DEVELOPMENT OF SOURCE SPECIATION PROFILES FROM THE TNRCC 2000 POINT SOURCE DATABASE

Final Report

Prepared by:

Pacific Environmental Services, Inc.
5001 S. Miami Boulevard, Suite 300
Research Triangle Park, NC 27709-2077

Under Subcontract to:

Environ Corporation
101 Rowland Way, Suite 220
Novato, California 94945-5010

Prepared for:

Ms. Jocelyn Mellberg
Texas Natural Resource Conservation Commission
Austin, TX

August 30, 2002

Disclaimer

This document was prepared by Pacific Environmental Services, Inc. under subcontract to Environ Corporation, for the Texas Natural Resource Conservation Commission (TNRCC). The opinions, conclusions, and recommendations are those of the authors, and do not necessarily reflect those of the prime contractor or TNRCC. The use of trade names does not imply an endorsement of any particular products by Pacific Environmental Services, Environ Corporation, or TNRCC.

Contents

Page

Tables

Figures

Page

INTRODUCTION

The Texas Natural Resource Conservation Commission (TNRCC) will be performing photochemical modeling of specific areas in Texas (primarily the Houston/Galveston area), for ozone episodes that occurred during the 2000 Texas Air Quality Study (TexAQS). An accurate inventory of emissions of the primary ozone precursors NO_x, and VOC are essential to generate reliable model results. In the case of VOC, many different specific organic compounds that represent a range of reactivity in ozone formation mechanisms can be present in the emissions mix representing a complex industrial base such as that in the Houston/Galveston area. Reactivity in this application represents a measure of the contribution that each individual chemical has in the overall process that results in the formation and accumulation of ozone in urban systems.

An accurate representation of the individual chemical species that are emitted from each source is required to achieve meaningful results in this type of modeling exercise. To meet this need, VOC emissions are allocated to specific organic compounds by the application of a speciation profile. The speciation profile represents the weight percent of the specific organic compounds that are typically emitted from a particular process. Speciation profiles are typically used to characterize emissions at the source classification code (SCC) level. Much of the data used to develop these SCC-average profiles are based on national-level information. Major point sources of VOC emissions, such as, chemical, petrochemical and petroleum refining operations can vary and the use of an average SCC profile can rarely be expected to represent any individual facility exactly. Therefore, development of source-specific speciation profiles is one approach to improve the overall VOC speciation for urban and regional modeling applications. Alternatively, the use of SCC-specific profiles based on data collected for sources in a particular geographic region will also result in improved speciation characteristics relative to the use of speciation profiles developed at the national-level.

In many areas of the United States where observed ozone concentrations exceed the National Ambient Air Quality Standard (NAAQS), the primary contribution of VOC is from mobile sources and dispersed stationary sources that are too small to track on an individual basis. Since point sources are not a major component of VOC in many areas, recent efforts at the national-level to develop more accurate VOC speciation profiles have focussed on area and mobile sources. In the Houston/Galveston area, however, significant point sources of VOC are present. The speciation profiles that are typically used for point sources in urban and regional modeling analyses are typically old and may be based on outdated measurement techniques. Add-on control devices and process changes have been implemented in many of the operations in the Houston/Galveston area that represent the largest point sources of VOC since the national default speciation profiles were developed. These activities have limited both the amount and reactivity of VOC emissions mix. For these reasons, many of the existing VOC speciation profiles that are routinely applied to point sources are not expected to accurately represent the VOC mix in the Houston/Galveston area, and therefore, may not accurately represent the reactivity of VOC in ozone formation processes.

Numerous studies using speciated ambient data available through the Photochemical Assessment Monitoring System (PAMS) network, along with assumed speciated emissions information have been completed in the past 5 to 10 years. In many of these studies, a significant discrepancy has been observed between species that are represented in the inventory and the species that are actually observed in ambient air downwind of the sources. Since these types of discrepancies have been observed and reported for low reactivity species as well as high reactivity species, it is not simply a result of reactions that occur between the source and the monitoring location, but rather a real lack of accuracy in the application of many of the common speciation profiles.

