

ON FOOD, SPECTROPHOTOMETRY, AND MEASUREMENT DATA PROCESSING (KEYNOTE LECTURE)

Roman Z. Morawski

Warsaw University of Technology, Faculty of Electronics and Information Technology, Institute of Radioelectronics
Warsaw, Poland
r.morawski@ire.pw.edu.pl

Abstract: Spectrophotometry is getting more and more often the method of choice not only in laboratory analysis of (bio)chemical substances, but also in the off-laboratory identification and testing of physical properties of various products, in particular – of various organic mixtures including food products and ingredients. Specialized spectrophotometers, called spectrophotometric analyzers are designed for such applications. This keynote lecture is on the state of the art and developmental trends in the domain of spectrophotometric analyzers of food with particular emphasis on wine analyzers. The following issues are covered: philosophical and methodological background of food analysis, physical and metrological principles of spectrophotometry, the role of measurement data processing in spectrophotometry, food analyzers on the market and their future. General considerations are illustrated with examples, predominantly related to wine analysis.

Keywords: spectrophotometry, chemometrics, spectral data processing, food analysis, wine analysis.

1. PHILOSOPHICAL BACKGROUND

The pragmatic Americans used to say: *"You are what you eat."* On the other hand, the elevation of mind over body, of reason over senses, of man over beast, and of culture over nature is a characteristic feature of the Western philosophical tradition. In particular, the dismissal of food as a proper subject for philosophical inquiry is deeply rooted in the history of thought. The denigration of food and cooking can be traced as far back as Plato for whom cooking was a simple practical skill, as opposed to a genuine art like medicine. There are at least two alternative explanations for such views, *viz.*: food and taste have been ignored because they have been mistakenly regarded as unimportant, or – they have been regarded as too important to be exposed to risky philosophical rumination. The situation is changing now in contemporary Western societies because there has been a significant shift in how we define ourselves: today we are inclined to link our identities much more to what we consume than to our social roles as we used to do a generation or two ago.

Despite philosophical abnegation of food issues, the practice of food preparation and refinement has flourished for centuries, and inspired research-and-invention-oriented minds. Today, we may speak about a fully developed discipline of science and technology. Enough to say that the Institute of Food Technologists – the largest international, non-profit professional organization involved in the advancement of food science and technology – is encompassing 23 000 members worldwide. Its Committee on Higher Education provided the following definition of food science: *"Food science is the discipline in which the engineering, biological, and physical sciences are used to study the nature of foods, the causes of deterioration, the principles underlying food processing, and the improvement of foods for the consuming public"* [1]. This definition is emphasizing the empirical nature of food science and – consequently – its strong affinity to measurement and instrumentation. This is the aspect of food science we are here interested in. A variety of measurement methods and techniques that are currently applied in food science and technology is encompassing both simple tools and procedures for measuring mass or temperature and very sophisticated ones, such as computer-based measuring systems of chromatography combined with mass spectrometry. This lecture is devoted to food applications of spectrophotometry, with particular emphasis on the methods for processing raw spectrophotometric data. The choice of that measurement technique is motivated by a rapid increase of its applications both in terms of the repertory of measurands and in terms of the number of installed spectrophotometric transducers, instruments and systems. This increase is a response to a growing demand for adequate methods of food analysis, driven by several factors. Despite the legislative efforts, aimed at protecting consumers and ensuring fair trade, food scandals are reported quite frequently. Their origin is most often related to pesticide residues, veterinary drug residues, endocrine disruptors, processing contaminants, packaging materials and natural toxins. On top of that, genetically modified food products and ingredients have become an issue, at least in the European Union.

2. PRACTICAL BACKGROUND

In July 2006, an exciting news appeared in the media, both popular and professional, announcing that a research team of *NEC System Technologies* and *Mia University* in Japan designed and prototyped a robot able to check whether the contents of a bottle of wine corresponds to its label [2]. The robot's performance was limited to several dozens of kinds of wine, but the promise of its database extension open prospects for further developments. Automatic analyzers of beverages, based on various measurement principles, had been known for some time, but that happening opened a new chapter in their development since wine testing is a task much more complex than testing of any other kind of alimentary product. Why? Because its aim is not only to differentiate between white and red wine, between port and champagne, or between French and Chilean *Cabernet Sauvignon* ... A robot for wine testing is expected to be able to state, with sufficiently low uncertainty, that the contents of a bottle has all the features promised by its label, while even a very modest label is providing a lot of information about the geographical origin and vintage of wine, the sort of grapes it has been made of, the type of barrels it has been aged in, *etc.* Much more information remains behind the scene: sensory properties of wine, identified by someliers using organoleptic means, and its chemical properties, identified by oenology expert equipped with sophisticated analytical instruments.

The robot for wine testing, launched in 2006, was a successful fusion of two powerful technologies: near-infrared spectrophotometry and measurement data processing. During last two decades, this approach has provided a quite universal solution for the analysis and testing of many alimentary products, especially outside analytical laboratories. Spectrophotometric analyzers for grain and fruits, milk and beer, chocolate and cheese, *etc.*, are today manufactured by numerous companies all over the world. Their intensive development is motivated by many factors, including growing awareness of the nutritional issues, and increased incidence of fraud in the domain of fabrication and distribution of such products as wine, olive oil or cheese. This keynote lecture contains an outline of a broad engineering and application background of that technology and a more detailed review of the state of the art in the domain of measurement data processing dedicated to spectrophotometric analyzers.

The R&D works on automatic wine testing are smoothly moving forward because the practical value of the results already achieved is opening prospects for wide application of robots in wine production and distribution, and – consequently attracting potential sponsors and investors. The motivation is, of course, of economic nature: a robot for checking the contents of a bottle of wine will become an effective tool for fighting the fraud on the wine markets. Let's consider a simple example in order to imagine the financial dimension of this social issue. Let's assume that 900 bottles of wine, priced EUR 20 per bottle, has been mixed with 100 bottles of wine, priced EUR 5 per bottle. The sold of 1000 bottles of this mixture, at the price of EUR 20 per bottle, will generate the extra income of EUR 1 500! That's why the robot is attracting an interest of

wine makers, wine dealers, custom officers... The detection of fraud by means of a robot will be much easier, quicker and cheaper than by means of today's procedures referring to time-consuming laboratory analyses and expensive experts' services.

The Japanese robot, whose size is comparable with two four-bottle wine containers, is able to analyze and identify a 5-ml sample of wine within 30 seconds. Its functioning is based on the principle of spectrophotometry in the near-infrared range and on the sophisticated algorithms for processing raw spectrophotometric data by means of an internal computer. This technological combination seems today to engender the most promising solutions in the domain of wine robotics.

3. SPECTROPHOTOMETRIC ANALYZERS

Spectrophotometric tools, and – consequently – the methods for interpretation of spectrophotometric data, are of increasing importance for analytical laboratories, as well as for environmental, biomedical and industrial monitoring. On the one hand, the development of corresponding applications is driven by a growing demand for this kind of tools, the demand implied by the advancement of standards related to environment protection, health care, individual and collective security, as well as by the widespread use of optical means for inspection of industrial procedures. On the other hand, however, this development is due to the market availability of miniature spectrophotometers: mini- and micro-spectrophotometers [3-9].

The term "spectrophotometric analyzer" is used here for various spectrophotometric sensors, devices, instruments, probes and testers dedicated to measuring physical and/or chemical parameters characterizing a pre-defined class of chemical or biochemical substances. Thus, a spectrophotometric transducer (ST) is the heart of any such analyzer. It is converting an optical signal into a sequence of raw data $\tilde{\mathbf{y}} = [\tilde{y}_1 \dots \tilde{y}_N]^T$ representative of the spectrum $x(\lambda)$ of that signal, where λ is wavelength, x is light intensity, N is the number of data, and the tilde placed over y_1, \dots, y_N is to indicate that the data are subject to various disturbances of ST-external and ST-internal origin. The wavelength values cover a broader or narrower subrange $[\lambda_{\min}, \lambda_{\max}]$ of one of the following standard intervals: 200–300 nm – middle-ultraviolet radiation (MUV), 300–380 nm – near-ultraviolet radiation (UV), 380–750 nm – visible radiation (Vis), 750–2,500 nm – near-infrared radiation (NIR) or 2.5–10 μm – middle-infrared radiation (MIR). Several physical principles and corresponding devices may be used for ST design:

- a dispersive element (a grating or a linear variable filter) that enables separation of spectral components in space;
- a tuneable filter that enables separation of spectral components in time;
- an optical heterodyne that enables shifting the spectrum in a wavelength range where its analysis is easier;
- an interferometer providing the data whose Fourier transform is representative of the spectrum.

The use of one of them, most frequently applied today, *viz.* of a dispersive element, is explained in Fig. 1:

- An optical signal to be converted into a digital representation of its spectrum is separated into N optical signals, corresponding to narrow subranges of wavelength λ by a dispersive element, *e.g.* a grating.
- Each of those signals is reaching a corresponding photodetector PD converting it into a current proportional to its intensity.
- The output currents of the photodetector matrix are scanned by an analogue-to-digital converter and, one by one, coded.

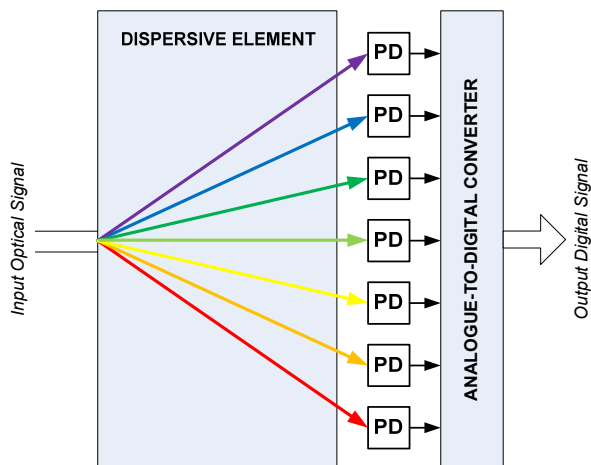


Fig. 1. A spectrophotometric transducer (ST) based on a dispersive element – the principle of functioning.

