Quality Assurance and Estimation of Emissions Activity Data in the National Emissions Inventory

Andy Bollman, Holly Chelf and Randy Strait
E.H. Pechan & Associates, Inc., 3622 Lyckan Parkway, Suite 2002, Durham, NC 27707
abollman@pechan.com

Leif Hockstad
U.S. Environmental Protection Agency, 1200 Pennsylvania Ave., NW, Washington, DC 20460
hockstad.leif@epa.gov

ABSTRACT

The U.S. Environmental Protection Agency (EPA) prepares national emission estimates for the criteria air pollutants and ammonia in the National Emissions Inventory (NEI). The majority of NEI criteria point source category estimates are compiled from data submitted by State/Local/Tribal (S/L/T) agencies. Use of information from the NEI is often hampered by the lack of complete reporting of emissions activity data, which is the "driver" behind emissions estimates. Often these data are not reported due to concerns about confidentiality.

In its efforts to investigate additional sources of activity data for the U.S. greenhouse gas emissions inventory, the EPA's Clean Air Markets Division (CAMD) instructed E.H. Pechan & Associates, Inc. (Pechan) to compile fossil fuel consumption activity estimates from Version 2.0 of the NEI for industrial boiler/internal combustion (IC) engine source categories. Because emissions activity data are often missing for these categories, Pechan developed a "decision tree" of procedures for estimating emissions activity values for industrial boiler/IC engine source category records.

In conducting research in support of this effort, Pechan discovered numerous industrial boiler/IC engine records with suspect emissions activity and/or control information. Therefore, Pechan also developed quality assurance (QA) procedures for identifying suspect NEI emission activity, emission factor, and control information and for replacing the information in these fields with more reasonable default values.

This paper describes the decision tree of rules and procedures that were developed in this study. Also included are comparisons of fossil fuel consumption estimates derived from NEI Version 2.0 data with consumption estimates reported by the Department of Energy. This paper concludes with a discussion of the major findings and limitations of the study, and areas for future research.

INTRODUCTION

Background

The EPA prepares national emission estimates for the criteria air pollutants and ammonia in the NEI. The emissions of each pollutant are estimated for many different source categories, which collectively account for all anthropogenic emissions of these pollutants. Emissions are estimated for all 50 of the United States and these estimates are developed using a variety of methods under the leadership and direction of the Emission Factor and Inventory Group in EPA's Office of Air Quality Planning and Standards. The 1999 NEI point source estimates are compiled from emission inventory data submitted by S/L/T agencies and supplemented with data from Version 1.5 of the NEI.[1]

Use of information from the NEI is often hampered by the lack of complete reporting of the data that reflect the level of emissions activity. These data are often not reported due to expressed concerns about confidentiality. However these data can be used to serve many important functions, including QA

of S/L/T submitted data and calculation of emission estimates for pollutants not supplied by S/L/T agencies.

Purpose

The purpose of this study was to develop methods for estimating missing NEI activity data. The EPA's CAMD authorized Pechan to develop methods for estimating NEI emissions activity (fuel consumption) data for industrial boiler/IC engine source categories. These categories were selected because of their large contribution to total point source fuel combustion carbon monoxide (CO) emissions in Version 2.0 of the 1999 NEI. CO emissions were selected as a reasonable surrogate for fuel combustion activity because CO emissions are directly related to fuel combustion, and these emissions, relative to other pollutant emissions, are least likely to be controlled. The latter fact is important because the NEI reports actual emissions that incorporate the effects of emission controls. Because EPA's CAMD was interested solely in fossil fuel combustion, industrial wood/bark waste combustion categories were excluded from the study.

This paper describes the development of the methodologies used to estimate missing NEI activity data. Because a review of NEI Version 2.0 data identified outlier values that did not appear to be reasonable, Pechan also developed methods for identifying and replacing emissions activity outliers with reasonable estimates. The results of implementing the emission activity estimation decision rules are provided along with comparisons of the revised 1999 NEI activity data (after implementing the decision rules) with industrial energy consumption estimates from (a) the Energy Information Administration (EIA)'s *State Energy Data Report* (*SEDR*) *1999*;[2] and (b) EIA's 1998 Manufacturing Energy Consumption Survey (MECS).[3] This paper concludes with a discussion of the major limitations of the analysis and recommendations of areas for future research.

METHODOLOGY

Pechan developed procedures for estimating missing NEI activity data by first reviewing the activity and control information supplied in Version 2.0 of the 1999 criteria air pollutant NEI.[4]

Review of 1999 NEI Activity Data

Pechan compiled a summary of the industrial boiler/IC engine fuel consumption estimates reported in the final NEI Version 2.0, which was completed in October 2002. As part of the QA of the final NEI, EPA removed S/L/T throughput and emission factor (EF) data that were reported in units other than those permitted by NEI Input Format (NIF) Version 2.0. To assist in this effort, Pechan included the throughput values from the draft NEI in cases where they had been removed for the final NEI. Pechan first converted the draft NEI throughput values to the standard units for each source classification code (SCC) whenever such conversions could be made. In some cases, however, it was not possible to utilize the draft NEI throughput values because they were reported using units that are not relevant to the fuel for the given source category (e.g., coal throughput reported in million cubic feet rather than tons).

