

STATISTICAL WEIGHTS IN SPECTROPHOTOMETRIC ANALYSIS OF MULTICOMPONENT MIXTURES

O. SCHIESSL^a and J. ŠUSTEK^b

^a State Forest Product Research Institute, 891 99 Bratislava and

^b Research Institute of Agrochemical Technology, 810 04 Bratislava

Dedicated to Professor S. Stankoviánsky on the occasion of his 70th birthday.

Received November 25th, 1976

Statistical weights were employed during spectrophotometric analyses of multicomponent mixtures. The mathematical relations derived were applied to a five-component model system simulated by means of a computer. The necessity of the use of statistical weights is discussed for the case that systematic errors occur.

In the preceding works¹⁻⁴ dealing with the use of overdetermined systems of linear absorption equations in spectrophotometric analyses of multicomponent mixtures, all the analytical positions were deemed equivalent and a unit statistical weight was assigned to them. In real conditions, however, the absorbance measurement accuracy is affected by various factors, such as the fact that the error of absorbance measurement at the absorption band maximum is lower than that at the band side, the effect of the absorption band overlap, *etc.*; the experimental absorbances measured at various analytical positions (wavelengths, wavenumbers) should be then treated with different statistical weights. This more general case of data processing is the concern of this work.

THEORETICAL

For simplicity's sake we shall consider the problem of calculation of component concentrations in a single sample. The equation for the sample absorbance residues written* in the matrix form as³

$$E_B = KX - B \quad (1)$$

reduces then to the vector equation

$$e_b = Kx - b. \quad (2)$$

* For the meaning of the mathematical symbols see the attached List of Symbols.

If the absorbances are measured at k analytical positions, this represents a set of k equations

$$e_{b,g} = \sum_{j=1}^n (k_{gj} \cdot x_j) - b_g, \quad (3)$$

whence

$$\frac{\partial e_{b,g}}{\partial x_{j^*}} = k_{gj^*}. \quad (4)$$

Assume different statistical weights to be associated with the absorbance measurements at different analytical positions. According to the method of least squares we seek for a solution \mathbf{x} such that for any $j^* = 1, 2, \dots, n$ the sum of the weighted squares of absorbance deviations be minimum, *i.e.*

$$\frac{\partial \sum_{g=1}^k e_{b,g}^2 \cdot w_g}{\partial x_{j^*}} = 0. \quad (5)$$

After substitution of Eq. (3) in Eq. (5) and differentiation and with regard to Eq. (4) we have the system of n equations

$$2 \sum_{g=1}^k w_g \cdot \left[\sum_{j=1}^n (k_{gj} \cdot x_j) - b_g \right] \cdot k_{gj^*} = 0. \quad (6)$$

Introducing the diagonal weight matrix

$$\mathbf{W} = \begin{bmatrix} w_1 & 0 & 0 & \dots & 0 \\ 0 & w_2 & 0 & \dots & 0 \\ 0 & 0 & w_3 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & w_k \end{bmatrix} \quad (7)$$

we can rewrite the system of equations (6) in the matrix form;

$$\mathbf{K}^T \mathbf{W} \mathbf{K} \mathbf{x} = \mathbf{K}^T \mathbf{W} \mathbf{b}; \quad (8)$$

multiplication by the symmetrical matrix $(\mathbf{K}^T \mathbf{W} \mathbf{K})^{-1}$ from the left yields the expression

$$\mathbf{x} = (\mathbf{K}^T \mathbf{W} \mathbf{K})^{-1} \mathbf{K}^T \mathbf{W} \mathbf{b}. \quad (9)$$

If the elements of the matrix \mathbf{W} are independent of the absorbance values of the samples (the weights are a function of the analytical positions solely), the relation (9) can be generalized to

$$\mathbf{X} = (\mathbf{K}^T \mathbf{W} \mathbf{K})^{-1} \mathbf{K}^T \mathbf{W} \mathbf{B}. \quad (10)$$

The expression

$$(\mathbf{K}^T \mathbf{W} \mathbf{K})^{-1} \mathbf{K}^T \mathbf{W} = \mathbf{Q}_w, \quad (11)$$

occurring in the relations (9) and (10), represents a matrix analogous to the matrix $\mathbf{Q} = (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{K}^T$ for the calculation of the concentrations, derived in the preceding paper³.

