Comments and Responses to May 6, 2009 Version of BAAQMD Multi-Pollutant Evaluation Method Technical Document

In May 2009, the Air District sent a preliminary draft version of the Multi-Pollutant Evaluation Method (MPEM) Technical Document to experts in this area, including Dr. Jane Hall of Cal State Fullerton, Donald McCubbin of UC Davis, Dr. Robert Harley of UC Berkeley, and Dr. Bart Ostro of the Office of Environmental Health Hazard Assessment (OEHHA). Written comments were provided via email by reviewers Hall, McCubbin, and Harley. Their comments and Air District staff responses are summarized in this document. *Comments provided by reviewers are provided in italics*. Air District responses in normal font. These comments have been addressed in the June 2009 version of the Draft MPEM Technical Document which is posted on the Air District website at:
http://www.baaqmd.gov/Divisions/Planning-and-Research/Plans/Clean-Air-Plans/Resource-and-Technical-Documents.aspx

Comments submitted by Jane Hall, Ph.D.
Professor of Economics, California State University Fullerton:

Regarding thresholds for criteria pollutants, in the absence of agreement (or documentation) in the health literature, I would use background, even knowing that this is not the regulatory target. The reason is that benefits do result from going below the standards, and meeting the standards pushes concentrations closer to background in some locations and during some seasons of the year.

Assuming health effects to the background is essentially what we're doing. We're assuming no threshold for toxics or $PM_{2.5}$ and a 50 ppb threshold for daily maximum 1-hour ozone. This is at the high end of the background ozone distribution. One difficulty with lowering this value is the limitations of our current modeling. We've only modeled a limited number of high ozone days; our estimates of the change in ozone per change in precursors may not work well for days with ozone near background. Specifically, there are days and locations where ozone concentrations at some District sites are below background. Reducing precursors may well increase ozone on those days, but using our model results – which extrapolates down from ozone episode days – may predict further ozone reductions.

Regarding the NOx emissions- ozone concentrations relationship, there is a large and somewhat contentious literature on this. You need an atmospheric scientist to look at this. Indeed there are some ranges over which the isopleth is backward bending in some locations.

Our modeling shows NOx disbenefits in some parts of the Bay Area for some episodes. Ambient monitoring data also suggest that NOx reductions may cause increases in ozone at some sites under some conditions. Our modeling data is limited, but this is the best we have at the moment.

Here is a citation to an article on GHG mitigation and related health effects, specifically relevant to criteria pollutants. It is a very useful reference.
http://www.sciencemag.org/cgi/content/summary/293/5533/1257

This suggests a potentially valuable application for the MPEM. We do talk about how the MPEM can show the benefits and disbenefits of tradeoffs, but this suggests we should also talk

about possible synergies. We added the following bullet item in "Applications" section of the Executive Summary:

- Evaluating the benefits GHG measures in reducing criteria pollutants

Rather than rely on referring to the CARB 2008 C-R function for PM-related mortality, I would go back to the underlying studies. You end up in the same place, but it is more robust – especially when you use the original articles as the basis for end points otherwise.

This is a good addition to our list of changes for Methodology Version 2.

Comments submitted by Donald McCubbin, Institute of Transportation Studies
UC Davis, CA

On page 13: "Generally, we want to consider joint effects, but the model does not provide these directly." And on page 14 "Since we are concerned with all precursors, we attempt to estimate the marginal effects." You might say a bit more about this (rationale, pros, cons), because this drives what you do next with your regression equation. Do you want to consider joint effects because generally you are reducing multiple precursor emissions with the same policy? If you are looking at ammonia from agricultural this might not be a very good assumption, since agriculture is not a big emitter of much else other than ammonia. Joint approach is fine, just might discuss it a bit.

We agree; the rationale was not clear. It was based on the argument for using coefficients from a multiple regression rather than individual regressions, namely that the multiple regression coefficient represents the effect of one predictor with the others held constant.

Also, we should have used "partial" rather than "marginal".