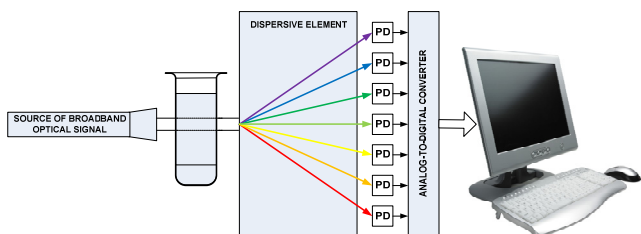


Fig. 2. A spectrophotometer based on a dispersive element – the principle of functioning.

A functional diagram of a spectrophotometric analyzer, based on the ST from Fig. 1, is shown in Fig. 2. Its functioning may be explained as follows:

- First, the source optical signal is converted into the data \tilde{y}_0 representative of its intensity spectrum $x_0(\lambda)$.
- Next, the same source optical signal is passed through a cuvette, containing a sample of the substance to be analyzed, and converted into the data \tilde{y}_1 representative of its intensity spectrum, modified by a sample, $x_1(\lambda)$.
- Finally, the intensity data \tilde{y}_0 and \tilde{y}_1 are processed by a digital processor (a controller, a microprocessor, a digital signal processor or a personal computer) in order to obtain an estimate of the target parameters of the sample, such as concentrations c_1, \dots, c_J of selected compounds of that sample.

The estimation of target parameters is, as a rule, an ill-conditioned numerical problem of *measurand*

reconstruction. Therefore, its solution may require not only simple conversion of codes and estimation of the transmittance or absorbance spectrum on the basis of two intensity spectra, but also some more sophisticated operations, aimed at extraction of information hidden in the data, such as estimation of the optical noise level, estimation of the spectrum with a better resolution than the resolution of the ST or estimation of concentrations on the basis of pre-processed data. In some cases, estimation of some intermediate or auxiliary parameters, such as positions and magnitudes of spectral peaks, may be necessary, because the positions of peaks in the absorbance spectrum may facilitate identification of the compounds which are present in the sample; and their magnitudes may facilitate estimation of the concentrations of the identified compounds.

The complexity of spectrophotometric data processing significantly depends on the technological and functional imperfections of the ST. Their external manifestations – such as noise-type disturbances in the data or blurring and overlapping of spectral peaks – must be taken into account during data processing. Therefore, the corresponding algorithms, as a rule, require considerable amount of information on the mathematical model of the ST (or on the mathematical model of the ST data at its output). This information is acquired during an operation, performed off-line, called *calibration of the ST*, which is including not only scaling of wavelength axis and intensity axis, but also the identification of a forward or of an inverse mathematical model of the ST – the identification based on the use of the ST data corresponding to some standards of the spectrum or of a class of substances the analyzer is designed for.

The interpretation of the intensity spectra $x_0(\lambda)$ and $x_1(\lambda)$, in terms of concentrations (or other physical or chemical parameters) characterizing a sample under study, may be simplified if those spectra are transformed into the transmittance spectrum defined as:

$$x^{Tr}(\lambda) = \frac{x_1(\lambda)}{x_0(\lambda)} \quad (1)$$

The data \tilde{y}_0 and \tilde{y}_1 should be then converted into the transmittance data, representative of $x^{Tr}(\lambda)$, according to the formula:

$$\tilde{y}_n^{Tr} = \frac{\tilde{y}_{1,n} - b_n}{\tilde{y}_{0,n} - b_n} \quad \text{for } n = 1, \dots, N \quad (2)$$

where $\mathbf{b} = [b_1 \dots b_N]^T$ is the ST response to the zero-intensity optical signal. The interpretability of spectra may be further improved by using the absorbance scale defined as follows:

$$x^{Ab}(\lambda) = -\log_{10}(x^{Tr}(\lambda)) \quad (3)$$

and the corresponding absorbance data:

$$\tilde{y}_n^{Ab} = -\log_{10}(\tilde{y}_n^{Tr}) \quad \text{for } n = 1, \dots, N \quad (4)$$

An example of transmittance data is shown in Fig. 3. The result of their transformation into absorbance data – in

Fig. 4. The absorbance data emphasize the most informative parts of the spectrum, *viz.* absorption peaks whose parameters, positions and magnitudes, are determined by the qualitative and quantitative composition of the sample under study. The image of those data is a finger-print enabling the identification of the sample.

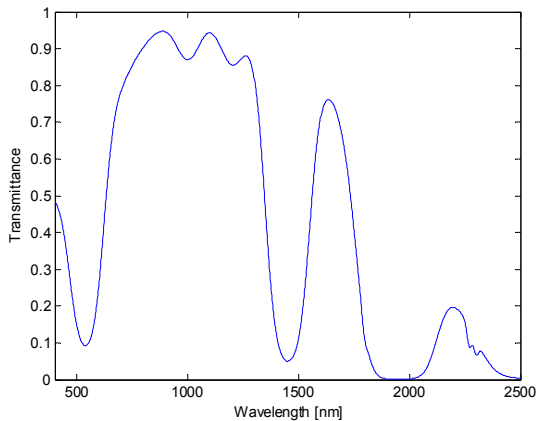


Fig. 3. The data representative of the transmittance spectrum of red wine.

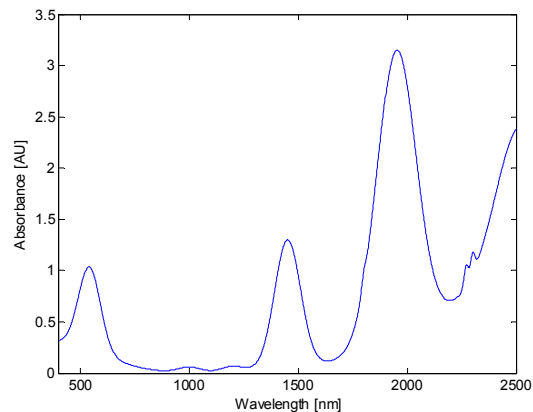


Fig. 4. The data representative of the absorbance spectrum of red wine.

4. SPECTROPHOTOMETRIC ANALYZERS OF FOOD

The term *spectrophotometric analyzer of food* (SAF) is used here for various spectrophotometric sensors, devices, instruments, probes and testers dedicated to measuring physical and (bio)chemical parameters characterizing food products (including beverages) and ingredients.

As mentioned in Section 2, NIR spectrophotometry is of particular usefulness for food analysis because the spectra of organic samples comprise broad bands arising from overlapping absorption peaks corresponding to C-H, O-H and N-H chemical bonds. The main advantage of NIR spectrophotometry, when applied for routine analysis of food, is its simplicity and speed: usually no sample preparation is necessary and the time of analysis is not greater than 1 minute. Another advantage of NIR spectrophotometry is that it allows several constituents to be identified concurrently. Finally, the relatively weak absorption due to water enables one to analyze high-moisture food products and ingredients.

The driving forces behind the development of SAFs may be easily identified on the basis of the general structure of the food business, shown in Fig. 5, and of the specification of needs of its main actors, provided in Table 1.

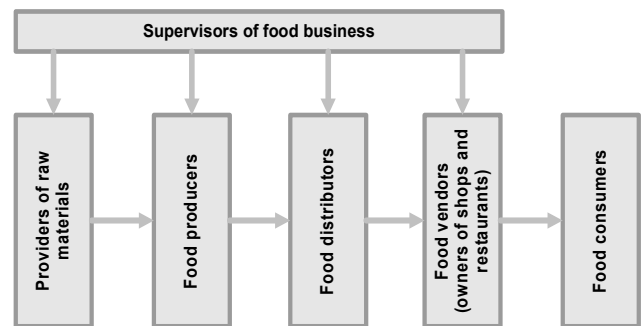


Fig. 5. The general structure of the food business.

Table 1. The motivation for using spectrophotometric analyzers by the main actors of the food business.

	Specific examples of SAFs	Reasons to use SAFs
Providers of raw materials	Analyzers of fruits ripeness Analyzers of moisture content in grains Analyzers of meat tenderness Analyzers of raw milk	Checking the quality of raw materials Specification of raw materials enabling their labelling Precise classification of raw materials enabling their better pricing
Food producers	Analyzers of olive oil: chemical contents Analyzers of olive oil: geographical origin Analyzers of chocolate products Analyzers of cheese	Monitoring of the production from the verification and selection of raw materials to the final products Providing data necessary for manufacturing control Specification of final products enabling their labelling Precise classification of final products enabling their better pricing
Food distributors	Analyzers of flour Analyzers of honey: chemical contents (sugars) Analyzers of honey: floral origin Analyzers of coffee grains	Checking the quality of products to be distributed Monitoring of the quality of stored products Precise classification of products to be distributed, enabling their better pricing Increased productivity
Food vendors	Analyzers of water (for restaurants) Analyzers of bread Analyzers of dairy products Analyzers of fruits (aging)	Checking the quality of products to be sold Monitoring of the quality of stored products Precise classification of products for sale, enabling their better pricing
Food consumers	Analyzers of mineral water Analyzers of wine Analyzers of beer Analyzers of tea	Checking the quality of purchased products Monitoring of the quality of stored products Improved selectivity in purchasing food products (based on better quality price evaluation)
Supervisors	Analyzers for microbiological evaluation of food Analyzers for toxicological evaluation of food	Checking the quality of food products <i>in situ</i> Increased effectiveness of supervision

All those actors are potential users of SAFs, but the measurement needs are today better defined and the use of SAFs is more common among the food producers and supervisors of food business than among the providers of raw materials, food distributors, food vendors and individual consumers. It may be predicted, however, that the significant lowering of the prices of SAFs will quickly increase the demand among the members of the latter group of food business actors. As shown in Table 1, there are several reasons for using SAFs by the main actors of the food business. They may be summarized as follows:

- checking the quality of food;
- monitoring of the food production process;
- providing data necessary for production control;
- specification of food products necessary for their labelling;
- precise classification of food products enabling their better pricing.

There are more-or-less evident economic benefits behind each of them. Precise classification, for example, enables one to minimize losses due to aging of food by replacing a worst-case approach with a realistic-case approach since the selection and grading of food may be based on the objective measurement results rather than on the "best-before" date.

4. MATHEMATICAL MODELLING OF SPECTROPHOTOMETRIC DATA

The contents of this and two following sections are based on an author's review paper, devoted to spectrophotometric data processing [10], where many more details and explanations may be found.

