

SPECIATE DATABASE UPDATE

WORK PLAN

EPA Contract No. 68-D-00-265
Work Assignment No. 3-28

November 14, 2003

Prepared for:

Mr. Lee Beck
U.S. Environmental Protection Agency
Mail Drop E305-02
4930 Old Page Road
Research Triangle Park, NC 27711

Prepared by:

Dr. Frank Divita
E.H. Pechan & Associates, Inc.
5528-B Hempstead Way
Springfield, VA 22151

A. BACKGROUND AND INTRODUCTION

SPECIATE is EPA's repository of Total Organic Compound (TOC) and particulate matter (PM) source profiles. These profiles are used to: 1) create speciated emissions inventories for regional haze, PM_{2.5}, and O₃ modeling; 2) estimate hazardous and toxic air pollutant emissions from total PM and TOC primary emissions; 3) provide input to Chemical Mass Balance (CMB) source apportionment receptor models; 4) verify profiles derived from ambient measurements by multivariate receptor models (e.g., factor analysis, positive matrix factorization, UNMIX).

The current software is SPECIATE v3.2 which was completed in late 2002. Updates provided in SPECIATE v3.2 centered around improvements in the software functionality and display, but did not include updates to the source profile databases.

As discussed by an expert panel at the 2002 American Association for Aerosol Research (AAAR) meeting in Charlotte, NC, the SPECIATE database needs to: 1) identify chemical and physical characteristics of primary PM and TOC emissions; 2) tabulate and document fractional abundances and variabilities of specified chemical and physical components in primary PM and TOC emissions; and 3) provide data interfaces to receptor source apportionment models and speciated emission inventories.

The panel agreed that:

- C Mass fractions of specified chemical and physical properties must normalize to a known and documented quantity. For PM, the normalization mass should be size specific (typically PM_{2.5} and/or PM₁₀). The normalized mass can be measured or the weighted sum of major chemical components (sulfate, nitrate, ammonia, soil elements with assumed or measured oxides, organic carbon, elemental carbon, and sea salt). For TOC, normalization might be mass (g/g) or volume carbon (ppbC/ppbC) specific. The preferred normalization is to the sum of PAMS species that are typically measured in canisters. Other common normalizations are to the sum of measured components, total hydrocarbons, or total GC-elutable TOCs.
- C Hot gas compliance samples do not represent TOC or PM profiles that have cooled to ambient temperatures. Test methods must include this cooling and must be documented with each profile.
- C TOC profiles should include, at a minimum, the 55 (now 56) PAMS species, plus whatever additional species are available (additional light

hydrocarbons from canisters, heavy hydrocarbons from TENAX, and carbonyls from DNPH).

- C PM profiles should include, at a minimum, major elements (at least those reported by the IMPROVE and PM_{2.5} Speciation Trends networks), major water-soluble ions (sulfate and nitrate at a minimum, preferably ammonium, potassium, sodium, chloride, fluoride, phosphate, calcium, magnesium), carbon fractions (total carbon, organic carbon, and elemental carbon), preferably with other fractions such as the eight IMPROVE fractions, and carbonate carbon). Organic fractions, isotopic abundances, organic compounds, and single particle properties should be included where they are reported, well-defined, and can be normalized to PM or TOC mass. Since non-equivalent measurement methods (e.g., IMPROVE and STN carbon) are used, measurement methods must be specified with each profile.
- C Profiles must have chemical abundances (the fraction of PM mass emissions or total TOC) and their variabilities. Priority should be given to profiles that express an average and standard deviation of individual profiles for representative samples. Analytical uncertainties for individual source profiles should be identified as such and are typically much smaller than the variability of several samples taken at different times from the same emitter or from samples taken from different, but similar, emitters. The method for estimating variabilities should be specified with the profile.
- C The primary reference for each profile should be cited as documentation, not secondary references that have compiled profiles from primary references. Secondary references should be cited in addition when original profiles have been modified (i.e., by aerosol aging, different sample compositing, etc.) along with the original reference.