Here is the revised paragraph:

" **Individual vs. Partial Effects:** There are potential differences between how the concentration of a species is affected by the change in a precursor when considered in isolation vs. when considered with other precursors, analogous to the difference between a simple derivative and a partial derivative. Since we are concerned with all precursors, we attempt to estimate the partial effects. That is, we attempt to estimate the effect of changing a given precursor *with the others held constant*, rather than its effect considered in isolation. If the effect on concentration of various precursors is linear, then using these partial coefficients will correctly estimate the joint effect on concentration of a given set of precursor reductions."

Minor typo in footnote (9). Should be 1am - 8am (not 8pm). I would probably write it: 1:00am-8:59am.
Response: 9 now reads: "Ozone measurements are collected as 1-hour averages. To compare with the standard, running 8-hour averages are computed for each day starting from midnight-7:59am, 1:00am-8:59am, and so on, then the maximum among these is obtained. These values are computed on a site-by-site basis."

Page 20: "The factors for adjusting from individual to joint effects were 0.85 for ammonia, 0.90 for sulfur gases and 1.03 for sulfate." I am curious as to the physical significance of these coefficients. I guess NH3 depends on sulfate (whether directly emitted or formed from SOx gases), so if there is less sulfate then NH3 would have less of an effect, hence the 0.85 coefficient. I am not so sure about sulfur gases. Does the model allow for partially neutralized sulfate (NH4SO4) and sulfuric acid aerosols? You might mention these, if only to say that sulfates are generally completely neutralized in the Bay area, or whatever.

We can't answer the physical significance question. We believe that the model allows for partially neutralized sulfate and sulfuric acid aerosols but these should be a negligible component in the Bay Area. Ammonium sulfate is a smaller fraction of $PM_{2.5}$ than it is in the East and it forms preferentially to ammonium nitrate. We haven't measured sulfuric acid, but we found negligible amounts of nitric acid in the one study we did.

Page 21, the results for ammonium nitrate are very curious. I could see SOx reductions increasing (Ansari & Pandis, 1998) ammonium nitrate because it releases ammonium that could then react with gaseous HNO3. Is there an explanation? Also, I am curious how VOC affects ammonium nitrate. Minor quibble, you might mention that the 80/62 is the ratio of the molecular weights. I looked at Appendix E. I did not find any explanation for these results. What was done by way of model validation? Maybe this is discussed and I missed it. The modeling is so important that more effort should be made to show that one should believe the model results. By the way, the table (E6) with the changes in ammonium nitrate, might be more intuitive if the positive were associated with increases in ammonium nitrate and the negative with decreases.

Since sending you the 5/6/09 draft we've done considerable analysis of ambient data including comparing weekend nitrate concentrations to weekday. The weekend levels were significantly lower, in contrast to the model predictions. Based on this, and other analysis we are replacing the model results with a positive coefficient, specifically, we are using an across-the-board factor of 0.032 to convert NOx (ug/m^3) to ammonium nitrate (ug/m^3). The details are presented in Appendix F.

The daytime chemistry of ammonium nitrate formation has similarities to the formation of ozone. This includes reactions between VOCs and NOx.

Response to "minor quibble": We added footnote: " Nitrate, NO3, has atomic weight 62. Each nitrate molecule combines with an ammonium (NH4) molecule, for an atomic weight of 62 + 18 = 80."

*Page 22: "Our method estimates 'backyard' exposure.." And on page 23: "In the Bay Area, at least, air quality monitors tend to be placed in areas with above-average concentrations. Thus, monitored values are likely to overestimate average backyard concentrations." I agree that monitors tend to be put in problem areas. However, you appear to be using these very same monitor values for your interpolation and the model values are used to calculate a percent change from this interpolated value. There may be more details, but I think that is the essence of it. So the argument that your backyard exposure estimates lead to a downward bias falls flat. Unless, perhaps, the interpolation method uses the model values to scale the monitor values? (E.g., if a monitor is 80 ppb, the model cell containing that monitor is 60 ppb, and the target model cell is 30ppb, then interpolated value would be 40ppb = 80 * [30/60].)*

The paragraph on page 23 was unclear. There are two different issues. One is that the health studies that produce the C-R functions generally rely on monitored data and *these* values may exceed what those same studies would have found if they had had unbiased backyard exposure values available.

