BNL-65965

ESTIMATING UNCERTAINTY OF EMISSIONS INVENTORIES: WHAT HAS BEEN DONE/WHAT NEEDS TO BE DONE

Carmen M. Benkovitz
Environmental Chemistry Division
Department of Applied Science
Brookhaven National Laboratory
Upton, NY 11973-5000

October 1998

Presented at the
IPCC Meeting on Managing Uncertainty in
National Greenhouse Gas Inventories,
Paris, France
October 13-15, 1998

This research was performed under the auspices of the U.S. Department of Energy under Contract No. DE-AC02-98CH10886.

BACKGROUND

Developing scientifically defensible quantitative estimates of the uncertainty of atmospheric emissions inventories has been a “gleam in researchers’ eyes” since atmospheric chemical transport and transformation models (CTMs) started to be used to study “air pollution”. Originally, the compilation of these inventories was done as part of the development and application of the models by researchers whose expertise usually did not include the “art” of emissions estimations. In general, the smaller the effort spent on compiling the inventories the more effort could be placed on the model development, application and analysis. Yet model results are intimately tied to the accuracy of the emissions data; no model, however accurately the atmospheric physical and chemical processes are represented, will give reliable representation of air concentrations if the emissions data are flawed.

In parallel with these scientific studies governments started to be concerned with the effects of anthropogenic emissions to the atmosphere on air quality and human health. Organizations such as the United States Environmental Protection Agency (USEPA) were given the power to collect information on these anthropogenic emissions. This information was collected in “emissions inventories”, and was used to draft and enforce legislation to limit the amounts emitted and thus protect both human health and environmental resources. Initial efforts were made to develop quantitative estimates of the uncertainty of emissions data; a few examples are the sensitivity analysis work of Ditto et al. [1973], and the work of PEDco Environmental Inc. [1978; 1974]. These and other initial efforts were of limited scope and the methodologies used were not entirely suitable.

In the 1970s large, multi organization scientific programs were started to rigorously study the effects of anthropogenic emissions on air quality. One example (and only one, there are many others) was the MultiState Atmospheric Power Production Pollution Study (MAP3S), funded by the U.S. Department of Energy, which brought together researchers from different National Laboratories and universities. The objective of this program was to study the impact of emissions from power generation in the United States. With this and other programs researchers stated to realize that the compilation of regional detailed emissions inventories is resource intensive, and realized that cooperation with those compiling inventories for legislative and enforcement purposes was not only desirable but necessary. Within the MAP3S program, we at Brookhaven National Laboratory (BNL) took this route, and started obtaining emissions data from the USEPA and processing it for input to the atmospheric models used by MAP3S researchers.

By the late 1970s researchers realized that: a) transport of emissions in the atmosphere does not recognize national borders, and b) research into the effects of atmospheric emissions requires a wide range of scientific expertise and adequate resources. Thus in 1980 legislation creating the National Acid Precipitation Assessment Program (NAPAP) was passed by the US Congress. This program was to last ten years and was to be carried out as a collaboration of multiple agencies within the government. Within this program cooperation with Canadian research efforts were also established, and so NAPAP became not only a multi agency but a multinational program. NAPAP research was coordinated by approximately ten Task Groups; each group addressed different aspects of the problem and was composed of experts from several organizations.

NAPAP formalized the use of emissions inventories compiled for legislative mandate in scientific research. Task Group I, Emissions and Controls, brought together the USEPA and the US Department of Energy (USDOE), two agencies in charge of compiling emissions data in the U.S. This group directed the research to prepare emissions inventories for use with the Regional Acid Deposition Model (RADM), which was developed by other Task Groups within NAPAP. NAPAP attempted to generate quantitative estimates of the uncertainty of emissions, but the emphasis placed and the resources available for this task were not sufficient to allow the detail studies needed to develop scientifically defensive estimates.

The European community was also addressing the problems of compiling emissions inventories. In comparison with the U.S. and Canada, two large bordering countries with federal governments and a common language, the work in Europe involved a multiplicity of nations, languages and governing philosophies. Initial estimates of emissions in Europe were developed within each country, but researchers quickly came to realize that a common approach was necessary if the issues of transboundary air pollution were to be properly addressed. As with the US and Canada, the immediate problem was to develop the emissions estimates, and evaluation of the uncertainty of this information was not of primary importance. The initial program to develop harmonized inventories in Europe was CORINAIR, which developed inventories for western European countries for 1985 base year. Some work on the estimation of the uncertainty of national inventories has been done, but the methodologies used were not applicable to all inventories. The compilation of harmonized inventories for Europe has been further formalized under the European Environment Agency, and the scope of the inventories expanded to include most European nations so that inventories for the 1990 base year are much more complete.

