UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NC 27711

JUN 30 2015

OFFICE OF
AIR QUALITY PLANNING
AND STANDARDS

MEMORANDUM

SUBJECT: Supplemental Information for EPA's 2009 Draft Report regarding Reassessment of IWAQM Phase 2 Recommendations

FROM: Tyler J. Fox, Group Leader
Air Quality Modeling Group, C439-01
Office of Air Quality Planning and Standards

TO: Proposed Regulatory Docket No. EPA-HQ-OAR-2015-0310

This memorandum provides supplemental information to formally complete the EPA's 2009 draft report, "Reassessment of the Interagency Workgroup on Air Quality Modeling (IWAQM) Phase 2 Summary Report: Revisions to the Phase 2 Recommendations", as it is considered in support of EPA's proposed revisions to the *Guideline on Air Quality Models*, published as Appendix W to 40 CFR part 51. In 2005, EPA formed a CALPUFF workgroup to help identify issues with the existing guidance concerning the application of the CALPUFF model for use in regulatory long-range transport (LRT) modeling contained in the 1998 report, "Interagency Workgroup on Air Quality Modeling (IWAQM) Phase 2 Summary Report and Recommendations for Modeling Long Range Transport Impacts". This 1998 report provided a series of recommendations that supported the revisions in 2003 to the *Guideline* that established CALPUFF as EPA's preferred model for long-range transport (*i.e.*, source-receptor distances of 50 to several hundred kilometers) modeling applications for purposes of demonstrating compliance with Class I PSD increments.

Specifically, this memorandum provides the following sections (as attached) that were not completed at the time of issuing the 2009 draft report:

- Section 4.0: Evaluation Studies and Findings
- Appendix A: CALMET Recommendations
- Appendix B: Summary Comparisons of CALPUFF Modeling System Updates

The information provided in Section 4.0 that summarizes LRT performance evaluations is taken from the 2012 EPA report, "Documentation of the Evaluation of CALPUFF and Other Long Range Transport Models using Tracer Field Experiment Data" that was presented as part of the 10th Modeling Conference.

Recycled/Recyclable • Printed with Vegetable Oil Based Inks on 100% Recycled Paper (40% Postconsumer)

4.0 EVALUATION STUDIES AND FINDINGS

The U.S. EPA is exploring different alternatives for performing single-source dispersion modeling over longer distances to address Class I and Class II area air quality and AQRV issues. Such issues include PSD pollutant concentrations, including SO_2, NO_2 and $PM_{2.5}$ concentrations, visibility and sulfur and nitrogen deposition. Ozone is also becoming a pollutant of increasing importance. Important components of visibility and deposition are sulfate (SO_4) and nitrate (NO_3) that are secondarily formed PM species from gaseous SO_2 and NO_x emissions, respectively. Thus, the correct depiction of chemistry is an important feature of LRT dispersion models.

Although CALPUFF became the EPA-recommended LRT dispersion model in 2003 for distances beyond 50 km and chemically inert pollutants, it has several limitations and issues:

- The chemical conversion algorithm in the regulatory versions of CALPUFF (Version 5.8) is almost three decades old (developed in 1983) and has been shown to be inconsistent with our current knowledge on secondary $PM_{2.5}$ formation chemistry (Morris et al., 2003; 2005; 2006).
- The CALPUFF chemistry algorithm was recently updated in Version 6.4 (Karamchandani, Chen and Seigneur, 2008), but still does not contain photochemical reactions, which are important to simulate secondary PM formation.
    - Sulfate and nitrate formation is formed through a complex set of photochemical reactions that require the correct depiction of the radical cycle, including hydroxyl (OH) and perhydroxy (HO_2) radicals that are driven by organic and inorganic species.
- CALPUFF does not estimate ozone formation from single emissions sources.
- CALPUFF sensitivity modeling using alternative CALMET meteorological inputs found wide variations in the CALPUFF model estimates.

