration to which people are exposed, even in locations near the monitor, compared to the measured concentration at the monitor, can differ? This could be clearer. Line 14: "indoor reaction rate" – this issue has not been addressed and needs to be – at least a paragraph at an appropriate location in the document. Lines 16-18: "dose" is mentioned here but the distinction between exposure and dose has not been clearly stated. Also, what is meant by "extent" of exposure (define in glossary). This sentence, however, seems out of place here.

Page 3-40, line 6: what are "minute ventilations"? why are no references cited in this paragraph?

Page 3-40, lines 8-14. References need to be cited in this paragraph. The term "gender" appears to be understood to refer to gender identity, and thus is not the right term to use here. Che et al. (2014) analyzed the effects of different diary sampling methods on exposure estimates and this work may be relevant here.

Page 3-41, section 3.3.3.2. Text here is often unclear, and lacks adequate citations to literature.
Page 3-42, line 7: "for $SO_2$" – this is unclear. Does this refer to $SO_2$ *ambient concentration*? Goodness of fit with respect to what?

Page 3-42, last paragraph – this is ahrd to read. How can *pollutants* have error? Error in what, specifically? (pollutant concentration?). "Bias in $SO_2$" .. bias in what, exactly? Bias in the estimated relationship of health effect to ambient concentration? This paragraph needs a rewrite to remove awkward phrases, and to be more clear and precise regarding the basis of statements made.

Page 3-43: distance to a monitor is useful but not sufficient. For example, downwind location of plume "touchdown" would depend also on stack weight, wind speed, wind direction, atmospheric stability

Figure 3-2 “from to” in caption is confusing.

Page 3-47: Variations among seasons is not exactly the same as “temporal variability”. “Temporal variability” could imply hourly, diurnal, daily, weekly, etc. Seasonal is only one possible interpretation of this term. Inter-annual variability is probably not considered by many to be the same as “temporal variability” Line 18 “the two methods”? Prior text does not clearly define and identify each of two methods, so the antecedent here is unclear.

Page 3-48: Section 3.3.3.3 - the reader wonders why we have another section on infiltration since there was a prior section on infiltration.

Page 3-48, section 3.3.3.4 – “moderately correlated” needs to be defined quantitatively to avoid being vague. The discussion of implications of data below detection limit was also discussed quite a bit in the review of the draft ISA in the recent NOx review. Please be sure that the treatment of this issue is consistent between the SOx and NOx ISAs.

Page 49: up to 3.3.4 text here is sloppy. Concentration is not the same as exposure. Instrument error is not only because of interference, although this may be one type of error that is important in some cases. The text should distinguish between relative versus absolute errors as appropriate. The text lacks adequate citations to literature.

Page 49, Section 3.3.4 – “confounding” needs to be defined. The notion of effect modifiers should also be introduced and discussed. A clearer distinction between confounders and effect modifiers would be useful.

Page 50 – first paragraph, text switches between ambient concentration, exposure concentration, exposure, and effects – this is very confusing and these distinctions are not clearly made or well organized.

Page 50 – second paragraph: no references are cited (!). could write more about mechanistic expectation of collinearity or correlation, in terms of emissions, atmospheric processes. Averaging times need to be clearer.

Also, missing are some complex inter-relationships. For example, coal-fired power plants emit both $SO_2$ and $NO_x$ in often significant quantities. Thus, there is a common general exposure pathway to both of these pollutants, although the specifics differ in terms of atmospheric transformations. On the other hand, SO2 and NOx may be emitted also from vehicles, although the SO2 emissions from vehicles are usually small in magnitude.

Page 3-51: a correlation of 0.41 is described as “moderate” here, calling into question the meaning of other statements about moderate correlations that are unquantified elsewhere in the document. A correlation of 041 might be considered by some to be “weak.” Thus, definition of the meaning of these terms will help both the authors and the readers.

Page 3-51, lines 14-21: seems like this isn't likely to be adequately interpreted – e.g., howmuch of the data were below detection limits? What are the confidence intervals on the correlations? (Were the correlations statistically significant?).

Page 3-51, line 27: no basis is given (no reference is cited).

Table 3-9: Please clarify the averaging time – is this is what is meant by "duration"? What was the time frame of each of these studies? E.g., if they are based on daily averaging times, how many days of data were used? (or indicate sample size and the meaning of a sample – e.g., n=365 daily averages). To the extent possible, please indicate if reported correlations are statistically significant. Some insight regarding the proportion of data below detection limit would be helpful.

Page 3-58, section 3.3.4.2: what are the implications and relevance of this information?

Page 3-58, Section 3.3.5: Implication of *what* for epidemiologic studies? Here, we finally get a definition of exposure error and its components. Yet, some text here seems repetitive of earlier text. References should be cited for classical and Berkson error. More clarity on errors related to bias versus errors related to imprecision would be helpful.

Page 3-62, Table 3-10: This table is rather difficult to decipher. This appears to be a case study based on one paper. The meaning of the rows and columns should be clearer. For the columns, please include text descriptors of what is the intended purpose or meaning of each of these metrics. The rows labeled as averages are confusing- average of what for what location and time period? This table is not self-explanatory, but it should be. If all of this is model-based, what really is the point of including this in so much detail in the ISA?
Page 3-63, line 34: $SO_2$ is not a nonreactive pollutant.

Page 3-65, line 17: were these simulations for a reactive air pollutant? If not, how are these simulations "relevant to $SO_2$"? Not clear.

Page 3-66, line 4: clarify that the lack of spatial variability does not imply lack of variation over time.

Page 3-66, section 3.3.5.3: With regard to the first paragraph, while it is the case that stochastic population based models are "not designed" for comparison with panel studies, this does not mean that comparisons cannot be made. For example, comparisons could be made for the frequency distributions predicted by the model versus the frequency distribution measured in the panel study for the same averaging time, time period, and geographic scope. Such a comparison will be subject to imprecision. For example, the confidence intervals of the empirical frequency distribution from the panel study could be estimated using bootstrap techniques.

Section 3.4:
Page 3-69, line 12 – this is not correct.
Page 3-69, lines 18-27. Sloppy use of "$SO_2$" when the correct meaning may be "SO2 ambient concentration"

**References Cited**

Chang, H.H., M. Fuentes, and H. C. Frey, "Time Series Analysis of Personal Exposure to Ambient Air Pollution and Mortality Using an Exposure Simulator," Journal of Exposure Science and Environmental Epidemiology, 22, 483-488 (2012) doi:10.1038/jes.2012.53. Sept/Oct 2012.

Che, W., H.C. Frey, and A. Lau, "Assessment of the effect of population and diary sampling methods on estimation of school-age children exposure to fine particles," Risk Analysis, 34(12):2066-2079 (December 2014), DOI: 10.1111/risa.12238

Che, W., H.C. Frey, and A.K.H. Lau, "Comparison of Sources of Variability in School Age Children Exposure to Ambient $PM_{2.5}$," Environmental Science and Technology, 2015, 49(3):1511–1520

Mannshardt, E., K. Sucic, W. Jiao, F. Dominici, H.C. Frey, B. Reich, and M. Fuentes, Comparing Exposure Metrics for the Effects of Fine Particulate Matter on Emergency Hospital Admissions, Journal of Exposure Science and Environmental Epidemiology, 23:627-636 (doi:10.1038/jes.2013.39). (2013)

## Dr. Steven Hanna

Note that my expertise is primarily in atmospheric transport and dispersion modeling and analysis of observed concentrations, and my comments focus on those areas. However, I do have some experience in other areas of the report, such as emissions categorization and distribution, background concentration determination, and exposure modeling, and provide some comments on those topics where appropriate.

*Similarly, are there topics for which discussion should be shortened or removed? Does the Panel have opinions on how the document can be shortened without eliminating important and necessary content?*

I notice that there are whole paragraphs that are identical in the Executive Summary and in Chapter 1 – Summary of the ISA. I checked only the topics that I am familiar with. For example, the Chapter 1 paragraphs in lines 25-34 on p 1-7, lines 3-34 on p 1-8, and lines 1-14 on p 1-9 are nearly identical to what is found in the Executive Summary. An attempt should be made to avoid such long examples of duplication.

Also, the material on dispersion modeling in Chapter 2 is inconsistent with that in Chapter 3. I prefer the write-up in Chapter 2 and have several issues with what is in Chapter 3. Different people probably wrote those sections and they did not communicate enough with each other. The material on dispersion modeling in Chapter 3 could be eliminated and absorbed into the section in Chapter 2.

Another general comment is that the averaging times and sampling times and adjectives (peak, daily, etc.) need always to be clearly defined, and should be consistently used throughout the report. Often some key information is left out and the resulting statement or conclusion is ambiguous.

It is said that the uncertainty is what is not explained by models. The meteorological (and other environmental) communities have a different view – namely that there is some inherent or natural uncertainty in the system that can never be explained by models. For example, this inherent uncertainty in wind speed model estimates is about 1 m/s. For concentration predictions, the inherent uncertainty is at least 10 % even for the best of scenarios.

*Question 1: The Executive Summary is intended to provide a concise synopsis of the key findings and conclusions of the SOx ISA for a broad range of audiences. Please comment on the clarity with which the Executive Summary communicates the key information from the SOx ISA. Please provide recommendations on information that should be added or information that should be left for discussion in the subsequent chapters of the SOx ISA.*

See my earlier response regarding the fact that some paragraphs are identical in the Executive Summary and in Chapter 1.

*Question 2: Chapter 1 summarizes key information from the Preamble about the process for developing an ISA. Chapter 1 also presents the integrative summary and conclusions from the subsequent detailed chapters of the SOx ISA and characterizes available scientific information on policy-relevant issues.*

a. *Please comment on the usefulness and effectiveness of the summary presentation. Please provide recommendations on approaches that may improve the communication of key findings to varied audiences and the synthesis of available information across subject areas. What information should be added or is more appropriate to leave for discussion in the subsequent detailed chapters?*

It seems odd to have a summary chapter following the Executive Summary. Is this required? If so, please try to use different wording, since there is much redundancy.

P 1-7, line 27 – Insert "1-hour" before "maximum". This is just one example of the topic in paragraph 3 in my comments on the introductory material. Please be very clear about averaging times, peaks, etc. and use the same terms throughout the report.

*Question 3: Chapter 2 describes scientific information on sources, atmospheric chemistry, and measurement and modeling of ambient concentrations of gaseous sulfur oxides.*

*To what extent is the information presented regarding sources, chemistry, and measurement and modeling of ambient concentrations accurate, complete, and relevant to the review of the $SO_2$ NAAQS?*

*Please comment on the extent to which available information on the spatial and temporal trends of ambient $SO_2$ concentrations at various scales has been adequately and accurately described. In particular, what is the extent to which the analyses of recently available 5-min $SO_2$ concentration data are informative in considering relationships between 5-min and 1-hr $SO_2$ concentrations?*

*How informative is the analysis of correlations between $SO_2$ and co-occurring pollutant concentrations for interpretation of epidemiologic studies?*

This chapter (2) is on topics most closely aligned with my area of expertise (analysis of observed concentrations and dispersion modeling).

The siting criteria for $SO_2$ monitors seem rather arbitrary and are not consistent with EPA and NOAA and NRC guidance for other types of instruments. For example, in line 11 of p 2-32 – I do not see why the monitor should be placed on the windward side of a building, and the text gives no scientific justification. On p 2-33 line 1, I do not see how a building can "scavenge" $SO_2$. Trees and other vegetation may be able to scavenge $SO_2$, but most building materials are rather inert.

