UDC 543.422

Novel Approaches to Spectrophotometric Analyzing the Unseparated Mixtures of Organic Substances

V. I. VERSHININ

Dostoevsky Omsk State University,
Pr. Mira 55A, Omsk 644077 (Russia)

E-mail: vyvershinin@yandex.ru

Abstract

We discuss the tasks and methodological problems of molecular absorption spectrometry as a method for the analysis of unseparated mixtures of organic compounds (petroleum products, pharmaceuticals, and others), which method in some cases can successfully compete with chromatography. Using the chemometric algorithms, in particular the method of projection to latent structures, allows successfully solving three major problems: to eliminate the influence of inter-reference spectral overlapping, to take into account a non-additive character of absorption, to level the detection sensitivity with respect to similar analytes. A brief review of relevant studies carried out within recent years in the Omsk State University is reported.

Key words: analysis of organic substances, analysis of unseparated mixtures, molecular absorption spectrometry, inter-reference overlapping, non-additivity, integrated parameters

INTRODUCTION

The quantitative analysis of unseparated mixtures by molecular absorption spectra in the UV, IR or visible region has a long and complicated history [1]. The spectrum of a mixture of organic compounds rarely allows finding the regions (wavelengths) where the only one component, one kind of molecules exhibits absorption. This complicates the molecular absorption analysis of mixtures in comparison with atomic absorption. Many research works were devoted to solving the mentioned problem, starting from the studies made by the founder of spectrophotometric analysis Karl Vierordt (1873). Techniques were developed for directly determining organic compounds in pharmaceuticals, petroleum products and other mixtures. Usually, a sample as solubilised to measure the absorbance at several wavelength values, with further calculating the concentration of components to solve the system of algebraic equations [2, 3]. Thus, it is possible to calculate also the total content of a number of compounds under determination (analytes), which is important for ecology and medicine.

Often, before the registration of the sample spectrum, researchers use to perform the derivatization of analytes with the separation or masking of the interfering components. For example, phenols contained in natural waters are distilled with steam, to convert them to intensely coloured quinonimine compounds [4]. However, the corresponding photometric reactions are insufficiently selective. Many organic compounds could be determined without derivatization, according to own absorption thereof in the UV region. In the IR spectrometry, the derivatization technique is almost not used.

Taking into account the unpredictability of the influence of inter-reference overlapping, the spectrometry of unseparated mixtures is rarely used in practice. The spectrometry has become a precise method for the quantitative analysis of real objects only within the last years of the 20 century, after the developing of chemometric algorithms and software for processing multivariate data [5]. The "mathe-

maticized" variants of multi-wavelength spectrometry today are used not only in scientific research [6], but also at industrial analytical laboratories. So, using previously obtained multivariate calibration, from UV absorption spectrum of a pharmaceutical preparation on could determine separately active components thereof even in the presence of filling agents [7]. According to the absorption spectrum of gasoline in the near-infrared region, one could determine the structural-group composition, octane number and other characteristics thereof [8]. The multi-wave spectrometry of unseparated mixtures can in some cases successfully compete with chromatography [9], providing approximately the same accuracy with a less consumption of time and money.

Unfortunately, the unique capabilities of "computerized" spectrometry are unknown to the most of chemists and technologists, whereas the method is not yet been described in the university textbooks. Methodological problems arising in the course of analyzing any mixtures are usually considered only for the mixtures of either one type. The ideas and achievements of professionals engaged, for example, in the analysis of petroleum products, are insufficiently used for the analysis of pharmaceutical preparations, and *vice versa*. In is obvious that in addition to the specific methods of analyzing the mixtures of each type one should develop a general methodology for the analysis of mixtures. The problems of the qualitative analysis of unseparated mixtures were generally considered by the authors of monograph [10] and in a number of reviews. The general methodology for the quantitative analysis of unseparated mixtures is only under development [3, 11, 12]. The centre of relevant studies in Russia within recent years is presented by the Analytical Chemistry Chair of the Omsk State University (OmSU), but in the studies there are also involved experts from other organizations: the Institute of Hydrocarbons Processing of the SB RAS (Omsk), the Tomsk Polytechnic University, the Kuban State University (Krasnodar), the North (Arctic) Federal University (Arkhangelsk), and others. The methods spectrometric analysis of unseparated mixtures are developed at the Saratov State University [13–15]. Very interesting papers by foreign researchers were published in journals "Journal of Chemometrics" and "Chemometrics and Intelligent Laboratory Systems" (see review [16]).

This paper summarizes the results of fundamental research in the field of spectrometric analyzing the unseparated mixtures of organic compounds. These studies were performed at the interface of analytical chemistry, chemometrics and molecular spectroscopy. The regularities revealed within recent years have allowed us to develop and put into practice a number of express methods for the analysis of petroleum products, pharmaceuticals, food, and drinking water. Being familiar with the methodology and examples of analyzing such mixtures would help researchers to successfully develop this promising method with respect to new objects (polymers, biomaterials, catalysts, *etc.*).