EXPERIMENTAL

The weight matrix. The elements of the diagonal weight matrix \mathbf{W} were evaluated according to the relation

$$w_g^* = f_g^* / \sum_{g=1}^k f_g^*, \quad (12)$$

where

$$f_g^* = 1/s_{A_{g^*}}^2. \quad (13)$$

The values of the estimates of the standard deviations $s_{A_{g^*}}$ of the mixture absorbances at the g^* -the analytical position were calculated according to the relation

$$s_{A_{g^*}} = \left[\sum_{i=1}^m e_{A_{g^*}i}^2 / (m - n) \right]^{1/2}, \quad (14)$$

where $(m - n)$ represents the degree of freedom of the sum of squares of absorbance residues at the g -th analytical positions. The values $\sum_{i=1}^m e_{A_{g^*}i}^2$ are the elements of the main diagonal of the matrix $(\mathbf{E}_A \mathbf{E}_A^T)$, where

$$\mathbf{E}_A = \mathbf{K} \mathbf{C} - \mathbf{A}. \quad (15)$$

The model system. All calculations including the model system preparation, calculations of the calibration data and component concentrations in the model sample as well as the statistical evaluation of the results were performed on a computer GIER (Regnecentralen). The programs were ALGOL (version GIER-ALGOL 4).

The effect of the statistical weights on the precision and accuracy of analysis of a multi-component mixture was studied with a five-component model system, simulated by means of a computer. The preparation of the model system II, model system III, and the model sample has been described⁵.

RESULTS AND DISCUSSION

The averages of ten repeated determinations of concentrations of the components (j) in the model sample calculated according to the relations

$$\mathbf{X} = \mathbf{QB} \quad (16)$$

and

$$\mathbf{X} = \mathbf{Q}_w \mathbf{B} \quad (17)$$

are given in Tables I and II, respectively. In the two tables, the known concentrations, c_j , of the sample components, the standard deviations of the concentration determinations calculated according to the relation

$$s_j = \left[\sum_{i=1}^m (c_j - x_{ji})^2 / (m - 1) \right]^{1/2} \quad (18)$$

and the relative accuracy of concentration (%),

$$\delta_j = 100 \cdot (c_j - \bar{x}_j) / c_j, \quad (19)$$

are given as well.

A comparison of the results of Tables I and II indicates that the introduction of the weights is accompanied by a considerable improvement of the precision and accuracy of determination of the component concentrations for the model system II, while for the model system III no appreciable changes appear.

In the model system III, random errors were added to the absorbances at all analytical positions. These errors were picked out in a random way from a set possessing the Gauss distribution. When this normal distribution of errors exists, the least squares method affords correct result along with the error estimate, and thus the introduction of the weights, which approached unity, virtually did not affect the concentration determination.

In the model system II the random errors were added to absorbances at five analytical positions only. These selected analytical positions represented absorption band maxima of the individual components. This corresponds to the case, encountered in practice, where systematic errors occur at particular analytical positions; for instance, the sample may contain components not considered during the calibration. The errors of the model system II do not follow a Gauss distribution, which showed up in the various weights differing as much as by several orders of magnitude. This means that the analytical positions, to which the errors were added, are virtually not taken into account.

TABLE I

Evaluation of the Analysis of the Model Sample
The relation $\mathbf{X} = \mathbf{QB}$ is used for the evaluation.

Model system	Parameter	Component, j				
		1	2	3	4	5
II	c_j	0.6500	0.1000	0.0500	0.0500	0.1500
	\bar{x}_j	0.6488	0.1026	0.0478	0.0537	0.1497
	s_j	0.0041	0.0064	0.0032	0.0081	0.0055
	δ_j (%)	0.19	-2.63	4.44	-7.30	0.21
III	\bar{x}_j	0.6499	0.1022	0.0513	0.0483	0.1479
	s_j	0.0116	0.0067	0.0039	0.0066	0.0088
	δ_j (%)	0.01	-2.21	-2.60	3.44	1.38

TABLE II

Evaluation of the Analysis of the Model Sample
The relation $\mathbf{X} = \mathbf{Q}_w\mathbf{B}$ is used for the evaluation.

