JOURNAL OF CHEMOMETRICS

Received: 24 October 2011, Revised: 7 March 2012, Accepted: 17 March 2012, Published online in Wiley Online Library: 12 April 2012

(wileyonlinelibrary.com) DOI: 10.1002/cem.2445

Process analytical technology: a critical view of the chemometricians

Alexey L. Pomerantsev[a,b] and Oxana Ye. Rodionova[a]*

The role of chemometrics in process analytical technology (PAT) solutions development is presented in the review on the basis of publications from 1993 to 2011. Main areas of application, stages of implementation, instruments, and chemometric methods used for the PAT implementations are reviewed. Generally speaking, PAT is considered to be an approach applicable not only in pharmaceutical industry but also in any production area such as food industry and biotechnology. PAT is claimed to be a new flexible manufacturing concept that accounts for variability and adapts the process to fit it. Copyright © 2012 John Wiley & Sons, Ltd.

Keywords: PAT; pharmaceuticals; biotechnology; food industry; feasibility; QbD; PLS; PCA; SIMCA; MIA; MSPC; filtering; MCR; TOS; N-way; DoE

1. INTRODUCTION

The term *process analytical technology (PAT)* has become widely popular (in a rather small community) since 2004 when the Food and Drug Administration (FDA) published the *Guidance for Industry* [1]. We presume that the meaning of PAT is well known to the readers of this review; therefore, only a short quotation from the aforementioned guidance [1] is provided. "PAT is a system for designing, analyzing, and controlling manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes, with the goal of ensuring final product quality."

The history of this document could be found in a review [2] that gives a detailed account of the extensional and intentional aspects of the PAT phenomenon. The scope of the current paper is much more specific as it illuminates just a single issue being chemometrics in PAT. The word *chemometrics* is never mentioned in [1]. Instead, the term *multivariate mathematical approaches* is used. These approaches are declared to include "the design of experiments, response surface methodologies, process simulation, and pattern recognition tools". However, it is clear now that chemometrics constitutes the essence of the matter. We can definitely state that such PAT was *de facto* used by many chemometricians, who, for more than 40 years, have been applying PAT without knowing it [3].

The PAT approach could be viewed in both narrow and wider scopes. In the first case, PAT may be applied to pharmaceutical industry only, and its methods are limited by a rather scanty list of the aforementioned PAT tools. In this case, PAT is an approach that has been exhaustively defined by the Food and Drug Administration functionaries, and therefore, it may not be interpreted in a broader sense. The wider view is not restricted to drug manufacturing alone, and so PAT can be used in any production area, that is, in biotechnology or food industry. PAT instruments, methods, and tools are being developed endlessly, and their list cannot be limited in principle. According to this interpretation, PAT is an approach that is continuously developed through collective input of researchers who are contributing by means of numerous publications. This can be seen in line with the old definition [4] that PAT is what *PATtioners* do.

Needless to say, this review represents the wider scope of PAT. Therefore, we consider that the first paper [6] on PAT can be traced back to 1984. The numerous publications on multivariate statistical process control (MSPC) [6–10] followed shortly after. There is no doubt that MSPC is a PAT tool, and the *Guidance for Industry* [1] confirms this. It is even more striking that only 5% of papers that describe tools used in PAT apply MSPC. Process analytical chemistry (PAC) [11–22] is a direct forerunner of PAT. Sometimes, it is difficult to distinguish between the PAC and PAT approaches as they both aim at manufacturing control and quality improvement. One could say that PAT inclines to pharmaceutical manufacturing, whereas PAC is not restricted to any production area in particular. We suppose that the method of the process analysis can be seen as the discriminating rule: off-line and at-line monitoring is for PAC, and strictly on-line and in-line monitoring is for PAT [23]. This, however, could be disputed – as late as 1984, some in-line methods has been proposed [11] and applied [5].

Working on this review, we have analyzed about 690 relevant papers, which have been published between 1993 and 2011, with some exclusion. Fishing out the PAT papers was a difficult task, as many authors claim having conducted a PAT research but merely present roundabout reflections. On the other hand, there are plenty of studies, for example [24–26], which represent good PAT research but never mention this term. Among the 690 papers, directly or indirectly related to PAT, only 30% employ chemometrics. These 245 chemometrics-related PAT (CRPAT) papers constitute a representative set that is analyzed further in this review. Figure 1 demonstrates the distribution of the

* *Correspondence to: Oxana Ye. Rodionova, Semenov Institute of Chemical Physics RAS, Kosygin Str. 4, 119991 Moscow, Russia.*
E-mail: rcs@chph.ras.ru

a *A. L. Pomerantsev, O. Y. Rodionova*
Semenov Institute of Chemical Physics RAS, Kosygin Str. 4 119991 Moscow, Russia

b *A. L. Pomerantsev*
State South Research and Testing Site RAS, Teatral'naya 8a 354000 Sochi, Russia

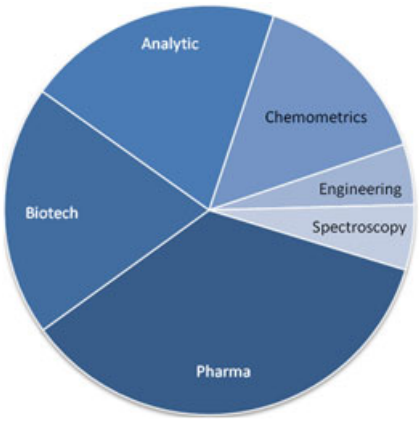


Figure 1. Distribution of the chemometrics-related process analytical technology papers among the journals grouped in six categories.

papers among 63 journals, which are grouped in six categories. About 100 of CRPAT papers are cited in this review.

The review is organized as follows. Section 2 presents three main production areas and the PAT solutions utilized in these areas. Section 3 is mainly devoted to indispensable stages of PAT solution implementation from the chemometric, not procedural, point of view. In Section 4, we provide a brief overview of instruments that are fit for in-line and on-line monitoring and are used in PAT solutions. Section 5 is the most important one as it presents the main mathematical/chemometric methods applied or sometimes specially designed for PAT solutions.

2. OBJECTS

Traditionally, PAT is considered as a specific pharmaceutical tool in line with the PAT Guidance document. The latter claims that the purpose of PAT is intended to support innovation and efficiency in pharmaceutical development, manufacturing, and quality assurance [1]. We prefer to see a bigger role for PAT solutions, as there are numerous applications of it in biotechnology, food industry, etc. In fact, from a chemometric point of view, it is impossible to separate, for example, the analysis of manufacturing of yeasts used in drug production and yeasts used for feed production. Therefore, we consider PAT as a universal methodology regardless of the area it is used in. The distribution among various industries that (currently) employ PAT is presented in Figure 2.

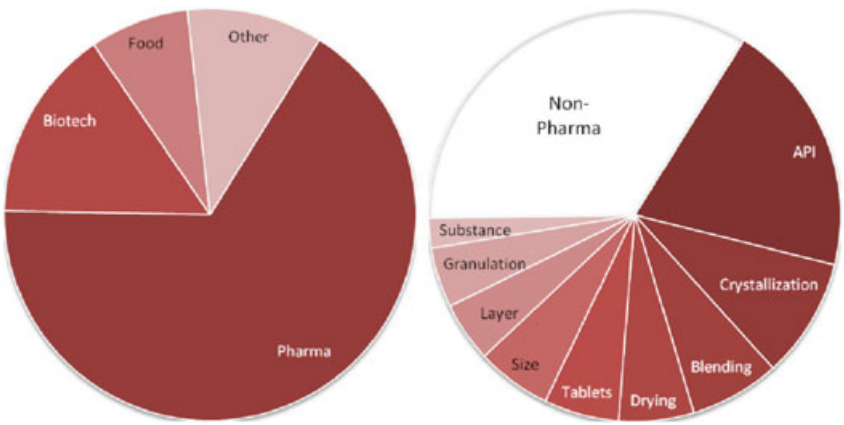
2.1. Pharmaceutical industry

The majority of PAT applications (about 70%) are developed for pharmaceutical industry mainly because of the requirements of the regulatory bodies in the USA and the EU. About 25% of the solutions focus on the identification and quantification of the active pharmaceutical ingredient (API) in the course of manufacturing, for example, during granulation [27], as well as in the finished products [28,29]. The necessity to assure the intended product quality during a production process encourages the implementation of on-line and in-line control of critical manufacturing stages (Figure 2b). A number of applications investigate model development for prediction of the mixing [30] and blending [31,32] homogeneity under different process conditions. The purity of the end product and its physical, chemical, and biological behaviors are often affected by the crystallization stage, which is one of the important manufacturing units [33–36]. The implementation of the general PAT principles during drug manufacturing at a small hospital pharmacy is described in [37]. An on-line process control was performed for the identification of raw material, blend uniformity analysis, and final content uniformity of capsules.