Examples of the misuse of terms related to averaging times and maxima are seen in section 2.5.1. I think that the current report is often comparing apples to oranges. Ensure in Fig. 2-28 that the 5-minute maximum concentrations for any hour correspond with the hourly average concentrations for that same hour (i.e., it should be impossible for data points to be below the 1:1 line). Since the current standard is based on the peak one-hour average concentration on a given day, then why not be comparing that number with the peak 5 min average concentration during that day? Note that the peak 5 min concentration could occur during a different hour in the day than the peak one-hour concentration. The basic theory (see Slade 1968 and Turner 1970) describes the variation of peak concentrations with averaging time during a given sampling period (e.g., the sampling duration may be one year and the peak concentrations may be determined for averaging times of 5 min, 1 hr, 1 day, and one month).

The ISA report authors should include discussions of the extensive historical literature on peak-to mean concentration ratios. The literature includes theoretical analysis (e.g., Gifford 1960) and analysis of observed peak to mean data (e.g., Slade 1968 and Turner 1970). The current analysis in Chapter 2 is presented as if the current ISA authors are the first to look at this subject. The literature from 50 years ago suggests that peak concentration is inversely proportional to averaging time ($T_a$) to the 1/5 power. Thus peak C (5 min)/peak C (1 hr) would be about 1.6. Figure 2-28 in Chapter 2 shows that hourly peak 5-min values are 1-3 times higher than the corresponding hourly values, with a peak-to-mean ratio (PMR) of 1.6 for 5-min averages. This is clearly consistent with the "0.2 power law" relating PMR to averaging time. Additional references on peak-to-mean concentrations include Dourado et al. (2012), Schauberger et al. (2012), and Venkatram, A. (2002).

I am further confused in Section 2.5.2 (spatial variability) where the quantities being studied are not clearly-defined. (see lines 4-16 on p 2-35). Also note Table 2-7 (and other tables), which need the entries to be completely defined.

The authors should review analyses of other environmental data, which usually show the long tailed distributions (with a small average and a few large maxima) as discussed, for example, at the bottom of p 2-36.

The section on urban mesoscale variability (2.5.2.2) should refer to 40-year-old EPA studies such as the EPA St Louis Regional Air Pollution (RAPS) study in the 1970s, which generated many detailed EPA reports and journal articles by what is now the ORD Air Modeling Group. I Googled EPA RAPS and ended up with a long list of EPA reports that the ISA authors can review. Also, rather than first running the statistical software and then trying to interpret the results, it would be better to first list a few basic facts or hypotheses that are expected to be verified by the data. For example, it is expected that concentrations will be relatively high for monitors closer to large sources (e.g., see bottom of p 2-46 and lines 6-9 on p 2-49). This can also be shown by application of a one line Gaussian plume model.

Seasonal trends discussion in Section 2.5.3.2 – Differences are also due to the effects of stack height (for large industrial sources) and plume buoyancy flux.

Section 2.5.3.3 – I cannot find a definition of deil and it is not listed in the acronyms.

Section 2.5.3.4 on 5 min data – See my comments on the previous page. Also, we know that for summer daytime periods, the 5 min fluctuations can be large since that is the period of convective eddies in the boundary layer. A stack plume can be brought down to the ground for 5 to 10 minutes, and then be carried upwards for the next ten minutes.

Section 2.5.4 on background – This is a difficult subject. I would like to see a discussion comparing the variability in background for observed concentrations and for modeled concentrations (e.g., by MOZART).

Section 2.6.1 – Dispersion Modeling – I found this section to be very good. I would like to see more discussion of Lagrangian puff models such as CALPUFF and SCICHEM, though, since they are very useful for the nearfield out to regional scales. They can handle winds and stabilities that vary in space and time. Note that I do wonder where the statement comes from on p 2-86 lines 11-12 that "the terms "dispersion model" and "Gaussian model" are associated with each other". I have never read this

anywhere. It is especially puzzling because the EPA should be aware of their own additional dispersion models such as CALPUFF, SCICHEM, and CMAQ.

Section 2.7 – Summary – Isn't $H_2SO_4$ a PM problem?

*Question 4 - Chapter 3 describes scientific information on exposure to ambient $SO_2$ and implications for epidemiologic studies.*

*To what extent is the discussion on methodological considerations for exposure measurement and modeling clearly and accurately conveyed, appropriately characterized, and relevant to the review of the $SO_2$ NAAQS?*

*Please comment on the accuracy, level of detail, and clarity of the discussion regarding exposure assessment and the influence of exposure error on effect estimates in epidemiologic studies of the health effects of $SO_2$.*

My comments focus on Section 3.2.2 Modeling. As I said in my initial comments, the dispersion model discussion here and that in Chapter 2 are not consistent. This report should present a consistent view, and I greatly prefer the discussion in Chapter 2. Please revise Chapter 3 so that it reflects the material in Chapter 2. For example, include CALPUFF and SCICHEM (Lagrangian puff models), which have more extensive chemical mechanisms than AERMOD. Mention that OAQPS is currently testing a "plume-in-grid" version of CMAQ in which SCICHEM is used for the plumes.

The chapter makes a major point that CMAQ has a mean bias. I checked with my contacts in the EPA CMAQ development group and they say that the mean bias for $SO_2$ was corrected years ago. Biases in the secondary pollutants (formed from $SO_2$) have also been removed, and this work is described in several papers and presentations. So apparently the Chapter 3 authors are using outdated information.

The chapter includes several statistical "models" that involve curves fit to concentration observations. Hybrid approaches are preferable, where basic science relations are used to guide the statistical curve fits. For example, on line 6 of p 3-7, it is said that concentration decreases with distance from the source, x, following a power law of $x^{-2/3}$. However basic dispersion principles would yield an $x^{-3/2}$ power for near-ground point sources. The $x^{-3/2}$ relation has been verified with many field studies. Theory would support an $x^{-2/3}$ power law for line sources such as busy roads, where the lateral dispersion component is less important.

Regarding inverse distance weighting (discussed on p 3-10), there are numerous references over several decades supporting the use of the inverse square relation in the atmospheric boundary layer, and these should be listed.

Pp 3-10 – middle of 3-13 - Revise this overview of different kinds of dispersion models to be consistent with Chapter 2.

**References:**

Dourado, H., J. Santos, N. Reis Junior, A.M. Melo, 2012: The effects of atmospheric turbulence on peak-to-mean concentration ratio and its consequence on odour impact assessment using dispersion models. Chem. Engin. Trans. 163-168.

Gifford, F.A., (1960) Peak-to-average concentration ratios according to a fluctuating plume model. Int. J. Air Poll. 3, 253-260.

Schauberger, G. et al., 2012: Concept to estimate the human perception of odour by estimating short-term peak concentrations from one-hour mean values. Atmos. Environ. 54, 624-628.

Slade, D.H. (1968). Meteorology and Atomic Energy, U.S. Atomic Energy Commission (No. TID-24190). *U.S. Atomic Energy Commission,* pp 109-111, 154-156.

Turner, D.B. (1970). Workbook of Atmospheric Dispersion Estimates (No. 999-AP-26). *U.S. Environmental Protection Agency,* pp 37-38.

Venkatram, A., 2002: Accounting for averaging time in air pollution modeling. Atmos. Environ. 36, 2165-2170.

# Dr. Jack Harkema

**Comments on the Executive Summary and Chapter 1**

General Comments:

Currently, the summarized material (and format) in these introductory sections of the ISA appropriately highlights and summarizes the important information provided in the subsequent chapters. Overall the key scientific information, to *date*, that have dictated the authors' causality determinations has been appropriately highlighted. The authors have nicely used summary tables and chapter references to synthesize and streamline the study results that were key for their determinations of causality. There are, however, some important revisions that will have to be made in the individual later chapters that may need to be captured or revised in Chapter 1 and the Executive Summary. For example, see my comments/suggestions for Chapter 4 below.

In addition, the Executive Summary contains some technical jargon and phrasing that could be modified (eliminated or better defined) to be more readable for a wider sector of the public interested in this area.

Specific Comments:

In the Executive Summary, *Sources and Human Exposure to Sulfur Dioxide* section is quite long and should be shortened by referring the more technical details to Chapter 1 or an appropriate later Chapter.

xliv, line 1. *The strongest evidence* . . . should be deleted and rewritten (e.g., *Strong scientific data indicates that* . . . .).

xlv, Iine 6. Does *additional studies* refer to more *recent studies* (since the last ISA)?

1-13, 1.5.2 Airway mucosal injury should be defined. Airway remodeling (epithelial mucous cell metaplasia/hyperplasia, intramural fibrosis) is another important pathological feature of asthma (COPD) and in the airways of laboratory animals repeatedly exposed to SO2.

1-14, 1.2, line 16-18. A clearer definition of *new asthma onset* is needed to set it apart from *asthma exacerbation.*

1-14, 1.6, line 31-33. For setting NAAQS to prevent health effects of sulfur dioxide, is it really appropriate to define short-term exposures as long as 1 month?

1-15, 1.6.1, line 6. Rewrite first sentence – see comment above xliv, line 1.

1-15, 1.6.1, line 16. Replace . . . *under increased ventilation conditions* . . . with . . .*when exercising.*

1-15, 1.6.1, lines 26-33. Not sure what is meant by *most severe asthmatics have not been tested* when *moderate-severe* asthmatics were reported to have larger changes in lung function. This needs clarification.

1-18, lines 1-2. The sentence *There is a lack . . . . across disciplines* is not clear and should be rewritten.

1-19, lines 4-6. Define *supportive evidence for health outcomes*. This is not clear in the context of the sentence and the paragraph.

**Comments on Chapter 4 - Dosimetry and Mode of Action**

In general, this chapter is very well written with appropriate dosimetry and mode of action information, and of suitable depth, for the known health effects of inhaled sulfur dioxide. A few specific comments are listed below.

Specific Comments:

Since the nasal mucosa is the tissue receiving the greatest SO2 exposure (95-99% absorption in the nose), the SO2 metabolizing capabilities should be discussed in this chapter including a comparison with other tissues in the respiratory tract and elsewhere. For example, it has been reported that the activity of sulfite oxidase is comparable in the nose, trachea, and proximal and medium bronchi, but lower in lung parenchyma of laboratory dogs. As a start, see the paper by Maier et al. 1999 on xenobiotic-metabolizing enzymes in the canine respiratory tract for more specifics (1).

1. Maier KL, Wippermann U, Leuschel L, Josten M, Pflugmacher S, Schröder P, Sandermann H Jr, Takenaka S, Ziesenis A, Heyder J. Xenobiotic-metabolizing enzymes in the canine respiratory tract. Inhal Toxicol. 1999. Jan;11(1):19-35. PubMed PMID: 10380157.

In section 4.3.1 (*Activation of Neural Reflexes*) the authors should consider replacing the term *sensory irritation* with *nasal irritation* when comparing to *pulmonary (*better yet *bronchial) irritation*. Both are referring to SO2-induced sensory nerve activation but at different anatomical sites and different sensory nerves (trigeminal vs c fibers). In addition, are there any studies of airway TRP receptor activation with SO2 exposure? Recognition of these important airway (and vasculature) receptors in response to irritating agents is important when developing a mode of action for an upper respiratory tract irritant.

