Data Evaluation in High-Performance Liquid Chromatography-Diode-Array Detection-Fluorescence Detection by Information Theory

Victor David* and Andrei Medvedovici

Department of Analytical Chemistry, Faculty of Chemistry, University of Bucharest, Sos. Panduri, No. 90, 76.235, Section 5, Bucharest, Romania

Received November 7, 1999

The concepts of information theory were applied to the high-performance liquid chromatography (HPLC) technique, with diode-array (DAD) and/or fluorescence (FLD) detections. The information amount for a complete analysis can be computed as a function of analytical parameters, such as the number of analytes, level of concentration, and standard deviation of determinations. By means of the proposed method, the information content of a qualitative and quantitative analysis accomplished by HPLC-DAD-FLD was estimated, and sensitivity was optimized taking into account a maximum information content, while the detection limit was estimated considering that at this level of concentration the information content approaches zero.

INTRODUCTION

Evaluation of analytical data is usually focused on the precision and accuracy of determinations during an analytical process. Depending on their sizes, and a criterion chosen to satisfy a standard quality,[1,2] these statistical parameters can be used in optimization of experimental parameters to yield maximum information from the process. Thus, the information theory is deeply involved in analytical chemistry, mainly as a consequence between an analytical and a communication process.[3] Moreover, the advantage of using the new possibility of data processing with the aid of a computer in-line with the analytical instrument can be used in evaluation of the analytical information. Those techniques, such as GC−MS, GC−FTIR, HPLC−MS, or HPLC−DAD (DAD = diode-array detection), which finally report an analytical result as a raw of components identified in a sample together with the so-called similarity index, can be easily discussed in terms of information theory. Thus, in a previous paper we proposed a new method for estimation of analytical data from GC−MS, with very good results in the case of very complex mixtures.[4]

High-performance liquid chromatography (HPLC), with one or two detections, diode-array detection (DAD) and fluorescence detection (FLD), is an analytical technique with large applications in separation and determination of nonvolatile compounds from different mixtures. Due to the recent development in the field of data acquisition and processing, the analytical results can be reported from qualitative and quantitative points of view. For this reason, these results can be characterized by means of information theory. By the application of the information theory in data evaluating a certain analytical process, we usually estimate quantitatively its information amount (content) dependent on the analytical parameters and optimize them to obtain as much information as possible. Up to the present, there have been several attempts in the literature to apply the information theory in the field of HPLC techniques. According to Matsuda et al[5,6] the information theory can be used to select the most efficient HPLC conditions for dissolution testing of multiingredient pharmaceuticals in the framework of a total chromatographic optimization procedure. Information content, sample complexity, physicochemical detectors, and chromatographic techniques are investigated in the analysis of plant extracts.[7] Even the homogeneity of the distribution of an analyte in a matrix has been characterized in terms of information theory.[8] Nevertheless, there is not now a comprehensive treatise of this analytical technique by means of the information theory. Therefore, it is the purpose of the present study to apply this theory to this field in view of measuring the information content obtained from such a determination.

INFORMATION ENTROPY

An analytical process compared with a communication process can be represented by means of the following two fields of probabilities:[9]

$$\text{sample: } \underset{\{X_i,\, p_i,\, i = 1, 2, \ldots n\}}{\text{a priori}} \rightarrow \text{analytical process} \rightarrow \underset{\text{result: } \{X_i,\, q_i,\, i = 1, 2, \ldots, n\}}{\text{a posteriori}}$$

In this representation the main role is played by the events X_i, the number of which being denoted by n. The event X_i can be in HPLC-FLD-DAD a certain compound to be identified with the aid of its UV−vis absorption spectrum, a value of the measured absorbance (in DAD), or a value of the emission intensity (in FLD). The main role of the HPLC process played from the information point of view is to improve the values of a posteriori probabilities, assigned to each chemical compound to be found in the sample, due to the separation process that takes place during this analysis.

