

ON THE EVALUATION OF HOMOGENEITY OF STANDARDS

P. BOHÁČEK, J. NOVÁK and N. RYŠAVÁ

*Physical Institute,
Czechoslovak Academy of Sciences, Prague 8*

Received February 17th, 1971

The contribution of inhomogeneity of a standard alloy sample to the total variance of values measured analytically is evaluated and the dependence of the variance of inhomogeneity on the size of the sampling used for analysis is determined.

As a consequence of the development of analytical methods which impose increasingly more severe requirements as to the correctness and exactness of data about composition of standards, the claims concerning their homogeneity become also more grave. Standard samples are necessary for the determination of the error relevant to an analytical method. However, when its accuracy attains the limit given by the possibilities of the manufacturer of standards so that the inaccuracy due to errors of measurement is comparable with that due to insufficient homogeneity of the standard, a relation between the mentioned inaccuracies is needed together with a corresponding terminology.

The errors in chemical measurements have been systematically dealt with in the literature¹, while attempts to evaluate statistically the material homogeneity are rare^{2,3} and some of them are only of an introductory character⁴. Definitions of basic ideas pertinent to the homogeneity of materials were already proposed⁵.

We therefore attempted to evaluate quantitatively the variance of measured values, *i.e.* to distinguish the part of the variance due to the inaccuracy of measurement from that due to material inhomogeneity. Furthermore, we derived and verified experimentally the dependence of the variance of inhomogeneity on the size of the sampling.

If the concentration of an element in a set of n samplings of the analyzed material is determined, the average value and standard deviation σ or variance σ^2 can be determined. The latter is given as

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n x_i^2, \quad (1)$$

where x_i denotes deviation of i -th measured value from the average and is equal to the sum of the error of measurement, x_{mi} , and a quantity, x_{hi} , called for convenience "deviation in composition". The latter involves the deviation of the true content of the element to be determined from the average composition as well as the deviation of all decisive material factors from the average. On introducing x_{mi} or x_{hi} into Eq. (1) we obtain the quantity σ_m^2 characterizing the accuracy of the method of measurement and σ_h^2 characterizing the inhomogeneity of the material or, more precisely, the sum of composition and interference inhomogeneities⁵. We shall denote σ_m^2 as variance of the measuring method, σ_h^2 variance of inhomogeneity and σ^2 variance of measured values or total variance.

In the present paper we derive the relation among the quantities σ^2 , σ_h^2 and σ_m^2 and propose a method of their separation; further we determine both theoretically and experimentally the dependence of σ_h^2 on the size of the sampling, *i.e.* amount of material used for a single determination⁵.

THEORETICAL

We assume a distribution, $f(x_h)$, of relative frequencies of deviations x_h in the mentioned set, which may but need not correspond to the Gauss function. On introducing x_h into Eq. (1) and replacing summation by integration, we obtain

$$\sigma_h^2 = \int_{-\infty}^{+\infty} x_h^2 f(x_h) dx_h. \quad (2)$$

The distribution $f(x_h)$ as well as the variance σ_h^2 apply for a measurement without errors. In reality, the errors of measurement change this distribution so that every column $f(x_h) dx_h$ is scattered around x_h according to the Gauss distribution:

$$g(y - x_h) = (2\pi\sigma_m^2)^{-1/2} \exp[-(y - x_h)^2/2\sigma_m^2].$$

In the point y , the contribution of the original relative number of deviations in the point x_h will be $f(x_h) g(y - x_h) dx_h$. The resulting variance after the measurement will be in analogy with Eq. (2) given by

$$\sigma^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y^2 f(x_h) g(y - x_h) dx_h dy = \int_{-\infty}^{+\infty} f(x_h) \left(\int_{-\infty}^{+\infty} y^2 g(y - x_h) dy \right) dx_h.$$

We set $y^2 \equiv (y - x_h)^2 - x_h^2 + 2x_h y$ and make use of the fact that Eq. (2) gives for a distribution g a variance σ_m^2 and that the integral of the distribution function is equal to one, whereas the integral of the odd function $z g(z)$ is equal to zero:

$$\begin{aligned} \int_{-\infty}^{+\infty} y^2 g(y - x_h) dy &= \int_{-\infty}^{+\infty} (y - x_h)^2 g(y - x_h) dy - x_h^2 \int_{-\infty}^{+\infty} g(y - x_h) dy + \\ &+ 2x_h \int_{-\infty}^{+\infty} y g(y - x_h) dy = \sigma_m^2 - x_h^2 + 2x_h \int_{-\infty}^{+\infty} (z + x_h) g(z) dz = \sigma_m^2 + x_h^2. \end{aligned}$$

