

# Characterization of the chemical homogeneity of solid-state standard materials by chemometric methods\*

V. Liebich, G. Ehrlich, U. Stahlberg, and W. Kluge

Academy of Sciences of the GDR, Central Institute of Solid-State Physics and Materials Research, Helmholtzstrasse 20, DDR-8027 Dresden, German Democratic Republic

## Homogenitätscharakterisierung fester Standardproben unter Verwendung chemometrischer Methoden

**Summary.** For characterizing the homogeneity of a copper standard for optical emission spectrometry, the results of spark source mass spectrographic (SSMS) milliprobings have been treated by diverse chemometric methods. The proposed strategy includes visualization of the data, global homogeneity testing for each element, testing the anisotropy of elemental distributions, search for correlation between elements, and statements on the homogeneity with respect to other analytical procedures. The copper standard was proved to be non-homogeneous for SSMS but homogeneous for techniques of optical emission spectroscopy having larger sampling volumes.

## 1 Introduction

Nowadays, distribution analysis is one of the most outstanding fields of solid-state analysis. It comprises the determination of concentration profiles (e.g. depth-profiles) as well as the analysis of preselected small regions of a sample (micro-areas, surfaces, interfaces). Characterization of the chemical homogeneity may be considered as the lowest level of distribution analysis as far as it aims only at stating whether there are significant deviations from the mean composition within the samples and – in some cases – to give rough information on the “structure” of inhomogeneity, e.g. stochastic or periodical fluctuations or monotonous concentration gradients called “trends”.

When investigating the homogeneity of a solid, one has to bear in mind that homogeneity is a relative property of a solid-state material depending on the parameters of the analytical test procedure and on the ultimate purposes of the material, respectively [1, 2].

Recently, various approaches to computer-aided evaluation of spatially resolved analytical measurements for detecting chemical inhomogeneities have been published [3–9]. It results that each of them has its particular

advantages and restrictions. Therefore, several variants should be applied complementarily to get relevant information.

In the following some well-known and some unconventional approaches will be applied to characterize the homogeneity of a standard material. Generally, a standard material is considered to be homogeneous, if the variance of analytical results, obtained by repetitive sampling, does not exceed significantly the procedural random error. That means the aptitude of a given material as standard material depends on the precision and the spatial resolution (e.g. sampling volume) of the procedure to be calibrated. Thus, often the question arises, whether a standard material certified for a given analytical procedure may be used for another one having different procedural parameters.

As an example a reference material for optical emission spectrometry has been tested with respect to its suitability for three spectroscopic techniques of different precision and sampling volume: conventional spark source mass spectrography (SSMS), SSMS milliprobings, and glow-discharge optical emission spectrography (GD-ES).

## 2 Methodological

### 2.1 Experimental

The disc-shaped copper standard sample under investigation (CuII/2, ASMW/DDR) contains Cr, Mn, Fe, Ni, Zn, Ag, Sn, Sb, Pb, and Bi in the concentration range between 15 and 600 ppma.

For homogeneity characterization of this standard sample the SSMS milliprobe technique was used under working conditions listed in Table 1. When analysing this standard sample by conventional SSMS (i.e. sparking between two sample rods), sampling volume and procedural standard deviation are comparable with that of SSMS milliprobings. However, a considerably larger sampling volume and a lower procedural error are valid in GD-ES.

Figure 1 shows the arrangement of the measuring points (sub-samples) along polar co-ordinates ( $\alpha$ ,  $r$ ) for SSMS milliprobings. Since relative concentration values are sufficient for homogeneity characterization, quantification was done according to [10, 11]. The data sets obtained in this way for Bi, Ni, Ag, and Zn, as representatives of the 10 elements included in this study, are given in Table 2.

\* Dedicated to Prof. Dr. G. Tölg on the occasion of his 60th birthday

Offprint requests to: V. Liebich

**Table 1.** SSMS conditions

<i>Analysis facility</i>	
Mass spectrometer	MS 7 (AEI Manchester)
Ionization	rf spark source
<i>Electrode configuration</i>	
Sample electrode	point to plane
Probe electrode	plane
Gap	tantalum tip
	≤ 0.08 mm
<i>Operating parameters</i>	
Main slit width	0.04 mm
Spark variac	40%
Spark pulse length	0.1 ms
Spark pulse repetition rate	30 Hz
Accelerating voltage	20 kV
Electrostatic analyser	2 kV
Pressure in the analyser	≤ 10 <sup>-4</sup> Pa
Exposition	1.5 × 10 <sup>-9</sup> As
<i>Detection</i>	photoplates ORWO UV 2
<i>Crater dimensions</i>	
Diameter	0.49 mm
Depth	0.034 mm
Volume	0.0064 mm <sup>3</sup>
<i>Analytical lines</i>	
	<sup>52</sup> Cr <sup>+</sup> <sup>55</sup> Mn <sup>+</sup>
	<sup>56</sup> Fe <sup>+</sup> <sup>60</sup> Ni <sup>+</sup>
	<sup>67</sup> Zn <sup>+</sup> <sup>107</sup> Ag <sup>+</sup>
	<sup>120</sup> Sn <sup>+</sup> <sup>121</sup> Sb <sup>+</sup>
	<sup>208</sup> Pb <sup>+</sup> <sup>209</sup> Bi <sup>+</sup>
<i>Reference line</i>	<sup>50</sup> V <sup>+</sup> (external standard)
<i>Procedural RSD</i>	0.2

## 2.2 Chemometrical

For elaborating a strategy widely applicable to analogous practical problems of homogeneity characterization, it is assumed that only one set of measurements without replicates is available. The strategy will be divided into five steps:

- visualization of the measured data set for each element, e.g. by a so-called three-dimensional (“3D”) plot or as isoconcentration lines,
- testing the global homogeneity for each element,
- testing anisotropy of the elemental distributions,
- search for correlation between the elements,
- extrapolation of results to other analytical procedures.