4.1. Mathematical modelling of intensity data

The intensity data, provided by an ST, may be modelled using a white-box approach, a black-box approach, or a grey-box approach combining some advantages of white-box and black-box approaches [11]. 50 years of experience behind modelling of spectrophotometric data seems to support the conclusion that the approximation power of the *Wiener operator*, *i.e.* of a superposition of a linear integral operator with a nonlinear algebraic operator, is sufficient for adequate modelling of the relationship between the intensity spectrum $x(\lambda)$ and the corresponding raw data $\tilde{\mathbf{y}}$. Let the variable \hat{y}_n denote the mathematical model of the "noise-free" version of the datum \tilde{y}_n . Then this operator may be given the following form:

$$\hat{y}_n = F \left(\int_{-\infty}^{+\infty} g_n(\lambda_n - \lambda) x(\lambda) d\lambda; \mathbf{a}_n \right) \text{ for } n = 1, \dots, N \quad (5)$$

where:

- $g_n(\lambda)$ is the response of the ST, measured at the output of the n th photodiode, to a tuneable monochromator producing an optical signal whose spectrum is close to $x(\lambda) \equiv \delta(\lambda - l)$ where l is moving from λ_{\min} to λ_{\max} ;
- $F(\bullet; \mathbf{a}_n)$ is an *a priori* known function (*e.g.* an algebraic polynomial, a cubic spline) whose parameters,

$\mathbf{a}_n = [\alpha_{n,1} \alpha_{n,2} \dots]^T$, have to be determined during the ST calibration.

4.2. Mathematical modelling of transmittance data and their dependence on concentrations

The reasoning presented for the intensity-domain data may be applied to the transmittance-domain data. As a rule, due to the compensation of some irregularities by the division of corresponding intensity data, the function $F(\bullet; \mathbf{a}_n)$ may be less complex in this case, and the variability of \mathbf{a}_n and $g_n(\lambda)$ along the wavelength axis – less important. Consequently, in many cases this variability may be ignored, and the corresponding model of the transmittance-domain data may be simplified by fixing: $\mathbf{a}_n \equiv \mathbf{a}$ and $g_n(\lambda) \equiv g(\lambda)$ for $n = 1, \dots, N$. This analysis justifies the use of the following mathematical model of the relationship between the transmittance spectrum $x^{Tr}(\lambda)$ and the corresponding data $\tilde{\mathbf{y}}^{Tr} = [\tilde{y}_1^{Tr} \dots \tilde{y}_N^{Tr}]^T$:

$$\tilde{y}_n^{Tr} = \int_{-\infty}^{+\infty} g(\lambda_n - \lambda) x^{Tr}(\lambda) d\lambda + \eta_n \text{ for } n = 1, \dots, N \quad (6)$$

where the additive residuals η_n represent the total uncertainty of data modelling.

For the purpose of modelling the dependence of the data on concentrations, one may use the Lambert-Beer laws of absorption to relate $x^{Tr}(\lambda)$ in Eq.(6) to the vector of concentrations $\mathbf{c} = [c_1 \dots c_J]^T$:

- the absorbance of a solution of a single component is proportional to its concentration;
- the absorbance spectrum of a multicomponent solution:

$$x^{Ab}(\lambda) = -\log_{10} [x^{Tr}(\lambda)] \quad (7)$$

equals the linear combination:

$$x^{Ab}(\lambda) = c_1 x_1^{Ab}(\lambda) + \dots + c_J x_J^{Ab}(\lambda) \quad (8)$$

of the normalized absorbance spectra of the components $x_1^{Ab}(\lambda), \dots, x_J^{Ab}(\lambda)$.

Eq.(6), Eq.(7) and Eq.(8) should undergo discretization to become a useful basis for the development of numerical methods for estimation of concentrations. Thus, Eq.(6) is replaced with:

$$\tilde{\mathbf{y}}^{Tr} = \mathbf{G} \cdot \mathbf{x}^{Tr} + \boldsymbol{\eta} \quad (9)$$

where:

$$\boldsymbol{\eta} = [\eta_1 \dots \eta_N]^T \quad (10)$$

$$\mathbf{x}^{Tr} = [x^{Tr}(\lambda'_1) \dots x^{Tr}(\lambda'_M)]^T \quad (11)$$

$$\lambda'_m = \lambda_{\min} + (m-1) \frac{\lambda_{\max} - \lambda_{\min}}{M-1} \text{ for } m = 1, \dots, M \quad (12)$$

and \mathbf{G} is a rectangular matrix whose values depend on the values of the function $g(\lambda)$ and on the chosen method of numerical integration. The discretization of Eq.(7) and Eq.(8), consistent with Eq.(9), yields:

$$\mathbf{x}^{Ab} = [-\log_{10}(x^{Tr}(\lambda_1)) \dots -\log_{10}(x^{Tr}(\lambda'_M))]^T \quad (13)$$

$$\mathbf{x}^{Ab} = c_1 \mathbf{x}_1^{Ab} + \dots + c_J \mathbf{x}_J^{Ab} = \mathbf{X}^{Ab} \mathbf{c} \quad (14)$$

where:

$$\mathbf{x}_j^{Ab} = [x_j^{Ab}(\lambda_1) \dots x_j^{Ab}(\lambda'_M)]^T \quad \text{for } j=1, \dots, J \quad (15)$$

$$\mathbf{X}^{Ab} = [\mathbf{x}_1^{Ab} \dots \mathbf{x}_J^{Ab}]^T \quad (16)$$

The up-to-now developed model of the relationship $\mathbf{c} \rightarrow \tilde{\mathbf{y}}^{Tr}$ characterizes this relationship for one particular pair of \mathbf{c} and $\tilde{\mathbf{y}}^{Tr}$. It should be generalized on the populations of admissible vectors \mathbf{c} and $\tilde{\mathbf{y}}^{Tr}$ to enable the use of probabilistic tools. This may be achieved by introducing random vectors (denoted hereinafter with underlined symbols) for modelling unknowns. The randomization of the vector $\boldsymbol{\eta}$, representative of modelling uncertainty, yields:

$$\tilde{\mathbf{y}}^{Tr} = \mathbf{G} \cdot \mathbf{x}^{Tr} + \boldsymbol{\eta} \quad (17)$$

where $\boldsymbol{\eta} = [\eta_1 \dots \eta_N]^T$. The model of the relationship between concentrations and absorbance spectrum, corresponding to Eq.(14), takes on the form:

$$\underline{\mathbf{x}}^{Ab} = \mathbf{X}^{Ab} \underline{\mathbf{c}} + \underline{\boldsymbol{\varepsilon}} \quad (18)$$

where $\underline{\mathbf{c}} = [c_1 \dots c_J]^T$ is the random vector of the concentrations to be estimated, $\underline{\boldsymbol{\varepsilon}} = [\varepsilon_1 \dots \varepsilon_M]^T$ is the random vector of the residual spectrum corresponding to unexpected or neglected components of the analyzed sample, and CAT is the operator of absorbance-to-transmittance conversion:

$$\underline{\mathbf{x}}^{Tr} \equiv [\underline{x}_1^{Tr} \dots \underline{x}_M^{Tr}]^T = \text{CAT}[\underline{\mathbf{x}}^{Ab}] = [10^{-\underline{x}_1^{Ab}} \dots 10^{-\underline{x}_M^{Ab}}]^T \quad (19)$$

The substitution of Eq.(18) to Eq.(19) and then of Eq.(19) to Eq.(17) yields a compact-form, discretised and randomized model of the data:

$$\tilde{\mathbf{y}}^{Tr} = \mathbf{G} \cdot \text{CAT}[\mathbf{X}^{Ab} \underline{\mathbf{c}} + \underline{\boldsymbol{\varepsilon}}] + \boldsymbol{\eta} \quad (20)$$

4.3. Soft modelling in latent variables

In many practically important cases, the model of absorbance data, corresponding to Eq.(18):

$$\underline{\mathbf{y}}^{Ab} = \text{CAT}^{-1}[\mathbf{G} \cdot \text{CAT}[\mathbf{X}^{Ab} \underline{\mathbf{c}} + \underline{\boldsymbol{\varepsilon}}] + \boldsymbol{\eta}] \quad (21)$$

may be linearized :

$$\tilde{\mathbf{y}}^{Ab} = \mathbf{H}[\mathbf{X}^{Ab} \underline{\mathbf{c}} + \underline{\boldsymbol{\varepsilon}}] + \boldsymbol{\eta} \quad (22)$$

The linearity of the relationship between absorbance data $\tilde{\mathbf{y}}^{Ab}$ and concentrations \mathbf{c} is a justification for the widespread use of the following inverse model for estimation of concentrations:

$$\mathbf{c} = \mathbf{P}\mathbf{y} \quad (23)$$

where \mathbf{y} is a vector of selected elements of the vector $\tilde{\mathbf{y}}^{Ab}$, and \mathbf{P} is a matrix of parameters to be estimated during calibration on the basis of the reference data:

$$\langle \mathbf{c}_1, \mathbf{y}_1 \rangle, \langle \mathbf{c}_2, \mathbf{y}_2 \rangle, \dots \quad (24)$$

usually arranged into two matrices:

$$\mathbf{Y} = [\mathbf{y}_1 \mathbf{y}_2 \dots] \text{ and } \mathbf{C} = [\mathbf{c}_1 \mathbf{c}_2 \dots] \quad (25)$$

and assumed to satisfy the equations:

$$\mathbf{c}_1 = \mathbf{P}\mathbf{y}_1 + \boldsymbol{\gamma}_1, \mathbf{c}_2 = \mathbf{P}\mathbf{y}_2 + \boldsymbol{\gamma}_2, \dots \quad (26)$$

where the vectors $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \dots$ denote realizations of zero-mean random noise "buffering" the uncertainty of mathematical modelling.

