



EPA'S PARTICULATE MATTER (PM) AUGMENTATION TOOL: AUTOMATING QUALITY ASSURANCE AND PM SPECIATION TO GENERATE A MODEL-READY INVENTORY



Jonathan G. Dorn* and David M. Cooley
Abt Associates, Inc., Durham, NC, USA

Roy Huntley
U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

1. INTRODUCTION

The Consolidated Emissions Reporting Rule requires that state, local, and tribal agencies (SLT) submit particulate matter (PM) emissions for point and nonpoint sources to EPA's National Emissions Inventory (NEI). Since PM is both a National Ambient Air Quality Standard (NAAQS) pollutant and a major contributor to visibility impairment, the PM_{2.5} NAAQS and the Regional Haze Rule each emphasize emissions inventory development for the PM species required in regional air quality modeling. While submission of PM emissions to the NEI by SLT agencies should include filterable and primary PM (PM₁₀-PRI, PM₁₀-FIL, PM₂₅-PRI, and PM₂₅-FIL) along with condensible PM (PM-CON), SLT agencies often submit incomplete speciation. Augmentation of the PM species in the NEI point and nonpoint source inventories is necessary to ensure completeness of the PM inventories for air quality modeling and to ensure that SLT inventories do not contain erroneous pollutant reporting. This paper describes EPA's PM Augmentation Tool used to correct inconsistencies in SLT-submitted PM data and to gap-fill missing PM data in version 1 of the 2011 NEI. One recent modification to the PM Augmentation Tool is the expansion of the augmentation procedure to cover nonpoint source categories. The tool involves three main steps: 1) pre-screening and self-consistency checks; 2) trivial updates, where missing data can be calculated using addition or subtraction of filterable or condensible PM; and 3) non-trivial updates, where missing data must be estimated using PM calculators. Each of these steps is described in more detail below.

*Corresponding author: Dr. Jonathan Dorn, Abt Associates, Inc., Central Park West, 5001 S. Miami Blvd, Suite 210, Durham, NC 27703; e-mail: jonathan_dorn@abtassoc.com; phone: 919-294-7763

2. PRE-SCREENING STEPS AND SELF-CONSISTENCY CHECKS

In the first step of the PM Augmentation tool, SLT submitted PM data are reviewed and inconsistencies are removed where applicable. There are five PM pollutants that can potentially exist and be submitted for every source in the point and nonpoint source inventories, including:

PM₁₀-FIL – filterable particulate matter less than or equal to 10 microns in aerodynamic diameter.

PM₂₅-FIL – filterable particulate matter less than or equal to 2.5 microns in aerodynamic diameter.

PM₁₀-PRI – primary particulate matter less than or equal to 10 microns in aerodynamic diameter (includes condensible emissions).

PM₂₅-PRI – primary particulate matter less than or equal to 2.5 microns in aerodynamic diameter (includes condensible emissions).

PM-CON – condensible particulate matter, that matter which exists as a vapor at stack conditions but exists as a liquid or a solid after exiting the stack and being cooled by ambient conditions. For some sources this value will equal zero.

The pre-screening checks in the tool are based on the physical and mathematical relationship that exists between these five PM pollutant species. For example, PM₁₀-PRI is the sum of PM₁₀-FIL and PM-CON. Also, PM₂₅-PRI is the sum of PM₂₅-FIL and PM-CON. Other relationships are that the value for PM₁₀-FIL must be greater than or equal to that of PM₂₅-FIL and the value for PM₁₀-PRI must be greater than or equal to that of PM₂₅-PRI. When PM data submitted to EPA by SLT agencies do not conform to these relationships, then the data need to be corrected before PM augmentation is performed to prevent error propagation.

The PM calculator was originally a standalone computer program developed by EPA (but no longer supported) that would calculate PM₁₀-FIL

and/or PM25-FIL data from PM-FIL, and vice versa. Necessary inputs to the programs are process descriptions, fuel information, and control devices, indicated by EPA's source classification code (SCC) and control device codes. The PM Calculator used size fractionation data from Appendices B.1 and B.2 of AP-42 to calculate the various PM sizes from PM total data for a large number of processes. Since the PM Calculator computer program is no longer supported by EPA, the size fractionation data were converted to a set of MS Access® tables, and these tables are used in the methodology described in this paper. Due to concerns regarding file size, the Access® tables were split into three separate databases:

1. pm_calculator_for_all_sccs_except_if_start_with_3.accdb
2. pm_calculator_for_sccs_starting_with_3_to_30399999.accdb
3. pm_calculator_point_sccs_start_with_3_and_greater_than_30399999.accdb

An updated interface was added to the tool to make it more user-friendly. The interface allows for the tool to be run with the click of a single button, rather than requiring the user to run several queries and macros in each of the individual databases.

