Vol. 16 No. 1 January/March 2020

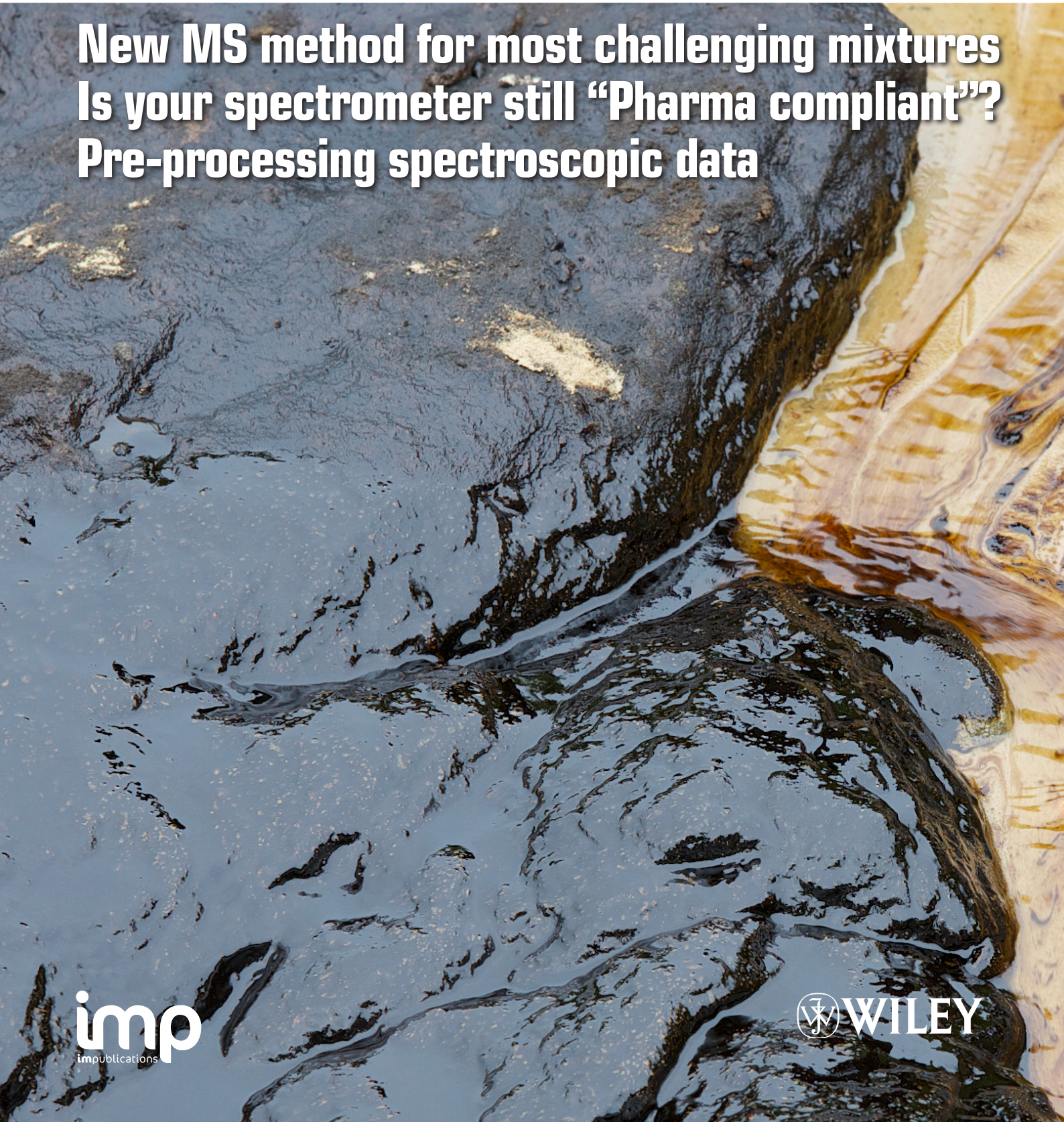
SPECTROSCOPY asia

The essential magazine for spectroscopists in the Asia/Pacific region

New MS method for most challenging mixtures

Is your spectrometer still "Pharma compliant"?

Pre-processing spectroscopic data



imp
impublications

WILEY

The analysis of highly complex mixtures presents significant analytical challenges. Amongst the range of such samples, petroleum is considered one of nature's most complex mixtures. A new mass spectrometry method has the answer and is described in the article starting on page 6.

Publisher
Ian Michael
(ian@impublications.com)

Advertising Sales UK and Ireland
Ian Michael
Spectroscopy Asia, 6 Charlton Mill, Charlton, Chichester, West Sussex PO18 0HY, United Kingdom. Tel: +44-1243-811334, Fax: +44-1243-811711, E-mail: ian@impublications.com

Americas
Joe Tomaszewski
John Wiley & Sons Inc.
Tel: +1-908-514-0776
E: jtomaszews@wiley.com

Europe and the Rest of World
Vanessa Winde
Wiley-VCH Verlag GmbH & Co. KGaA, Boschstraße 12, 69469 Weinheim, Germany
Tel: +49 6201 606 721
E-mail: vanessa.winde@wiley.com

Circulation
Spectroscopy Asia Subscriptions, IM Publications, 6 Charlton Mill, Charlton, Chichester, West Sussex PO18 0HY UK.
E-mail: subs@impublications.com

Spectroscopy Asia is a joint publication of John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, UK, and IM Publications, 6 Charlton Mill, Charlton, Chichester, West Sussex PO18 0HY, UK.