In this article I will briefly summarize some of the work done to develop quantitative estimates of the uncertainty of emissions inventories. The summary is not meant to be extensive; my aim is to acquaint the reader with some of the previous work and some of the strengths and weaknesses. Based on our experiences at BNL in the development and use of emissions inventories, I will then present what we currently believe is needed if we are to develop scientifically defensible quantitative estimates of the uncertainties of emissions data.

REVIEW

Here I will review the general methodology used to develop estimates of atmospheric emissions.

Emissions from a source are estimated as:

$$E_i = \left(\overline{EF}\right)_c \times A_i \times P_1 \times P_2 \times \ .....$$

Where:

E_i = emissions from source i of category c, $\left(\overline{EF}\right)_c$ emission factor for category c
A_i = activity rate(ex: fuel use), and
P_1, P_2, = additional parameters needed (ex: % S in fuel).

To calculate total emissions by source type, by country, etc.:

$$E_{tot} = \sum_i E_i$$

The most important characteristics of emissions parameters pertinent to uncertainty estimation are:

- Probability distribution of parameters is not known.
- Parameters may not be independent (ex: same emission factor used within each source in category).
- Measurement data used to evaluate parameters are usually not available.

These characteristics must be taken into account when selecting the statistical methodologies to be used in estimating emissions uncertainty.

WHAT HAS BEEN DONE

To date projects compiling inventories of atmospheric emissions of trace gases have spent from 99.9% to 100% of their resources in developing better methodologies and data to estimate emissions. The uncertainties associated with these estimates have either been ignored altogether or have been addressed in a cursory manner. The latter generally used existing statistical methodologies even if the emissions parameters did not fit the criteria for the use of such methodologies. Thus no scientifically defendable quantitative estimates of the uncertainty of emissions inventories are currently available.

The following is a list of the major methodologies that have been used to estimate the “uncertainty” of emissions inventories. This is not an all inclusive list but it gives a “flavor” of the work to date.

Qualitative methods

- Expert opinion determines the classification. Examples of this are the AP-42 emission factor rating and the GEIA anthropogenic inventories of SO_2 and NO_x for base year 1985.

Semi-quantitative methods

- Data attribute rating. This technique is designed to rank inventories on the basis of attributes that affect the accuracy, appropriateness, and reliability of an estimate.

- Expert opinion. An example of the use of expert opinion to obtain "quantitative" estimates of emissions uncertainty can be found in the paper by Dickson et al. [1991]. Delphi techniques were used to gather subjective estimates regarding the uncertainty of emissions data; these subjective estimates were then translated to quantitative estimates of uncertainty using several methodologies such as the Lognormal Method, the Probability Method, Error Propagation Method. In general, these methods do not account for covariance between emissions estimation parameters, use of averaged parameters, etc.

Quantitative methods

- Standard statistical methods for error propagation. These methods will not give reliable estimates of the uncertainty of emissions because the parameters used to calculate emissions fail several points of the criteria for applying these methods (normal distribution, independence of parameters).

- Numerical simulations. Numerical simulations are now within easy reach due to the increase in power and decrease in price of personal computers. Monte Carlo simulations and bootstrap simulations are two numerical simulation techniques. Examples of these studies are given in the work of Frey and Rhodes [1998] and Frey [1996]. However, these techniques are not directly applicable to estimate the uncertainty of emissions inventories because:

  - Most require that the probability distribution of parameters be known.
  - Many require independence between parameters.
  - Some require point estimates for certain inputs.

Methodology Studies

Here I will summarize two studies conducted at BNL mainly because I participated in this work so I am familiar with the methods and results. There are probably similar studies in the published literature which should come to the forefront of any intensive literature search. I would appreciate being referred to any past or current studies of this type (e-mail: cmb@bnl.gov).

The BNL studies started to test the impact of a few of the characteristics of emissions parameters in the estimation of the uncertainty of the emissions. Mathematical statistics were used to:

- Study the statistical implications of the autocorrelation and covariance of emissions parameters. This study addressed the impact of the autocorrelation and the covariance of emissions parameters on the calculation of the uncertainty of emissions estimates. Conclusions state:

- Autocorrelation of parameters has little effect on average bias but covariance of parameters has large effects on average bias.

- The Mean Square Error (MSE) is dominated by the bias when covariance of the parameters is > 0.2.