* Corresponding author. E-mail: david@chem.unibuc.ro.

10.1021/ci990139x CCC: $19.00
Published on Web 06/02/2000

With these two probabilities the information entropies can be computed by Shannon's formula:

$$H_i(\text{a priori}) = -p_i \log_2 p_i - (1 - p_i) \log_2(1 - p_i) \quad (1)$$

$$H_i(\text{a posteriori}) = -q_i \log_2 q_i - (1 - q_i) \log_2(1 - q_i) \quad (2)$$

The information content, ΔH_i, will be the difference between the two entropies:

$$\Delta H_i = H_i(\text{a priori}) - H_i(\text{a posteriori}) \quad (3)$$

Depending on the values of p_i and q_i, we can encounter one of the following situations with corresponding significance.[10]

(1) $p_i = 0$; $q_i = 0$ ($\Delta H_i = 0$). In this case the process is proving our knowledge about the analyte X_i: without any doubt it is not possible to be found in the sample. The process confirms that the sample does not possibly contain analyte X_i, and thus it does not carry out any information with regard to the analyte X_i.

(2) $p_i = 1$; $q_i = 1$ ($\Delta H_i = 0$). This is a situation similar to that described above, but with another conclusion: the analyte X_i is undoubtedly present in the sample.

(3) $p_i = 1$; $q_i = 0$ (ΔH_i questionable). This is a wrong supposition by assigning to the analyte X_i a sure presence in the sample, while the determination process does not prove this presence. In this rare case it is difficult to estimate the information content of the process.

(4) $p_i = 0$; $q_i = 1$ (ΔH_i questionable). This is the same situation as described above, but the supposition that X_i is not found in the sample seems not to be true after the determination.

(5) $0 < p_i = q_i < 1$ ($\Delta H_i = 0$). The uncertainty assigned to the compound X_i to be present in the sample has not been changed, making the analytical process inopportune in the determination of this compound.

(6) $0 < p_i < q_i < 0.5$ ($\Delta H_i < 0$). Although the probability that X_i is found in the sample has increased, the uncertainty about this compound has also increased, leading to a negative information content of the determination.

(7) $0 < p_i < q_i < 0.5$ ($\Delta H_i > 0$). The certainty for X_i not to be found in the sample has increased, and consequently a positive value of the information content of the determination is obtained.

(8) $0.5 < p_i < q_i < 1$ ($\Delta H_i > 0$). The certainty that X_i is found in the sample has increased.

(9) $0.5 < q_i < p_i < 1$ ($\Delta H_i < 0$). The uncertainty that X_i is found in the sample has increased.

The main conclusion arising from these situations is that an analytical process is useful for determining an analyte X_i if and only if the a posteriori probability q_i is closer to the values 0, or 1, than the a priori probability p_i. In the case of a maximum uncertainty a priori to the analysis, i.e., $p_i = 1/2$ for all the events $i = 1, 2, ..., n$, and a lack of any uncertainty a posteriori to the analysis, i.e., $q_i = 0$, or 1, for all the events $i = 1, 2, ..., n$, the value of the information amount is n bits. Moreover, if the uncertainty of the a posteriori probability field is zero, then the information amount obtained from analysis is given by the a priori information entropy.

INFORMATION CONTENT

The semiquantitative analysis of a sample consists of setting up a number of intervals [C_j(lower), C_j(upper)] within which the concentration of the analyzed component of the analytical sample is situated. These concentration intervals within which the value of the concentration of a certain compound X_i identified to be present in a sample by a DAD detection could be practically represented as follows:

$$C_j(\text{lower}) = 10^{2-j}\%, \quad C_j(\text{upper}) = 10^{3-j}\%, \quad j = 1,2, .., \nu \quad (4)$$

The probability q_{ij} that the value of the concentration of the component i lies in one of the ν intervals is

$$q_{ij} = 1/\nu; \sum_{j=1}^{\nu} q_{ij} = 1, \quad \text{i} = 1, n \quad (5)$$

Therefore, the information amount corresponding to the probability field of a semiquantitative analysis for one component i and a lack of any uncertainty in a final result is given by

$$\Delta H_i(\text{semiquant}) = \log_2 \nu \quad (6)$$

For instance, if a semiquantitative analysis for a certain compound is carried out down to the parts per million level (10^{-4}%, i.e., $\nu = 6$) by means of DAD detection, the information content of this analysis has the value of 2.57 bits, whereas if the analysis is carried out down to the parts per billion level ($\nu = 9$) by means of FLD detection, the value is 3.14 bits.

