United States
Environmental Protection
Agency

Environmental Sciences Research
Laboratory
Research Triangle Park NC 27711

Research and Development

EPA-600/S3-83-023c June 1983

EPA

Project Summary

Houston Oxidant Modeling Study—1978: Volume III. Characterization of Data Quality

J.R. Martinez

During the period 15 September through 12 October 1978, the U. S. Environmental Protection Agency (EPA) conducted a special program that monitored air quality and meteorology in the Houston, Texas area. The objectives of the program were to obtain a comprehensive data base suitable for use with photochemical air-quality simulation models and to provide a detailed body of data that can be used to investigate Houston's air quality in general and photochemical oxidants (HOMS Study) and aerosols (HACS Study) in particular.

The objectives of this contract were (a) to evaluate the suitability of the 1978 Houston data base for photochemical modeling application, (b) to analyze spatial and temporal patterns of pollutant concentrations, (c) to archive data in a manner suitable for use with air quality simulation models, (d) to analyze and characterize the quality of the gaseous pollutant measurements, and (e) to use the data to evaluate the performance of the EKMA photochemical model. Results from the data evaluation and archiving work are reported in a three-volume report. Results from the EKMA evaluation work are reported in a separate report.

This report, Volume III of the three-volume report, covers the effort to analyze and characterize the quality of the gaseous pollutant measurements obtained in the 1978 HOMS Study. The analysis is based on data obtained in a number of field audits of the instruments used in the HOMS. The audits were performed independently by the EPA and by the Research Triangle Institute under the sponsorship of the EPA. The audit data were analyzed statistically to derive overall bounds for measurement accuracy and to define the accuracy and precision of individual instruments.

This Project Summary was developed by EPA's Environmental Sciences 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

This is the third volume of a three-volume report describing the results of a study of the quality of the pollutant measurements made during the 1978 Houston Oxidant Modeling Study (HOMS). In the course of the HOMS, field audits were conducted to check the performance of the instruments that measured gaseous pollutants. In this study, the audit data will be analyzed to determine the accuracy and precision of the HOMS air quality data.

The HOMS was conducted from 15 September through 12 October 1978. Its primary objective was to obtain a data base suitable for use with photochemical air quality simulation models. The description of the data base and the analysis of the air quality and meteorological data are reported in Volumes I and II.

Audit Procedures

Concern was with the instruments that monitored the following gaseous pollutants:

- Ozone (O_3)
- Nitric oxide (NO)

- Nitrogen dioxide (NO_2)
- Oxides of nitrogen (NO_x)
- Carbon monoxide (CO)
- Total hydrocarbons (THC)
- Methane (CH_4)
- Sulfur dioxide (SO_2)

Field audits of these instruments were performed independently by two organizations: the U. S. Environmental Protection Agency (EPA) and the Research Triangle Institute (RTI). RTI performed the audits under the sponsorship of the Quality Assurance Branch of EPA.

The audit procedure consisted of feeding a known concentration (reference concentration) of the gaseous pollutant to the instrument and recording the instrument's response. The reference concentrations usually included a zero level and several nonzero values.

Reference ozone concentrations were produced using an ultraviolet ozone generator. The EPA determined the reference ozone level using NBKI, whereas RTI established the reference concentration using gas phase titration.

Test atmospheres of NO, NO_x, and NO_2 were generated by both EPA and RTI using standard-gas cylinder dilution and gas phase titration. The standard gases were referenced to NBS-SRM 1683 (RTI) and to NBS-SRM 1684 (both EPA and RTI). Standard-gas cylinders were also used to generate reference concentrations of CO, CH_4, and SO_2. EPA used CO and CH_4 cylinders from Scott Environmental.

Statistical Analysis of Audit Data

The statistical analysis of quality assurance audit data from the HOMS includes: (1) obtaining frequency distributions for percent relative error (percent relative error = 100 x [measured - reference] / reference); and (2) performing regression analyses of measured and reference concentrations for individual instruments. The analysis was performed separately for the EPA and RTI data because of the differences in their audit procedures.

The audit data have been analyzed in two ways. One type of analysis examines the total variation of measurement errors for each variable (e.g., ozone) by aggregating the data for all the instruments. This allows us to investigate the error distribution and to estimate overall bounds of measurement error. Although these error bounds do not necessarily apply to a specific instrument at a particular time, they quantify the size and frequency of the errors likely to be found in the totality of measurements. Aggregating the audit data also provides an easy way to identify large errors and possible anomalies.

The second analysis considers each instrument individually, using regression techniques to define the relationship between measured and reference concentrations.

Conclusions and Recommendations

The premise of an audit program is that random checks of an instrument can provide a good indication of the performance of that instrument. During the HOMS, most instruments were audited at least once but not more than twice by either EPA or RTI or both. Such an audit rate for individual instruments is too low to provide a data base from which to draw statistically valid conclusions regarding the overall performance of a particular instrument.

Because the audits were relatively few in number, the data were considered in two ways: by pooling all the audits for each type of instrument and by treating each instrument individually. Pooling the data enlarges the data base and allows general conclusions to be drawn about the overall quality of the data. Aggregating the audit data also helps to identify performance anomalies. By its nature, the results of the analysis of pooled audit data cannot be applied to specific instruments. Thus, we can make statements such as, "The audits indicate that about 70% of the measurements of pollutant X were accurate to within 20%." But we cannot state that 70% of the measurements of a specific instrument were accurate to within 20%.