· Study the statistical implications of using averaged parameters (ex: emission factors) for emissions calculations. This study addressed the impact of using averaged parameters on the calculation of the uncertainty of emissions estimates. Conclusions state:

- If sources in a category are homogeneous, using average emission factors *may* result in slightly better estimate of emissions from individual source than using source-specific emission factors.

- Errors in individual source estimates using average emissions factors cumulate in the sum of emissions.

This study also developed an analytical expression to estimate MSE of emissions sums when the individual emissions are the product of two parameters.

At the conclusion of this work our project was redirected to other tasks.

Model Runs to Evaluate Inventories

In these studies monitored air concentrations are used in conjunction with inverse modeling techniques to estimate emissions. These emissions are then compared with the emissions given by the inventory under evaluation. An example of these studies is the work of Veltkamp et al. [1995]; there are several other examples in the literature. The reasons for the differences of the emissions estimated by these models and the emissions in the inventories used can be tricky to elucidate. Are the differences caused by the inappropriateness of the parameterizations in the inverse models? because the monitoring data are not truly representative of the time and locations being modeled? because of the errors in the emissions inventories used? In any case, results of these studies do not generally provide quantitative evaluations of the uncertainty of the emissions data.

WHAT NEEDS TO BE DONE

The importance of having scientifically defensible quantitative estimates of the uncertainty of emissions inventories has now become critical. International treaties call for emissions reductions of a certain magnitude, and enforcement of such treaties, possible emissions trading, etc. need to be based on inventories of known and accepted quality. Most of the stipulations of these treaties will be phased in over a period of time, so is time to *bite the bullet*; i.e., the estimation of the uncertainty of emissions inventories needs to be **SERIOUSLY** addressed **NOW**.

My recommendation is that

The estimation of the uncertainty of emissions inventories should be defined and funded as an independent research project.

This project must be implemented as a collaboration between emissions inventory experts and statisticians. Main responsibility for the development of the project should fall on a “core” emissions expert/s and a “core” statistician/s. The emissions expert/s would serve as link to and synthesize the opinions of the larger community of emissions experts, and with the help of the larger community of statisticians the core statistician/s would be responsible for the selection and/or development of the appropriate methodologies to be used in estimating the uncertainty of emissions inventories. Collaborators in this project **must** come from the **international community**.

Next I present some **VERY PRELIMINARY** thoughts and some of the tasks I envision for this proposed project. The first issue to be addressed is the development of strict definitions of terms: uncertainty, variability, etc. There are several sources for definitions of these and other key terms, for example the International Organization of Standardization, IPCC, and USEPA; these would serve as a starting point or selected if appropriate.

The next few tasks would be iterative. One task would address the development of minimum criteria for emissions parameters that a) are followed by or could be followed by existing or new data, and b) allow calculation of emissions uncertainty. This work would proceed in parallel with another task addressing the selection and/or development of statistical methodology that, when combined with the developed criteria, would allow quantitative estimation the uncertainty of emissions inventories. The next task would address the selection of an existing or the development of a subset of emissions estimation parameters to be used as testbed for the developing methodologies, to be followed by the estimation of the uncertainty of the testbed cases. Iterations would continue until an applicable set of criteria and methodologies is developed.

After this initial phase the uncertainty calculations would be extended to other emissions data that satisfy the defined criteria. Results here might indicate that the criteria and /or the methodologies developed need modification; if so, the iterative process would continue. Recommendations for the work needed to bring emissions data that do not follow the establish criteria to “qualifying” status would be developed, and the uncertainty calculations extended to these new data.

It would be naive to expect that scientifically defensible, quantitative estimates of the uncertainties of all emissions data can be obtained with information currently available. However, it behooves us to make a commitment to develop a program which can address this problem directly and to provide needed resources on a long-term basis so that a) quantitative uncertainty estimates of current inventories are developed whenever possible and b) criteria are established to insure that the continuing development of emissions inventories includes the acquisition of the information needed for quantitative estimates of the uncertainties associated with these inventories.

REFERENCES

Baars, H.P., Accuracy of Emission Inventories: Methodology and Preliminary Results of the Dutch NO_x Inventory, TNO Division of Technology for Society, Delft, The Netherlands, 1990.

Beck, L.L., and D. Wilson, EPA's Data Attribute Rating System, in *Specialty Conference on Emissions Inventory: Planning for the Future*, pp. 176-189, Air and Waste Management Association, Pittsburgh, PA, 1998.

Benkovitz, C.M., MAP3S/RAINE Emissions Inventory Progress Report, Brookhaven National Laboratory, Upton, NY, 1981.