Table 1 presents a summary of 1999 fuel consumption estimates compiled from the emission activity data reported in the NEI. This table also displays comparisons of the NEI estimates with fuel consumption values from two EIA sources: the 1999 *SEDR* and data tables summarizing the results of the 1998 MECS. The third column in this table displays the total of the 1999 NEI energy consumption estimates for the industrial boiler/IC engine fuel combustion SCCs. Because the 1999 *SEDR* represents data for Standard Industrial Classification (SIC) codes 01-39 and many of the NEI industrial fuel combustion SCC records are identified with SIC codes outside this range, the fourth column displays the NEI consumption estimates excluding these records. The fifth column presents the *SEDR* industrial

energy consumption estimates. Because the MECS source represent data for SIC codes 20-39, the sixth column displays the NEI industrial fuel SCC consumption estimates for records associated with these SIC codes. The final column presents the 1998 MECS estimates.

Table 1. Comparison of pre-augmentation 1999 NEI, 1999 *SEDR*, and 1998 MECS national energy consumption estimates.

			SIC Codes 01 - 39		SIC Codes 20 - 39	
Fuel	**Units**	**1999 NEI Total**	**1999 NEI**	**1999 SEDR**	**1999 NEI**	**1998 MECS**
Natural Gas	billion cu ft	28,528	13,638	10,067	9,356	6,481
Coal and Coke	(million short tons)	21	13	96	11	78
Residual Oil	(million bbl)	139	138	40	138	57
Distillate Oil	(million bbl)	90	76	185	65	26
LPG	(million bbl)	23	4	624	4	38
Gasoline	(million bbl)	0.03	0.01	29	0.01	N/A

N/A - Not available.

Note: Unlike the MECS data, the *SEDR* data includes nonfuel use of energy -- this largely explains why *SEDR* estimate is much higher than MECS for LPG

In addition to the MECS consumption estimates representing a different year from the NEI, the *SEDR* consumption estimates are not strictly comparable to the NEI industrial fuel consumption estimates because the *SEDR* data represent total energy consumed by the industrial sector, including energy used as feedstocks (i.e., raw materials used in manufacturing). For energy sources that are extensively used as feedstocks, the *SEDR* values do not provide a meaningful comparison to the NEI values for industrial boiler/IC engine SCCs. There are also a number of fuel types that are reported in the NEI that are not available from the MECS or *SEDR*. Of particular importance is the "Process Gas" NEI category. The NEI reports industrial sector combustion of nearly 13 trillion cubic feet of process gas in 1999. The draft NEI Version 2.0 reported one plant with process gas consumption of nearly 11 trillion cubic feet. This outlier was replaced with a more reasonable value when Pechan implemented the emissions activity QA/estimation decision rules described in the following section.

Emissions Activity QA/Estimation Decision Rules

Pechan relied upon its NEI experience in developing a set of methods for estimating missing activity and control data. While testing these methods on industrial natural gas source categories, Pechan identified two cases which call for the replacement of NEI reported activity and control data with default values: (1) values that are associated with units that are not valid for the source category; and (2) values that appear unreasonable when compared with information from other sources. An example of the former case is the use of acre-year units to describe the amount of natural gas consumption. An example of the latter case is an NEI record whose emissions and activity data represent an implicit EF (calculated by dividing emissions by throughput) that is 10 times the EF EPA recommends for the source category. Pechan, therefore, developed a set of decision rules to both estimate missing NEI emissions activity and control data and identify and replace outlier NEI data with more reasonable values.

Pechan first extracted the 1999 NEI industrial boiler/IC engine category values for the NEI fields identified in Table 2. Pechan next removed NEI records that reported emissions as zero (it is assumed that these records are associated with emission units that were not operating in 1999, and therefore, have no throughput associated with them). For NEI records that report activity and EFs in units other than the standard units for each SCC (e.g., reporting coal consumption in pounds rather than tons), Pechan converted the activity and EF values into values representing the standard units. In addition, the initial analyses were conducted on uncontrolled emission records because these records eliminate the possibility of invalid/outlier NEI control information to enter the procedure. In general, the methods defer to use of the NEI data. As described below, however, the methods also are used to replace some suspect NEI data with more reasonable values.

Table 2. NEI fields of interest.

Table	Field(s)
EP	State_FIPS, County_FIPS, Site_ID, Emission_Unit_ID, Process_ID, SIC_Unit_Level, SCC, Ash_Content, Sulfur_Content
SI	SIC_Primary
PE	Actual_Throughput, Throughput_Unit_Numerator
EM	Pollutant_Code, Emission_Ton_Value, Emission_Numeric_Value, Emission_Unit_Numerator, Factor_Numeric_Value, Factor_Unit_Numerator, Factor_Unit_Denominator, Rule_Effectiveness, Control_Status
CE	Pollutant_Code, Primary_Pct_Control_Efficiency, Pct_Capture_Efficiency, Total_Capture_Cntrl_Efficiency, Primary_ Device_Type, Secondary_Device_Type, Control_System_Description, Third_Control_Device_Type, Fourth_Control_Device_Type

Figure 1 displays the "decision tree" of NEI emissions activity QA and estimation procedures for industrial boiler/IC engine source categories. Due to constraints on the length of conference papers, it is not possible to describe each step in detail (readers interested in additional details, can obtain additional information on the procedures from the primary author). Because NEI industrial boiler/IC engine source category records generally report emissions for multiple pollutants, it was necessary to prioritize the selection of pollutants for which the QA/estimation procedures would be implemented. Pechan selected pollutants for each source category based on the number of records with throughput in standard units (i.e., for each category, Pechan first selected the pollutant with the largest number of such records).