These recommendations serve as a starting point for improving the utility of the SPECIATE software and its contents. The objectives of this EPA project are to: 1) improve SPECIATE software and documentation; 2) obtain, evaluate, and document additional source profiles; and 3) acquire and integrate metadata (e.g., references, reports, test method descriptions) into the database structure. Under this project, primary technical work will be performed by E.H. Pechan & Associates, Inc. (Pechan). The Desert Research Institute (DRI) and Dr. David Allen, as subcontractors to Pechan, will provide assistance in performing Tasks 3 and 8. The technical approach for each Task is described below.

B. TECHNICAL APPROACH

The technical approach to this WA will be performed in eight tasks. Pechan's proposed approach to each of these tasks is presented below.

Task 1: Kickoff Meeting to Prepare Work Plan

On October 27th, Pechan participated in the project meeting with EPA. The participants included: Lee Beck (the EPA WAM), Tom Pace, David Mobley, Ron Ryan, and Dennis Beauregard (all of EPA) and Frank Divita, Pechan's proposed Project Manager.

A brief synopsis of the meeting was provided to the meeting participants via e-mail.

Task 2: Preparation of Detailed Work Plan; General Requirements

This task entails all management-related activities for this WA including: preparation of the project work plan and monthly progress reports, staffing the project and providing needed resources, review and quality assurance of all work products, cost tracking, and interaction with the EPA WAM and Project Officer. Pechan will inform the WAM by telephone of any problems that may impede project performance, along with any corrective actions needed by Pechan or EPA to solve the problem. Also, as part of this task, Pechan will notify EPA when 75 percent of the approved hours for this assignment have been expended, as well as provide any other contractually-required reporting.

The project management activities may also include, where necessary and/or appropriate: database administration, other specific quality assurance/review tasks, management of subcontractors, and/or any other specific WA or WA Amendment requirements. In addition, Pechan will insure that all products developed under this work assignment will comply with relevant data standards established by EPA's Office of Environmental Information, including the use of EPA-standardized chemical names and ID's.

Task 3: Database Design for PM and TOC Speciation Profiles

Pechan will review the structure of the current SPECIATE databases and design a new relational database (with tables, data elements, and linkages) that will accommodate present and future PM and TOC profiles. Data tables will include at a minimum: 1) source profile abundances; 2) PM size fractions, where applicable; 3) normalization methods; 4) source test methods; 5) analytical method(s); 6) source category and sub-category definitions; and 6) references. In addition, data elements we add include: measurement year, additional quality assurance (QA) flags (e.g., Data Attribute Rating System [DARS] codes), Chemical Abstracts Service (CAS) numbers, Source Classification Codes (SCC), Storage and Retrieval of Aerometric Data (SAROAD) codes, and uncertainty calculation

methods. Flags to distinguish between individual and composited profiles will also be added as needed.

Pechan will ensure that the database design will take into account the data requirements defined under Tasks 4 - 8 below, and those identified by the SPECIATE expert panel.

Pechan will prepare a draft technical memorandum describing the proposed database structure and provide it to DRI and Dr. Allen for review and comment. DRI and Dr. Allen will provide written suggestions for revision and improvement to Pechan. Pechan will incorporate these changes and submit the database design to EPA in the form of a technical memorandum.

If requested by EPA, Pechan, DRI, and Dr. Allen will participate in a conference call with EPA to discuss the proposed database design.

Task 4: Criteria for Inclusion and Classification of Profiles in the Database

Under this task, Pechan will prepare criteria for inclusion of new data into the SPECIATE database. Consideration for inclusion will include: scientific merit, likeliness of use by researcher, source type, collection method, availability of required data elements (including meta data and uncertainties) and the degree to which the profile is raw, averaged, or normalized data. In addition, Pechan will also propose a scheme for classifying and flagging profiles so the SPECIATE user can easily limit queries to profiles applicable to their intended use. As suggested by EPA, and because of different user needs, it likely will be appropriate to have more than one set of inclusion criteria, so that profiles useful to some, but not all, users are included.

Pechan will prepare a draft technical memorandum under this task and provide it to DRI and Dr. Allen for review and comment. DRI and Dr. Allen will provide written suggestions for revision and improvement to Pechan. Pechan will incorporate these changes and provide the proposed inclusion and classification scheme to EPA for review and comment.

If requested by EPA, Pechan, DRI, and Dr. Allen will participate in a conference call with EPA to discuss the proposed inclusion and classification scheme.

