Revised Preliminary Comments from Dr. Stephen Schwartz

This review is focused mainly on Chapter 2 dealing with sources, atmospheric processes, and deposition.

GENERAL COMMENTS

Precision of standard

I expect that there are historical reasons why the nitrogen oxide secondary standard is set at 0.053 ppm (annual average), and not a round number such as 0.05 or 0.06, the implied precision being some 2%. (this precision contrasts with the several other standards: 0.5 ppb for SO_2, 15, 35, and 150 µg m^{-3} for PM). This precision would seem to require a justification. As is made abundantly clear in the document there are issues of measurement, modeling, source strengths, deposition rates, chemical reaction rates and the like that have much greater uncertainty than the 2% implied in the precision of the standard. The threshold for damage to natural and managed ecosystems, structures and the like also has much greater uncertainty. Thus a standard with such precision seems to me virtually impossible to defend. In a rational world one would set the standard to a rounder number, but perhaps there are institutional or legal reasons why such a change would be difficult to support, it taking more justification to change a standard than to leave it unchanged. That said, perhaps some explanation could be given as to why such precision should be maintained.

Following up on that, I note that in Section 2 page 2-54 estimates are given for uncertainty in deposition flux of particulate matter:

> Uncertainties in depositional flux estimates in this approach result from the combined uncertainties in the satellite-derived surface concentrations and model-derived deposition velocities used in the flux calculations; average relative uncertainties are estimated to be ~30 % for both NO_2 and SO_2 over land.

Similar uncertainty is given for dry deposition of NO_2 and SO_2, page 2-78. Such uncertainties, which in my opinion may well be optimistic, certainly raise question to the appropriateness of a standard with the precision implied by a numerical value of 0.053 ppm.

A perhaps more general question is whether a standard on atmospheric abundance (mixing ratio) is the most appropriate *form* of the standard, as opposed to, say, deposition flux. (The four elements of a standard are nicely delineated at page xlviii, line 13.) ES page lxi states that the main findings of this ISA are related

to N and S deposition. I suspect that there is good reason for retaining atmospheric abundance as the form rather than deposition flux (ability to measure mixing ratio, versus great difficulty in measuring deposition flux); also phytotoxicity is probably more related to atmospheric abundance than to deposition flux. But the question of form does not seem to be explicitly addressed in the document.

It seems as if some justification is required for the very different averaging times for the two substances, specifically for NOx, annual arithmetic average whereas SO_2 is 3 hour average, not to be exceeded more than once per year.

Presentation of uncertainties

It is fundamental in science that in presentation of measurement results or model results, it is imperative to provide uncertainties associated with the measured or modeled quantities. I thus express the concern here that that requirement is frequently not met in presentation in the Chapter under examination. I elaborate on this concern below. I am particularly concerned with maps presented showing detailed geographical distribution of deposition (e.g., Figure 2-21, page 2-63) without accompanying map showing uncertainty in the quantity. In a few instances, such as Figure 2-33, page 2-79, such maps of uncertainties are presented. Such uncertainty maps should accompany all maps of concentrations or deposition.

An alternative means of assessing uncertainties associated with concentrations and deposition fluxes calculated by chemical transport models (alternative to examining the effects of propagated uncertainties in model inputs) is comparison of the results of two or more models. Although this approach is not given prominence in the Chapter, the comparisons between results from CMAQ and CAMx reported on page 2-80 indicate substantial differences:

> On an annualized basis, mean normalized errors (MNEs) in gas-phase concentrations ranged from ~25 to ~100%. MNEs in dry deposition were much larger and ranged from ~50 to >300% and MNE in wet deposition ranged from ~40 to ~100% with no clear preference for one model over another. MNE for NH_{4+} in dry and wet deposition ranged from ~35 to 70%.

Unless one or the other model can be established to be erroneous, inter-model differences serve as a measure of uncertainty in modeled quantities. To the extent that results from either model are used as a basis for standard setting, cognizance should be given to the uncertainties implied by these differences.

Separability of particulate matter from sulfur and nitrogen oxides

The title of the present document is "Integrated Science Assessment for Oxides of Nitrogen, Oxides of Sulfur, and Particulate Matter— Ecological Criteria" and it is stated (Executive summary, p. lxi) that

> The ecological effects of forms of PM, which are not related to N or S deposition, are included in Appendix D. The nonecological welfare effects associated with PM, such as visibility, climate, and materials effects, are considered as part of a separate review of PM (81 FR 87933, December 6, 2016).

This separation may be required for reasons not explicitly stated here. However, much of the generation of secondary organic particulate matter derives from reactions of natural and anthropogenic gaseous organic compounds reacting with oxidants (ozone, hydroxyl radical, NO_3 radical) whose concentrations are greatly enhanced by anthropogenic nitrogen oxides. Likewise, virtually all of the atmospheric sulfate results from anthropogenic emissions of SO_2 and sulfate, and much of the atmospheric oxidation of SO_2 to sulfate is influenced by oxidants (OH, H_2O_2) whose concentrations are enhanced by nitrogen oxides. Consequently there is an intrinsic coupling of the atmospheric chemistry governing sub-micrometer particulate matter and the nitrogen oxides and sulfur oxides under examination here. This coupling may require re-examination of the role of these oxides in formation of PM when the nonecological welfare effects of PM, which may be result in standards that are more restrictive than those related to ecological effects, are examined in the separate review referred to above.