Vol. 16 No. 1
January/March 2020

ISSN 1743-5277

CONTENTS

Spectroscopy Asia is a controlled circulation journal, published four times a year and available free-of-charge to qualifying individuals in the Asia–Pacific region. Others can subscribe at the rate of €120 (Europe), £83 (UK), $160 (ROW, postage included) for the four issues published in 2016. All paid subscription enquiries should be addressed to: *Spectroscopy Asia*, John Wiley & Sons Ltd, Journals Administration Department, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, UK.

Printed in Singapore by C.O.S. Printers Pte Ltd.

WILEY

imp impublications

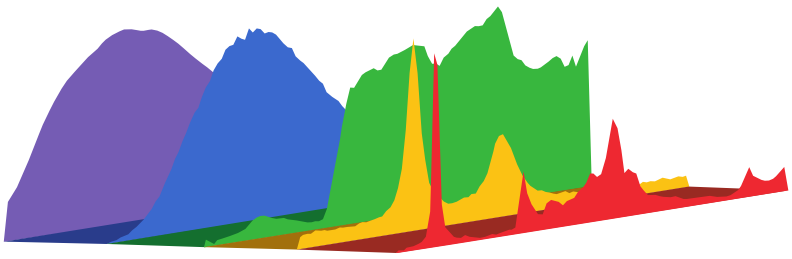
EDITORIAL

Our first article is "New mass spectrometry method for characterisation of the most challenging complex mixtures", written by Diana Lozano and Mark Barrow from the University of Warwick, UK. They have developed a clever solution to be able to provide constant resolution, high-resolution mass spectrometry data across the spectrum. The resolving power afforded by their method, called OCULAR, is sufficient to resolve peaks separated by a difference equivalent to only the mass of one electron. They have assigned 244,779 unique elemental compositions: a record.

Our second article, by Nathan Hulme and John Hammond, is titled "Is your spectrophotometer still 'Pharma compliant'? A review of the new European Pharmacopoeia 10th Edition". The latest edition of the European Pharmacopoeia on ultraviolet and visible spectroscopy has become mandatory as of 1 January 2020, so those of you who need to comply with its requirements will find this of particular interest. Nathan and John pick apart the significant changes with a view to their practical application for instrument users. Cells, control of equipment performance, wavelength accuracy, absorbance accuracy, photometric linearity, stray light and resolution, system suitability and reference materials are all covered.

Tony Davies, Jan Gerretzen and Henk-Jan van Manen consider "Pre-processing spectroscopic data: for good or ill?". They offer many useful points to consider when using pre-processing techniques. In particular, think before you pre-process: is it really necessary and what is most suitable?

The Sampling Column is from Kim Esbensen and Brad Swarbrick, and looks at "Sampling for spectroscopic analysis: consequences for multivariate calibration". Kim and Brad point out that incorrect sampling is irreversible: no amount of chemometrics or further samples will be able to produce a valid model if the sampling is not representative. This applies in flowing PAT analyses as much as in static.



NEWS

New graphene amplifier for THz radiation

A team of physicists has created a new type of optical transistor—a working THz amplifier—using graphene and a high-temperature semiconductor. The physics behind the simple amplifier replies on the properties of graphene, which is transparent and is not sensitive to light and whose electrons have no mass. It is made up of two layers of graphene and a superconductor, which trap the graphene mass-less electrons between them, like a sandwich. The device is then connected to a power source. When the THz radiation hits the graphene outer layer, the trapped particles inside attach themselves to the outgoing waves giving them more power and energy than they arrived with.

Professor Fedor Kusmartsev, of the University of Loughborough's Department of Physics, said: "The device has a very simple structure, consisting of two layers of graphene and superconductor, forming a sandwich. As the THz light falls on the sandwich it is reflected, like a mirror. The main point is that there will be more light reflected than fell on the device. It works because external energy is supplied by a battery or by light that hits the surface from other higher frequencies in the electromagnetic spectrum. The THz photons are transformed by the graphene into massless electrons, which, in turn, are transformed back into reflected, energised, THz photons. Due to such a transformation the THz photons take energy from the graphene—or from the battery—and the weak THz signals are amplified."

The research was carried out by researchers from Loughborough University, in the UK; the Center for Theoretical Physics of Complex Systems, in Korea; the Micro/Nano Fabrication Laboratory Microsystem and THz Research Center, in China and the AV Rzhanov Institute of Semiconductor Physics, in Russia. It has been accepted for publication in *Physical Review Letters*.

The team is continuing to develop the device and hopes to have prototypes ready for testing soon. Professor Kusmartsev said they hope to have a working amplifier ready for commercialisation in about a year.

Is Raman winning the non-invasive glucose monitoring race?

MIT scientists have now taken an important step toward making Raman spectroscopy a practical tool for diabetic patients to use to monito their blood sugar levels without a needle prick. They have shown that they can use it to directly measure glucose concentrations through the skin, as described in a paper in *Science Advances* (https://doi.org/10.1126/sciadv.aay5206). Until now, glucose levels had to be calculated indirectly, based on a comparison between Raman signals and a reference measurement of blood glucose levels. While more work is needed to develop the technology into a user-friendly device, this advance shows that a Raman-based sensor for continuous glucose monitoring could be feasible.

