

For the MCl_2 , the mean ionic activity coefficients of each valence type were averaged from tabulated values (6). The average single ion activities in Table IV were estimated using the pH convention (7). Although this convention is properly applicable only at ionic strengths less than 0.1M, these activity coefficients are adequate for the present application.

LITERATURE CITED

- (1) R. J. Levins, *Anal. Chem.*, **43**, 1045 (1971).
- (2) R. J. Levins, *Anal. Chem.*, **44**, 1544 (1972).
- (3) D. Ammann, E. Pretsch, and W. Simon, *Anal. Lett.*, **5**, 843 (1972).
- (4) J. Kielland, *J. Am. Chem. Soc.*, **59**, 1675 (1937).

- (5) R. G. Bates, B. R. Staples, and R. A. Robinson, *Anal. Chem.*, **42**, 867 (1970).
- (6) R. A. Robinson and R. H. Stokes, "Electrolyte Solutions," 2nd ed (Revised), Butterworths, London, 1970.
- (7) R. G. Bates, *Pure Appl. Chem.*, **36**, 407 (1973).

Elizabeth W. Baumann

E. I. du Pont de Nemours & Co.
Savannah River Laboratory
Aiken, SC 29801

RECEIVED for review October 7, 1974. Accepted January 17, 1975. Work performed under USAEC Contract No. AT(07-2)-1.

AIDS FOR ANALYTICAL CHEMISTS

Program ELAL: An Interactive Minicomputer Based Elemental Analysis of Low and Medium Resolution Mass Spectra

Joseph E. Evans and N. B. Jurinski

U.S. Army Environmental Hygiene Agency, Aberdeen Proving Ground, Md. 21010

In recent years, the popularity of the combination gas chromatograph-mass spectrometer (GC-MS) has increased markedly. This is due, in part, to the more stringent regulations set forth by OSHA and EPA requiring definitive identification of environmental contaminants. Another factor that contributed to this increase was the advent of the minicomputer based dedicated data system. This allowed for an efficient means of collection, reduction, and analysis of large volumes of data, thus making feasible, for many laboratories, an instrument that was once considered only a tool of research.

This rapid increase, however, has created a deficiency in the number of qualified mass spectrometrists needed to operate the large number of new instruments and interpret the data they provide. Therefore, many laboratories rely heavily upon the use of libraries of reference spectra and computerized search routines for definitive identification of unknowns (1). Although these are useful tools that aid in identification, invalid conclusions can often result when these methods are used alone. The practice of confirming the identification of unknown spectra by analyzing a reference standard is an excellent but not always practical procedure. The diligent use of elemental analysis can help the typical support laboratory reach more reliable conclusions, while significantly reducing turnaround time.

Although elemental analysis will not provide complete structural information, it does provide information as to the amounts of each element present. High resolution mass spectrometry is the most powerful method of determining elemental composition; however, the effects of the abundances of naturally occurring isotopes in low and medium resolution spectra can also lead to the elemental composition of many ions.

Many mass spectrometrists, because of inexperience and the time involved, fail to compute an elemental analysis on each unknown analyzed in the laboratory. With a dedicated

minicomputer at their disposal, this chore can be readily handled. It was for this reason, program ELAL was developed. The program will compute an elemental analysis on low or medium resolution spectra for C, N, H, Cl, Br, F, S, Si, O, and P.

EXPERIMENTAL

ELAL was designed to be a callable subroutine that can be included as a part of existing GC-MS software or as a separate program that can be used when the data system is not in use. The program asks for only five items of information: the mass and intensity of the (A) peak and the intensities of the (A + 1), (A + 2), and the (A + 4) peaks. Either normalized or raw data may be entered. Computation is based on a modification of the manual method of elemental analysis described by McLafferty (2). With this method, elements are categorized as (A), (A + 1), or (A + 2) elements depending upon the prevalence of isotopes having 1 or 2 additional mass units. Thus, Cl with isotopes of 35 and 37 amu is considered an (A + 2) element. The common (A) elements are H, F, and P; the (A + 1) elements are C and N; and the (A + 2) elements are Cl, Br, S, Si, and O.