The problem of estimation of the parameter matrix \mathbf{P} is numerically ill-conditioned because of the collinearity in the data matrix \mathbf{Y} [12]. Three kinds of measures are applied to remediate this difficulty, viz.: selection of samples whose spectral data are used for modelling, selection of wavelength values defining the spectral data that are included in the vectors $\mathbf{y}_1, \mathbf{y}_2, \dots$, and so-called *soft modelling*. The latter consists in expressing the data in terms of a new, more informative, orthogonal variables, called *latent variables*. This is most frequently achieved by a mathematical conversion of the data to new axes based on covariance in the data \mathbf{Y} . The latent variables, being linear combinations of the independent spectral data, are created to describe independent sources of the observed variation in the the data. They are, as a rule, ordered after the size of the independent sources of variation that they explain – in order to facilitate the decision on the number of latent variables that are needed to adequately explain the systematic variation in the data \mathbf{Y} , and eliminate non-informative and less-informative latent variables. There are numerous references containing detailed reviews of the methodology and practice of soft modelling, both for experts and for beginners [13-16].

4.4. Identification of mathematical models of spectrophotometric data

Any procedure of spectrophotometric data interpretation refers to their more or less complex mathematical model. The identification of this model is the goal of ST calibration. There are two fundamentally different approaches of the problem of calibration that imply substantially different solutions for data interpretation, aimed at estimation of concentrations: the forward-model-based approach and the inverse-model-based approach. The forward-model-based approach is aimed at identification of a forward model of the ST, i.e. estimation of the parameters of an operator, mapping

the space of real-valued functions $x(\lambda; \mathbf{c})$ or the space of vectors \mathbf{c} into the space of vectors $\tilde{\mathbf{y}}$. The inverse-model-based approach is aimed at identification of an inverse model of the ST, *i.e.* at estimation of the parameters of an operator, mapping the space of vectors $\tilde{\mathbf{y}}$ into the space of real-valued functions $x(\lambda; \mathbf{c})$ or the space of vectors \mathbf{c} . In both cases a set of, reference data:

$$\tilde{\mathbf{D}}^{cal} = \left\{ \tilde{\mathbf{c}}_v^{cal}, \tilde{\mathbf{y}}_v^{cal} \mid v = 1, 2, \dots, N^{cal} \right\} \quad (27)$$

is necessary to attain the required validity of the model. It should contain concentrations $\tilde{\mathbf{c}}_v^{cal}$, characterizing samples used for calibration, as measured by a reference instrument, and the corresponding spectral data $\tilde{\mathbf{y}}_v^{cal}$ acquired by means of the calibrated ST.

5. FROM CALIBRATION TO INTERPRETATION OF SPECTROPHOTOMETRIC DATA

5.1. General considerations

The absorption spectrum of a sample $x(\lambda; \mathbf{c})$ is sensitive to the concentrations \mathbf{c} of the compounds that sample is composed of; it may be used, therefore, for measuring those concentrations. In most cases, the dimension of the vector \mathbf{c} is significantly smaller ($J = 1-25$) than the number of the raw measurement data in the vector $\tilde{\mathbf{y}}$ ($N = 100-1000$). As a rule, the data provided by STs are highly correlated from one wavelength value to the next and from one sample to another. This correlation decreases the apparent information in the data; but – if an approach is used that takes it into account – it is possible to turn the correlation to a benefit by taking advantage of the redundancy. Therefore, a methodology for solving those problems, as a rule, comprises two steps: the compression of the data $\tilde{\mathbf{y}}$, *i.e.* their transformation into an estimate $\hat{\mathbf{p}}_y$ of a low-dimensional vector of informative parameters \mathbf{p}_y :

$$\tilde{\mathbf{y}} \rightarrow \hat{\mathbf{p}}_y \quad (28)$$

and subsequent estimation of the concentrations on the basis of $\hat{\mathbf{p}}_y$:

$$\hat{\mathbf{p}}_y \rightarrow \hat{\mathbf{c}} \quad (29)$$

The complexity of both steps depends on the number of informative parameters: the greater this number, the larger is the time required both for ST calibration and for estimation of concentrations. The ST calibration, performed on the basis of the reference data $\tilde{\mathbf{D}}^{cal}$, is aimed at estimation of the parameters of the operator of data compression, defined by Eq.(28), and of an operator of measurand estimation, defined by Eq.(29), provided the structure of both operators is chosen in advance. A large variety of algorithms can be generated by combining various techniques of data compression with various techniques of estimation. The compression consists, as a rule, in the selection of the most informative among the parameters computed according to

various principles, such as: projection of the data $\tilde{\mathbf{y}}$ on a set of linearly independent (preferably orthogonal) vectors; parameterization of the data $\tilde{\mathbf{y}}$, *i.e.* their approximation by means of a known function $\hat{y}(\lambda_n; \mathbf{p}_y)$ with unknown parameters \mathbf{p}_y ; estimation of the moments of $\tilde{\mathbf{y}}$; soft modelling techniques, such as principal component analysis (PCA) or partial least squares (PLS). The most self-imposing tool for estimation of concentrations is a neural network being a universal approximator [17], [18], but – for the reason of efficiency – multidimensional B-splines and wavelet transforms are viable alternatives in many instances of calibration. For historical reasons, the algorithms for estimation of concentrations are most studied within the discipline called *chemometrics*.

Once a soft model, described in Subsection 4.3, has been developed, the estimation of the concentrations is a sequence of simple algebraic operations. The quality of the result of estimation has been already decided during ST calibration. All the efforts must be, therefore, concentrated on the measurement and computational operations related to the procedure of calibration. Both complexity and reliability of this procedure depends on the formulation of the measurement problem it is designed for; the following elements of its formulation should be distinguished:

- all or only selected compounds of a mixture are to be determined;
- all or only some pure standards are available.

If all pure standards are available, then it is possible to optimize calibration process by designing reference mixtures for calibration purposes.

5.2. Multivariate ST calibration

The univariate calibration is, by definition, related to modelling the dependence of the concentration of a single compound on a single spectral data point [19], while multivariate calibration might involve determining the concentrations using the spectral data acquired for tens or hundreds of wavelength values. A general scheme of multivariate calibration, applied in practice, comprises the following steps:

- selection of the reference samples (mixtures) to be used for identification of the model;
- acquisition and visual evaluation of the spectral data, used for calibration, before and after preprocessing;
- a first modelling trial to decide whether it is possible to attain the expected quality of the model and whether non-linearity should be introduced in this model;
- refinement of the model, *e.g.* by considering elimination of possible outliers with respect to the model, selecting the model complexity;
- final validation of the model.

The optimum complexity of the model, resulting from calibration, is a key issue. As the complexity of the model is increasing, the prediction error determined for the data set $\tilde{\mathbf{D}}^{cal}$ is diminishing. To determine the optimum number of significant components in a series of spectral data, one may look at how the error is decreasing when that number is increasing [20-23]. By increasing the complexity of the model, one may reduce this error almost to zero, but at the

same time make the model produce artefacts when it is applied to unknown samples. Validation is the step in which the prediction with the chosen model is tested independently. An additional set of data \tilde{D}^{val} , independent of \tilde{D}^{cal} but having the same structure, should be used for this purpose:

$$\tilde{D}^{val} = \left\{ \tilde{\mathbf{c}}_v^{val}, \tilde{\mathbf{y}}_v^{val} \mid v = 1, 2, \dots, N^{val} \right\} \quad (30)$$

Any calibration procedure includes the estimation of model parameters [24], [25]. For the time being, no single method for estimation is considered the best in all spectrophotometric applications. Sometimes, the most traditional tool, *i.e.* the ordinary least squares estimator is sufficient. Its main drawback is its sensitivity to collinear components in the spectral data. Therefore, several alternative methods have been developed to remediate for this drawback. The most commonly used among them is the partial least squares estimator. It has gained enormous popularity among the researchers, and – consequently – has been enhanced and modified in various ways, including the incorporation of some nonlinearity and constraints, such as the positivity of spectral data and the positivity of concentrations of compounds [26-40]. Another alternative method developed to remediate for ill-conditioning of the parameter estimation, is ridge regression [41], [42] and the generalized ridge regression, where each variable gets a slightly different ridge parameter [43].

It should be noted that the linear-regression approach may fail when considerably nonlinear relationships are modelled. Quite a number of linearization techniques have been proposed to deal with such situations: some of them work on the data and others work on the model or on some model parameters. In general, the artificial neural networks are natural tools for dealing with nonlinear models. They are used more and more frequently in spectrophotometric data processing, as a rule in combination with various data compression techniques, such as principal component analysis, to avoid excessive overfitting [44].

An important step, related to any method referring to the idea of latent variables, is the selection of the most informative variables or elimination of non-informative variables – *cf.* examples of techniques developed for this purpose [45-48].

5.3. Selection of samples and wavelength values

The multivariate calibration requires standards, *i.e.* samples for which the estimates of concentrations, obtained by a reference method, are known. Because of the cost or time of the reference technique, the number of such "standards" cannot be very large (usually not greater than 50 or 100). Because the model has to be used for the prediction on new samples, all possible sources of variation that can be encountered later must be included in the calibration data \tilde{D}^{cal} . This means that the chemical compounds, present in the samples to be analyzed, must be included in the samples used for calibration, and the range of variation in their concentrations should be at least as wide as that expected of the samples to be analyzed. There is, however, a practical limit on what is available. It is therefore necessary to

achieve a compromise between the number of samples to be analyzed and the prediction error that can be attained. When it is possible to artificially generate a number of samples, experimental design can and should be used to decide on the composition of the calibration samples. In most cases only real samples are available, so that an experimental design is not possible. This is the case for the analysis of food products and ingredients.

There are several strategies available for selection of the calibration samples representative of the analysis problem to be solved. The simplest is random selection, but it implies a risk that some sources of variation will be lost. Another possibility is based on *a priori* knowledge about the analysis problem: if all the sources of variation are known, then samples can be selected on the basis of that knowledge. A more detailed analysis of the problem of samples selection may be found in the literature [49], [50].

For each sample, the spectral data are recorded for many wavelength values. An important opportunity to improve the numerical conditioning of estimation problem is the proper selection of those values. It can be based on *a priori* knowledge of the most informative wavelength values, derived from previous experience, or may be performed by means of special techniques of selection such as stepwise ordinary least squares regression.