MIT's Laser Biomedical Research Center has been working on Raman-spectroscopy-based glucose sensors for more than 20 years. The NIR laser beam used for Raman spectroscopy can only penetrate a few millimetres into tissue, so one key advance was to devise a way to correlate glucose measurements from the interstitial fluid to blood glucose levels. However, another key obstacle remained: the signal produced by glucose tends to get drowned out by the many other tissue components found in skin.

"When you are measuring the signal from the tissue, most of the strong signals are coming from solid components such as proteins, lipids and collagen. Glucose is a tiny, tiny amount out of the total signal. Because of that, so far we could not actually see the glucose signal from the measured signal", Kang says.

To work around that, the MIT team has developed ways to calculate glucose levels indirectly by comparing Raman data from skin samples with glucose concentrations in blood samples taken at the same time. However, this approach requires frequent calibration, and the predictions can be thrown off by movement of the subject or changes in environmental conditions. For the new study, the researchers developed a new approach that lets them see the glucose signal directly. The novel aspect of

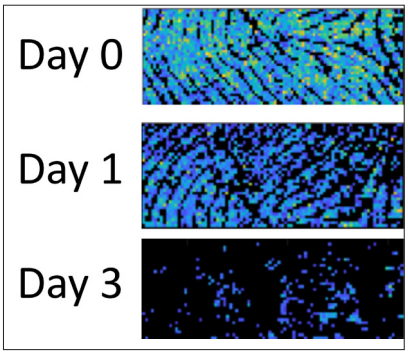
their technique is that they shine NIR light onto the skin at about a 60° angle, but collect the resulting Raman signal from a fibre perpendicular to the skin. This results in a stronger overall signal because the glucose Raman signal can be collected while unwanted reflected signal from the skin surface is filtered out.

"This is the first time that we directly observed the glucose signal from the tissue in a transdermal way, without going through a lot of advanced computation and signal extraction", So says.

Further development of the technology is needed before the Raman-based system could be used to monitor people with diabetes, the researchers say.

MS imaging to detect age of fingerprints

Police have long relied on fingerprints to identify suspects, however, there has been no way to tell how



Levels of an unsaturated triacylglycerol decline in fingerprints from an individual from day 0 (top) to day 1 (middle) and day 3 (bottom). Credit: adapted from Analytical Chemistry 2020, doi: https://doi.org/10.1021/acs.analchem.9b04765

long ago those prints were left behind: information that could be crucial to a case. A preliminary study in *Analytical Chemistry* (doi: https://doi.org/10.1021/acs.analchem.9b04765) reports that compounds contained in fingerprints could be linked to their age. Scientists have already started mining fingerprint residues for clues to the identity of the person who made them, but timing has proven more difficult to reliably pin down. Notably, past research has shown that a gas chromatography-mass spectrometry method succeeded in determining if prints were more or less than eight days old; however, investigators often need more precision. To get a better idea of when prints were deposited, Young Jin Lee and colleagues looked to reactions already suspected to take place in these residues, when ozone in air reacts with unsaturated triacylglycerols left by a fingertip.

Using prints collected from three donors, the researchers tracked shifting levels of triacylglycerols using mass spectrometry imaging. They found they could reliably determine the triacylglycerol degradation rate for each person over the course of seven days. But the rate differed among individuals, with one person's triacylglycerols declining more gradually than the others. The researchers attribute this difference to higher levels of lipids in that individual's fingerprints. The method also worked on residues that had been dusted with forensic powder. The researchers say that although a large-scale study is needed to better understand how lipid levels affect triacylglycerol degradation, this analysis is a first step toward developing a better fingerprint ageing test.

ICP-MS used for the characterisation of microplastics

A team of researchers from Ghent University (UGent) and VITO (an independent Flemish research organisation in the area of cleantech and sustainable development) has now developed a method based on inductively couple plasma-mass spectrometry (ICP-MS) for the characterisation of microplastics (MP). The approach relies on the ultra-fast monitoring of transient signals (with a detector dwell time of only 100 µs) when using a quadrupole-based ICP-MS instrument in single-event mode and registering the signal spikes produced by individual microparticles by monitoring the signal intensity at a mass-to-charge ratio (m/z) of 13 ($^{13}C^+$). Spherical polystyrene microspheres of 1 µm and 2.5 µm—to mimic MPs coming from plastic waste—have been detected using ICP-MS, thus demonstrating the potential of the technique for providing information on the mass concentration (concentration of C per volume of water), particle number density (number of particles per volume of water) and size distribution of the MPs present. Further research is required before the newly introduced method can be used routinely, or for detecting and characterising MPs of even lower sizes (hence also addressing the nanoparticles). Development of adequate sample preparation techniques for separating plastic microparticles from fragments of animal or plant origin is also required. Despite the need for further optimisation, the introduction of this novel method is considered a breakthrough as the technique has the potential to provide crucial information needed in studies on the environmental impact of MPs and their influence on human health, while demonstrating a high sample throughput.

The work is reported in *Journal of Analytical Atomic Spectometry* (doi: https://doi.org/10.1039/C9JA00379G).