ANALYTICAL OBJECTIVES AND METHODOLOGICAL PROBLEMS

Within the framework of the analysis of mixtures one could define four typical problems: 1) the determination of a single component; 2) the determination of a number of separate components; 3) the determination of the total content of a group of similar components; 4) separate determination of several groups. All these problems could be solved either with the separation or without the separation of components. Currently, the mixtures of organic compounds are being analyzed *via* separating all the components in a chromatographic column, but in some cases it is impractical. No complete separation of a mixture required in the course of the determination of a single compound or a group of compounds. Taking into account the advantages of the spectrometry such as rapidity, minimal impact on the sample, the simplicity and cheapness of analysis, for to solve the first and third objectives it is better to use this method. If it is necessary to separately determine the content of a number of components in a sample or a group composition, it is preferable to use chromatography. However, owing to the instability of certain analytes in the course of separation thereof and a long time of the chromatographic analysis it is sometimes advisable to use the

TABLE 1

Methodological problems appearing in fulfilling different type tasks problems in the analysis of mixtures

Objective type	What is determined	Methodological problems
1	Content of analyte (*X*)	Effect of foreign substances
2	Content of a number of analytes (*X*, Y, *Z etc.*)	Inter-reference overlapping, optical absorption non-additivity, effect of foreign substances
3	Total content of a group of one-type analytes (ΣX)	Different sensitivity in determining the analytes, optical absorption non-additivity, not completely known qualitative sample composition, choosing a standard substance, the influence of foreign substances
4	Total content of a number of groups (ΣX, ΣY, ΣZ)	Intergroup spectral overlapping, intragroup difference between absorption coefficients for one-type analytes, optical absorption non-additivity, not completely known qualitative sample composition choosing a standard substance, the influence of foreign substances

spectrometry also for the objectives of the second and fourth types.

The methodological problems those are faced by spectroscopists in the analysis of unseparated mixtures are determined by the type of problem to be solved rather than by the nature of the mixture under investigation and choosing either region of the spectrum. It is natural that the number and complexity of these problems in the series of the mentioned objective exhibits an increase (Table 1). The main problems in all the cases consist in the inter-reference of spectral overlapping and light absorption non-additivity. In the case of determining the total content values one faces the problem of different sensitivities in determining single-type components. The analysts of the OmSU place high emphasis on just these three issues. Serious difficulties can arise in connection with a not completely known qualitative composition of the sample under investigation and an unpredictable influence of foreign substances. However, basing on data published, no systematic research in these fields were performed.

The mentioned factors result in systematic errors in the analysis those could be eliminated or reduced *via* applying the mathematical treatment of spectral data.

The mathematical data processing is aimed at extracting the maximum amount of information concerning the composition of the system under study and at the evaluation of the reliability of this information. Let us consider the methodological problems inherent in the spectrometric analysis of the mixtures of organic compounds, and the ways to solve these problems based on the use of chemometric algorithms.

Taking into account the inter-reference overlapping

The problem of inter-reference overlapping is particularly urgent in the case of simultaneous and separate determining a number of components in a mixture (the objective of the second type). Figure 1 demonstrates the absorption spectra of six vitamins in the UV region. To determine the content of each vitamin in the spectrum of such a mixture, at first glance,

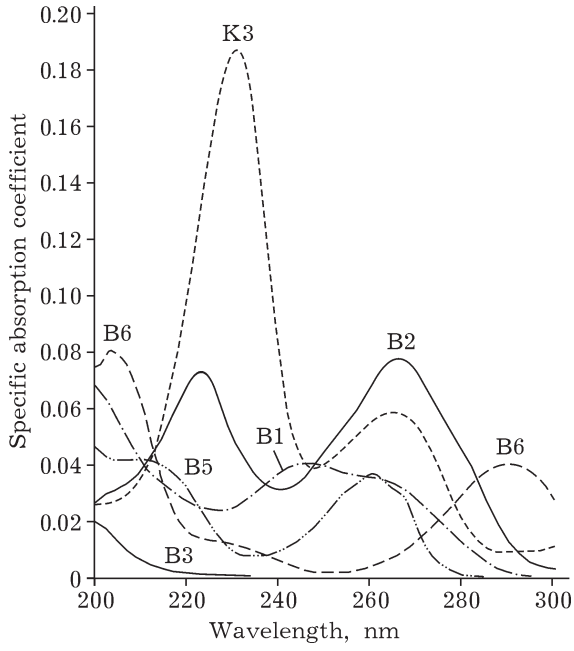

Fig. 1. Absorption spectra of some water-soluble vitamins.

seems impossible. Therefore, in developing the methods for the analysis of multivitamin mixtures, our predecessors concentrated the attention either on searching for photometric reactions inherent in each component or on the separation of vitamins by means of HPLC. However, with appearing the chemometric algorithms the inter-reference overlapping have ceased to be an insurmountable obstacle. Using an algorithm of projections to the latent structures, the spectrum of multivitamin mixture nowadays allows researchers to determine the concentration of each vitamin and total content thereof [17]. In this case, chemical reactions are not performed. In the presence of a previously prepared multivariate calibration the analysis of the mixture takes only 20–30 min.

There are many chemometric algorithms known suitable for the analysis of mixtures under the conditions of inter-reference overlapping, the method of main components analysis (MC), multiple linear regression (MLR) method, the method of projection to latent structures (PLS), many spectral decomposition methods and others [18]. For the quantitative analysis of mixtures there are Vierordt's method [2], linear programming method [3] and other algebraic methods widely used. The influence inter-reference overlapping is greatly reduced as the result of spectra differentiation of the. To solve a particular analytical task one should only choose the most appropriate method for mathematical processing the spectral data.

Unfortunately, the possibility of different chemometric algorithms as applied to the spectral analysis of mixtures is almost unknown. It is not known whether it is possible to use either method, when the qualitative composition of the mixture is not completely known. What should be done when the light absorption of a mixture of non-additive? What an algorithm should be chosen when there are no standards or the Lambert–Beer–Bougue law is not held, when the optical properties or the concentration of the components differ to a great extent? The researchers of the OmSU first of all had to examine and compare the analytical capabilities of different algorithms. At the same time, as the model objects there were used the mixtures of organic compounds, from binary to ten-component ones.