6. PREPROCESSING OF SPECTROPHOTOMETRIC DATA

6.1. Aims of spectrophotometric data pre-processing

The estimation of the concentrations, directly from the acquired spectrophotometric data, is in many practically important cases problematic due to instrumental imperfections of the ST and some quantum phenomena in the analyte that produce blurring in the data representative of the measured spectrum. Both effects may be eliminated or reduced by means of the algorithms of deconvolution or generalized deconvolution. As a rule, those algorithms are designed and assessed using the criteria of the quality of deconvolution, which are not specific of spectrophotometry, such as the root-mean-square errors of approximation. Consequently, the estimates of the concentrations, determined on the basis of the results produced by those algorithms, may turn out to be very poor. On the other hand, the variational algorithms, used for estimation of concentrations, may be very inefficient if not provided with a good initial guess of the sought-for estimates [51]. A significant improvement may be attained by taking into account that the positions of spectral peaks are carrying information on the compounds of the analyzed substance, and the magnitudes of peaks – on their concentrations. In a series of publications [52-54], a methodology for estimation of the concentrations, on the basis of spectrophotometric data, has been developed; it is comprising the compression of spectral data to the estimates of the positions and magnitudes of spectral peaks, and qualitative identification of the compounds on the basis of those estimates. Not only parameters of spectral peaks, but also some other morphological features of the spectrum of an analyte may be used for identification of its contents. In the paper by Benjathapanun *et al.* [55], for example, the concentrations

are estimated using a procedure based on an assumption that the list of possible compounds is known *a priori*: first, the spectral data are twice differentiated and partitioned into segments along the wavelength axis; next, the presence or absence of the compounds from the pre-defined list is checked on the basis of the binary codes of the segments, using a neural network NN1; finally, the concentrations are estimated on the basis of the data corresponding to the most informative parts of the spectrum, using a neural network NN2.

The main aims of spectrophotometric data preprocessing may be summarized as follows:

- the elimination or suppression of those features of the spectral data that are not related to the concentrations;
- the enhancement of those features of the spectral data that are carrying information on the concentrations;
- the improvement of the resolution.

6.2. Suppression of undesirable features of spectrophotometric data

The first goal is traditionally attained by reduction of noise present in the data, by elimination of the non-informative part of the spectrum called *baseline* or *background*, and by elimination of outliers. More recent are methods for elimination of systematic variations in the spectral data, being not correlated to the concentrations, regardless of the nature of those variations: orthogonal signal correction [56] and prewhitening of spectral data by their covariance-weighted filtering [57].

The noise in spectral data is most frequently reduced by means of moving-window averaging filters [58], in particular – by means of the Savitzky-Golay filter. Alternatively, it may be suppressed with the filters defined in the domain of the discrete Fourier transform or in the domain of wavelet transform [59], [60]. Another way of reducing noise is the repeated acquisition of the data and their averaging [61]; the signal-to-noise ratio (SNR) increases then with the square root of the number of data vectors. The tutorial information on the noise reduction in spectral data may be found in the series of internet papers authored by H. Mark and J. Workman [62], more advanced – in the Ph.D. thesis by D. Brown [63].

The baseline may be informally defined as a component of the data, slowly-changing with wavelength, whose subtraction from the data leaves spectral peaks placed on the zero level of intensity or absorbance. There are numerous heuristic methods for baseline subtraction, referring to this informal definition. For example: constant baseline differences can be eliminated by using offset correction (each vector of spectral data is corrected by subtracting either the absorbance value at a selected wavelength or the mean value in a selected wavelength range); the constant and linear components of the baseline may be removed by double differentiation of the data. Further examples may be found in the literature [64-67].

Outliers are extreme, very large or very small, measurement results. If they are retained in the data, one may falsely conclude that they do not follow a normal distribution. There are various ways for overcoming this problem, the simplest being to remove the outliers. But one can also use different criteria for fitting a model.

6.3. Enhancement of informative features of spectrophotometric data

The second aim of data preprocessing is attained mainly by numerical differentiation which is enhancing spectral differences. Both first and second derivatives are used, but the second derivative seems to be applied more frequently. This is not the case for the first derivative. An important drawback of the use of derivatives is the decrease of the SNR implied by the amplification of noise (for that reason, smoothing is needed before differentiation). The higher the degree of differentiation used, the higher the degradation of the SNR. Another drawback is that models, obtained as a result of calibration, based on spectral data preprocessed by differentiation, are sometimes less robust to instrumental changes, such as wavelength shifts which may occur over time. The tutorial information on the methods of data differentiation may be found in the references [68], [69], a review of recent developments in this area – in [70], [71], and examples of applications – in [72-75].

6.4. Improvement of spectral data resolution

If a method, used for estimation of the vector of concentrations \mathbf{c} , is ignoring the limitations of the optical resolution of the ST, then the preprocessing of the raw data, increasing their resolution, may reduce the uncertainty of the estimates $\hat{\mathbf{c}}$. This operation may be viewed as estimation of the spectrum $x(\lambda)$ on the basis of the data $\tilde{\mathbf{y}}$, called *spectrum reconstruction*. The model of spectrometric data, defined by Eq.(6), and – consequently – the methods of deconvolution, may be used for this purpose. In the majority of practically interesting cases, however, the problem of spectrum estimation is numerically ill-conditioned, *i.e.* oversensitive to the errors in the data. Consequently, the solution of Eq.(6) is, as a rule, buried in the noise "inherited" from the data with an amplification coefficient reaching the values of hundreds, thousands or millions. Thus, the problem of spectrum reconstruction, although very simple from the logical point of view, is not trivial and far from being definitively solved in practice. Numerous sophisticated algorithms for dealing with it have been developed since 1931 when P. H. van Cittert proposed a first iterative algorithm of deconvolution for improving resolution of spectrometric measurements [76]. None of them, nevertheless, has turned out to be the best in all applications under all possible measurement conditions. Consequently, a complex methodology for solving problems of spectrum reconstruction has been developed, and comparative studies of the methods have been continually carried out [77-81]. This methodology refers to the idea of a numerical mechanism, called regularization, that consists in such a modification of the initial problem that makes it less sensitive to the errors in the data. There are various general-purpose methods for regularization of the problems of spectrum reconstruction – all consist in constraining the set of admissible solutions. They may be classified into the following groups: direct methods, variational methods, probabilistic methods, iterative methods, parametric methods, and transformation-based methods [82]. As a rule, the methods of reconstruction, met in practice, refer to some combinations of these elementary methods of regularization.

The rapid increase in computing power, available for dealing with practical tasks of measurement and instrumentation, has radically changed views on the applicability of various mathematical methods and algorithms for solving problems of spectrum reconstruction. The speed and accuracy of computing – offered by general-purpose computers, digital signal processors and application-specific processors – make possible implementation of the algorithms which for decades have been considered too complex for practical applications. First of all, optimization-based approaches of reconstruction problems have become very common. They produce efficient variational algorithms of calibration or reconstruction – due to the strong regularizing properties, the possibility of combining all elementary mechanisms of regularization, and the convenience of incorporating additional constraints of the set of feasible solutions, deduced from *a priori* information on the solution $x(\lambda)$ and on the disturbances in the data \tilde{y} . A variational algorithm of spectrum reconstruction is a numerical implementation of the operation:

$$\hat{x}(\lambda) = \arg_{x(\lambda) \in \mathbb{X}} \text{opt} \left\{ \mathcal{J}[x(\lambda); \hat{\mathbf{p}}] \mid x(\lambda) \in \mathbb{X} \right\} \quad (31)$$

where \mathcal{J} is a criterion of optimization, and \mathbb{X} is a set of admissible solutions, defined on the basis of *a priori* information derived from physical or technical context of the measurement problem under consideration. A key issue in this case is the choice of the criterion \mathcal{J} in such a way as to mathematically express our answer to the question what a good spectrum reconstruction means...

7. EXAMPLE: SPECTROPHOTOMETRIC ANALYSIS OF WINE

Instruments, designed for measurement of ethanol concentration and sugar concentration in wine, were used by wine makers already in the first half of the XIXth century, and started to be their common tools after 1857 when Louis Pasteur explained the biochemical nature of fermentation. However, more complete analysis of wine had to wait *ca.* 100 years for the development of such modern analytical techniques as liquid chromatography (separation of mixture components) and mass spectrometry (elemental analysis). The time of spectrophotometry came by the end of the 1980's.

From chemical point of view, wine is a water solution of numerous (*ca.* 1000) organic and non-organic substances. The most important groups of them are quantified in Fig. 6. Even this incomplete list of wine components enables one to imagine in how many parameters (variables) wines may differ, depending on their geographical origin, sort of grapes and technology used for their fabrication. The task to be performed by a spectrophotometric wine analyzer (estimation of those parameters or recognition of a pattern they belong to) is extremely difficult because the spectra of very different wines may differ insignificantly. It may be successfully carried out only if very high requirements, concerning both metrological features of the ST and the in-built capacity to process the raw data, are satisfied. The

progress in this domain is, thus, conditioned by parallel advancements of micro- and optoelectronics, on the one hand, and methods for measurement data processing, on the other. The corresponding R&D works are carried out by numerous academic and industrial centres employing biochemists and enologists, as well as computer and measurements experts.

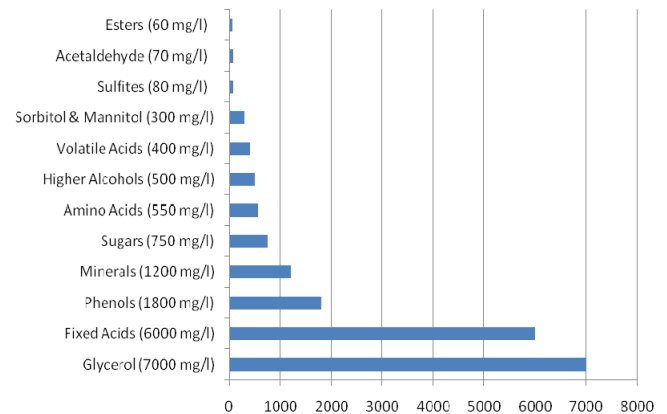


Fig. 6. The classification of main components of red wine [83].

As a rule, the industrial R&D works are oriented on prototyping complete wine analyzers, their testing and production. The Danish company *Foss* is an example of a manufacturer involved in this kind of activities. Ten years ago, it launched an instrument which is now distributed in three versions: *WineScanTM FT120 Basic* [84], *WineScanTM Auto* [85] and *WineScanTM Flex* [86]. It is based on NIR spectrophotometry (with an option to include Vis region of wavelength) and is able to determine: ethanol, glucose and fructose, malic acid, volatile acid, total acid and pH in finished wine or must under fermentation. Although it is rather large and heavy (*ca.* 70 kg), its functioning requires an external computer with a special software package for data processing. Similar instruments are manufactured by *Brimrose Corporation of America*, *Bruker Optics*, *NIR Technology Australia*, *Thermo Fisher Scientific* and *LT Industries*.

