MASTER

ERROR ANALYSIS OF MASS SPECTROMETER DATA

R. W. Ashcraft

DEVELOPMENT DIVISION

OCTOBER - DECEMBER 1976

Normal Process Development
Endeavor No. 241

Mason & Hanger-Silas Mason Co., Inc.
Pantex Plant

P. O. BOX 30020
AMARILLO, TEXAS 79177
806-335-1581

operated for the
ENERGY RESEARCH AND DEVELOPMENT ADMINISTRATION
under
U. S. GOVERNMENT Contract EY 76 C 04 0487

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately-owned rights.

NOTICE
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

ERROR ANALYSIS OF MASS SPECTROMETER DATA

R. W. Ashcraft

DEVELOPMENT DIVISION

October - December 1976
Endeavor No. 241

ABSTRACT

A computer program utilizing weighted least squares has been used to analyze the mass spectra of gas samples. The reliability of the partial pressures has not been established because the magnitude, nature, and effect of errors in the cracking pattern and sensitivity data have not been confirmed. A series of special mass spectrometric analyses was designed to isolate these errors. A sample data case was analyzed using the nominal values of the cracking pattern and sensitivity. The nominal values were then perturbed according to the results of the special runs and the data case was analyzed again. There was not a significant change in the solutions to the system of equations by the introduction of errors to the cracking pattern and sensitivity data.

DISCUSSION

BACKGROUND

The use of weighted least squares to analyze the mass spectra of gas samples has been discussed in a previous report(1). That report contains a mathematical description of the system of equations to be solved, definitions of the quantities appearing in the equations, and an application of the analysis for a known gas sample. It was shown in the report that the spectrum of a gas sample can be represented as

$$C_j = \sum_{i=1}^N k_{ij} f_{ij} P_i$$

where C_j is the number of counts for the j^{th} mass to charge ratio, k_i is the sensitivity calibration for the i^{th} gas, f_{ij} is the cracking pattern coefficient for the j^{th} mass to charge ratio and i^{th} gas, and P_i is the (unknown) partial pressure of the i^{th} gas.

Solving for the partial pressures is analogous to obtaining least squares estimates of model parameters in multiple linear regression. The products $k_{ij} f_{ij}$ are analogous to observed values of the independent variables, which, for the classical regression model, are assumed to have no error. When the conditions for the classical model are met, confidence limits may be computed for the solutions, i.e., reliability of the solutions can be established.

Conditions for the classical model are not met with the mass spectrometric data because the values of k_i and f_{ij} are subject to error. The cracking pattern coefficients (f_{ij}) have a tendency to "drift" with time and require periodic updating to maintain accuracy. The sensitivity calibration values (k_i) are updated daily, but not by direct measurement. Only nitrogen is measured daily and the remaining gases are adjusted according to the behavior of the nitrogen. The k_i have also been suspected of having a functional dependence on sample pressure. In order to establish reliability of the solutions (partial pressures) it was necessary to obtain the magnitude of the errors and verify their existence.

(1) R. W. Ashcraft, "Analysis of Mass Spectrometer Data by Weighted Least Squares," MHSMP-75-20V (January-March 1975).

DESCRIPTION OF DATA SETS

A series of special mass spectrometric analyses was designed to study the errors in the system. The analyses, performed by the Quality Gas Lab over a two week period in the spring of 1976, were as follows:

- Data Set 1: Triplicate runs on CO₂ at 50 μ m Hg each morning after the nitrogen calibration runs.
- Data Set 2: Single runs on CO₂ at 40, 30, 20, and 5 μ m Hg every other day at various times on these days.
- Data Set 3: A single run on butane at 50 μ m Hg each day at various times of the day.

DATA SET 1

This data set consisted of triplicate runs on CO₂ at 50 μ m Hg daily. By averaging the 3 runs (the method used for the nitrogen calibration), values of CO₂ sensitivity were obtained for comparison with the values calculated from the nitrogen runs. Table I gives a listing of these data as well as the differences and percentage errors. The largest discrepancy occurred on 5/28/76 with a percentage error of -5.48, but this was due to equipment malfunction. The remainder of the percentage errors were treated as a random sample from a normal distribution whose parameters were estimated by the sample mean and unbiased standard deviation. Hence, the percentage error in CO₂ sensitivity due to nitrogen calibration was represented as a normal distribution with mean -0.3588 and standard deviation 0.6929.