We agree that if monitored values are used to interpolate to backyard values, then this would cancel out the bias. This is the approach we used for ozone. For toxics and PM2.5 we modified our approach, using modeled concentrations. Thus, for toxics and $PM_{2.5}$ the point still seems valid. The paragraph now reads:

"In the Bay Area, at least, air quality monitors tend to be placed in areas with above-average concentrations. To the extent that this is true in the areas where C-R functions have been calculated, this would cause an underestimation of the response for a given concentration, provided that unbiased estimates of backyard concentrations were used. For $PM_{2.5}$, we used modeled values, which may be unbiased. Thus, for $PM_{2.5}$ C-R functions, the response may be underestimated. For ozone, monitored values were interpolated to backyard values thereby approximately canceling the bias. That is, the backyard ozone values are likely to overestimate actual backyard ozone concentrations, thereby roughly canceling the presumed underestimate in the C-R functions."

Page 23: "This is a known regression theory result." You might give a cite for this. (I know this has been noted in the epidemiological literature, but I can never find this particular cite when I want it!)

We were unable to find a citation. We did our own analysis suggesting that, in fact, the result may not always hold, but does hold in a large class of cases. See Appendix G for our analysis. We modified that paragraph as follows:

"Exposures were estimated with error. If exposure were estimated without bias[1], but with error, then the C-R coefficient would tend to underestimate the effect of the pollutant on the health endpoint. This is a regression theory result, where if the independent variable, x, is measured with error: x* = x + error, and the error has zero mean, then the fitted regression slope, b*, of the regression of y on x* will tend to be less in magnitude than the true slope, b, from the regression of y on x (had it been known). That is $|b^*| < |b|$. See Appendix G."

Page 24: "The health impacts from air pollutants are generally estimated in lab analyses where the risk is compared with true exposure, yielding an estimate of risk assuming given a true average lifetime exposure." This might be more carefully written. Two things come in mind. First, the unit risk estimate for benzene is not lab-based -- I think it is based on occupational exposure data, which is notoriously bad. Also, diesel is based in large part, I believe, on occupational studies (http://www.arb.ca.gov/toxics/dieseltac/de-fnds.pdf). So, that's two out of the five toxics. Second, even if the unit risk estimate were based on lab results, the lab results are for rats. To extrapolate to humans is not very precise to say the least, so the discussion about "true exposure" and the difference between backyard exposure and indoor exposure seems to be missing the elephant in the room.

[1] Bias is a systematic over- or under-estimation, like a scale that always reads 3lbs lighter than you really are. Error means the difference between the measured and true value. So a scale might be unbiased but sometimes read 2lbs more than the real weight and sometimes 2lbs less than the real weight, so the error is plus or minus 2 lbs.

Agreed. We've revised the paragraph as follows:

"The opposite relation may exist with our estimates of cancer effects, but the exposure bias is dwarfed by other uncertainties. The health impacts from toxic air pollutants are estimated from occupational studies or studies of lab animals. In occupational studies, exposure estimates are very rough. In studies of lab animals, the exposure may be well-controlled, but the low-dose extrapolation and extrapolation from other species to human introduces large uncertainties.

"In addition to these large uncertainties, there is likely to be a modest overestimate of exposure: We believe our models do a reasonable job of estimating backyard exposure…"

Page 26, the discussion of the C-R function for toxics might elaborate a bit on the function used. I have used exactly the function you describe and have gotten in discussions in the past with public health scientists who strongly disliked my assumption that all that matters is life-years of exposure. That is, the risk of one person exposed to 70 years at a constant concentration is the same as 70 people exposed for 1 year at this same level. The assumptions underlying the unit risk model are very restrictive, and whenever we use unit risk estimates in real world situations to calculate deaths, we violate these assumptions.

We did not understand this comment. Can you elaborate?