TNRCC has recognized this weakness and the effects it can have on modeling analyses and has implemented a program to improve the situation. The major point sources in Texas are encouraged to report individual VOC species directly from each source in addition to the total VOC emission. These estimates of emissions of individual species can be aggregated to form speciation profiles for specific process-level point sources and for SCC-level processes representative of the conditions that exist in the Houston/Galveston area. These data have been compiled in the Point Source Data Base (PSDB).

OBJECTIVE

The objective of this project is to review the existing data from the 2000 TNRCC PSDB and develop source specific and SCC specific VOC speciation profiles that can be used to improve the point source speciation of VOC emissions for future modeling episodes. The work completed in this project follows work that was conducted in a previous project that used a 1999 PSDB to generate source and SCC specific profiles. The goal of analyzing the 2000 PSDB is to build and improve on knowledge generated from the 1999 PSDB analysis.

In addition to the profile development using the 2000 PSDB, PES was also tasked to evaluate the TexAQS "Special Study Data" and use it to develop temporal data for a limited number of sources. After several discussions between TNRCC, Environ, and PES it was determined that this task should be cancelled because an insufficient amount of data had been collected. After the first set of 2000 PSDB profiles were generated by PES in March 2002, and the "Special Study Data" task was cancelled, PES was instructed that additional 2000 PSDB data was to be collected by TNRCC. The revised plan entailed PES completing another round of analysis on the revised PSDB instead of the "Special Study Data" task.

SUMMARY

TNRCC provided a data summary from the 2000 PSDB for analyses in this study. In previous work, PES had been supplied with separate files, one representing accounts within the Houston/Galveston area and the other representing accounts in the rest of the State. It was determined that performing the PSDB analysis on separate files for the purpose of separating the Houston/Galveston area from the rest of the State was not necessary. Therefore in this project, the PSDB was supplied to PES in one file containing points from the entire State of Texas. If it is necessary to analyze a particular area of Texas, the account numbers associated with the profiles generated in this effort can be linked to the correct area of Texas.

The focus of the current effort was to use the 2000 PSDB to develop speciated profiles for use by TNRCC in photochemical modeling. This portion of the project was to be conducted similarly to the work that had been performed on the 1999 PSDB while incorporating several improvements based on discussions with TNRCC and Environ personnel. A list of recommendations was based on these discussions and included in PES's report on the 1999 PSDB analysis.

PES completed a first-round analysis of the 2000 PSDB and submitted the resulting profiles and a detailed description of the profiles to TNRCC in March 2002. Following this analysis, TNRCC collected additional data and modified the 2000 PSDB. PES was supplied with a revised copy of the 2000 PSDB in July 2002. The summary of the 2002 PSDB discussed in this summary will address the procedures that were developed and applied to both the first and second analysis of the 2000 PSDB, but the specific references to profiles, and overall results pertain only to the second analysis.

The 2000 PSDB as supplied in July 2002 included data for 24,406 individual points at 1911 separate accounts.

Individual source-specific profiles and SCC-average profiles were constructed from data from those points that reported greater than 75% of the VOC emissions as species that can be identified as individual species. Since many of the "species" reported in the PSDB were not immediately identifiable as individual compounds or elements (e.g. crude oil), an effort to apply a recognized SPECIATE (EPA's repository for VOC and PM profiles) profile was undertaken to resolve as many of the undefined species as possible. This effort will be discussed in more detail later in this report. After resolving most of the species reported in the 2000 PSDB, source-specific profiles based on known VOC species that can be identified with a SAROAD/AIRS pollutant code were developed for 9,116 points. Additionally, 900 SCC-average profiles were developed from that database. The data have been compiled in a series of ACCESS® tables. The tables are included on a Compact Disk accompanying this report.

The methodology used to evaluate the original databases and develop the profiles is discussed. A more detailed presentation of the results is then provided, followed by a list of recommendations that could be implemented in the future to expand and improve the data available from the PSDB. Finally, a description of the ACCESS® table structures is presented as an attachment to the report.

METHODOLOGY

Data were received from the TNRCC from the 2000 Point Source Data Base (PSDB) in an ACCESS® file. The information contained in these files is summarized in Table 1.