ARTICLE

New mass spectrometry method for characterisation of the most challenging complex mixtures

Diana Catalina Palacio Lozano[a] and Mark P. Barrow[b,*]
[a]Department of Chemistry, University of Warwick, Coventry, CV4 7AL, United Kingdom.
https://orcid.org/0000-0001-5315-5792
[b]Department of Chemistry, University of Warwick, Coventry, CV4 7AL, United Kingdom.
E-mail: M.P.Barrow@warwick.ac.uk, https://orcid.org/0000-0002-6474-5357

Introduction

The analysis of highly complex mixtures presents significant analytical challenges. Amongst the range of such samples, petroleum is considered one of nature's most complex mixtures. As a result, there is a strong need to better understand the chemistry of heavy petroleum and this becomes a driving force to improve analytical methodologies.

The molecular characterisation of petroleum and its derivatives by mass spectrometry has become known as "petroleomics". During the course of our research within this field, we have developed an approach for the characterisation of the most challenging complex mixtures. Based upon a combination of custom experiments and an in-house data processing algorithm, we used Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) to successfully analyse a non-distillable fraction of a heavy petroleum sample, representing one of the most complex mixtures investigated to date. The approach enabled the acquisition of data at constant ultrahigh resolving power (approximately 3 million FWHM, full-width at half maximum) across the entire mass range and, in the process, we assigned 244,779 unique elemental compositions, setting a new record.

Ultrahigh resolution mass spectrometry is key

A broad view of the chemical components of heavy crude oil is key in the development of new upgrading technologies, and analytical chemistry, therefore, remains at the research forefront of petroleum upgrading. Heavy crude oils are extremely complex and most conventional analytical techniques do not offer the level of performance required to provide sufficiently detailed molecular insights into crude oil behaviour under upgrading conditions.

Ultrahigh resolution mass spectrometry, offering resolving power in the region of 10^5–10^6 FWHM and sub part-per-million mass accuracy, is perhaps the only analytical technique able to separate and assign individual compositions in complex mixtures such as crude oils. In particular, FT-ICR mass spectrometers (Figure 1) offer the highest available performance and, therefore, are the ideal analytical technique for obtaining insights into the molecular composition of crude oils. Due to the complexity of such samples, many compositions can be observed at each nominal mass, differing by small mass defects as a result of contributions from the different elements for each composition. Ultrahigh resolving power (for example, 500,000 FWHM at *m/z* 400) is often required to separate species with very small mass differences.

Figure 1. FT-ICR mass spectrometer used by the authors.

As a well-known example, two elemental compositions can differ through the presence of C_3 rather than SH_4—a difference of only 0.00337 Da! It is essential to be able to resolve this mass difference (or "mass split") in order to observe and assign sulfur-containing compounds that can poison catalysts, act as atmospheric pollutants and prove to be corrosive at elevated temperatures.

Traditional FT-ICR-based analysis can be used to assign the individual molecular species across a wide mass-to-charge ratio (*m/z* 150–1600), which includes components that are sufficiently heavy to be less amenable for gas chromatography. It is well known, however, that the ability to resolve individual compositions using FT-ICR MS decreases as the *m/z* increases; for a given *m/z* range, the

resolving power is inversely proportional to the *m/z*. Therefore, small mass differences might be not individually determined at higher mass. Unfortunately, it is also at higher *m/z* that it is possible to have a greater number of compositions per nominal mass, and so this is the region where ultrahigh resolving power is most needed.

FT-ICR instruments must also deal with a balance that determines the complexity of the data they can produce. Such mass spectrometers typically need approximately 100 ions per *m/z* (i.e per ion cloud) to successfully record a peak, but there is also an upper limit on the number of ions that can be stored within an ICR cell during an experiment: typically of an order of a few million ions, before space–charge effects become too severe. This essentially means that FT-ICR mass spectrometers can detect a maximum of tens of thousands of peaks within a single experiment, thus limiting the dynamic range.

Overcoming space–charge limitations

Those performance limitations are particularly important when heavier, low volatility fractions of heavy crude oil are analysed. The aim of our work was to analyse the "maltenes" (a fraction of petroleum soluble in *n*-heptane) of a truly non-distillable fraction (boiling point above 687°C at atmospheric equivalent temperature) of a heavy crude oil. Following experiments using different experimental set-ups and methods of sample preparation, successful acquisition of a full mass spectrum remained elusive.

The reason for this turned out to be the extraordinary complexity of the sample, basically overwhelming the instrument. In order to detect signal, it was necessary to start by isolating and detecting a very narrow *m/z* range. In this experimental set-up, a narrow *m/z* range of ions is transmitted through a quadrupole and accumulated in a collision cell, prior to transfer to the ICR cell for excitation and detection. This decreases the number of ions been detected in a single experiment, reduces the deleterious space–charge effects and increases the dynamic range. By performing this experiment, it was found that approximately 300 individual molecular compositions could be detected per nominal mass. By acquiring data using multiple, narrow windows, it is possible to "stitch" the data to generate a single mass spectrum. This method has been previously demonstrated and led to the previous record number of unique compositional assignments for a single sample: 126,264 molecular species for an "asphalt volcano" sample.[1]

Although space–charge effects are reduced and dynamic range is increased, the resolution achieved by traditional stitching methods, however, suffers from the well-known decrease in resolving power with increasing *m/z*, in similarity with traditional broadband experiments. This means such stitching experiments do not offer sufficient improvement in performance to resolve and assign compositions at the higher *m/z* region.