DETAILED REVIEW OF CHAPTER 2

This chapter starts out, appropriately, with emissions. But it does not state why emissions are required (e.g., as input to models for concentrations or deposition), how they are required (location, seasonal, time of day) and how accurate they need to be. All of these requirements need to be set out quantitatively prior to any reporting of current emissions. Then emissions estimates should be presented with time and space scales appropriate to the requirements, and with uncertainties specified. I would think that for primary particle emissions it is insufficient simply to specify mass emission rates, but rather that a more differentiated picture needs to be presented, at minimum number and mass distributions of emissions. I would think that for some considerations size-distributed composition is important. Also optical properties, such as absorption coefficient. The appropriate unit might be $m^2\ s^{-1}$, evaluated as $m^2\ g^{-1} \times g\ s^{-1}$, which would be very pertinent to black carbon emissions from diesels, on- and off-road vehicles, combustion facilities, wildfires. To my thinking emissions is squarely in the bailiwick of EPA as other

agencies rely on emission inventories from EPA, for example in modeling the climate effects of tropospheric aerosols. To my thinking the treatment of emissions in the present report falls far short, either by inclusion or by reference, of adequately describing both needs for emissions inventories and the present status of such inventories relative to those needs.

Page 2-2, line 13 and throughout: "major species include..." Use of "include" here and in general is sloppy and reflects laziness. It implies that the authors have not done the necessary work to provide an exhaustive list. Better: "Major species are..." Then the reader has confidence that these are the major species and there are not others lurking out there. Even better if some indication of whether there are other species that the authors are aware of, and how much they might be contributing.

Page 2-3, line 3. Nitrous oxide is not included because of lack of reactivity, as stated correctly at page 2-1, line 19.

Table 2-1 presents emissions of several species by process, often to two significant figures, sometimes with a large leading digit, e.g., 0.74 for NOx by off-highway vehicles, which implies rather precise knowledge of the quantity, 1 part in 74, or 1.4%, which seems highly optimistic to me. It seems essential that such a table include estimates of uncertainty with clearly stated meaning of the uncertainty range. I call attention to the notation developed by IPCC in AR5 in which uncertainty ranges are qualified as to "likely" by which it is meant that it is likely that the actual value lies within the range (66% likelihood, i.e., central 66% of the pdf of the quantity, roughly ± 1 standard deviation) or "very likely (central 90% of the pdf, roughly 1.6 s. d.). It is wholly unacceptable not to present uncertainty estimates here. And it is essential that there be a transparent chain of reasoning and citations going back to the process by which these emissions are estimated. That is, it needs to be made clear that the quantities presented a summation of emissions calculated as emission factor times process rate, summed over processes (at least that is my supposition). So for each addend in the sum there needs to be an uncertainty, itself reckoned as the uncertainty in the emission factor times the uncertainty in the process rate. These uncertainties would then be propagated into the totals presented.

Further on emissions, the table is introduced (page 2-3, line 18) by the statement that the emissions are compiled from the NEI (National Emissions Inventory), with indication that the numbers come from states and other entities, but without indication of the *process* whereby the NEI numbers are generated. This should be explicitly laid out. The same para states that "For most sources, estimates are generally available

for all 50 states." This raises the question of what is the magnitude of omitted emissions, requiring statement more explicit than "most". Is it estimated that the emissions not included are 10%, 20%? This needs to be stated and justified. Line 19 states "emission estimates developed by U.S. EPA from measurements by source sector " but without any indication of what is meant by "source sector". Caption to Table 2-1 refers to "source categories". Are sectors and categories two words for the same thing, or different things? If the same, then decide which word to use. If different, then both have to be defined and distinguished.

Page 2-5, line 7, states "emissions from mobile sources have been overestimated by ~50 to 75% in the 2005 NEI " but does not get into the implications of this, other than the vague "raise concerns" at line 18. One hopes that this is addressed later. Para should state where in this document this issue is returned to and examined.

Staying with emissions some statement needs to be made at the top of the section as to why one wants total US annual emissions in the first place. Is it as input to models? In that case emphasis should be made on emissions as a function of location (and season, time of day), not just aggregated emissions. Last para on page 2-6 refers to chemical transport models in which emissions are optimized by minimizing a cost function containing contributions from the difference between model predictions and observations. This may well be a better approach, but it requires an inversion to yield emissions by location and activity to compare with bottom up models. If this is done, my guess is that the results are not highly constrained. And if it is not done, it seems to rule out any possibility of sensibly comparing modeled concentrations with measured, as the modeled concentrations are essentially derived from measured. So this needs to be spelled out. Some of these issues are recognized in that para. But the hard assessment of the accuracy of emissions relative to the requirements is not done here.

page 2-7 The para that introduces the discussion of atmospheric chemistry should make it clear *why* the discussion of atmospheric chemistry of nitrogen and sulfur species is being presented. Just says "included here"; "briefly recounted here".