The data provided by TNRCC were reviewed, and statistics on the content of these databases were prepared. A large fraction of the emissions in these databases are expressed as a general designation representing a collection of organic materials, as process raw materials, as industrial products, or as generic chemical classes. Examples of the general designations are non-methane VOC, and VOC gas mixture. Similarly examples of entries characterized as raw materials and products include crude oil and gasoline. Finally, examples of generic chemical classes represented in the database include alcohols undifferentiated, and aromatics undifferentiated. The databases also include entries for species that are not VOC (e.g., ethane, trichloroethane, acetone, hydrogen cyanide, etc.) Table 2 is a summary of the content of the original database as supplied to PES in July 2002.

The first step in the analysis of the 2000 PSDB was to delete all the chemicals that do not meet the definition of VOC from the data. There were 42 chemicals representing only 2.2% of the total mass in the original PSDB that were eliminated from the data before performing analysis. This list of chemicals is included in the ACCESS® file supplied in conjunction with this report in the table titled "Chemical List Master" with the "Unspecd" field designated as "R" for "Remove".

After deleting all the non-VOC chemicals, the generic chemical classes were examined and grouped in such a way that they could be further speciated using existing SPECIATE and other similar profiles. The list of 83 generic chemicals was evaluated to determine the number of occurrences in the PSDB, and the amount of each generic chemical in the data. Based on the amount of each generic chemical in the data and the number of occurrences, the generic chemicals were evaluated to determine a speciation profile that could be used to define the chemical. Speciation profiles were used to replace 41 of the 83 generic chemicals resulting in resolution of 53.54370 of the 71.44270 tons of generic chemicals (75%). Following resolution of the 41 chemicals, only 3.7% of the PSDB remained generic. The 174.39360 tons of Nonmethane VOC-U could not be resolved using speciation profiles directly.

The speciation profiles used to resolve the generic chemicals are included in the ACCESS® file as the following tables:

- Spec Crude Oil
- Spec Gasoline
- Spec Naphthas
- Spec Refinery
- Spec Stoddard Solvents

These speciation profiles have been incorporated into the final 2000 PSDB profiles, and are explicitly included in the data files for your review only. The generic species "Crude oil" with contam_cod "59001" occurred in the PSDB 508 times and represented more than 16 tons of reported VOC (the highest of any of the generic chemicals). An examination of the sources associated with the reported emissions of crude oil consisted mostly of large crude oil storage tanks and pipelines. SPECIATE profile number 2487 titled "Composite of 7 Emission Profiles from Crude Oil Storage Tanks – 1993" was used to resolve all the occurrences of crude oil in the PSDB.

The second highest representation by a generic species was by "Gasoline" – contam-cod "59003" which occurred in the PSDB 544 times and represented nearly 12 tons of reported VOC. A speciation profile for Houston area gasoline from the summer of 2000 was supplied by ENVIRON to resolve the occurrences of "gasoline" in the PSDB. The profiles were supplied by ENVIRON in three formats: 1. Liquid gasoline, 2. Gasoline vapor, 3. Tunnel emissions. The profile for gasoline vapor was used, since the PSDB represents actual emissions from point sources and emissions of gasoline would be in the form of vapor.

Several generic species could be characterized as naphthas, and were resolved using a profile for naphthas supplied in the ENVIRON report – "Speciated VOC Emissions for the Dallas/Fort Worth Nonattainment Area", October 1997.

Several generic species could be characterized as stoddard solvents/mineral spirits, and were resolved using SPECIATE profile number 1193 titled "Drycleaning", which was developed from a composite of the headspace from 5 mineral spirit samples.

An evaluation of the remaining generic species in the PSDB showed that most could be described as refinery by-products, in-process refinery feedstocks, and general emissions from refinery processes or storage tanks. Generally, most of these generic chemicals are not defined well enough to speciate individually, but could be speciated effectively as a group – all associated as refinery emissions. Therefore a number of the generic species were resolved using SPECIATE profile number 2457 titled "Composite of 10 Emission Profiles - Misc. Chemical and Refining Plants in Houston – 1993".