RESULTS AND DISCUSSION

After the requested data are entered, ELAL normalizes them and then uses the abundance of the (A + 1) peak to compute the number of carbons present. By means of a binomial expansion, the contribution of carbon to the (A + 2) peak is determined. The (A + 2) elements are then computed in order of decreasing isotopic abundance. If sulfur and/or silicon are found to be present, ELAL corrects for their contribution to the abundance of the (A + 1) peak and then completely recalculates the (A + 1) and (A + 2) elements. The contribution of oxygen to the (A + 1) peak is usually insignificant, so a correction need not be performed. Checkpoints are provided throughout the program to ensure that ELAL has not calculated an erroneous number of atoms for any one element. Chlorine and bromine are computed simultaneously by subtracting the abundances of

```

PROGRAM ELAL
AN ELEMENTAL ANALYSIS OF MASS SPECTRA

INPUT AM, A, A+1, A+2, A+4
?95, 2.4, .06, 0, 0

INPUT NORMALIZED
A= 100
A+1= 2.5
A+2= 0
A+4= 0

A+1 REMAINDER = .4

MASS OF A REMAINING = 71
MASS OF A REMAINING = 9

OVERSATURATED
C 2   H 9   N 0   CL 0   BR 0   F 0
S 0   SI 0   O 0   P 2

R+DB= -1.5

AM=95
MW=95
INTERACTION? TYPE 1 IF YES, 2 IF NO
?1
INPUT CN AND NN
?2, 1
INTERACTION WITH OTHER ELEMENTS?
TYPE 1 IF YES, 2 IF NO
?2

A+1 REMAINDER = .04

MASS OF A REMAINING = 57
C 2   H 9   N 1   CL 0   BR 0   F 3
S 0   SI 0   O 0   P 0

R+DB= 2

AM=95
MW=95
INTERACTION? TYPE 1 IF YES, 2 IF NO
?2

```

Figure 1. An elemental analysis of trifluoroacetonitrile. (A + 1) remainder and fractional "rings + double bonds" calculation indicates presence of nitrogen

the (A), (A + 2) and (A + 4) peaks from an abundance table within a stored array and then comparing the squares of the differences with a minimum acceptable value. This algorithm allows analysis of compounds containing up to 4 chlorine atoms and/or 3 bromine atoms. The data array can be easily expanded to include the (A + 6) and (A + 8) peaks which would allow the analysis of compounds containing up to 7 chlorine and/or 4 bromine atoms. The program itself places no numerical limits on the number of atoms of other elements that can be analyzed. The only constraints are those imposed by the maximum possible abundance of the isotope peaks. For example, each carbon atom contributes 1.08% to the abundance of the (A + 1) peak, thus limiting the maximum number of carbon atoms that can be measured to 92. Similar constraints exist for the other (A + 1) and (A + 2) elements; however, they are infrequently encountered.

The (A) elements are computed last since they contribute minimally, if at all, to the abundances of other peaks. Their computation is based on the mass remainder and saturation of the ion after the (A + 1) and (A + 2) elements have been determined.

During the analysis, ELAL lists certain pertinent data about the series of mass peaks it is analyzing. These data are extremely helpful to the operator in evaluating the accuracy of the elemental analysis being performed.

The first output is a list of the normalized input data. This assists the operator by presenting the data in an easily recognizable form whereby the most abundant peak is assigned a value of 100 and all others are normalized to it. Following this, the remainder of the abundance of the (A + 1) peak after carbon calculation is printed. This allows a check on round-off error that could affect the number of carbons calculated. Since nitrogen is also an (A + 1) element, it gives the user information as to its possible presence. If chlorine and/or bromine is detected, ELAL responds with a Cl/Br fit factor. The smaller this number,

the closer the fit and the greater the probability that the number of halogen atoms, predicted by ELAL, is correct. Next, the remainder of the abundance of the (A + 2) peak is displayed and is continually updated each time an (A + 2) element is encountered. After all (A + 1) and (A + 2) elements have been computed, the remaining mass of the ion is printed and then the (A) elements are computed.