Page 2-9 starts off: "Reactions producing more oxidized forms of nitrogen (NO_Z) " but in fact has reactions on which NO_2 is on the left hand side, reacting to form still more oxidized substances. The key reaction producing NO_2 is not even in the list of reactions: $NO + O_3 \rightarrow NO_2 + O_2$. This reflects poorly on the document. That said, it is not clear why these reactions are presented here in the first place.

The table of Henry's law coefficients 2-2 is accurate in the main, but misleading, as the uptake of acidic or basic species such as HONO and NH_3 is governed not just by Henry's law solubility of the gaseous species but also by the ionic dissociation reaction, which is quite rapid and the extent of which depends on solution pH. The Henry law coefficient given for HNO_3, 2.6 x $10^{6,}$ probably reflects acid dissociation. Values about an order of magnitude lower are given in the compilation by Sander Atmos. Chem. Phys., 15, 4399–4981, 2015 (note different units). To the extent that this number is used in models, this discrepancy should be noted and any implications discussed.

Page 2-15 presents observations regarding organic nitrogen but not clear why. Are the results of Cornell and Jickells consistent or inconsistent?

The top para on page 2-16 is qualitative and speculative. Is there any evidence to support this, and what fraction?

Section 2.4 commencing on page 2-16. Not clear why this is being presented. It seems very qualitative and old (figure from 1968). Seems like material being rehashed. What is the point?

Page 2-24, l 9 ff refers to so-called hybrid approach using satellite data together for column NO_2 with a model to get surface mixing ratio relative to measured column amount. (This method is used also for SO2, below). What is missing is comparisons of surface NO_2 mixing ratio by this approach versus in-situ measurements to assess the accuracy and biases, whether these biases depend on mixing ratio, etc., which can be examined and displayed only by suitable graphical comparisons.

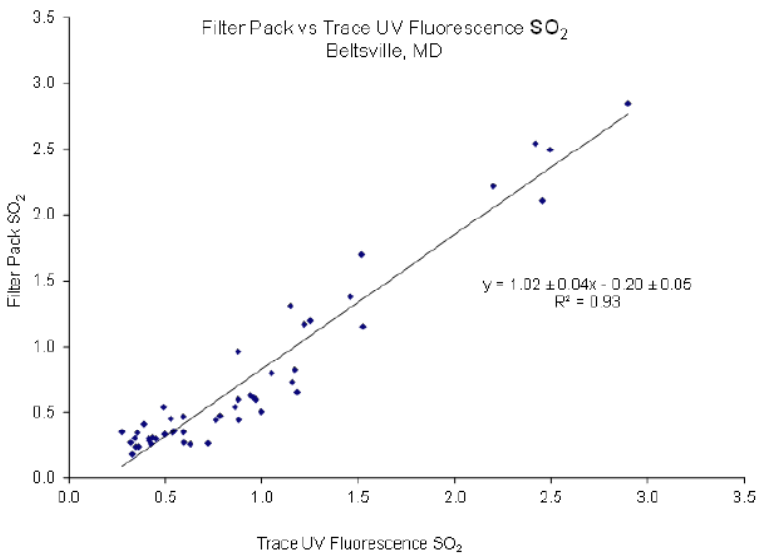
Page 2-30 presents results at rural areas noting that the pulse fluorescence method gives large relative errors at low mixing ratio of SO_2. I would ask whether this matters for purpose of compliance with standards which are much higher. Discuss.

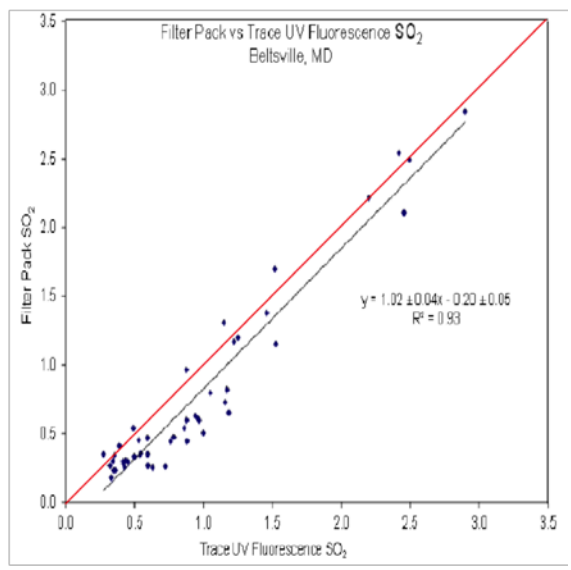
Page 2-31, line 3 states: "As can be seen in Figure 2-8, SO_2 is measured by the CASTNET filter pack by IC analysis of extracts from the cellulose filters. ". I cannot see this from the figure, which compares SO_2 measurements by two techniques. I might add that the figure scales should give the unit (ppb?). The language introducing the figure is hardly informative:

As can be seen in Figure 2-8, SO_2 is measured by the CASTNET filter pack by IC analysis of extracts from the cellulose filters. Because the nylon filter adsorbs some of the SO_2 (Sickles et al., 1999; Sickles and Hodson, 1999), SO_4^{2-} is also measured on nylon and added to the SO_2 (expressed as SO_4^{2-}) collected on the backup cellulose-fiber filters.