Thus, we obtain the necessary relation

$$\sigma^2 = \int_{-\infty}^{+\infty} f(x_h) (\sigma_m^2 + x_h^2) dx_h = \sigma_m^2 + \sigma_h^2. \quad (3)$$

Here only the variance σ^2 of the measured values is known. To determine σ_h^2 and σ_m^2 we shall derive the dependence of σ_h^2 on the size of the sampling. We shall assume that

the distribution of deviations x_h from homogeneity (further denoted as x) is given by the Gauss function $f(x)$. We shall consider two sets, A and B , of equal number of samples of equal size. If the number of samples is sufficiently large, their mean values, variance and distribution functions are the same in each set: $\sigma_h^2(A) = \sigma_h^2(B)$ and

$$f_A(x) = f_B(x) = [2\pi\sigma_h^2(A)]^{-1/2} \exp[-x^2/2\sigma_h^2(A)].$$

On pairing one sample from the group A with one randomly chosen from B we obtain a set of double-sized samples which we denote as $A + B$ and its variance as $\sigma_h^2(A + B)$. If the deviations from the average for the first and second samples are x_A and x_B , after their pairing, the resulting deviation will be $x = \frac{1}{2}(x_A + x_B)$. Introducing an auxiliary variable Δ by $x_A = x - \Delta$ (and hence $x_B = x + \Delta$) we obtain for the distribution of relative frequencies of the deviation x in the set $A + B$ the expression

$$f_{A+B}(x) = \frac{\int_{-\infty}^{+\infty} f_A(x - \Delta) f_B(x + \Delta) d\Delta}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_A(x - \Delta) f_B(x + \Delta) d\Delta dx} =$$

$$= (\pi\sigma_h^2(A))^{-1/2} \exp[-x^2/\sigma_h^2(A)] = [2\pi\sigma_h^2(A + B)]^{-1/2} \exp[-x^2/2\sigma_h^2(A + B)].$$

Hence $\sigma_h^2(A + B) = \frac{1}{2}\sigma_h^2(A)$, the variance of doubly sized samples is one half as large. Analogously, it can be shown that

$$\sigma_h^2(kA) = \sigma_h^2(A)/k. \quad (4)$$

This equation can be used to determine the unknowns σ_h^2 and σ_m^2 in Eq. (3) as follows. We select from the analyzed material a set of samples of the size (weight) kA . The variances $\sigma^2(A)$ and $\sigma^2(kA)$ of the measured values fulfil the Eq. (3):

$$\sigma^2(A) = \sigma_h^2(A) + \sigma_m^2(A), \quad \sigma^2(kA) = \sigma_h^2(kA) + \sigma_m^2(kA). \quad (5)$$

The relation between $\sigma_h^2(A)$ and $\sigma_h^2(kA)$ is given by Eq. (4). The amounts A and kA are, of course, chosen so that the analysis can be carried out by the same method. Then the error of the determination is the same in both cases: $\sigma_m^2(A) = \sigma_m^2(kA)$. Thus the second Eq. (5) takes the form $\sigma^2(kA) = \sigma_h^2(A)/k + \sigma_m^2(A)$ and the unknowns $\sigma_h^2(A)$ and $\sigma_m^2(A)$ can be expressed as

$$\begin{aligned} \sigma_h^2(A) &= k[\sigma^2(A) - \sigma^2(kA)]/(k - 1), \\ \sigma_m^2(A) &= [\sigma^2(A) - k\sigma^2(kA)]/(1 - k). \end{aligned} \quad (6)$$

The remaining unknown $\sigma_n^2(kA)$ is calculated from Eq. (4). Eqs (6) enable to determine separately the variances of inhomogeneity, σ_n^2 , and of the measuring method, σ_m^2 . However, the assumptions inherent in the derivation of Eq.(4) must be observed, e.g. that of the Gauss distribution of the deviations in composition, $f(x_n)$. Examining this, one must take into account that a solid material (including standards) contains relatively homogeneous regions of the size of grains or smaller, the composition of which can vary appreciably (precipitates of admixtures in the matrix). If the dimensions of the analyzed samples are comparable with those of the mentioned regions or are even smaller (microprobe) then the measured values correspond mostly to individual phases in the standard so that the distribution $f(x_n)$ is more of a discrete character and cannot be even approximately considered as normal (Gauss type).