The information content corresponding to the events of a semiquantitative analysis for all the components of the sample has the expression

$$\Delta H(\text{semiquant}) = \sum_{i=1}^{n} \Delta H_i(\text{semiquant}) = n \log_2 \nu \quad (7)$$

By carrying out a quantitative analysis for a certain component, a confidence interval $\bar{c} \pm \mu\sigma$ is set up, within which the value of the concentration of the analyzed component can be situated; here σ is the standard deviation of the analytical result c, and μ is a number, set after analysis, which indicates the magnitude of the confidence interval in σ units (usually the value $\mu = 3$ is considered). The value of the concentration of a certain component i, after the analysis has been carried out, is $\bar{c}_{ij} \pm \mu_{ij}\sigma_{ij}$, $\bar{c}_{ij} \in [c_j(\text{lower}), c_j(\text{upper})]$, $j = 1, \nu$. Hence, the events of the probability field, before a quantitative analysis has been carried out, are certain concentration microintervals, with a magnitude of $2\mu_{ij}\sigma_{ij}$, $j = 1,\nu$, $i = 1, n$. Besides, the events of the probability field corresponding to a quantitative analysis depend on the events of a semiquantitative analysis.[11]

According to this procedure, the semiquantitative intervals could be chosen such that σ_{ij} is constant over one of these intervals.

If $2\mu_{ij}\sigma_{ij}$ is the magnitude of the confidence interval for the quantitative results, relative to the chemical component i and to the concentration j, the interval j can be divided into s_{ij} microintervals of magnitude $2\mu_{ij}\sigma_{ij}$; their number is given by the relation

$$s_{ij} = \frac{C_j(\text{upper}) - C_j(\text{lower})}{2\mu_{ij}\sigma_{ij}} = \frac{9 \times 10^{2-j}}{2\mu_{ij}\sigma_{ij}} \quad (8)$$

Since the obtained microintervals are independent of one another, the probability p'_{ij} of a microinterval from the interval j, for a component i, is the reverse of s_{ij}:

$$p'_{ij} = \frac{2\mu_{ij}\sigma_{ij}}{9 \times 10^{2-j}} \tag{9}$$

The described procedure of dividing before the analysis an interval j into microintervals of magnitude $2\mu_{ij}\sigma_{ij}$ implies another probability, after the analysis has taken place, namely, the probability that the corresponding microinterval represents the true concentration, denoted by $\bar{C}_{ij}$. The probability that a result C_{ij} deviates from the true value C_{ij} is a Gauss function. Therefore, the probability that a certain obtained result $C_{ij} \in [C_j(\text{lower}), C_j(\text{upper})]$ belongs to an interval $\bar{C}_{ij} \pm \mu_{ij}\sigma_{ij}$ can be estimated by the following relation:

$$p''_{ij} = \frac{1}{\sigma_{ij}(2\pi)^{1/2}} \int_A^B \exp\left[-\frac{(C_{ij} - \bar{C}_{ij})^2}{2\sigma_{ij}^2}\right] dC_{ij} \tag{10}$$

$$(A = \bar{C}_{ij} - \mu_{ij}\sigma_{ij}; \quad B = \bar{C}ij + \mu_{ij}\sigma_{ij})$$

After the variable change $z = (C_{ij} - \bar{C}_{ij})/\sigma_{ij}$ is performed, the integral (10) becomes

$$p''_{ij} = 2\left[\frac{1}{(2\pi)^{1/2}} \int_0^{\mu_{ij}} \exp\left(-\frac{z^2}{2}\right) dz\right] 2\Phi(\mu_{ij}) \tag{11}$$

where $\Phi(\mu_{ij})$ is the Laplace function, which is known dependent on μ.