Pechan initially placed NEI records into one of three categories: (1) **Uncontrolled**, which included records with control efficiencies reported as null or zero and control device codes reported as null or uncontrolled (i.e., 000); (2) **Controlled**, which included NEI records with control efficiency and control device code information not categorized as "Suspect"; and (3) **Suspect**, which included NEI records identified with a control device that was presumed not to control the pollutant with which it was associated in the NEI. There were two methods used to identify invalid control device/pollutant combinations. The first method was based on a review of the control device/pollutant combinations for which at least 75 percent of NEI records did not report a control efficiency value. Each of these control device/pollutant combinations was then compared to the controls identified in EPA's ControlNET data base,[5] *AP-42,*[6,7] and a point source emission control report prepared for the California Air Resources Board (ARB).[8] Combinations that did not appear in any of these sources were considered to be invalid. The second method focused on control devices that, while predominantly listed in the NEI as controlling one pollutant, were much more rarely reported as controlling additional pollutants. These rare control device/pollutant combinations were considered invalid when the NEI reported all pollutants controlled by a given control device, but did not report any control efficiencies for these combinations.

Additional QA procedures were then conducted on the initial set of Uncontrolled records for potential reclassification into the Suspect category. The following records were reclassified as Suspect: (a) records with invalid throughput values (i.e., throughput reported as null or in units that could not be converted to the standard units for the SCC) AND not reporting EFs or with invalid EFs (i.e., EFs reported in non-standard units or with a EF value not within the SCC's assumed lower and upper bounds); and (b) records with valid throughput data for which the calculated implicit EF is not within the SCC's assumed lower and upper EF bounds. To determine acceptable lower and upper bound uncontrolled EF values for each SCC/pollutant combination, Pechan utilized the emission factor quality rating reported in EPA's Factor Information Retrieval (FIRE) data system.[9] For those EFs identified with a quality rating of "A" or "B, "Pechan computed lower and upper bound uncontrolled EFs by multiplying the FIRE EF by 0.75 and 1.25, respectively. The resulting lower/upper bound EFs reflect the assumption that uncontrolled EFs should be within 25 percent of the value reported in FIRE. A 50 percent assumption was used for EFs identified with a quality rating of "C" or "D," and a 75 percent

Figure 1. Decision Tree of Emissions Activity QA/Estimation Procedures.

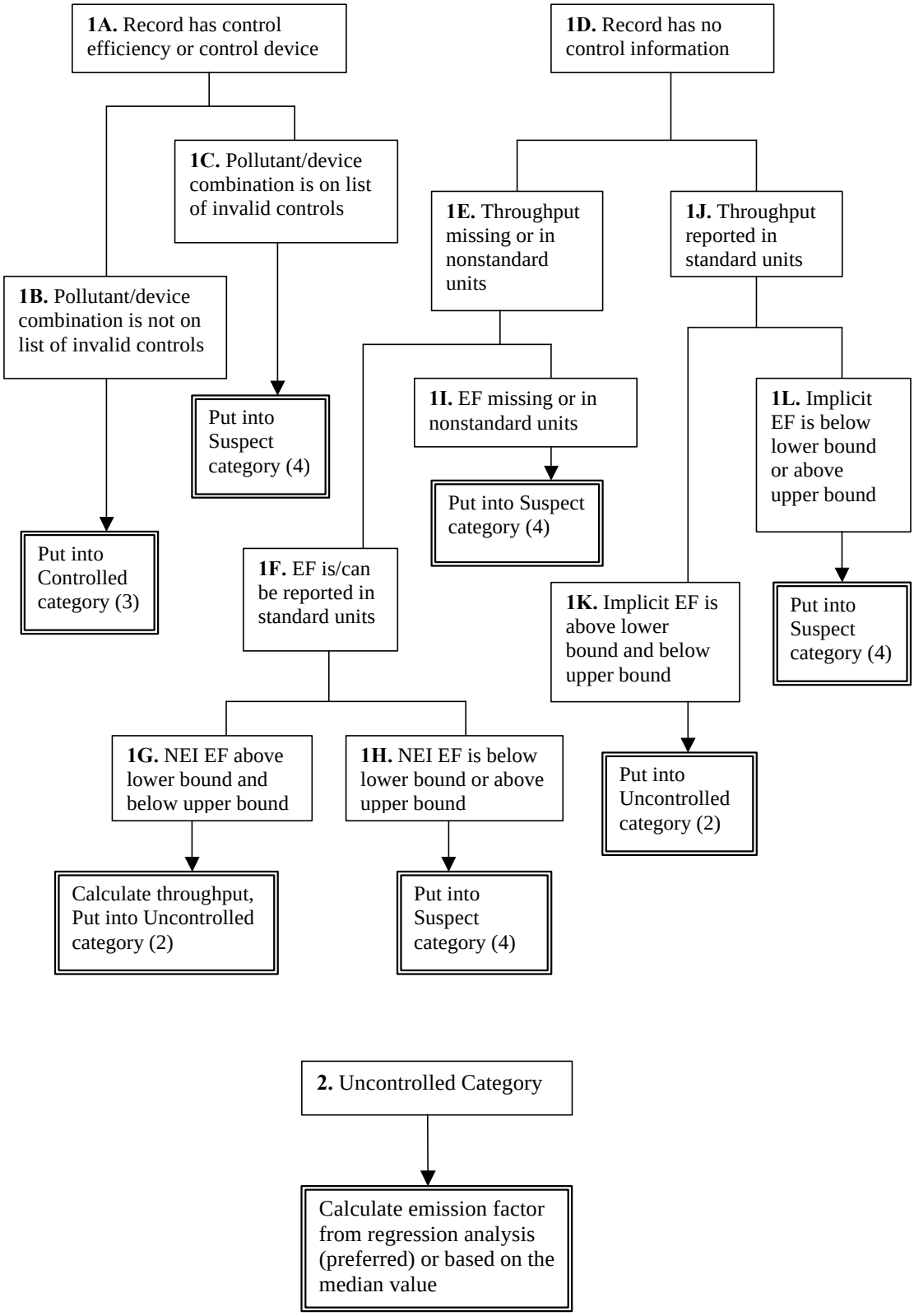


Figure 1. (continued)