Results of an intercomparison of weekly average SO_2 data (ppbv) collected by the CASTNET filter pack and trace level SO_2 monitors during all of 2014 at Bondville, IL and Beltsville, MD are shown in Figure 2-8. (AMEC Environment & Infrastructure, 2015).

What is required is a statement of whether the two methods are consistent, inconsistent, which one is thought to be more accurate, and why. As a matter of style in graphics, the figure that compares two techniques should be squared up so that the physical lengths of the ordinate and abscissa scales are the same and the 1:1 line should be drawn so that one can see whether the data fall above or below that line. Then the implications should be discussed. Is there a systematic bias? How much? Is it important? Which is thought to be more accurate. Just presenting the figures is not very informative.





Left is as presented in the draft document; right is after scales are adjusted to be equal; red line denotes 1 to 1 line. Note how this presentation clearly shows the bias of one measurement against the other not evident in the original.

The document presents a lot of facts, for example, page 2-31:

In addition to the above in situ methods, satellite-based measurements have also been used to measure tropospheric SO_2 and to infer surface SO_2 concentrations with the aid of the GEOS-Chem chemistry-transport model (Nowlan et al., 2014; Lee et al., 2011). Tropospheric column abundances of SO_2 are obtained by the Ozone Monitoring Instrument (OMI) on the Aura satellite or the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) on Envisat and are combined with results from the GEOS-Chem, global-scale, three-dimensional, chemistry-transport model to derive surface concentrations of SO_2 (as they are for NO_2). Lee et al. (2011) associated annual mean surface mixing ratios of SO_2 derived from the hybrid satellite/model technique with ambient measurements of SO_2, (R_2 = 0.66 and 0.74, slope = 0.70 and 0.93, n = 121 and 115, for OMI and SCIAMACHY, respectively).

But the significance of those facts is not presented. What are the implications of the slopes being 0.7 and 0.9. Are they consistent or not? Which is right? The model or the satellite? Can the satellite measurements replace surface networks? Why does this matter? Of all the studies in the published literaure, why are the authors presenting the ones they have selected?

Reference is made to comparisons of annual mean surface mixing ratio from the hybrid technique compared to ambient measurements. One would certainly like to see the xy comparisons plotted as the example shown above. values of n are given: 121, 115 for two satellites; one wonders whether these are for different locations. One would wish to see whether there is bias at the high values of mixing ratio; for standard setting it would seem that the low values are less important. But perhaps even more instructive would be to see individual point measurements at specific times compared to satellite to get a sense of the spread of the data. Just stating the R^2 value and the slope conveys little information.

page 2-33 refers qualitatively to sources of error in retrieval of column measurements. But what needed is the sysematic and random error in the surface mixing ratio derived from the column measurement. And an assessment of the utility of the saellite measurements for the inferring the surface mixing ratio given those uncertainties together with issues of the verticcal profile. Just listing sources of uncertainties is insufficient. If this method is to be used with confidence, it is essential to see the comparisons. The discussion of the satellite measurements concludes, page 2-33:

> The errors in the column measurements result mainly from uncertainties in the vertical profiles of NO_2 and SO_2, cloud fraction, cloud pressure, surface reflectivity, and particles used in the calculation of air mass factor. A correction is required to account for NO_2 in the stratosphere (produced from N_2O oxidation and cosmic ray interactions dissociating with N_2). The SO_2 offset correction refers to a global background correction arising from issues in spectral fitting, such as spectral correlations with O_3 and stray light within the instrument.

But it is not clear what to make of this. What is the intended use of the satellite measurements, especially as the standards in question are for surface mixing ratios, not column abundances.

Again section 2.5.6, p 2-33 to 2-35 presents a lot of facts. But what is not clear is why. Is the question whether data from the other networks are of sufficient quality that they can be used to ascertain whether standards are being met? And after all the presentation of the methods, how do the measurements compare with EPA approved methods when side by side? Are the other methods good enough? I don't think the reader wants to know if it is a nylon filter or deionized water. The reader wants the bottom line. And that is missing.

Section 2.6 page 2-36 is headed "**Geographic Distributions of Species Relevant for Deposition**, which gives a hint to the motivation of the section. First sentence line 1 says "Maps of the average distribution of atmospheric species are presented in this section. " but it needs to state why they are presented. This should be explicitly spelled out. Once the motivation is presented then that dictates the requirements. How accurately must mixing ratios or concentrations be known, and of what substances and for what purposes (compliance or ascertaining of exceedances; calculation of dry deposition; human health exposure?), on what sorts of time scales? Deposition velocities are very much a function of atmospheric stability, surface roughness, and for vegetation, photosynthetic activity and the like. It would be a great error to calculate a deposition flux as the product of a long-term mean concentration times a single deposition velocity. Authors need to make clear the requirements.

