# Multi-Modal-Spectroscopy and Multivariate Data Analysis as a Tool for Non-Invasive Process Analysis

R.W. Kessler[1], K. Rebner[1] and W. Kessler[2]
[1]Process Analysis & Technology (PA&T), Reutlingen University, Alteburgstr. 150,
72762 Reutlingen, Germany,
rudolf.kessler@reutlingen-university.de
[2]Steinbeis Transferzentrum Prozesskontrolle und Datenanalyse,
Herderstr. 47, 72762 Reutlingen, Germany,

**Abstract:**

The aim of Process Analytical Technology (PAT) is essentially to understand and control a process. Online and inline process spectroscopy is a well-established and widely used technique to non-invasively control complex reactions in industry on a molecular basis. In combination with standard multivariate data analysis (MVA), the spectral information is correlated to response variables. However, often these MVA results into empirical calibration models which are based on pure statistical correlations and therefore are lacking causality and first principle interpretation. This paper describes the procedure how to use multi-modal spectroscopy as a tool to separate multi-dimensional information. This allows e.g. to separate scatter from absorption when opaque systems are investigated. Furthermore, by means of a pushbroom imaging system it is possible to identify the distribution of the chemical components on a surface. Pushbroom imaging with multi point detection ability can reduce the costs for online or inline PAT in the near future.

**Key words:** process analysis, chemometrics, PAT, knowledge based production, multi modal optical spectroscopy,

## Introduction

The role of PAT is essentially to promote an understanding of the process on a basic mechanistic level and introduce knowledge-based production. In January 2011 a new "Guidance for Industry: Process Validation: General Principles and Practices" was introduced by the FDA [1]. The Guideline recommends that producers should achieve an understanding of the process based on first principles during Stage 1 (Process Design) and develop a process control strategy during the development phase. In Stage 2 (Process Qualification), qualification and verification should be carried out on the process in order to ensure reproducible results. Finally during Stage 3 (Continued Process Verification), in-line control, and spectroscopic techniques in particular, help manage and continually improve the process. In this guidance it is strongly emphasized, that the process should be understood from first principles and from a scientific and engineering point of view. Cost pressure, globalization and quality assurance will undoubtedly stimulate significant demand for process analytical technology (PAT) and PAT-automation in the future. The following strategies have already become apparent in the future roadmap of PAT:

• Rationalization of processes that are not already automated

• Proactive process and quality management: Integration of quality assurance into production along with increased system availability and reliability

• Enhanced product safety and increased production efficiency

The increasing complexity of high quality products and the cost pressure demand new strategies. In the future, the focus will not only be on process optimization but also on tailoring property profiles of products according to specifications and individual desires of customers. This will allow the fast creation of operating instructions including the choice of materials and process parameters to design a product with a given preference profile. The term multimodal relates to the possibility of

collecting information on chemical and physical parameters at the same time and also includes the use of the fingerprint information of a process trajectory [2]. Multimodal methods have matured to an accepted new technology especially in the field of medicine [3] and they will also enable the realization of multipurpose optimization in processing [4].

## Knowledge Based Production

There are a number of possible approaches to knowledge-based process development and production [4,5,6]. Instrument-based process analytical technology expands the traditional spectrum of variables (e.g. temperature and pressure) to optimize control of sophisticated production processes. The progress of a reaction or process can be controlled most effectively when this information is directly accessible. Besides the continued technological development of process analytical technology, another goal is to increase confidence in the complex technology. This is an issue which researchers, equipment manufacturers and users will have to address together.

Equipment, process and approval data that already exists in plants and labs provide another route to knowledge-based production. These data are currently not being fully exploited to achieve a thorough understanding of production. The goal in this instance is to deploy information management systems in lab, production and company environments for structured acquisition of data which transcends the process chain. Statistical analysis can then identify possible ways of improving the process. Analytical information, set points and device messages functioning as soft sensors are already being evaluated as a combined data set in some initial process control applications without the need for additional software investment.