EPA evaluated CALPUFF and five other LRT dispersion models using data from four atmospheric tracer field experiments including:

1) 1980 Great Plains Field Experiment (GP80),
2) 1975 Savannah River Laboratory Field Experiment (SRL75),
3) Cross Appalachian Tracer Experiment (CAPTEX), and
4) European Tracer Experiment (ETEX).

The LRT dispersion modeling was performed by EPA from 2008 to 2010 and builds off several previous LRT dispersion modeling studies that evaluated models using tracer study field experiments. The EPA tracer test comparison evaluated the transport and dispersion components of the LRT models and raised additional questions regarding the CALPUFF LRT dispersion model.

The following are some of the key conclusions of the LRT dispersion model tracer test field experiment evaluation.

CALPUFF/CALMET Concentration Predictions are Highly Variable: Use of alternative CALMET input options within their range of reasonableness can produce wide variations in the CALPUFF concentration predictions. Given the regulatory use of CALPUFF, this result points toward the need to have a standard set of recommended CALMET settings for regulatory application of CALPUFF to assure consistency and eliminate the potential of selecting CALMET options to obtain a desired outcome in CALPUFF. No one CALMET configuration consistently produced the best CALPUFF model performance, although use of MM5 data with CALMET did tend to improve CALPUFF model performance with 36 and 12 km MM5 data being better than 80 km MM5 data.

Comparison of Current CALPUFF Model Performance with Previous Studies: The comparison of the model performance for current version of CALPUFF with past CALPUFF evaluations from the 1998 EPA study (EPA, 1998a) using the GP80 and SRL75 tracer study field experiments was mixed. For the GP80 100 km receptor arc, the current and past CALPUFF model performance evaluations were consistent with CALPUFF tending to overestimate the plume maximum concentrations and underestimate plume horizontal dispersion. The current version of CALPUFF had difficulty in reproducing the good performance of the past CALPUFF application in estimating the tracer residence time on the GP80 600 km receptor arc. Only by invoking the CALPUFF slug option, as used in the 1998 EPA study, was CALPUFF/CALMET able to reproduce the tracer residence time on the 600 km receptor arc. As the slug option is for near-source modeling and is a very non-standard option for LRT dispersion modeling, this result questions the validity of the 1998 CALPUFF evaluation study as applied for CALPUFF LRT modeling. The CALPUFF/MMIF was less sensitive to the slug option and more sensitive to puff splitting than CALPUFF/CALMET. For consistency, the current and EPA 1998 study CALPUFF evaluation approach both used the fitted Gaussian plume model evaluation methodology, along with angular plume centerline offset and tracer receptor arc timing statistics. The fitted Gaussian plume evaluation approach assumes that the observed and predicted concentration along a receptor arc has a Gaussian distribution. At longer downwind distances such an assumption may not be valid. For the CALPUFF evaluation using the SRL75 tracer field experiment, there was a very poor fit of the Gaussian plume to the observations resulting in some model performance statics that could be misleading. We do not recommend using the fitted Gaussian plume evaluation approach in future studies and instead recommend using approaches like the ATMES-II statistical evaluation approach that is free from any a priori assumption regarding the observed tracer distributions.

EPA-FLM Recommended CALMET Settings from the 2009 Clarification Memorandum: The EPA-FLM recommended CALMET settings in the 2009 Clarification Memorandum produces wind field estimates closest to surface wind observations based on the CAPTEX CALMET modeling. However, when used as input into CALPUFF, the EPA-FLM recommended CALMET settings produced one of the poorer performing CALPUFF/CALMET configurations when comparing CALPUFF predictions against the observed atmospheric tracer concentrations. Given that the

CALMET wind evaluation is not an independent evaluation because some of the wind observations used in the evaluation database are also input into CALMET, the CALPUFF tracer evaluation bears more weight. Other aspects of the EPA-FLM recommended settings generally produced better CALPUFF tracer model performance including use of prognostic meteorological data as input to CALPUFF. The CALPUFF evaluation also found better CALPUFF performance when 12 km grid resolution is used in MM5 or CALMET as opposed to 80 or 36 km.