In section 4.3.2. (Injury of Airway Mucosa). There is a noticeable lack of review of SO2-induced airway remodeling that should include mucous cell metaplasia of airway epithelium (and associate mucus hypersecretion) and intramural fibrosis. These are key hallmarks of COPD and chronic asthma in humans that are modeled in animals repeatedly exposed to SO2 alone – see 2006 paper by Wagner et al. (2). These pathological changes are recognized key events in the pathway of chronc airway diseases caused by inhaled SO2 in animals, which have been known for at least 30 years.

2. Wagner U, Staats P, Fehmann HC, Fischer A, Welte T, Groneberg DA. Analysis of airway secretions in a model of sulfur dioxide induced chronic obstructive pulmonary disease (COPD). J Occup Med Toxicol. 2006 Jun 7;1:12. PubMed PMID: 16759388; PubMed Central PMCID: PMC1559628.

On 4-21, the title of 4.3.4, *Transduction of Extrapulmonary Effects*, does not adequately convey the information in the text and a more appropriate title should be used (e.g., Induction of Systemic Effects).

Figure 4-1 does not reflect the text – Airway hyperresponsiveness is enhanced bronchoconstriction to a pharmaceutical agent, as noted on 4-25 lines 1-4. This needs to be better illustrated in this diagrammatic progression of key events.

In the *Mode of Action* section (and possibly elsewhere in this chapter) some mention of the potential role of airway innate immune responses (i.e., innate lymphoid cells group 2) in the induction (or enhancement) of type 2 immunity that is related to innate and adaptive allergy (even though this may not yet been examined with SO2). This is an important newly recognized pathway that at least should be mentioned in the text. See references below (3, 4, 5) that have found these cells to be important with inhaled irritants (e.g., O3). Restricting the discussion to only conventional Th1/Th2 pathways are no longer enough (see editorial by Koyasu, Int Immunol 28:1-2, 2015).

3. Kumagai K, Lewandowski R, Jackson-Humbles DN, Li N, Van Dyken SJ, Wagner JG, Harkema JR. Ozone-induced Nasal Type 2 Immunity in Mice is Dependent on Innate Lymphoid Cells. Am J Respir Cell Mol Biol. 2015 Nov 11. [Epub ahead of print] PubMed PMID: 26559808.

4. Yang Q, Ge MQ, Kokalari B, Redai IG, Wang X, Kemeny DM, Bhandoola A, Haczku A. Group 2 innate lymphoid cells mediate ozone-induced airway inflammation and hyperresponsiveness in mice. J Allergy Clin Immunol. 2015 Aug 15. pii: S0091-6749(15)00936-7. doi: 10.1016/j.jaci.2015.06.037. [Epub ahead of print] PubMed PMID: 26282284.

5. Palm NW, Rosenstein RK, Medzhitov R. Allergic host defences. Nature. 2012 Apr 25;484(7395):465-72. doi: 10.1038/nature11047. PubMed PMID: 22538607; PubMed Central PMCID: PMC3596087.

6. Holtzman MJ, Byers DE, Alexander-Brett J, Wang X. The role of airway epithelial cells and innate immune cells in chronic respiratory disease. Nat Rev Immunol. 2014 Oct;14(10):686-98. doi: 10.1038/nri3739. Epub 2014 Sep 19. Review. PubMed PMID: 25234144.

7. Koyasu S. Introduction: Innate Lymphoid Cells Special Issue. Int Immunol. 2016 Jan;28(1):1-2. doi: 10.1093/intimm/dxv073. PubMed PMID: 26712897.

## Dr. Daniel Jacob

Chapter 2 describes scientific information on sources, atmospheric chemistry, and measurement and modeling of ambient concentrations of gaseous sulfur oxides. I found it to be overall well-informed and informative. With regard to the charge questions,

a. *To what extent is the information presented regarding sources, chemistry, and measurement and modeling of ambient concentrations accurate, complete, and relevant to the review of the SO2 NAAQS?*

I think it's good but I have a lot of minor comments, listed below.

b. *Please comment on the extent to which available information on the spatial and temporal trends of ambient SO2 concentrations at various scales has been adequately and accurately described. In particular, what is the extent to which the analyses of recently available 5-min SO2 concentration data are informative in considering relationships between 5-min and 1-hr SO2 concentrations?*

I think that this is done very well and that the correlations between 5-min and 1-h concentrations provide the information needed to model 5-min exposure on the basis of 1-h concentrations.

c. *How informative is the analysis of correlations between SO2 and co-occurring pollutant concentrations for interpretation of epidemiologic studies?*

I think it's good enough – correlations are weak, as one would expect.

Below are specific comments related to charge question a). (Page, line)

1. (2.1, 14) Smelters are also important.
2. (2.9) Figure 2.5 has strange left and right y-axes title/labels (which is correct?). I don't understand the green line labeled as "fuel comb. Industrial" – it doesn't fit with any of the other information and why would it be as large as the EGUs? That's not consistent with Table 2.1. A general problem with the different source type attribution estimates presented in figures and tables in this section is that they are not internally consistent because they come from different sources – this needs to be either fixed or acknowledged.
3. (2.10, 5). Hydrogen disulfide is not atmospherically relevant. Dimethylsulfide is and would be worth mentioning.
4. (2.11) Figure 2.6: $CS_2$, not CS. DMS source is inconsistent with other numbers given in tables and text. Volcanoes also emit $H_2S$. Not sure that it's worth mentioning primary sea salt and dust sulfate sources – they can't make $SO_x$. And I would put a number on the volcano source because it's no more uncertain than DMS. It's variable year-to-year but that's a different matter.
5. (2.12, 5) They're not difficult to estimate, they're just variable year to year.
6. (2.15, 6) These may be modeling studies but their specification of volcanic emissions is not based on their model.
7. (2.16, 5) Volcanoes release $H_2S$ in larger amounts than COS or $CS_2$.
8. (2.18, 14) COS loss is mainly by OH, see Table 4b of Chin and Davis (JGR 1993).

9. (2.19, 1-2) Awkward, suggests that NO3 is dominant sink for DMS.
10. (2.20, 10) Radical, not ion
11. (2.21, 1-3) Real issue with Criegees as SO2 oxidants is their hydrolysis; see Kim et al., ACP 2015. http://acmg.seas.harvard.edu/publications/2015/kim2015_seac4rs.pdf
12. (2.21) Equation 2.4 should be written as 3-body reaction.
13. (2.21, 14-15). Major sulfur species in cloud would be sulfate. Also, "acid dissociation" is better terminology than "react with water".
14. (2.28, 26) To say that satellites don't measure SO2 directly is gratuitous. The same could be said for other methods, in particular DOAS.
15. (2.28, 30) Real problem with SO2 sensitivity is air scattering, not ozone column.
16. (2.28, 36) Probably should mention detection of emissions from oil sands in western Canada (McLinden C A, Fioletov V, Boersma K F, Krotkov N, Sioris C E, Veefkind J P and Yang K 2012 Air quality over the Canadian oil sands: A first assessment using satellite observations Geophys. Res. Lett. 39 L04804)
17. (2.29, 1-3) Throwaway statement. Cut.

# Dr. Farla Kaufman

Overall the document is a good compilation of the available information and the presentation is clear and well-written. However, the quality of the evaluation, integration, and synthesis of evidence varies across the document (see comments below). The causal framework also overall seems to be applied appropriately.

The evidence for respiratory effects of $SO_2$ pertaining to the lowest concentrations seems appropriately characterized in the ISA.

**Comments on Chapter 5**

In most sections of this chapter the previous and recent evidence is comprehensive and well presented.

Many summary sections provide good integration and evaluation of the body of evidence. However, there are other summary paragraphs for individual sections that report inconsistent findings and/or lack of multipollutant models without additional synthesis of the information. Further comparison of the studies on specific characteristics (e.g. study design, analytical approach, measurement error, confounding factors and the likelihood of bias), would help to elucidate strengths and limitations and could be used to provide a more detailed evaluative and integrative assessment of the evidence. This would aid in identifying studies of higher quality as well as contribute to helping understand the basis for inconsistent findings. In addition, these summaries and conclusions, which serve to identify the specific body of evidence forming the basis for determining the classification of causality could then be cited as the key references and key evidence. Currently the degree to which this evidence is provided in these tables varies across sections.

Although summarizing information in tabular format is very helpful, there seems to be room in certain sections for reformatting, combining and/or condensing the tables, and perhaps placing some of the tables in a supplementary materials section. It is recognized that this could be a trade-off in certain sections. However, combining tables that present study characteristics with those that present results, as done in Table 5-28, would be helpful in evaluating the evidence and in shortening the document. To the degree possible, consistency across tables in what information is presented and how it is presented could aid in the readability of the tables as well as the document.

Providing more detail about individual studies would be appropriate at times. A simple example is for the study by Rubinstein et al. (1990). The text states $NO_2$ induced greater airway responsiveness in only one subject; however, no information was included as to the total number of subjects tested. Knowing that one subject constituted >10% of those tested would be important information.

As it is recognized that spatial variability is an important factor in exposure assessment for $SO_2$, it would be useful to include information on the distance to the monitor, whenever possible, in tables describing epidemiology studies.
Although there are only a few controlled studies on the effects of mixtures, there is evidence of interactions with other pollutants. Could this be given further consideration within the current framework.

Currently our U.S. population includes a significant percentage of children who are obese and a large percentage who are asthmatic, and many disadvantaged children live in conditions exposing them to allergens. For children in general, especially for those who are overweight or obese, increased exercise is the recognized goal. In this context, perhaps a summary of findings across all respiratory effects for potential "at risk" populations could be helpful given the evidence such as: 1) the, at times substantial, variability in responses to exposures; 2) greater decrements in lung function from exposure to mixtures; 3) the recognition that previous controlled studies of exercise in asthmatics were conducted in subjects who, by today's standards, would be classified as moderate, not moderate/severe; 4) greater sensitivity conferred in the context of allergic inflammation; and 4) greater vulnerability of children.
(Some studies on asthma and obesity are included below).

Pg 5-35 Table 5-8, also page 5-36 line 19 - Study by Spira-Cohen et al. (2011). Mean $SO_2$ and upper concentration level is NR (not reported). However, the Table references an email exchange between Dr. Spira-Cohen and Dr. Patel that suggests Dr. Spira-Cohen sent these data to U.S. EPA.

Pg 5-240 Section 5.4 Although the literature may be limited in the area of neurobehavioral effects, relevant studies should be included in this review of developmental toxicity (see studies listed below).

5-241 - line 30 - Geer et al (2012) also evaluated $SO_2$ adjusted for $O_3$.

Pg 5-242 Table 5-44 Information on reproductive and developmental studies should include more study details (e.g. study design, study population, statistical analysis, covariates, etc.).

Pg 5-247 Section 5.4.1.3 There are many studies that have examined $SO_2$ exposure during pregnancy and fetal growth. More of the older studies could be included in the section. Most studies report an increased risk of fetal growth restriction with $SO_2$ exposure. Thus, perhaps this endpoint warrants further discussion. Consideration of factors such as potential exposure misclassification, temporal and spatial variability, may help to weigh the study findings and evaluate inconsistency of findings across the body of literature.

5-250 line 16 Table 5S-13 Brauer et al. (2008), the mean $SO_2$ concentration for this study is listed as 5.3 ppb. However, the mean value is actually ~2.2 ppb, as in the publication $SO_2$ is 5.3 µg/m3. 2.2 ppb is close to the limit of detection for this study (2 ppb). Thus, exposure misclassification in this study might have decreased the likelihood of detecting an effect if one had been present.

