Development of Factors to Estimate 4-digit SIC Employment from 6-digit NAICS Employment Data for Selected Nonpoint Emission Inventory Categories With Example VOC Emission Calculations

Robert E. Wooten, Jr., P.E.
North Carolina Division of Air Quality, 1641 Mail Service Center, Raleigh, NC 27699-1641, bob.wooten@ncmail.net

Abstract

For a number of nonpoint source (previously area source) categories, emission estimates of VOC can be made by multiplying a factor times employment in certain SIC categories. Employment is no longer reported by SIC categories (since 1997). This paper describes development of a conversion crosswalk to use the appropriate portions of 6-digit NAICS based employment data to construct estimates of the various 4-digit SIC employment values needed to calculate selected nonpoint source emissions. Example emission calculations are shown. Access database files and an Excel file demonstrating crosswalk development, use of the crosswalk, and processing of employment data are offered.

Introduction

A number of nonpoint (area source) emission inventory categories may be estimated by multiplying employment in particular Standard Industrial Classification (SIC) categories by an appropriate emission factor. However, the SIC approach for recording and reporting employment data was discontinued after 1997. Since 1997, employment is reported in accordance with the North American Industry Classification System (NAICS). For the year 1997 both systems were used. Data from 1997 is used to construct a crosswalk file to estimate equivalent SIC employment from data reported under the NAICS system.

An Access file used to develop the SIC to NAICS crosswalk accompanies this paper to illustrate the development process. An Excel file used to make a working file from 2004 County Business Patterns data is provided. Finally, a second Access file is provided which illustrates the use the crosswalk file and the County Business Patterns employment data to calculate emissions for several inventory categories. These files should be available to view when reading this paper.

Factor Development and Use

The NAICS to SIC crosswalk was developed for the SIC codes corresponding to the following nonpoint source categories:

- Industrial Surface Coating: Furniture and Fixtures, Metal Containers, Automobiles (new), Machinery and Equipment, Appliances, Other Transportation Equipment, Sheet Strip and Coil, Factory Finished Wood, Electrical Insulation, and Marine Coatings

- Solvent Cleaning: Automobile Repair Cold Cleaning, Manufacturing Cold Cleaning, Electronics and Electrical Vapor and In-Line Cleaning, and Other Vapor and In-Line Cleaning

- Dry Cleaning: Coin Operated and Commercial / Industrial

- Automotive Refinishing

The SIC codes and emission factors pertaining to these categories may be found in the EIIP documents[1].

Crosswalk files in .dbf format were downloaded from the U.S. Census Bureau web site[2]. These files may be seen as tables E97B1 and E97B2 in the Access file SICtoNAICS.mdb[3]. These files contained extraneous material that was removed through queries. There were columns with numeric codes and letter codes for which there was no explanation given. Mr. Paul Zeisset of the U.S. Census Bureau explained these items[4].

When viewing SICtoNAICS.mdb it may be helpful to view the properties of the files and queries. Notes are provided that will help understand the origins and relationship of the queries and tables. The query SIClist3 was used to create table SIClist4 which has only the SIC codes needed for the area source categories under consideration. The query EMPLOYMENTfraction creates the table SICfractOFnaics that is the final crosswalk table. In this table the column SIClist4ADJ_EMPF contains indicators of employment estimates which were necessary because the Census Bureau must maintain confidentiality for individual establishments. A note such as “f est” beside the number 750 in the column to the left means that 750 is the midpoint estimate of the range of employment given. A note such as “l est*” beside 56338 indicates that a better than midpoint of range estimate was made by considering the data found on the Census Bureau web pages showing the bridge between NAICS and SIC[2].

The nonpoint source categories have the issue of potential overlap with point sources. No information was available concerning the number of employees at point sources. Therefore it was decided that employment of less than 100 at an establishment would place the establishment in the nonpoint category. Since estimates of employment based on ranges used the midpoints, the employment at establishments with 50 to 99 employees was estimated to be 75.

In order to make use of the NAISC to SIC crosswalk it is necessary to have county level employment data and to remove the establishments with 100 or more employees. The 2004 county business pattern data for North Carolina (2005 data was not available at the time) was downloaded from the Census web site[5] and processed to make an Excel file. It

was necessary to make estimates of employment from ranges provided, in most cases. The file ModCBP04NC.xls[6] may be examined for methods of best employment estimates. The tab ModCBP04NC was the starting point. Tab EmplDisc_noLarge shows actual employment given where there were no establishments with 100+ employees. Tab EstbExcdEqlEmploy shows estimates where number of employees were equal or less than establishments (numbers were added). Tab EMPtotNotDisc shows center of range estimates where no exact employment was given. Tab totlsDISCO2copy shows situations with total employment was stated and there were establishments with employment 100 and greater. In some cases, middle of range numbers for less than 100 employee establishments were adjusted down to allow sufficient employees for the larger establishments at the minimum of their size ranges. Tab AdjEmp1_99 combines the appropriate worksheets and provides the final employment estimates for establishments with fewer than 100 employees.

The NAICS to SIC crosswalk and the employment data are used in the file EMP_EMIS_2004cbp.mdb[7] to calculate emissions for the nonpoint source categories. This file also was used for processing the 2004 County Business Pattern data to make the Excel file used to make the small establishment employment estimates. The properties notes for the tables and queries should be examined for information about their purpose. The table named CategoriesCodes is where emission factors and the Source Classification Codes (SCC) used to link to projection factors are found. Also, the days per week assumptions (for daily emissions if wanted) and the estimated emission reductions are here. The final output is from query PROJemis2005.

The NAICS codes were revised in 2002 but none of the revisions involved the code numbers used for the work shown here. Future NAICS revisions should be investigated in case revisions to the NAICS to SIC crosswalk are necessary.

Conclusions

The method of producing a NAICS to SIC crosswalk file for certain nonpoint source categories has been shown. Also demonstrated are the adjustments needed to use employment data and the calculation of estimated emissions. Any revisions to NAICS codes should be examined for effects on the crosswalk numbers.

References

1. Emission Inventory Improvement Program (EIIP), Technical Report Series Volume III, Area Sources, http://www.epa.gov/ttn/chief/eiip/techreport/volume03/
2. http://www.census.gov/epcd/ec97brdg/
3. SICtoNAICS.mdb, A file accompanying this paper.
4. Paul Zeisset, 2007, U.S. Census Bureau, *personal communication*
5. http://www.census.gov/econ/www/index.html
6. ModCBP04NC.xls, A file accompanying this paper.
7. EMP_EMIS_2004cbp.mdb, A file accompanying this paper.