The research works carried out by universities and other non-industrial institutions are, as a rule, oriented on solving selected problems related to partial analysis of wine rather than on the design of a complete analyzer, *e.g.* [87-103]. A typical objective of such research works is to attain – by means of spectrophotometry – similar metrological outcomes as those attainable by means of expensive systems of liquid chromatography combined with mass spectrometry. The research works of this type are carried out, *i.a.*, by the following institutions:

- *Australian Wine Research Institute* (Adelaide) and *Cooperative Research Centre for Viticulture* (Glen Osmond) in Australia;
- *EMBRAPA Instrumentação Agropecuária* (São Carlos) in Brazil;
- *Institut National d'Optique* (Sainte-Foy, Québec) in Canada;
- *College of Biosystems Engineering and Food Science, Zhejiang University* (Hangzhou) in China;

- *Laboratoire Génie de l'Environnement Industriel, Ecole des Mines d'Alès* (Alès) and *Unité Mixte de Recherche Sciences pour l'Oenologie, INRA* (Montpellier) in France;
- *Laboratory of Food Chemistry, University of Ioannina* (Ioannina) in Greece;
- *Forschungsanstalt Geisenheim, Institut Oenologie und Getränke-technologie* (Geisenheim) in Germany;
- *Istituto Nazionale di Ricerca per gli Alimenti e la Nutrizione* (Rome) and *Istituto Sperimentale per L'Enologia* (Asti) in Italy;
- *Departamento de Fitotecnia, Universidade de Évora* (Évora) and *Instituto dos Vinhos do Douro e Porto* (Porto) in Portugal;
- *Kmetijski Institut Slovenije* (Ljubljana) and *Biotehniška Fakulteta, Univerza v Ljubljani* (Ljubljana) in Slovenia;
- *Departamento de Química Analítica, Nutrición y Bromatología, Universidad de Santiago* (Santiago de Compostela) and *Department of Chemistry, University of La Rioja* (La Rioja) in Spain;
- *Eidgenössische Versuchsanstalt für Obst-, Wein- und Gartenbau* (Wädenswil) in Switzerland;
- *Department of Viticulture and Enology, University of California* (Davis) in USA.

The activity of the countries, traditionally involved in wine making and distribution (France, Greece, Italy, Spain, Portugal) is evident. We have been also accustomed, during last 20 years, to the presence of Australian, Californian or Chilean wines on our tables. We may be, however, slightly surprised by the presence of China, a country which – after spectacular successes in some other domains of economy – is joining the league of mass producers of wine.

Although the objectives of research works, carried out in various regions of the world, are similar; their motivation may be, however, slightly different. In the countries, traditionally associated with wine culture, the improvement of laboratory-dedicated techniques of analysis seems to be the main driving force. In the countries where industrial wine production is aimed (Australia, Brazil, China), the motivation of research works is related to the demand for automated monitoring of wine production at all its stages. The common motive is the need to counteract the adulteration of wine by means of robots mentioned in Section 2. Their development may be based on two approaches:

- the translation of spectral data into the language of wine contents, followed by the translation of wine contents into the language of wine labels;
- the direct identification of wine on the basis of spectral data and appropriately compiled libraries.

The latter approach requires concerted efforts of interdisciplinary teams of scientists (biochemists, mathematicians, metrologists) and practitioners (engineers and someliers) to teach the robots the skills which are up to now reserved for human beings. Useless to say, that many wine experts, especially in Europe, look with scepticism at all the efforts aimed at automatic wine testing.

8. CONCLUSIONS

The research problems related to the spectrophotometric analysis and testing of food ingredients and products – such as olive oil, cheese, grains, bread, meat, milk or chocolate – are approached in a similar way as those related to wine analyzers and robots. The existing instruments are rather bulky (both in terms of external dimensions and mass), but their miniaturization seems to be imminent due to the rapid advancements of micro-technologies applicable in spectrophotometry. The number of commercially-available miniature and relatively cheap spectrophotometric transducers has been growing quickly. Their dimensions are in the range of millimetres and – consequently – their metrological performance is on the whole far below that of laboratory instruments, but this deficiency may be at least partially compensated by sophisticated algorithms of measurement data processing. It seems very probable that the on-going progress in the domain of microspectrophotometry and digital signal processing will lead, in the perspective of 5-10 years, to the widespread use of hand-held spectrophotometric analyzers in food manufacturing and distribution, next – in restaurants, bars and private kitchens. A wine analyzer of size comparable with that of a cellular phone and priced below EUR 300, would attract many buyers even today...

The overview of methods for spectrophotometric analysis of food, presented in this lecture, is demonstrating richness and diversity of approaches and techniques used for solving relevant problems of food analysis. Although the methods for solving those problems have been studied mathematically for a long time, only recently they have found mass applications. This is mostly due to the progress in computer technology which has provided the designers of the analyzers with miniaturized and computationally powerful means of data processing, such as application-specific integrated circuits or digital signal processors and digital signal controllers. But this is also due to the above-mentioned proliferation of microspectrometers whose metrological potential may be fully exploited provided appropriate advances in measurement data processing are accomplished. Hence the motivation for further development of the methods and techniques for measurement data processing, dedicated to food analysis, including calibration of spectrophotometric transducers, especially of microtransducers. The main trends involve generalization of the existing methods towards problems described by more sophisticated models of the object of measurement, spectral data and their relationship. Apart from the increased dimensionality, which is of crucial importance, the nonlinearity and nonstationarity of the models is considered. When looking to the future, one should also mention the methods of measurement data processing, which are already in use but whose applicability potential seems to be far being exhausted, *viz.*: methods of time-frequency analysis, including wavelet transforms and Bayesian methods as means for estimation of concentrations or parameters of spectrometric transducers.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Science and Higher Education, Poland (grant # N505 011 31/1428).

REFERENCES

- [1] D. R. Heldman: "IFT and the Food Science Profession." *Food Technol.*, October 2006. p. 11
- [2] "Introducing a Personal Robot with a Sense of Taste", *NEC System Technologies Ltd. News*, Aug. 1, 2006, <http://www.necst.co.jp/english/news/20060801/index.htm>.
- [3] <http://www.oceanoptics.com/Products/spectrometers.asp> [as of May 2008]
- [4] <http://www.controldevelopment.com/index.php?topgroupid=5> [as of May 2008]
- [5] <http://www.arcoptix.com/arcspectro-nir.htm> [as of May 2008]
- [6] <http://www.jobinyvon.co.uk/SiteResources/Data/Templates/Idivisional.asp?DocID=544&vIID=&lang=5> [as of May 2008]
- [7] http://microspectralsensors.com/Product_Details.asp?ProductID=R860 [as of May 2008]
- [8] <http://www.matchbox-spectrometers.co.uk> [as of May 2008]
- [9] <http://www.StellarNet-Inc.com> [as of May 2008]
- [10] R. Z. Morawski: "Spectrophotometric Applications of Digital Signal Processing", *Measurement Science and Technology*, Vol. 17, August 2006, pp. R117-R144.
- [11] M. Wiśniewski, R. Z. Morawski, and A. Barwicz: "Modeling the Spectrometric Microtransducer", *IEEE Trans. Instrum. Meas.*, June 1999, Vol. 48, No. 3, pp. 747-752.
- [12] T. Næs, and B.-H. Mevik "Understanding the collinearity problem in regression and discriminant analysis", *J. Chemometr.*, Vol. 15, No. 4, pp. 413-426.
- [13] A. DeJuan, E. Casassas, and R. Tauler: "Soft modeling of analytical data", in *Encyclopedia of Analytical Chemistry* (Ed. R. A. Myers), Vol. 11, J. Wiley & Sons, Chichester 2000, pp 9800-9837.
- [14] J. E. Jackson: *A User's Guide to Principal Components*, Wiley-Interscience, New York 1991.
- [15] B. K. Lavine: "Chemometrics: Fundamental Review", *Anal. Chem.*, Vol. 72, 2000, pp. 91R-98R.
- [16] H. Martens, and T. Næs: *Multivariate Calibration*, J. Wiley & Sons, Chichester 1989.
- [17] P. Cardaliaguet: "Approximation of a Function and its Derivative with a Neural Network", *Neural Networks*, Vol. 5, 1992, pp. 207-220.
- [18] S. Geva, and J. Sitte: "A Constructive Method for Multivariate Function Approximation by Multilayer Perceptrons", *IEEE Trans. Neural Networ.*, Vol. 3, No 4, 1992, pp. 621-624.
- [19] N. Francois, B. Govaerts, and B. Boulanger: "Optimal designs for inverse prediction in univariate nonlinear calibration models", *Chemometr. Intell. Lab.*, Vol. 74, 2004, pp. 283-292.
- [20] C. E. Alciaturi, M. E. Escobar, and I. Esteves: "The use of the autocorrelation function in modeling of multivariate data", *Anal. Chim. Acta*, Vol. 553, 2005, pp. 134-140.
- [21] N. Benoudjit, D. Francois, M. Meurens, and M. Verleysen: "Spectrophotometric Variable Selection by Mutual Information", *Chemometr. Intell. Lab.*, No. 74, 2004, pp. 243-251.
- [22] N. Benoudjit, E. Cools, M. Meurens, and M. Verleysen: "Chemometric Calibration of Infrared Spectrometers – Selection and Validation of Variables by Non-Linear Models", *Chemometr. Intell. Lab.*, No. 70, 2004, pp. 47-53.
- [23] M. C. Pontes, R. Kawakami Harrop Galvao, M. C. Ugulino Araujo, P. N. Teles Moreira, O. D. Pessoa Neto, G. E. Jose, T. C. Bezerra Saldanha: "The successive projections algorithm for spectral variable selection in classification problems", *Chemometr. Intell. Lab.*, Vol. 78, 2005, pp. 11-18.
- [24] J. Kalivas: "Interrelationships of multivariate regression methods using eigenvector basis sets", *J. Chemometr.*, Vol. 13, 1999, pp. 111-132.
- [25] P. Geladi, and E. Dabakk: "Computational methods of analysis and chemometrics in near-infrared spectroscopy", in: *Encyclopedia of Spectroscopy and Spectrometry* (Eds. J. Lindon, G. Holmes, and J. Holmes), Academic Press, London 1999, pp. 343-349.
- [26] R. Bro: *Multi-way Analysis in the Food Industry – Models, Algorithms, and Applications*, Ph.D. Thesis, Universiteit van Amsterdam, 1998.
- [27] M. A. Alonso-Lomillo, O. Dominguez-Renedo, and M. J. Arcos-Martinez: "Resolution of Binary Mixtures of Rifamycin SV and Rifampicin by UV/VIS Spectroscopy and Partial Least-Squares Method (PLS)", *Chem. Biodivers.*, Vol. 1, 2004, pp. 1336-1343.
- [28] J. A. Arancibia, A. Rullo, A. C. Olivieri, S. Di Nezio, M. Pistonesi, A. Lista, and B. S. Fernández Band: "Fast spectrophotometric determination of fluoride in ground waters by flow injection using partial least-squares calibration", *Anal. Chim. Acta*, Vol. 512, No. 1, June 2004, pp. 157-163.
- [29] T. Azzouz, A. Puigdomènech, M. Aragay, and R. Tauler: "Comparison between different data pre-treatment methods in the analysis of forage samples using near-infrared diffuse reflectance spectroscopy and partial least-squares multivariate calibration method", *Anal. Chim. Acta*, Vol. 484, No. 1, May 2003, pp. 121-134.
- [30] D. Cozzolino, M. J. Kwiatkowski, M. Parker, W. U. Cynkar, R. G. Damberg, M. Gishen, and M. J. Herderich: "Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy", *Anal. Chim. Acta*, Vol. 513, No. 1, June 2004, pp. 73-80.
- [31] Y. P. Du, Y. Z. Liang, J. H. Jiang, R. J. Berry, and Y. Ozaki: "Spectral regions selection to improve prediction ability of PLS models by changeable size moving window partial least squares and searching combination moving window partial least squares", *Anal. Chim. Acta*, Vol. 501, No. 2, Jan. 2004, pp. 183-191.
- [32] J. A. Fernandez-Pierna, L. Jin, F. Wahl, N. M. Faber, and D. L. Massart: "Estimation of partial least squares regression prediction uncertainty when the reference values carry a sizeable measurement error", *Chemometr. Intell. Lab.*, Vol. 65, 2003, pp. 281-291.
- [33] J. Ghasemi, N. Shahabadi, and H. R. Seraji: "Spectrophotometric simultaneous determination of cobalt, copper and nickel using nitroso-R-salt in alloys by partial least squares", *Anal. Chim. Acta*, Vol. 510, No. 1, May 2004, pp. 121-126.
- [34] I. Hiroaki, N. Toyonori, and T. Eiji: "Measurement of pesticide residues in food based on diffuse reflectance IR spectroscopy", *IEEE Trans. Instrum. Meas.*, Vol. 51, No. 5, Oct. 2002, pp. 886-890.
- [35] J. Huang, D. Brennan, L. Sattler, J. Alderman, B. Lane, and C. O'Mathuna: "A comparison of calibration methods based on calibration data size and robustness", *Chemometr. Intell. Lab.*, Vol. 62, No. 1, April 2002, pp. 25-35.
- [36] S. Kasemsumran, Y. P. Du, K. Murayama, M. Huehne, and Y. Ozaki: "Near-infrared spectroscopic determination of human serum albumin, gamma-globulin, and glucose in a control serum solution with searching combination moving