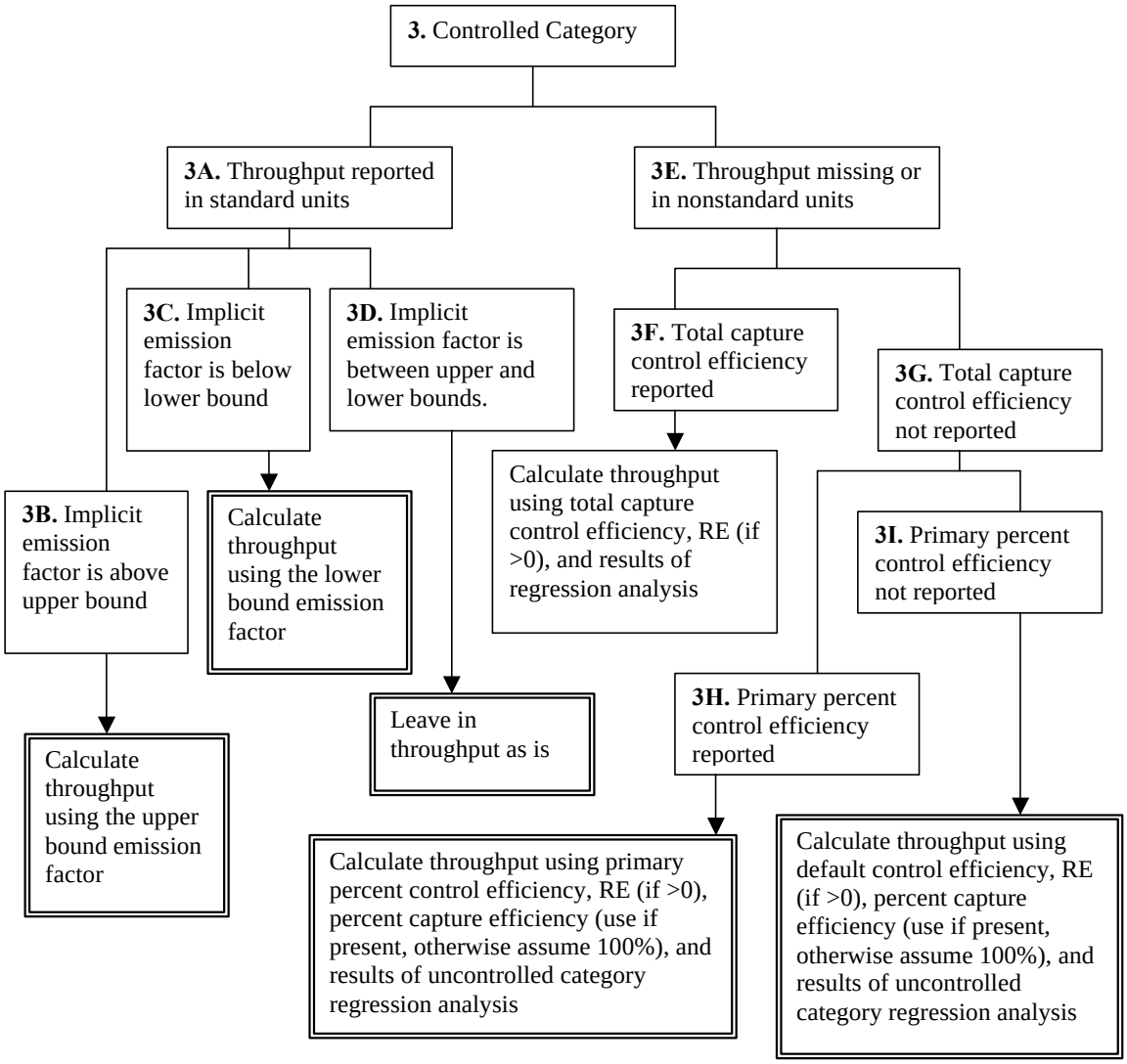


Figure 1. (continued)