Task 5: Assessment and/or Classification of Profiles Now Present in SPECIATE

Under this Task, Pechan and its subcontractors will review the profiles currently in SPECIATE v3.2 and apply the approved criteria defined under Task 4 to identify those that are appropriate for retention in upcoming versions of

SPECIATE. SPECIATE v3.2 currently contains 376 PM source profiles and 567 TOC profiles. Pechan anticipates that many of these will not meet the Task 4 inclusion criteria.

Pechan will prepare a draft technical memorandum under this task and provide it to EPA for review and comment. The memorandum will include clear recommendations of which profiles to retain for upcoming versions of SPECIATE.

If requested by EPA, Pechan, DRI, and Dr. Allen will participate in a conference call with EPA to discuss the proposed inclusion and classification scheme.

Task 6: Inclusion of New Profiles

Additional TOC and PM source profiles will be identified, procured, and documented. At a minimum, these will include: 1) PM profiles associated with DRI source characterization projects conducted over the past two decades; 2) TOC profiles acquired during the Houston-Galveston Air Quality Science Evaluation project; and 3) California Air Resources Board PM and TOC profiles that do not overlap with 1 and 2 (<http://www.arb.ca.gov/emisinv/speciate/speciate.htm>).

For the PM profiles, DRI staff will prepare data tables compatible with the software and structure defined in Task 3 for source profiles measured in previous air quality studies. Both individual and composite profiles will be included, with documentation of which individual profiles are included in each composite. Project reports referring to methods and source characteristics will be scanned and provided in PDF format. A bibliography and copies of journal publications reporting profiles and methods will also be provided. A table summarizing the different projects and sources from which the profiles were derived will be provided as well. As shown in Table 1, more than 1,200 source profiles are available from the DRI database.

For the TOC profiles, Pechan staff will prepare data tables compatible with the software and structure defined in Task 3 for source profiles measured in previous air quality studies. These include about 600 official CARB source profiles, most of which are TOC, in addition to the profiles obtained from the Houston-Galveston Air Quality Science Evaluation project.

If adequate resources exist after completion of the work specified above under this Task, DRI will prepare data tables compatible with the software and structure defined in Task 3 using data contained in their in-house TOC source profile database (more than 700 TOC source profiles).

Task 7: Protocol for Inclusion of Future Profiles

Under this task, Pechan will prepare a draft protocol for including new data into the database. The purpose of this protocol, accompanied by the inclusion criteria developed in Task 4, is (1) to guide profile data collectors on how to collect and present profile data to maximize their utility to SPECIATE users, (2) to assist future SPECIATE managers in assessing whether data should be incorporated, and (3) to facilitate the mechanics of the actual inclusion.

In developing this profile inclusion protocol, Pechan will consider the data exchange protocol for speciation profiles developed by the Regional Planning Organizations (RPOs), whenever that exchange protocol is relevant to the profile inclusion protocol.

Pechan will prepare a draft technical memorandum under this task and provide it to EPA for review and comment. If requested by EPA, Pechan, DRI, and Dr. Allen will participate in a conference call with EPA to discuss the proposed inclusion scheme.

Task 8: Recommendations for the User-Friendly Front End

Pechan will provide recommendations to EPA for features that should be available in the front end that EPA plans to eventually create for the database.

Pechan will prepare a draft technical memorandum under this task and submit it to EPA for review and comment. If requested by EPA, Pechan, DRI, and Dr. Allen will participate in a conference call with EPA to discuss the proposed recommendations.