The PM Calculator replacement tables in the tool are used to calculate controlled PM10-FIL and controlled or uncontrolled PM2.5-FIL using uncontrolled PM10-FIL. In addition, condensible emissions are estimated by applying conversion factors developed by Strait et al. (1999) to SLT reported PM species or PM species derived from the replacement tables or by applying emission factor ratios.

Before performing any calculations to create new PM data, the PM data submitted to EPA by SLT agencies are reviewed and corrected for mathematical inconsistencies, as noted above. The following set of checks and fixes are performed on all SLT submitted PM data to ensure self-consistency. These checks are repeated after the augmentation process to ensure that no inconsistencies were introduced by the augmentation process:

1. If PM10-FIL>PM10-PRI when both are reported by SLT, PM10-PRI is replaced with Null and the remaining augmentation steps are performed as if it had not been reported. For the PM10-PRI record, the EmissionsComments field is populated with this comment. "___-reported

emissions for PM10-PRI are recalculated using PM-Calculator data because ___ reported PM10-FIL>PM10-PRI." (Note: In the EmissionsComments field examples, ___ is the program system code for the SLT data, which identifies the reporting agency.)

2. If PM25-FIL>PM25-PRI when both are reported by SLT, PM25-PRI is replaced with Null and the remaining augmentation steps are performed as if it had not been reported. For the PM25-PRI record, the EmissionsComments field is populated with this comment. "___-reported emissions for PM25-PRI were recalculated using PM-Calculator data because ___ reported PM25-FIL>PM25-PRI."
3. If PM10-FIL+PM-CON>PM10-PRI, PM10-PRI is replaced with PM10-FIL+PM-CON to eliminate inconsistencies. For the PM10-PRI record, the EmissionsComments field is populated with this comment. "___-reported emissions for PM10-PRI were replaced with ___-reported PM10-FIL+PM-CON because ___ reported PM10-FIL+PM-CON>PM10-PRI."
4. If PM25-FIL+PM-CON>PM25-PRI, PM25-PRI is replaced with PM25-FIL+PM-CON to eliminate inconsistencies and a check is performed to ensure that the new PM25-PRI value is less than PM10-PRI. For the PM25-PRI record, the EmissionsComments field is populated with this comment. "___-reported emissions for PM25-PRI were replaced with ___-reported PM25-FIL+PM-CON because ___ reported PM25-FIL+PM-CON>PM25-PRI."
5. If PM25-PRI>PM10-PRI when both were reported by SLT, PM10-PRI is replaced with PM25-PRI. For the PM10-PRI record, the EmissionsComments field is populated with this comment. "___-reported emissions for PM10-PRI were replaced with ___-reported PM25-PRI because ___ reported PM25-PRI>PM10-PRI."
6. If PM25-FIL>PM10-FIL when both were reported by SLT, PM10-FIL is replaced with PM25-FIL. For the PM10-PRI record the EmissionsComments field is populated with "___-reported emissions for PM10-FIL were replaced with ___-reported PM25-FIL because ___ reported PM25-FIL>PM10-FIL."

7. If PM-CON > PM25-PRI or PM-CON > PM10-PRI when both are reported, two cases are considered:
 - a. The SLT reports PM-CON that is much higher than PM10-PRI, which is inconsistent with the definition. For this case, (PM-CON > 110% PM10-PRI) PM-CON is replaced with Null and the remaining augmentation steps are performed as if PM-CON had not been reported. For the PM-CON record, the EmissionsComments field is populated with this comment: “___-reported PM-CON emissions were recalculated using PM Calculator data because ___ reported PM-CON > PM10-PRI/PM25-PRI.”
 - b. The SLT was trying to say that all the PMxx-PRI is PM-CON but PM-CON is rounded to a different accuracy and ends up slightly higher than PMxx-PRI. For this case, (110% PMxx-PRI > PM-CON > PMxx-PRI) PM-CON is replaced with PMxx-PRI, effectively setting PMxx-FIL to zero. For the PM-CON record, the EmissionsComments field is populated with “___-reported PM-CON emissions are replaced with PM10-PRI/PM25-PRI because ___ reported PM-CON > PM10-PRI/PM25-PRI.”