The concentrations of components in the photometric solution ranged within 10^{-7}–10^{-4} mol/L, the concentration ratio amounted up to 20 : 1. The modified algorithms and tested with respect to model mixtures were used for the development of the methods for the analysis of real objects. So, the group headed by I. V. Vlasova was engaged in analyzing pharmaceutical preparations, the group headed by T. V. Antonova studied the analysis of hydrochemical objects, the group headed by T. G. Tsyupko analyzed food. The other groups analyzed petroleum products, the mixtures of the products of laboratory-scale organic synthesis and so on. The studies performed have demonstrated that using chemometrics algorithms researchers could obtain accurate (±5 rel. %) results in the course of analyzing unseparated mixtures even when the UV or IR absorption spectra of the components overlap completely. However, some algorithms (Vierordt's method) are suitable only in the simplest cases (2–3 components), whereas the second others (MLR) could be used for more complicated mixtures (4–5 components), and the third others (PLS) are able to give good results even in the course of analyzing ten-component mixtures [19].

Many chemometric algorithms require for the preliminary construction of a mathematical model (multivariate calibration). Such models are calculated from the spectra of mixtures with a known qualitative and quantitative composition, those constitute a so-called training sample. However, it is unclear what the mix-

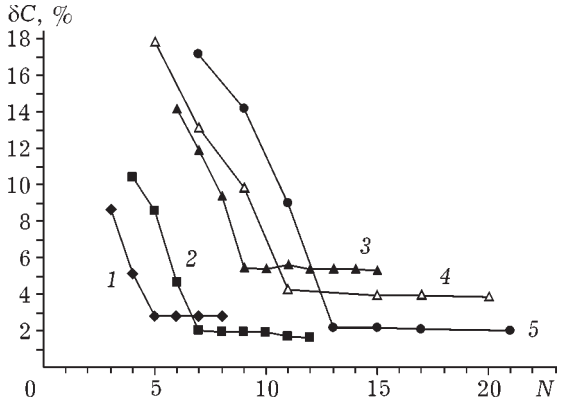

Fig. 2. Generalized analytical error (δC) for model mixtures depending on the training sample number (*N*): *1–5* – number of simultaneously determine of components (*n*), equal to *1–5*, respectively.

tures should be included in the training sample, and how many mixtures are required in each case. It is believed that the more the better. But in some cases it is not so. Our studies have demonstrated that increasing the amount of a training sample result in the fact that the error values of subsequent analyses vary in a non-monotonic manner (Fig. 2): at first, the errors exhibit a decrease, whereas further they remain approximately constant [20]. The necessary and sufficient number of mixtures with known composition (the optimal training sample number) depends on the complexity of the problem, especially on the number of analytes under simultaneous determination. As it is seen from Fig. 2, for the construction of multivariate calibration by means of PLS method one should use training samples containing $N = 2n + 1$ mixtures with known composition, where n is the number of simultaneously determined components. Further increasing the parameter N does not result in a higher accuracy. However, in some cases the optimum value of N considerably exceeds $(2n + 1)$ [21]. The excess is required in the following cases: a) the spectra of the components are very similar (authentically correlated); b) the macro and micro components of the sample being simultaneously determined; c) both strongly and weakly absorbing compounds being simultaneously determined.

Researchers also succeeded in establishing what the mixtures should be included in the training sample to provide the most accurate subsequent analyzes in the implementation of MLR method. To solve this problem, there was matrix analysis used and special software created, to automatically calculate the optimum composition of the training samples [22]. The user should only specify the lower and upper limits of the range of each analyte content in the testing sample, as well as the number of compounds in the training sample. Rapid procedures for the analysis of pharmaceutical preparations developed basing on these recommendations have been metrologically certified to be listed in the Federal Register of analytical techniques. An example could be presented by the method of simultaneous determining the active components of pharmaceutical preparation “Papazol” according to the absorption spectrum of an aqueous solution in the UV region (Table 2).

There was MLR algorithm used in the modification of the indirect calibration, i.e. the regression coefficients for papaverine and dibasol were determined from the spectra of model mixtures ($N = 5$). This technique has been included in the Federal Register (FR 1.31.2010.07919); the accuracy thereof is not worse as to compare the pharmacopeia procedure, but the result is achieved more quickly and with no use of toxic reagents.

To control the quality of pharmaceutical preparations basing on the absorption spectra in the near-infrared region, other algorithms are also suitable (in particular, the method of MC); those also involve the construction of multivariate calibrations [7]. An alternative approach to solving the problems associated with the separate determination of a number of components is under development by the Saratov researchers [15]. They use chemometric algorithms for decomposing the spectrum of a mixture. The subspectra isolated by a computer are very similar to the spectra of the mixture components taken separately; in the case of properly choosing the algorithm and the wavelength range, the corresponding correlation coefficients are almost equal to 1. This makes it possible to identify the components and to conduct the qualitative analysis of mixtures containing up to six components. Further, one determines the contribution of each component to the total absorbance of the mixture at the analytical wavelength, and, finally, calculates the concentration of components.

TABLE 2

Metrological characteristics of methods for the analysis of pharmaceutical preparation “Papazol”

Analytes	Content range, mg/tab.	Repeatability level, %	Reproducibility level, %	Accuracy level, %
Bendazol	20–40	2.0	3.0	7.0
Papaverine hydrochloride	20–40	1.5	2.0	6.0

Inter-reference overlapping hinders also the spectrometric determination of the group composition of objects (the fourth type problem). Correctly choosing the training sample composition and the spectral range for the construction of adequate multivariate calibrations is much more difficult in this case. A serious problem also consists in the background absorption of light by side compounds (not the ones those should be determined). We first faced this problem in the analysis of pharmaceutical and vitamin preparations containing fillers. In such cases, the mixtures used for constructing multivariate calibration should contain the same set of interfering substances as those contained the sample under investigation. The compliance should be performed also according to the quantitative content of impurities, which requires for a reliable reference technique of analysis or for a stable preliminary known composition of impurities.