DATA SET 2

This data set consisted of single runs on CO₂ at pressures of 40, 30, 20, and 5 μ m Hg. The runs were made every other day at random times on those days. These data are listed in Table II and presented graphically in Fig. 1. As seen in Fig. 1, the data exhibit an increase in CO₂ sensitivity as sample pressure is decreased, although the relationship is somewhat less evident for the data collected on 5/24/76. Least squares lines were fitted to each of the five sets to obtain estimates of the slopes; these values are given in Table II and are all negative. Thus, CO₂ sensitivity (k) appeared to have a linear relationship with sample pressure (P) given by

$$k = k_{50} + m (P - 50) \quad (1)$$

where k_{50} is the sensitivity of CO₂ at 50 μ m Hg, and m is the slope estimated by averaging the five individual slopes. For this data set, m had a value of -0.80485 counts/ μ m Hg/ μ m Hg.

Table I. Results of Triplicate Runs on CO₂ at 50 μ m Hg

<u>Date</u>	<u>Average CO₂ Sensitivity^a By Direct Measurement</u>	<u>Average CO₂ Sensitivity By N₂ Calibration</u>	<u>Difference</u>	<u>Percentage Error</u>
05/24/76	1259.8720	1263.7000	+ 3.8280	+0.3038
05/25/76	1241.8550	1228.0222	-13.8328	-1.1139
05/26/76	1260.5454	1258.4640	- 2.0814	-0.1651
05/27/76	1268.0758	1260.6558	- 7.4200	-0.5851
05/28/76	1332.8283	1259.8035	-73.0248	-5.4789 ^b
06/01/76	1274.3991	1270.8843	- 3.5148	-0.2758
06/02/76	1264.3674	1269.1795	+ 4.8121	+0.3806
06/03/76	1277.3388	1257.3681	-19.9707	-1.5635
06/04/76	1273.8599	1275.7550	+ 1.8951	+0.1488

Note: The mean and standard deviation of the percentage errors are -0.3588 and 0.6929, respectively.

^a*Sensitivity values are given in units of counts/ μ m Hg.*

^b*This value was not included in the calculation of the mean and standard deviation of the percentage errors because it occurred during a period of equipment malfunction and is not representative of the system.*

Table II. Results of Runs on CO₂ at 40, 30, 20, and 5 μ m Hg

<u>Date</u>	<u>Pressure^a</u>	<u>Sensitivity^b</u>	<u>Maximum Sensitivity Change (%)</u>	<u>Slope</u>
05/24/76	40.72	1259.8232	0.79	-0.10155 ^c
	30.03	1288.7113		
	20.34	1274.3363		
	5.19	1269.7495		
05/26/76	40.66	1271.5199	3.19	-1.10339
	30.07	1277.0203		
	20.72	1274.1313		
	5.48	1312.0438		
05/28/76	40.57	1279.2704	2.85	-0.99347
	30.54	1283.5625		
	20.09	1284.2210		
	5.13	1315.7895		
06/01/76	40.79	1233.1454	3.06	-0.95365
	30.61	1254.4920		
	20.08	1253.4861		
	5.28	1270.8333		
06/03/76	40.92	1292.7663	2.65	-0.87220
	30.86	1302.6572		
	20.47	1295.5545		
	5.81	1327.0224		

^aSample pressure is given in units of μ m Hg.

^bSensitivity is given in units of counts/ μ m Hg.

^cThis is the value of the slope of the regression line and is given in units of counts/ μ m Hg/ μ m Hg. The average value of the five slopes is -0.80485.

Fig. 1. CO₂ Sensitivity as a Function of Sample Pressure for Five Different Days

In order to relate equation (1) to other gases, it was necessary to change the dependent variable from absolute sensitivity to relative sensitivity. This was accomplished by dividing the sensitivity values on a given day by the value observed at 40 $\mu\text{m Hg}$. Least squares lines were fitted to the five resulting data sets and the five slopes were averaged as before to get an effective slope of $-0.000634603 (\mu\text{m Hg})^{-1}$. By choosing 50 $\mu\text{m Hg}$ as the reference point, the equation for relative sensitivity became

$$k/k_{50} = 1 - 0.000634603 (P - 50) \quad (1a)$$

where k_{50} is the sensitivity of the gas at 50 $\mu\text{m Hg}$, P is the sample pressure, and k is the sensitivity of the gas at pressure P . Table III gives a listing of the data for the derivation of equation (1a).