3. TRIVIAL UPDATES

After the pre-screening checks and fixes, trivial updates are made. These include updates where missing data can be calculated based on the addition or subtraction of condensible or filterable PM values. The trivial updates performed by the tool include:

1. If both PMxx-FIL and PM-CON are available, but not PMxx-PRI, then $PMxx-PRI = PMxx-FIL + PM-CON$.
2. If both PMxx-PRI and PM-CON are available, but not PMxx-FIL, then $PMxx-FIL = PMxx-PRI - PM-CON$ if the difference is positive, or zero if the difference is negative.
3. If both PMxx-PRI and PMxx-FIL of either diameter are available, but not PM-CON,

then $PM-CON = PMxx-PRI - PMxx-FIL$, with preference to PM10-PRI - PM10-FIL if all four non-condensable pollutants are reported.

In addition, if SLT agencies report a pollutant as zero that by definition must have higher emissions than another pollutant, and the zero value has been determined to be correct in the pre-screening checks, then the following steps are taken:

1. If PM10-PRI is reported as zero, all other pollutants are changed to zero.
2. If PM25-PRI is reported as zero, PM25-FIL and PM-CON are changed to zero.
3. If PM10-FIL is reported as zero, PM25-FIL is changed to zero.

After performing the trivial updates, the pre-screening steps are repeated to ensure that no inconsistencies had been introduced by the trivial update step.

4. NON-TRIVIAL UPDATES

A mapping of all possible reported pollutant combinations to non-trivial cases is provided in Table 1. A set of pollutants is considered equivalent to a non-trivial case if after applying the trivial updates, the same set of pollutants have populated values. A set of pollutants is considered trivial if after applying the trivial updates, all five pollutants have populated values.

For the processes in which missing pollutants are not able to be gap-filled by applying trivial updates, the processes are divided into uncontrolled and controlled processes. Both controlled and uncontrolled processes are further split into three SCC groups corresponding to the three Access® databases derived from the PM Calculator. The purpose of the split to three Access® databases is to manage the file size. Cases are then assigned to the SLT agency reported pollutants as shown in Table 2 below.

The PM Calculator databases have certain variables that depend on EPA's Source Classification Code (SCC), primary control, and secondary control. For the dataset containing the SLT process level PM data, primary and secondary control codes are assigned to every process. Beginning with the control information from the facility inventory in the EIS, obsolete control codes are replaced with the map-to equivalents from the EIS control measure code

table. Duplicates created from the mapping are eliminated. In addition, controls are evaluated to determine if the controls would have an impact on PM emissions. If a control is determined to not impact PM emissions, the control code is reassigned to 999. If only one control existed, it is considered primary, and the secondary code is set to 999. If only two existed, the first code is assigned as primary and the last code as secondary. For more than two controls, the control code 99 for “miscellaneous other” controls is eliminated. Then if there are two or more controls remaining, the first is assigned as primary and the last as secondary. A small number of processes had more than two controls excluding the “miscellaneous other” category. The assignments for these processes are arbitrary, as only two controls can be matched to the PM Calculator databases.

Emissions for the missing pollutants are calculated using the formulas in Table 2. Gap-

filling all missing pollutants for some processes using the Access® databases from the PM Calculator is not possible since the databases do not contain all SLT reported SCCs. The term "pmcalc_" in Table 2 refers to factors derived from dividing emissions output data from the PM Calculator databases. Where the table refers to (un)controlled factors, there are two separate fields in the PM Calculator databases for controlled and uncontrolled which were used depending on the status of the unit associated with the SLT reported emissions. The controlled factors refer to the matching of the primary and secondary controls as described above. The following terms identify the conversion factors discussed in the Strait et al. paper (1999): pm10pri_to_pm10fil and pm10fil_to_pmcon. Factors to convert PM25-PRI to PM25-FIL were not developed; therefore, SLT supplied PM25-PRI emissions are converted to PM25-FIL emissions using the conversion factors for pm10pri_to_pm10fil.