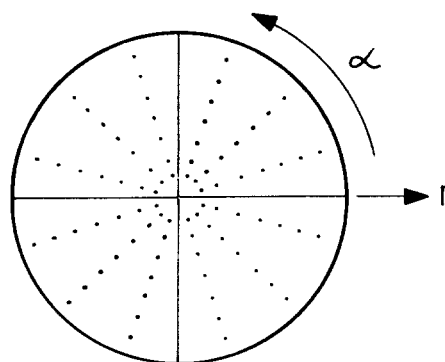
Only peculiarities of these steps will be noted here. All statistical tests were carried out assuming normal distributed, statistical independent measurements, a probability of 0.95, and stability of the analytical procedure.

### 2.2.1 Visualization

In our case, the 3D-plot – a diagonal parallel projection – is based on the “hidden line” algorithm. Isoconcentration lines were plotted using cubic splines.

### 2.2.2 Global tests

Global tests may be carried out by methods used in time series (“process”) analysis. Therefore, each data matrix of Table 2 was linearized into a “vector” having 84 “components”. This was done by joining either column by column (called here process of type I) or row by row (process of

**Fig. 1.** Subsample arrangement for SSMS milliprobng

type II). In both cases, the data vector of a homogeneous sample realizes a stochastic process. The test to be applied depends on the structure of the inhomogeneity.

*Trends* in a process can be revealed by the CUSUM (cumulative sum) technique [12]. It is based on the fact that a plot of the cumulative sum of the differences between single readings  $c_i$  and a “target value”  $c$  scatters about  $c$  so long as the process remains stationary. When, however, some consecutive readings systematically deviate from  $c$ , then the cusum will steeply increase/decrease, thus indicating a possible trend earlier than the original values  $c_i$ . Using the significance coefficient [13]

$$H_c = \frac{D_{\text{ext}}}{(0.165n + 1.71) \sqrt{\frac{\sum_{j=1}^n (D_j - D_{j+1})^2}{2(n-1)}}}, \quad (1)$$

where  $D_j = \sum_{i=1}^j (c_i - c)$  is the cusum,  $D_{\text{ext}} = \text{Max}(|D_j|)$  the most extreme cusum, and  $n$  the total number of readings  $c_i$ , the trend is significant, i.e. the sample is non-homogeneous, if holds  $H_c > 1$ . When applying CUSUM to our problem – contrary to a real time series – the process can be evaluated retrospectively, and thus all parameters in Eq. (1) are easily to be computed. Particularly, the sample mean  $\bar{c}$  may be used as an estimator for  $c$ .

*Periodicities* in a process can be treated by means of the theory of generalized stochastic processes. Therefore, the computer program STOPRO had been developed to evaluate data from a microprobe linescan considered as a realization of a stochastic process [14]. The current version [15] enables visualization of the autocorrelation function and the periodogram (spectral density after Fourier transformation) which hint to possible periods. Furthermore, computed periods are tested for significance. Thus, periodicities are revealed which can never be discerned directly from the original data.

Applied to our “puzzled” data vector, periodicities may be caused by trends along the co-ordinate axes. Nevertheless, any significant period means sample inhomogeneity.

*Noise* (stochastical fluctuations) in a stationary process can be proved by testing the relative standard deviation (RSD) of the whole data set  $s_r$  against a given RSD of the analytical procedure  $\sigma_r$  for the element under investigation

**Table 2.** Data matrices: concentrations/ppma of Bi, Ni, Ag, and Zn measured at different positions (polar co-ordinates)

r/mm	$\alpha$											
	15°	45°	75°	105°	135°	165°	195°	225°	255°	285°	315°	345°
<b>Bi</b>												
1	28	46	55	45	60	36	42	98	63	54	59	48
4	45	46	44	38	35	48	50	65	42	41	50	46
7	36	32	34	33	46	39	45	33	33	41	29	34
10	31	24	18	28	19	43	34	15	17	32	16	41
13	22	19	17	31	31	30	21	17	55	19	20	42
16	33	26	25	39	26	35	27	19	40	34	24	32
19	37	21	26	57	24	41	24	26	25	39	51	38
<b>Ni</b>												
1	240	320	300	480	630	450	450	580	360	510	560	450
4	240	240	240	320	330	480	340	430	210	290	320	460
7	330	300	220	310	310	300	320	190	220	310	320	300
10	370	230	260	290	320	320	300	290	300	250	280	270
13	300	260	240	310	410	420	290	300	260	140	310	340
16	350	280	260	350	390	340	210	300	250	260	320	370
19	330	250	280	370	350	310	210	220	230	260	360	340
<b>Ag</b>												
1	340	680	500	870	630	780	660	920	1100	1000	520	1000
4	480	590	600	500	530	560	700	700	430	490	500	480
7	1100	820	750	510	580	550	660	320	410	470	450	380
10	1300	890	340	500	310	610	470	240	260	390	220	510
13	820	390	820	590	500	600	430	230	520	180	350	540
16	960	360	900	730	560	590	450	240	460	340	380	420
19	780	260	740	800	430	670	520	260	290	400	420	430
<b>Zn</b>												
1	720	570	630	990	830	1000	710	1000	930	1000	910	960
4	780	830	840	760	780	670	1100	1100	600	1000	1100	990
7	890	530	500	750	700	710	1000	440	600	950	960	630
10	760	430	340	710	370	830	690	420	440	560	400	720
13	470	390	370	590	550	620	570	370	780	420	540	860
16	750	370	460	840	570	720	580	410	860	620	550	620
19	610	300	430	940	520	800	560	390	430	710	850	710

("analyte"). Using the significance coefficient derived from  $\chi^2$ -test

$$H_x = \frac{s_r^2 \cdot f}{\sigma_r^2 \cdot \chi^2(P, f)} \quad (2)$$

$[\chi^2(P, f)$  quantile of the  $\chi^2$ -distribution, P probability, f degree of freedom (d. f.)], non-homogeneity is proved if holds  $H_x > 1$ .