DATA SET 3

This data set consisted of single runs on butane at 50 $\mu\text{m Hg}$. The runs were made daily and were done at random times during the day. Butane was selected for this set because it has many (43) peaks with a wide range of relative size (1.0 to 0.00026), and it is easily evacuated from the mass spectrometer after analysis. The mean, standard deviation, and coefficient of variation of the cracking pattern coefficients were computed for each of the 43 peaks and are listed in descending order in Table IV. A plot of coefficient of variation (V) versus the mean of the cracking pattern coefficients (f) is given in Fig. 2. As seen in Fig. 2, V increases linearly as f decreases from 1 to 0.01. At this point a slight discontinuity in the data occurs and a second set of points shows a linear trend for V as f decreases from 0.01 to 0.0007. Below 0.0007, V tends to oscillate between 4.5% and 21%, indicating that these small peaks are more unstable than those above 0.0007. The data were put into the following functional form:

$$V = 0 \quad \text{if } f = 1 \quad (2)$$

$$V = 0.30524 - 0.20821 \log f \quad \text{if } 0.01185 \leq f < 1 \quad (3)$$

$$V = -0.28015 - 1.24379 \log f \quad \text{if } 0.00075 \leq f < 0.01185 \quad (4)$$

$$V = 14.96670 \quad \text{if } 0 < f < 0.00075 \quad (5)$$

Equation (2) results from the fact that the main peak always has a coefficient of 1.0. Equations (3) and (4) result from linear fits to the specified regions. Equation (5) is an attempt to describe the data in the region with unstable small peaks. It was decided that the average of the six values of V was the best way to represent this region. The standard deviation (S) could now be given as the following function of f :

$$S = 0 \quad \text{if } f = 1 \quad (6)$$

$$S = f (0.0030524 - 0.0020821 \log f) \quad \text{if } 0.01185 \leq f < 1 \quad (7)$$

$$S = f (-0.0028015 - 0.0124379 \log f) \quad \text{if } 0.00075 \leq f < 0.01185 \quad (8)$$

$$S = 0.14967 f \quad \text{if } 0 < f < 0.00075 \quad (9)$$

Table III. Results of Least Squares Analysis of Relative Sensitivity of CO₂

<u>Date</u>	<u>Pressure^a</u>	<u>Relative Sensitivity^b</u>	<u>Slope^c</u>
05/24/76	40.72	1.0000	-0.000081
	30.03	1.0229	
	20.34	1.0115	
	5.19	1.0079	
05/26/76	40.66	1.0000	-0.000868
	30.07	1.0043	
	20.72	1.0021	
	5.48	1.0319	
05/28/76	40.57	1.0000	-0.000777
	30.54	1.0034	
	20.09	1.0039	
	5.13	1.0285	
06/01/76	40.79	1.0000	-0.000773
	30.61	1.0173	
	20.08	1.0165	
	5.28	1.0306	
06/03/76	40.92	1.0000	-0.000675
	30.86	1.0077	
	20.47	1.0022	
	5.81	1.0265	

^aSample pressure is given in units of $\mu\text{m Hg}$.

^bRelative sensitivity is a unitless quantity.

^cThe slope of the regression line is given in units of $(\mu\text{m Hg})^{-1}$. The average value of the five slopes is $-0.000635 (\mu\text{m Hg})^{-1}$.

Table IV. Mean, Standard Deviation, and Coefficient of Variation for Butane Cracking Pattern Coefficients Observed from 5/24/76 to 6/4/76

<u>M/e</u>	<u>Mean</u>	<u>Standard Deviation</u>	<u>Coefficient of Variation</u>
43	1.0000	0.00000	0.00000
29	0.43327	0.00143	0.32892
27	0.40020	0.00210	0.52367
28	0.32463	0.00072	0.22287
41	0.29914	0.00065	0.21684
58	0.15288	0.00110	0.72011
39	0.14991	0.00060	0.40080
42	0.12688	0.00101	0.79671
26	0.07600	0.00002	0.32156
15	0.07101	0.00062	0.87405
44	0.03318	0.00015	0.44492
57	0.02601	0.00011	0.43986
38	0.02472	0.00013	0.52398
40	0.01963	0.00012	0.61691
50	0.01731	0.00006	0.35864
1	0.01705	0.00017	0.97242
14	0.01404	0.00014	0.98023
37	0.01358	0.00015	1.07330
51	0.01315	0.00004	0.27717
55	0.01055	0.00023	2.17514
30	0.00952	0.00020	2.13560
53	0.00894	0.00020	2.19386
56	0.00834	0.00017	2.09095
25	0.00691	0.00017	2.41590
59	0.00669	0.00025	3.68876
49	0.00563	0.00011	1.87653
13	0.00477	0.00012	2.43728
25.5	0.00448	0.00014	3.02234
52	0.00309	0.00006	1.87685
12	0.00270	0.00007	2.72917
54	0.00214	0.00007	3.16307
16	0.00180	0.00006	3.23308
2	0.00166	0.00006	3.62854
32	0.00128	0.00005	3.60841
36	0.00123	0.00004	3.22265
48	0.00092	0.00003	3.64837
24	0.00075	0.00002	3.30520
18	0.00062	0.00013	20.62723
19	0.00049	0.00010	19.67721
45	0.00039	0.00002	4.67404
20	0.00034	0.00003	10.20406
19.5	0.00026	0.00005	17.59477
17	0.00026	0.00004	17.02291