Problem of non-additivity

When solving problems related to the analysis of unseparated mixtures, a particular problem is presented by optical absorption non-additivity, usually associated with the interaction between the components. It is a source of systematic errors [12]. In the case of the non-additivity of optical absorption the optical density (absorbance) of the mixture (A_Σ) is significantly different from the sum of the optical densities of the components measured separately at the same wavelength (ΣA). Deviation $\Delta A = A_\Sigma - \Sigma A$ can be either positive or negative. It should be noted that the systematic studies of the deviations from optical absorption (OA) additivity were not conducted earlier, although statistically significant OA quite often occur in the practice of the analysis of unseparated mixtures. On the other hand, the measurements of the OA could provide valuable information concerning chemical interactions in solution (including supramolecular interactions).

To all appearance, the development of methods for the analysis of unseparated mixtures, checking the additivity of analytical signals should be as mandatory and trivial procedure, as the sensitivity evaluation or selectivity verification. It is advisable to check the optical absorbance additivity for model mixtures under different conditions, *via* varying the concentration of the analyte and reagents, the pH value, the wavelength and so on [23]. The deviations could be easily revealed *via* registering the spectrum of a mixture to visually compare it with the superposition of the spectra of the components. An example could be presented the spectrometric determination of the total content of two antioxidants such as ascorbic acid and quercetin according to the reaction with iron (III) and phenanthroline (Fig. 3, *a*). In this case there appear negative OA reaching of approximately 40 % in modulo [24]. This results in an underestimation of the total con-

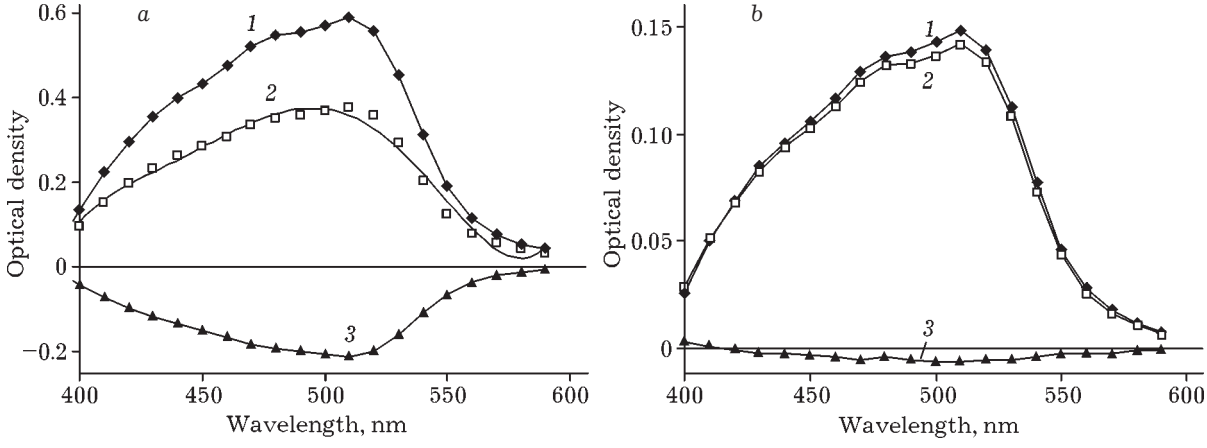

Fig. 3. Spectra of the products of interaction between the antioxidant mixtures and a complex reagent (C_{Fe} = 60 μmol/L, $C_{o\text{-phen}}$ = 100 μmol/L, T = 25 °C): *a* – ascorbic acid (10 μmol/L) + quercetin (3 μmol/L); *b* – ascorbic acid (2.5 μmol/L) + rutin (2.5 μmol/L): *1* – ΣA, *2* – A_Σ, *3* – ΔA.

TABLE 3

Deviations from additivity within the IR region according $3S$ criterion [26]

Composition of the mixture	ν, cm^{-1}	A_Σ	ΣA	$\lvert\Delta A\rvert$	$3S$	Significance ΔA
Naphthalene + anthracene	879.7	0.219	0.219	0	0.006	−
	3055.7	0.291	0.289	0.002	0.019	−
Ethylenediamine + benzoic acid	883.5	0.226	0.285	0.059	0.001	+
	1597.3	0.237	0.110	0.127	0.001	+
cis-R*+ *trans*-R*	1500.9	0.059	0.057	0.002	0.007	−
	1709.2	0.100	0.098	0.002	0.002	−

*R = *N*-[(2-oxohexyliden)methyl]-2-chloroacetamide.

tent of antioxidants. For another pair such as ascorbic acid and rutin the OA is not significant (see Fig. 3, *b*), whereas the total concentration can be determined correctly.

To reveal any small but statistically significant OA at a fixed wavelength is quite difficult: it is required for the statistical processing of data obtained in the course of a multiple preparation and photometry of solutions, and for the subsequent evaluation of the significance of the null hypothesis. There are no standard methods for such verification. Earlier, there were complicated and laborious methods proposed for this purpose based on the use of the correlation or regression analysis [23]. We have developed and tested three ways to identify statistically significant OA [25]. They are applicable not only to spectrometry, but also to other instrumental methods of analysis. The easiest way consists in checking according to the $3S$ criterion. To knowingly additive mixtures (*e. g.*, the mixtures of arenes) the ΔA value in modulo does not exceed triple standard deviation for the optical density (absorbance) of the mixture. In other cases (the mixture of ethylenediamine and organic acids) $|\Delta A| > 3S$, which indicates that the OA is statistically significant (Table 3).