### 2.2.3 Testing anisotropy

If homogeneity has been rejected by any global test, often the question arises, whether the deviations from homogeneity depend on the spatial direction.

This question can be answered by analysis of variance (ANOVA) or by regression [3]. These methods were applied to the data matrices as shown in Table 2 and to the transposed matrices of them.

*Two-way ANOVA* is based on the model

$$c_{ij} = \bar{c} + a_i + b_j + w_{ij} + e. \quad (3)$$

Here  $c_{ij}$  are the measured values,  $\bar{c}$  is the overall mean,  $a_i$  and  $b_j$  are contributions of the factors  $i$  and  $j$ ,  $w_{ij}$  represents

interaction of the two factors, and  $e$  is the random error of the analytical procedure.

If, as in our case, there are no replicates,  $w_{ij}$  and  $e$  are mixed to an error term  $e_{ij}$ . Therefore, the absence of interchanges between the factors must be proved by special tests [16], or one takes into account that the test may become less rigorous. The latter was done by us using the wide-spread computer program ABSTAT.

Even *one-way ANOVA* may be used to investigate the homogeneity in one spatial direction assuming homogeneity in the other one. Since this supposition is not fulfilled exactly in most cases, usually the test will be still more insensitive.

For ANOVA a significance coefficient, called "coefficient of homogeneity", was introduced by Danzer et al. [4] according to

$$H_k = F_k / F(P, f_1, f_2), \quad (4)$$

where  $F_k$  is the well-known value for  $k$ -way ANOVA ( $k = 1; 2$ ),  $F(P, f_1, f_2)$  the quantile of Fisher's  $F$ -distribution, P the probability,  $f_1$  and  $f_2$  the d.f.

The  $\chi^2$ -test applied separately to both directions is a further approach for testing anisotropy. In this case the

numerator of Eq. (2) is to be replaced by the numerator of the test value  $F$  of two-way ANOVA which reads – noted for  $j$  –

$$s_r = \frac{1}{\bar{c}} \sqrt{\frac{n_j \sum_{i=1}^{n_i} (\bar{c}_i - \bar{c})^2}{(n_i - 1)}} \quad \text{with } \bar{c}_i = \frac{1}{n_j} \sum_{j=1}^{n_j} c_{ij} \quad (5)$$

( $i = 1, 2, \dots, n_i$  indicates the  $r$  axis and  $j = 1, 2, \dots, n_j$  the  $\alpha$  axis or vice versa).

*Polynomial regression* is useful if trends in a direction have been suggested e.g. by visualized data. The sample is considered to be non-homogeneous, if the null hypothesis (all coefficients are equal to zero) is rejected in any order of the polynomial. Additionally, the polynom providing the maximum  $F$ -value and the minimum variance of residuals, respectively, approximates best the concentration profile along the direction under investigation, thus also describing the structure of inhomogeneity. At this, one has to bear in mind that the results of regression depend on the “starting point”, which especially in  $\alpha$ -direction may be chosen arbitrarily. In the case of regression, the significance coefficient  $H_R$  is calculated analogous to Eq. (4) using as numerator the  $F$ -value from regression

$$F_R = \frac{q^2}{1 - q^2} \cdot \frac{f_2}{f_1} \quad \text{with } f_1 = m, f_2 = n - m - 1, \quad (6)$$

where  $q$  is the correlation coefficient,  $f_1$  and  $f_2$  d.f.,  $m$  order of the polynomial ( $m = 1$  means a straight line),  $n$  the total number of measurements (data pairs of the analyte).

#### 2.2.4 Correlation between elements

A direct information on correlations between elements with respect to its spatial distribution in the sample will be provided by the *correlation matrix* containing the correlation coefficients of each element with each other.

The *hierarchical cluster analysis* (HCA) [17, 18] may serve for grouping the elements by their correlation. Agglomerative HCA was applied to a data matrix, in which each row contains the process vector of one element. Thus, the columns of the matrix represent the variables or features (concentrations in the sub-samples), whereas the rows correspond to the objects (elements). HCA was carried out using the distance measure derived from correlation coefficients

$$d_{jk} = 1 - \frac{1}{n} \sum_{i=1}^n x_{ij} x_{ik} \quad j, k = 1, 2, \dots, p \quad (7)$$

with

$$x_{ij} = \frac{c_{ij} - \bar{c}_{.j}}{\sqrt{\frac{\sum_{i=1}^n (c_{ij} - \bar{c}_{.j})^2}{n - 1}}}$$

where  $j, k$  indicate the  $p$  elements and  $i$  denotes one of the  $n$  subsamples (components of the data vector). From this definition follows  $d = 0$  for  $q = 1$  (strong correlation),  $d = 1$  for  $q = 0$  (no correlation) and  $d = 2$  for  $q = -1$  (“anti-correlation”). The single linkage mode used by us facilitates the interpretation of the dendrogram, since the linkage levels are directly proportional to the “distances” according to Eq. (7).