Another source of information on a molecular basis is spectroscopy. Nowadays, online and inline process spectroscopy is a well-established and widely used technique to control chemical reactions in industry on a molecular basis [6,7,8,9]. In combination with standard multivariate data analysis (MVA) the spectral information is correlated to response variables which can be the quality and the concentration of a chemical component in a mixture. Ultraviolet- and visible (UV/Vis) spectroscopy is a highly sensitive technique for electronic transitions while mid infrared (MIR) spectroscopy is specific for vibrational transitions. NIR-spectroscopy is less sensitive due to lower yields of the higher order

vibrational transition probabilities. However, although not easily directly interpretable, the major advantage with NIR is that even at higher concentrations no sample preparation (e.g. dilution) is needed. It is important to emphasize that both NIR and MIR spectroscopy are highly sensitive to water absorption. In recent years, Raman spectroscopy has developed into a highly sensitive and versatile technique and therefore has proven a very suitable process monitoring tool. For special applications fluorescence spectroscopy is certainly one of the most sensitive techniques in inline analysis. Even more, 2D- fluorescence spectra (also labeled as excitation-emission spectra) allow separating optically multicomponent systems without the necessity of a separation by chromatography.

## PAT in scattering systems: a new strategy

In scattering systems, the interaction of light (photons) is complex and includes refraction, specular and diffuse reflectance and/or transmission as well as absorption and scattering simultaneously. Due to the diffusion of photons, even the spatial identification and attribution to a defined spatial coordinate in x and y may diminish. Samples that contain phase boundaries simultaneously display absorption and scattering effects and therefore the system cannot be characterized by a single measurement.

E.g. Raman spectroscopy (standard, confocal or spatially resolved), is certainly a perfect technique to follow highly sensitive changes in the chemical and to a certain extent also morphological changes in nanoparticle synthesis. However, quantification of Raman spectra in light scattering media is difficult because of path length variations of the exciting laser and emitted Raman scatter. Laser scattering methods, particle video microscopy or dynamic light scattering techniques can act as reference methods to characterize the morphology of the particulate system, but again, there are numerous sources of errors which may lead to misinterpretations of the data. The combination of UV/Vis and NIR spectroscopy may indirectly provide information on particle size especially in those areas where no absorption is measured.

The state of the art approach in Multivariate Data Analysis of spectroscopic data from particulate systems is to exclude the scattering information from the spectral features by data pre-treatment procedures like Standard Normal Variate (SNV), Multiplicative Scattering Correction ((extended) MSC) or Orthogonal

Signal Correction (OSC) to obtain unperturbed quantitative information [10,11]. In this case, the information scatter is often regarded as unwanted and therefore eliminated instead of being used as supplementary information on the morphology of the substrate.

Thus, the knowledge based process control of multi-component and multi-phase systems should ideally use both information of absorbance (which describes the chemistry) and of scatter (which describes the morphology as particle size, size distribution and porosity) to exploit the full potential of the spectral information. An ideal approach would be for multi-component and multi-phase systems to use the information of the absorbance (which describes the chemistry) and of the scatter (which describes indirectly the morphology e.g. as particle size, size distribution, agglomeration and porosity). To follow this idea in the case of characterizing opaque systems, more than one measurement is needed to separate absorbance from scatter as described by Rayleigh and Mie-Theory, or the Radiative Transfer Equations (RTE) and Kubelka Munk (KM) theory. When this "first principle spectroscopy is integrated into most modern MVA methods like Multivariate Curve Resolution (MCR), causality is obtained between the data and response variables. This closes the gap between empirical correlative and first principle process information and control [12-14]. An example is shown in figure 1. Here the spectra of an aspirin tablet in transmission and reflection is shown and the scatter spectrum as well as the absorption spectrum was calculated by means of the Kubelka Munk function [6,9,15]. It can clearly be demonstrated that the separated absorption spectrum can easily be integrated into a multivariate data analysis procedure with a robust calculation of the API concentration. In multiblock analysis, all information spectra can be used for correlation analysis.