If multiplied by p''_{ij}, the probability p'_{ij} gives the probability p_{ij} of obtaining a microinterval of magnitude $2\mu_{ij}\sigma_{ij}$, which would represent the true value C_{ij}. Hence, p_{ij} is given by

$$p_{ij} = \frac{4\mu_{ij}\sigma_{ij}\Phi(\mu_{ij})}{9 \times 10^{2-j}} \tag{12}$$

The probability of obtaining by quantitative analysis a microinterval of the interval j, of magnitude $2\mu_{ij}\sigma_{ij}$, when the interval k has been determined previously by semiquantitative analysis, represents the probability $\pi(2\mu_{ij}\sigma_{ij}|k)$ of an event of the quantitative analysis conditioned by the event of a semiquantitative analysis, having the expression

$$\pi(2\mu_{ij}\sigma_{ij}/k) = p_{ij}\delta_{jk} \tag{13}$$

where δ_{jk} is Kronecker's symbol ($\delta_{jk} = 1$ if $j = k$, and $\delta_{jk} = 0$ if $j \neq k$). In the case of an in-line analysis, HPLC-DAD-FLD, this conditional probability has an important significance: it gives the probability of a fluorescent compound to give an analytical signal in FLD, when it gives a signal in DAD, at a proper spectral domain, where it can absorb UV radiation.

The information content of the probability field that has as events the microintervals of magnitude $2\mu_{ij}\sigma_{ij}$ and the probabilities p_{ij} conditioned by the event k of a semiquantitative analysis is given by

$$\Delta H(2\mu_{ij}\sigma_{ij}/k) = -s_{ij}\pi(2\mu_{ij}\sigma_{ij}/k) \log_2 \pi(2\mu_{ij}\sigma_{ij}/k) \tag{14}$$

The information content of the probability field corresponding to a HPLC quantitative analysis of component i, conditioned by the probability field corresponding to a semiquantitative analysis, denoted by ΔH_i(quant/semiquant), has the expression

$$\Delta H_i(\text{quant/semiquant}) = \sum_{k=1}^{\nu}\sum_{j=1}^{\nu} q_{ij}H(2\mu_{ij}\sigma_{ij}/k) = -\sum_{k=1}^{\nu}\sum_{j=1}^{\nu} s_{ij}q_{ij}p_{ij}\delta_{ij} \log_2 p_{ij} \tag{15}$$

Replacing expressions 5 and 12 for q_{ij} and p_{ij}, respectively, the following expression is obtained for the information content of the probability field corresponding to the events of a quantitative analysis of component i, depending on the probability field of the events of a semiquantitative analysis:

$$\Delta H_i(\text{quant/semiquant}) = \frac{2}{\nu}\sum_{j=1}^{\nu} \Phi(\mu_{ij}) \log_2 \frac{9 \times 10^{2-j}}{4\mu_{ij}\sigma_{ij}\Phi(\mu_{ij})} \tag{16}$$

Taking into account all the components of the sample, expression 16 becomes

$$\Delta H_i(\text{quant/semiquant}) = \frac{2}{\nu}\sum_{i=1}^{n-1}\sum_{j=1}^{\nu} \Phi(\mu_{ij}) \log_2 \frac{9 \times 10^{2-j}}{4\mu_{ij}\sigma_{ij}\Phi(\mu_{ij})} \tag{17}$$