2.2. Biotechnology

Bioprocesses for the API production by fermentation (e.g., antibiotics) belong to one of the segments where PAT solutions help in monitoring, control, and significant enhancements of process understanding. In [38], authors show that combining different types of multistage process information provides the possibility to predict the final titer of a nominal antibiotic fermentation. The potential impact is in the improvement of the process analysis, diagnostics, as well as redesigning both monitoring and sampling schemes.

The PAT approach is successfully used to monitor and understand biotechnical processes, which are far afield from the pharmaceutical industry, for example, fermentation processes in biogas plants. Study [39] focuses on a mesoscale biogas plant unlike most studies devoted to the investigations of



a. Areas using PAT methodology b. PAT in Pharmaceutical Industry

Figure 2. (a) Areas using process analytical technology methodology, (b) process analytical technology in the pharmaceutical industry.

laboratory-scale or minor pilot plants only. A robust multivariate calibration/prediction model for the anaerobic digestion processes is established. This enables designing reliable low-cost PAT monitoring systems in biogas production.

2.3. Food and feed

In the food industry, PAT approach fully incorporates methods of improvement of the end-product quality on the basis of the investigation of raw material properties and process variable. Monitoring of the fed-batch cultivation of filamentous fungi used both in pharmaceutical and food industry, by means of in-line multiwavelength fluorescence measurements and chemometric analysis, is presented in [40].

A number of studies were conducted before 2004 when PAT initiative was not claimed. The general approach is to consider an industrial experimental problem starting with the definition of the production stages, utilization of an appropriate experimental design, and employment of measurements and parameters that reflect the desired phenomenon. A study such as this, together with a real-world example of low-fat mayonnaise production as a four-stage process, is presented in [19]. Fluorescence spectroscopy coupled with chemometric techniques have been identified as ideal PAT tools for process monitoring and quality control of sugar production [41].

3. STAGES

In this section, we present the essential stages of the PAT design and implementation. In this context, the term *stage* relates to PAT development rather than to manufacturing steps. However, the former partly overlaps with the production stages. The analysis of CRPAT papers enables us to suggest the following sequence of stages: feasibility, quality by design (QbD), and before, in, and after the process actions. Figure 3 shows the distribution of the CRPAT papers among the stages. This split is of course ambiguous, as a paper can relate to many stages, but even a not-perfect classification helps us to disclose the tendencies and trends in the current PAT research.

Half of all papers are devoted to feasibility studies. The reasons for this are clear; the laboratory studies are easier to conduct, and the results of such studies are usually published. As to manufacturers, they do not always welcome publication of the results to avoid the disclosure of technical information. We suppose that many more interesting PAT implementations exist, but we either do not know about them or only heard little something from the fellow chemometricians.

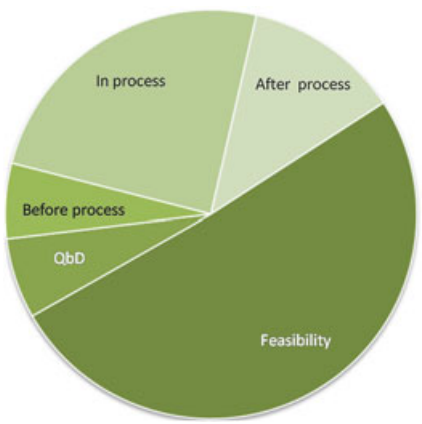


Figure 3. Process analytical technology stages studied in the chemometrics-related process analytical technology papers.

3.1. Feasibility studies

Before a PAT solution is suggested to a manufacturer, it is important to test it in the laboratory to assure that the designed approach works well enough. Often, there is another goal: to obtain lab-scale results that help convince the manufacturer of the necessity of the proposed solution. These preliminary studies investigate three different issues.

The first set of studies evaluates the properties' quantification potential or monitors specific operations. The goal of study [42] is the extraction of the meaningful information in real-time near infrared (NIR) measurements of coagulation milk. In [43], two techniques, the NIR spectroscopy and acoustic chemometrics, were investigated as means to monitor the maize silage spiked biogas process. A potential ability of the microwave resonance technology sensor to predict the final granule size was studied in [44]. The feasibility step can be considered as the "simplest" as deviations in material properties and experimental conditions are ordinarily supervised by an analyst.

Papers of the second type analyze the impact of variations in the real manufacturing conditions, such as temperature, pressure, and flow turbulence, as well as variations in the input material. Paper [45] analyzes and models the effect of temperature variation on NIR spectra and nuclear magnetic resonance (NMR) relaxation data. In [46], a special method that can efficiently model the external nonlinear effects is presented. Investigation of the effects of scatter on quantitative analysis of the chemical composition is presented in [28]. An algorithm that addresses the effect of multiplicative light scattering is proposed in [47]. Study [48] demonstrates how multiple external factors can be removed from the spectral data by orthogonalization.

Feasibility papers of the third type are devoted to the development of the reliable calibration models. In [49], a systematic prediction error correction method is proposed. This method aims at maintaining the predictive abilities of calibration models in the cases where spectrometer or measurement conditions are altered. In [10], a two-stage partial least squares (PLS) methodology for the monitoring of processes that are known to be affected by sources of variation is presented. These variations are an inherent part of routine operations. A new methodology for construction of calibration sets is developed in [50]. It is based on a selection of laboratory samples encompassing the same variability sources as the production samples.

It is hard to say how many feasibility studies have been, to any degree, implemented in full-scale plants, but this laboratory stage is, of course, indispensable.

3.2. Quality by design

One of the basic ideas of the PAT Guidance [1] is that quality of the end-product "should be built-in or should be design" and that PAT procedures "would be consistent with the basic tenet of quality by design". At the same time, the studies on the QbD procedures are emerging as specific area with its own vocabulary represented by a set of abbreviations, such as QTPP (quality target product profile), CPP (critical process parameter), DS (design space), and CQA (critical quality attribute). More and more papers use QbD along

with PAT in their titles or keywords. This could attract additional attention to publications, but from our point of view, this is nothing but a wordplay. As far as QbD is concerned, this is of course an inherent part of the design and development of well-understood processes. Usually, the QbD tools, for example, design of experiments (DoE) and response surface analysis, help develop DS and reveal the sources of variability of the process. Afterward, a PAT tool can be used to control the process in the established DS.

Paper [51] uses the QbD approach to study calibration design that had large variations in drug concentration (15%–85% w/w) and compaction pressure (100–500 MPa). These properties are used as process variables in the development of a pharmaceutical formulation. Such a large variation in both chemical and physical variables does not allow building a PLS model with good predictive ability; therefore, the three-way calibration method was employed.

In [52], the QbD approach is applied both to end products and process development. The authors consider a manufacturing process that involves five unit operations: high-shear wet granulation, milling, blending, compression, and coating. DoE and response surface analysis were used to evaluate the influence of three designed factors (water amount, wet massing time, and lubrication time) on response variables and to establish design space. In addition, many other variables, including both process variables and quality attributes, were considered across all unit operations. The application of multivariate analysis (principal component analysis (PCA) and PLS) helps predict the impact of material properties and process variables on the intermediate and final product quality attributes. As it is fairly noted in [52], QbD and multivariate data analysis are complementary tools. The level of understanding would not be achieved with either approach alone.

3.3. Before the process

Routine control of the raw material properties is an essential part of PAT. In pharmaceutical industry, such a procedure is mainly used for confirmation of a satisfactory quality of each lot of the input materials. In [53], a procedure for the routine testing of pharmaceutical substances directly in warehouses is presented. NIR measurements are conducted by remote fiber optic probe through the closed polyethylene (PE) bags. Direct PCA approach to classification failed because about 25% of the drums with good substance were classified as outliers. This was a great disappointment both for process engineers and instrument supplies. To make the whole procedure reliable, a special trichotomy classification procedure, which takes into account possible spectra distortion by PE bag folders, was designed.

Paper [54] investigates the application of fiber-optic probes with NIR spectrometers for the measurements of powder samples in process streams. It is shown that the magnitude of the spectral residuals depends on both the sample speed and the sample area presented to the probe.

In the food industry, variations in material properties often originate from the biological nature of the raw stock. These unwanted variations influence the end product. Sorting raw stock into homogeneous groups helps stabilize the end-product quality. Three such methods are presented in [20] with application to baking.

Sometimes, it is necessary to measure specific properties of input materials to tune the process parameters. In biotechnology, the at-line determination of input material properties is presented in [21]. NIR measurements are applied to rapidly determine moisture content in the forest biomass that was used for energy production.