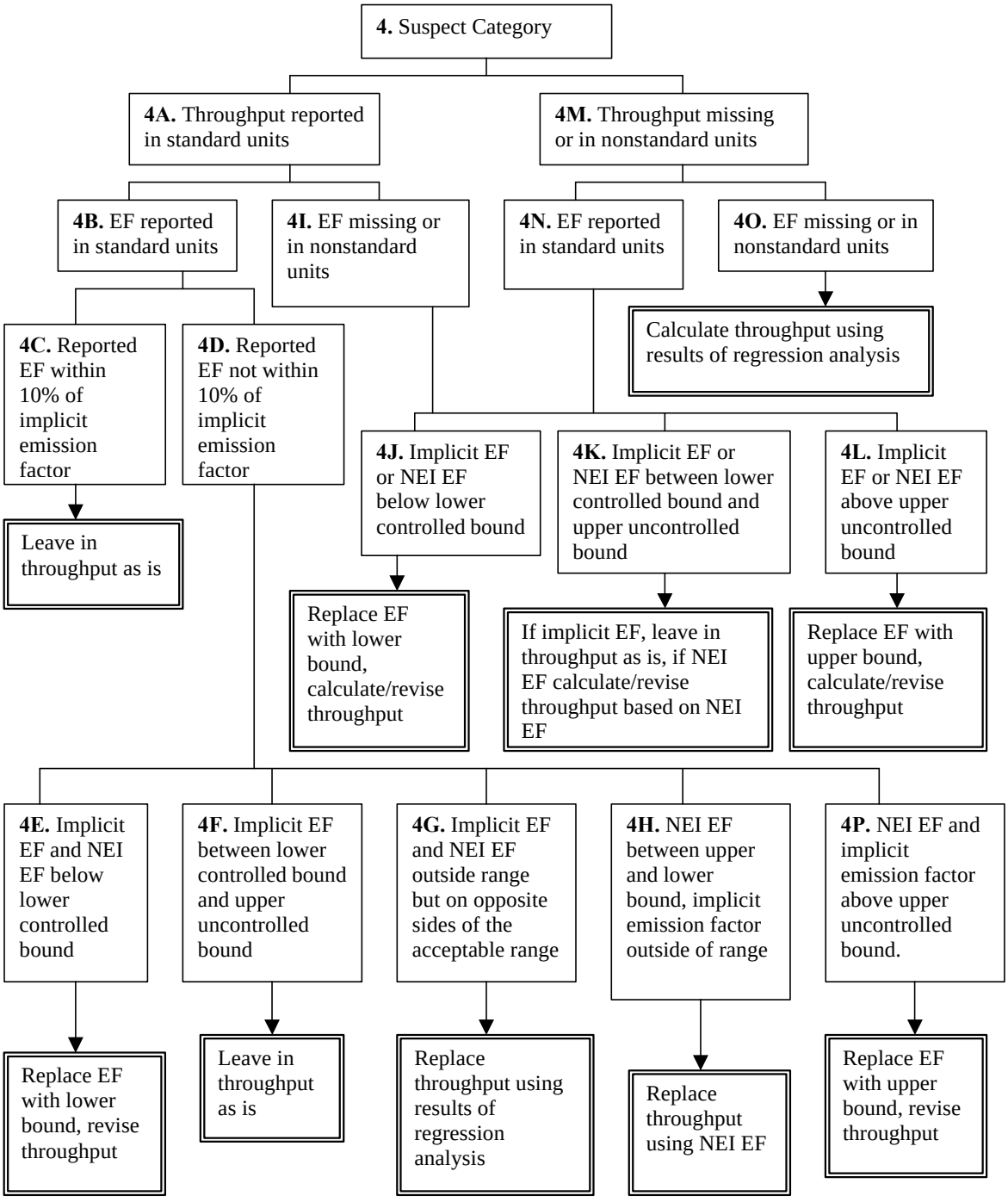
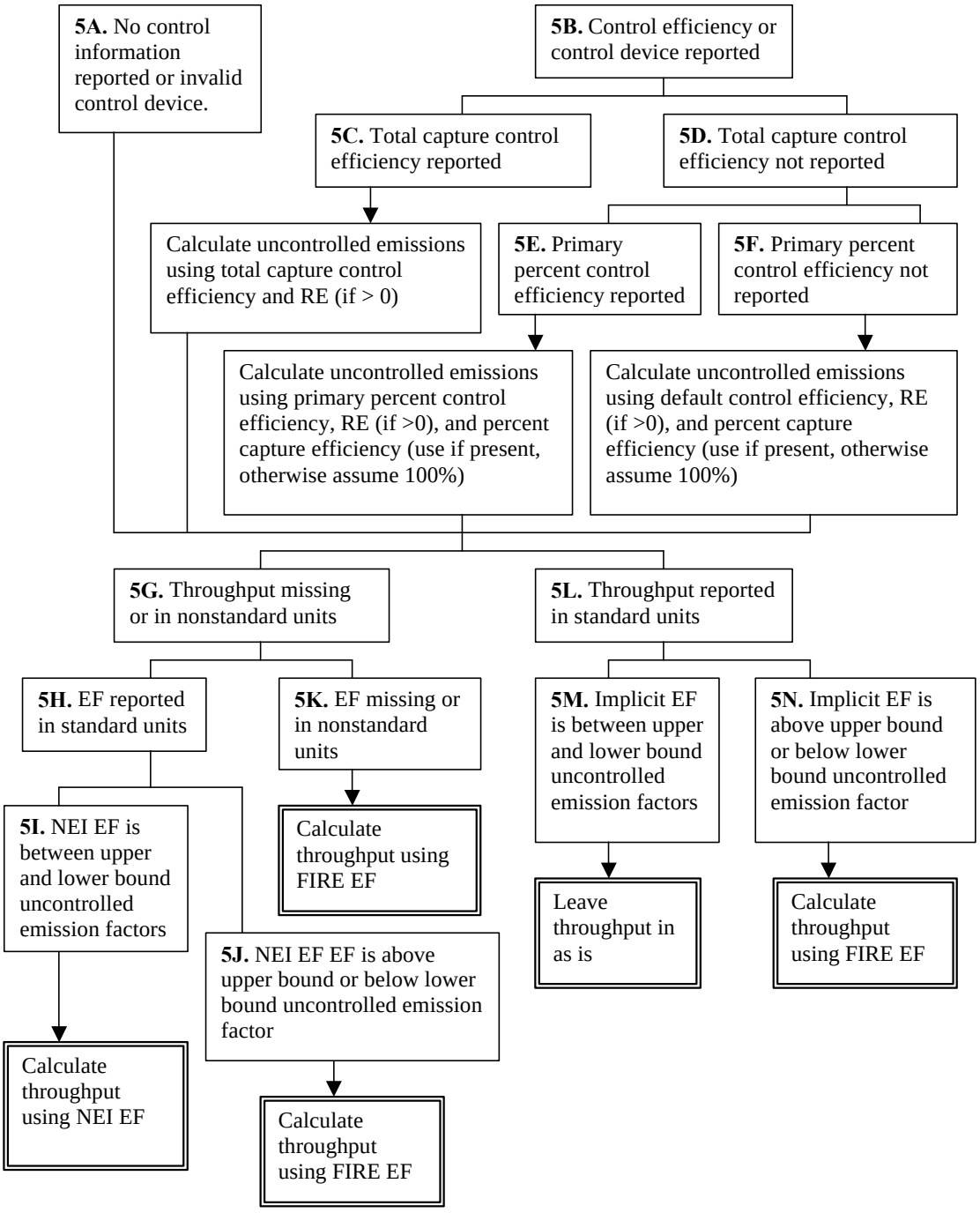


Figure 1. (continued)

Simplified approach used for remaining emission units



assumption was used for EFs identified with quality ratings below "D," as well as all EFs without a quality rating.