5-250 Section on Birth Weight - It would seem that an expanded discussion of this outcome is warranted as there are a number of newer studies that should be viewed along with some more of the older studies (e.g. Lin et al. (2004).

5-251 line 26 As mentioned above, the summary sections should be expanded to include an integrative evaluations of these studies including the strengths and limitations. In addition, evaluating studies on the basis of important aspects such as the spatial and temporal variability of exposure is important in understanding how much weight to place on a particular study. Important factors such as a very low level of exposure, as noted above in the study by Brauer et al. (2008), or very low variability in $SO_2$ levels as in the study by Bell et al. (2007) provides a context for the study results.

5-252 Section on Birth Defects - There are two meta-analyses available but they were not included here, not sure why.

This section could also benefit from some evaluative summary such as to the characteristics of the studies in order to discern reasons for inconsistent findings, i.e. level of exposure, covariates considered, likelihood of exposure misclassification, classification of defects, etc.

Table 5-46 – "Under Key Evidence" it states that "associations not evaluated in co-pollutant models" . I am not clear which studies this is referring to since the three studies cited under "key references" (Liu et al., 2003; Le et al., 2012; and Sagiv et al. 2005) all evaluated co-pollutants, and found that associations with $SO_2$ remained significant. In addition, this was the only limitation mentioned in the table for this outcome; however, no mention of co-pollutant models was made in the summary for the corresponding section. The table should reflect the text for the corresponding section and reference the relevant studies.

In the table on 5-244 - Uncertainty around the timing of exposure across studies for the BW and LBW outcomes is noted in Table 5-46. However, for outcomes such as birth weight and low birth weight it is conceivable that different mechanisms could be influencing the outcome at varying timing of exposure, including early in pregnancy and late in pregnancy.

P5-256 as stated in the ISA -
*"The state of California, under the auspices of Proposition 65, the California Safe 22 Drinking Water and Toxic Enforcement Act of 1986, has listed sulfur dioxide as a 23 chemical known to cause reproductive toxicity based on evidence from laboratory animal 24 studies and epidemiologic studies. However, much of this evidence is from toxicological 25 studies with exposure to SO2 at 5,000 ppb or greater (beyond the scope of this ISA); 26 effects seen at the higher doses include male reproductive effects on sperm and fecundity, 27 as well as oxidative damage to the male reproductive organs, changes in birth weight or 28 litter size, delayed reflexes in early life, and aberrant behavior of pups after in utero 29 exposure. Epidemiologic evidence used for this listing is also evaluated under differing 30 criteria than are employed for the ISA."*

Correction about the basis for the listing of $SO_2$ under Prop 65 -
The listing for $SO_2$ under the State of California's Proposition 65 is for developmental toxicity. The particular type of developmental toxicity noted by the Developmental and Reproductive Toxicant Identification Committee that formed the basis for this listing was intrauterine growth restriction. Only two toxicological studies were considered by the Committee for this outcome, and, as noted in the ISA, they were conducted at exposures greater than 5,000 ppb. Although studies relevant to male reproductive toxicity were considered, the DART IC committee did not find sufficient evidence of toxicity to list this endpoint under Proposition 65.

I am also assuming that the final document will have a Table of Contents that includes the various headings and sub-headings for the chapters and there will be separate lists for tables and figures.

<u>Studies on Obesity and Asthma:</u>

Childhood body mass index and subsequent physician-diagnosed asthma: a systematic review and meta-analysis of prospective cohort studies.
EEgan et al. BMC Pediatrics 2013, 13:121gan et al. BMC Pediatrics 2013, 13:121

General and abdominal obesity and incident asthma in adults: the HUNT study.
Brumpton B1, Langhammer A, Romundstad P, Chen Y, Mai XM.
Eur Respir J. 2013 Feb;41(2):323-9. doi: 10.1183/09031936.00012112. Epub 2012 May 31.

Obesity and asthma: location, location, location.
Shore SA.
Eur Respir J. 2013 Feb;41(2):253-4. doi: 10.1183/09031936.00128812. No abstract available.

Obesity and bronchodilator response in black and Hispanic children and adolescents with asthma.
McGarry ME, Castellanos E, Thakur N, Oh SS, Eng C, Davis A, Meade K, LeNoir MA, Avila PC, Farber HJ, Serebrisky D, Brigino-Buenaventura E, Rodriguez-Cintron W, Kumar R, Bibbins-Domingo K, Thyne SM, Sen S, Rodriguez-Santana JR, Borrell LN, Burchard EG.
Chest. 2015 Jun;147(6):1591-8.

Childhood obesity and asthma control in the GALA II and SAGE II studies.
Borrell LN, Nguyen EA, Roth LA, Oh SS, Tcheurekdjian H, Sen S, Davis A, Farber HJ, Avila PC, Brigino-Buenaventura E, Lenoir MA, Lurmann F, Meade K, Serebrisky D, Rodriguez-Cintron W, Kumar R, Rodriguez-Santana JR, Thyne SM, Burchard EG.
Am J Respir Crit Care Med. 2013 Apr 1;187(7):697-702.

Additional Studies of Reproductive Toxicity, etc.

Yao et al., Feb 2015
Chronic $SO_2$ inhalation above environmental standard impairs neuronal behavior and represses glutamate receptor gene expression and memory-related kinase activation via neuroinflammation in rats.

Yun et al., 2013
$SO_2$ inhalation causes synaptic injury in rat hippocampus via its derivatives in vivo.

Lin et al., 2014
Multilevel analysis of air pollution and early childhood neurobehavioral development.
Jung et al., 2013
Air pollution and newly diagnostic autism spectrum disorders: a population-based cohort study in Taiwan.

Vrijheid et al., 2011
Ambient air pollution and risk of congenital anomalies: a systematic review and meta-analysis.

Tong et al., Jan 2015
The association between air pollutants and morbidity for diabetes and liver diseases modified by sexes, ages, and seasons in Tianjin, China.

# Dr. David Peden

*Chapter 4 characterizes scientific evidence on the dosimetry and modes of action for SO2. Dosimetry and modes of action are bridged by the absorption and reaction of SO2 in the epithelial lining fluid to form SO2-derived products (e.g., sulfite and/or S-sulfonates) that are widely distributed throughout the body.*

a. *To what extent is the discussion of the chemistry of inhaled SO2 and the processes of absorption, distribution, metabolism, and elimination accurate, complete, and relevant to the review of the SO2 NAAQS?*

It is likely that decreased upper airway airflow in children and obesity (in all ages) will modify deposition of SO2 or $H_2SO_4$ in the airway. This may be covered in section 5 (though I did not find it). Using radiolabeled particle uptake data assessment of deposited marker particles, children had decreased nasal contribution to breathing and increased deposition to the lower airway (1). In a comparable study, it was found that children with increased BMI also had increased central airway deposition of inhaled agents (2). There is a gap in the literature for this as well regarding SO2. However, the data in these reports l demonstrate changes in airflow associated with nasal and orophergygeal breathing in children (though there are data derived from labeled particles), which may be germane for $SO_2$ exposure in the form of $H_2SO_4$. It seems likely that these factors may be as much of a physiological parameter as asthma in directing deposition of SO2 onto specific airway sites.

Extrapolating the robust literature on obesity and obstructive sleep apnea, it is very likely that functional nasal obstruction, with increased oropharyngeal breathing will occur in obese and overweight individuals of all ages. This will likely be true for breathing at rest (see note above). Thus the lower airway distribution/deposition of $SO_2$ (or any pollutant) will likely be greater in obese individuals. Given the rate of obesity/overweight in the US, this is an important consideration, especially as much of the data on radiolabelled SO2 uptake is fairly dated (when obesity was a much less prevalent issue) (refs 3-12).

c. *To what extent are the discussion and integration of the potential modes of action underlying the health effects of exposure to sulfur oxides presented accurately and in sufficient detail? Are there additional modes of action that should be included in order to fully characterize the underlying mechanisms of sulfur oxides?*

One area that may need better definition is the impact of genetic factors on response to SO2. A cursory search identifies a study in which persons with the AA genotype for the -308 TNF promoter region have increased airway response to $SO_2$ (13). As increasing data are accumulated showing the effect of SO2 on signaling and inflammatory mechanisms, there is an increased likelihood of identifying genetic factors in response to SO2 (14,15).

There has been an association with decreased death due to emphysema and decreased SO2 exposure, suggesting SO2 may impact structural integrity of the lung parenchyma (16)

There are a number of studies that suggest that So2 contributes to airway infections (17-19) and there are other several studies (some relatively dated) that indicate an impact of SO2 on mucociliary

clearance, and subsequent effect of this on infections, including otitis media and bronchitis (20-22). The study by Carson et al demonstrates the structural effect of SO2 on cilia (23).

Systemic effects do occur with SO2 (15). SO2 has effects on psychosis, cardiovascular effects and risks of respiratory effects. While the gender and age specific aspects of the effect of SO2 vary by endpoint, there are certainly data showing gender and age based differences impact susceptibility to SO2, with women having more effect of SO2 on CVD and age being a determinant of asthma effects of SO2 (24-26). These suggest specific mechanistic considerations as well.

**References:**


1. Bennett WD, Zeman KL. Effect of body size on breathing pattern and fine-particle deposition in children. J Appl Physiol (1985). 2004 Sep;97(3):821-6. Epub 2004 Apr 23. PubMed PMID: 15107416.

2. Bennett WD, Zeman KL, Jarabek AM. Nasal contribution to breathing and fine particle deposition in children versus adults. J Toxicol Environ Health A. 2008;71(3):227-37. PubMed PMID: 18097948.

3. Bazzano LA, Hu T, Bertisch SM, Yao L, Harville EW, Gustat J, Chen W, Webber LS, Shu T, Redline S. Childhood obesity patterns and relation to middle-age sleep apnoea risk: the Bogalusa Heart Study. Pediatr Obes. 2016 Jan 19. doi: 10.1111/ijpo.12103. [Epub ahead of print] PubMed PMID: 26780975.2:

4. Kassim R, Harris MA, Leong GM, Heussler H. Obstructive sleep apnoea in children with obesity. J Paediatr Child Health. 2016 Jan 8. Doi 10.1111/jpc.13009. [Epub ahead of print] PubMed PMID: 26748912.

5. Jalilolghadr S, Yazdi Z, Mahram M, Babaei F, Esmailzadehha N, Nozari H, Saffari F. Sleep architecture and obstructive sleep apnea in obese children with and without metabolic syndrome: a case control study. Sleep Breath. 2015 Dec 28. [Epub ahead of print] PubMed PMID: 26711131.

6. Franklin KA, Lindberg E. Obstructive sleep apnea is a common disorder in the population-a review on the epidemiology of sleep apnea. J Thorac Dis. 2015 Aug;7(8):1311-22. doi: 10.3978/j.issn.2072-1439.2015.06.11. Review. PubMed PMID: 26380759; PubMed Central PMCID: PMC4561280.
7. 5: Wilhelm CP, deShazo RD, Tamanna S, Ullah MI, Skipworth LB. The nose, upper airway, and obstructive sleep apnea. Ann Allergy Asthma Immunol. 2015 Aug;115(2):96-102. doi: 10.1016/j.anai.2015.06.011. Review. PubMed PMID: 26250769.