CALPUFF Model Performance using CALMET versus MMIF: The CALPUFF tracer model performance using meteorological inputs based on the MMIF tool versus CALMET was mixed. The variations of the CALPUFF model predictions using MMIF were much less than when CALMET was used and the CALPUFF/MMIF model performance was usually within the range of the performance exhibited by CALPUFF/CALMET. Specific examples from the tracer tests are as follows:

- For the GP80 100 km receptor arc, the CALPUFF/MMIF exhibited better fitted plume observed tracer model performance statistics than all of the CALPUFF/CALMET configurations except when CALMET was run using MM5 and surface meteorological observations but no upper-air meteorological observations.
- CALPUFF/CALMET using no MM5 data and just meteorological observations exhibited the best plume centerline location on the GP80 100 km receptor arc with CALPUFF/CALMET using just MM5 data and no observations and CALMET/MMIF exhibiting the worst plume centerline location.
- For the GP80 600 km receptor arc, the CALPUFF/MMIF fitted plume model performance statistics are in the middle of the performance statistics for the CALPUFF/CALMET configurations.
- The slug option was needed for CALPUFF/CALMET to produce good 600 km receptor arc tracer residence time statistics but had little effect on CALPUFF/MMIF. However, use of puff splitting greatly improved the CALPUFF/MMIF tracer residence time statistics.
- Of all the CALPUFF sensitivity tests examined, CALPUFF/MMIF using the slug option and puff splitting produced the best CALPUFF fitted plume tracer model performance statistics for the GP80 600 km receptor arc.
- In an opposite fashion to the GP80 100 km receptor arc, for the SRL75 100 km receptor arc the best plume centerline offset was achieved when CALPUFF was run with just MM5 data and no meteorological observations (either with CALMET or MMIF) with performance degraded when meteorological observations are used with CALMET.
- The CALPUFF model performance using the MMIF tool and 36 and 12 km MM5 data performed better than all of the CALPUFF/CALMET sensitivity tests for the CAPTEX CTEX3 experiment. However, the CALPUFF/MMIF using 36 and 12 km MM5 data performed worse than all of the CALPUFF/CALMET sensitivity tests for the CAPTEX CTEX5 experiment.

Comparison of Model Performance of LRT Dispersion Models: Six LRT dispersion modeled were evaluated using the CAPTEX Release 3 and 5 tracer database and five LRT dispersion models

were evaluated using the ETEX tracer test field experiment. In each case the same MM5 meteorological data were used as input into all of the dispersion models, although different MM5 configuration options were selected for each tracer experiment.

The CAMx and CALGRID Eulerian photochemical grid models, FLEXPART Lagrangian particle model, HYSPLIT Lagrangian particle, puff and particle/puff hybrid model and CALPUFF and SCIPUFF Gaussian puff models were evaluated. For all three tracer experiments (CTEX3, CTEX5 and ETEX), the CAMx model consistently ranked highest when looking across all of the model performance statistics or when using the RANK composite performance statistic. For the CTEX3 field experiment, the RANK composite performance statistic gave consistent rankings of model performance with the suite of statistical metrics with CAMx being the highest RANK score (1.91) followed by SCICHEM (1.71).

The rankings of the models using all of the statistics versus the RANK composite statistic were inconsistent for the CTEX5 experiment. Both approaches showed CAMx and HYSPLIT were the highest ranking LRT dispersion model for the CTEX5 field experiment. However, the RANK statistic ranked CALGRID as the 3rd best performing model, whereas when looking at all the performance statistics it was the worst performing model because it exhibited a large spread underestimation bias, had no correlation with the observations and little skill in reproducing the spatial distribution of the observed tracer. The CTEX5 LRT model evaluation points out the need to examine all performance statistics and not rely solely on the RANK composite statistic. It also points out the need to define a RANK-type composite statistic that focuses on the regulatory application of LRT dispersion models where an underestimation bias is undesirable. Of the three top performing LRT dispersion models, CAMx had the highest RANK composite statistic and scored the highest for most (64%) of the other ATMES-II statistical model performance metrics, with HYSPLIT scoring the highest for 27% of the metrics. Additional findings of the ETEX tracer test evaluation are as follows:

- The model performance rankings were preserved closer to the source (e.g., within 300 km) as well as further downwind.
- CALPUFF puff splitting sensitivity tests had little effect on CALPUFF model performance.
- CAMx vertical mixing and horizontal advection solver sensitivity tests found that use of the
- MM5CAMx CMAQ-like vertical mixing diffusion coefficients and the PPM advection solver produced the best tracer test model performance. Similar results were seen in the CTEX3 and CTEX5 sensitivity modeling.
- HYSPLIT sensitivity tests using solely particle, solely puff and hybrid particle/puff and puff/particle combinations found that the hybrid configurations performed best and the puff configuration performed worst, with the CTEX3 and CTEX5 sensitivity test producing similar results.

APPENDIX A

CALMET RECOMMENDATIONS

Section 2 of this report provided recommended CALMET settings with an objective to try and "pass through" the WRF/MM5 meteorological model output as much as possible for input into CALPUFF. However, further testing of CALMET and CALPUFF by EPA's CALPUFF workgroup found that these recommended CALMET settings did not achieve the intended result to "pass through" the WRF/MM5 meteorological variables as CALMET still re-diagnosed some and modified other meteorological variables thereby degrading the WRF/MM5 meteorological fields. Based in part on CALMET evaluations using tracer test field study databases, EPA determined interim CALMET settings that produced the best meteorological model performance and on August 31, 2009 released a Clarification Memorandum "Clarification on EPA-FLM Recommended Settings for CALMET" with new recommended settings for CALMET (Available at: http://www.epa.gov/ttn/scram/guidance/clarification/CALMET%20CLARIFICATION.pdf). In this Clarification Memorandum, EPA reiterated the desire to "pass through" meteorology from the WRF/MM5 prognostic meteorological models to CALPUFF, but the CALMET model at this time was incapable of achieving that objective.

APPENDIX B

SUMMARY COMPARISON OF CALPUFF MODELING SYSTEM UPDATES

As with any modeling system, periodic updates are anticipated as part of the standard software life cycle to address bugs that are identified, as well as enhancements that may be needed to address new data formats or other needs that may arise. To address the need for a systematic process to assess impacts of modifications to the CALPUFF modeling system, EPA established a standard "Protocol for Updating the CALPUFF Modeling System" and developed a "CALPUFF Assessment Tool" to support that process. Such a process is vital to preserving the integrity of the preferred status of models recommended by EPA in the *Guideline on Air Quality Models* (40 CFR Part 51, Appendix W).

The current EPA regulatory version of the CALPUFF Modeling System includes:

- CALPUFF version 5.8.4, level 130731
- CALMET version 5.8.4, level 130731
- CALPOST version 6.221, level 080724

For every update of the "EPA-Approved" version of the CALPUFF Modeling System, a consequence analysis is performed using an update protocol that identifies what model changes have been made and their implications based on the analysis results. For this purpose, EPA developed the CALPUFF Assessment Tool that consists of 11 scenarios designed to test the modeling system across a range of possible applications in terms of modeling domain, meteorological data options, and source types. This tool prepares summaries of differences in predicted concentrations between two versions of the CALPUFF Modeling System, the "Base" version referring to the current EPA-approved version, and the "Beta" version referring to the updated version of the modeling system that is the subject of the assessment. The CALPUFF Assessment Tool has been successfully applied to support EPA's adoption of updates to the CALPUFF modeling system since its promulgation in 2003 with summary reports and comparison results provided for each update on EPA's SCRAM website at: http://www.epa.gov/ttn/scram/dispersion_prefrec.htm#calpuff.