3.4. In the process

Implementation of PAT solutions designed in a laboratory (or even tested in a small-scale plant) faces various difficulties in the course of transferring these solutions to full-scale manufacturing. Obvious challenges are posted by the nonideal conditions compared with those in a laboratory, limited possibilities of varying manufacturing conditions, unfeasibility to test PAT solutions at extreme fault conditions, etc. An active role of product engineers is also of great importance. As a rule, manufacturing staff is extremely busy solving daily problems, for example, spoilage abatement. Under such pressure, they have no time to analyze historical process data (already recorded and bulked) and to implement prospective PAT solutions. In connection to this, it is worth referring to paper [55], which describes the whole chain from feasibility study to real-time on-line application. From this paper, we can conclude that company engineers greatly help in project implementation. The aim of the solution is to develop an on-line process monitoring and control system for the automatic dosing of ammonia for production of low methoxylated amidated by using on-line optical-fiber NIR transmittance spectroscopy of the amidation liquid. The authors describe all challenges related to transferring the laboratory NIR measurements and synthetic calibration model to the full-scale plant. In particular, the paper discusses the adjustments to nonideal conditions of real-time manufacturing. The implementation was very successful and resulted in multiple benefits of the improved control of the ammonia dosing system.

3.5. After the process

Application of the PAT solutions with real-time release testing can replace the end-product testing, for example, in-line monitoring of the pellet coating thickness can replace the time-consuming final drug release test [56]. At the same time, this does not replace the quality control step. The aim of the latter is not to investigate the end product properties in details but to confirm that all batches are of the same satisfactory quality. This may be implemented as a solution for a classification problem of attributing units of the specific end product belonging to the class/classes of product with satisfactory quality, as it is performed in [57,58]. The absence of a pertinent quality control step at one of the Russian pharmaceutical factories led to a fatal accident in February 2009. A batch of ampoules for anesthesia, Lysthenon, was wrongly marked as mildronate, a remedy for heart and blood vessel diseases. Two persons died after injections because of respiratory standstill [59]. At the same time, NIR measurements of closed ampoules, as presented in [60], provide rapid on-line control and could prevent such accidents.

Another approach is to predict key end-product properties such as API content [61,62].

There is one more important issue that should be taken into account. Many circumstances can decrease model predictability or even make it invalid. There could be numerous reasons for this, for example, deterioration of manufacturing equipment, aging of sources, probes, and detectors in measuring instruments, and environmental influences, such as temperature and humidity variation. Thus, the essential part of each PAT solution is continuous model validation. Methods that provide multivariate

calibration transfer are often applied to fulfill this purpose. A detailed review is presented in [63]. In paper [55], described in Section 2.4, the authors also mentioned this stage in their PAT solution. Despite the fact that the predictions are regarded with high confidence at all levels of the organization, process operators still acquire samples for titrations. This is performed to both validate the NIR predictions and maintain the ability to continue production in the event of a breakdown of the ammonia dosing system.

4. INSTRUMENTS

Instruments applied in PAT solutions should be able to reflect the state of the process in a real-time scale. Various process analyzers are incorporated as on-line and in-line process monitoring devices. The distribution among different instrumentation methods applied in PAT is shown in Figure 4. The majority of analyzers are spectroscopic instruments, which are built into manufacturing process by means of flow cells, quartz windows, and immersion probes. An overview of spectroscopic analyzers is presented by E. Skibsted and S. B. Engelsen [23]. The authors picturesquely explain the differences between a traditional industrial process control and the advantages of remote sensing. General requirements for spectroscopy analyzers and various spectroscopic techniques, such as Fluorescence, Visual, NIR, IR, Raman, and H^1NMR, are reported with their areas of application as well as the pros and cons in implementation.

4.1. Near infrared

Near infrared instruments rank first among PAT analyzers. NIR spectra can be recorded without sample preparation, and this provides possibilities to conduct on-line and in-line measurements. The ability to penetrate inside of a sample provides NIR spectra monitoring and calibration with both chemical and physical properties, such as viscosity [64] or crystallinity [36]. A fiber-optic cable for an NIR instrument can be up to 100 m long and enables remote spectra measurements in hazardous or super-sterile environment. In industrial applications, spectra acquisition and observed samples are far from the well-controlled laboratory conditions. Samples may be highly heterogeneous as in the case of moisture determination of forest biomass [21], or samples/substance may be packed in closed PE bags [53]. Various external conditions, and first of all temperature fluctuations, should be taken into account [65].

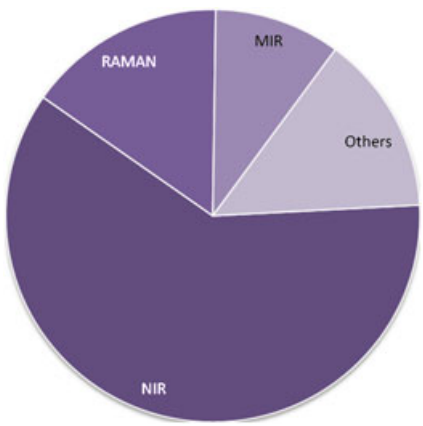


Figure 4. Instruments and analyzers used in process analytical technology applications.

4.2. Raman spectroscopy

There is an increasing interest in using Raman spectroscopy. Similarly to NIR, Raman spectroscopy can be used without sample preparation and implemented directly on the production line by using optical fibers with probes or quartz windows. Paper [66] presents the first Raman application to API in tablets. In paper [67], it was shown that Raman spectroscopy used by means of fiber-optical immersion probe is suitable for the in-line real-time monitoring of the blending process. Feasibility studies devoted to identification of polymorphic forms and phase changes monitoring are presented in [68,36]. The employment of Raman spectroscopy for quantitative assessment of chemical content is described in several publications. Powder samples are analyzed in [29], whereas API in intact tablets is predicted in [61]. The possibility for elucidating chemical reaction information from the Raman data arrays is shown in [69,70]. High selectivity to inorganic substances and polymorphs and low sensitivity to water make Raman spectroscopy a promising instrument for in-line and on-line process monitoring.

4.3. Middle infrared

Middle infrared spectroscopy is widely used in analytical practice both for scientific studies and routine analysis in the laboratories. One of the main obstacles for industrial application is strong absorption of water in 2500–4000 nm range. This can be surmounted by application of the attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR) technique. The most widely used ATR-FTIR application in PAT is the in-line monitoring of crystallization, which is an important purification unit operation in manufacturing of pharmaceuticals. In paper [34], the ATR immersion probe was employed to monitor the solute concentration throughout the crystallization process of sulfathiazole from different mixtures. In-line monitoring of the batch cooling crystallization of two monosodium glutamate and L-glutamic acid is described in [33]. Another obstacle pointed out in paper [23] is the inability to use an optical fiber probe. We foresee that this problem can be avoided with the application of new technologies, such as chalcogenide infrared fibers for 1500–6000 nm spectral range and polycrystalline infrared fibers for 4000–18 000 nm spectral range.

4.4. Other instruments

There is a variety of other analytical methods that were shown to be applicable in PAT solutions.

Fluorescence analysis, in particular, two-dimensional excitation–emission spectroscopy, is a sensitive in-line monitoring tool used in food [17] and biotechnological industries [40,71]. Visual imaging is used to determine particle size and mean layer thickness in monitoring of fluid bed pellet coating process [72,27]. Hyperspectral imaging [21] is applied to distinguish pine from spruce and to identify bark. It is shown that such an approach can replace slow laboratory analysis. A feasibility study, which applies acoustic measurements to fluidized bed granulation of a fertilizer product at a semi-industrial pilot plant, is presented in [73]. In paper [29], the Terahertz (50–300 cm^{-1}) pulse spectroscopy is shown to be a high-throughput technique with many areas of potential application in pharmaceutical

industry. X-ray diffraction [74] is one of the most widely used methodologies for the *in situ* analysis of kinetic processes involving crystalline solids. Application of the X-ray diffraction together with smoothed principal component analysis can significantly improve the signal-to-noise ratio and hence lower the detection limit. In [75], the authors propose a procedure that can be implemented as the high-performance liquid chromatography (HPLC) analysis that runs according to PAT concept aiming for real-time release. Paper [76] discusses the application of NMR, which has a huge future potential as PAT analyzer.

The necessity of rapid in-line and on-line monitoring of process data, innovations in instruments, and design of new multivariate data processing methods will inevitably change the diagram presented in Figure 4 in the future.

5. TOOLS

In this section, we consider the "multivariate tools for design, data acquisition, and analysis" [1] or, in other words, chemometric methods employed to find a PAT solution. Figure 5 presents the frequencies of a particular method usage in PAT applications. It is worth mentioning that the length of the related subsections that follow in this chapter does not correspond to the size of the pie-chart slices in Figure 5. We consider that some rarely used methods are quite promising and therefore deserve closer attention.

The present part of the review cannot be used as a tutorial as we do not explain the methods in detail. For the latter purpose, we could recommend an excellent paper on chemometrics in PAC [77] or the newest tutorial on multivariate data analysis in pharmaceutics [78].

