Preliminary Comments on the REA Planning Document from Dr. James Boylan

Ambient Air Concentrations

1. The use of an AERMOD model-based approach to predict hourly concentrations at all receptor locations within selected study areas. [Sections 3.3.2, 4.1.3.3]

The model-based approach to predict hourly concentrations at all receptor locations within the selected study areas is appropriate and will better quantify the spatial variation in concentrations compared to using observations alone. AERMOD is an appropriate model for predicting SO2 concentrations in ambient air. The approach that is describes includes running AERMOD to obtain 1-hour SO2 concentrations at all receptors and all hours, then using the 5-minute SO2 observations to convert the 1-hour AERMOD results into continuous 5-minute results. If AERMOD is performing well, this is a valid approach.

Page 4-18 of the REA states that "Model performance (e.g., comparison with available monitor data) can be evaluated using procedures outlined in the EPA Protocol for Determining Best Performing Model (U.S. EPA, 1992)." A summary of the specific model performance approach that will be implemented and the model performance "acceptance" criteria needs to be included. The ISA states, "For models intended for application to compliance assessments (e.g., related to the 1-h daily max SO2 standard), the model's ability to capture the high end of the concentration distribution is important. Measures such as robust highest concentration (RHC) (Cox and Tikvart, 1990), and exploratory examinations of quantile-quantile plots (Chambers et al., 1983) are useful. The RHC represents a smoothed estimate of the top values in the distribution of hourly concentrations. In contrast, for dispersion modeling in support of health studies where the model must capture concentrations at specified locations and time periods, additional measures of bias and scatter are important."

Most published AERMOD model performance evaluations are associated with using the model for compliance assessments. In the REA, the model is being used to support health studies where spatial and temporal accuracy is much more important compared with compliance assessments. The model results need to be evaluated against observations paired in time and space. The REA needs to discuss "acceptable" model performance criteria and options for correcting the model results if there are significant biases (over predictions or under predictions) in the modeling results. For example, would it be acceptable to over predict the SO2 concentrations by a factor of 2 or under predict the SO2 concentrations by 50%? If not, what will be done with these modeling results?

Also, the overall model performance may look good for the entire year due to compensating errors, but the daily temporal profiles and/or seasonal profiles might be way off. One option that could be used in the REA would involve scaling the model results up/down to match the SO2

observations. This would keep the relative spatial distributions identified by the model in place, but would adjust the concentration levels to match the observations. This would minimize impacts from poor model performance on the modeled 1-hour and 5-minute SO2 concentrations.

2. The use of SO2 measurements at ambient air monitors within and near the study areas to estimate continuous 5-minute concentrations, where appropriate (e.g., filling missing values, for AERMOD hourly predictions). [Sections 3.3.1, 4.1.3.1, 4.1.3.2]

The comprehensive 2011-2015 1-hour average and 5-minute SO2 data sets will help estimate continuous 5-minute concentrations. On page 4-13, proposed step #3 states "Substitute any hours not having measurements for any year with a value of zero (0)." However, it might be more appropriate to substitute missing values with the lower detection limit for the SO2 monitor (1 ppb) or interpolate between the values before and after the missing value(s). Equation 4-1 and Equation 4-3 seem appropriate. However, Equation 4-2 should be updated since the cases where the 5-min max is $\geq$2 x 1-hour average produces 5-min values for $C_6 - C_1$ that are lower than the cases where the 5-min max is <2 x 1-hour average. Conceptually, a linear or exponential ramp is appropriate for cases where the 5-min max is $\geq$2 x 1-hour average. However, a minimum value should be assigned. Options for setting an appropriate minimum value could include using 50% of the current 1-hour average SO2 concentrations or using the 1-hour average concentration for the next hour (assuming it is significantly lower than the current hour).

3. The proportional approach selected for adjusting ambient concentrations to simulate air quality that just meets the existing standard. [Section 4.1.3.4]

In general, the proportional approach selected for adjusting ambient concentrations to simulate air quality that just meets the existing standard seems to be appropriate. On page 4-20, it is stated "For the planned REA for the current review, in each study area, F will be calculated by dividing 75 ppb by the DV and will be used to adjust all SO2 concentrations in a study area by this factor to simulate just meeting the existing standard." In order to just meet the standard, it would seem appropriate to only adjust the SO2 concentrations for those receptors with a design value above 75 ppb and leave the receptors with a design value already below 75 ppb unchanged.

Study Areas and Time Periods (Section 4.1.2.1)

It seems appropriate to model three consecutive years to evaluate variability in exposures across a 3-year period since this is consistent with the form of the existing standard. Also, modeling domains with receptors within a 10-km radius of all nearby SO2 emission sources greater than 100 tons/year is appropriate.

Four of the nine candidate study areas listed in Table 4-1 were selected for detailed analysis. It would be helpful to state the specific reasons why each of the other five areas were not selected. Finally, the maps in Figures 4-2, 4-3, 4-4, and 4-5 are difficult to read and interpret. I suggest

showing DVs 0-75 (blue), 76-100 (yellow), and >100 (red). For the large SO2 emission sources, it would be better to represent the emission sources with different size bubbles that are representative of the size of the SO2 emissions (small bubble for low emissions, large bubble for high emissions). Also, I would suggest break points of 100-500, 500-1000, 1000-2000, 2000-5000, >5000 tons.

Also, it is not clear how the four study areas will be used to calculate the number and percent of the population across the country experiencing 5-minute SO2 exposures at or above benchmark levels of concern, the number of occurrences of lung function decrements in the at-risk populations across the country, and the number and percent of the at-risk populations across the country estimated to experience single or multiple occurrences of those lung function decrements.

Other Potential Standard Levels

With so many improvements being implemented to better characterize the 5-minute SO2 concentrations in ambient air (new 2011-2015 SO2 measurements and enhancements to AERMOD) and the improvements to the exposure assessment (APEX and CHAD), it would seem appropriate to re-evaluate other potential standard levels (50 ppb and 100 ppb) besides the current level (75 ppb).