Table 1. Examples of Source Profiles to be Assembled

Study		SCENIC Denver Haze Study	CARB California Source Characterization Study	Nevada PM ₁₀ Source Apportionment Study	Santa Barbara PM ₁₀ Source Apportionment Study	Phoenix PM ₁₀ & Visibility Study	Ohio PM ₁₀ Study	Bay Area (San Jose) PM ₁₀ Study	Imperial Valley/ Maxicali PM ₁₀ Study	San Francisco Bay Area PM ₁₀ Study	Mt. Zirkel Visibility Study	Las Vegas PM ₁₀ Study	Antarctica (Mc Murdo) PM ₁₀ Study	NFRAQS (Denver) Visibility Study	Robbins Particulate Study	South Africa Source Contribution Study	Mexico City PM ₁₀ Study	CARB Fugitive Dust Characterization Study	Tuscarora Mountain Tunnel Diesel Emission Study	BRAVO (Big Bend National Park) Visibility Study	California Regional PM Air Quality Study (CRPAQS)*	GEEER Petroleum Comb. Emission*	Total	
Year		87-88	87-88	89	89	89-90	90	91-92	92-93	92-93	94-95	95-96	95-97	96-97	97	97	97-98	97	99	99-00	01-03	98-03		
Geological	Paved road dust ^a	7	35 ^{^,c}		2 ^e	6 ^d	9 ^d	2 ^e	13 ^e	7 ^e	9 [#]	11 ^e		6	6 ^b		8 ^b	5 ^{^,e}		14 ^{^,b}				
	Unpaved road dust	1	35 ^{^,c}			12 ^d	6 ^d			7 ^e	3 [#]	3 ^e			6 ^b		4 ^b	10 ^{^,e}		8 ^{^,b}				
	Soil	1	90 ^{^,c}	8 ^e	9 ^e	12 ^d	15 ^d		7 ^e	11 ^e	6 ^{+,#}	14 ^{+,e}	1 ^e		10 ^b	10 ^b	30 ^b	57 ^{^,e}		16 ^{+,^,b}				
	Fly ash								7 ^e							2 ^b				4				
Motor Vehicle	Gasoline	16				1							1 ^e	56 ⁻		2 ^b	8		1 ⁻					
	Diesel	3	10 ^{^,c}			1				1			2 ^e	13 ⁻			21		1 ⁻					
	Gasoline & Diesel	7				10			42	2	18 [#]	27 ^e					20		18 ⁻	39 [^]				
Vegetative burning	Residential	9	15 ^{^,c}								4 [#]	31 ^e		18 ⁻							6			
	Agriculture		20 ^{^,c}						9												3			
	Forest Fire										5 [#]									29 [^]				
Coal Combustion	Residential									5 [#]					8 ^b									
Cooking	Commercial								2					5 ⁻			7			17 [^]				
Industry and Utility Emission	Coal	5									16 [#]										33 [^]			
	Oil		10 ^{^,c}						1				3 ^e							6 [^]				
	Manure								5															
	Gas	1																					26 ^{^,-,#}	
Geothermal Spring										5 [#]														
Regional Background					2					3 [#]														
Others		1 ⁱ	10 ^{^,c,j}										2 ^{e,k}			4 ^{b,n,p}				14 ^{^,m}				
Total		51	225	10	11	44	30	2	86	28	74	86	9	98	22	28	98	72	20	180	9	26	1209	

- | | | | |
|---|---|---|--|
| a | Paved road and parking lot | + | including 1 overall geological profile |
| b | PM _{2.5} , PM ₁₀ | ^ | with carbonate |
| c | PM ₁ , PM _{2.5} , PM _{coarse} , PM ₁₀ , TSP | ~ | with organics |
| d | PM _{2.5} , PM _{coarse} , PM ₁₀ | # | with SO ₂ and NH ₃ |
| e | PM ₁₀ | * | profiles coming soon |
| i | Fluidized catalyst cracker regenerator | m | cement |
| j | Incinerator | n | lime kiln |
| k | C130 and UH-1 airplane | p | residential paper waste burning |

D. DELIVERABLES AND SCHEDULE

Table 4 lists the milestones and deliverables for the WA. The schedule has been adjusted to reflect DRI's assertion that the PM source profiles required under Task 6 have to be assembled and verified from many old project reports, and the time associated with doing so.

Table 4
Schedule of Milestones and Deliverables

Task Number	Task Description	Schedule Date
1	Kickoff Meeting	October 27, 2003
2	Work Plan	November 14, 2003
3	Relational Database Design Recommendations	December 5, 2003
4	Criteria for Inclusion/Classification of Profiles	December 5, 2003
5	Assessment/Classification of Profiles Currently in SPECIATE	January 9, 2004
6	Complete Inclusion of New Profiles	
	PM and TOC Profiles by DRI	June 30, 2004
	TOG and ARB Profiles by Pechan and Dr. Allen	January 16, 2004
7	Protocol for Inclusion of Future Profiles	June 30, 2004
8	Recommendations for User-Friendly Front End Features	June 30, 2004