8. Ryan CM, Pillar G. Adolescent obesity, adenotonsillar hypertrophy, and obstructive sleep apnea. Am J Respir Crit Care Med. 2015 Jun 1;191(11):1220-2. doi: 10.1164/rccm.201504-0677ED. PubMed PMID: 26029834.

9. Shetty S, Parthasarathy S. Obesity Hypoventilation Syndrome. Curr Pulmonol Rep. 2015 Mar 1;4(1):42-55. PubMed PMID: 26029497; PubMed Central PMCID: PMC4444067.

10. Teodorescu M, Broytman O, Curran-Everett D, Sorkness RL, Crisafi G, Bleecker ER, Erzurum S, Gaston BM, Wenzel SE, Jarjour NN; National Institutes of Health, National Heart, Lung and Blood Institute Severe Asthma Research Program (SARP) Investigators. Obstructive Sleep Apnea Risk, Asthma Burden, and Lower Airway Inflammation in Adults in the Severe Asthma Research Program (SARP) II. J Allergy Clin Immunol Pract. 2015 Jul-Aug;3(4):566-75.e1. doi: 10.1016/j.jaip.2015.04.002. Epub 2015 May 21. PubMed PMID: 26004304; PubMed Central PMCID: PMC4500672.

11. Schwab RJ, Kim C, Bagchi S, Keenan BT, Comyn FL, Wang S, Tapia IE, Huang S, Traylor J, Torigian DA, Bradford RM, Marcus CL. Understanding the anatomic basis for obstructive sleep apnea syndrome in adolescents. Am J Respir Crit Care Med. 2015 Jun 1;191(11):1295-309. doi: 10.1164/rccm.201501-0169OC. PubMed PMID: 25835282; PubMed Central PMCID: PMC4476519.

12. Oliveira MC, Tufik S, Haddad FL, Santos-Silva R, Gregório LC, Bittencourt L. Systematic Evaluation of the Upper Airway in a Sample Population: Factors Associated with Obstructive Sleep Apnea Syndrome. Otolaryngol Head Neck Surg. 2015 Oct;153(4):663-70. doi: 10.1177/0194599815577598. Epub 2015 Mar 27. PubMed PMID: 25820578.

13. Winterton DL, Kaufman J, Keener CV, et al. Genetic polymorphisms as biomarkers of sensitivity to inhaled sulfur dioxide in subjects with asthma. Ann Allergy Asthma Immunol 2001;86:232–8.

14. Tadokoro T, Wang Y, Barak LS, Bai Y, Randell SH, Hogan BL. IL-6/STAT3 promotes regeneration of airway ciliated cells from basal stem cells. Proc Natl Acad Sci U S A. 2014 Sep 2;111(35):E3641-9. doi: 10.1073/pnas.1409781111. Epub 2014 Aug 18. PubMed PMID: 25136113; PubMed Central PMCID: PMC4156689.

15. Reno AL, Brooks EG, Ameredes BT. Mechanisms of Heightened Airway Sensitivity and Responses to Inhaled SO2 in Asthmatics. Environ Health Insights. 2015 Apr 1;9(Suppl 1):13-25. doi: 10.4137/EHI.S15671. eCollection 2015. Review. PubMed PMID: 25922579; PubMed Central PMCID: PMC4384764.

16. Kravchenko J, et al Long-term dynamics of death rates of emphysema, asthma, and pneumonia and improving air quality.. Int J Chron Obstruct Pulmon Dis. 2014 Jun 16;9:613-27. doi: 10.2147/COPD.S59995. eCollection 2014.

17. Tam WW, Wong TW, Ng L, Wong SY, Kung KK, Wong AH. Association between air pollution and general outpatient clinic consultations for upper respiratory tract infections in Hong Kong. PLoS One. 2014 Jan 23;9(1):e86913. doi: 10.1371/journal.pone.0086913. eCollection 2014. PubMed PMID: 24466288; PubMed Central PMCID: PMC3900683.

18. HEI Collaborative Working Group on Air Pollution, Poverty, and Health in Ho Chi Minh City, Le TG, Ngo L, Mehta S, Do VD, Thach TQ, Vu XD, Nguyen DT, Cohen A. Effects of short-term exposure to air pollution on hospital admissions of young children for acute lower respiratory infections in Ho Chi Minh City, Vietnam. Res Rep Health Eff Inst. 2012 Jun;(169):5-72; discussion 73-83. PubMed PMID: 22849236.

19. Wong CM, Thach TQ, Chau PY, Chan EK, Chung RY, Ou CQ, Yang L, Peiris JS, Thomas GN, Lam TH, Wong TW, Hedley AJ; HEI Health Review Committee. Part 4. Interaction between air pollution and respiratory viruses: time-series study of daily mortality and hospital admissions in Hong Kong. Res Rep Health Eff Inst. 2010 Nov;(154):283-362. PubMed PMID: 21446214

20. Shusterman D. Environmental nonallergic rhinitis. Clin Allergy Immunol. 2007;19:249-66. Review. PubMed PMID: 17153018.

21. O'Brien DW, Morris MI, Ding J, Zayas JG, Tai S, King M. A mechanism of airway injury in an epithelial model of mucociliary clearance. Respir Res. 2004 Aug 24;5:10. PubMed PMID: 15357883; PubMed Central PMCID: PMC517808.

22. Knorst MM, Kienast K, Riechelmann H, Müller-Quernheim J, Ferlinz R. Effect of sulfur dioxide on mucociliary activity and ciliary beat frequency in guinea pig trachea. Int Arch Occup Environ Health. 1994;65(5):325-8. PubMed PMID: 8175188.

23. Carson JL, Brighton LE, Collier AM, Bromberg PA. Correlative ultrastructural investigations of airway epithelium following experimental exposure to defined air pollutants and lifestyle exposure to tobacco smoke. Inhal Toxicol. 2013 Feb;25(3):134-40. doi: 10.3109/08958378.2013.763314. PubMed PMID: 23421485.

24. Tong L, Li K, Zhou Q. Season, sex and age as modifiers in the association of psychosis morbidity with air pollutants: A rising problem in a Chinese metropolis. Sci Total Environ. 2016 Jan 15;541:928-33. doi: 10.1016/j.scitotenv.2015.09.066. Epub 2015 Nov 11. PubMed PMID: 26457738.

25. Alhanti BA, Chang HH, Winquist A, Mulholland JA, Darrow LA, Sarnat SE. Ambient air pollution and emergency department visits for asthma: a multi-city assessment of effect modification by age. J Expo Sci Environ Epidemiol. 2015 Sep 9. doi: 10.1038/jes.2015.57. [Epub ahead of print] PubMed PMID: 26350981.

26. Qin XD, Qian Z, Vaughn MG, Trevathan E, Emo B, Paul G, Ren WH, Hao YT, Dong GH. Gender-specific differences of interaction between obesity and air pollution on stroke and cardiovascular diseases in Chinese adults from a high pollution range area: A large population based cross sectional study. Sci Total Environ. 2015 Oct 1;529:243-8. doi: 10.1016/j.scitotenv.2015.05.041. Epub 2015 May 26. PubMed PMID: 26022408.

# Dr. Richard Schlesinger

**Comments on Chapter 4**

*Chapter 4 characterizes scientific evidence on the dosimetry and modes of action for SO2. Dosimetry and modes of action are bridged by the absorption and reaction of SO2 in the epithelial lining fluid to form SO2-derived products (e.g., sulfite and/or S-sulfonates) that are widely distributed throughout the body.*

a. *To what extent is the discussion of the chemistry of inhaled SO2 and the processes of absorption, distribution, metabolism, and elimination accurate, complete, and relevant to the review of the SO2 NAAQS?*

The discussion of inhaled SO2 and processes are complete and relevant to review of the NAAQS.

b. *Please comment on the discussion comparing endogenously generated and ingested sulfite with that derived from ambient inhalation.*

The discussion of endogenous and ingested sulfite with that derived from inhalation of SO2 could be a bit more concise.

c. *To what extent are the discussion and integration of the potential modes of action underlying the health effects of exposure to sulfur oxides presented accurately and in sufficient detail? Are there additional modes of action that should be included in order to fully characterize the underlying mechanisms of sulfur oxides?*

See General Comment re section 4.1 below. SO2 may result in effects due to generation of sulfuric acid, and that is not made clear in the document.

<u>General Comments</u>

<u>Section 4.1</u>

It is noted here that understanding the mechanisms of biological responses may provide biological plausibility for effects observed in the epidemiological studies, which are described in Chapter 5. However, it seems that it would be best to move Chapter 4 after Chapter 5, since the former actually introduces health effects in the context of mechanisms that are the discussed later in Chapter 5. This results in much redundancy. As a toxicologist, I would prefer to read about the biological effects in Chapter 5 and then read potential mechanisms that would convince me of the biological plausibility, rather than the other way around.

An example of this issue is on page 4-15, where beginning on line 10 there is a discussion of health outcomes rather than mechanisms. Other examples are on page 4-16, beginning on line 12; page 4-17 lines 3-9 and beginning on line 29.

An alternative would be to limit the current Chapter 4 solely to dosimetry and then have a section at the end of Chapter 5 dedicated to mechanisms.

Section 4.2

The first paragraph of this section notes that few studies have been done since the last ACQD. However, there are detailed discussions of studies that were done prior to and incorporated into the prior ACQD. It would be best to summarize results of these earlier studies to make the chapter more concise.

Equation 4.1. The manner in which this is presented is a bit confusing.

Page 4.5, lines 4-7. Triggering of asthma may in fact reduce the amount of gas reaching the lower respiratory tract.

Page 4-7, lines 1-16. Line 4 notes that SO2 derived products are widely distributed throughout the body very early after exposure, yet line 16 notes that SO2 derived produces are only slowly absorbed into the blood, which would suggest that they are not widely distributed.

Page 4-12,
line 1. What is meant by the statement that SO2 is a reactive antioxidant gas?

Line 29. Penetration into which region of the respiratory tract?

Page 4-23, lines 29-32. If an endpoint is an effect measured in the clinic, how is this different from an outcome measured at the organism level?

Figure 4-1 Some of the evidence arrows are not consistent with discussion in the text as to what has been seen in animals vs humans.

# Dr. Elizabeth A. (Lianne) Sheppard

**Comments on Chapter 3**

From my perspective the primary purpose of this chapter is to allow readers to understand human exposure to SO2, as well as how it is measured and modeled, in order to appropriately interpret the epidemiologic study results in Chapter 5. Is this EPA's intent? Assuming it is, then I believe the primary objective has not been met. Overall I judge that this chapter needs considerable reorganization and revision. I find it less clear than other similar sections/chapters from previous ISAs. In numerous places the wording is confusing or misleading, making it difficult to grasp the key messages the chapter should convey.

I think an important objective of the chapter is to give readers an understanding of SO2 as a pollutant (in the context of human exposure) and the implications of this understanding for how well various epidemiological study designs will perform under different exposure metrics (focusing only on the ones relevant to the discussion in Chapter 5). It should clearly distinguish between the understanding of exposure that is generally applicable to any ambient air pollutant from insights that are specific to SO2. In addition this chapter should clearly covey new understanding since 2008 and distinguish this from the material discussed in the 2008 document.