- window partial least squares", *Anal. Chim. Acta*, Vol. 512, No. 2, June 2004, pp. 223-230.
- [37] H. Khajehsharifi, M. F. Mousavi, J. Ghasemi, and M. Shamsipur: "Kinetic spectrophotometric method for simultaneous determination of selenium and tellurium using partial least squares calibration", *Anal. Chim. Acta*, Vol. 512, No. 2, June 2004, pp. 369-373.
- [38] B. Muik, B. Lendl, A. Molina-Díaz, L. Pérez-Villarejo, and M. J. Ayora-Cañada: "Determination of oil and water content in olive pomace using near infrared and Raman spectrometry. A comparative study", *Anal. Bioanal. Chem.*, Vol. 379, No. 1, May 2004, pp. 35-41.
- [39] C. Pizarro, I. Esteban-Díez, A.-J. Nistal, and J.-M. González-Sáiz: "Influence of data pre-processing on the quantitative determination of the ash content and lipids in roasted coffee by near infrared spectroscopy", *Anal. Chim. Acta*, Vol. 509, No. 2, May 2004, pp. 217-227.
- [40] Qi Shen, Jian-Hui Jiang, Guo-Li Shen, and Ru-Qin Yu: "Variable selection by an evolution algorithm using modified Cp based on MLR and PLS modeling: QSAR studies of carcinogenicity of aromatic amines", *Anal. Bioanal. Chem.*, Vol. 375, No. 2, Jan. 2003, pp. 248-254.
- [41] I. Frank, and J. Friedman: "A statistical view of some chemometric regression tools", *Technometrics*, Vol. 35, 1993, pp. 109-148.
- [42] R. Sundberg: "Continuum regression and ridge regression", *J. Roy. Stat. Soc. B-Met.*, Vol. 55, No. 3, 1993, pp. 653-659.
- [43] E. Vigneau, M. Devaux, M. Qannari, and P. Robert, "Principal component regression, ridge regression and ridge principal component regression in spectroscopy calibration", *J. Chemometr.*, Vol. 11, 1997, pp. 239-249.
- [44] Li-Xian Sun, A. M. Reddy, N. Matsuda, A. Takatsu, K. Kato, and T. Okada: "Simultaneous determination of methylene blue and new methylene blue by slab optical waveguide spectroscopy and artificial neural networks", *Anal. Chim. Acta*, Vol. 487, No. 1, July 2003, *Anal. Chim. Acta*, Vol. 487, No. 1, July 2003, pp. 109-116.
- [45] A. Höskuldsson: "Selection of subsets of variables in linear regression", *Homepage of Chemometrics*, Editorial Nov. 2003, <http://www.acc.umu.se/~tnkjtg/Chemometrics/Editorial>.
- [46] H. Swierenga, F. Wulfert, O. E. de Noord, A. P. de Weijer, A. K. Smilde, and L. M. C. Buydens: "Development of robust calibration models in near infra-red spectrometric applications", *Anal. Chim. Acta*, Vol. 411, 2000, pp. 121-135.
- [47] S. Wold: "PAC, PAT and variable selection", *Homepage of Chemometrics*, Editorial Jan. 2004, <http://www.acc.umu.se/~tnkjtg/Chemometrics/Editorial>.
- [48] Xueguang Shao, Fang Wang, Da Chen, and Qingde Su: "A method for near-infrared spectral calibration of complex plant samples with wavelet transform and elimination of uninformative variables", *Anal. Bioanal. Chem.*, Vol. 378, No. 5, March 2004, pp. 1382-1387.
- [49] L. Petersen, P. Minkinen, and K. H. Esbensen: "Representative sampling for reliable data analysis: Theory of Sampling", *Chemometr. Intell. Lab.*, Vol. 77, 2005, pp. 261-277.
- [50] S. Serneels, C. Croux, P. J. van Espen: "Influence properties of partial least squares regression", *Chemometr. Intell. Lab.*, Vol. 71, 2004, pp. 13-20.
- [51] A. Gilbert: "The Resolution of Bands in Spectroscopy", in: *Analytical Applications of Spectroscopy II* (Eds. Davies A. M. C. and Creaser C. S.), Royal Soc. Chem., Cambridge (UK), 1991, pp. 275-285.
- [52] R. Z. Morawski, A. Miękina, and A. Barwicz: "Combined Use of Tikhonov Deconvolution and Curve Fitting for Spectrogram Interpretation", *Instrum. Science & Technol.*, Vol. 24, No. 3, 1996, pp. 155-167.
- [53] A. Miękina, R. Z. Morawski, and A. Barwicz: "The Use of Deconvolution and Iterative Optimization for Spectrogram Interpretation", *IEEE Trans. Instrum. Meas.*, Vol. 46, No. 4, 1997, pp. 1049-1053.
- [54] A. Miękina, and R. Z. Morawski: "Industrial Application of Algorithms for Interpretation of Spectrometric Data", *Proc. 6th IMEKO Symp. "Metrology for Quality Control in Production"* (Vienna, Austria, Sept. 7-10, 1998), pp. 441-447.
- [55] N. Benjathapanun, W. J. O. Boyle, and K. T. V. Grattan: "Binary Encoded 2nd-differential Spectrometry Using UV-Vis Spectral Data and Neural Networks in the Estimation of Species Type and Concentration", *IEE Proc. Sci., Meas. Technol.*, Vol. 144, No. 2, 1997, pp. 73-80.
- [56] J. Trygg: "Everything you need to know about Orthogonal Signal Correction (OSC) filters – and how they can improve interpretation of your data", *Homepage of Chemometrics*, Editorial, March 2002, <http://www.acc.umu.se/~tnkjtg/Chemometrics/Editorial>.
- [57] M. Høy: *Building simple, reliable and relevant multivariate data-analysis tools*, Ph.D. Thesis, NTNU Trondheim, 2002.
- [58] S. Morton: "Lecture course covering all practical, and interpretive aspects of electron spectroscopy", <http://www.uksaf.org/tutorials.html>.
- [59] J. Trygg: "Wavelets in Chemometrics - compression, denoising and feature extraction", *Homepage of Chemometrics*, Editorial May 2002, <http://www.acc.umu.se/~tnkjtg/Chemometrics/Editorial>.
- [60] Xiuqi Zhang, Jiye Jin, Jianbin Zheng, and Hong Gao: "Genetic algorithms based on wavelet transform for resolving simulated overlapped spectra", *Anal. Bioanal. Chem.*, Vol. 377, No. 7-8, Dec. 2003, pp. 1153-1158.
- [61] S. Morton: "Lecture course covering all practical, and interpretive aspects of electron spectroscopy", <http://www.uksaf.org/tutorials.html>.
- [62] H. Mark, and J. Workman: "Chemometrics in Spectroscopy – Analysis of Noise – Parts I-XIV", *Spectroscopy on line*; Vol. 15-18, 2000-2003; www.spectroscopyonline.com.
- [63] D. Brown: *Rational approaches to data preprocessing in multivariate calibration*, Ph.D. thesis, Dalhousie University, Halifax 2000.
- [64] C. Bosch-Ojeda, and F. Sanchez-Rojas: "Recent developments in derivative ultraviolet/visible absorption spectrophotometry", *Anal. Chim. Acta*, Vol. 518, No. 1-2, Aug. 2004, pp. 1-24.
- [65] Da Chen, Xueguang Shao, Bin Hu, and Qingde Su: "A background and noise elimination method for quantitative calibration of near infrared spectra", *Anal. Chim. Acta*, Vol. 511, No. 1, May 2004, pp. 37-45.
- [66] V. Mazet, C. Carteret, D. Brie, J. Idier, and B. Humbert: "Background removal from spectra by designing and minimizing a non-quadratic cost function", *Chemometr. Intell. Lab.*, Vol. 76, 2005, pp. 121-133.
- [67] Qu Hai-bin, Ou Dan-lin, and Cheng Yi-yu: "Background correction in near-infrared spectra of plant extracts by orthogonal signal correction", *J. Zhejiang Univ. Sci.*, Vol. 6B, No. 8, 2005 pp. 838-843.
- [68] H. Mark, and J. Workman: "Chemometrics – Derivatives in Spectroscopy – Parts I-III", *Spectroscopy on line*, Vol. 18-19, 2003-2004, www.spectroscopyonline.com.
- [69] S. Morton: "Lecture course covering all practical, and interpretive aspects of electron spectroscopy", <http://www.uksaf.org/tutorials.html>.
- [70] C. Bosch-Ojeda, and F. Sanchez-Rojas: "Recent developments in derivative ultraviolet/visible absorption