All remaining **Uncontrolled** records were used to calculate an average EF for each SCC. For SCCs with sufficient records for conducting a regression analysis, Pechan regressed uncontrolled emissions against fuel throughput. As indicated by Table 3, the regressions identified a strong relationship between the two variables. For these SCCs, Pechan developed the EF from the slope of the regression equation. For other SCCs, Pechan assigned EFs from the following sources in priority order: (1) EF reported in the EPA's Factor Information REtrieval (FIRE); (2) median EF value reported in the NEI; and (3) EF obtained using one of the above methods for a similar SCC.

Table 3. Results of emission factor regression analyses.

r^2 Value	Number of EF Equations
< 0.90	4
0.90 to <0.95	7
0.95 to <0.99	29
>0.99	101

For **Controlled** records for which throughput data were not reported in the NEI, Pechan estimated throughput based on uncontrolled emissions. Pechan used one of the following methods to calculate uncontrolled emissions and throughput for these records:

- For NEI records reporting total capture and control efficiency (and, if reported, rule effectiveness [RE], if not, 100 percent was assumed), Pechan used the reported values to calculate uncontrolled emissions and then estimated throughput by dividing uncontrolled emissions by the EF computed from the Uncontrolled records as described above;
- For NEI records that did not report total capture and control efficiency, but did report a primary control device efficiency, Pechan calculated uncontrolled emissions using this value, the percent capture efficiency (if reported, if not, 100 percent was assumed), and the NEI RE value (if reported, if not, 100 percent was assumed). Throughput was then estimated from uncontrolled emissions using the methods described above; and
- For NEI records for which control efficiency values were not reported, Pechan calculated uncontrolled emissions using a default control efficiency, the capture efficiency (if reported in the NEI, if not, 100 percent was assumed), and the RE value (if reported in the NEI, if not, 100 percent was assumed). Throughput was then estimated from uncontrolled emissions using the methods described above. Default control efficiencies were identified from a review of the efficiencies reported in ControlNET,[5] *AP-42,*[6,7] and a point source emission control report prepared for the ARB.[8] In cases where a control efficiency was not available from these sources, Pechan used the average of the NEI reported efficiencies.

For **Controlled** records reporting throughput data, Pechan reviewed the data for reasonableness by calculating uncontrolled implicit EFs and comparing these values to upper and lower bound uncontrolled EFs. The following three cases resulted:

1) For records with an implied EF above the upper bound uncontrolled EF, Pechan recalculated throughput using the NEI controlled emissions, the appropriate control efficiency, and the upper bound uncontrolled EF;
2) For records with an implied EF below the lower bound uncontrolled EF, Pechan recalculated throughput using the NEI controlled emissions, the appropriate control efficiency, and the lower bound uncontrolled EF; and

3) For records with an implied EF between the lower and upper bound uncontrolled EF, Pechan made no changes to the NEI throughput.

For **Suspect** records reporting throughput data, Pechan conducted additional tests on the reasonableness of the reported throughput. Pechan considered the NEI reported data to be valid if any of the following three conditions were met:

1) NEI implicit EF is within 10 percent of the NEI reported EF;
2) NEI implicit EF is not within 10 percent of NEI reported EF, and the implicit EF is between the lower bound controlled EF and the upper bound uncontrolled EF (to develop lower bound controlled EFs, Pechan used maximum control efficiencies for each pollutant based on values reported in the literature [e.g., CO = 99 percent]); and
3) NEI implicit EF is between lower bound controlled EF and upper bound uncontrolled EF and the NEI EF is missing or in units that could not converted to standard units.

For **Suspect** records with throughput data reported in non-standard units, or not reported at all, Pechan estimated throughput using one of three methods: (1) from NEI emissions and NEI EF if the NEI EF value was between the lower bound controlled EF and the upper bound uncontrolled EF; (2) from NEI emissions and the lower bound controlled EF if the NEI EF value was below the lower bound controlled EF; or (3) from NEI emissions and the upper bound uncontrolled EF if the NEI EF was above the upper bound uncontrolled EF.

For **Suspect** records for which NEI throughput is/can be reported in standard units, the NEI EF is/can be reported in standard units, and the NEI EF is not within 10 percent of the implicit EF:

1) Pechan replaced the NEI throughput with a value calculated from NEI emissions and the uncontrolled EF for records with both the implicit EF and the NEI EF outside the range of acceptable lower bound controlled and upper bound uncontrolled EFs;
2) Pechan replaced the NEI throughput with a value calculated by dividing NEI emissions by the lower bound controlled EF for NEI records with both the implicit EF and the NEI EF below the lower bound controlled EF;
3) Pechan replaced the NEI throughput with a value calculated by dividing NEI emissions by the upper bound uncontrolled EF for NEI records with both the implicit EF and the NEI EF above the upper bound uncontrolled EF; and
4) Pechan replaced the NEI throughput value with a value calculated by dividing NEI emissions by the NEI EF for NEI records with the implicit EF outside the range of reasonable EFs, and the NEI EF within the range of acceptable EFs.