The information content ΔH corresponding to the events of a complete HPLC analysis is given by

$$\Delta H = H(\text{qual}) + H(\text{semiquant}) + H(\text{quant}|\text{semiquant}) \tag{18}$$

or more explicitly

$$\Delta H = n \log_2 2\nu + \frac{2}{\nu}\sum_{i=1}^{n-1}\sum_{j=1}^{n} \Phi(\mu_{ij}) \log_2 \frac{9 \times 10^{2-j}}{4\mu_{ij}\sigma_{ij}\Phi(\mu_{ij})} \tag{19}$$

Expression 19 depends on four variables: n, ν, σ_{ij}, and μ_{ij} ($i = 1, n$; $j = 1, \nu$). The information amount ΔH is a linear function of n and a logarithmic function of ν and σ_{ij}; it is a convex function of μ, so that a value of μ can be found for which the function $\Delta H(\mu)$ has a maximum.

APPLICATIONS

Information Content in Qualitative Analysis. The main problem related to the computation of information amount obtained from a qualitative analysis by HPLC-DAD is given by the difficulty of establishing the a posteriori probability assigned to each component identified in a sample. This is translated into the problem of how much a spectrum of one component during a chromatographic run is similar to the real one, because many times small changes in chromatographic separation can influence an UV−vis spectrum, not deeply, but enough to introduce from the probability point of view an uncertainty a posteriori to the analysis. Our assumption is that the degree of overlapping of two spectra of one component (one being a reference, while the other being investigated) gives the probability of an experimental spectrum to belong to a certain component. Many times this degree of overlapping of two spectra is called a similarity

Table 1. Information Content for a Qualitative Analysis of 10 Analytes, Depending on the Value of a Posteriori Probability

q_i	H(a posteriori)	ΔH (bits)	q_i	H(a posteriori)	ΔH (bits)
0.55	9.909	0.091	0.80	7.206	2.794
0.60	9.691	0.309	0.85	6.087	3.913
0.65	9.323	0.677	0.90	4.681	5.319
0.70	8.796	1.204	0.95	2.859	7.141
0.75	8.098	1.902	1.00	0.000	10.00

index, and new analytical instruments facilitate to do that by means of special data processing. For example, Lepadatu[12] proposed a general method for comparing two spectra: one of them has to be searched, and the other is a known spectrum of the compound to be identified. The similarity between two spectra (D_n), which has the meaning of a distance between these two spectra, is given by the formula

$$D_n = \frac{\left(\sum_{k=1}^{n} \frac{A_k}{B_k}\right)^2}{n\sum_{k=1}^{n}\left(\frac{A_k}{B_k}\right)^2} \geq 1 \qquad (20)$$

where A_k and B_k are the values of absorbance from the unknown spectrum and a known spectrum, respectively. According to this method the more closed to 1 is the value of D_n, the more probable is it that the two searched spectra belong to the same compound. In terms of a posteriori probability, this one can be easily computed by the ratio $q = 1/D_n$. In a previous paper[13] we already proved that an estimating factor of the information content of a spectrum belonging to a certain compound can be used in the estimation of the degree of overlapping between two spectra. This factor was called a minimum information content, and it is similar to formula 20 given above.

In Table 1 we give several examples of hypotethical components found with an equal probability to be present in a sample.

Information Content in Quantitative Analysis. The standard deviation σ_{ij} for component $i = 1, n$ and concentration interval $j = 1, \nu$ may be expressed as a function of the proportionality of the lower limit of the corresponding concentration interval, C_j(lower) $= 10^{2-j}$, as follows:

$$\sigma_{ij} = \beta_{ij} \times 10^{2-j} \qquad (21)$$

In this case, the function ΔH can be written in a simpler form:

$$\Delta H = n \log_2 2\nu + \frac{2}{\nu}\sum_{i=1}^{n-1}\sum_{i=1}^{\nu}\Phi(\mu_{ij}) \log_2 \frac{9}{4\beta_{ij}\mu_{ij}\Phi(\mu_{ij})} \qquad (22)$$