I'm having some difficulty distinguishing between how to divide my responses to the two charge questions for this chapter. I interpret that the first one is focusing on exposure measurement and modeling, while the second on exposure assessment and measurement error in epidemiological studies. (Are these intended to address sections 3.2 (Methodological considerations for use of exposure data) and 3.3 (Exposure assessment and epidemiologic inference), respectively?) I would not separate exposure modeling from exposure assessment and measurement error. EPA's definition of exposure assessment (in Guidelines for Exposure Assessment, EPA/600/Z-92/001 May 1992) is "The determination or estimation (qualitative or quantitative) of the magnitude, frequency, duration, and route of exposure." This clearly includes exposure measurement and modeling. Thus I find it difficult to understand how EPA staff want us to approach these two questions. For the moment I have not attempted to separate my comments by charge question.

**Charge question draft responses**

*a. To what extent is the discussion on methodological considerations for **exposure measurement and modeling** clearly and accurately conveyed, appropriately characterized, and relevant to the review of the SO2 NAAQS?*

*b.* Please comment on the accuracy, level of detail, and clarity of the discussion regarding **exposure assessment and the influence of exposure error** on effect estimates in epidemiologic studies of the health effects of SO2.

The level of detail seems uneven. I think the total length of the chapter is reasonable in principle as 70 pages is sufficiently long to delve into important topics while not so long as to put off readers. However, as currently written, the chapter should be shortened to eliminate unnecessary content and redundancies.

The clarity and accuracy of the discussion are insufficient. There are many statements in the chapter that are inaccurate or unclear. There are a number of broad issues:

- There are topics covered that are not relevant to interpretation of epidemiologic studies, specifically several pages devoted to probabilistic exposure modeling. While some of this discussion is specifically noted as relevant for the risk and exposure assessment, and thus could be retained for that purpose, not all of the discussion of this topic is so carefully delineated. (e.g. p 3-41)
- There is confusion between 1) exposure measurement error that comes from the uncertainty of the exposure metric being used relative to an uncertain version of that metric, and 2) a different target parameter of interest in the epidemiologic study than one might prefer. The latter comes from the distinction between the exposure metric selected vs. the exposure of interest (e.g. the most common case is that the exposure metric selected is ambient concentration while the exposure of interest is personal exposure to ambient SO2). The entire discussion needs to be revised with this distinction in mind.
- There is discussion of dose (by bringing in the topic of ventilation) that is not clearly distinguished from exposure. Using EPA's own definition (in Guidelines for Exposure Assessment, EPA/600/Z-92/001 May 1992), exposure is "The condition of a chemical contacting the outer boundary of a human is exposure." In my opinion, ventilation is not at the outer boundary of a human. Thus most of Section 3.3.3.1 could be eliminated.
- In discussions of the impacts of mismeasured exposures it is not clearly articulated what is meant. For instance, when a mismeasured exposure "widens the confidence intervals", is this relative to using the true (albeit unknown) exposure? And how does this statement relate to understanding of the correct (i.e. nominal) coverage of the confidence intervals being reported? Related to this point, it should be clearly stated what happens when measurement error is ignored in the analysis vs. what happens when it is corrected for. (Specifically, what happens to the effect estimates, standard errors, and coverage of confidence intervals when measurement error is corrected? What are the implications for interpretation of studies that do not correct for measurement error?)
- Some of the studies are not clearly identified as simulation studies. The assumptions of the studies are treated as data and thus the reader can easily incorrectly infer that these assumptions are properties of SO2. (see e.g. p 3-49, paragraph starting on line 3)
- There are broad-brush statements made that may not be adequately worded or valid. Often they should be backed up with references to the literature or other sections of the document. Consider for instance the sentence on p 3-37 lines 6-8: "The use of SO2 measurements from central ambient monitoring sites is the most common method for assigning exposures in epidemiologic studies." How does EPA equate using measurements from central site monitors to a method for assigning exposure? Does this pertain to all epi study designs?

The organization of the chapter should be revisited. I think it is important to convey 1) a general understanding of exposure to ambient air pollutants along with 2) an appreciation for the specific characteristics of SO2 that relate to this understanding. These specific characteristics have implications for measurement and modeling of SO2 as well as for accurate inference about SO2 health effects in epidemiological studies. Similarly, there are general approaches to measuring criteria air pollutants (e.g. ambient concentrations at fixed locations, indoor-outdoor measurements at homes and other subject locations, and personal measurements) and specific technologies to measure SO2. These specific features of measuring SO2 further inform our understanding of epidemiologic studies and the SO2 exposure studies. I expect 70% or more of this chapter should consist of material that could appear in any ISA exposure chapter while the remainder of the text would be specific to SO2.

**More specific comments**: (I am skipping many problematic details to concentrate on larger issues. I encourage EPA to also attend to the many details that need correcting.)

- Section 3.2.2 needs to be completely revised. Its organization and grouping of models are problematic. There is a strange mixing of methods and results in this section. It is written with insufficient clarity making it difficult for a reader to make much sense of the literature or understand the key points.
    - What is the point of the discussion of time-activity models (Section 3.2.2.2)? What epi studies take these into account?
    - Why is Section 3.2.2.3 on models of air exchange rates and microenvironmental models needed?
- Table 3-2: While the intent of this table is helpful, only methods that have at least one epidemiological study of SO2 discussed in Chapter 5 should be listed. Furthermore, the list should be comprehensive for the important epidemiological studies and overlap with the models described in Table 3-3.
- Table 3-4: It is unclear why a table of this length is needed since 10 of 12 cited studies were available for the 2008 ISA. Furthermore, the table mixes reporting of ratios and slopes. These are not the same and it makes it difficult for readers to distill key points from the information presented.
- I found Section 3.3.3 (the opening paragraph) to be very broad-brush and speculative.
- The introduction of new rules for ultra-low sulfur diesel was brought up as potentially important factor in interpretation of reported correlations between pollutants (p 3-50 line 30). However, the subsequent discussion did not appear to take the timing of these studies into account.
- Important terms have not been clearly defined. Examples include the specific R2's reported in the text in Section 3.2.2.1 and in some entries of Table 3-3, and the definitions or data used to compute the correlations in Table 3-9 and the text discussing it. (For instance, for the copollutant correlations in long-term epi studies, are the correlations based on purely space or space and time? Also, are they based on predictions or measurements?) I will provide EPA with additional comments about R2 definitions if asked.

Throughout this chapter there are many sentences with poor wording. Experts other than the primary author(s) of each section should review the chapter and help improve the wording. To support this effort, below is some suggested text rewording for paragraphs that discuss my work or work of my UW colleagues.

*Proposed revised text to replace text that appears on page 3-59, lines 3-20*:
Measurement error for modeled exposures has been decomposed into Berkson-like and classical-like components, sharing some characteristics with Berkson and classical errors, respectively, but with key differences (Szpiro et al. 2011). Berkson-like errors occur when the modeled exposure does not capture all of the variability in the true exposure. Under ideal conditions, Berkson-like error increases the variability around the health effect estimate in a manner similar to pure Berkson error and does not induce bias, but Berkson-like error is spatially correlated and not independent of predicted exposures so it results in underestimation of standard errors. Szpiro and Paciorek (2013) analyzed the impact of Berkson-like error under more general conditions and found that it can bias health effect estimates either toward or away from the null. For example, in one simulation study where the spatial distributions of monitor and subjects locations were dramatically different, the health effect estimates were biased away from the null. In another example, where spatially structured covariates were included in the health model but not in the exposure model, the health effect estimates were biased toward the null. Hence,

Berkson-like error can lead to bias of the health effect estimate in either direction and should not be ignored. Classical-like error results from uncertainty in estimating exposure model parameters. It can add variability to predicted exposures and can bias health effect estimates in a manner similar to pure classical error, but it differs from pure classical errors in that the additional variability in estimated exposures is not independent across space.

*Proposed revised text to replace text that appears on page 3-65, lines 1-17:*
Minimization of error in the exposure estimate does not always minimize error in the health effect estimate. Szpiro et al. (2011) used simulation studies to evaluate the bias and uncertainty of the health effect estimate obtained when using correctly specified and misspecified exposure models. The correct exposure model was a LUR with three covariates while the misspecified model included only two of these three covariates. They estimated the exposure model parameters using monitor data and predicted exposures at subject locations. They studied two conditions: where the variation in the third covariate was identical in the monitor and subject data vs where it was much smaller in the monitor data than in the subject data. Szpiro et al. (2011) showed that prediction accuracy of the exposure estimate was always higher for the correctly specified model compared with the misspecified model. The health effect estimate had better properties (lower RMSE) for the correct model when the third covariate had identical variability in the monitor and subject data. However, when the third covariate was much less variable in the monitor data, then the health effect estimate had better properties for the misspecified model. The results of Szpiro et al. (2011) demonstrate one situation where use of a more accurately defined exposure metric does not improve the health effect estimate. (While the Szpiro et al. (2011) simulations were for a generic air pollutant , they are relevant to SO2.)

*Proposed revised text to replace text that appears on page 3-65, lines 18-36*:
Error correction is a relatively new approach to estimate the correct standard error and to potentially correct for bias in air pollution cohort studies. Szpiro and Paciorek (2013) and Bergen and Szpiro (2015) developed methods to correct for bias from classical-like measurement error by exploiting asymptotic properties of the variability in exposure model parameter estimates and propagating these variances through the health model by means of the delta-method. Valid standard error estimates are obtained by means of the non-parametric bootstrap. Methods have also been proposed to correct for bias from Berkson-like error, but these require stronger conditions including compatibility between subject and monitor locations and inclusion of spatially structured health model covariates in the exposure model. Spiegelman (2013) noted that the new measurement error correction methods developed by Szpiro and Paciorek (2013) are a version of regression calibration.

**New Reference**:

Bergen S, Szpiro AA: Mitigating the impact of measurement error when using penalized regression to model exposure in two-stage air pollution epidemiology studies. Environmental and Ecological Statistics. 2015, 22:601-631.

**Examples of other details to be revisited in Chapter 3:**

- P 3-2 sentence beginning line 11: Would it be appropriate to replace “some level” with “potentially substantial”?
- P 3-2 sentence beginning line 12: This statement needs a reference.
- P 3-2 line 20: Is it only uncertainty or also potentially bias?

- P 3-7 line 15: In what sense is the exposure error large?
- P 3-9 line 6: What is a spatial autocorrelation LUR model?
- P 3-10 lines 20-1: Do the authors intend to report odds ratios that are negative or essentially 0?
- P 3-13 lines 25-26: What studies use the population-level exposure assessments with time-activity diary data?
- P 3-13 line 31: Consider inserting the sentence: Thus most epidemiologic studies focus on ambient concentration rather than exposure.
- P 3-14 section 3.2.2.3: Please give some estimates of rate of entry and removal for SO2 so we can understand how important AER is for SO2. Overall it would be helpful in this section to discuss how much AER and microenvironmental exposures matter for SO2.
- P 3-26 section 3.3.1: Many other ISA's have covered this topic. Consider how to acknowledge and/or build on previously published material.
- P 3-27 sentence beginning line 26: Is this still true? Does it apply to all study designs?
- P 3-28, sentence beginning line 8: Do we need to speculate for SO2?
- P 3-28 line 15: Is this implying k is large? It would be good to include discussion of studies that show this and to relate to the quantities in the conceptual model for total personal exposure.
- P 3-41 lines 25-27: When is this assumption made?
- P 3-41 lines 37-28: It would help to reference this statement. Ditto p 3-42 line 3-4 (if it is true).
- P 3-42 lines 12-14: The size of the effect estimate varying by the increment size is not a real difference in estimates but rather a feature of the increment used for reporting.
- P 3-42 sentence starting line 15: I don't understand. References to back up the statement would help.
- P 3-48 section 3.3.3.3: Please explicitly justify including other gases in the discussion or omit them.
- P 3-49 section 3.3.4: Why not cite Paciorek (2010) in this section?
- P 3-50 sentence beginning line 9: This deserves a reference. It seems like a key point to demonstrate.
- P 3-58, section 3.3.5: A fair amount of reworking of this section is necessary. Some of my comments on the NOx ISA are relevant here. Each new ISA should build on the understanding developed in the previous one, particularly for concepts that are relevant to multiple pollutants.
- P 3-66: Is a discussion of panel studies even important for SO2? If none of the studies reviewed in chapter 5 are panel studies, then this section should be dropped.
- P 3-70 sentence starting line 6: Presumably in any individual study it is knowable how much certain (e.g. a minority of) sites may impact the inference in that study.
- P 3-70 line 33: At least for statistical models, it is not possible to derive models without measurements. So how are models an alternative to measurements?