The non-additive absorption of the mixtures of acids and amines within the infrared region could be explained by a protolytic interaction between the components. It should be noted that the verification of IR spectra additivity cannot be performed using the scale of transmissions: $\Sigma T \neq T_\Sigma$, even in the case of additive mixtures. One should not judge the additivity also basing on the position of peaks in the wavelength scale (wave numbers). Significant shifts of the peaks could appear also in the spectra of additive mixtures due to the superposition of the peaks of different substances positioned close to each other [26].

A more complex, but the most informative way of testing consists in the construction of statistical models for light absorption in the course of complete factorial experiment. This allows one to associate the value of OA with the mixture composition, *i. e.*, to predict the level of non-additivity with a fairly high forecasting accuracy: for the systems studied by the authors of [24, 25], the differences between predicted and real values ΔA did not exceed 20 rel. %. By means of using such models one could calculate the conditions for performing the analysis wherein the OA values should be negligible. This is possible even in the case when the cause of occurring the non-additivity is revealed. However, one should always try to find out those causes. In addition to protolysis, the formation of associates and the other interaction processes between solutes the OA could be caused by a competition between the analytes for a reagent (at a small excess of the reagent with respect to the stoichiometric amount thereof), by an inhibition or catalytic acceleration of photometric reactions, by a wrong calibration of a measuring instrument *etc.* [12]. With

TABLE 4

Determination errors (δC, %) for papaverine (P) and bendazol (B) in three non-additive binary mixtures using different calculation algorithms (FM, MLR, PLS)

Mixtures	FM		MLR		PLS	
	P	B	P	B	P	B
1	17.3	12.1	3.1	−0.8	0.5	−0.2
2	−3.4	1.9	−3.1	2.4	0.7	−0.2
3	2.1	−3.9	−1.1	−2.3	0.9	−0.6

this in mind, one could intentionally change the analysis technique (concentration conditions wavelengths, methods for measuring the signals), with preventing the deviation from additivity. Studies performed in Kuban State University (Krasnodar, Russia) on evaluating the total content of antioxidants in foods could serve as an example [27]. It is important to understand how the OA influence upon the results analyzing the mixtures. For some algorithms, this connection was established theoretically [28, 29], for others it was determined in the course of experiments with model mixtures. In any case, the systematic errors of the analysis increase rapidly with increasing the OA [19]. Comparison of the results of analysis for the same mixtures with the use of a number of algorithms demonstrated that Vierordt's method is the most sensitive with respect to OA whereas – the PLS method is sensitive to a least extent (Table 4).

Potentialities of accurate estimating the total content

The objectives associated with determining the total content of a number of compounds related structurally or functionally (problem type 3), are rather poorly understood in a theoretical aspect, despite their importance. Most often, the total content of one-type components in a mixture are determined "as calculated for the standard substance". Determining the "phenolic index" (PI) could serve as an example. This represents an express evaluation of the total content of volatile phenolic compounds in natural and waste waters, as calculated for the simplest phenol C_6H_5OH. In order to determine the PI the phenols are distilled with steam to convert them into coloured derivatives via the reaction with 4-aminoantipyrine and to measure a generalized analytical signal (absorbance at a fixed wavelength in the visible region) and, finally, to calculate the PI using a calibration graph plotted for the standard solutions of C_6H_5OH [30].

The express evaluations of the total content of similar compounds, as calculated for any standard substance are called integral parameters (IP) those are widely used in the analysis of environmental objects, petroleum products and foodstuff (Fig. 4). However, the results of the analysis of the same sample, as calculated for different standard substances significantly differ between each other and do not correspond to the real total content of the components under investigation (c_Σ). At the same time, the value of the integral parameter (c^*) could be several-fold different from the value of c_Σ. Taking into account the inaccuracy, subjectivity and theoretically unsubstantiated nature of many IP, famous analyst Valcárcel called the IP a "black hole" in the chemical metrology [31]. However, no systematic researches in this field were carried out.

We have demonstrated that the occurrence of systematic errors in the evaluation of the

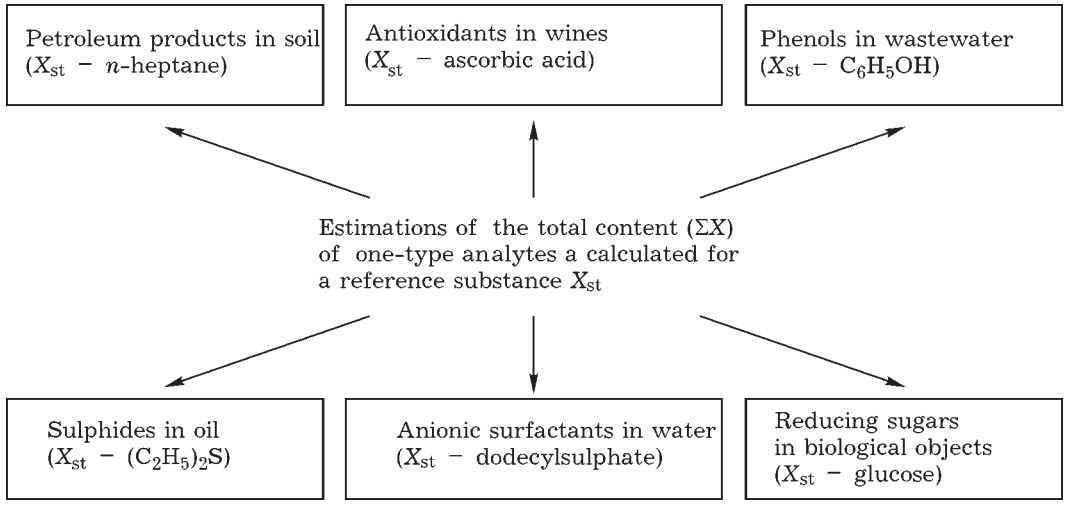

Fig. 4. Integrated parameters for estimating the total content of different analytes.