Benkovitz, C.M., Compilation of Source Emissions Data and Their Use in Air Quality Models, in *3rd Joint Conference on Applications of Air Pollution Meteorology,* American Meteorological Society, San Antonio, TX, 1982.

Benkovitz, C.M., and V.A. Evans, User Access to the MAP3S Source Emissions Inventory, Brookhaven National Laboratory, Upton, NY, 1981.

Benkovitz, C.M., and N.L. Oden, Individual Versus Averaged Estimates of Parameters Used in Large Scale Emissions Inventories, *Atmos. Environ.*, *23*, 903-909, 1989.

Benkovitz, C.M., M.T. Scholtz, J. Pacyna, L. Tarrasón, J. Dignon, E.V. Voldner, P.A. Spiro, J.A. Logan, and T.E. Graedel, Global Gridded Inventories of Anthropogenic Emissions of Sulfur and Nitrogen, *J. Geophys. Res.*, *101* (D22), 29239-29253, 1996.

Bouscaren, R., CORINAIR Technical Annexes Volume 2: Default Emission Factors Handbook, pp. 327, European Commission, Brussels, Belgium, 1995.

Bouscaren, R., and M.-H. Cornaert, CORINAIR Technical Annexes Volumes 1: Nomenclature and Software, pp. 126, European Commission, Brussels, Belgium, 1995.

Bouscaren, R., J.-P. Fontelle, M.-H. Cornaert, J. Maes, E.d. Roeck, and B. Lübkert, CORINAIR Inventaire des Emissions de Dioxyde de Soufre, D'Oxides D'Azote et de Composés Organiques Volatils dans la Communauté Européenne en 1985, pp. 205, Commission Européenne, Bruxelles, Belgique, 1995.

Dickson, R.J., V.M. Sadeghi, and W.R. Oliver, Uncertainty Estimation for Regional Emissions Inventories, in *Emissions Inventory Issues in the 1990s*, pp. 79-97, Air & Waste Management Association, Durham NC, 1991.

Ditto, F.H., L.T. Gutierrez, T.H. Lewis, and L.J. Rusbrook, Weighted Sensitivity Analysis of Emissions Data: Volume I and II, US Environmental Protection Agency, Research Triangle Park, NC, 1973.

Frey, H.C., Methods for Quantitative Analysis of Variability and Uncertainty in Hazardous Air Pollutant Emissions, in *Proceedings of the 91st Annual Meeting of the Air & Waste Management Association*, Air & Waste Management Association, Pittsburgh, PA, 1998.

Frey, H.C., and D.S. Rhodes, Characterizing, Simulating, and Analyzing Variability and Uncertainty: An Illustration of Methods Using an Air Toxics Emissions Example, *Human and Ecological Risk Assessment*, *2* (4), 762-797, 1996.

Hogg, R.V., and A.T. Craig, *Introduction to Mathematical Statistics*, 415 pp., Macmillan Publishing Co., New York, NY, 1970.

Mendenhall, W., R.L. Scheaffer, and D.D. Wackerly, *Mathematical Statistics with Applications*, 686 pp., Duxbury Press, Boston, MA, 1981.

Oden, N.L., and C.M. Benkovitz, Statistical Implications of the Lack of Independence in the Parameters Used for Calculations of Large Scale Emissions Inventories, *Atmos. Environ.*, *24A*, 449-456, 1990.

PEDco Environmental Inc., Development of an Emission Inventory Quality Assurance Program, pp. 46, US Environmental Protection Agency, Research Triangle Park, NC, 1978.

PEDCo Environmental Specialists Inc., Source Inventory and Emission Factor Analysis, Volumes I and II, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1974.

Saeger, M., J. Langstaff, R. Walters, L. Modica, D. Zimmerman, D. Fratt, D. Dulleba, R. Ryan, J. Demmy, W. Tax, D. Sprague, D. Mudgett, and A.S. Werner, The 1985 NAPAP Emissions Inventory (Version 2): Development of the Annual Data and Modelers' Tape, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1989.

U.S. Environmental Protection Agency, Compilation of Air Pollutant Emission Factors, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1985.

Veltkamp, A.C., R. Eisma, A.T. Vermeulen, W.M. Kieskamp, W.C.M.V.d. Bulk, B. Beemsterboer, O. Zwaagstra, J.J. Möls, A. Hensen, and G.P. Wyers, Validation of Methane Source Strengths, pp. 1-82, Netherlands Energy Research Foundation, Petten, The Netherlands, 1995.

Wagner, J.K., and J.E. Langstaff, Estimation of Uncertainty for the 1980 NAPAP Emissions Inventory, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1986.