United States
Environmental Protection
Agency

Air and Energy Engineering
Research Laboratory
Research Triangle Park NC 27711

Research and Development

EPA/600/S7-86/055 Apr. 1987

EPA

Project Summary

Estimation of Uncertainty for the 1980 NAPAP Emissions Inventory

John E. Langstaff and Janice K. Wagner

The National Acid Precipitation Assessment Program (NAPAP) has charged its Task Group on Emissions and Controls with developing comprehensive and accurate inventories of natural and anthropogenic emissions of substances thought to be important in acid deposition processes. This responsibility includes quantifying the degree of uncertainty associated with those estimates. This report presents the methods, assumptions, and results of an effort to develop quantitative estimates of emissions uncertainties for the 1980 NAPAP Emissions Inventory. Conventional statistical concepts were applied to sample uncertainty data. The shortcomings of these data and questions concerning the proper application of statistical methods to the problem at hand were deferred for later study. Results of this study consist of lessons learned and problems identified, and fall into three categories: uncertainty relationships at various levels of emissions disaggregation, statistical methodology questions, and gaps in input data. The computed uncertainty values are illustrative in nature and included for completeness; these values may be used only to identify uncertainty trends and relationships between the various pollutants at increasing levels of aggregation and allocation.

This Project Summary was developed by EPA's Air and Energy Engineering Research Laboratory, Research Triangle Park, NC, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

The National Acid Precipitation Assessment Program (NAPAP) has charged its Task Group on Emissions and Controls with developing comprehensive and accurate inventories of natural and anthropogenic emissions of substances believed to be important in acid deposition processes. The Task Group is developing estimates of past, present, and future emissions with adequate geographic, temporal, and source resolution to support the research requirements of NAPAP. The Task Group is also responsible for the quantification of the uncertainty associated with those estimates.

This report presents the methods, assumptions, and results of a study to develop quantitative estimates of emissions uncertainties for the 1980 NAPAP Emissions Inventory. Conventional statistical concepts, adapted to the emissions inventory by Brookhaven National Laboratory in previous work, were applied to illustrative uncertainty estimates supplied by a team of emissions inventory developers. The shortcomings of these data and questions concerning the proper application of statistical methods to the problem at hand were deferred for later study.

Approach

The calculation of emissions uncertainty for individual sources and source aggregation is performed by a computer program. Annual emissions estimates for each process are allocated temporally, spatially, and by pollutant subspecies. This allocation is carried out by merging process- and pollutant-specific allocation

factors to the individual emission records in the file. After these allocation factors have been assigned, the system calculates emissions by pollutant for a specified averaging time, source classification, or geographical region. The emissions uncertainty is calculated similarly. Uncertainty estimates, expressed as percent error values, are assigned to the emission records on a process- and pollutant-specific basis. By use of the error propagation formulas for products and sums, it has been assumed that emissions uncertainty can be computed directly for the various levels of source aggregation.

The initial set of uncertainty input values was provided by a panel of emissions inventory developers who were assembled at a 1980 NAPAP Emissions Inventory Uncertainty Workshop in May 1985 at the Air and Energy Engineering Research Laboratory, U.S. EPA, Research Triangle Park, NC. Details of the results of this Workshop are included in report Appendix A. At the Workshop, the panel was asked to provide consensus estimates of the 90% confidence interval surrounding nominal mean values for generalized classes of the variables used to calculate emissions at various levels of species, temporal, and spatial resolution.

Although members of this panel were experienced in developing emissions inventories and the methods for estimating emissions, they could not be considered experts in the calculation of emissions and associated uncertainty for all source types They were generally not experienced with the actual emitting processes under consideration or with the typical variability of emissions from those processes

This study was intended as a first step toward quantifying uncertainty in the 1980 NAPAP Emissions Inventory. Obvious shortcomings in the data set and doubts about validity of the methods were deferred to later study. Key assumptions in the uncertainty calculations are:

- Emission calculation parameters are independent; i.e., they do not covary.
- Emission factors represent true mean values.
- All estimates are unbiased.
- The emission parameters can be treated as random variables which are approximately normally distributed.
- No coding or transcription errors are present.
- The data are complete; no emissions data, emissions sources, or emissions source categories are missing.

Results from the Workshop were later extended to other pollutants, and otherwise interpreted, by contractor personnel and the EPA Project Officer. These extensions were performed to illustrate uncertainty results for all emission estimates in the inventory. For these reasons, the results of the uncertainty calculations should be used only to identify uncertainty relationships between the various pollutants at increasing levels of aggregation and allocation.

Results

Based upon the uncertainty estimates of the panel, the contractor, and the EPA Project Officer, and the methodology described in report Section 3, illustrative emissions uncertainty values have been calculated for sulfur dioxide (SO_2), oxides of nitrogen (NO_x), speciated oxides of nitrogen (S-NO_x), carbon monoxide (CO), volatile organic compounds (VOC), speciated volatile organic compounds (S-VOC), total suspended particulates (TSP), speciated total suspended particulates (S-TSP), lead (Pb), ammonia (NH_3), hydrogen chloride (HCl), hydrogen fluoride (HF), and sulfate (SO_4). (See Tables 1 and 2.) The percent uncertainties show an expected decrease as aggregation level increases, and an increase in uncertainty with disaggregation to finer levels of detail. Note that these phenomena result inevitably from the combination of assumptions about the character of the uncertainty input data and the workings of the error propagation formulas on data of that character; the particular values provided by the panel are not important in this effect.