Table 1. Reported Pollutants to Non-Trivial Case Lookup

SLT Reported Pollutants	Non-Trivial Case	Equivalent Non-Trivial Case
PM10-PRI	1	
PM10-FIL	2	
PM2.5-PRI	3	
PM2.5-FIL	4	
PM-CON	5	
PM10-PRI and PM2.5-PRI	6	
PM10-FIL and PM2.5-FIL	7	
PM10-PRI and PM2.5-FIL	8	
PM10-FIL and PM2.5-PRI	9	
PM10-PRI, PM10-FIL, and PM-CON	10	
PM2.5-PRI, PM2.5-FIL, and PM-CON	11	
PM10-FIL and PM-CON		10
PM10-PRI and PM10-FIL		10
PM10-PRI and PM-CON		10
PM2.5-FIL and PM-CON		11
PM2.5-PRI and PM2.5-FIL		11
PM2.5-PRI and PM-CON		11
PM10-FIL, PM2.5-FIL, and PM-CON		Trivial
PM10-PRI, PM10-FIL, and PM2.5-FIL		Trivial
PM10-PRI, PM10-FIL, PM2.5-PRI, and PM2.5-FIL		Trivial
PM10-PRI, PM10-FIL, PM2.5-PRI, and PM-CON		Trivial
PM10-PRI, PM2.5-PRI, and PM10-FIL		Trivial
PM10-PRI, PM2.5-PRI, PM2.5-FIL, and PM-CON		Trivial
PM10-PRI, PM2.5-PRI, and PM-CON		Trivial
PM10-PRI, PM10-FIL, PM2.5-FIL, and PM-CON		Trivial
PM10-PRI, PM2.5-FIL, and PM-CON		Trivial
PM10-PRI, PM2.5-PRI, and PM2.5-FIL		Trivial
PM10-FIL, PM2.5-PRI, PM2.5-FIL, and PM-CON		Trivial
PM10-FIL, PM2.5-PRI, and PM2.5-FIL		Trivial
PM10-FIL, PM2.5-PRI, and PM-CON		Trivial

Table 2. Calculation of PM species not reported by SLT

SLT reported Pollutants	Non-Trivial Case	Final Augmented Pollutants
PM10-PRI	1	PM10-PRI=Original PM2.5-PRI=PM10-PRIx(pm10pri_to_pm10fil)x(pmcalc_pm25fil(un)controlled/pmcalc_pm10fil(un)controlled)+PM10-PRI-PM10-PRIx(pm10pri_to_pm10fil) PM10-FIL=PM10-PRIx(pm10pri_to_pm10fil) PM2.5-FIL=PM10-PRIx(pm10pri_to_pm10fil)x(pmcalc_pm25fil(un)controlled/pmcalc_pm10fil(un)controlled) PM-CON=PM10-PRI-PM10-PRIx(pm10pri_to_pm10fil)
PM10-FIL	2	PM10-PRI=PM10-FILxpm10fil_to_pmcon+PM10FIL PM2.5-PRI=PM10-FILx(pmcalc_pm25fil(un)controlled/pmcalc_pm10fil(un)controlled)+PM10-FILxpm10fil_to_pmcon PM10-FIL=Original PM2.5-FIL=PM10-FILx(pmcalc_pm25fil(un)controlled/pmcalc_pm10fil(un)controlled) PM-CON=PM10-FILxpm10fil_to_pmcon
PM2.5-PRI	3	PM10-PRI=(PM2.5-PRIxpm10pri_to_pm10filxpmcalc_pm10fil_(un)controlled/pmcalc_pm25fil_(un)controlled)+PM2.5-PRIx(1-pm10pri_to_pm10fil) PM2.5-PRI=Original PM10-FIL=PM2.5-PRIxpm10pri_to_pm10filxpmcalc_pm10fil_(un)controlled/pmcalc_pm25fil_(un)controlled PM2.5-FIL=PM2.5-PRIxpm10pri_to_pm10fil PM-CON=PM2.5-PRIx(1-pm10pri_to_pm10fil)
PM2.5-FIL	4	PM10-PRI=PM2.5-FILx(pmcalc_pm10fil_(un)controlled/pmcalc_pm25fil_(un)controlled)+PM2.5-FILx(pmcalc_pm10fil_(un)controlled/pmcalc_pm25fil_(un)controlled)xpm10fil_to_pmcon PM2.5-PRI=PM2.5-FIL+PM2.5-FILx(pmcalc_pm10fil_(un)controlled/pmcalc_pm25fil_(un)controlled)xpm10fil_to_pmcon PM10-FIL=PM2.5-FILx(pmcalc_pm10fil_(un)controlled/pmcalc_pm25fil_(un)controlled) PM2.5-FIL=Original PM-CON=PM2.5-FILx(pmcalc_pm10fil_(un)controlled/pmcalc_pm25fil_(un)controlled)xpm10fil_to_pmcon
PM-CON	5	PM10-PRI=PM-CON+PM-CON/(pm10fil_to_pmcon) PM2.5-PRI=(PM-CON/(pm10fil_to_pmcon))x(pmcalc_pm25_fil_(un)controlled/pmcalc_pm10_fil_(un)controlled)+PM-CON PM10-FIL=PM-CON/(pm10fil_to_pmcon) PM2.5-FIL=(PM-CON/(pm10fil_to_pmcon))x(pmcalc_pm25_fil_(un)controlled/pmcalc_pm10_fil_(un)controlled) PM-CON=Original
PM10-PRI and PM2.5-PRI	6	PM10-PRI=Original PM2.5-PRI=Original PM10-FIL=PM10-PRIx(pm10pri_to_pm10fil) PM2.5-FIL=PM2.5-PRI-(PM10-PRI-PM10-PRIx(pm10pri_to_pm10fil)) PM-CON=PM10-PRI-PM10-PRIx(pm10pri_to_pm10fil)
PM10-FIL and PM2.5-FIL	7	PM10-PRI=PM10-FIL+PM10-FILxpm10fil_to_pmcon PM2.5-PRI=PM25-FIL+PM10-FILxpm10fil_to_pmcon PM10-FIL=Original PM2.5-FIL=Original PM-CON=PM10-FILxpm10fil_to_pmcon
PM10-PRI and PM2.5-FIL	8	PM10-PRI=Original PM2.5-PRI=PM25-FIL+PM10-PRIx(1-pm10pri_to_pm10fil) PM10-FIL=PM10-PRIxpm10pri_to_pm10fil PM2.5-FIL=Original PM-CON=PM10-PRIx(1-pm10pri_to_pm10fil)