For **Suspect** records for which NEI throughput is/can be reported in standard units and the NEI EF is missing or reported in units that can not be converted to standard units, Pechan replaced the NEI throughput with a calculated value. For records with an implicit EF below the lower bound controlled EF, Pechan replaced NEI throughput by dividing NEI emissions by the lower bound controlled EF. For records with an implicit EF above the upper bound uncontrolled EF, Pechan replaced NEI throughput with the value calculated by dividing NEI emissions by the upper bound uncontrolled EF.

Due to resource constraints, it was necessary to implement a simplified set of emissions activity QA/estimation procedures for a small proportion of total emission unit/pollutant combinations. These procedures are displayed on the final page of Figure 1.

RESULTS

For most fuel types, the emission activity QA and estimation procedures resulted in much higher fuel consumption estimates than the original NEI reported estimates (see Table 4). One major exception, however, is the process gas category, which declined from a total of nearly 13 trillion cubic feet in the pre-augmentation NEI to less than 500 billion cubic feet. This dramatic decline was due to the replacement of a draft NEI process gas consumption estimate of nearly 11 trillion cubic feet for one emission process, with a more reasonable consumption estimate. The consumption increases that

typically took place were anticipated because of the fact that most industrial boiler/IC engine records in the NEI do not report throughput values. Of particular note, however, is the uncharacteristically large increase in estimated residual and distillate oil consumption. For residual oil, for example, the total consumption estimate rose from 139 million barrels to 3,618 million barrels. When placed on a comparable SIC code basis with the *SEDR* and MECS estimates, the NEI-based residual and distillate oil estimates are more than 18 times the estimates reported by the EIA.

Table 4. Comparison of post-augmentation 1999 NEI, 1999 *SEDR*, and 1998 MECS national energy consumption estimates.

			SIC Codes 01 - 39		SIC Codes 20 - 39	
Fuel	**Units**	**1999 NEI Total**	**1999 NEI**	**1999 SEDR**	**1999 NEI**	**1998 MECS**
Natural Gas	billion cu ft	60,340	57,092	10,067	57,087	6,481
Coal and Coke	(million short tons)	85	62	96	62	78
Residual Oil	(million bbl)	3,618	3,541	40	3,541	57
Distillate Oil	(million bbl)	3,854	3,414	185	3,414	26
LPG	(million bbl)	138	10	624	10	38
Gasoline	(million bbl)	0.05	0.02	29	0.02	N/A

N/A - Not available.

Note: Unlike the MECS data, the *SEDR* data includes nonfuel use of energy -- this largely explains why *SEDR* estimate is much higher than MECS for LPG

To identify the reason for these dramatic increases, Pechan developed summaries of the throughput for each of these fuels by the estimation approach that was used to calculate throughput. The summaries indicated that the majority of total estimated throughput for both residual and distillate oil was computed using an estimation approach used for **Controlled** records. The specific approach was based on NEI records that reported control information, but did not report throughput. Pechan reviewed the NEI process records for which throughput were estimated using this approach. For residual oil, throughput for 230 processes were estimated using this approach. In all but one case, these processes were located in the State of Massachusetts. For distillate oil, throughput for 109 emission processes were estimated using this approach. In every case these processes were located in Massachusetts. In addition, about one-third of total distillate oil throughput was estimated based on the simplified approach identified in Box 5B of Figure 1 (i.e., the simplified approach used for **Controlled** records). The review of emission process records indicated that 39 of the 44 emission processes for which throughput was estimated using this approach were located in Massachusetts.

Because the NEI-based natural gas consumption estimates are considerably higher than the *SEDR* and MECS estimates, Pechan conducted a similar analysis of the natural gas throughput data. The vast majority of natural gas throughput was based on the simplified approach used for **Controlled** records (see Box 5B of Figure 1). Again, a review of emission process records indicated that 93 of the 136 processes for which throughput was estimated using this approach were located in Massachusetts.

These analyses point to the need for additional review of Massachusetts industrial boiler/IC engine point source data. It appears that the NEI emission and/or control information is incorrect for many or all Massachusetts industrial boiler/IC engine records.

Because of the concerns with the accuracy of Massachusetts NEI data, Pechan developed summaries of post-augmentation NEI throughput that excluded all Massachusetts records. These summaries are displayed in Table 5. These summaries indicate that the NEI estimates are much closer to the 1999 *SEDR* and 1998 MECS estimates than the Table 4 NEI estimates, and are more in-line with the EIA estimates than the original NEI data (see Table 1). In a few cases, however, the NEI-based estimates are still considerably higher than the estimates reported by the EIA. Further QA review of the NEI data would likely result in the identification of additional emission/control data-reporting anomalies

that are responsible for these discrepancies. It is suggested that these efforts initially focus on the control information reported in the NEI.

STUDY LIMITATIONS/FUTURE RESEARCH

The emission activity and control information estimates developed in this study are only as accurate as the underlying data on which they are based. Pechan relied on information from a combination of data sources in this study, including emission factors and control information from the NEI, EPA's FIRE data system, and the ControlNET data base. This study has identified a number of anomalies with key data that are reported to the NEI. These anomalies generally fall into one of the following categories:

Table 5. Comparison of post-augmentation 1999 NEI (excluding Massachusetts), 1999 *SEDR,* and 1998 MECS national energy consumption estimates.