Perhaps more importantly, in view of the use of models here, a short (two page?) description should be presented here of what the model consists of, what are the inputs, the processes modeled, the outputs, the uncertainties that can be ascribed to the modeled quantities. And clear distinction needs to be made between model output and observation. It seems to me that that is essential before presenting any blended product. Comparison between modeled and observed quantities is always good, including pointing out limitations of each. Perhaps then appropriate to show blended product.

Page 2-36 lines 2-4; The fact that this is modeled mixing ratio should be stated in first sentence; not in second sentence as a "However". Otherwise it is misleading to the reader. Start with "Modeled". Height (surface or height above surface) should be specified. It would seem a matter of taste whether to present modeled results before or after measurements, but I think that both should be presented, as well as any blended product.

Fig 2-9. It would be helpful to have a reminder what NO_Y consists of, especially for comparison with NO_2 in Fig 2-10. What fraction of NO_Y is NO_2? As Fig 2-9 is a model output and as individual species should be available from the model, why not show NO_2 explicitly so that there can be an apples to apples comparison with satellite derived NO_2 in Fig 2-10?

Page 2-36, line 8-11: Particles and gases have very different deposition velocities (up to several orders of magnitude), for gases strongly dependent on the substance and substrate, and for particles strongly dependent on particle size. The text implies that concentrations of nitrates are presented as TN = HNO_3 + pNO_3^-, gaseous nitric acid plus particulate nitrate. From a deposition perspective does it even make sense

to present the sum. Authors should report separately or justify. Again as fig 2-9 is a model output, show all NO_Y species individually as mixing ratio. Then it is straightforward to compute deposition flux for each of the several species, and it should become apparent that the flux per mixing ratio (mean deposition velocity over the period) differs substantially for the several species. Actually the mean dep velocity so calculated would be a kind of hybrid quantity because of correlations of mixing ratio and dep velocity. Again if the motivation is calculation of dep fluxes all this needs to be spelled out in detail.

page 2-37, line 5; reference is made to "hazard". Hazard by what mechanism? Rather strong word. but in any event if maps of each species one would not be speculating about the composition of the sum. At line 3 "subjected to concentrations < 1 ppb" seems rather inappropriate; "exposed" seems more neutral.

page 2-39 figure shows mass concentrations. Why not convert to mixing ratio so that the comparisons with mixing ratios in other figures is more transparent.

Figs 2-11 and 2-12 are the scales chosen to be the same to permit comparison? This might be stated. The text p 2-37 describing these and other figures in this section is much too sparse, frequently just stating what is evident in looking at the figure. What is the significance of the quantities being reported? The reader is at a loss because the motivation of the presentation of the geographical distribution has not been given. I contrast the same scales used in the above figures with the quite different scales (and mode of presentation) between Figures 2-13 and 2-14. Would it be useful to show a map of the fraction of NH_3 + NH_4^+ that is one or the other to make the point.

p. 2-40, line 9, increasing at a rate of over 5% per year. Is this worth a figure? One would like to see the trend over a period of time. Ditto for other quantities.

page 2-42 ; line 1: 15.07 µg m^{-3}. Four significant figures? More importantly, what is the distribution? Is it one high value; are the high values spread over the year or seasonal. A time series here would seem especially apt.

Page 2-44: "fusion of monitoring data obtained at Clear Air Status and Trends Network sites (black dots) and Community Multiscale for Air Quality model system results." Some explanation seems required here. How is this fusion done? One would hope that the values are anchored by the measurements and that the model is used to interpolate. Is the interpolation done on a day by day basis? Other details seem essential. What governs the scales here an in Fig 2-16. They are quite different. Is this to encompass the range of

observations or to facilitate comparison. What is the significance of the blank area at the Idaho Oregon border; also Big Bend area,? No conversion of Sulfate to ppb is given; why not? I would remind the authors that mixing ratio by mole is equally valid for particulate matter as for gases. But the reader should not be required to do this in his/her head. Suggest mixing ratio throughout.

Pages 2-48 to 2-52 give a nice précis of dry deposition theory and measurement approaches. About the appropriate length and detail for this Assessment. (This précis stands out in comparison to the lack of such a précis for the modeling on which presentations of results are given in figure 2-9.). The précis also speaks to the compromises and concerns with various measurement techniques. What seems to be absent, however, is an assessment of the uncertainties and systematic errors that result from the approaches taken. One is thus presented with Tables 2-4 and 2-5 without a sense of uncertainty associated with the quantities. For SO2 comparison of the two tables shows that the dep velocity ranges from 0 over snow in winter, 0.1 over grassland in winter, 0.6 for dry grassland in summer, 1.0 for wet grassland, 3.0 for wet deciduous forest in summer, and so on. So it is clear that the value 0.8 given in Table 2 as an average has quite a range of variability. One assumes similarly for other species. So the question that does not seem to be addressed is how is this variability treated. Clearly there are two or three terms that contribute to the variability: surface properties, canopy properties, and atmospheric properties (mainly stability). So the variation in reported deposition velocities is a consequence of all three. There may be compensation point issues with respect to NH_3. One would hope that the calculations do not simply use the seasonal and vegetation type overall deposition velocities, but work with the variations in the governing resistances to calculate fluxes. The text (p 2-53, line 1) acknowledges the possibility of positive or negative correlations of concentrations with dep velocity and resultant errors, but does not seem to quantitatively address the magnitudes or consequences in calculated deposition fluxes. Some estimation of these would seem essential.