An examination of the table "2000 Unspecified Chemical Summary" in the ACCESS® database shows the analysis of the generic chemical list from the 2000 PSDB. The "Spec_code" entry shows which speciation profile the generic chemical was assigned to. "Mass_spec" shows the mass in tons of the generic chemical in the PSDB. "Pct_tmass" shows the percentage of the generic chemical in the PSDB. "Occurrence" shows the number of times the generic chemical appears in the PSDB.

The master chemical table "Chemical List Master" was populated with SAROAD codes and CAS codes where those codes could be determined. PES contacted EPA personnel in an attempt to obtain a master list of SAROAD codes, but EPA is not maintaining such a list at this time. As there exists no recognized standard procedure for assignment of SAROADs, one had to be adopted to complete the SAROAD assignment exercise. The following hierarchical approach was obtained from ENVIRON and used to make the SAROAD assignments:

1. Find an exact match, e.g., toluene assigned to toluene.
2. Match to a general VOC category that includes the specific VOC isomer, e.g., 2,4,4-trimethyl-1-pentene assigned to "c7 olefins."
3. Match to a similar isomer, e.g., 2,2,4-trimethylhexane assigned to 2,2,5-trimethylhexane. Approach 2 is preferred to 3 because 2 shows more clearly that a reassignment has been performed. In this case, the assignment to a similar isomer is preferred over a general "c8 alkane" assignment because it permits the presence of a tertiary carbon to be identified, which impacts the resulting CB4 split.
4. Match to the surrogate with the most similar properties, e.g., 2,4-dimethyl-1-pentene assigned to 3-ethyl-2-pentene. Here the surrogate is selected to show an alkene branched at the double bond.

Additionally, 74 chemicals were added to the chemical list to accommodate chemicals introduced in the stoddard solvents profile. These chemicals were assigned to contam_codes beginning "C" to identify them as codes introduced as a result of this work.

After resolving the generic species and assigning the SAROAD and contam_codes, the process of developing the profiles began. The PSDB was regenerated with the resolved specie substitutions replacing the generic species. The quantity of emissions from each individual point was totaled and the percent of emissions represented by unknown species was calculated for each point. Profiles were only generated for those points with 75% or greater of the total VOC emissions represented by known species.

Separate profiles were created for each individual point and for the aggregate SCC level. The profiles that contained unknown species were normalized to develop a profile using the remaining known species. For example, if a profile contained 90% known species and 10% unknown species (e.g., non-methane VOC), the unknown specie would be removed from the profile, and each of the known species would be increased by 10%. In that way, 100% of the mass would be represented in the profile and all of the emissions could be associated with some particular chemical compound.

Table 1. Summary of 2000 PSDB Information

Parameter	Description	Notes
Area	TNRCC Area Code	3-character code used to identify one of 11 areas in Texas as recognized by TNRCC
Pollutant	Reportable Criteria Pollutant	In this case, VOC
SIC	Standard Industrial Classification Code	Standard code to represent industry type, recently replaced by NAICS North American Industrial Classification System
SIC_clss	Industry Sector Identifier	Example: Organic Chemicals
Business	Subset of Industry	Example: Petrochemicals and Polymers
SCC	Source Classification Code	Code used to describe specific processes that result in air emissions
Fac_name	Process Description	Related to SCC
Account	Identifier code for a particular plant or industrial facility	Unique code used by TNRCC
FIPS	State and County FIPS codes	Nationally recognized State and County codes
Plant	Plant identifier	TNRCC Plant identifier
Fac_id	Equivalent to an individual emission point	Identifier to represent different emission processes at an account
Point_id	Equivalent to an individual emission stack	Individual source resulting in emissions
Ozone Season	Emissions expressed in tons per day during the ozone season for specific contaminant	Emissions are specific for each process and contaminant
Contam code	Identifier used by TNRCC to represent individual chemical emissions species	Unique code used by TNRCC
CAS	Chemical Abstract Service Number	Nationally recognized pollutant code
Species	Preferred species name	Mnemonic identifier

Table 2. Overall Summary of the 2000 PSDB

Unique species represented	465	
Total VOC emissions	481.88 (tpd)	
Generic and unknown species		84
Generic and unknown species emissions	245.84 (tpd)	51.0%
Known species		381
Known species emissions	236.04 (tpd)	49.0%

RESULTS

General

The raw 2000 PSDB represents a total of 484.12 tons per day (tpd) of VOC emissions from 1911 accounts or plants. Although the database is intended to represent speciated emissions, 174.39 tpd or 36% of the total emissions were reported simply as non-methane VOC undifferentiated. Figure 1 is a map of the Texas showing the locations of all PSDB accounts, with added detail to show the accounts that supplied enough speciated data to develop at least one profile using the 75% known criteria. At least one profile was developed for 940 of the 1911 accounts. Figure 2 shows the location of all the accounts with more than 1 tpd of VOC emissions, with added detail to show the accounts that supplied enough speciated data to develop at least one profile using the 75% known criteria.