Additionally, traditional stitching methodologies have influenced the overall mass envelope as a consequence of different instrument parameters for the different windows. Anticipating molecular compositions of up to *m/z* 2000 present in the truly non-distillable fraction, a traditional stitching method will not have enough resolving power to individually assign molecular species with a mass difference of less than 0.0011 Da, for example, above *m/z* 1000. A fundamental modification of the stitching method was therefore necessary.

The mass resolving power of FT-ICR MS instruments can be increased by increasing the magnetic field, which means buying larger magnets and is extremely expensive, or by increasing the acquisition time of the experiment, as long as the signal can be sufficiently long-lived. To overcome the limitations of resolving power, narrow *m/z* windows were acquired but after prior planning, a target resolving power was calculated and the segments were acquired with increasing time domain data length with increasing *m/z* (i.e. the data were detected for longer periods for consecutive windows). By doing so, it was possible to ensure each *m/z* window was produced at the same resolving power (Figure 2).

The new approach also incorporates in-house software named Rhapso[2] that trims the data, determines the best position for overlapping the many segments, corrects the relative abundances of the ions in the segments (due to "edge effects" for windows when using a quadrupole for the isolation) and then automatically stitches the segments. The result

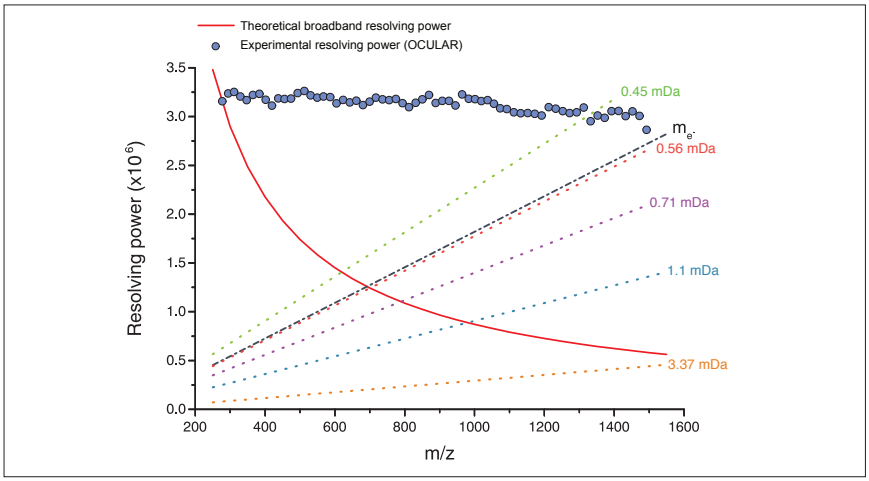


Figure 2. Red line: mass resolving power calculated for a mass spectrum acquired in broadband mode under ideal conditions, such as a perfect vacuum and no space-charge effects. The dotted lines represent the calculated minimum mass resolving powers required to resolve two peaks of comparable abundance, separated by the mass difference listed. The data points mark the mean experimental resolving power per 20 Da window. The resolving power afforded by the OCULAR method is sufficient to resolve peaks separated by a difference equivalent to only the mass of one electron (m_e, continuous black line; 0.0005485 Da) across the full mass range. Reproduced from Reference 3 under a Creative Commons Attribution licence.

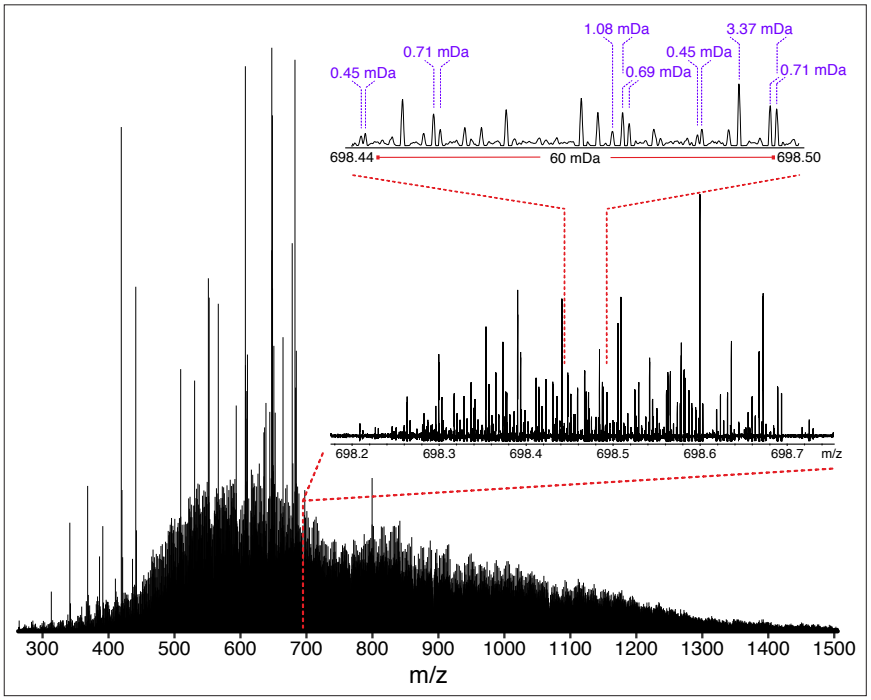


Figure 3. Stitched mass spectrum of the non-distillable fractionof a heavy crude oil, with peaks spanning *m/z* 260–1505. The spectrum was obtained by stitching 65 windows, each with an *m/z* width of 24, with steadily increasing acquisition time per segment to result in near constant resolving power across the *m/z* range. An enlarged region, showing assignments for *m/z* 698, is shown. Reproduced from Reference 3 under a Creative Commons Attribution licence.