5.1. Partial least squares

Partial least squares is the most popular chemometric tool in PAT. However, the majority of such PLS/PAT solutions can be presented by a simple formula

$$\mathrm{API} = \mathrm{PLS}(\mathrm{NIR}).$$

Publications based on the aforementioned formula present the routine applications of PLS modeling; therefore, they are not reviewed here. An exception could be made for paper [31] that claims the old-fashioned ordinary least squares (OLS) algorithms took advantage of pure component scans to produce the most sensitive calibrations when compared with PLS regression. Because the authors luckily avoid their own interpretation of sensitivity, which is a very ambivalent concept, therefore, we cannot evaluate the efficiency of the NIR calibration method applied for pharmaceutical blend monitoring.

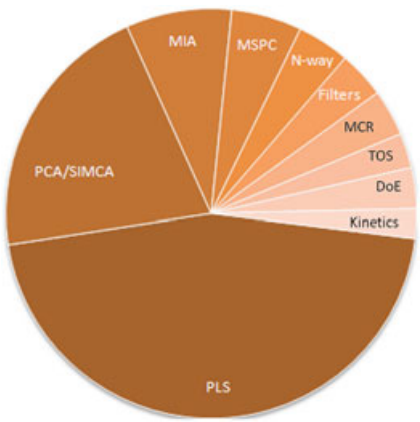


Figure 5. Chemometric tools used in process analytical technology applications.

An interesting challenge in the PLS analysis of the process data is combining the predictor variables' blocks originated from the different sources, for example, spectral readings with process parameters [24] or NIR spectra with Raman spectra [79], in one regression model. A similar problem arises in simultaneous modeling of different process stages [80]. A plain augmentation method does not work as it reduces the influence of (very) important but scanty process variables, which are typically counted in dozens whereas spectral readings are counted in thousands. Alternative approaches are LS-PLS [25] (least squares - partial least squares) regression and path modeling [81]. The LS-PLS algorithm starts with regressing the response on the design variables by using OLS. The residuals are then regressed on the spectra by using PLS regression. PLS scores are combined with the design variables in a new OLS regression, and updated regression coefficients and residuals are calculated. This procedure is iterated until a stable solution is obtained. This approach has been applied in optimizing a fish feed production process [24].

Paper [82] aims at investigating the potential benefits of combining NIR and MIR spectral regions for employment in calibration development of soybean flour quality properties. Initially, both spectra blocks are analyzed separately using single PLS, and then, they are combined in modeling by multiblock PLS in a parallel mode (MB-PLS) or in a serial mode (S-PLS). The results seem to be moderately advantageous, as each multiblock approach just reduces the number of PLS components, whereas root-mean square error of cross-validation values remain rather similar to the single NIR/PLS model.

Basic methods of path modeling are concerned with a network of data blocks, where certain matrices (data blocks) are defined as input data and some others as output data blocks. Path modeling suggests that the regression should be carried out in steps. At each step, a weighing procedure, which reflects the emphasis of the analysis, should be applied. For example, there should be a weight vector found such that the score vector has certain optimal properties [81].

On the other hand, it is problematic when single block of predictors **X** is regressed on multiresponse block **Y**. Each sample in the **Y** block could be a continuous curve representing a specific property of the sample, for example, the particle size distribution or the drug release profile. In [56], a new method for the prediction of drug release profiles during a running pellet coating process from in-line NIR measurements has been developed. The NIR spectra are acquired during a manufacturing process through an immersion probe. Yet the pellets sampled at the process time points are subjected to dissolution tests. In this case, the drug release profiles have a sigmoid form, which is modeled by the autocatalytic kinetics by using nonlinear regression. The estimated kinetic constants are then used as the new responses in PLS regression. This two-step approach enables prediction of the release profiles from the process NIR data. It was used to monitor the final pellet quality in the course of a coating process.

5.2. Principal component analysis and soft independent modeling of class analogy

Principal component analysis is the earliest approach in chemometrics but still popular for classification, especially when PCA is strengthened by the soft independent modeling of class analogy (SIMCA) method. This combination has a wide usage in PAT (Figure 5), in particular, for the MSPC applications [34].

In the SIMCA method, two distance measures can be employed. These are the space distance (SD, a.k.a. leverage) and the orthogonal distance (OD, a.k.a. the residual variance), which characterize a sample position with respect to a PCA model. For each of the distances, the critical membership levels can be established. Therefore, as soon as a new candidate object is considered, it is projected onto the PCA subspace, and its own SD and OD values are calculated. They are further compared with the known critical levels to make a decision on the membership of the class [83].

This approach, however, is often restricted to the consideration of OD value only. A typical example is paper [32], in which a halved SIMCA was applied to predict the blend homogeneity of independent blend samples under different processing conditions. In this paper, PCA and SIMCA are applied to the NIR spectra collected at-line for the designed samples. In this simple case, the OD measure alone was enough to acceptably determine homogenous samples.

It is interesting that paper [67] has the same objective. It demonstrates that Raman spectroscopy can be used for the in-line and real-time endpoint monitoring and understanding of a powder blending process. The noninvasive, in-line monitoring is a more complicated method of data acquisition, which reveals the progressive increase of homogeneity in a batch process. Several last spectra of each batch are chosen as reference spectra, which are used to build the SIMCA model. The spectra collected in the beginning of a batch are considered as the test set. The goal is to verify whether they belong to the model class or, in other words, whether or not the blend can be considered homogeneous. SIMCA analysis of the test set revealed large deviations both in OD and SD measures at the initial stage of blending. Thus, the full-scale SIMCA model based on two distance measures is crucial for the accurate determination of the process endpoint.

The critical point in the SIMCA application is the proper choice of the acceptance levels of both OD and SD values.

5.3. Multivariate image analysis

Multivariate image analysis (MIA) takes third place among the chemometric tools for PAT. This approach becomes more and more popular. However, in PAT applications, it tends to be an ancillary tool used mainly for visual monitoring or measuring some geometrical properties. At the same time, there are many other important aspects of the process samples' appearance, besides measurable distances, that can carry valuable information.

In paper [72], the MIA methods are applied to at-line monitoring of fluid bed pellet coating process. The quantitative description of images of pellet samples, taken from different process stages, has been obtained using two different approaches: wavelet decomposition and angle measure technique. Both methods reveal a strong correlation between image features and process parameters with some advantage of angle measure technique. It has been shown that pellet images, taken with a conventional digital camera, can be used for at-line monitoring of the process, in particular, for the control of the pellets' growth during coating. An algorithm for precise counting of pellets has been developed. Combined with the sample weighing, it enables an accurate determination of the mean added pellets' weight.

Multivariate image analysis is frequently used at the postprocessing stage, for example, as a reliable and rapid analytical method to identify API in finished pharmaceutical products [84,85]. Another area is combating counterfeit drugs. Paper [85] suggests a concept that combines single-point near-infrared spectroscopy (NIRS) and near-infrared chemical imaging (NIR-CI) with statistical variance analysis. The chemometric method is based on (1) summation and unfolding of multidimensional predicted classification scores, which results in a linear image signature, and (2) multivariate linear image signature data analysis. This procedure not only represents an approach for the identification of counterfeiting but is also applicable to determine the product variability.

5.4. Multivariate statistical process control

The main MSPC concept is, firstly, to apply historical data on performance attributes (**X** matrix) for construction of a linear (calibration) model, which explains how the final results (**y** vector) depend on the observed **X** variables and, secondly, to verify that the process remains in a 'state of statistical control'. Strictly speaking, MSPC is not a tool, but a concept that employs various chemometric methods. For example, paper [34] employs PLS, PCA, SIMCA, and orthogonal signal correction for MSPC of the polymorphic composition of sulfathiazole. Another paper [86] aims at control of the particles size distribution employing the PLS and PCA methods, both applied to the variablewise and the batchwise decomposed matrix of the process variables.

Conventional MSPC approach provides a *post factum* optimization, whereas the most important issue in manufacturing is an *in situ* optimization, which prescribes immediate actions in the course of production to correct its current state and to improve the future performance. Now, the traditional MSPC concept is being extended to develop an approach for in-line process optimization [81,87,88].

5.5. Filtering

Data preprocessing is always necessary in PAT. Along with the well-known methods of data filtering, such as multiplicative scatter correction [21] and orthogonal signal correction [89], new algorithms are suggested for usage in PAT [90].

Several data preprocessing tools have been developed for temperature spectral correction. For example, extended inverted signal correction [91], which was developed mainly for spectral scatter correction, is a potential alternative to account for variations in NIR data induced by temperature (to a certain extent). The main advantage of this approach is that knowledge of the temperature is not required. Another method of filtering, termed extended loading space standardization, can successfully model both the temperature-induced spectral variations and multiplicative effects caused by the fluctuations of other measurement conditions. This is achieved by standardizing spectra measured at different temperatures into an arbitrarily selected reference temperature and then estimating the multiplicative parameters from the standardized spectra [46].