After ELAL completes its analysis, a tabular listing of the elemental data is displayed. A (rings + double bonds) calculation is then performed on the listed data. This calculation provides still more information as to the correctness of the analysis by providing the total number of rings and double bonds necessary to attain the correct degree of unsaturation for the calculated analysis.

At this point, the program enters a subroutine which allows operator interaction. ELAL needs help in computing an elemental analysis when nitrogen is present. Because nitrogen contribution to the (A + 1) peak is often equal to the experimental error present during the calculation of the number of carbons, the presence or absence of this element is best left to the judgment of the operator. Information provided by the (A + 1) remainder, the (rings + double bonds) calculation, and visual inspection of the spectrum will usually enable the operator to make a correct determination. ELAL asks the operator for the expected number of nitrogens and then recomputes a complete elemental analysis based on this input. It is also possible for the operator to correct any other mistake ELAL may have made. This is done by inputting what the operator thinks is the correct number of atoms of any of the elements. ELAL then reruns the analysis based upon these corrections.

The analysis run time on a minicomputer is extremely short and is basically governed by the type of output device utilized. When a CRT is used, the program appears almost transparent and is completed within 5 to 6 seconds. When sulfur and/or silicone is present, the run time is increased by 2 to 3 seconds because of an abundance correction loop. Use of a teletype requires about 45 seconds for the same output. A similar run time for each output device is required for the interactive stage. Figure 1 represents a typical elemental analysis.

CONCLUSIONS

ELAL will perform an accurate analysis as long as the mass spectrometer is in a reasonable state of tune. The constants used in the program were derived assuming a maximum mass abundance error of 5% of the theoretical abundance. Almost all instruments, if well adjusted and maintained, will operate within these limits. These constants can and should be changed if the program is to be used with an instrument that continuously maintains a higher degree of accuracy. Alternate constants have been derived for 2% error and can be supplied with the program. ELAL is written both in Hewlett-Packard Basic and Fortran. (The Fortran or Basic listing of the program and a users guide is available from the address listed above.) The Fortran used is compatible with many of the minicomputers used with GC-MS data systems. Hewlett-Packard Basic is a truly conversational language and can easily be translated into the Basic languages supported by other minicomputer manufacturers. The core requirement for the Basic interpreter and ELAL is 8K words. The compiled Fortran version requires 4K words.

ACKNOWLEDGMENT

The authors thank R. M. Tuggle for his assistance during the development of this program.

LITERATURE CITED

- (1) H. S. Hertz, R. Hites, and K. Biemann, *Anal. Chem.*, **43**, 681-691 (1971).
 (2) F. W. McLafferty, "Interpretation of Mass Spectra," 2nd ed., W. A. Benjamin, Inc., Reading, Mass., 1973, pp 16-27.

RECEIVED for review August 23, 1974. Accepted December 9, 1974. This material is presented for information only.

Random Error Propagation by Monte Carlo Simulation

Lowell M. Schwartz

Department of Chemistry, University of Massachusetts—Boston, Boston, MA 02125

In general, a result or set of results y_i are calculated from a number of experimental quantities x_j but the x_j are uncertain due perhaps to determinate and/or random (indeterminate) errors. To project these errors into the resultant y_i is a topic treated in many analytical and physical chemistry texts under the title of "error propagation". I wish to focus here on the particular problem of calculating the "most probable" uncertainty or error δy_i in y_i due to random errors δx_j in the x_j . The "most probable" uncertainty is to be distinguished from the "maximum probable" error, this latter quantity estimating the maximum excursion expected of the results y_i from the (unknown) true values which can be rationalized by the δx_j . This maximum by nature overestimates the most probable error.

A formula for the most probable propagated random error is commonly given (1, 2) as

$$(\delta y_i)^2 = \sum_j \left(\frac{\partial y_i}{\partial x_j} \right)^2 (\delta x_j)^2 \quad (1)$$

The validity of this expression is based on a number of assumptions: (a) the various δx_j are statistically uncorrelated; (b) the x_j are functionally independent; and (c), if the y_i are not linear functions of the x_j , then each δx_j must be sufficiently small relative to the corresponding mean values of the x_j so that the functions can be reasonably linearized about the means. If one or more of these assumptions is invalid, Equation 1 can be suitably corrected but at the expense of computational effort. A more difficult problem arises when the functional forms $y_i(x_j)$ are differentiable only with great difficulty or perhaps not at all. This problem, however, could be attacked by numerical methods.