p 2-53, lime 7, "relatively short." Always better to be quantitative. For example, if transport velocity is 5 m s-1, then the distance scale is 500 km; maybe relatively short on a continental scale, but relatively long on an urban scale and relatively long compared to time scales of dispersion of point source plumes. Ditto "nearby sources" line 8. Much better if one were to say that that distance sets the scale of influence of sources.

p 2-53 line 9 to 2-54, line 3. These points (regarding temporal variability and correlation) are quite important, but what is missing is a description of how cognizance of these points is reflected in modeling

of deposition pertinent either to calculation of the rate of removal of substances from the atmosphere or to calculation of deposition fluxes.

page 2-54, line 14. The authors are commended for presenting estimates of uncertainties here. One might like to see apportionment of the uncertainty to atmospheric concentration and deposition velocity. Also some description of how the time variation of both, referred to in the previous para, is dealt with in the calculations. Uncertainty estimates such as these get to the heart of the question of whether the implied precision of a standard expressed as 0.053 ppm is justified.

page 2-55, Figure 2-18. I am puzzled by the figure from Lin. It would appear that the deposition velocities should be proportional to the ratio of flux to concentration; as all quantities are on log scale, that ratio should be proportional to the distance between the curves for mass flux and mass conc. This does not seem to be the case in this figure, so perhaps it is not the best figure to illustrate the point.

Page 2-56 – 2-57. The assessment of the state of understanding and model representation of deposition velocity of particles seems fair. The question is then the implications of these uncertainties.

Page 2-58, line 4: The acidity may be much more a function of geographical location (Arizona vs New England) than of whether the clouds are precipitating.

Page 2-58. Not clear why the discussion of throughfall; suggest motivate or omit.

Page 2-61, line 5-6, corrections such as those noted here for stickiness, here of nitric acid on inlet tubing by a factor of 1.62 should be viewed with caution. Is this important in the interpretation of deposition? If so, perhaps it should be flagged.

Page 2-61, line 7 ff. This section presents deposition maps generated by the hybrid approach (model and observations). There seems to be much merit in this approach. One is interested in the deposition as a function of time and location, but the measurements are sparse. Hence the value of using a model as an interpolation mechanism. However, If the approach described here is important to the assessment, then it would seem to require much closer scrutiny. Terms like "bias corrected modeling results" need to be fully explained and the approach assessed. Ditto "fusion of data" from different networks. There are a variety of statistical tests to assess bias between different data sets. It would seem necessary to assess such bias before fusing the data sets. Systematic differences would need to be resolved or otherwise would

contribute to measurement uncertainty. What seems to be missing in all this is estimate of bias and uncertainty, at time and space scales relevant to decision making on suitability of standards.

I am concerned about the approach used in Schwede and Lear (2014a) that is the basis of the results presented here and that presumably are used for standard setting. Examination of Figure 2 of that paper shows that weekly average dry deposition velocities (from CMAQ output) are used with weekly observed air concentrations to calculate weekly average dry deposition for each species. Because of correlations (anticorrelations) between deposition velocity and concentration over such extended periods there are inevitable errors associated with such a procedure. The question would be the magnitude of such error. I could well anticipate that it could be factor of 2 or more. This sort of question is quite amenable to examination from the time series of the deposition velocity from the model together with time series of mixing ratios of say NO_2 or SO_2 available from real-time instruments. Such an analysis would seem essential either to be included in the present document or by reference to primary literature. The magnitude of the uncertainty and bias resulting from such a procedure seems essential to inform any standard setting based on this procedure.

Page 2-61, line 26 ff. Similar concerns as with the estimation of dry deposition amount. The statement "estimates of dry deposition could be obtained using CMAQ evaluated by comparison with monitoring results" raises the question of whether this has been done; what are the results. Again the magnitude of uncertainty and bias seems essential to the use of the results.

Page 2-62, line 1 ff. I note concern regarding the maps of deposition produced by the procedure described in the foregoing paragraphs and in Schwede and Lear. Evidently the authors of the assessment are similarly concerned (lines 7-8): "it should be remembered when viewing these maps that model estimates are subject to uncertainty, and for many parameters, comparison to observations is still needed." That said, it is essential that such comparisons as are available be shown here. Only by such comparisons can one get a sense of the magnitude of errors and biases in the approach. Further, it seems to me that a statement that comparisons are needed is inappropriate here; this document is meant to be an Assessment of present knowledge and understanding pertinent to standard setting, not a statement of required work.