**Additional comments on other chapters**:

- P 5-40: The analyses corresponding to the Sheppard references were done in Seattle WA. I did not revisit the health effects reported.
- Tables 5-32 & 5-38: Here is an opportunity to ensure the terminology for the exposure metric is consistent within the tables and across chapters (aligning with chapter 3).
- P 5-300: There are a number of topics relevant to Chapter 3 that also appear in this section of the annex.

- What is meant by central site measurements? Is it necessary for them to be correlated with personal exposure to be relevant? The exposure quantity of interest is ambient source personal exposure, rather than total personal exposure. It is unrealistic in most epi studies that central site measurements will be combined with time-activity data. Drop this reference to central site measurements only being useful if combined with time-activity information?
- What studies use exposure metrics that combine time-activity and microenvironment information? Have any such studies been published and discussed in Chapter 5?
- I am not aware of work that has been done to show the measurement error properties of atmospheric models are better (i.e. less severe) than for exposure metrics that rely on monitoring data. In my opinion I don't think EPA should advocate (even implicitly) the use of these models in epi studies in place of exposure monitoring data. We don't know how the uncertainty from these models will affect health effect estimates in epi studies.
- The relatively larger spatial variation in SO2 **should** be given more weight in long-term epi studies. And of course it is necessary to use exposure metrics that adequately capture this spatial variation. Exposure modeling approaches will vary in this regard, due not only to the approach, but also due to the amounts and representativeness of locations in the input data.

- P xxxix paragraph starting on line 26: This section needs work.
  - Broadly, it should be rewritten
  - More complex approaches can have more exposure measurement error due to greater uncertainty in all the inputs or components of the complex metrics, in comparison to simple metrics. This should be acknowledged.
  - I am not understanding the point intended in the last sentence, or given one interpretation, convinced of its veracity.

- P xl:
  - Generally review and improve the wording of the paragraph starting on line 29.
  - lines 18-20: This sentence is mixing several issues. Proximity to monitors assumes that one is using those monitors directly in the health analysis. Spatiotemporal variation of SO2 will be more or less of an issue depending on how much monitoring has been done. Time-activity is more or less relevant depending on whether one is focusing on ambient concentration or ambient exposure as the exposure metric in the epi study.
  - Verify the wording on lines 32-36.
  - P xli lines 2-3: I believe one could come up with a counter-example to the statement.
- P xlvi lines 29-31: Not a terribly helpful summary statement.
- P 1-8
  - line 1: A key point could be that the existing monitoring network is insufficient to capture SO2 variation for epidemiology. I'm not sure this came out in the document. This point should be relevant for time series studies as well since they typically rely on a single monitor or an average, instead of spatially-resolved monitors.
  - P 1-8 line 4: Clarify that the correlations being reported are for co-located monitors.
  - Line 15: Not only did this not come across clearly in CH 3, but I don't agree that this is generally an appropriate use of dispersion models. (This opinion may not be broadly shared. Regardless, an appropriate perspective about dispersion modeling for epi studies should be conveyed.)

- P 1-9
  - line 3: I would say personal monitors are only suitable for panel studies and for exposure validation studies. Also, are there any epi studies we are reviewing that use personal monitoring, or is this purely hypothetical? If the latter, this can be removed from this chapter.
  - Lines 5-7: I don't think this statement is terribly helpful for interpreting the health results.
  - Additional points pertinent to the first paragraph are the same in the executive summary.
  - Line 19: Wouldn’t lack of indoor sources make different studies more likely to produce more similar results? How would indoor deposition vary across building?
  - Line 21: How would personal activities affect indoor-outdoor relationships for SO2?
  - Line 23: It seems that bringing in lack of indoor sources of SO2 is relevant to this point.

# Dr. Frank Speizer

**Comments on the Preamble to the ISA**

I would like to complement the Staff on the current state of this document. Although I have minor thoughts as to some of the wording, I have watched this document evolve of the last 15 years or so and believe it now is an impressive consensus of both structure and form that is critical for going forward for the whole Clear Air Act, CASAC and standard setting process.

**Comments on Chapter 5**

General Comment: The organizational structure of the Chapter follows a logical sequence. A vast amount of data and publication were reviewed effectively and summarized. Although as indicated below there are some inconsistencies in the way the tables are presented I suspect that most of these arise out of the lack of details presented in many of the paper. One comment that I suspect will not be appreciated is my concern that as an ISA document we have fallen back into the pattern that lead from a Criteria Document to an ISA, which is to consider almost all publications irrespective of quality on an equal "need to include" basis.

Table 5.1: Suggest some slight reformatting of the table. I recognize that the studies are presented alphabetically by first author, it turns out that only the studies of Koenig et al summarized on page 5.9 have subjects under age 18 (need to check need to check Jones and Magnussen 1990, and Linn, 1983a; need also age range for Tunnicliffe 2003). All the other studies are in adults. Suggest you break out by 12-<18 and 18 and above as I believe that "life course" will become a defining issue as we move forward with more recent studies and if we do not have enough will be part of the remaining uncertainty.

Page 5-17. These two paragraphs don't match up. In paragraph 1 the discussion focuses on the fact that you have "responders and non-responders". In paragraph 2 if I read it correctly the authors are rejecting the data because the sum of the studies discussed in paragraph 1 are not homoscedastic. By the nature of responses they shouldn't be and thus this seems the wrong reason for giving Goodman et al no further consideration.

Page 5-21. Although the data are summarized elsewhere, it would be useful at the end of this, and I assume other section, to offer a concluding sentence as what the interpretation of the previous work is.

Page 5-25. In the Dales et al, last column it is not clear how to interpret the effect of Diurnal change related to 12 hour averaging time., Which 12 hours?

Page 5-26 - 5-27 In this study there are lots of tests reported with few being statistically significant. Do the authors (either of the paper or this document have a clear interpretation of whether the results are negative or positive? The text simply reports the findings.
Page 5-29-32. We are slipping into an old pattern of compulsive reviewing of the world literature, without selection: I see no reason for Table 5-7, as surely this is a table that appeared in the 2008 document and the results are described in paragraph above and last sentence in section on page 5-33.

Page 5-33, last paragraph. Too many decimal places are presented in OR and Confidence Intervals

Page 5-37-38. The discussion of the toxicological finding without presenting a table in contrast to the extensive review tables for the human studies seems striking, particularly because the tox finding sound impressive. Suggest add a table summarizing these studies discussed.

Page 5-39, Figure 5-2. Suggest reorder or divide into two separate figures to separate children only from adults and mixed. The details provided in Table 5-9 would also have to be reordered.

Page 5-42, para 1. Issue of ED visits potentially being less serious and caution on children less that 5. I suggest this be reconsidered. First the practice in EDs has been to watch and observe in the ED rather than admit does not mean that these are less serious cases. Particularly in children hospital admission are always traumatic and physicians have learned to manage very serious conditions, particularly asthma, on the ED floor. With regard to the <5 this is not an issue in the ED or hospital admission (as it might be on a death certificate) as clearly "wheeze" is wheeze when a physician sees a child. (this is discussed further on page 5-51 and might be included or modified here).

Page 5-43 48, Table 5-10. This table seems to belong in Chapter 3 rather than here as no biologic data are presented. In fact the text describing results from these sites uses summary statistics that really do not appear in table.

Pages 5-63-66. Another example of compulsive completeness 3 pages to say 2 studies are null!

End of this section before starting section 5.2.1.3: I find no summary that highlights the important observations discussed on page 5-56 on the delayed hyperallergenic response to prior SO2 exposure. I would have thought it should have been brought forward in summary of acute responses.

Page 5-71, end of line 5. Give reference for this statement. (assume it is Pencock et al).

Ditto: Could have left out lines 26-32.

Page 5-88 Summary of Respiratory Infections. I believe the interpretation is more conservative that it should be particularly as related to young children. Too many caveats are presented to minimize importance of findings. (Eg lack of multiple studies of same outcomes, effects of copollutants, and non discussion of lag structures, all issues that in the past had little impact of conclusions about major outcomes, and when subsequently considered have had modest modifying effects.

Page 5-105,part of Table 5-20. For the study of Steinvil, the effect estimates need more detail. I assume the FEV and FVC changes are in ml., however the numbers in the ratios of FEV/FVC make no sense since units are not specified.

Page 5-108, Table 5-21. Give units for effect in study. For Altuq study indicate what lung function was used.

Page 5-120 Table 5-24, for Zhao et al. Last entry on last column something missing for "Furry 3pet or pollen"

Page 5-122, line 1-3. The statement does not match up with the details presented in Table. Wheeze and nocturnal attacks of breathlessness are significant in the Table and are stated as no association in text.

Page 5-140 Conclusion. Mention should be made of the increased sensitivity to allergens as part of the conclusion.

Page 5-167-169 and Table 5-31. The discussion and table are incomplete in that what is left out is the evidence for Asthma exacerbation, particularly in children. I think another section after Asthma Development needs to be added as I believe the evidence for exacerbation is more than suggestive and actually causal. The summary of the evidence presented on page 5-169, lines 15-23 set up the individual with conditions that contribute to recurrent exacerbation, and thus are part of the causal chain for chronic effects.

Page 5-204 207, Table 5-39. In last column under Copollutants in each case where the SO2 association is mentioned as being robust or remaining positive it would be useful to add the word "significantly" if that is the case and if not add "not significantly"

Page 5-254, line 1-13. Reference is made to Table 5-44. It is true that the data are summarized there but because the respiratory effect are so different from the birth defects it would seem better to move this material to the respiratory section, or to make a separate sub-table so that it is not lost in the larger table of inconsistent and negative results.

**Comments on Chapter 6**

Overall I found this chapter almost useless. Most of the information presented is selected from previous chapters and much more adequately discussed in those chapters. The tables are for the most part incomplete in either not being fully representative of the data from the various studies, not presented in a consistent manner, and highly selective on which studies are included. I fear that the chapter will lead to more controversy than clarity and suggest it be left out. Table 6-16 could just as easily been presented as part of chapter 5 with the saving of 40 pages of text.

Page 6-1, line 20. Suggest change "increase risk" to "modified risk". Risk may be increased or decreased by factors. If accept idea may have to reward next paragraph as well.