Model system	Parameter	Component, j				
		1	2	3	4	7
II	c_j	0.6500	0.1000	0.0500	0.0500	0.1500
	\bar{x}_j	0.6500	0.0997	0.0500	0.0499	0.1501
	s_j	0.0000	0.0003	0.0000	0.0001	0.0001
	δ_j (%)	0.00	0.27	0.07	0.13	-0.09
III	\bar{x}_j	0.6510	0.1020	0.0515	0.0484	0.1478
	s_j	0.0092	0.0069	0.0041	0.0061	0.0080
	δ_j (%)	-0.15	-2.03	-3.06	3.26	1.47

Our experience warrants the statement that if the degree of overdetermination in the number of analytical positions is sufficiently high, it is more convenient to use some of the methods of selection of the optimum analytical positions⁵⁻¹⁰, thereby reducing this degree to about a threefold or a fourfold overdetermination. If the number of the analytical positions available is not high enough for a sufficient over-

determination, or if systematic errors appear, it is more convenient and more correct to apply the above calculation procedure introducing the statistical weights.

LIST OF SYMBOLS

- $g, g^* = 1, 2, \dots, k$ sequence of the analytical position (wavelength, wavenumber)
 $i = 1, 2, \dots, m$ sequence number of the calibration mixture or sample
 $j, j^* = 1, 2, \dots, n$ sequence of the component
 k total number of analytical positions (wavelengths, wavenumbers)
 m total number of calibration mixtures or samples
 n total number of components
A absorbances matrix of the calibration mixtures ($k \cdot m$ matrix)
B absorbances matrix of the sample absorbances ($k \cdot m$ matrix)
C matrix of the component concentrations in the calibration mixtures ($n \cdot m$ matrix)
 E_A matrix of the absorbance residues of the calibration mixtures ($k \cdot m$ matrix)
 E_A^T matrix transposed to the matrix E_A ($m \cdot k$ matrix)
 E_B matrix of the absorbance residues of the samples ($k \cdot m$ matrix)
K absorptivities matrix ($k \cdot n$ matrix)
 K^T matrix transposed to the matrix **K** ($n \cdot k$ matrix)
Q matrix for the calculation of the concentrations ($n \cdot k$ matrix)
 Q_W matrix for the calculation of the concentrations involving the weight factors ($n \cdot k$ matrix)
W weight matrix ($k \cdot k$ matrix)
X matrix of the component concentrations in the samples ($n \cdot m$ matrix)
b absorbances vector of the sample
 e_b vector of the absorbance residues of the samples
x vector of the component concentrations in the sample
 b_g element of the vector **b** (sample absorbance at the g -th analytical position)
 c_j the known concentration of the j -th component
 e_{Agi} element of the matrix E_A (absorbance residue of the i -th calibration mixture at the g -th analytical position)
 $e_{b,g}$ element of the vector e_b (absorbance residue at the g -th analytical position)
 k_{gj} element of the matrix **K** (absorptivity of the j -th component at the g -th analytical position)
 w_g element of the matrix **W** (statistical weight of the absorbance measurement at the g -th analytical position)
 x_j the determined concentration of the j -th component
 \bar{x}_{ji} element of the matrix **X** (concentration of the j -th component in the i -th sample)
 \bar{x}_j average of the determined concentrations of the j -th component
 s_{Ag} estimate of the standard deviation of the absorbance at the g -th analytical position
 s_j standard deviation of determination of the concentration of the j -th component with respect to the known concentration c_j
 δ_j relative accuracy of the concentration mean of the j -th component with respect to the known concentration c_j

REFERENCES

1. Vasilev A. F.: *Zavod. Lab.* 31, 677 (1965).
2. Przybylski Z.: *Chem. Anal. (Warsaw)* 13, 453 (1968).
3. Šustek J., Livař M., Schiessl O.: *Chem. Listy* 66, 168 (1972).

4. Junker A., Bergmann G.: *Fresenius' Z. Anal. Chem.* 272, 267 (1974).
5. Šustek J.: *Anal. Chem.* 46, 1676 (1974).
6. Wojdala T.: *Chem. Anal. (Warsaw)* 13, 289 (1968).
7. Przybylski Z.: *Chem. Anal. (Warsaw)* 14, 1047 (1969).
8. Kac M. D.: *Zh. Prikl. Spektrosk.* 18, 508 (1973).
9. Kac M. D.: *Zavod. Lab.* 39, 160 (1973).
10. Junker A., Bergmann G.: *Fresenius' Z. Anal. Chem.* 278, 191 (1976).

Translated by P. Adámek.