			SIC Codes 01 - 39		SIC Codes 20 - 39	
Fuel	**Units**	**1999 NEI Total**	**1999 NEI**	**1999 SEDR**	**1999 NEI**	**1998 MECS**
Natural Gas	billion cu ft	13,023	10,210	10,067	8,467	6,481
Coal and Coke	(million short tons)	85	62	96	59	78
Residual Oil	(million bbl)	124	107	40	107	57
Distillate Oil	(million bbl)	187	145	185	137	26
LPG	(million bbl)	136	8	624	8	38
Gasoline	(million bbl)	0.05	0.02	29	0.02	N/A

N/A - Not available

Note: Unlike the MECS data, the *SEDR* data includes nonfuel use of energy -- this largely explains why *SEDR* estimate is much higher than MECS for LPG

- control devices that are designed to control a single pollutant that are identified as controlling every pollutant reported for a given emission process (e.g., gravity collectors identified as controlling CO emissions);
- EF and throughput values that are reported using inappropriate units (e.g., reporting natural gas throughput in acre-years);
- NEI reported EF values that differ significantly from the implicit EF calculated from the record's throughput and uncontrolled emissions;
- extraordinarily large throughput values (e.g., a single plant reporting nearly 12 trillion cubic feet of process gas combustion);
- control devices that are reported with atypically high control efficiencies; and
- records whose emissions and throughput values appear to indicate that emissions are controlled although they do not report any control information–this appears to be of particular concern for NO_x sources that may, for example, be of low-NO_x design, although no controls are associated with the source in the NEI.

Other limitations of this study include the general lack of information in the NEI (e.g., missing ash and sulfur content information for use with EFs that are based on these values), a lack of default EFs in the literature for certain source categories (e.g., requiring the use of factors from other similar SCCs), and the use of EF thresholds for identifying suspect NEI values–these thresholds were determined based on arbitrary upper and lower bound percentages applied to the EFs in FIRE.

This initial effort to estimate missing emissions activity and control information from the NEI highlights the importance of the accuracy of the NEI data that are reported in the estimation process. The NEI is a comprehensive air pollutant emissions inventory prepared with emission inventories that S/L/T agencies submit to EPA. With completion of the final NEI Version 2.0, the EPA is turning its attention to the development of Version 3.0 of the 1999 NEI, which will incorporate S/L/T inventories

not included in Version 2.0. The information obtained from this and any other studies that begin to utilize the NEI should inform future efforts for diagnosing QA issues with the S/L/T inventory submittals. It is hoped that this effort will inform the emission inventory community about the importance of the emission activity and control information reported to the NEI.

There are two major areas identified for future related research. The first area pertains to refinement of the emissions activity estimation procedures. Specific efforts in this area would include comparisons of the uncontrolled EFs developed in this study with those reported in FIRE. Investigations of any significant discrepancies may help pinpoint potential estimation procedure improvements. In addition, similar emissions activity estimation studies should be undertaken for additional point source categories of interest. These studies would provide insights into whether these procedures can be easily transferred to other portions of the NEI and assist in refining the existing methods.

A second major area for future research involves the development and implementation of additional NEI QA procedures. Specific efforts in this area would include the development of comprehensive lists of valid control device/pollutant combinations and SCC/throughput units and SCC/emission factor units combinations. The control device/pollutant combination list would facilitate identification of invalid control information that would assist in improving both the emission activity data developed from these estimation techniques as well as the QA procedures that are used to review S/L/T inventory submittals. Comprehensive lists of valid throughput units and emission factor units by SCC would serve a similar role.

REFERENCES

1. Strait, R.P., R. Huntley, S. Boone, D. McKenzie, and S. Connor, "Incorporation of State, Local, and Tribal Emission Inventory Data into the 1999 National Emission Inventory", Presented at the 11th Annual Emission Inventory Conference of the U.S. Environmental Protection Agency, Atlanta, GA, April 2002.

2. Energy Information Administration, *State Energy Data Report 1999: Consumption Estimates*, U.S. Department of Energy, Washington, DC, available from http://www.eia.doe.gov/emeu/sedr/contents.html, accessed September 2002.

3. Energy Information Administration, 1998 Manufacturing Energy Consumption Survey data tables, U.S. Department of Energy, Washington, DC, available from http://www.eia.doe.gov/emeu/mecs/contents.html, accessed September 2002.

4. U.S. Environmental Protection Agency, “Criteria Inventory Documentation - *Final* 1999 NEI Version 2,” Office of Air Quality Planning and Standards, Research Triangle Park, NC, available from http://www.epa.gov/ttn/chief/net/index.html#final2, accessed September 2002.

5. E.H. Pechan & Associates, Inc. “Control Measure Evaluations: The Control Measure Data Base for the National Emission Trends Inventory (ControlNET), Draft Report,” prepared for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, September 1999.

6. U.S. Environmental Protection Agency, “Compilation of Air Pollutant Emission Factors - Volume 1: Stationary Point and Area Sources, Fifth Edition,” *AP-42*, Chapter 1, External Combustion Sources, Office of Air Quality Planning and Standards, Research Triangle Park, NC, September, 1998.

7. U.S. Environmental Protection Agency, “Compilation of Air Pollutant Emission Factors - Volume 1: Stationary Point and Area Sources, Fifth Edition,” *AP-42*, Chapter 3, Stationary Internal Combustion Sources, Office of Air Quality Planning and Standards, Research Triangle Park, NC, August 2000.

8. E.H. Pechan & Associates, Inc. “Identification of Point Source Emission Controls and Determination of Their Efficiencies and Costs,” Prepared for California Air Resources Board and the California Environmental Protection Agency, Sacramento, CA, April 30, 1998.

9. U.S. Environmental Protection Agency, “Factor Information REtrieval (FIRE) v.6.23,” Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, available from http://www.epa.gov/ttn/chief/software/fire/index.html, October 2000.

Key Words

Quality Assurance
Emission Activity
Emission Factor
Emission Control
Fuel Consumption