The technique of Monte Carlo simulation is an alternative approach to be considered when Equation 1 or its corrected form is inconvenient and when digital computing facilities are available to handle the substantial and repetitive calculations. It has been applied to various problems in engineering, economics, and industrial operations (3), but is relatively unfamiliar to chemists.

Calculational Method. The random error propagation calculation under consideration in this paper is generally attempted only when the δy_i are required but when these cannot be determined directly by repeating the experiment enough times for statistical validity. Clearly, if the experiment could be done repeatedly to generate a reasonable statistical population of each y_i value, then the information on the random uncertainty in the y_i would be in hand. The digital computer offers a convenient means of simulating the repetition. It is only necessary to generate new sets of x_j data values, calculate the resultant y_i values and store these for later statistical analysis. The repeated sets of x_j could be generated in accordance with any distribution function, but the Gaussian (normal) serves to represent fluctuation in most physical situation.

Distribution by this Agency does not imply endorsement of any commercial product referred to herein. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

Experienced scientific programmers should have little difficulty with this straight-forward calculation, but for those readers wishing to adapt a written program to their own problem, one is hereby offered with the following features and limitations: 1) Written in FORTRAN IV language. 2) Accepts $j \leq 20$ x_j values and their uncertainties δx_j , assumed to be uncorrelated standard deviations. 3) Calculates $i \leq 20$ y_i results and their standard errors (estimated standard deviations) δy_i . 4) Simulates $n \leq 100$ repetitive experiments based on the assumption that the x_j values scatter normally. 5) Prints out calculated mean values and estimated standard deviations of the y_i .

A program listing and operating instructions will be sent without charge upon request to the author. The operator, in addition to writing his own subroutine relating his y_i to x_j , must select the number n of simulated repetitions. This choice is to be based on a trade-off between computer running time and the degree of certainty required of the calculated δy_i values. A quantity which may be helpful in making this choice is the standard error (estimated standard deviation) in δy_i which is given by $(\delta y_i)^2 \sqrt{2/(n-1)}$ and which is derived for normal distributions (4) but is an approximation for others. Reference (4) also describes how confidence limits may be estimated for δy_i .

Sample Calculation. The following problem illustrates application of the technique. The primary dissociation constant pK_1 of a dibasic acid is to be calculated from a potentiometric pH titration of 50.00 \pm 0.05 ml of 0.0563 \pm 0.0005*F* solution of the disodium salt Na_2A with HCl solution beyond the first equivalence point. The second dissociation constant K_2 is known to be 0.00538 \pm 0.00012. After addition of 7.00 \pm 0.01 ml of 0.634 \pm 0.003 *F* HCl, the pH is 1.525 \pm 0.002. Each of these six expressed uncertainties δx_j are measured or estimated standard deviations. The desired result pK_1 is calculated from the following set of coupled equations which represent the equilibria, stoichiometry, and activity coefficient correlation.

$$pK_1 = -\log K_1 \quad K_1 = \frac{\gamma_{\pm}^2 [H^+][HA^-]}{[H_2A]}$$

$$[H^+] = \frac{10^{-pH}}{\gamma_{\pm}} \quad \log \gamma_{\pm} = \frac{-0.51\sqrt{I}}{1.0 + 2.0\sqrt{I}}$$

$$[HA^-] = \frac{2CV(V+v) - [H^+] - C_{HCl}v/(V+v)}{1 + \frac{2K_2}{[H^+]\gamma_{\pm}^2}}$$

$$[A^{2-}] = \frac{K_2[HA^-]}{[H^+]\gamma_{\pm}^2}$$

$$[H_2A] = \frac{CV}{V+v} - [HA^-] - [A^{2-}]$$

$$I = \frac{1}{2} \left([H^+] + [HA^-] + 4[A^{2-}] + \frac{C_{HCl}v + 2CV}{V+v} \right)$$