If at least one dimension of the sample is so large as to involve a number of the mentioned regions, the distribution $f(x_n)$ loses its discrete character and approaches the more the normal one the larger is the dimension of the sample. As a rough estimate, at least one dimension of the sample should be by an order of magnitude larger than the mean distance of precipitates differing by the concentration of the element to be determined. A more detailed analysis of the structural problems will form the subject of a further work.

The second assumption is implicitly involved in the derivation of Eq. (4): the sample of the set $A + B$ is formed by combining two parts the composition of which is independent of each other. This means that the sample of the size kA can be divided into k parts the composition of which is independent of one another and the shape of which is the same as with samples of the set A . This condition can be easily fulfilled if all three dimensions of the sample are sufficiently large. If, however, one of them is not large enough in the sense of the preceding paragraph (flat samples – quantometer), one must take into account that the composition of two samples taken close to each other in the direction of the small dimension can be more or less interrelated since both samples can involve a certain number of the same grains. In such a case it is necessary before applying Eq. (4) to check independence of the neighbouring samples in advance. For example, two subsequent samplings are taken in a number of points on a given material. From the results one calculates the variance σ_1^2 of the set of first samplings and the variance σ_{12}^2 of the differences between the results of analyses of both samplings taken in the same point (again for all points). If $\sigma_{12}^2 = 0$, the neighbouring samplings are entirely interdependent. If $\sigma_{12}^2 = 2\sigma_1^2$, they are entirely independent.

If the condition of independence (with flat samples) is not fulfilled its fulfilment can be achieved by modifying the procedure so that the set kA will consist of samples of equal thickness as the set A but of an area larger by the factor of k . This case comes into consideration especially with methods which do not require taking samples (X-ray fluorescence).

A method in which the assumptions involved in Eq. (4) are fulfilled is the wet

analysis, where the sample weights are several tens of milligrams or more, so that it can be assumed that they contain a sufficiently large number of grains. The shape of the sample is isometrical. In this case Eq. (4) gives the dependence of the variance of inhomogeneity on the size of the sample. Eqs (6) are also fulfilled so that variances σ_h^2 and σ_m^2 can be separated. However, in this case it is not the only way since σ_m^2 can be determined separately by performing a series of measurements on the same stock solution. Nevertheless, the method of separation of σ_h^2 and σ_m^2 based on Eqs (6) is more general and principally applicable even with analytical methods which do not make use of a homogeneous stock solution.

EXPERIMENTAL

To check the derived dependence of the variance of inhomogeneity on the size of the sample we used as a material to be analyzed an alloy "silumin" of the composition (in percent) Cu 0.81, Si 10.08, Mn 0.6, Fe 1.02, Ni 1.27, Mg 0.94, Ti 0.015, Zn 0.66 and the remainder Al. We determined the content of Cu, Mn and Fe from samplings of three sizes, 0.04, 0.2 and 1.0 g, photometrically 15 times for each size. Copper was determined by its reaction with tetraethylthiuramdisulphide in ethanolic solution⁶, iron with thiocyanate⁷ and manganese with formaldoxim, the Fe(III) ions being screened with cyanide⁸. Since our measurements showed this masking to be unsatisfactory the results were checked by titration with periodate⁶. The alloy was dissolved in a mixture of hydrochloric and nitric acids, evaporated with perchloric acid until white fumes were evolved, the undissolved silicon was filtered off, the filter washed and the filtrate made up to 100, 500 and 1000 ml with the 0.04, 0.2 and 1.0 g samples, respectively, to obtain stock solutions for the analyses.