On page 63, it is not clear why sulfate has such a large effect. Direct SO4 is, as far as I know, a small percentage of SOx emissions, perhaps in the range of 3%. Of course, not all SO2 goes to SO4, still it is surprising to me that a 20% reduction in SOx gas has less effect than a 20% reduction in direct SO4.

In our modeling, emissions of SO2 and other sulfur gases were 53 tons/day and emissions of direct sulfate were 3 tons per day, or about 6% of SOx emissions. The Bay Area is very different from the East in this respect, however. In the East, the principal source ammonium sulfate is SO2 from coal-fired power plants. In the Bay Area, SOx emissions are mainly from ships and refineries.

From what we have seen, direct sulfate emissions are *not* a large percentage of ship emissions. But one study suggests that primary sulfate may indeed be a large source for ammonium sulfate in areas with a lot of ship traffic from the National Academy of Sciences:
www.pnas.org/cgi/doi/10.1073/pnas.0805255105
This article posits that the "long atmospheric residence time" of primary sulfate emissions may account for this. Perhaps this is being captured in our model.

Page 66: "The chemistry has similarities with how ozone is formed, and some of the same 'irrational' behavior can occur." The article by Ansari and Pandis (1998) talks of non-linearities in ammonium nitrate formation. Rather than say "irrational" I would suggest a more precise term. Irrational is an unfortunate choice as it makes me wonder a bit about whether the air quality model results are believable to begin with. Having a little more discussion about the formation of ammonium nitrate would be helpful. I am not aware of why VOC is relevant. Regarding the unimportance of a 20 percent reduction in sulfate, it could be that sulfate levels are so low to begin with that a 20 percent reduction is negligible; the model results should be able to show this.

We changed the term “irrational” to “counter-intuitive.”

You might include some maps of the modeling results.

That is a good suggestions. We've added a couple of figures.

Ansari, A. and S. Pandis (1998). "Response of inorganic PM to precursor concentrations." Environ. Sci. Technol 32(18): 2706-2714.

****

Comments submitted by Dr. Robert Harley, Professor and Group Leader for Environmental Engineering
Department of Civil and Environmental Engineering
University of California at Berkeley

1. I am very supportive of the idea that BAAQMD will be considering multiple pollutants together in a unified planning & analytical framework, to me that is a major advance over the old way of doing things. It allows for weighing of trade-offs in situations where for example there is an ozone disbenefit from a control measure that also reduces toxic emissions from diesel exhaust.

2. I note NOx and black carbon (BC) are relevant not only to local air quality but also to global change (NOx is a precursor to tropospheric ozone in the background atmosphere; ozone itself is a GHG). BC control is probably rather expensive in the U.S. as a GHG reduction strategy, but still desired for local air quality improvement in any case. My main point here is just to note synergies between local/regional air quality and global climate change for these pollutants.

We agree. We're not sure where we could add this point in the Methodology write-up.

We added a new section in the Methodology, titled " Potential future enhancements to multi-pollutant evaluation method". We've incorporated some of your comments there.

3. The health effects discussion seems out of line with my understanding of what the relevant effects of concern are. I understand the main concern about human exposure to fine particles to be heart attack/stroke (i.e., cardiovascular effects), you emphasize lung cancer on page 8. Current standards for PM2.5 count all mass equally, regardless of chemical composition. I suggest reviewing current evidence to see if water-soluble neutral aerosols such as ammonium nitrate and ammonium sulfate are really harmful and worth all the effort you are proposing to expend on modeling them. Also you know that wood smoke is a major contributor to wintertime PM in the Bay area, but it's not flagged in the same way that diesel exhaust has been (i.e., as a toxic air contaminant). Are you really confident that while diesel exhaust PM is harmful, wood smoke is benign?

Regarding treating all $PM_{2.5}$ equally, this is really the state of the art at this time. Dr. Bart Ostro at OEHHA has studied the components if $PM_{2.5}$ looking for differences in health impacts. See, e.g., "The Effects of Components of Fine Particulate Air Pollution on Mortality in California: Results from CALFINE", Ostro et al., Environmental Health Perspectives 115, January 2009.