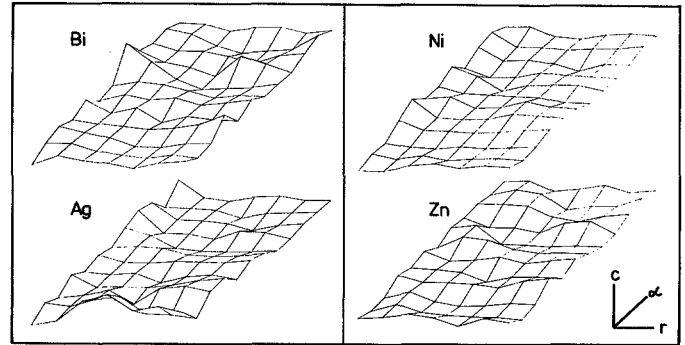


Fig. 2. 3D-plot of normalized (divided by sample mean) concentrations for the representative elements Bi, Ni, Ag, and Zn

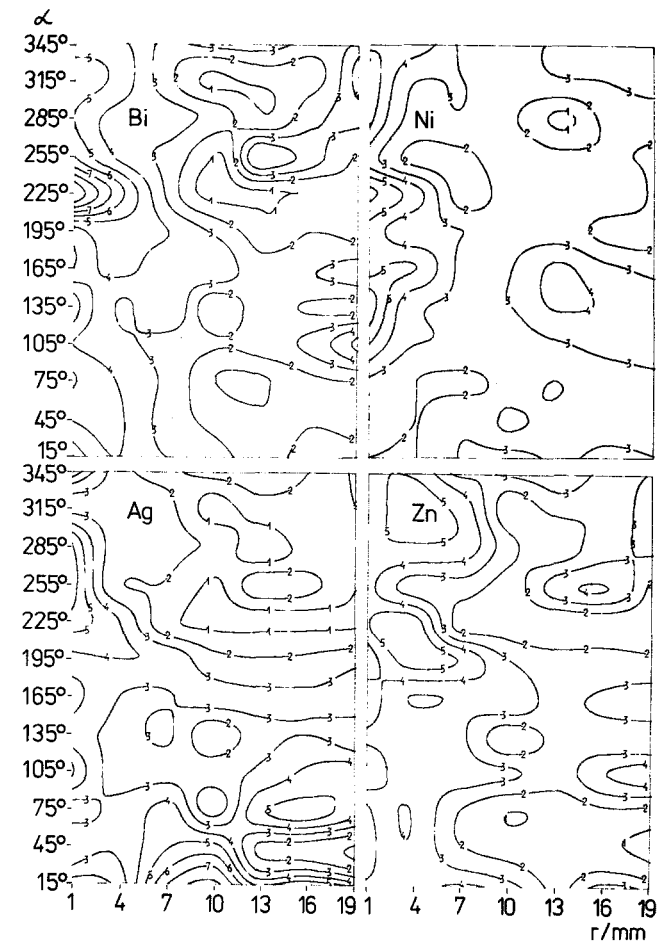


Fig. 3. Isolines plot of normalized concentrations. The isoline markers correspond to the following values of normalized concentrations: 1 0.5; 2 0.75; 3 1.0; 4 1.25; 5 1.50; 6 1.75; 7 2.0; 8 2.25; 9 2.5; 10 2.75

#### 2.2.5 Extrapolation of results to other analytical procedures

Principally, the RSD values obtained by a higher (spatially) resolving procedure may be transformed into corresponding values of a lower resolving procedure (or vice versa) by means of the program STOPRO [14]. However, in our case this program was unable to transform the RSD from SSMS to GD-ES, mainly because of the “low-pass filter action” due to the wide distances between the measuring points.

A useful approach in this case is the approximate equivalence of an increased sampling volume and an increased number of replicates. It yields

$$s_{r, inh, o} = \frac{s_{r, inh, m}}{\sqrt{V_o/V_m}} \quad (8)$$

with

$$s_{r, inh, m} = \sqrt{s_{r, tot, m}^2 - \sigma_r^2}$$

where m, o refer to the measurement procedure and the optional procedure, respectively,  $s_{r, tot}$  is the total RSD,  $s_{r, inh}$  the RSD from inhomogeneity,  $\sigma_r$  the procedural RSD, V the sampling volume.

### 3 Results and discussion

The 3D-plot (Fig. 2) and the isolines plot (Fig. 3) visualize the distributions of normalized (by sample mean divided) concentrations of Bi, Ni, Ag, and Zn in polar co-ordinates. The figures suggest that there may be some inhomogeneities, especially along the radial co-ordinate r.

The visual impression is objectively confirmed by the results of global tests (Table 3). For the process of type II, the CUSUM-technique as well as the  $\chi^2$ -test yield signifi-

cance coefficients > 1 for each element, the former proving significant trends in the "process", the latter showing overall RSDs significantly higher than procedural error of SSMS. Therefore, the general conclusion can be drawn that the sample is non-homogeneous with respect to SSMS. The CUSUM technique applied to the process of type I is less sensitive because of more alternating values due to the shorter "periods" in this case. Figures 4 and 5 illustrate

**Table 3.** Significance coefficients of global tests

Element	CUSUM-technique H <sub>c</sub>	$\chi^2$ -test H <sub>x</sub>
Cr	1.15	1.71
Mn	1.50	1.21
Fe	1.13	2.21
Ni	1.61	1.49
Zn	2.14	1.90
Ag	1.40	3.28
Sb	1.95	3.66
Sn	2.03	1.94
Pb	2.21	2.99
Bi	2.17	2.91

**Table 4.** Periods computed by analysis of stochastic processes

Process L	Period	frequency	Type I											Type II														
			Cr	Mn	Fe	Ni	Zn	Ag	Sn	Sb	Pb	Bi	Cr	Mn	Fe	Ni	Zn	Ag	Sn	Sb	Pb	Bi						
1	84.000	0.01190	*	+	+																	*	*		*	+	*	*
2	42.000	0.02381	+	+	+	*	+																					
4	21.000	0.04762	*				*	*																				
5	16.800	0.05952														+		+	-						+	+		
6	14.000	0.07143								*								+	-		+	+	+					
7	12.000	0.08333	*	+	*													+		+	+							
8	10.500	0.09524																							+	-	-	
10	8.400	0.11905	*																									
11	7.636	0.13095			+					+																		
12	7.000	0.14286	+	+		*	*	-	*	*	*	*																
13	6.462	0.15476								+																		
14	6.000	0.16667								-						+	*	+	*	*								
15	5.600	0.17857	+	+						+															+	+	+	-
16	5.250	0.19048	+	+	+																							
20	4.200	0.23810	+	+	+																							
22	3.818	0.26190																										
23	3.652	0.27381		+																								
24	3.500	0.28571	-	+	*	*																						
25	3.360	0.29762								+																		
26	3.231	0.30952																									+	+
27	3.111	0.32143																									+	+
28	3.000	0.33333													*		+	+	+									
29	2.897	0.34524															+	+	-						+	+	-	
30	2.800	0.35714								*																		
31	2.710	0.36905																										
32	2.625	0.38095	+																									
36	2.333	0.42857	*	-	+	*				-	-																+	
40	2.100	0.47619																										