SLT reported Pollutants	Non-Trivial Case	Final Augmented Pollutants
PM10-FIL and PM2.5-PRI	9	PM10-PRI=PM25-PRIxpmcalc_pm10fil(un)controlled/pmcalc_pm25fil(un)controlled PM2.5-PRI=Original PM10-FIL=Original PM2.5-FIL=PM25-PRI-(PM25-PRIxpmcalc_pm10fil(un)controlled/pmcalc_pm25fil(un)controlled-PM10-FIL) PM-CON=PM25-PRIxpmcalc_pm10fil(un)controlled/pmcalc_pm25fil(un)controlled-PM10-FIL
PM10-PRI, PM10-FIL, and PM-CON	10	PM10-PRI=Original or Trivial Update PM2.5-PRI=PM10-FILx(pmcalc_pm25fil(un)controlled/pmcalc_pm10fil(un)controlled)+PM-CON PM10-FIL=Original or Trivial Update PM2.5-FIL=PM10-FILx(pmcalc_pm25fil(un)controlled/pmcalc_pm10fil(un)controlled) PM-CON=Original or Trivial Update
PM2.5-PRI, PM2.5-FIL, and PM-CON	11	PM10-PRI=PM2.5-FILx(pmcalc_pm10fil_(un)controlled / pmcalc_pm25fil_(un)controlled)+PM-CON PM2.5-PRI=Original or Trivial Update PM10-FIL=PM2.5-FILx(pmcalc_pm10fil_(un)controlled / pmcalc_pm25fil_(un)controlled) PM2.5-FIL=Original or Trivial Update PM-CON=Original or Trivial Update

All three groups of uncontrolled processes and all three groups of controlled processes are consolidated into a single table, and the pre-screening steps repeated to ensure consistency. In addition, the trivial updates are rerun with checks to ensure that the updates did not introduce any inconsistencies that would violate the pre-screening criteria.

To ensure that all PM10-PRI and PM25-PRI records contained values, any remaining null records for these PM species are populated according to the schema presented in Table 3. All null PM25-PRI records are populated prior to gap-filling null PM10-PRI records. For example, if PM25-PRI is null, PM25-FIL is null and PM10-PRI is not null, then PM25-PRI would be set equal to PM10-PRI.

Non-trivial and trivial emissions for all pollutants not reported by an SLT agency are consolidated into the EIS format. Results are split into two separate files, one containing overwrites and one containing additions. Overwrites are cases where the SLT reported emissions for a process/pollutant that were overwritten or deleted in the pre-screening because of inconsistencies. The overwrites file contains some null values where the SLT provided inconsistent information that was made null in the pre-screening and there was no replacement value since the SCC/control combination was not in the PM Calculator. Additions are the remaining process/pollutant level records that were added in the trivial or non-trivial updates sections.

Table 3. Gap-filling Schema for Null PM10-PRI and PM25-PRI Records

Pollutant with Null Record	Gap-filling Priority List (Null Record Set Equal to First Non-null Value in List)
PM25-PRI	1. PM25-FIL 2. PM10-PRI 3. PM10-FIL 4. PM-CON
PM10-PRI	1. PM10-FIL 2. PM25-PRI

5. REFERENCES

Strait, R., D. McKenzie, and R. Huntley, 1999: PM Augmentation Procedures for the 1999 Point and Area Source NEI.

AP 42, Fifth Edition, Compilation of Air Pollutant Emission Factors, Volume 1: Stationary Point and Area Sources, Appendices B.1 and B.2, <http://www.epa.gov/ttn/chief/ap42/index.html>