His analyses show correlations between ammonium nitrate and mortality. This is a result we also found when I looked at mortality vs. air pollution in Santa Clara County. [Table 1b, pg 101 in "Revised analyses of time-series studies of air pollution and health Part II," Health Effects Institute, May 2003 finds that when both $PM_{2.5}$ and nitrate are in a regression with non-accidental mortality as the response, nitrate continues to be statistically significant.]

We added the following (including footnote): "The $PM_{2.5}$ components considered in this study constitute over 90% of the anthropogenic $PM_{2.5}$ in the Bay Area. In line with other health benefit studies, we assume that the impact of the various $PM_{2.5}$ components on health is the same – depending only on mass, not composition or size provided the size is < 2.5 microns.[2]"

We are very confident that wood smoke is harmful too. But, unlike diesel, there is no established risk function. This is definitely something we will include once a risk function is established.

4. Your universe of toxic air contaminants is a small one. With RFG and LEV program, benzene concerns have been mostly addressed now except for cigarette smokers. As other examples, what about highly reactive unsaturated aldehydes such as acrolein and crotonaldehyde? What about ultrafine particles (particle number, PN, as opposed to particle mass, PM). To my knowledge these are not measured in the district's current air sampling programs, but they are examples of other pollutants that may be toxic contaminants and worthy of your attention.

We added the following: "For every pollutant included in the above list there are literally dozens that are not included. The main reasons for limiting the pollutants to those in Table 1 is pragmatic – to include dozens more would require a major effort both in developing the methodology and in rule development.

"We note some key omissions. We only considered a limited number of carcinogenic toxics. There are other carcinogens and also toxics that have other serious health effects, e.g., acrolein, lead, mercury. In addition, there are risks that undoubtedly exist but have not been quantified. An example is the carcinogenicity of wood smoke, which is very similar chemically to tobacco smoke, a known carcinogen. We will continue to monitor the health effects literature and, we hope, update the methodology with some of these omitted pollutants."

*5. Ozone health effects seems focused on acute health effects and decreased lung function when trying to exercise the same day. You talk about a threshold below which you stop caring about this issue. But there are also concerns about chronic effects of ozone exposure, specifically mortality due to respiratory causes. Imagine lifetime exposure to ozone leads to more rapid decline in lung function over lifespan, then when older person gets challenged with flu, pneumonia or other respiratory ailments they are less able to fight back and recover, so they die prematurely in some sense. Impaired lung function due to chronic ozone exposure could be an important concern even below the threshold of 50 ppb you specify. See Jerrett et al. (New England J Medicine 2009), he is one of the first to my knowledge to publish epi study results separating the mortality effects of *chronic* exposure to ozone and PM2.5.*

[2] The question of what components, sizes and aspects of PM are more harmful is an area of current research. There is mounting evidence that ultra-fine particles are more harmful than larger particles, but the results are not yet definitive.

The suggestion to add respiratory mortality from chronic ozone exposure is a good one. We will consider this as a potential future enhancement to the methodology.

Regarding thresholds, Jerrett et al. actually find some evidence for a threshold. Their best-fitting model, with a threshold of 56 ppb, had a p-value of 0.06 relative to a non-threshold model. For the Bay Area, another issue with assuming a no-threshold model is that not only is the ozone background non-zero, but often ozone values in the Bay Area are *below* the threshold. The impact of our controls over the years has, in general, been to *increase* those ozone values. Thus, if we assume a no-threshold model, we may have to conclude that our efforts have been counterproductive. Absent clear-cut evidence to the contrary, we make the assumption that there is a threshold.

6. Some control measures are targeted at very specific point sources/locations. On page 3 you say emission reductions will be distributed throughout the bay area according to baseline distribution of emissions from all sources in the current inventory. This could be significantly in error and understate local benefits of potential measures focused on addressing environmental justice issues near Port of Oakland, Oakland & SFO airports, refineries, etc. For example, ARB has adopted various rules that focus on ports specifically.

Re: comment: "This could be significantly in error and understate local benefits of potential measures…" We are aware of this shortcoming, but it is unclear how to do better short of running the models for every control measure, which is impractical due to resource constraints.