is a single mass spectrum which was produced at near constant resolving power across the full *m/z* range. This approach has been termed "operation at constant ultrahigh resolving power" or "OCULAR". Using OCULAR, it becomes possible to acquire a mass spectrum with a constant ultrahigh resolving power, improved mass accuracy, reduced space–charge effects and increased dynamic range (Figure 3). In the published example, it was possible to assign a total of 244,779 individual molecular compositions in a truly non-distillable fraction of a heavy crude oil: the largest number of unique elemental compositions detected in a single sample to date.[3]

The approach is flexible and the performance that can be achieved using OCULAR depends on the number of windows and the target resolving power. For instance, a mass spectrum with an increased resolving power could be achieved by the acquisition of a minimum of two segments and, in this case, the mass spectra could be acquired in approximately 30 min. Heavier fractions of crude oil, however, will require higher performance and, therefore, a higher number of narrow windows in order to sufficiently boost performance. In the example of the truly non-distillable fraction, 65 windows were acquired with a mass range of 24 Da each in order to produce the complete mass spectrum.

A constant ultrahigh resolving power of approximately 3 million FWHM was achieved across the range of *m/z* 260–1505. This level of performance has enabled the resolution of species which differ in mass defect by less than the mass of an electron and has allowed the assignment of molecular compositions with sub-ppb (part-per-billion) mass accuracy. The extraordinary number of compositions spanned dozens of heteroatomic compositions, contained up to 114 carbon atoms, and represented up to 51 double bond equivalents (DBE). The high number of species with heteroatomic compositions (species containing sulfur, nitrogen and/or oxygen atoms) helped to explain the extremely low volatility of the truly non-distillable fractions of the crude oil and highlight the enormous challenges that will need to be addressed to upgrade heavier crude oils.

Further work

While the OCULAR approach has been used to successfully characterise a petroleum sample which had previously been too challenging due to its complexity, its applicability is not confined to petroleum samples. With the flexibility of the method, it is also suitable for the analysis of other complex mixtures and challenging samples, with anticipated applications in the fields of metabolomics, medical research, polymers, environmental analysis and renewable energy, amongst others. At the University of Warwick, usage and development of OCULAR is on-going and further publications utilising OCULAR will follow.

Acknowledgements

This work was supported by a Newton Fund award (reference number 275910721), Research Agreement No. 5211770 UIS-ICP, and COLCIENCIAS (project No. FP44842-039-2015).

References

1. L.C. Krajewski, R.P. Rodgers and A.G. Marshall, "126 264 assigned chemical formulas from an atmospheric pressure photoionization 9.4T Fourier transform positive ion cyclotron resonance mass spectrum", *Anal. Chem.* **89,** 11318–11324 (2017). https://doi.org/10.1021/acs.analchem.7b02004
2. R. Gavard, D.C. Palacio Lozano, A. Guzman, D. Rossell, S.E.F. Spencer and M.P. Barrow, "Rhapso: automatic stitching of mass segments from Fourier transform ion cyclotron resonance mass spectra", *Anal. Chem.* in press (2019). https://doi.org/10.1021/acs.analchem.9b03846
3. D.C. Palacio Lozano, R. Gavard, J.P. Arenas-Diaz, M.J. Thomas, D.D. Stranz, E. Mejía-Ospino, A. Guzman, S.E.F. Spencer, D. Rossell and M.P. Barrow, "Pushing the analytical limits: new insights into complex mixtures using mass spectra segments of constant ultrahigh resolving power", *Chem. Sci.* **10,** 6966–6978 (2019). https://doi.org/10.1039/C9SC02903F

ARTICLE

Is your spectrophotometer still "Pharma compliant"? A review of the new European Pharmacopoeia 10th Edition

Nathan Hulme and John Hammond
Starna Scientific Ltd Hainault Business Park, 52/54, Fowler Rd, Ilford IG6 3UT, UK

Introduction

European Pharmacopoeia (EP) Chapter 2.2.25 on ultraviolet and visible spectroscopy (or spectrophotometry) has been extensively revised in both detail and scope and the new Edition 10.0[1] (*10.0*) is mandatory from 1 January 2020. A major change is that the scope is now extended to include high-performance liquid chromatography (HPLC) detectors and process analytical technology (PAT) as applications of ultraviolet/visible (UV/vis) spectrophotometry. This is a considerable divergence from the latest US Pharmacopeia (USP) Chapter <857>[2] on Ultraviolet-Visible Spectroscopy, mandatory from 1 December 2019, that specifically *excludes* HPLC detectors from its scope. HPLC and PAT are both more dynamic and system-specific techniques than basic spectrophotometry, with more variables to consider, so, for reasons of simplicity, this article covers the new regulations only in so far as they apply to basic spectrophotometry. The new Edition introduces some new approaches to instrument qualification and suggests new reference materials for qualification measurements.