Source Specific and SCC Average Profiles

It was possible to develop 9,116 point specific profiles and 900 SCC average profiles using the data from points with at least 75% resolved chemical data. Some of those profiles contain species that are either unknown or are species for which no SAROAD/AIRS pollutant code was available. Therefore, all profiles were normalized to represent only known species for which SAROAD identifiers were available.

Each of the SCC average profiles can be used to speciate any emissions source having an identical or similar SCC. This is the similar method used commonly to speciate emissions records using national-level speciation profiles. These profiles will improve the speciation of sources particularly when applied to the specific regions from which they are developed.

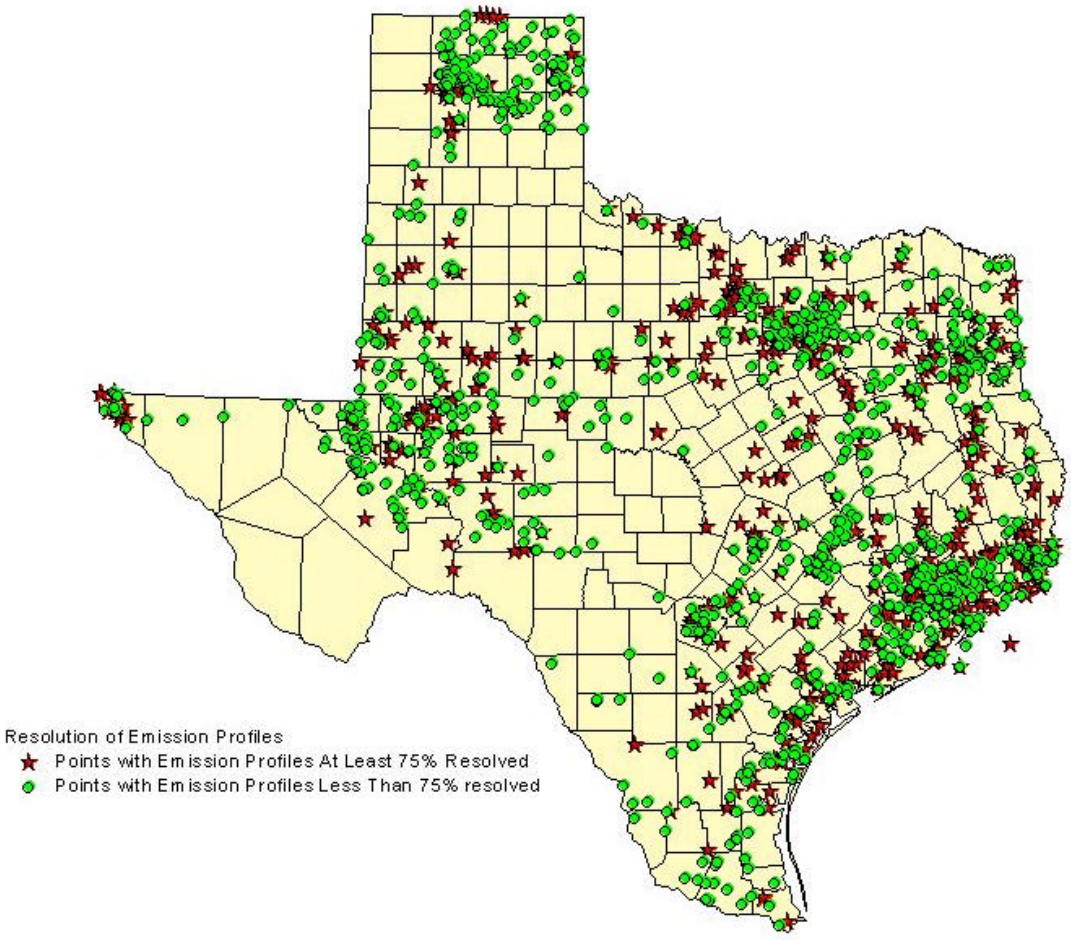

Figure 1. Location of All Accounts Represented in the 2000 PSDB

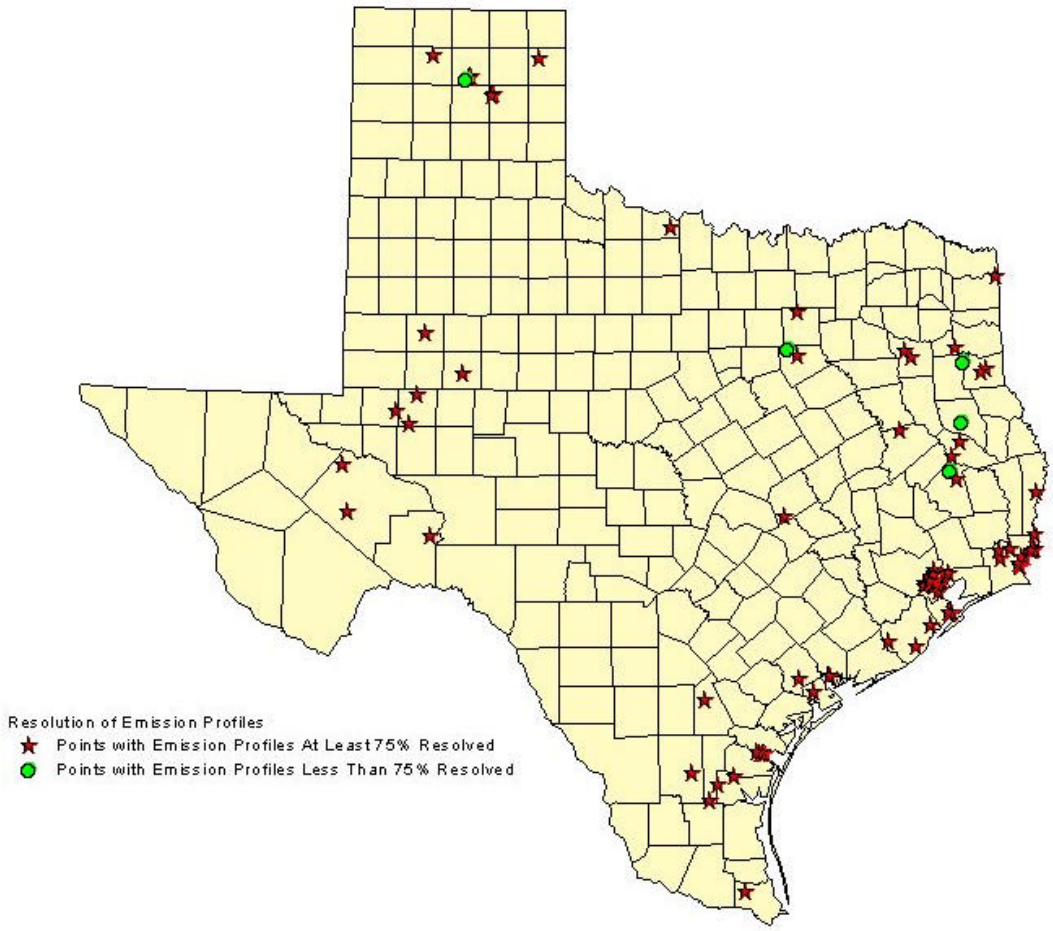

Figure 2. Location of Accounts Reporting More Than 1 TPD of Total Emissions

Sources of Ethylene and Propylene

Currently, sources of ethylene and propylene are of particular interest to TNRCC and are receiving attention in analyses of emissions information. Therefore, the data representing the contributions of ethylene and propylene were extracted and summarized. This exercise offers a good opportunity to use the data for an actual analysis of interest to demonstrate the utility of the PSDB.