A series of new preprocessing methods has been developed by Z.-P. Chen and J. Morris. A new method, systematic prediction error correction, has been developed to maintain the predictive abilities of multivariate calibration models when, for example, the spectrometer or measurement conditions are altered [49]. A unique calibration strategy called multiplicative effects correction was proposed to separate the Raman contributions because of changes in analyte concentration from those caused by the multiplicative confounding effects of the sample's physical properties [68]. Optical path-length estimation and correction method [47] was developed to linearize spectral data contaminated by multiplicative effects. Without any additional requirements imposed on spectral data, optical path-length estimation and correction can effectively correct nonlinear multiplicative effects.

5.6. Multivariate curve resolution

The multivariate curve resolution (MCR) method is still an underestimated approach in PAT. This is rather odd because MCR has a unique ability for the simultaneous qualitative and quantitative recovery of reaction constituents in real time. A good exemplar is paper [92] that demonstrates a cheaper, easier, and faster method of monitoring of the vinyl acetate monomer process. The advantage of this approach is a vast reduction in the calibration time and sample set. Therefore, this strategy could be used for reactions for which it is impossible to prepare mixtures of known composition because of the absence of isolated stable reference materials.

Paper [50] has already been mentioned among the feasibility studies focusing on calibration sets construction. Its second goal is the use of the multivariate curve resolution-alternative least squares (MCR-ALS) algorithm to examine potential polymorphic transformations of the API during granulation. It is shown that the MCR-recovered spectra of the powder and granulated samples are very similar to those for the pure API. By contrast, correlation with the spectrum for the amorphous form is much lower, which suggests that it undergoes no polymorphic transformation during granulation.

Interesting research has been performed in paper [28] that is devoted to the analysis of the effects of scatter on quantification of chemical composition. The authors suggest a novel approach that consists of two steps. First, separate scattering spectral fingerprints, denoted as S spectra, and absorption spectral fingerprints, denoted as K spectra, are determined. Second, the two spectra are used as hard model constraints for the MCR-ALS algorithm to account for the spectral distortions due to the interaction of scatter and absorption. The approach has been employed to design standard tablets prepared by mixing of three components at three different compaction levels. The samples are measured by an ultraviolet/visual/near infrared (UV/Vis/NIR) spectrometer in the range from 500 up to 2100 nm. In comparison with PLS modeling with EMSC pretreatment of spectra, the hard model constrained MCR-ALS algorithm results in an improved prediction of the concentration of the API together with a higher robustness of the calibration models.

5.7. N-way

Multiway (N-way) methods are not very popular in PAT – they constitute near 4% among other chemometric approaches. The comprehensive N-way models (such as PARAFAC and Tucker3 [17]) are seldom in PAT. Often, researchers employ a more straightforward approach that unfolds the three-dimensional array into a two-dimensional matrix, preprocesses it if necessary, and then performs PCA or PLS on the unfolded two-dimensional matrix. An example is paper [93] that presents the use of experimental design, optimization, and multiway principal component analysis to investigate the root cause of tablet dissolution shift (slow–down) upon stability and to develop control strategies for a drug product during formulation and process development.

Another example is described in publication [45] that develops a quantitative method for simultaneous measuring of the flavor and water contents in model spray-dried flavor delivery systems by using the NIR and NMR relaxation spectra. In this study, three-way data arrays are obtained by considering each sample measured at different temperatures T as an independent sample. Different calibration methods are applied to the data: two bilinear PLS (with implicit and explicit inclusion of T) and one trilinear PLS. Overall, NIR spectroscopy and NMR relaxometry are identified as complementary techniques rather than competitive methods in the characterization of encapsulated flavor systems.

A three-way calibration strategy has been used in paper [51] to develop a simple method for drug quantitation in intact pharmaceutical tablets. Parallel factor analysis (PARAFAC) is applied to deconvolute the NIR spectra in scores associated with drug concentration variation and loadings related with wavelength and compaction ranges. The PARAFAC is followed by multiple linear regression to obtain a more robust calibration model than using a traditional two-way modeling technique.

5.8. Theory of sampling

Sampling is an essential part of the PAT solution. Esbensen and co-workers have been pioneers for systemizing the sampling situation in theory of sampling (TOS).

In paper [94], the didactic data sets are used to illustrate a "how to do" representative process sampling. It is demonstrated how selected process data lead to diverse variogram expressions with different systematics. Following variogram data analysis leads to a fundamental decomposition into zero-dimensional sampling versus one-dimensional process variances, on the basis of the three principal variogram parameters: range, sill, and nugget effect. All presented cases of variography either solved the initial problems or served to understand the reasons and causes behind the specific process structures revealed in the variograms.

Paper [95] represents a practical application of TOS to monitor industrial bioenergy anaerobic digestion processes. The analytical methods investigated are at-line NIR and image analysis. Optimized sampling on four different scale levels allows acceptable PLS prediction models for the process parameters. It was found that PAT approach is critically dependent on representative reference calibration sampling, which has to be fully compliant with the TOS.

5.9. Design of experiments

The PAT Guidance [1] puts statistical DoE and response surface methodologies in the first place among multivariate mathematical approaches, which should be used for the PAT benefit. Furthermore, it is remarkable that few papers in our collection of the PAT papers report the usage of DoE (Figure 5).

In the DoE context, two of the latest publications [64,96] can be mentioned. They form a two-paper set representing work that was conducted in the framework of a QbD project involving the production of a pharmaceutical gel. By using historical data from the previously manufactured batches, two QTPPs as well as five CQAs are identified. The CQAs are considered as the predictor variables. They are used to construct a *D*-optimal design aimed to optimize QTTPs, treated as the response variables *Y*. Each response is described by a second-order polynomial, which included only binary interactions. The higher-order terms are highly unlikely under the considered experimental conditions. By identifying the best conditions for simultaneously optimizing both QTTPs, a desirability function is used.

In the second paper [64], the NIR spectroscopy is applied. The primary aim is to develop the PLS calibration models for the in-line determination of temporal changes in the CQA values, which affect the product quality. The study is completed by using the batch MSPC method to compare the batches included in the experimental design with the exemplary batches, conducted under the normal operational conditions.

The nonlinear principal component regression is applied for optimization of a hybrid binder formulation that includes a water solution of sodium silicate (water glass) and polyisocyanate [97]. Optimization is performed with respect to 10 output quality characteristics. Calibration modeling is performed as a two-step procedure. At first, PCA is applied to the **X** block for variable reduction. Then, a polynomial regression is used to predict a particular quality characteristic as a function of score vectors. The input variables' reduction enables to choose an optimal binder formulation that meets the predefined quality requirements.

5.10. Kinetics

It could be expected that kinetic studies play an important role in the PAT applications. Both hard and soft kinetic modelings are capable to disclose essential information about the manufacturing process. Moreover, being combined with spectroscopic data in a style of "gray modeling", the kinetic approach could be a powerful tool for the process control. However, in this area, the PAT approach is overlapping with another popular technique – chemical engineering. This could be a reason for this rather poor representation.

In kinetic modeling, two concepts can be distinguished. The first is a "hard" approach where a model is built on the basis of the substantial physical and chemical consideration of the phenomenon in question. Such a model is typically presented in the form of differential equations accounting for consumption and generation of the reaction constituents. The "soft" kinetic model has no solid background and is typically represented by a nonlinear data-driven function. The "hard" model may be used for the out of data prediction (extrapolation), whereas the "soft" approach is valid just inside the data area (interpolation) [98]. To illustrate these concepts, two papers are considered.

The first paper [99] aims at basic understanding of an industrial catalytic hydrogenation of API. It reports on combining the at-line monitoring technique with a kinetic model of the process. NIR spectroscopy and PLS regression are used to monitor the most relevant reaction constituents. The obtained concentration values are then employed to fit the proposed kinetic model, which is capable of describing the industrial process under diverse operating conditions. The model is presented in a "hard" form as a system of differential equations. The estimates of the kinetic parameters are found with the help of nonlinear regression.

Another paper [42] considers the extraction of meaningful information from NIR measurements of coagulating milk in real-time scale. This information is then used to develop automatic cutting time determination. The NIR spectra are compressed by PCA, and the scores values are considered as a function of process time. Milk coagulation includes three stages: κ-casein proteolysis, micelle aggregation, and network formation. Two kinetic models are proposed to describe coagulation: a total model for the entire process and a composite model for the three individual stages. Each model has a typical "soft" form being a combination of the logistic equation (autocatalytic reactions), an exponential term (order reactions), and an offset. It was shown that the total model suffers from multicolinearity; therefore, the composite model is preferable.