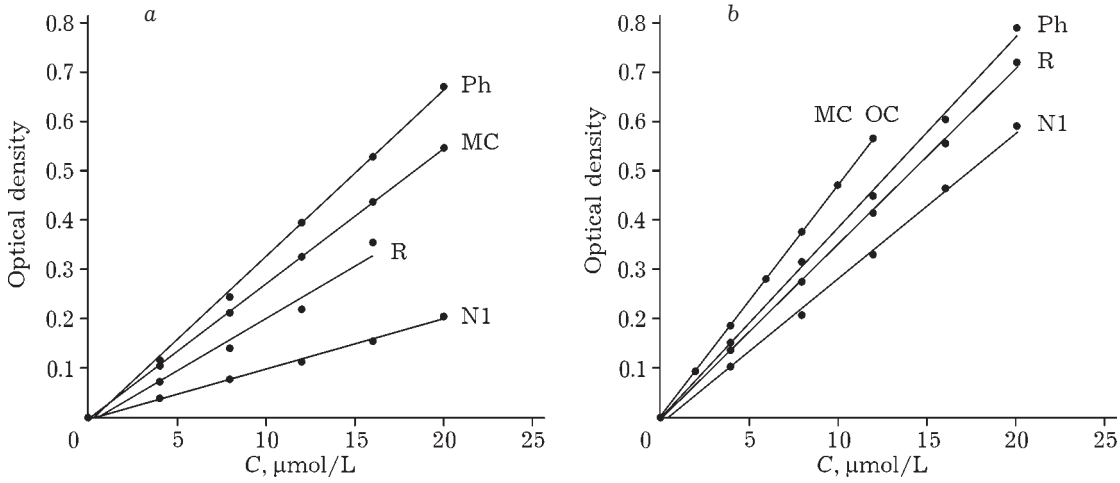

Fig. 5. Calibration fans in the case of determining phenols *via* the reaction with sulphanilic acid: *a* – before optimizing the technique (λ = 450 nm, pH 8.3, T = 3.4), *b* – after optimizing the technique (λ = 360 nm, pH 7.3, T = 1.5); Ph – phenol (carbolic acid), MC – *meta*-cresol, OC – *ortho*-cresol, R – resorcinol, N1 – 1-naphthol.

total content as calculated for the standard substance is caused by unequal detection sensitivity inherent in different components of the mixture, which is manifested in the form of a "fan of calibrations" (Fig. 5). An algorithm was developed for the prediction of relevant errors, taking into account the possible composition of a mixture under investigation and the nature of a standard substance [32]. It was found that the relative error $\delta C = (c^* - c_\Sigma)/c_\Sigma$ depends on the choice of the standard substance and a ratio between the components in the mixture, rather than on the total content thereof:

$$\delta C = \sum p_i R_i - 1 \qquad (1)$$

where $p_i = K_i/K_{st}$ is detection sensitivity with respect to i-th analyte normalized to the standard substance; R_i is the fraction of the i-th analyte in the mixture of compounds determined jointly. The prediction results for δC according to formula (1) are in a good agreement with the experiment. Appropriate verification was performed for different model mixtures with the use of different spectrometric methods, refractometry and conductometry.

The algorithm developed could be used for optimizing the methods of analysis (choosing the standard substance, the wavelength and so on). However, often it is not known exactly what the analytes from the given group at what the concentration ratio are present in the samples under investigation; it should be noted that without this information it is impossible to use the algorithm reported by the authors of [32]. Recently, the mentioned limitation has been overcome. We have begun to predict not the errors themselves, but the limiting values thereof, those do not depend on the qualitative and quantitative composition of either mixture, being dependent on the level of difference between the sensitivity coefficients inherent in different analytes of this group, *i. e.* on the fact how wide is the fan of calibrations. As the optimization parameter, it is suitable to use variable T equal to a ratio between the maximum and minimum sensitivity coefficients in determining the analytes of the given group:

$$T = K_{max}/K_{min} \qquad (2)$$

The wider is the fan of calibrations, the greater (in absolute value) could be the errors of the analysis of an undivided mixture. The total content of one-type analytes as calculated for some standard substance is determined with the errors falling within the range of

$$Q^{-1} - 1 \le \delta C \le W - 1 \qquad (3)$$

where $Q = K_{st}/K_{min}$, $W = K_{max}/K_{st}$, $QW = T$.

The greater is the value of T, the mentioned range is wider. The second factor *i. e.* the choice of a standard substance plays a role, too. As it was demonstrated by the authors of [33], the optimal value should be $K_{st} = 0.5(K_{min} + K_{max})$, *i. e.*, the calibration curve for the standard substance should be in the middle of the fan of calibrations obtained under the assay conditions for all the compounds in this group. This rec-

TABLE 5

Total content of polyphenoltype antioxidants (ΣC_{AO}) in five-component model mixtures using MLR algorithm

ΣC_{AO}, μmol/L		δC, %	s_r
Introduced	Found		
66	67.3	1.9	0.005
67	70.1	4.7	0.002
69	70.8	2.7	0.018
70	73.0	4.3	0.001
73	74.9	2.6	0.002
77	74.8	−2.8	0.030
78	74.1	−5.0	0.015
81	80.8	−0.2	0.006
83	83.5	0.6	0.006
85	86.0	1.2	0.004
89	90.5	1.7	0.005

ommendation was confirmed in the experiment, in particular, in the case of the spectrometric determination of the total phenol content in model mixtures with different composition. One should also minimize the width of the fan of calibrations varying the wavelength whereat the total absorption of mixtures is measured. One could also vary the pH of the solution or change the method of calculation of the integral parameter. After the optimization the parameter T could be reduced by 2–3 times, whereby the accuracy of estimates for the total content increases to a considerable extent: the errors exhibit a 3–4-fold decrease [27, 33].