- spectrophotometry", *Anal. Chim. Acta*, Vol. 518, No. 1-2, Aug. 2004, pp. 1-24.
- [71] D. Brown: *Rational approaches to data preprocessing in multivariate calibration*, Ph.D. thesis, Dalhousie University, Halifax 2000.
- [72] I. Badea, D. Moja, and L. Vladescu: "Determination of para-aminobenzoic acid, a degradation product of procaine hydrochloride, by zero-crossing first-derivative spectrometry", *Anal. Bioanal. Chem.*, Vol. 374, No. 1, Sep. 2002, pp. 51-53.
- [73] E. Ergun, B. Demirata, G. Gumus, and R. Apak: "Simultaneous determination of chlorophyll a and chlorophyll b by derivative spectrophotometry", *Anal. Bioanal. Chem.*, Vol. 379, No. 5-6, Jul. 2004, pp. 803-811.
- [74] S. S. Kharintsev, M. Kh. Salakhov: "A simple method to extract spectral parameters using fractional derivative spectrometry", *Spectroch. Acta Part A*, Vol. 60, 2004, pp. 2125-2133.
- [75] Z. Moldovan, and L. Alexandrescu: "Derivative spectrophotometry in the determination of phenyl- \hat{I}^2 -naphthylamine used as an antioxidant in rubber mixtures", *Anal. Bioanal. Chem.*, Vol. 374, No. 1, Sep. 2002, pp. 46-50.
- [76] P. H. van Cittert: "Zum Einfluss der Spaltbreite auf die Intensitätsverteilung in Spectrallinien II", *Z. Phys.*, Vol. 69, 1931, pp. 298-311.
- [77] P. A. Jansson: *Deconvolution of Images and Spectra*, Academic Press, London 1997.
- [78] P. B. Crilly: "A Quantitative Evaluation of Various Iterative Deconvolution Algorithms", *IEEE T. Instrum. Meas.*, Vol. 40, No. 3, 1991, pp. 558-562.
- [79] M. S. Milewski, R. Z. Morawski, and L. Szczeciński: "Comparative Study of Some Iterative and Variational Algorithms of Measurand Reconstruction", *Proc. Int. IMEKO-TC1&TC7 Coll. (London, UK, Sept. 8-10, 1993)*, pp. 305-310.
- [80] A. Miękina, R. Z. Morawski, and A. Podgórski: "Dynamic Reconstruction of Measurands and Calibration of Measuring System", *Measurement - J. IMEKO*, Sept. 1994, Vol. 14, No. 1, pp. 63-72.
- [81] R. Z. Morawski, L. Szczeciński, and A. Barwicz: "Deconvolution Algorithms for Instrumental Applications: A Comparative Study", *J. Chemometr.*, Vol. 9, 1995, pp. 3-20.
- [82] R. Z. Morawski: "Unified Approach to Measurand Reconstruction". *IEEE Trans. Instrum. & Meas.*, April 1994, Vol. 43, No 2, pp. 226-231.
- [83] <http://waterhouse.ucdavis.edu/winecomp/comp2.htm>
- [84] *WineScan™ Basic*, Foss datasheet P/N 580231, Issue 6 GB, Oct. 2006, <http://www.foss.us/Solutions/DownloadProductDataSheets.aspx> [as of Jan. 2008]
- [85] *WineScan™ Auto*, Foss datasheet P/N 1025826, Issue 1 GB, Oct. 2005, <http://www.foss.us/Solutions/DownloadProductDataSheets.aspx> [as of Jan. 2008]
- [86] *WineScan™ Flex*, Foss datasheet P/N 1025825, Issue 1 GB, Oct. 2005, <http://www.foss.us/Solutions/DownloadProductDataSheets.aspx> [as of Jan. 2008]
- [87] "Application of spectroscopy for non-destructive 'in-bottle' measurement of wine composition and quality", *CRCV Technology Application Note – April 2007*, www.crcv.com.au/viticare/vitinotes/Viti-Notes/.
- [88] "Near infrared spectroscopy in the Australian grape and wine industry", *CRCV Technology Application Note – April 2007*, www.crcv.com.au/viticare/vitinotes/Viti-Notes/.
- [89] A. N. Araujo, J. L. F. C. Lima, A. O. S. S. Rangel, and M. A. Segundo: "Sequential injection system for the spectrophotometric determination of reducing sugars in wines, *Talanta*, Vol. 52, 2000, pp. 59-66.
- [90] C. J. Bevin, R. G. Damberg, A. J. Fergusson, and D. Cozzolino: "Varietal discrimination of Australian wines by means of mid-infrared spectroscopy and multivariate analysis", *Anal. Chim. Acta*, 2007, doi:10.1016/j.aca.2007.10.042.
- [91] D. Cozzolino, M. J. Kwiatkowski, M. Parker, W. U. Cynkar, R. G. Damberg, M. Gishen, and M. J. Herderich: "Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy", *Anal. Chim. Acta*, Vol. 513, No. 1, June 2004, pp. 73-80.
- [92] D. Cozzolino, M. Parker, R. G. Damberg, M. Herderich, and M. Gishen: "Chemometrics and Visible-Near Infrared Spectroscopic Monitoring of Red Wine Fermentation in a Pilot Scale", *Biotechnol. & Bioeng.*, Vol. 95, No. 6, 2006, pp. 1101-1107.
- [93] D. Cozzolino: "Non-Destructive Analysis by Vis-NIR Spectroscopy of Fluid(s) in its Original Container", *PCT Patent Application #WO2007006099*, published on Jan. 18, 2007.
- [94] Fei Liu, Yong He, Li Wang, and Hongming Pan: "Feasibility of the use of visible and near infrared spectroscopy to assess soluble solids content and pH of rice wines", www.aseanfood.info/scripts/count_article.asp?Article_code=11020326 [as of Dec. 2007].
- [95] A. P. Ferreira, T. P. Alves, and J. C. Menezes: "Monitoring Complex Media Fermentations with Near-Infrared Spectroscopy – Comparison of Different Variable Selection Methods", *Biotechnol. & Bioeng.*, Vol. 91, No. 4, 2005, pp. 474-481.
- [96] J. Herrera, A. Guesalaga, and E. Agosin: "Shortwave-near infrared spectroscopy for non-destructive determination of maturity of wine grapes", *Meas. Sci. & Technol.*, Vol. 14, 2003, pp. 689-697.
- [97] L. Liu, D. Cozzolino, W. U. Cynkar, R. G. Damberg, L. Janik, B. K. O'Neill, C. B. Colby, and M. Gishen: "Preliminary study on the application of visible-near-infrared spectroscopy and chemometrics to classify Riesling wines from different countries", *Food Chemistry*, Vol. 106, 2008, pp. 781-786.
- [98] J. L. Moreira, and L. Santos: "Spectroscopic interferences in Fourier transform infrared wine analysis", *Analytica Chimica Acta*, 2004, Vol. 513, No. 1, pp. 263-268.
- [99] K. A. Riganakos, P. G. Veltsistas: "Comparative spectrophotometric determination of the total iron content in various white and red Greek wines", *Food Chemistry*, Vol. 82, 2003, pp. 637-643.
- [100] M. J. Saiz-Abajo, J. M. Gonzalez-Saiz, and C. Pizarro: "Prediction of organic acids and other quality parameters of wine vinegar by near-infrared spectroscopy – A feasibility study", *Food Chemistry*, Vol. 99, 2006, pp. 615-621.
- [101] P. Serapinas, P. R. Venskutonis, V. Aninkevičius, Z. Ežerinskis, A. Galdikas, and V. Juzikiene: "Step by step approach to multi-element data analysis in testing the provenance of wines", *Food Chemistry*, Vol. 107, 2008, pp. 1652-1660.
- [102] K. Skogerson, M. Downey, M. Mazza, and R. Boulton: "Rapid Determination of Phenolic Components in Red Wines from UV-Visible Spectra and the Method of Partial Least Squares", *Amer. J. Enol. & Viticult.*, Vol. 58, No. 3, 2007, pp. 318-325.
- [103] A. Soriano, P. M. Perez-Juan, A. Vicario, J. M. Gonzalez, and M. S. Perez-Coello: "Determination of anthocyanins in red wine using a newly developed method based on Fourier transform infrared spectroscopy", *Food Chemistry*, Vol. 104, 2007, pp. 1295-1303.