From the measured values we calculated the averages and variances $\sigma^2(0.04)$, $\sigma^2(0.2)$ and $\sigma^2(1)$ according to Eq. (1). The variance of inhomogeneity, $\sigma_h^2(0.04)$, was calculated from the couple $\sigma^2(0.04)$ and $\sigma^2(0.2)$, $\sigma_h^2(0.2)$ from $\sigma^2(0.2)$ and $\sigma^2(1)$, and $\sigma_h^2(1)$ from $\sigma^2(1)$ and $\sigma^2(0.04)$ according to Eqs (6) and (4). For each couple, the variance of the measurement, $\sigma_m^2(AB)$, was also calculated. Its average value, σ_m^2 , for a given element is compared with the variance σ_M^2 of a series of 15 measurements carried out on a single stock solution of silumin. The results are shown in Table I. Both used methods of determination of the variance of measurement lead to nearly identical results ($\sigma_m^2 \approx \sigma_M^2$). The dependence of the variance of inhomogeneity, σ_h^2 ,

TABLE I
Measured and Calculated Variances (parts per 10^4)²

Variance	Cu	Fe	Mn	Variance	Cu	Fe	Mn
$\sigma^2(0.04)$	75.45	41.82	119.2	$\sigma_m^2(0.04, 0.2)$	10.55	18.47	15.65
$\sigma^2(0.2)$	23.57	23.15	36.7	$\sigma_m^2(0.2, 1)$	8.25	7.65	34.67
$\sigma^2(1)$	11.31	10.75	35.0	$\sigma_m^2(0.04, 1)$	8.64	9.46	31.48
$\sigma_h^2(0.04)$	64.90	23.35	103.55	σ_m^2	9.17	11.86	27.27
$\sigma_h^2(0.2)$	15.32	15.50	1.70	σ_M^2	9.55	11.80	21.40
$\sigma_h^2(1)$	2.67	1.29	3.52				

on the size of the sample corresponds to Eq. (4) only with copper. With iron and manganese, the measured dependence differs from the theoretical, probably because 15 measurements is too little to make a statistics. Nevertheless, assuming that the true form of the dependences, which we measured only roughly, is the same for all three elements (except for a constant coefficient), we can calculate their average course relative to the absolute values obtained for copper:

$$\sigma_h^2(0.04) = 68.48 \quad \sigma_h^2(0.2) = 14.03 \quad \sigma_h^2(1) = 3.01,$$

$$\sigma_h^2(0.04)/\sigma_h^2(0.2) = 4.88 \quad \sigma_h^2(0.2)/\sigma_h^2(1) = 4.66.$$

Both last values are very close to $k = 5$ calculated from Eq. (4). Therefore, we conclude that the experimental results are in spite of the limited number of measurements in accord with Eq. (4) derived theoretically.

The dependence of the variance of inhomogeneity on the size of the sample can be, in our opinion, made use of in practice since it enables the material intended as an analytical standard to be more properly characterized by its manufacturer — mainly for wet analysis. Namely, it is sufficient to indicate the variance of inhomogeneity, σ_h^2 , of the given material for a certain size of the sample (e.g. 0.1 g). For another size, the analytical chemist can calculate the value of σ_h^2 from Eq. (4) and after its subtraction from the total variance of measured values, σ^2 , obtain the variance of the measuring method, σ_m^2 . Furthermore, the value of $\sigma_h^2(0.1)$ indicated on the standard (the variance of inhomogeneity for a 0.1 g sample) would enable to select both a suitable analytical method and a size of the sample corresponding to the desired degree of accuracy.

REFERENCES

1. Eckschlager K.: *Errors of Measurement and Results in Chemical Analysis*. Van Nostrand — SNTL, London—Prague 1969.
2. Yackowitz H., Vieth D. L., Heinrich K. F. H., Michaelis R. E.: *Nat. Bur. Standards, Misc. Publ.* 260—10 (1965).
3. Yackowitz H., Vieth D. L., Michaelis R. E.: *Nat. Bur. Standards, Misc. Publ.* 260—12 (1966).
4. Hoffman J. I.: *Anal. Chem.* 31, 1934 (1959).
5. Novák J., Bičovský K.: *Hutnické listy* 26, 66 (1971).
6. Bičovský K., Kutý V.: *Soubor analytických předpisů pro metalurgické laboratoře*, p. 89. Výzkumný ústav ČKD, Prague 1966.
7. ČSN 420670, *Chemický rozbor hliníku a slitin hliníku*.
8. Van Schouwenburg J. C.: *J. Agric. Sci.* 14, 127 (1966).

Translated by K. Micka.