## Multimodal Spectroscopy: "All in One Sensors"

As described before, complex systems can only be described by multiple information systems. In principle, the implementation of a multimodal inline process control approach can be achieved in three different ways:

The first implementation strategy uses different wavelength ranges. The UV range is more suitable for scattering effects (s $\sim 1/\lambda^n$), while the short near-infrared range has a much higher penetration depth because of the low absorption cross sections.
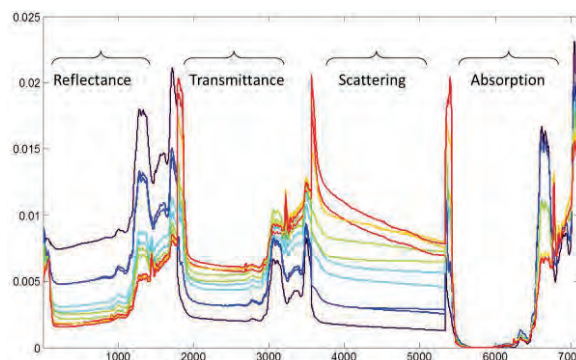


Fig.1: Spectra of an aspirin tablet in cellulose excipient of different particle size, measured in transmittance and reflectance in the wavelength range between 200nm until 2300nm. The scatter spectrum as well as the absorption spectrum is calculated by means of the Kubelka Munk function [6,9,15]. The x-axis is in channel numbers, the y-axis is in arbitrary units to present the multiblock information.

Fluorescence techniques for excitation and emission spectroscopy can be used to separate molecules due to their different spectral signatures. Raman Spectroscopy as a light scattering technique can be used as a finger print method to identify molecular structures or bonding effects.

The second implementation strategy involves using different optical set ups within the same wavelength ranges. For example, diffuse reflectance and transmittance in the UV or NIR can be used to separate morphological scattering from chemical absorption. Using (inverse) Monte Carlo simulations and integrating first principles like e.g. Mie-Theory and Photon Diffusion Theory it is possible to separate the morphological from the chemical information [14-18]. In addition Attenuated Total Reflection (ATR) Spectroscopy can be used to analyze highly concentrated samples without prior dilution and avoiding interference by scatter.

The third implementation strategy deals with laterally resolved measurements to achieve the desired differentiation. Angular resolved spectral measurements or line scans with a pushbroom imaging system (see next chapter) lead to different penetration depths, which are highly specific for particulate systems and can be used to describe particle properties. Besides the chemical information, also parameters like homogeneity, particle size, particle distribution and density may be analyzed indirectly.

**Multivariate Data Analysis: Integrating first principles**

The new approach as described above assumes to determine a separate scatter-ing spectral fingerprint and a separate absorption spectral fingerprint exploiting the information of the multimodal optical spectra. In a first step the (inverse) Monte Carlo simulation is used to calculate from the superposed spectra the s- and the k-spectrum. In a second step, the two spectra may then be used as input to an alternating least square (ALS) algorithm in Multivariate Curve Resolution (MCR) in order to account for the spectral distortions and thus simultaneously integrating first principles [10,11,15]. The major reason of increasing interest in MCR solved by alternating least squares (MCR-ALS) is its ability to extract from a complex spectral feature a) the number of involved components b) to attribute the resulting spectra to chemical compounds and c) to quantify the individual spectral contributions. Thus for a chemist interpretable loadings which represent original spectra are obtained.

Another approach is to use the concept of multivariate process trajectories [2]. This relatively new technique is suitable when the system shows a high complexity and first principles fail to describe the process. A trajectory is the intrinsic sig-nature of a process and is surrounded by a space that is constituted of an ensemble of variables selected to represent the process course. In our case this could be all the sensor measurements, process variables as well as the complete multimodal spectroscopic information thus combining information of molecular markers (chemical and morphological) from the spectroscopic measurements with standard process data. The trajectory is then the projection of the entire process phenomenon onto a chosen space. Chemometric tools like PCA, PLS or MCR are used to reduce the dimensionality of the space and to visualize the course of the trajectory.

The perfect way to relate product variability to process and raw material changes is to use a Design of Experiments (DoE) strategy. However, it is sometimes difficult to select the appropriate parameters and parameter settings for the design. A parameter is a measurable value which can describe the characteristics of a system, e.g., temperature, pressure etc. Very often, these parameters may be the factors (= independent variables) which predominantly influence the process and product quality ("critical process parameter or factor", CPP). DoE allows to identify the relevant factors and to quantify their relative importance.