Pages 6-6, Table 6-3 and discussion on pages 6-7, 8. Footnote a does not apply to all the studies, since outcomes are not all changes in FEV. I find this section as presented extremely weak for documenting the causal association with asthma. The table suggests that "convincing evidence of a causal relation is based on clinical studies in 152 asthmatic subjects (some of whom were not exposed to SO2 below .5 ppm)! This is followed by Table 6-4 which is not discussed at all. The causal argument is much better made in Chapter 5 and if it is necessary to be convincing in this Chapter would have to repeated in greater detail. I suggest greater consideration be given to the combination of all the data rather than suggesting "the primary evidence is from controlled exposure studies". (See specifically Figure 5-2 on page 5-39).

Page 6-9, Table 6-5. Footnote "a" seems to have been copied from the previous table and doesn't belong here as stated.

Table 6-6 This needs to be reviewed. States that dealing with older adults with pre-exisiting CVD, yet states dealing with adults 5-75. Either a serious typo of something else.

Page 6-10, line 8. What does Section A.B.C mean?

Page 6-12, line 17. As stated this is almost the opposite of what is said on lines 3 and 4 on the same page.

Table 6-8 (as an example of inappropriate selection of studies to report) Although I agree that the data in Chapter 5 are suggestive at best the studies selected here to put in the table would not lead one to make that conclusion. In fact the results are more consistent with a positive effect.

Table 6-10. Results as presented are quite misleading. When comparing older to younger children or children to adults for asthma admission or ED visits these may not be fair comparisons and since many of these same studies are discussed in greater detail in chapter 5, presented here in brief summary form is confusing at best and appears to be somewhat biased toward the null in the selection of the studies (particularly since in Chapter 5 the conclusion is that the association is causal.

# Dr. James Ultman

## Overall ISA

The EPA staff and its consultants should be commended for a well written document that lays out in detail what is known about the inhalation toxicity of SO2 and its relation to the AAQS. The overall revisions that I suggest are:

## Executive Summary

Shorten the Executive Summary to eliminate unnecessary details that are covered in Chapter 2.

General Comments

I believe that an executive summary should be to the point and contain "bottom line" information. As it stands, the Executive Summary restates to much of what is covered in chapter 1. For example, lines 11-25 on page xxxix concerning dispersion modeling in the Executive Summary are essentially the same as lines 8-24 on page 1-8 of chapter 1. I suggest that unnecessary detail be eliminated from the Executive Summary.

Specific Comments

Sources and Human Exposure:

- I could find no mention of background SO2 levels or their implication for standard setting.
- On page xxxix, it may be worthwhile to add a sentence regarding the correlation between a 5-minute maximum SO2 concentration and a 1-hr maximum SO2 concentration and its relevance to the use of a 1-hr standard.

Dosimetry and Mode of Action

- On page xli, is it possible to say something about the relative retention of sulfites in respiratory compared to systemic tissues and what this implies about the relative toxicity of inhaled SO2 to those tissues (see lines 22-26 on page 1-12 in chapter 1)?

## Chapter 1

Perhaps, there should be a more direct discussion of uncertainties in the Conclusions section of this chapter.

## Chapter 4

General Comments:

Response to charge questions: This is a well-written chapter that generally meets the objectives of a) being a complete and accurate description that provides a good context for chapter 5 on the health effects; b) contains a good discussion and comparison of inhaled, injested and endogenous pathways for

SO2/sulfite; c) provides modes of action for short-term and long-term respiratory as well as for extra-pulmonary effects. Following are some suggestions for reorganization and clarification:

In section 4.2.1 on SO2 chemistry, the aqueous chemistry of SO2 is described (Eq. 4-1), but a discussion of SO2 destruction of disulfide bonds in amino acids and proteins (sulfitolysis) is postponed until section 4.2.3 on SO2 distribution. Auto-oxidation reactions of SO2 to generate free radicals is not mentioned until section 4.3 on mode of action. Enzymatic detoxification of sulfite by its oxidase is not described until section 4.2.4 on metabolism. Perhaps a complete discussion of these different reactions placed in section 4.2-1 would be better.

Section 4.2.2 on absorption begins with a comparison of the absorption distribution of SO2 and O3. This discussion is based on the assertion that the upper to lower distribution of the two gases is determined by their relative solubility. This is not entirely true since the relative diffusion coefficient and relative reaction rate also play a role. For example, for a mucous layer that is relatively thick (*i.e.* large Damkohler number), the SO2 uptake rate is proportional to the solubility times the square roots of the diffusion coefficient and pseudo first order reaction rate constant. I think that a conceptual model of respiratory gas transport is needed at the beginning of this section; this should include a discussion of how overall uptake and uptake distribution depend on inhaled dose (concentration×minute volume) and gas properties.

Also in section 4.2.2 on absorption, ventilation is given in units of $m^3$/kg-day and $m^3$/min. The document points out that "ventilation rates normalized to body mass should not be confused with median daily ventilation" but the possible significance of normalization by body mass is not discussed. Since inhaled dose is the product of ventilation and inhaled concentration, units of $m^3$/min are an indicator of the maximum total dose that the respiratory tract as a whole will receive. Assuming that body mass is an indicator of lung size, normalization of ventilation by body mass is a crude indicator of the average dose per unit area that the airway film receives. I think that a brief discussion of the two ways in expressing ventilation would be worthwhile.

In the summary paragraph on absorption on the bottom of page 4-4, there is no mention of the effect of exposure concentration on percent absorption. I think it is worth saying that there is some experimental evidence indicating that exposure concentration plays a very small role compared to ventilation and mode of breathing (nasal or oral).

Specific Comments

| page | line/figure | Comment |
|---|---|---|
| 4-2 | 5-8 | To clarify Henry's Law and its relationship to solubility, I suggest replacing the sentence "Henry's Law shows…in the gas phase" with "The Henry's Law constant, defined as the ratio of partial pressure or concentration of SO2 in a gas phase to SO2 dissolved in a liquid phase, is an inverse measure of solubility." |
| 4-3 | 7 | Please add the phrase ", expressed as a percentage of amount inhaled," before the term "absorption" to ensure that the reader understands exactly what it meant. |
| 4-4 | 5 | The meaning of this sentence would be clearer if "also breath oronasally" |

was replaced with “also show an increased contribution of oral breathing”

4-11 29-34 Unnecessarily redundant with previous sections. I think that this portions of this paragraph could be moved to the beginning of section 4.2.1 on chemistry.

4-12 28-29 It’s not clear to me that “higher concentrations lead to greater penetration.” Is this statement based on experimental observation, a respiratory transport model or is it pure conjecture?

4-13 19 I don’t understand what is specifically meant by “mirrors.” It would help to be more specific.

4-22 20 Should “exogenous” be corrected to “endogenous?”

**Chapter 5**

As I read this chapter, I found it particularly difficult to see the “forest through the trees.” I think that this arises because the extreme length of the chapter. Perhaps it would benefit the ISA to divide the material in chapter 1 into two chapters. Perhaps, there could be one chapter on motality and morbidity effects of inhaled SO2 and a second chapter on extra-pulmonary effects of SO2 inhalation.

# Dr. Ronald Wyzga

**Comment on Chapter 5**

p. 5-61, ll 14-31.: Exposure error could also be due to differences in time averaging. The human clinical studies find responses occurring after exposures as brief as 5 minutes. Most acute epidemiological studies consider a 24 hour average. If the 24 hour period includes short exposure windows of 5 minutes or so with higher SO2 levels, these could be masked in a 24-hour average; hence the 24-hour average contains considerable measurement error in estimating the relevant short (5-10 minute) exposures that could be responsible for any health response.

p. 5-62, ll. 17-20: Measurement error can also distort the estimated shape of the dose-response function and mask the existence of any threshold.

p. 5-77, ll. 1-13: In consideration of the correct co-pollutants, there may be guidance in terms of looking at the relevant sources of SO2. For example, it is diesel emissions, there is a greater need to consider polllutants such as EC or OC. If it is powerplants, these co-polllutants are less important.

p. 5-99, ll 17-20: It could be misleading to generalize from this one study.

p. 5-103: This section could also include Klemm, R. J.; Lipfert, F. W.; Wyzga, R. E.; Gust, C., Daily mortality and air pollution in Atlanta: Two years of data from ARIES. Inhalation Toxicology 2004, 16, 131-141.

p. 5-141, l. 1: Is ).95 and OR or RR?

p. 5-143, Table 5-28: is that a minus sign in from 1.47?

p.5-144, Table 5-28: For Nishimura et al., are we uniformly talking about ORs?

p. 5-175, l. 6: Can you be more explicit about what is meant by “some”?

p. 5-179, Table 5-33: Is the risk or odds ratio for Ballester et al, for a single pollutant model or after adjustment? Similar question for Cheng et al.

p. 5-191 Szyszkowicz et al, were any organics considered in multi-pollutant models?

p. 5-208, ll. 12-29: See comments for p. 5-77.
l. 25: please give confidence interval.
l. 28: ‘ ‘ ‘ ‘

p. 5-219, Table 5-40: do we have results for other lags for Dubowsky et al.?
Section 5.5.1. you may want to add the following reference to the studies considered:
Klemm, R. J.; Lipfert, F. W.; Wyzga, R. E.; Gust, C., Daily mortality and air pollution in Atlanta: Two years of data from ARIES. Inhalation Toxicology 2004, 16, 131-141.

p. 5-315: l. 1: It is often difficult to distinguish between "suggestive" evidence and "inadequate" evidence. I struggle with the choice between these two for the cancer endpoint. There are a couple of studies (imperfect) that suggest a possible relationship between SO2 exposure and cancer; on the other hand, the available information is limited and frankly inadequate to make any judgment. More guidance on the distinction between these two categories would be helpful.

**Comments on Chapter 6**

*Charge Question: To what extent has the available scientific evidence from epidemiologic, controlled human exposure, and toxicological studies been integrated to inform conclusions on at-risk populations and/or lifestages? Is there information available on other key at-risk factors that is not included in the draft SOX ISA and should be added?*

Overall comments: By and large this chapter utilizes the information from Chapter 5 to identify populations and lifestages potentially at risk. I have three major comments on this chapter. First of all the chapter largely ignores the fact that indoor SO2 levels are largely very low; hence populations who spend more time outdoors have greater potential to be exposed to higher levels of SO2. Secondly behavioral factors can add to this potential; for example, asthmatics appear to be more sensitive while exercising and there is some evidence that asthmatic medication can be protective of response. These issues should be mentioned here and to some extent they may explain while children could be more responsive as they exercise more frequently and probably spend more time outdoors as well. The last issue concerns the definitions of the "Direction of Effect" symbols in the tables. They need better definition and perhaps they should be expanded to be more informative. For example, I doubt the "-" symbol really means "no difference" in the absolute sense; for example, Table 6-4 presents a "-" for the study by Amadeo et al.. If we go back to Chapter 5 (Table 5-21) we note that there are non-zero differences between exposed and non-exposed subjects, but these differences are far from being statistically significant. My question is what is the dividing line between a "–" and an up or down arrow? Is it statistical significance or some other measure? I also wonder if it would more informative to have additional symbols, such as a double arrow when the difference is very large (to be defined). The various tables or text could also indicate which direction of effect is considered adverse. For example, in Table 6-8 we see contrasting directions.

Chapter 5 gives different weightings to North American studies and those undertaken overseas. This chapter does not do likewise. If such a distinction is made, it could change some of the inferences. For example, Table 6-11, looks at the older adult lifestage; of the 17 studies considered, only 3 are from North America, and two of the 3 show no difference in direction of effect. The third study finds no difference for those 65-74, but it does for those >74 years.