The map of deposition shown in Figure 2-21 is exemplary of many maps shown in the body of this chapter and in the Appendix. Rather high spatial resolution showing patterns over a variety of geographical scales as described in the text on page 2-63. However ***the recipient of this document should not be misled into***

ascribing the quality of the formatting of the map to the quality of the data being presented in the map. ***It is essential when presenting results of measurements or calculations or hybrid quantities such as those presented here to show the associated uncertainty. It is thus essential to show a map of the uncertainty associated with the quantity itself. I would assert the necessity of such an uncertainty map for each deposition map that is presented in this chapter.*** The paper of Schwede and Lear that serves as the basis for the deposition maps presented here is likewise silent on the magnitude of uncertainty associated with the calculated quantities. Although systematically examining uncertainty that results from a procedure such as that presented by Schwede and Lear is non trivial, nonetheless it is essential that such uncertainties be estimated and presented along with the results.

I note that uncertainties are given for dry deposition flux in Figure 2-33. This proves it can be done.

Page 2-64: Are these two figures complements of each other? If so omit one.

Page 2-66, line 7: "uncertainties for dry deposition are likely much larger than for wet deposition. " This statement really calls for quantitative assessment specifying the estimated uncertainty and justifying the estimate.

2-66, line 9 . "the assumption was made that 80% of pNO_3^- is in the fine mode and 20% is in the coarse mode ". Presumably this is based on observations, but it might be expected that this proportion is not a constant but varies with time and space. This would suggest the utility of ascertaining whether the deposition flux in critical areas is appreciably affected by this assumption, as noted in the remainder of that para. What is missing is the consequences of the assumption, which goes beyond the additional uncertainty noted in the conclusion of the para. Yes there is uncertainty in the actual deposition, but there should be certainty in the consequences of the assumptions in the model, and an assessment, ultimately, of sensitivity of the proposed standard to those assumptions.

Page 2-67, Figure 2-24. This is a very informative figure. The implication is that for most of the CONUS the great majority of deposition is due to explicitly modeled species, presumably NO, NO2, NH3, HNO3, nitrate. It would be valuable to have the text explicitly state this rather than the reader having to infer from the list of what is not explicitly modeled. The text at line 2 refers to "oxidized nitrogen species, whereas the figure caption says total nitrogen; this needs to be clarified. But my read of the figure is that deposition by the explicitly modeled species is at least 70% of the total in most of the CONUS. As the modeling is

probably no better than that 30%, perhaps to first order deposition by species whose deposition is not explicitly modeled can be neglected. Perhaps the assessment might explicitly state that. The text does state (lines 4-5) that the not explicitly modeled species can contribute substantially to the total in the vicinity of large urban areas. It does not address the consequences of this for the purpose of setting the standard. I am somewhat surprised at the finding, as most of the species whose deposition is not explicitly modeled are secondary, including organic nitrogen species, which I would have expected to be a larger proportion of the total well downwind of source regions. Perhaps this can be discussed.

Page 2-69, Figure 2-26. I am rather surprised at the high proportion of sulfur deposition by dry deposition coming out of the model calculations. It would be valuable here to have references back to observations that support this conclusion.

Page 2-69. The two page spreads showing the 2000-2002 and 2011-2013 panels would be more effective with a third panel showing the difference. As above some indication of uncertainty should be shown. Because of cancellation of systematic errors it might be that the uncertainty in the difference would be less than in either of the quantities themselves. Such difference plots are provided for wet deposition, Figures 2-28 – 2-31.

Page 2-71, Figure 2-27. ***This figure is an astonishing tribute to the effectiveness of the clean air act amendments, and should be an icon to the effectiveness of this legislation and to EPA***.

Page 2-79, Figure 2-33. I commend the authors on including estimates of uncertainties here. It appears as if dry deposition flux uncertainty is roughly proportional to flux. In this case would it make sense to express as fractional uncertainty to get a much smoother field? Appears to be about 30% for NO2; 50% for SO2, similar to statements at page 2-78, line 11. It would be useful to state how annual average dep velocities are calculated. And for that matter is annual dep flux calculated as mean dep velocity times mean conc, or is it the sum of dep velocity times conc for shorter intervals; and if so, how short, and how is the shorter term anticorrelation dealt with? To what extent is the uncertainty in dep flux due to uncertainty in conc, and to what extent uncertainty in dep velocity. This assessment might guide future research.

Page 2-80. I am quite uneasy over the utility of transference ratios based on annual average concentrations and deposition. I suggest that any such results be carefully scrutinized. Evidently the authors of the

Assessment are similarly skeptical. The strong difference in transference ratios between the two models in Figure 2-34 is further suggestion that the approach not be used in standard setting.

Setting aside the transference ratio approach, with respect to the use of modeled concentrations in the Assessment, I note the observation at line 30 of mean normalized errors in gas phase concentrations, apparently between simulations in CMAQ and CAMx of 25 to 100%, and in dry dep 50 to 300%. These errors should on face be the source of grave concern in using the modeled quantities for standard setting. I am concerned at the apparent lack of taking cognizance of differences between the models of such magnitude throughout the chapter. It would seem essential to examine the differences between the modeled concentrations and deposition fluxes in scatter plots similar to those given in Figure 2-34 for transference ratios. This seems essential. It would be useful also to show maps of the concentrations calculated by each of the models together with a map of the differences.