\*  $P \geq 0.95$  (significant)  
 +  $0.95 > P > 0.5$   
 -  $P \leq 0.5$

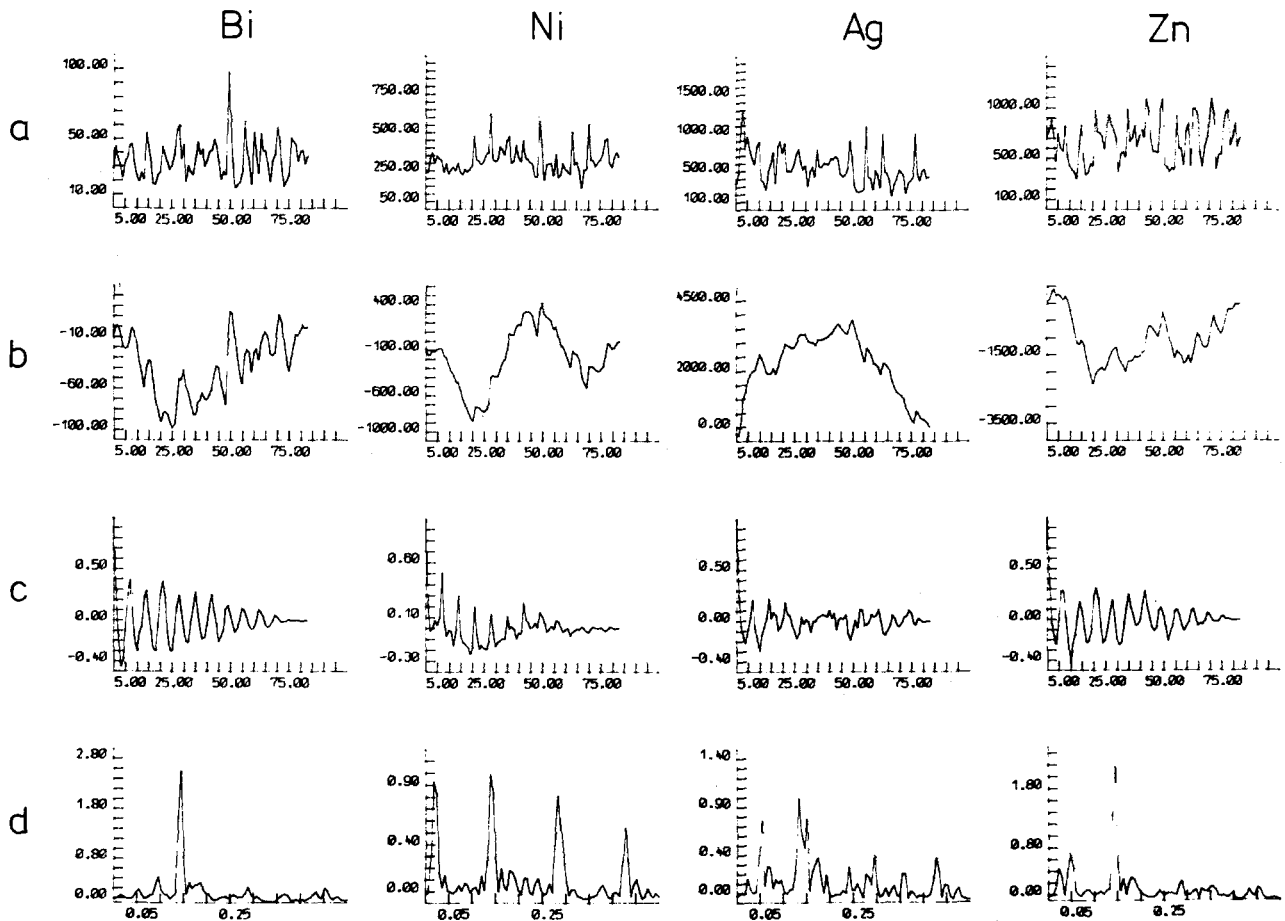
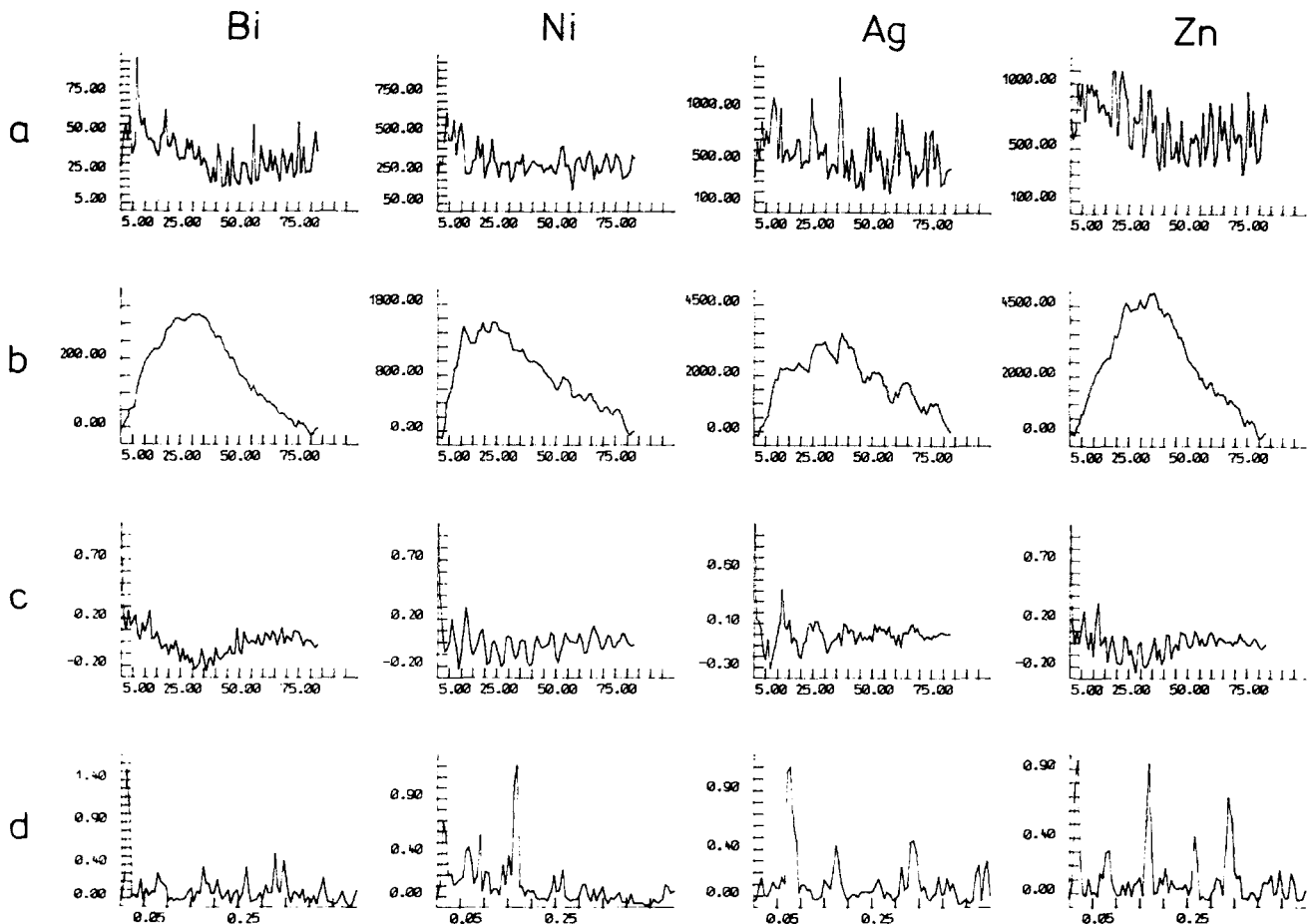