Analyzing the audit performance of individual instruments can only serve either to assuage or raise doubts about the quality of the measurement of a given instrument. On the one hand, a good audit performance may cause us to feel confident about the instrument, recognizing, however, that this does not necessarily imply that all the measurements of that instrument are equally good. On the other hand, a poor audit performance raises doubts about the instrument, but again does not necessarily imply that other measurements are equally poor. Hence, one should consult historical calibration and repair records to obtain a better indication of whether the poor (or good) audit represents a persistent or a one-time condition. The results of this study provide guidance about the potential performance of individual instruments, but because the analysis is restricted to the audit data, this study cannot separate persistent conc tions from one-time events. Each audit is one-time event, and only many audits ca fill in the picture.

Table 1 is a summary of the over measurement accuracy for the variou pollutants, based on the analysis of th pooled audit data. The table shows a estimated range for the percentage of th measurements that are accurate to with ±10%, ±20%, and ±30%. For exampl for ozone Table 1 shows that betwee 50% and 60% of the data were accurate t within ±20%, and 85% to 95% had erroi bounded by ±30%. Table 1 indicates th for ozone, NO, NO_x, NO_2, and SO_2 at leas 70% of the data were accurate to withi ±20%. Methane has the largest fraction (the data accurate to within ±20%. Carbo monoxide and THC can be considered t be the least accurate because they hav the smallest fraction of the data in th ±20% error band. The THC measurement were found to be subject to large ir accuracies, which explains the relativel low percentage of the data within th ±30% error band. There is cause fc concern about the quality of the TH measurements and, hence, the NMH measurements in the HOMS.

The regression analysis of individua instruments showed that almost all dem onstrated exceptionally good linearity o response. The precision of the measure ments, as indicated by the standard erro of the regressions, was also good in nearl all cases. The accuracy, as reflected in th slope of the regression line, was variable Specifically, the regression analyses in dicated potential accuracy problems at th following monitoring sites:

- O_3: Aldine, Pasadena, Channelview Seabrooke
- NO: Parkhurst, Fuqua, Pearland
- NO_x: Mae Drive, Parkhurst, Fuqua
- NO_2: Fuqua
- CO: Parkhurst, Pearland, Jackrabbit EPA Mobile Lab.
- CH_4: Parkhurst
- THC: Crawford, Clinton, Parkhurst
- SO_2: Aldine

All the instruments named above ha regression lines whose slope was greate than 1.25 or smaller than 0.75 in at leas one of the audits. The history of thes instruments should be investigated t establish whether or not the bias was a recurring phenomenon. If the bias recurred then it will be necessary to define cor rection factors to be applied to the routine ly collected data.

Future monitoring programs shoul continue to include field audits of th

Table 1. *Characterization of Overall Measurement Accuracy*

Variable	*Percent of Measurements Accurate to* ±10 %	±20 %	±30 %
Ozone	50-60	70-80	85-95
Nitric oxide	50-70	70-80	80-95
Nitrogen oxides	60-70	70-85	85-95
Nitrogen dioxide	45-55	70-75	80-90
Carbon monoxide	35-65	60-70	80-95
Methane	65-80	85-90	90-95
Total hydrocarbons	20-45	50-60	55-60
Sulfur dioxide	30-40	70-80	80-90

instruments. However, the procedures followed should be changed to facilitate the interpretation of the audit data. In particular, the audit procedure should be designed to allow the analyst to draw statistically sound general inferences about the accuracy of individual instruments from the audit data. It was not possible to do so in this study because the number of audits was too small. Hence, it is recommended that future audit programs include in their design a determination of the number of audits required to describe the performance of individual instruments over the life of the monitoring program. This is a problem in statistical experiment design that is beyond the scope of our investigation.

Although it is desirable to have more than one agency conduct field audits, differences between EPA and RTI audit protocols produced some inconsistent results (e.g., EPA ozone audits showed a tendency toward underestimation, and RTI audits showed the opposite). Thus, in this case procedural differences obscured the effects that the audits were intended to uncover. This situation should be avoided in future programs by having all parties adhere to a standard audit protocol. In this context, the audit procedures used in the present program resulted in several cases of same-day audits of the same instruments by both EPA and RTI. Such overlap is desirable to check the consistency, or lack thereof, of the results obtained by the two auditing agencies. However, same-day audits of the same instrument essentially constitute a single audit, which acts to reduce the total number of tests. Future programs should ensure that overlapping audits do not lower the total number of checks performed.

J. R. Martinez is with SRI International, Menlo Park, CA 94025.
***B. Dimitriades** is the EPA Project Officer (see below).*
The complete report, entitled "Houston Oxidant Modeling Study—1978: Volume III. Characterization of Data Quality," (Order No. PB 83-194 217; Cost: $11.50, 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:
Environmental Sciences 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

Postage and
Fees Paid
Environmental
Protection
Agency
EPA 335

Official Business
Penalty for Private Use $300

PS 0000329
U S ENVIR PROTECTION AGENCY
REGION 5 LIBRARY
230 S DEARBORN STREET
CHICAGO IL 60604