WHAT IS MISSING FROM THIS CHAPTER

What seems to be missing from the Assessment is bottom line analysis of present status (mixing ratios, deposition fluxes) relative to a situation that meets various standards or other requirements such as avoidance of some level of negative effects on ecosystems. Are concentrations in compliance with present standards or out of compliance, and by how much? What are the implications? To achieve compliance do emissions need to be reduced, or alternatively, is there latitude to allow some increase in emissions. Answering the latter questions is a most suitable application for models, provided cognizance is taken of uncertainties. Almost certainly there are multiple ways in which compliance can be achieved: trade-offs between more or less stringent emission requirements in different regions. To my thinking it would be of enormous value to the policy-making community that is the customer for this assessment that the assessment show how far we need to go, and how to get there in order to achieve compliance with present or proposed concentration standards or maximum deposition fluxes (or alternatively, by how much and where emissions can be increased without incurring exceedances). As well, the assessment should provide an evaluation of the current state of the art in such modeling, e.g., 10%, factor of 2, or the like, taking into account uncertainties in the parameters in the models and various structural uncertainties in the models.

Terminology and style

Page 1-12, footnote. The footnote appropriately justifies the use of the term concentration to denote abundance expressed either as mass per unit volume or mixing ratio as "firmly entrenched in the literature."

Page 2-1 line 10. In an attempt to define nitrogen species incorrectly defines "nitrogen oxide" as NO. The common chemical nomenclature for NO is "nitric oxide." In fact the report gets it right usually, e.g., Table 2-1, Figure 2-1, Table 2-2, but occasionally reverts to "nitrogen oxide" section head for 3.5.1, page 3-14. These errors or inconsistencies can only lead to confusion and to an appearance of lack of attention to detail in the document. Please fix.

The term "acid deposition" is to be preferred to "acidic deposition". Here "deposition" is a noun formed from the verb "deposit"; the acid is the implicit object of the verb deposit. This is to be distinguished from deposition being a noun qualified by the adjective "acidic" as to what kind of deposition it is.

Throughout: Concentration seems to be used interchangeably with mixing ratio seems to be used. For example page 2-37, "broader areas of high concentrations (>~5 ppb). " At worst this can lead to confusion; at best it is an indication of sloppiness in presentation. Especially in the context of conversion between concentration and mixing ratio, e.g. p. 2-39: "Concentrations of nitric acid (µg/m3) can be converted to mixing ratios (parts per billion) to rough approximation at normal temperature and pressure by multiplying by 0.38. "

"Elevated levels"; egg p 2-40, line 4. Care needs to be taken so that the reader does not think one is speaking of vertical dependence. Ditto page 2-59, line 29.

Page 2-57, line 26 "cloudwater chemistry"; better "cloudwater composition". Still better: "Cloudwater composition and occult deposition have been measured..." That said, the term "occult deposition" is deprecated; better "Cloud drop impaction (on vegetation)". I note at page 2-62, line 18 the term used is "cloud deposition", much to be preferred.

Figure color bars: In many of the figures, e.g., 2-26, the scale runs from 0 at the top (blue) to large number (here 100) at the bottom (red). The color scale is consistent with expectation, but one generally expects such a scale to run from low values at the bottom to high values at the top, as for y-axes on graphs.

Citations. Citations are generally appropriate. However an exception is the citation Sutton et al. (2011) on page 2-9, which is to a table of Henry's law coefficients adapted from a table in the book by Sutton, for which the citation is

Sutton, MA; Howard, C M; Erisman, J W; Billen, G; Bleeker, A; Grennfelt, P; van Grinsven, H; Grizzetti, B. (2011). The European nitrogen assessment: Sources, effects and policy perspectives. In M A Sutton; C M Howard; J W Erisman; G Billen; A Bleeker; P Grennfelt; H van Grinsven; B Grizzetti (Eds.), The European Nitrogen Assessment: Sources, Effects and Policy Perspectives (pp. 664). Cambridge, UK: Cambridge University Press.

As this was an obscure source and as I wished to check a particular value I went to some effort to get the book. When I got the book, I found the table on page . The table gives as its source the widely used compilation by Sander, of which the most recent version is readily available,

Compilation of Henry's law constants (version 4.0) for water as solvent. Atmos. Chem. Phys., 15, 4399–4981, 2015 www.atmos-chem-phys.net/15/4399/2015/ doi:10.5194/acp-15-4399-2015

So a lot of trouble to get a readily available citation if the authors had cited the paper by Sander. Much better in general to cite readily available sources rather than much less available book citations.

Production question

In reading the pdf file on the screen I notice pop-up windows that provide explanatory and/or qualifying information regarding the figures. Is there some intent to make this information available in hard copies? Which is the copy of record? Hard copy or electronic?