Fig. 4a-d. Process plots — type I (radial scans connected): a measured concentrations; b cumulative sums; c autocorrelation functions; d spectral densities (periodogram). Abscissa: sample distance of 1 for a-c, frequency for d



**Table 5.** Significance coefficients of tests in two spatial directions

Element	co-ordinate	1-way ANOVA $H_1$	2-way ANOVA $H_2$	re- gression $H_R$	$\chi^2$ -test $H_x$
Cr	r	1.45	1.96	1.57	2.84
	$\alpha$	1.47	1.85	0.79*	2.73
Mn	r	1.86	2.34	3.23	2.46
	$\alpha$	1.12	1.50	0.78*	1.60
Fe	r	1.35	1.78	1.07	3.49
	$\alpha$	1.42	1.74	0.02*	3.46
Ni	r	3.16	4.36	3.73	4.39
	$\alpha$	1.15	1.96	0.31*	2.00
Zn	r	3.46	4.70	6.74	5.94
	$\alpha$	1.06	1.88	1.24	2.41
Ag	r	0.90	1.06	1.65	3.68
	$\alpha$	1.13	1.24	3.45	4.36
Sb	r	2.20	2.08	4.20	8.39
	$\alpha$	0.28	0.37	0.31	1.50
Sn	r	3.44	3.23	6.05	6.02
	$\alpha$	0.22	0.34	0.06	0.64
Pb	r	4.50	3.98	7.04	10.89
	$\alpha$	0.09	0.15	0.04	0.40
Bi	r	4.66	4.35	7.05	10.81
	$\alpha$	0.19	0.32	0.50	0.80

Inhomogeneity proved for  $H > 1$ .

$H_R$  values tabled for linear regression (polynom of 1th order); the wild card (\*) marks significance in 4th and 5th order

Test values:

1-way ANOVA:

$F(0.95,6,77) = 2.21$  along r,  $F(0.95,11,77) = 1.91$  along  $\alpha$

2-way ANOVA:

$F(0.95,6,66) = 2.24$  along r,  $F(0.95,11,66) = 1.94$  along  $\alpha$

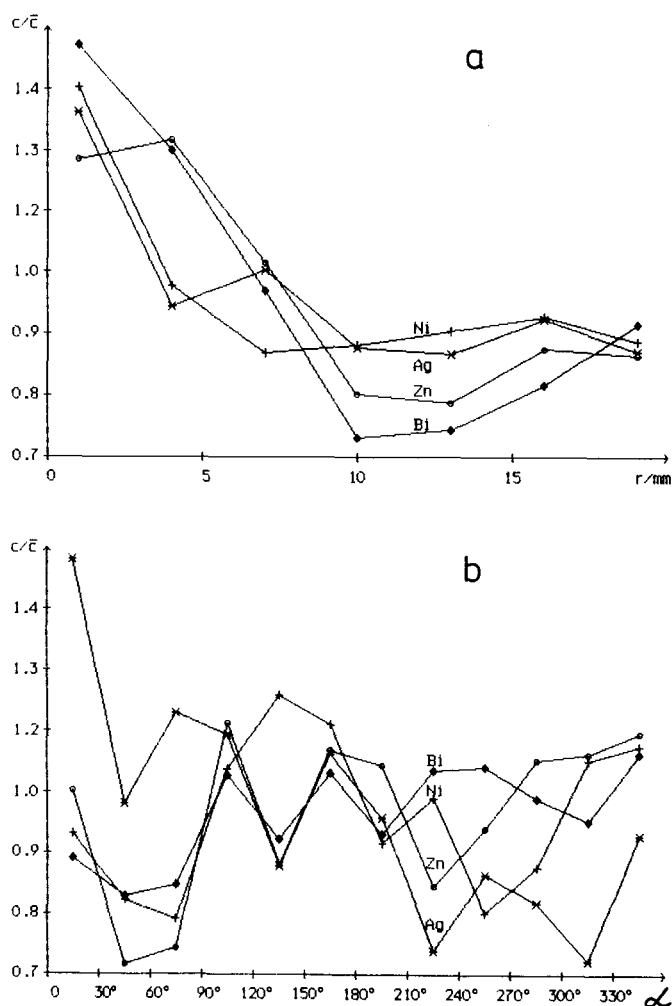
Regression:

$F(0.95,1,82) = 3.96$  for polynoms of the 1th order

$\chi^2$  test:

$\chi^2(0.95,6) = 12.59$  along r,  $\chi^2(0.95,11) = 19.68$  along  $\alpha$

for the process of type I and II, respectively, the measured concentration values (a), the cusum values (b), the autocorrelation functions (c), and the spectral densities (d) of Bi, Ni, Ag, and Zn. For a–c, the abscissae are scaled in subsample distances of 1. On this way the process ranges from 1 to 84. Only in the periodograms d the reciprocals, called frequencies, are used. Whereas the original concentration values (Figs. 4a and 5a) fluctuate nearly stochastically, the cumulative sums (Figs. 4b and 5b) as integrated values show smoothed, element-specific courses. In the process of type II all cusums turn one-sided to positive values. Contrary to it, for processes of type I almost the cusum values are more alternating and nearer to zero. The autocorrelation functions (Figs. 4c and 5c) hint more or less clearly to (spatial) “oscillations” in the processes. Some well-defined frequencies arise in the most periodograms (Figs. 4d and 5d). The detected frequencies were tested for significance. The results are summarized in Table 4. In accordance to [14], the detected periods range from 84 – the length of the process – down to the value of 2 – the twofold of the sampling distance. Moreover, all detected periods are self-explaining as upper harmonics of the value of 84 (84 divided by natural numbers  $l = 1, 2, \dots$ ). Obviously, the periodicities



**Fig. 6a, b.** Normalized concentration distributions: **a** mean values over azimuthal scans vs. radii r; **b** mean values over radial scans vs. azimuthal angles  $\alpha$

of 7 and 12 are induced by the “puzzling” of the process vector. That means the periods of 7 are caused by inhomogeneity along the co-ordinate r, whereas periods of 12 result from azimuthal inhomogeneities (along the co-ordinate  $\alpha$ ).

The individual tests in the two spatial directions  $\alpha$  and r (Table 5) prove the elemental distributions to be somewhat anisotropic. Significant deviations are detected for radial direction in each case with the only exception of 1-way ANOVA applied to Ag.

In azimuthal direction, however, the judgement about homogeneity differs from element to element and sometimes from method to method. There is no objection against homogeneity of the elements Sn, Pb, and Bi. For Sb in  $\alpha$ -direction the homogeneity is only rejected by the  $\chi^2$ -test. Ag and Zn are non-homogeneous by each method. For Cr, Mn, Fe, and Ni inhomogeneity in  $\alpha$ -direction is detected by all methods, if one accepts the results of regression in the 4th and 5th order.

The results of the tests in two spatial directions are better understandable by Fig. 6, showing the normalized concentrations averaged over  $\alpha$ , i.e. depicted vs. r (Fig. 6a), and vice versa (Fig. 6b). From Fig. 6a it becomes clearly that linear regression (1th order) along the r-axis yields

**Table 6.** Correlation coefficients between the elements

	Cr	Mn	Fe	Ni	Zn	Ag	Sb	Sn	Pb	Bi
Cr	1									
Mn	0.93	1								
Fe	0.87	0.85	1							
Ni	0.87	0.90	0.86	1						
Zn	0.36	0.60	0.28	0.52	1					
Ag	0.18	0.35	0.43	0.41	0.45	1				
Sb	0.41	0.61	0.48	0.53	0.66	0.75	1			
Sn	0.51	0.71	0.54	0.67	0.76	0.71	0.92	1		
Pb	0.50	0.68	0.45	0.57	0.75	0.49	0.86	0.87	1	
Bi	0.51	0.69	0.44	0.58	0.76	0.43	0.80	0.85	0.96	1

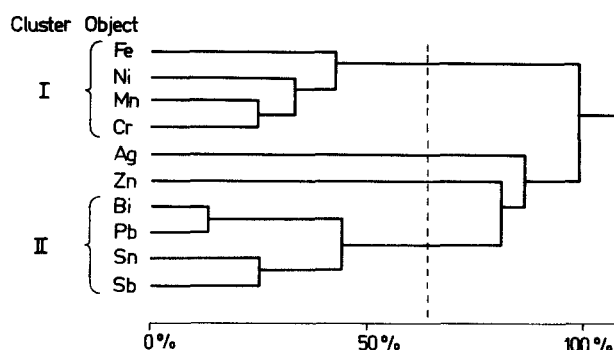
**Table 7.** Extrapolation of the inhomogeneity-caused RSD  $s_{r,inh}$  from measurement procedure (index m) to an optional one (index o)

Procedure	SSMS	GD-ES	
Sampling volume/mm <sup>3</sup>	0.064		0.6
RSD Element	$s_{r,tot,m}$	$s_{r,inh,m}$	$s_{r,inh,o}$
Cr	0.294	0.216	0.022
Mn	0.248	0.146	0.015
Fe	0.335	0.269	0.028
Ni	0.275	0.189	0.020
Zn	0.311	0.238	0.025
Ag	0.408	0.356	0.037
Sb	0.431	0.381	0.039
Sn	0.313	0.241	0.025
Pb	0.389	0.334	0.034
Bi	0.384	0.328	0.034

coefficients significant deviating from zero, whereas Fig. 6b explains why these coefficients are almost non-significant in  $\alpha$ -direction. Obviously, the test by regression can only be sensitive, if the model is adequate to the process.