Much more accurate estimates concerning the total content could be obtained in the case of registering the entire absorption spectrum of a sample, with further calculating the total content according to preliminarily constructed multivariate calibrations [34, 35]. In this case, the error in the analysis of mixtures with a known qualitative composition does not exceed 5 rel. % (Table 5).

In the course of the relevant studies there were results important for practice obtained. In particular, at the Omsk State University there have been developed "chemometric" methods for the structural group analysis of gasoline species based on the registration of a sample absorption spectrum in the near-infrared region and processing thereof using a PLS-1 algorithm [9]. Using novel techniques instead of the standard ASTM D.5134–98 method involving a complete separation of sample components, allows one to reduce the duration of the analysis by more than 20 times without decreasing the accuracy (Table 6). The techniques developed are used for the operational control and correction of technological processes at the Omsk Oil Refinery (Russia).

The only drawback of such methods consists in the fact that for each type of samples (*e. g.*, virgin gasoline species) researchers need to make and use its own multivariate calibration. Any attempts to apply the same calibration for the analysis of another type samples (for example, for the reforming of gasoline fractions) result in great errors. In this regard, in the course of the analysis, the correspondence between the next sample spectrum and the calibration used is automatically verified. Otherwise, for the analysis of this sample it is necessary to use a different model (preliminary loaded to the computer memory). The selection of adequate training samples, constructing and testing the calibrations require for a consider-

TABLE 6

Structural group composition of virgin gasoline species, determined by means of the standard (GLC) and the developed (NIR) methods

Sample	Naphthenes, mass %			Paraffins, mass %			Arenes, mass %		
number	GLC	NIR	δ, %	GLC	NIR	δ, %	GLC	NIR	δ, %
1	48.03	48.08	0.1	43.27	43.34	0.2	8.69	8.52	−2.0
2	48.25	48.02	−0.5	42.71	43.02	0.7	9.04	8.82	−2.5
3	48.94	48.81	−0.3	43.70	43.85	0.3	7.34	7.48	1.9
4	48.61	47.80	−1.7	45.00	44.49	1.1	7.38	7.77	5.3

Note. δ is relative deviation of the results, rel. %.

able labour and time consumption. However, the further analysis of each sample takes only 5?10 min. In is obvious that the spectrometric determination of the group composition of gasoline species is worthwhile in the case of the mass analysis of the samples of the same type, whereas for individual analysis of different samples it is better to use the GLC technique.

The qualitative composition of the objects under analysis is usually far from being completely known, and it is not clear what the compounds at what a ratio should be used in the formation of the training sample. This equivocation makes it difficult to use chemometric algorithms in order to estimate the total content of similar analytes. One could get out of the situation in the two ways: 1) to include in the training sample as many mixtures as possible with a qualitative composition not completely known, but to determine the total content of analytes according to a standard (reference) method in each the mixture; 2) to create and to use an information-redundant mathematical model. The first approach is used when there is a reliable reference method. So, the total content of paraffins, naphthenes and arenes in "training" samples of gasoline were preliminary determined in advance according to ASTM D.5134–98 using GLC technique [9], whereas the total content of polyphenols in tea was determined by means of Folin–Ciocalteu method [34].

The second approach could be used when there is no appropriate reference technique. In this case, the "training" mixtures are prepared using accurately weighed sample portions of individual analytes, as much as possible in the number thereof. Next, one calculate a multivariate calibration (information-redundant model). Using this model, basing on the real sample spectrum one can calculate the concentration of the compounds used in the formation of the model. The concentration of the compounds those are absent in the sample, appear zero or negative, whereas the values for the compound present therein are positive [35]. This allows one to specify the qualitative composition of the mixture, whereas further summing the concentrations of the identified components gives a fairly accurate (±10 rel. %) estimate of the total content of analytes.

CONCLUSIONS

It should be noted that the development of the direct spectrometric analysis with respect to unseparated mixtures using chemometric algorithms is just under beginning, but the fact that the novel method is promising, gives already no rise to doubt. Of course, the mentioned method would never replace the chromatographic analysis, but it could and should be an important supplement, and in some cases a good alternative of the latter.

For the development of spectrometric analysis of unseparated mixtures, of particular interest are the following promising areas of research:

– Improving the accuracy of analyzing the unseparated mixtures in adverse situations (similarity of reference spectra, a large excess of either analytes in comparison with others, and so on);

– The use of multi-wavelength spectroscopy and chemometric algorithms for studying the chemical reactions in solutions;

– Improving the accuracy of estimating the total content of analytes using chemometric algorithms and integral parameters, particularly in the analysis of environmental objects.

Successfully solving the three major problems of the spectrometric analysis of unseparated mixtures (inter-reference overlapping, non-additivity, different sensitivity in the detection of similar compounds) paves the way to develop many good methods for express analysis of a variety of objects. In particular, the "computerized" spectrometry could be used to control the composition of liquid and gaseous reaction mixtures in the implementation of catalytic reactions, as well as to study of the kinetics and mechanism of corresponding chemical reactions.