Fig. 2, Coefficient of Variation Versus the Mean of Butane Cracking Pattern Coefficients

SAMPLE DATA CASE

The nature and magnitude of the errors in the CO₂ sensitivity data and the cracking pattern of butane had now been determined; all that remained was to test the significance of the errors and their effect on the reliability of the solutions. A data case having most of the common gases present was selected for this test^a. The set was first run using the cracking pattern and sensitivity data as supplied by the Gas Lab. The next step was to perturb the data according to the findings established in the analysis of the three special data sets. It was necessary to assume that all the gases would behave in the same manner as butane and CO₂. The sensitivity of each gas was first altered according to the relationship given in equation (1a). A Gaussian random number generator was used to further perturb these values according to the normal distribution derived from Data Set 1. Finally, each cracking pattern coefficient was selected at random from a Gaussian distribution with mean equal to the nominal value of the coefficient and standard deviation as determined by equations (6) through (9).

Table V gives a comparison of the solutions as obtained from the nominal data and those obtained from the perturbed data. As seen in the table, almost no difference exists between the solutions and the differences decrease as the amount of gas present in the system decreases. Even the worst case (He) is reliable to 0.25%. Thus, the errors present in the cracking pattern and sensitivity data do not have a significant effect on the solutions of the equations.

CONCLUSIONS

The percentage error in CO₂ sensitivity due to nitrogen calibration can be represented as a normal distribution with mean -0.3588 and standard deviation 0.6929. The sensitivity of CO₂ at a pressure P, relative to the sensitivity at 50 μ m Hg, can be represented as

$$k/k_{50} = 1 - 0.000634603 (P - 50).$$

The peaks of the cracking pattern of butane can be treated as normal distributions with means equal to the nominal values of the peaks and the standard deviations given by

$S = 0$	if $f = 1$
$S = f (0.0030524 - 0.0020821 \log f)$	if $0.01185 \leq f < 1$
$S = f (-0.0028015 - 0.0124379 \log f)$	if $0.00075 \leq f < 0.01185$
$S = 0.14967 f$	if $0 < f < 0.00075$

^aThe data case selected was the aliquot sample from container X-28 in P. A. Foster's LX-14 compatibility study. The aliquot was from trap No. 5 on mass spectrometer CEC 21-104 No. 2 and was analyzed on 7/27/76.

Table V. Comparison of Solutions^a of X-28-05 Aliquot Sample Obtained from Nominal and Perturbed Data

<u>Gas</u>	<u>Percentage From Nominal Data</u>	<u>Percentage From Perturbed Data</u>	<u>Difference</u>
H ₂	0.38	0.37	0.01
He	89.31	89.56	- 0.25
CH ₄	Trace	Trace	-
Ne	Trace	0.01	-
CO	0.00	0.00	-
N ₂	7.51	7.34	0.17
NO	0.35	0.34	0.01
O ₂	0.08	0.08	-
Ar	0.14	0.13	0.01
N ₂ O	0.00	0.00	-
CO ₂	2.22	2.16	0.06

^aThe solutions of the equations are partial pressures of the gases which are converted to percentages by the equation

$$A_i = \frac{100 P_i}{\Sigma P_i}$$

where A_i is the percentage of the i^{th} gas, P_i is the partial pressure of the i^{th} gas, and ΣP_i is the total pressure of the sample.

When all the gases are assumed to behave in the same manner as CO₂ and butane, and the errors are introduced into the cracking pattern and sensitivity data, the calculated percentages of gases were found to be reliable to within 0.25%. Hence, the errors do not cause a significant change in the solutions. Special runs should be made for gases other than CO₂ and butane to obtain a more complete description of the system.

ACKNOWLEDGEMENTS

The author wishes to extend special thanks to L. C. Myers of the Development Division for his support of this project, and to D. K. Remling of the Quality Division for his efforts to provide "realistic data" in the series of special mass spectrometric runs.