In the 2000 PSDB, the total emissions of ethylene and propylene are 19.48270 tpd and 10.92390 tpd respectively. The combined total of the ethylene and propylene emissions represent 6.2% of the total mass reported in the PSDB. After resolving the generic species and normalizing the profiles to total 100% by removing the unresolved species such as Nonmethane VOC-U, the total emissions of ethylene and propylene are 18.87390 tpd and 9.75941 tpd respectively in the 75% known profiles. The combined total of the ethylene and propylene emissions represent 11.6% of the total mass in the 75% known profiles. Figure 3 represents the location of all the accounts represented in the emission profiles that emit ethylene and/or propylene.

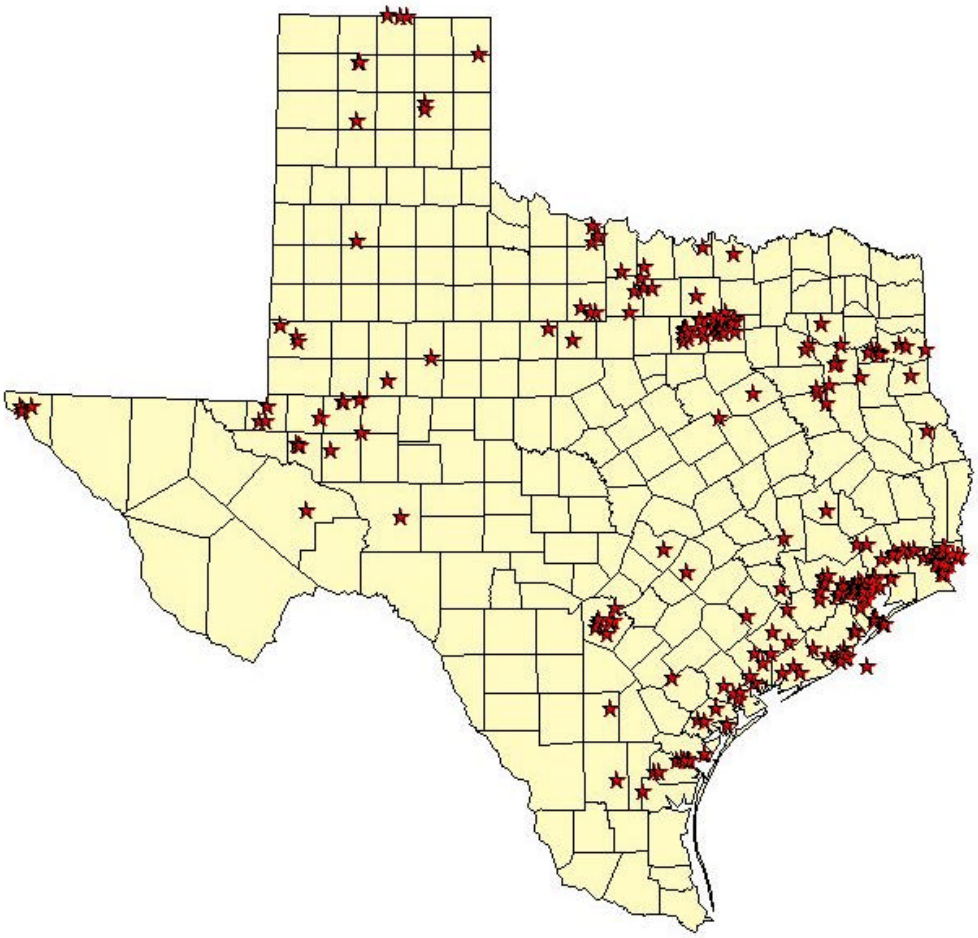

Figure 3. Location of Accounts Emitting Ethylene and/or Propylene

Recommendations

Several improvements were incorporated into the analysis of the 2000 PSDB that were not applied to the 1999 PSDB analysis. A considerable effort to resolve generic species in the 2000 PSDB was completed. The resolution involved the application of existing profiles to chemical categories, so that all the chemicals in the profiles could be identified as individual chemicals and not groups. Chemicals that were identified as not meeting the definition of VOC were removed from the PSDB. The resolution of most of the chemical data in the PSDB paved the way for creation of profiles for sources that previously would not have had the resolution necessary to create the profiles. Although a 75% known criteria was determined as the cut-off for profile creation, the criteria was applied after resolution of the chemical data, resulting in a more robust collection of profiles for the State of Texas.