The significant changes to the standard, and their practical implications for instrument users are discussed below.

General measurement principles

This topic was largely absent from the previous Edition 9.2[3] (*9.2*) but has been extensively re-written and expanded for *10.0*. While much of this section describes well-known aspects of UV/vis measurement, there are some new specific points to note:

- Definition of UV/vis: For the purposes of the EP, the UV region is now defined as from 180 nm to 400 nm and the visible from 400 nm to 800 nm.
- The user is recommended to: "Define the measuring conditions to obtain a satisfactory signal-to-noise ratio and to select the scan range, scan rate and slit-width that provide the necessary optical resolution for the intended application without losing the required signal-to-noise ratio or the linearity of the analytical method." This is, of course, just good practice, but it is also suggested that when using diode array instruments, "there is no need to adjust the beam size, scan range, scan rate or slit-width since the optical resolution is typically fixed and the full spectrum is always recorded". What is to be done if these fixed parameters do not yield a suitable signal-to-noise or linearity is not explained.

Cells (cuvettes)

Requirements for the optical quality of cells have been revised. A path length tolerance of ±0.005 cm was specified in *9.2*. This is amended to ±0.5%, which of course equates to ±0.005 cm for a 1-cm cell but becomes problematic when applied to much shorter path length cells.

While the ±0.5% tolerance for a 10-mm path length is well within the practical capabilities of most cuvette suppliers, as the path length reduces the ±0.5% tolerance becomes impractical. This would mean the tolerance on a 1-mm path length cell would be ±5 µm, where even the most reputable suppliers only quote a tolerance of ±10 µm. While ±5 µm is possible, it would add considerably to the cost, making such cells uneconomic as a day-to-day tool. Furthermore, taken to its logical conclusion a cuvette with a path length of 10 µm would have an unmeasurable tolerance of ±0.05 µm. This puts the user in a difficult position simply because applying a simple percentage does not work in practice. There was also a requirement in *9.2* that "When filled with the same solvent, the cells intended to contain the solution to be examined and the compensation liquid must have the same transmittance". The term "the same" is not quantifiable and is now clarified in *10.0*: "Cell absorbance <0.093 A @240 nm for a quartz cell, <0.035 A @650 nm for a glass cell; and when rotated 180° in the holder, an absolute difference <0.005 A."

Control of equipment performance

The instrument qualification required for compliance is defined in *10.0* by the purpose of the analysis being carried

Table 1. Minimum tests to be carried out for the control of equipment performance (Reproduced from Table 2.2.25.-1 in Reference 1).

Purpose	Method	Wavelength accuracy	Absorbance accuracy	Photometric linearity	Stray light	Resolution/spectral bandwidth
Quantitative or limit test	Based on measurement of the absorbance at one or more identified wavelengths (e.g. assay or impurities test)	X	X	X	X	If required in the monograph
Identification test	Based on wavelength of absorption maxima and minima	X	–	–	X	–
	Based on absorption measurement and wavelength of absorption maxima	X	X	–	X	–
	Based on comparison of spectrum with that of reference substance	X	X	–	–	–

out as shown in Table 1, taken from the standard.

Table 1 implies that instrument bandwidth is not important for qualitative analysis, but it should be remembered that if the spectra being examined contain sharp or complex absorption bands, the measured wavelength and absorbance of the peaks may be dependent on the resolution of the spectrophotometer, and may appear to shift simply due to the ability of the instrument, or lack of it, to resolve adjoining spectral features. Caution should, therefore, be exercised in such cases and it may be that a resolution qualification process is to be recommended.

The previous Edition of the standard contained a simple set of tests to evaluate an instrument's performance for wavelength and absorbance accuracy, stray light and resolution. If an instrument passed these tests it could be claimed to be "pharmacopoeia compliant". This approach has the potential weakness that an instrument qualification carried out under one set of operating conditions might not be valid for an analysis carried out using different conditions. For example, a qualification carried out in the UV using a deuterium lamp as source might not describe what would happen if the actual analysis were to be performed in the visible region using a tungsten halogen source. While the new standard requires the same parameters to be qualified, the requirement is now to demonstrate that the instrument has the necessary performance to carry out the actual analysis. This has always been a general requirement of GxP protocols, but not explicitly stated until now. The user must, therefore, determine the range of parameter values over which the system will be used in the analysis and demonstrate compliance over that range. One consequence of this is that the simplistic approach often adopted in the past—one qualification test for each parameter—may not suffice. Indeed, the standard now also requires that photometric linearity be qualified; this will certainly mean that more than one reference material with accurate absorbance values will be needed. The standard also recommends that the assigned parameter values of the references used for qualification should "bracket" the values to be used in the proposed analysis, so that a laboratory conducting several different assays may need to choose a range of different reference materials to demonstrate full compliance. These may be either purchased "certified reference materials" (CRMs) such as solid filters or liquid filters in appropriate sealed cells, or "solutions prepared in the laboratory". CRMs have several advantages over laboratory-prepared solutions, and this will be discussed later.

Control of wavelength accuracy

The user is required to:

"Control the wavelength accuracy of an appropriate number of bands in the intended spectral range using one or more reference materials"

and

"It is recommended to test at least 2 wavelengths that bracket the intended spectral range"

A selection of reference materials is proposed, with peak wavelengths (see Table 2).