**Multimodal Spatially Resolved and Multipoint Spectroscopy:**

Another approach to get simultaneously information on the chemical composition and the morphological variation is photon diffusion spectroscopy [19,20]. In this case, the substrate is spot illuminated and photons diffuse through the particulate system. The photons can then be measured spatially and spectrally resolved using a pushbroom imager. Figure 2 shows the principle of photon diffusion spectroscopy and the optical set up of a pushbroom imager.
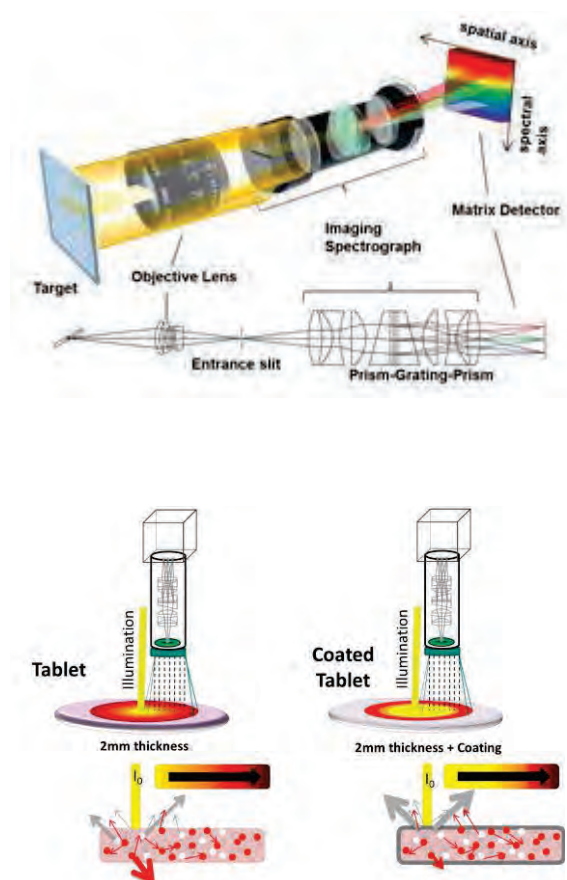


Fig. 2: top: principle of a Pushbroom Imager (courtesy of Specim), bottom: principle of photon diffusion spectroscopy example: API in a tablet resp. coated tablet, the probability to measure bulk information increases with th distance from the illumination spot

In pushbroom imaging, transmission or reflectance spectra are registered through an e.g. fixed prism-grating-prism optics with a two dimensional CCD camera attached to it. Thereby the x-axis of the CCD array corresponds to the spatial resolution and the y-

axis of the camera provides the full spectra of the sample (see figure 2).

Instead of using a single spectrometer at each individual production step, pushbroom imaging systems with numerous attached fiber bundles allow separate control of the quality at each intermediate and the final production step. Many fibers per spectrometer can be used for simultaneous measurements. In addition, different spectrometer technologies e.g. UV-Vis-NIR or even fluorescence can be combined. Thereby the probe becomes a multi information system, where the spectra describe the processing line in an ideal and complete way as a trajectory. This allows also a tremendous cost reduction as a single pushbroom imager can be used along the whole process chain [9,20,21].

## Conclusions

PAT is a highly inter- and trans-disciplinary concept and will highly beneficiary to the industry. The role of the persons involved in this new strategy is to bring the different disciplines together. The lack of qualified personnel makes realization of this goal difficult. The only way forward is to develop a new educational and training strategy for industry and increase the level of cooperation be-tween universities and the industrial sector. It is vital to foster the trialogue between universities, equipment manufacturers and end users and embed this knowledge into every facet of day-to-day operations.

## References

[1]  U. S. Department of Health and Human Services, Food and Drug Administration, Guidance for Industry, "Process Validation: General Principles and Practices", (Revison January 2011)

[2] A. Bogomolov, Multivariate Process Trajectories, Capture, Resolution and Analysis, Chemometrics and Intelligent Laboratory Systems 108, 49 -63 (2011)

[3] E. Salomatina, A. Muzikansky, V. Neel, A. N. Yaroslavsky, Multimodal Optical Imaging and Spectroscopy for the Intraoperative Mapping of Nonmelanoma Skin Cancer, Journal of Applied Physics 105, 102 - 110  (2009)

[4] R. W. Kessler: Perspectives in Process Analysis, Journal of Chemometrics,  in print  (2013), online: DOI: 10.1002/cem.2549

[5] trend report No. 8 ACHEMA 2012, Process Analytical Technology: http://presse.achema.de/achema_presse/en/trendreports.html

[6] R. W. Kessler (Ed.) Prozessanalytik. Strategien und Fallbeispiele aus der industriellen Praxis, 2006, Wiley – VCH, Weinheim, ISBN 3-527-31196-3

[7] J. Workman, B. Lavine, R. Chrisman, M. Koch, Process Analytical Chemistry, Analytical Chemistry 83, 4557 – 4578  (2011)

[8] K. A. Bakeev [ed.], Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries, 2$^{nd}$ Edn, John Wiley and Sons, Chicester, UK (2010).