Paper [56] studied the process of drug dissolution by using autocatalytic kinetics that was initially proposed just as the "soft" model aimed at evaluation of kinetic parameters, which were then used as the responses in PLS regressing of the process NIR data. Afterward, on the basis of the specific properties of the dissolution profiles fitting, the model's meaning was reconsidered and a guess on the "hard" nature of the model has been claimed.

6. CONCLUSIONS

We are fully convinced that a review (simple analysis) of publications cannot provide a complete understanding of the tendencies and challenges connected with the PAT implementation. We presented a chemometrically inclined view of the phenomena, recognizing the fact that other viewpoints are equally applicable. This, however, cannot stave us off a temptation to claim several somewhat biased conclusions, which could be considered as too sound.

The first conclusion follows from Figure 6 that represents the number of PAT papers (separately CR and non-CR) published in a corresponding year. The fast growth in 2004–2006 now changes to a rather slow increase or even stabilization. Nevertheless, an interest in PAT is still high. This indicates that the PAT approach is not a fashion phenomenon but a response to the demands of modern manufacturing process. The increasing share of CR papers confirms chemometrics to be the working horse of PAT.

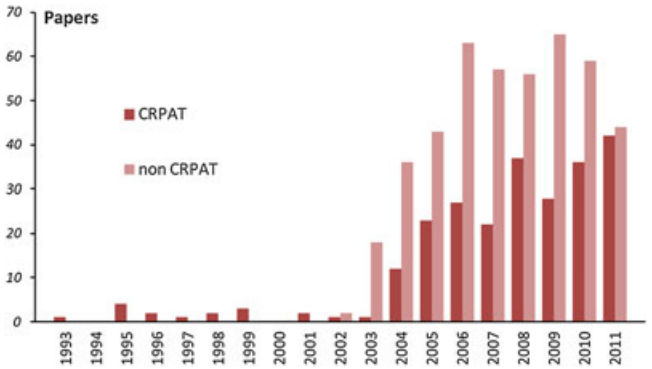


Figure 6. Papers on process analytical technology.

The second conclusion is of general (or even philosophical) nature. We suppose that PAT approach could become a new industrial paradigm broadly applicable as a new manufacturing model and going far beyond pharmaceutical (or related) areas. The old but still active industrial paradigm is based on the principles of differentiation and standardization. This means that the whole production process should be divided into as many different stages as possible, and each stage must be fully standardized to reduce variability with a goal to fit with the following stage. This concept made a revolution in the steam-based industry and is still working well in the electricity-based one. The well-known examples are the conveyer belt, Six Sigma, GMP, and ISO 900X documents. In 1950, W. Edwards Deming [100] wrote "If I had to reduce my message for management to just a few words, I'd say it all had to do with reducing variability." However, modern industrial processes are aimed at microsized and nanosized production by using the principles of biotechnology and bionics. This type of manufacturing cannot be standardized in principle. Variability is not an enemy but the inherent feature of the materials and processes in use. The old manufacturing methodology struggled with variations aiming to push them into the rigid process framework. A new manufacturing concept has to use a flexible framework that accounts for variability and adapts the process to fit them. In our view, PAT approach could be such a methodology as it fully satisfies all these demands because of various control tools: raw materials inspection, in-line monitoring, post process tuning, etc.

The PAT method may not be viewed as an approach that facilitates a process engineer's work, making it less intensive. On the contrary, PAT requires permanent attendance and treatment both at the development and implementation stages. The main PAT benefits are the additional degrees of freedom that come from the flexibility of this approach. Actual in-line monitoring of the process enables engineers to adjust it in real time with wide possibilities to keep the process in the optimal state with respect to the designed quality of the end product.

REFERENCES

1. U.S. Food and Drug Administration, Guidance for Industry, PAT – a framework for innovative pharmaceutical development, manufacturing, and quality assurance, September 2004.
2. Chew W, Sharratt P. Trends in process analytical technology. *Analytic. Meth.* 2010; **2**: 1412–1438.
3. Molière J-B. Le Bourgeois gentilhomme, Édition Louandre, 1910, tome 3.
4. Laitinen H. Editorial. What is analytical chemistry? *Anal. Chem.* 1966; **38**: 673–673.
5. Frank IE, Feikema J, Constantine N, Kowalski BR. Prediction of product quality from spectral data using the partial least-squares method. *J. Chem. Inf. Comput. Sci.* 1984; **24**: 20–24.
6. Wise BM. Using the PLS_toolbox in PAC applications. *Process Control Qual.* 1993; **5**: 73–85.
7. MacGregor JF, Kourti T. Statistical process control of multivariate processes. *Contr. Eng. Pract.* 1995; **3**: 403–414.
8. Nomikos P, MacGregor JF. Multivariate SPC charts for monitoring batch processes. *Technometrics* 1995; **37**: 41–59.
9. Kourti T, Lee J, Macgregor JF. Experiences with industrial applications of projection methods for multivariate statistical process control. *Comput. Chem. Eng.* 1996; **20**: S745–S750.
10. Martin EB, Morris AJ, Zhang J. Process performance monitoring using multivariate statistical process control. *IEE Proc. Contr. Theor. Appl.* 1996; **143**: 132–144.
11. Callis JB, Illman DL, Kowalski BR. Process analytical chemistry. *Anal. Chem.* 1987; **59**: 624A–637A.
12. Miller CE. The use of chemometric techniques in process analytical method development and operation. *Chemometr. Intell. Lab. Syst.* 1995; **30**: 11–22.
13. Kowalski BR. Identification of multiple sensor disturbances during process monitoring. *Anal. Chem.* 1997; **69**: 5031–5036.
14. Munck L, Norgaard L, Engelsen SB, Bro R, Andersson CA. Chemometrics in food science – a demonstration of the feasibility of a highly exploratory, inductive evaluation strategy of fundamental scientific significance. *Chemometr. Intell. Lab. Syst.* 1998; **44**: 31–60.
15. Wikstrom C, Albano C, Eriksson L, Friden H, Johansson E, Nordahl A, Rannar S, Sandberg M, Kettaneh-Wold N, Wold S. Multivariate process and quality monitoring applied to an electrolysis process. Part I. Process supervision with multivariate control charts. *Chemometr. Intell. Lab. Syst.* 1998; **42**: 221–231.
16. Workman J, Jr., Veltkamp DJ, Doherty S, Anderson BB, Creasy KE, Koch M, Tatera JF, Robinson AL, Bond L, Burgess LW, Bokerman GN, Ullman AH, Darsey GP, Mozayeni F, Bamberger JA, Greenwood MS. Process analytical chemistry. *Anal. Chem.* 1999; **71**: 121R–180R.
17. Bro R. Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis. *Chemometr. Intell. Lab. Syst.* 1999; **46**: 133–147.
18. Wise BM, Gallagher NB, Butler SW, White DD, Jr., Barna GG. A comparison of principal component analysis, multiway principal component analysis, trilinear decomposition and parallel factor analysis for fault detection in a semiconductor etch process. *J. Chemometr.* 1999; **13**: 379–396.
19. Sahni NS, Isaksson T, Naes T. The use of experimental design methodology and multivariate analysis to determine critical control points in a process. *Chemometr. Intell. Lab. Syst.* 2001; **56**: 105–121.
20. Berget I, Naes T. Sorting of raw materials with focus on multiple end-product properties. *J. Chemometr.* 2002; **16**: 263–273.
21. Geladi P, Sethson B, Nystrom J, Lillhonga T, Lestander T, Burger J. Chemometrics in spectroscopy: Part 2. Examples. *Spectrochim. Acta Part B At. Spectrosc.* 2004; **59**: 1347–1357.
22. Workman J, Jr., Koch M, Veltkamp D. Process analytical chemistry. *Anal. Chem.* 2005; **77**: 3789–3806.
23. Skibsted E, Engelsen SB. Spectroscopy for process analytical technology (PAT). In *Encyclopedia of Spectroscopy and Spectrometry* (2nd edn, Vol. **3**), Lindon J, Tranter G, Koppenaal D (eds). Elsevier: Oxford, 2010; 2651–2661.
24. Mage I, Naes T. Optimising production cost and end-product quality when raw material quality is varying. *J. Chemometr.* 2007; **21**: 440–450.
25. Jorgensen K, Naes T. The use of LS-PLS for improved understanding, monitoring and prediction of cheese processing. *Chemometr. Intell. Lab. Syst.* 2008; **93**: 11–19.
26. Mage I, Mevik B-H, Naes T. Regression models with process variables and parallel blocks of raw material measurements. *J. Chemometr.* 2008; **22**: 443–456.
27. Laitinen N, Antikainen O, Rantanen J, Yliruusi J. New perspectives for visual characterization of pharmaceutical solids. *J. Pharm. Sci.* 2004; **93**: 165–176.
28. Kessler W, Oelkrug D, Kessler R Using scattering and absorption spectra as MCR-hard model constraints for diffuse reflectance measurements of tablets. *Anal. Chim. Acta* 2009; **642**: 127–134.
29. Wu H, Heilweil EJ, Hussain AS, Khan MA. Process analytical technology (PAT): quantification approaches in terahertz spectroscopy for pharmaceutical application. *J. Pharm. Sci.* 2008; **97**: 970–984.
30. Skibsted ETS, Boelens HFM, Westerhuis JA, Witte DT, Smilde AK. Simple assessment of homogeneity in pharmaceutical mixing processes using a near-infrared reflectance probe and control charts. *J. Pharm. Biomed. Anal.* 2006; **41**: 26–35.
31. Zacour BM, Igne B, Drennen JK, III, Anderson CA. Efficient near-infrared spectroscopic calibration methods for pharmaceutical blend monitoring. *J. Pharmaceut. Innovat.* 2011; **6**: 10–23.
32. El-Hagrasy AS, Delgado-Lopez M, Drennen JK, III. A process analytical technology approach to near-infrared process control of pharmaceutical powder blending: part II: qualitative near-infrared models for prediction of blend homogeneity. *J. Pharm. Sci.* 2006; **95**: 407–421.
33. Chen Z-P, Morris J, Borissova A, Khan S, Mahmud T, Penchev R, Roberts KJ. On-line monitoring of batch cooling crystallization of organic compounds using ATR-FTIR spectroscopy coupled with an advanced calibration method. *Chemometr. Intell. Lab. Syst.* 2009; **96**: 49–58.