The values of uncertainty estimates for national levels of aggregation appear to be unreasonably small The emission inventory data do not warrant this degree of confidence. Therefore, the assumptions, data, and methodology must be reexamined for reasonableness. Preliminary indications are that the methodology must be expanded to incorporate other sources of error (e.g., bias, coding, and emission errors). Other elements of the uncertainty estimates are also counterintuitive and should be corrected by improved assumptions, data, and methodology Thus, the uncertainty estimates presented are only illustrative of the results achieved from applying the sample data set to the proposed methodology.

Despite the limitations noted above, this experiment represents a helpful step in learning how best to approach quantification of uncertainty in the NAPAP Emissions Inventories. Lessons learned here will guide further efforts toward an improved level of scientific validity and credibility. Investigation of the assumptions, methodology, and required input data is continuing.

Table 1. *Annual/Hourly Illustrative Uncertainty of the 1980 NAPAP Emissions Inventory*

	Percent Error[a]												
	Annual												
	SO_2	NO_x	SNO_x[b]	CO	VOC	SVOC[c]	TSP	STSP[d]	Pb	NH_3	HCl	HF	SO_4
Nation point & area	1	1	1	1	1	2	3	5	2	4	1	9	1
Region average[e] point & area	4	2	3	4	4	7	9	14	7	14	6	8	5
State average point & area	6	4	5	8	8	14	16	25	13	26	13	13	12
County Average point & area	16	16	19	29	32	54	73	110	54	60	57	34	62
point	24	27	32	31	36	79	30	72	33	74	24	19	34
area	16	16	19	30	33	56	79	120	56	60	94	100	73
Grid average point & area	25	18	21	30	31	52	74	110	45	50	25	20	34
area	20	19	22	33	37	60	82	120	59	63	97	110	75
SCC average point	39	45	54	45	52	120	47	120	88	100	99	99	89

	Percent Error[a]												
	Hourly												
	SO_2	NO_x	SNO_x[b]	CO	VOC	SVOC[c]	TSP	STSP[d]	Pb	NH_3	HCl	HF	SO_4
Nation point & area	2	1	2	3	2	3	4	6	4	5	4	13	3
Region average[e] point & area	8	4	5	9	6	9	12	17	13	19	19	21	9
State average point & area	14	8	9	17	12	18	22	31	21	35	34	34	21
County Average point & area	32	29	31	51	48	68	100	140	82	82	94	74	88
point	48	48	51	66	68	100	61	95	80	100	65	61	68
area	32	29	31	52	49	70	110	150	83	82	130	140	99
Grid average point & area	48	30	32	49	46	64	99	130	68	67	68	64	61
area	33	31	33	54	53	74	110	150	86	84	130	140	100
SCC average point	74	78	83	78	100	160	99	150	130	140	140	140	130

[a] With 90% confidence.
[b] Speciated NO_x.
[c] Speciated VOC.
[d] Speciated TSP.
[e] EPA Region.
Up to two significant digits have been reported to facilitate comparisons, and does not imply a corresponding level of precision.

Table 2. *Daily/Seasonal Illustrative Uncertainty of the 1980 NAPAP Emissions Inventory*

	Percent Error[a]												
	Daily												
	SO_2	NO_x	SNO_x[b]	CO	VOC	SVOC[c]	TSP	STSP[d]	Pb	NH_3	HCl	HF	SO_4
Nation point & area	1	1	1	2	1	2	3	5	2	4	1	10	2
Region average[e] point & area	5	3	3	5	4	7	9	14	8	14	9	10	6
State average point & area	8	5	6	9	8	14	17	26	14	27	17	18	13
County average point & area	21	19	22	33	34	56	77	120	58	63	63	42	65
point	31	33	38	37	41	83	35	75	42	78	32	28	40
area	21	20	22	33	36	58	82	120	59	63	98	110	77
Grid average point & area	32	21	23	33	33	53	77	110	49	52	34	29	38
area	23	22	25	36	39	61	86	130	63	65	100	110	78
SCC average point	50	55	62	55	61	130	56	120	94	110	110	100	95

	Percent Error[a]												
	Seasonal												
	SO_2	NO_x	SNO_x[b]	CO	VOC	SVOC[c]	TSP	STSP[d]	Pb	NH_3	HCl	HF	SO_4
Nation point & area	1	1	1	1	1	2	3	5	2	4	1	10	1
Region average[e] point & area	4	2	3	4	4	7	9	14	7	14	6	8	5
State average point & area	6	4	5	8	8	14	16	25	13	26	14	14	12
County average point & area	17	16	19	30	32	55	74	110	55	61	58	35	62
point	25	28	33	32	36	80	31	73	35	75	25	20	35
area	17	16	20	31	34	56	79	120	56	61	94	100	74
Grid average point & area	26	18	21	30	31	52	74	110	46	50	27	22	34
area	20	20	22	33	37	60	82	120	59	63	98	110	75
SCC average point	41	47	55	47	53	120	49	120	89	100	100	100	90

[a] With 90% confidence.
[b] Speciated NO_x.
[c] Speciated VOC.
[d] Speciated TSP.
[e] EPA Region.
Up to two significant digits have been reported to facilitate comparisons, and does not imply a corresponding level of precision

J. Langstaff and J. Wagner are with Alliance Technologies Corporation, Bedford, MA 01730.
***J. David Mobley** is the EPA Project Officer (see below)*
The complete report, entitled "Estimation of Uncertainty for the 1980 NAPAP Emissions Inventory," (Order No. PB 87-145 397/AS; Cost: $13.95, subject to change) will be available only from:
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650
The EPA Project Officer can be contacted at:
Air and Energy Engineering Research Laboratory
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

United States
Environmental Protection
Agency

Center for Environmental Research
Information
Cincinnati OH 45268

EPA/600/S7-86/055