Some of the results suggest a certain similarity between groups of analytes. Table 6 contains the correlation coefficients for all pairs of elements as objective measures of the similarity. A more comprehensive information gives the dendrogram from agglomerative hierarchical clustering (Fig. 7). One recognizes two main groups of similar elements. Group I contains Fe, Ni, Mn, and Cr, whereas group II includes Bi, Pb, Sn, and Sb. As mentioned above, the former are non-homogeneously distributed in both directions. For the latter, inhomogeneity in  $\alpha$ -direction is almost not proved. Moreover, Ag and Zn are linked to the other elements only on a very high level, thus demonstrating their behaviour to be considerably different from all the other elements. All these correlations were also found for conventional SSMS analyses of rod-shaped samples of the same standard material. The local correlation between the elements Bi, Pb, Sn, and Sb in the ASMW copper standard confirms analogous results obtained earlier from a NBS copper standard [6].

Summarizing, it can be stated that the copper sample under investigation is not sufficiently homogeneous for calibration of SSMS analysis without hesitation. Therefore the question was to be answered, under which conditions

**Fig. 7.** Dendrogram from agglomerative hierarchical clustering by the correlation measure, single linkage mode. The broken line marks the level for stable, reasonable clusters

this sample can be notwithstanding used for calibration purposes. The results of extrapolation to the higher sub-sample volume of 0.6 mm<sup>3</sup>, which is typically for our GD-ES analysis, are given in Table 7. It follows, that the copper standard should be sufficient for procedural calibration if a RSD from inhomogeneity of 0.05 is acceptable. However, because of the anisotropy, especially in radial direction, a preferred sampling of central parts is to be avoided. Also, the sample may be used for calibration of SSMS when scanning the whole surface. Stochastically positioning a high number (>25) of measuring points may sufficiently approximate this condition.

#### 4 Conclusions

Characterizing a standard material with respect to application in analytical techniques of a different sampling volume led to the following general ascertainments:

Selecting the most suitable chemometric methods for homogeneity characterization of solids demands information on the "structure" of presumed inhomogeneities.

However, even without a-priori information, an optimal strategy can provide global statements on homogeneity as well as rough impressions of the structure of inhomogeneities (trends, noise, periodicities).

Besides visualization of the results of spatially resolved measurements the strategy should include tests based on analysis of variance, regression models, statistics of stochastic processes (time series analysis), and multivariate data analysis.

Particularly, transforming the data matrix to a vector (process) by joining together rows or columns enables simple statistical tests for global homogeneity of the whole data set which can be carried out within some minutes by personal computers. However, the results depend on the starting point what may be overcome by cyclic data change.

In regression analysis, which is recommended for testing each spatial direction separately, polynomial models of higher orders should be taken into consideration.

A generalized test strategy, valid in most practical cases and including multivariate methods, is still to be elaborated.

*Acknowledgements.* The authors thank Doz. Dr. sc. nat. H. Zwanziger, TH "Carl Schorlemmer" Leuna-Merseburg, Sektion Chemie, for handing over the computer program CUSUM, Dr. U. Herrmann, VEB Spuremetalle Freiberg, for computational assis-



tance and Chem.-Ing. G. Pietzsch and Mrs. B. Liebich for carrying out the SSMS analyses.

## References

1. Danzer K, Doerffel K, Ehrhardt H, Geißler M, Ehrlich G, Gadow P (1979) *Anal Chim Acta* 105:1–7
2. Danzer K, Kuchler L (1977) *Talanta* 24:561–565
3. Danzer K, Ehrlich G (1984) *Tagungsberichte TH Karl-Marx-Stadt, Heft 12, Bd 2*:547–572
4. Singer R, Danzer K (1984) *Z Chem* 24:339–341
5. Danzer K, Singer R (1985) *Mikrochim Acta* I:219–226
6. Liu XD, Michiels F, Van Espen P, Adams F (1986) *Mikrochim Acta* III:49–70
7. Liu XD, Van Espen P, Adams F (1987) *Fresenius Z Anal Chem* 327:659–664
8. Nikolova L, Krasnobaeva N (1988) *Fresenius Z Anal Chem* 322:22–27
9. Adeberg V, Brügemann L, Doerffel K, Moenke L (1989) *Fresenius Z Anal Chem* 333:143
10. Ahearn AJ (Ed) (1972) *Trace analysis by SSMS*. Academic Press, New York London
11. Mai H, Stahlberg U (1976) *Jenaer Rundsch* 21:316–322
12. Marshall RAG (1977) *Anal Chem* 49:2193–2196
13. Doerffel K (1989) 6. Merseburger Analytiktagung, Prozeßanalytik; *Wiss Z Techn Hochschule*, in press
14. Ehrlich G, Kluge W (1989) *Mikrochim Acta* 1989 I:145–158
15. Ehrlich G, Kluge W, Liebich V, to be published
16. Tukey JW (1949) *Biometrics* 5:232–236
17. Massart DL, Vandeginste BGM, Deming SN, Michotte Y, Kaufman L (1988) *Chemometrics: a textbook*. Elsevier, Amsterdam Oxford New York Tokyo
18. Henrion G, Henrion A, Henrion R (1988) *Beispiele zur Datenanalyse mit BASIC-Programmen*. VEB Dt. Verlag der Wissenschaften, Berlin

Received July 31, 1989