34. Pollanen K, Hakkinen A, Huhtanen M, Reinikainen S-P, Karjalainen M, Rantanen J, Louhi-Kultanen M, Nystrom L. DRIFT-IR for quantitative characterization of polymorphic composition of sulfathiazole. *Anal. Chim. Acta* 2005; **544**: 108–117.
35. Pollanen K, Hakkinen A, Reinikainen S-P, Rantanen J, Minkkinen P. Dynamic P. CA-based MSPC charts for nucleation prediction in batch cooling crystallization processes. *Chemometr. Intell. Lab. Syst.* 2006; **84**: 126–133.
36. Norgaard L, Hahn MT, Knudsen LB, Farhat IA, Engelsen SB. Multivariate near-infrared and Raman spectroscopic quantifications of the crystallinity of lactose in whey permeate powder. *Int. Dairy J.* 2005; **15**: 1261–1270.
37. Paris I, Janoly-Dumenil A, Paci A, Mercier L, Bourget P, Brion F, Chaminade P, Rieutord A. Near infrared spectroscopy and process analytical technology to master the process of busulfan paediatric capsules in a university hospital. *J. Pharm. Biomed. Anal.* 2006; **41**: 1171–1178.
38. Lopes JA, Costa PF, Alves TP, Menezes JC. Chemometrics in bioprocess engineering: process analytical technology (PAT) applications. *Chemometr. Intell. Lab. Syst.* 2004; **74**: 269–275.
39. Holm-Nielsen JB, Esbensen KH. Monitoring of biogas test plants-a process analytical technology approach. *J. Chemometr.* 2011; **25**: 357–365.
40. Haack MB, Lantz AE, Mortensen PP, Olsson L. Chemometric analysis of in-line multi-wavelength fluorescence measurements obtained during cultivations with a lipase producing Aspergillus oryzae strain. *Biotechnol. Bioeng.* 2007; **96**: 904–913.
41. Nørgaard L. Classification and prediction of quality and process parameters of thick juice and beet sugar by fluorescence spectroscopy and chemometrics. *Zuckerindustrie* 1995; **120**: 970–981.
42. Lyndgaard CB, Engelsen RB, van den Berg FWJ. Real-time modeling of milk coagulation using in-line near infrared spectroscopy. *J. Food. Eng.* 2012; **108**: 345–352.
43. Lomborg CJ, Holm-Nielsen JB, Oleskowicz-Popiel P, Esbensen KH. Near infrared and acoustic chemometrics monitoring of volatile fatty acids and dry matter during co-digestion of manure and maize silage. *Bioresour. Technol.* 2009; **100**: 1711–1719.
44. Lourenco V, Herdling T, Reich G, Menezes JC, Lochmann D. Combining microwave resonance technology to multivariate data analysis as a novel PAT tool to improve process understanding in fluid bed granulation. *Eur. J. Pharm. Biopharm.* 2011; **78**: 513–521.
45. Andrade L, Farhat IA, Aeberhardt K, Rasmus BRO, Engelsen SB. Modeling of temperature-induced near-infrared and low-field time-domain nuclear magnetic resonance spectral variation: chemometric prediction of limonene and water content in spray-dried delivery systems. *Appl. Spectrosc.* 2009; **63**: 141–152.
46. Chen Z-P, Morris J. Improving the linearity of spectroscopic data subjected to fluctuations in external variables by the extended loading space standardization. *Analyst* 2008; **133**: 914–922.
47. Chen Z-P, Morris J, Martin E. Extracting chemical information from spectral data with multiplicative light scattering effects by optical path-length estimation and correction. *Anal. Chem.* 2006; **78**: 7674–7681.
48. Hansen CL, den Berg FV, Rasmussen MA, Engelsen SB, Holroyd S. Detecting variation in ultrafiltrated milk permeates – infrared spectroscopy signatures and external factor orthogonalization. *Chemometr. Intell. Lab. Syst.* 2010; **104**: 243–248.
49. Chen Z-P, Li L-M, Yu R-Q, Littlejohn D, Nordon A, Morris J, Dann AS, Jeffkins PA, Richardson MD, Stimpson SL. Systematic prediction error correction: a novel strategy for maintaining the predictive abilities of multivariate calibration models. *Analyst* 2011; **136**: 98–106.
50. Blanco M, Bautista M, Alcala M. Preparing calibration sets for use in pharmaceutical analysis by NIR spectroscopy. *J. Pharm. Sci.* 2008; **97**: 1236–1245.
51. Alcala M, Ropero J, Vazquez R, Romanach RJ. Deconvolution of chemical and physical information from intact tablets NIR spectra: two- and three-way multivariate calibration strategies for drug quantitation. *J. Pharm. Sci.* 2009; **98**: 2747–2758.
52. Huang J, Kaul G, Cai C, Chatlapalli R, Hernandez-Abad P, Ghosh K, Nagi A. Quality by design case study: an integrated multivariate approach to drug product and process development. *Int. J. Pharm.* 2009; **382**: 23–32.
53. Rodionova OY, Sokovikov YV, Pomerantsev AL. Quality control of packed raw materials in pharmaceutical industry. *Anal. Chim. Acta* 2009; **642**: 222–227.
54. Andersson M, Svensson O, Folestad S, Josefson M, Wahlund K-G. NIR spectroscopy on moving solids using a scanning grating spectrometer – impact on multivariate process analysis. *Chemometr. Intell. Lab. Syst.* 2005; **75**: 1–11.
55. Zachariassen CB , Larsen J, van den Berg F, Engelsen SB. Use of NIR spectroscopy and chemometrics for on-line process monitoring of ammonia in low methoxylated amidated pectin production. *Chemometr. Intell. Lab. Syst.* 2005; **76**: 149–161.
56. Pomerantsev AL, Rodionova OY, Melichar M, Wigmore AJ, Bogomolov A. In-line prediction of drug release profiles for pH-sensitive coated pellets. *Analyst*, 2011; **136**: 4830–4838.
57. Rodionova O, Pomerantsev AL. NIR-based approach to counterfeit-drug detection. *TRAC-Trend. Anal. Chem.* 2010; **29**: 795–803.
58. Mark J, Andre M, Karner M, Huck CW. Prospects for multivariate classification of a pharmaceutical intermediate with near-infrared spectroscopy as a process analytical technology (PAT) production control supplement. *Eur. J. Pharm. Biopharm.* 2010; **76**: 320–327.
59. Inpharmacia. Pharma analytical report, issue 4(36), 2009 (http://www.pharmexpert.ru/analytics_files/INPHARMACIA_04_04ENG.pdf)
60. Rodionova O, Pomerantsev A, Houmoller L, Shpak A, Shpigun O. Noninvasive detection of counterfeited ampoules of dexamethasone using NIR with confirmation by HPLC-DAD-MS and CE-UV methods. *Anal. Bioanal. Chem.* 2010; **397**: 1927–1935.
61. Johansson J, Pettersson S, Folestad S. Characterization of different laser irradiation methods for quantitative Raman tablet assessment. *J. Pharm. Biomed. Anal.* 2005; **39**: 510–516.
62. Cogdill RP, Herkert T, Anderson CA, Drennen JK, III. Synthetic calibration for efficient method development: analysis of tablet API concentration by near-infrared spectroscopy. *J. Pharmaceut. Innovat.* 2007; **2**: 93–105.
63. Feudal RN, Woody NA, Tan H, Myles AJ, Brown SD, Ferre J. Transfer of multivariate calibration models: a review. *Chemometr. Intell. Lab. Syst.* 2002; **64**: 181–192.
64. Rosas JG, Blanco M, Gonzalez JM, Alcala M. Quality by design approach of a pharmaceutical gel manufacturing process, part 2: near infrared monitoring of composition and physical parameters. *J. Pharm. Sci.* 2011; **100**: 4442–4451.
65. Chen Z-P, Morris J, Martin E. Correction of temperature-induced spectral variations by loading space standardization. *Anal. Chem.* 2005; **77**: 1376–1384.
66. Dyrby M, Nørgaard L, Engelsen SB. Chemometric quantitation of the active substance (containing C=N) in a pharmaceutical tablet using near-infrared (NIR) transmittance and NIR FT-Raman spectra. *Appl. Spectrosc.* 2002; **56**: 579–585.
67. De Beer TRM, Bodson C, Dejaegher B, Walczak B, Vercruysse P, Burggraeve A, Lemos A, Delattre L, Heyden YV, Remon JP, Vervaet C, Baeyens WRG. Raman spectroscopy as a process analytical technology (PAT) tool for the in-line monitoring and understanding of a powder blending process. *J. Pharm. Biomed. Anal.* 2008; **48**: 772–779.
68. Chen Z-P, Fevotte G, Caillet A, Littlejohn D, Morris J. Advanced calibration strategy for in situ quantitative monitoring of phase transition processes in suspensions using FT-Raman spectroscopy. *Anal. Chem.* 2008; **80**: 6658–6665.
69. Chew W. Information-theoretic chemometric analyses of Raman data for chemical reaction studies. *J. Raman Spectros.* 2011; **42**: 36–47.
70. Assirelli M, Xu W, Chew W. Reactor kinetics studies via process raman spectroscopy, multivariate chemometrics, and kinetics modeling. *Org. Process Res. Dev.* 2011; **15**: 610–621.
71. Bogomolov A, Grasser T, Hessling M. In-line monitoring of saccharomyces cerevisiae fermentation with a fluorescence probe: new approaches to data collection and analysis. *J. Chemometr.* 2011; **25**: 389–399.
72. Kucheryavski S, Esbensen KH, Bogomolov A. Monitoring of pellet coating process with image analysis – a feasibility study. *J. Chemometr.* 2010; **24**: 472–480.
73. Halstensen M, de Bakker P, Esbensen KH. Acoustic chemometric monitoring of an industrial granulation production process – a PAT feasibility study. *Chemometr. Intell. Lab. Syst.* 2006; **84**: 88–97.
74. Chen ZP, Morris J, Martin E, Hammond RB, Lai X, Ma C, Purba E, Roberts KJ, Bytheway R. Enhancing the signal-to-noise ratio of X-ray diffraction profiles by smoothed principal component analysis. *Anal. Chem.* 2005; **77**: 6563–6570.