REFERENCES

1 Zolotov Yu. A., Vershinin V. I., Istoriya i Metodologiya Analiticheskoy Khimii, Akademiya, Moscow, 2007.

2 Bershtein I. Ya., Kaminskiy Yu. A., Spektrofotometricheskiy Analiz v Organicheskoy Khimii, Khimiya, Moscow, 1986.

3 Vasiliev A. F., Teoreticheskiye Osnovy Sovremennykh Metodov Kolichestvennogo Analiza Mnogokomponentnykh Sistem po Spektram Pogloshcheniya (Doctoral Dissertation in Chemistry), Moscow, 1976.

4 Korenman I. M., Fotometricheskiy Analiz. Metody Opredeleniya Organicheskikh Soyedineniy, Khimiya, Moscow, 1975.
5 Esbensen K., Analiz Mnogomernykh Dannykh, Chernogolovka, 2005.
6 Gribov L. A. (Ed.), Matematicheskiye Metody i EVM v Analiticheskoy Khimii (Treatises), Nauka, Moscow, Nauka, Moscow, 1989.
7 Arzamastsev A. P., Sadchikova N. P., Titova A. P., *Vopr. Biol., Med. Farm. Khim.*, 1 (2010) 16.
8 Ivanova L. V., Safieva R. Z., Koshelev V. N., *Vestn. Bashkir. Un-ta*, 13, 4 (2008) 869.
9 Vershinin V. I., Kopteva E. V., Troitskiy V. V., *Zavod. Lab. Diagn. Mater.*, 71, 11 (2005) 10.
10 Vershinin V. I., Derendyaev B. G., Lebedev K. S., Kompyuternaya Identifikatsiya Organicheskikh Soyedineniy, IKTs "Akademkniga", Moscow, 2002.
11 Khots M. S., in: Matematicheskiye Metody i EVM v Analiticheskoy Khimii (Treatises), in Gribov L. A. (Ed.), Nauka, Moscow, 1989, pp. 87–99.
12 Vlasova I. V., Vershinin V. I., Tsyupko T. G., *Zh. Anal. Khim.*, 66, 1 (2011) 24.
13 Monakhova Yu. B., Astakhov S. A., Mushtakova S. P., *Zh. Anal. Khim.*, 64, 5 (2009) 495.
14 Monakhova Yu. B., Astakhov S. A., Kraskov A. V., *Chem. Intel. Lab. Syst.*, 103, 2 (2010) 108.
15 Monakhova Yu. B., Mushtakova S. P., Kolesnikova S. S., Mushtakova S. P., *Anal. Bioanal. Chem.*, 397, 3 (2010) 1297.
16 Rodionova O. E., Pomerantsev A. L., *Usp. Khim.*, 75, 4 (2006) 302.
17 Vlasova I. V., Masyakova E. N., Koryagina A. V., *Zavod. Lab. Diagn. Mater.*, 76, 2 (2010) 18.
18 Brereton P. G., Chemometrics. Data Analysis for Laboratory and Chemical Plant, Wiley, NY, 2003.
19 Vlasova I. V., Vershinin V. I., Shelpakova A. S., *Vestn. OmGU*, 2 (2010) 14.
20 Vlasova I. V., Shelpakova A. S., Vershinin V. I., *Zavod. Lab. Diagn. Mater.*, 75, 4 (2011) 19.
21 Vlasova I. V., Buryukina P. A., *Vestn. OmGU*, 2 (2012) 123.
22 Vlasova I. V., Shelpakova A. S., Nagaev A. A., *Vestn. OmGU*, 2 (2010) 99.
23 Khots M. S., Nazarov V. I., Lyovin A. A., *Chem. Intel. Lab. Syst.*, 18, 3 (1993) 281.
24 Vershinin V. I., Vlasova I. V., Tsyupko T. G., Nikolaeva N. A., Kharkova M. A., *Zh. Anal. Khim.*, 66, 7 (2011) 708.
25 Vershinin V. I., Vlasova I. V., Tsyupko T. G., *Metody Obyekty Khim. Analiza*, 5, 4 (2010) 226.
26 Usova S. V., Bogza Yu. P., Goncharov D. S., Vershinin V. I., *Anal. Kontrol*, 15, 1 (2011) 78.
27 Tsyupko T. G., Petrakova I. V., Brilenok N. S., Nikolaeva N. A., Chuprynina D. A., Temerdashev Z. A., Vershinin V. I., *Anal. Kontrol*, 15, 3 (2011) 287.
28 Vlasova I. V., Vershinin V. I., *Zh. Anal. Khim.*, 64, 6 (2009) 571.
29 Vlasova I. V., Isachenko N. A., Shilova A. V., *Zh. Anal. Khim.*, 65, 5 (2010) 481.
30 Vorobyeva T. V., Terletskaya A. V., Kushchevskaya N. F., *Khim. Tekhnol. Vody*, 4 (2007) 370.
31 Baena J. R., Valcarcel M., *Trends Anal. Chem.*, 10 (2003) 641.
32 Vershinin V. I., Brilenok N. S., Tsyupko T. G., *Zh. Anal. Khim.*, 67, 7 (2009) 715.
33 Vershinin V. I., Kuleshova M. P., Isachenko N. A., Shiligin P. V., *Zh. Anal. Khim.*, 68, 6 (2013) 535.
34 Quansheng C., Jiewen Z., Muhua L., *J. Pharm. Biomed. Anal.*, 46, 3 (2008) 568.
35 Vlasova I. V., Tsyupko T. G., Shelpakova A. S., *Metody Obyekty Khim. Anal.*, 7, 1 (2012) 118.