[9] A. Kandelbauer, M. Rahe, R. W. Kessler, Process Control and Quality Assurance - Industrial Perspectives 1 - 69, in:  Handbook of Biophotonics, Vol.3: Photonics in Pharmaceutics, Bioanalysis and Environmental Research, 1$^{st}$ Edition by J. Popp, V. Tuchin, A. Chiou and S. Heinemann,  2012,  Wiley-VCH Verlag GmbH & Co. KGaA.

[10] S. Brown, R. Tauler, R. Walczak (Eds.), Comprehensive Chemometrics, Elsevier, Amsterdam (2009)

[11] W. Kessler, Multivariate Datenanalyse in der Bio- und Prozessanalytik. Wiley-VCH, Weinheim, Germany (2006).

[12] J. Jaumot, R. Gargallo, A. de Juan and R. Tauler, A graphical User-Friendly Interface for MCR-ALS: A New Tool for Multivariate Curve Resolution in MATLAB", Chemometrics and Intelligent Laboratory Syststems 76, 101-110 (2005)

[13] W. Kessler, R. W. Kessler, Multivariate Curve Resolution – Integration von Wissen in Chemometrische Modelle, Chemie Ingenieur Technik 82, 441 – 451 (2010)

[14] Z. Shi, C.A. Anderson, Pharmaceutical Applications of Separation of Absorption and Scattering in Near-Infrared Spectroscopy (NIRS), Journal of Pharmaceutical Science  99, 4766 – 4783 (2010)

[15] W. Kessler, D. Oelkrug, R. W. Kessler, Using Scattering and Absorption Spectra as MCR-Hard Model Constraints for Diffuse Reflectance Measurements of Tablets, Analytica Chimica Acta 642, 127 – 134 (2009)

[16] Z. Shi, C. Anderson, Application of Monte Carlo Simulation-Based Photon Migration for Enhanced Understanding of Near-Infrared (NIR) Diffuse Reflectance. Part I: Depth of Penetration in Pharmaceutical Materials, Journal of Pharmaceutical Science 99, 2399-2412 (2010).

[17] D. Oelkrug, M. Brun, K. Rebner, B. Boldrini, R. Kessler, Penetration of Light into Multiple Scattering Media: Model Calculations and Reflectance Experiments. Part I: The Axial Transfer, Applied  Spectroscopy  66(8), 934-943 (2012)

and: D. Oelkrug, M. Brun, P. Hubner, K. Rebner, B. Boldrini, R. Kessler,"Penetration of Light into Multiple Scattering Media: Model Calculations and Reflectance Experiments. Part II: The Radial

Transfer, Applied Spectroscopy 67(4), 385-395. (2013)

[18] D. Oelkrug, E. Ostertag, R. W. Kessler, Quantitative Raman Spectroscopy in Turbid Matter: Reflection or Transmission Mode? Analytical and Bioanalytical Chemistry 405, 367–3379  (2013)

[19] Bogomolov, S. Dietrich, B. Boldrini, R. W. Kessler, Quantitative Determination of Fat and Total Protein in Milk Base on Visible Light Scatter, Food Chemistry, 134, 412–4182012,

[20] B. Boldrini, W. Kessler, K. Rebner, R. W. Kessler, Hyperspectral Imaging: a Review of Best Practice, Performance and Pitfalls for Inline and Online Applications, Journal of Near Infrared Spectroscopy 20, 438–508 (2012)

[21] P. Trefz, B. Boldrini, R. W. Kessler, S. Löbbecke, Online-Analyse von Mikro-Reaktionsprozessen mittels Pushbroom Imaging, Chemie Ingenieur Technik 82, 525 – 530 (2010)