75. Laursen K, Frederiksen SS, Leuenhagen C, Bro R. Chemometric quality control of chromatographic purity. *J. Chromatogr. A* 2010; **1217**: 6503–6510.
76. Nordon A, McGill CA, Littlejohn D. Process NMR spectrometry. *Analyst* 2001; **126**: 260–272.
77. Miller CE. Chemometrics in process analytical chemistry. In *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*, Bakeev KA (ed). Wiley: Chichester, 2010; 353–438.
78. Rajalahti T, Kvalheim OM. Multivariate data analysis in pharmaceutics: a tutorial review. *Int. J. Pharm.* 2011; **417**: 280–290.
79. Bogomolov A, Engler M, Melichar M, Wigmore A. In-line analysis of a fluid bed pellet coating process using a combination of near infrared and Raman spectroscopy. *J. Chemometr.* 2010; **24**: 544–557.
80. Pomerantsev AL, Rodionova OY, Höskuldsson A. Process control and optimization with simple interval calculation method. *Chemometr. Intell. Lab. Syst.* 2006; **81**(2): 165–179.
81. Höskuldsson A, Rodionova O, Pomerantsev A. Path modeling and process control. *Chemometr. Intell. Lab. Syst.* 2007; **88**: 84–99.
82. Brás LP, Bernardino SA, Lopes JA, Menezes JC. Multiblock PLS as an approach to compare and combine NIR and MIR spectra in calibrations of soybean flour. *Chemometr. Intell. Lab. Syst.* 2005; **75**: 91–99.
83. Pomerantsev AL. Acceptance areas for multivariate classification derived by projection methods. *J. Chemometr.* 2008; **22**: 601–609.
84. Alvarenga L, Ferreira D, Altekruse D, Menezes JC, Lochmann D. Tablet identification using near-infrared spectroscopy (NIRS) for pharmaceutical quality control. *J. Pharm. Biomed. Anal.* 2008; **48**: 62–69.
85. Puchert T, Lochmann D, Menezes JC, Reich G. Near-infrared chemical imaging (NIR-CI) for counterfeit drug identification – a four-stage concept with a novel approach of data processing (linear image signature). *J. Pharm. Biomed. Anal.* 2010; **51**: 138–145.
86. Huang J, Goolcharran C, Utz J, Hernandez-Abad P, Ghosh K, Nagi A. A PAT approach to enhance process understanding of fluid bed granulation using in-line particle size characterization and multivariate analysis. *J. Pharmaceut. Innovat.* 2010; **5**: 58–68.
87. Henriksen HC, Naes T, Rodbotten R, Aastveit A. Simultaneous modelling of process variables and raw material properties as measured by NIR. A case study from cellulose production. *Chemometr. Intell. Lab. Syst.* 2005; **77**: 238–246.
88. Bogomolov A. Multivariate process trajectories: capture, resolution and analysis. *Chemometr. Intell. Lab. Syst.* 2011; **108**: 49–63.
89. Stenlund H, Johansson E, Gottfries J, Trygg J. Unlocking interpretation in near infrared multivariate calibrations by orthogonal partial least squares. *Anal. Chem.* 2009; **81**: 203–209.
90. Rinnan Å, van den Berg F, Engelsen SB. Review of the most common pre-processing techniques for near-infrared spectra, *TRAC-Trend. Anal. Chem.*2009; **28**: 1201–1222.
91. Dyrby M, Petersen RV, Larsen J, Rudolf B, Norgaard L, Engelsen SB. Towards on-line monitoring of the composition of commercial carrageenan powders. *Carbohydr. Polym.* 2004; **57**: 337–348.
92. Richards SE, Becker E, Tauler R, Walmsley AD. A novel approach to the quantification of industrial mixtures from the vinyl acetate monomer (VAM) process using near infrared spectroscopic data and a quantitative self modeling curve resolution (SMCR) methodology. *Chemometr. Intell. Lab. Syst.* 2008; **94**: 9–18.
93. Huang J, Goolcharran C, Ghosh K. A quality by design approach to investigate tablet dissolution shift upon accelerated stability by multivariate methods. *Eur. J. Pharm. Biopharm.* 2011; **78**: 141–150.
94. Esbensen KH, Friis-Petersen HH, Petersen L, Holm-Nielsen JB, Mortensen PP. Representative process sampling - in practice: variographic analysis and estimation of total sampling errors (TSE). *Chemometr. Intell. Lab. Syst.* 2007; **88**: 41–59.
95. Holm-Nielsen JB, Dahl CK, Esbensen KH. Representative sampling for process analytical characterization of heterogeneous bioslurry systems-a reference study of sampling issues in PAT. *Chemometr. Intell. Lab. Syst.* 2006; **83**: 114–126.
96. Rosas JG, Blanco M, Gonzalez JM, Alcala M. Quality by design approach of a pharmaceutical gel manufacturing process, part 1: determination of the design space. *J. Pharm. Sci.* 2011; **100**: 4432–4441.
97. Starovoitova IA, Khozin VG, Abdrakhmanova LA, Rodionova OY, Pomerantsev AL. Application of nonlinear PCR for optimization of hybrid binder used in construction materials. *Chemometr. Intell. Lab. Syst.* 2009; **97**(1): 46–51.
98. Pomerantsev AL, Rodionova OY. Hard and soft methods for prediction of antioxidants' activity based on the DSC measurements. *Chemometr. Intell. Lab. Syst.* 2005; **79**(1–2): 73–83.
99. Rodrigues LO, Lopes JA, Cardoso JP, Menezes JC. A PAT study of an industrial catalytic hydrogenation of an active pharmaceutical ingredient. *Computer Aided Chemical Engineering* 2004; **18**: 775–780.
100. Deming WE. *Some Theory of Sampling*. Dover Publications: New York, 1950. ISBN: 0-486-64684-X.