If TNRCC continues to compile the speciated PSDB in future years, it would improve the quality of data and resulting profiles if only speciated data are collected for inclusion in the database. The large amount of Nonmethane VOC-U reported in the PSDB only serves to dilute the "real" data that exists in the database.

Although the reason that the TexAQS "Special Study Data" was not used for development of temporal data was due to lack of data, improvements in the data structure could be made to make it more useful in the future. If the TexAQS Special Study is completed by TNRCC at a later date, it would be more compatible to include this type of temporal data in a database other than EXCEL®. An ACCESS® or FoxPro® database with a more defined data structure would allow for more efficient analysis and use of the data.

Further improvements in the profiles could be made in the future by doing additional research to locate existing emission profiles that could be used to more accurately process unresolved chemical data from the PSDB. Also, comparisons of data collected in future years could be compared to previously developed profiles to help with developing conclusions about the accuracy and consistency of the PSDB data and profiles.

ATTACHMENT 1

EMISSION PROFILE DESCRIPTIONS:

Point-Specific Speciation Profiles:

The data for the 2000 point-specific speciation profiles are contained in two tables. The table described as "2000 Point Profile ID" contains one record for each profile. In this table, each profile is identified by a unique number assembled from a combination of the account id + fac id + point id. The "2000 Point Profile ID" table also contains all other relevant information supplied in the PSDB, such as the facility id, point id, area, facility name, SCC, SIC, SIC class, and business. Several calculated fields are also provided to show additional relevant information about the profile. These fields include the total mass represented by the profile in the PSDB, the mass of Nonmethane VOC-U, unspeciated mass remaining after assignment of generic chemicals to speciation profiles, the effective mass used in determining the emission profile, and the percents of unspeciated mass, Nonmethane VOC-U, and known chemicals.

Point source emission profiles were generated for all the point sources that contained at least 75% total resolution following resolution of the generic chemicals in the data. Since it was not possible to speciate the 164 tons of Nonmethane VOC-U even after resolution of the generic chemicals many of the point sources in the PSDB did not contain enough resolved data to create a realistic point source profile. Due to resolution of most of the generic data in the PSDB, PES was able to generate emission profiles for 6,110 of the 20,502 points in the PSDB. Many of the remaining sources in the PSDB can be effectively characterized by the SCC profiles discussed in this memorandum. In the work that was done previously to develop profiles from the 1999 PSDB, the 75% resolution criteria was used without attempting to speciate the generic chemicals, and therefore was not as effective.

The table described as "2000 Point Profile Data" contains one record for each chemical specie for each profile identified in "2000 Point Profile ID". As in "2000 Point Profile ID", the unique profile number is included in this table, and is the key connecting the two tables. The chemical specie is identified by a unique contam_code, SAROAD number where available, CAS number, and the chemical name. "2000 Point Profile Data" also includes the amount of each chemical species emitted described by the "Tonperday" field, and the percentage of the chemical in the profile described by the "Prof_pct" field.

SCC-Specific Speciation Profiles:

The data for the SCC-specific speciation profiles are contained in two tables. The table described as "2000 SCC Profile ID" contains one record for each profile. In this table, each profile is identified by SCC. The "2000 SCC Profile ID" table also contains the total mass of compounds in tons per day emitted by the process. The SCC profiles were created using only data from facilities containing at least 75% known chemicals.

The table described as “2000 SCC Profile Data” contains one record for each chemical specie for each profile identified in “2000 SCC Profile ID”. As in “2000 SCC Profile ID”, the SCC is included in this table, and is the key connecting the two tables. The chemical specie is identified by a unique contam_code, SAROAD number where available, CAS number, and the chemical name. The amount of each chemical species emitted is described by the “Tonperday” field. “K-prof_tot” shows the SCC profile including only the resolved chemical species as supplied in the PSDB and derived using the speciation profiles.

For point sources that could not be sufficiently resolved, in many cases the SCC profile can be used in place of a point source profile to characterize that source.