In what follows, we will take some examples of estimating the values of ΔH, considering the ideal case in which β_{ij} and μ_{ij} are constant, having the values β and μ for any $i = 1, n$ and $j = 1, \nu$. The following expression of ΔH is thus obtained:

$$\Delta H = n \log_2 2\nu + 2(n-1)\Phi(\mu) \log_2 \frac{9}{2\beta\mu\Phi(\mu)} \qquad (23)$$

The values of the information content obtained for a set of $n = 10$ analytes as a function of β and μ parameters, for an HPLC analysis carried out down to the parts per million level ($\nu = 6$) by DAD, and to the parts per billion level ($\nu = 9$) by FLD, are given in Table 2. It can be seen that the information content of the analytical results increases with the increase of the precision of the measurements, represented here by the parameter β.

Table 2. Values of the Information Content (bits) for a Set of $n = 10$ Analytes, as a Function of β and μ, for a Quantitative HPLC Analysis Carried out Down to Parts per Million and Parts per Billion Levels

		$\beta = 0.1$		$\beta = 0.01$		$\beta = 0.001$	
μ	$\Phi(\mu)$	$\nu = 6$	$\nu = 9$	$\nu = 6$	$\nu = 9$	$\nu = 6$	$\nu = 9$
0.25	0.0987	53.32	59.17	59.22	65.07	65.12	70.97
0.50	0.1915	63.00	68.85	74.45	80.30	85.90	91.75
0.75	0.2734	69.20	75.05	85.55	91.40	101.90	107.75
1.00	0.3413	72.98	78.83	93.38	99.23	113.79	119.64
1.25	0.3944	74.98	80.83	98.56	104.42	122.15	128.00
1.50	0.4332	75.72[a]	81.57[a]	101.62	107.48	127.53	133.38
1.75	0.4599	75.62	81.47	103.13	108.98	130.63	136.48
2.00	0.4772	75.01	80.86	103.55[a]	109.40[a]	132.08	137.93
2.25	0.4878	74.11	79.96	103.28	109.13	132.45[a]	138.30[a]
2.50	0.4938	73.07	78.92	102.60	108.45	132.12	137.97
2.75	0.4970	72.00	77.85	101.72	107.57	131.44	137.29
3.00	0.4987	70.95	76.80	100.77	106.62	130.59	136.44

[a] Indicates the maximum of ΔH for constant ν and β.

Sensitivity. An important application of the proposed model consists of setting up an optimum value of the sensitivity of the measurements (noted by S, and defined as a change in the analytical response to the change in the concentration of an analyte[14]) by this technique, which is being used in the analytical process as a function of the standard deviation of the whole analytical process.[15] It results from the described model that the optimum value of the measurement sensitivity, denoted by S_{opt}, is that value which corresponds to a maximum information content of the analytical results as a function of μ_{ij} and σ_{ij}, that is, $S_{opt} = \mu_{ij}\sigma_{ij}$ so that $\Delta H(\mu_{ij}\sigma_{ij}) =$ maximum; for $S < \mu_{ij}\sigma_{ij}$ one would obtain a large number s_{ij} of concentration microintervals, but with a low probability of representing the true value, whereas for $S > \mu_{ij}\sigma_{ij}$ one would obtain a small number s_{ij} and a high probability of representing the true value of the concentration of the component i. In this way, one can select the analytical results as a function of standard deviation, so as to carry out a maximum possible information content.

To exemplify, we will consider here the quantitative determination of one component by this analytical technique, in the range of parts per million, that is [1, 1000] ppm, with a standard deviation of the measurements $\sigma = 0.1$ ppm, assumed to be constant over the entire given concentration range. According to Table 2, the value $\mu = 1.5$ corresponds to a maximum information content of the analytical results, so that the results carrying out a maximum information amount are those from the following series, or any other series having 0.15 ppm as a microinterval of variation:

[1; 1.15; 1.30; 1.45; 1.60; 1.75; ...; 1000] ppm

The above series of results has therefore an information content larger than that of the series of results given below, obtained with a similar standard deviation:

[1; 1.01; 1.02; 1.03; 1.04; 1.05; ...; 1000] ppm

Table 3. Values of the Detection Limits that Do Not Bring any Information to the Analytical Results, for $j = 7$ (Domain 0.1−1 ppm)

μ	C_{min} (ppm)	μ	C_{min} (ppm)
0.50	2.35	2.00	0.24
0.75	1.10	2.25	0.20
1.00	0.66	2.50	0.18
1.25	0.46	2.75	0.165
1.50	0.35	3.00	0.15
1.75	0.28	3.25	0.113

To conclude, by means of this procedure the value of the information entropy of the analytical results in HPLC-DAD-FLD can be estimated and optimized as a function of the number of analyzed components, the number of concentration domains, the standard deviation, and the sensitivity of the analytical process.

Detection Limit. The problem of estimating the detection limit (C_{min}) of a certain measurement should start from the assumption that the information content is continuously decreasing by closing to the detection limit, beyond which this amount becomes zero, or even a negative value. This fact can be easily seen from formula 17, where at the detection limit the value of the confidence interval will cover the concentration interval in which the analyte is situated; i.e., the denominator in that formula is equal to the nominator, and thus the ration is becoming 1, giving a value of 0 in a logarithmic base. For example in Table 3 several values for the detection limit were given for a DAD detection, in the concentration interval of 0.1−1 ppm, these values being available for the FLD detection, but in parts per billion units.

CONCLUSIONS

The HPLC-DAD-FLD technique can be discussed from the information point of view, because it provides qualitative and quantitative information at the same time. For a qualitative analysis the main conclusion is that the information content is positive when the a posteriori probability for a component is closer to the value 0, or 1, than the a priori probability of the same component. To estimate the information content for a quantitative determination, we proposed a procedure that takes into account the main parameters of the HPLC process. By means of this procedure it is possible to estimate the sensitivity and detection limit to yield analytical results carrying out maximum information content.

REFERENCES AND NOTES

(1) Soltzberg, L. J.; Wilkins, C. L.; Kaberline, S. L.; Lam, T. F.; Brunner, T. R.; *J. Am. Chem. Soc*. **1976**, *98*, 7139.
(2) Soltzberg, L. J.; Wilkins, C. L.; Kaberline, S. L.; Lam, T. F.; Brunner, T. R. *J. Am. Chem. Soc*. **1976**, *98*, 7144.
(3) Eckschlager, K.; Stepanek, V. *Information as Applied to Chemical Analysis*; John Wiley & Sons: New York 1979.
(4) David, V.; Medvedovici, A.; Tache, F. *Ann. Chim. (Roma)* **1998**, *88*, 577.
(5) Matsuda, R.; Hayashi, Y,; Suzuki, T.; Saito, Y. *Fresenius' J. Anal. Chem*. **1993**, *347*, 225.
(6) Matsuda, R.; Hayashi, Y.; Suzuki, T.; Saito, Y. *Chromatographia* **1991**, *32*, 233.
(7) Crozier, A.; Reeve, D. R. *Anal. Proc. (London)* **1992**, *29*, 422.
(8) Danzer, K.; Schubert, M.; Liebich, V. *Fresenius' J. Anal. Chem*. **1991**, *341*, 511.
(9) Guiasu, S.; Theodorescu, R. *Matematica si Informatica, Ed. Stiintifica*; Bucharest, 1965.
(10) David, V. *Rev. Roum. Chim*. **1993**, *38*, 527.
(11) David, V. *Rev. Roum. Chim*. **1987**, *32*, 539.
(12) Lepadatu, C. *Rev. Roum. Chim*. **1985**, *30*, 3.
(13) David, V. *Rev. Roum. Chim*. **1986**, *31*, 823.
(14) Liteanu, C.; Rica, I. *Statistical Theory and Methodology of Trace Analysis*; Ellis Harwood: Chichester, U.K. 1980.
(15) David, V. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 278.

CI990139X