All the solutions and solid filters are commercially available as CRMs. Note that the spectra of the rare earth elements used in these materials contain sharp peaks, so the measured peak wavelength may vary with instrument resolution. Good wavelength CRMs will have wavelengths certified at different bandwidth values, and the user should qualify the instrument using the bandwidth specified in the analytical monograph.

Holmium oxide solution has been used as a wavelength reference for many years, but for wavelengths below 240 nm cerium oxide solution, with peaks down to 201 nm, is now recommended for this "far UV" region.

Glass filters might be considered to be more robust than liquid references in cuvettes, but wavelength intensity values can vary slightly from melt

Table 2. Examples of wavelengths used for the control of wavelength accuracy (Reproduced from Table 2.2.25.-2 in Reference 1).

Material	Peak wavelengths (nm)
Solutions:	
Cerium in sulfuric acid	201.1; 211.4; 222.6; 240.4; 253.7
Didymium in perchloric acid	511.8; 731.6; 794.2
Holmium in perchloric acid	241.1; 287.2; 361.3; 451.4; 485.2; 536.6; 640.5
Solid filters:	
Didymium glass	513.5
Holmium glass	279.3; 360.9; 453.4; 637.5
Lamps:	
Deuterium	486.0; 656.1
Mercury (low pressure)	184.9; 253.7; 312.5; 365.0; 404.7; 435.8; 546.1; 577.0; 579.1
Neon	717.4
Xenon	541.9; 688.2; 764.2

to melt so such filters should be individually certified. Solution cell filters can be cleaned (with care), as an optically polished quartz surface can be returned to a "clean" optical characteristic; however, this is not recommended for glass filters as by definition, cleaning may change the characteristics of the optical surface, and thereby invalidate the certification.

Atomic spectral lines such as those of mercury, neon or xenon are a primary physical standard and the ultimate wavelength reference and as such are always cited as suitable for instrument qualification. Caution is needed, however, as the US Pharmacopeia Chapter <857> notes: "The arc of the atomic emission source, or its image, needs to be located on the same optical path as the image of the primary light source of the spectrometer; thus, it can be used only in spectrometers that can be operated in a single-beam intensity mode and practically should be implemented only on a system designed to accommodate these sources". The built-in deuterium and xenon lamps often used as spectrophotometer light sources are on the optical path and have emission lines that can provide a useful routine wavelength check if the instrument is capable of single-beam operation. Note, however, that only visible wavelengths are referenced, so they are unsuitable for UV qualification.

The list above is not prescriptive, so if qualification is required for which none of the recommended materials is suitable, other CRMs are available and can be used. For example, for those needing qualification at even lower UV wavelengths, a "Deep UV" CRM[4] is available from a leading Reference Material Producer (RMP), with certified peaks down to 191 nm. Some simple instruments having a wide spectral bandwidth may be unable to resolve the sharp bands of the listed references, and for such cases a specially formulated "Green dye solution"[4] offered by one RMP is a CRM that can be used to qualify wavelength (and absorbance) at bandwidths up to 12 nm.

Whatever references are used, the EP's permitted tolerance for benchtop spectrophotometers is ±1 nm for wavelengths below 400 nm, and ±3 nm for 400 nm and above.

Control of absorbance accuracy

This section of *10.0* introduces several changes to traditional practice and in places is open to interpretation.

Potassium dichromate solution in acidic media has been the absorbance reference material of choice for many years and was cited in *9.2* for qualification at 235, 257, 313, 350 and 430 nm. Laboratories had the option to use commercially available CRMs and most regulated laboratories will probably already have one or more of these references. It is, however, not cited in the latest Edition, which now suggests nicotinic acid solutions. The EDQM website also states that *10.0* includes:

"introduction of nicotinic acid as an alternative to potassium dichromate (REACH Annex XIV)[5] for control of absorbance accuracy".

This implies that potassium dichromate constitutes a hazard to operators, but a detailed review of the REACH regulations,[6] shows that the risk, even if preparing potassium dichromate solutions in the laboratory, is vanishingly small at the concentrations and quantities used for instrument qualification and is non-existent when using commercially supplied CRMs in permanently sealed cells—the form in which most laboratories already hold this reference.

Furthermore, nicotinic acid cannot be regarded as an "alternative" to potassium dichromate except in certain defined situations. First, potassium dichromate can be shown to be a more universal absorbance reference, as it can be certified at five well-spaced wavelengths over a much wider wavelength range (235–430 nm) compared to just two wavelengths for nicotinic acid, 213 nm and 261 nm. There is, therefore, more scope to "bracket" the analytical wavelength as recommended in the standard. Second, and perhaps more important, the nicotinic acid spectrum is significantly affected by spectral bandwidth. Figure 1 shows the effect of bandwidth on the measured values of nicotinic acid solutions and of potassium dichromate solutions at different bandwidth settings.

It can be seen that the absorbance value of the nicotinic acid peak at 261 nm, recommended here for instrument qualification, is severely affected by bandwidth—indeed the effect is much greater than the tolerance allowed for compliance. The values for potassium dichromate at similar wavelengths are affected much less.

It is important, therefore, that qualification measurements using nicotinic acid are made at the same bandwidth setting as those used to establish the values for the reference material. *10.0*

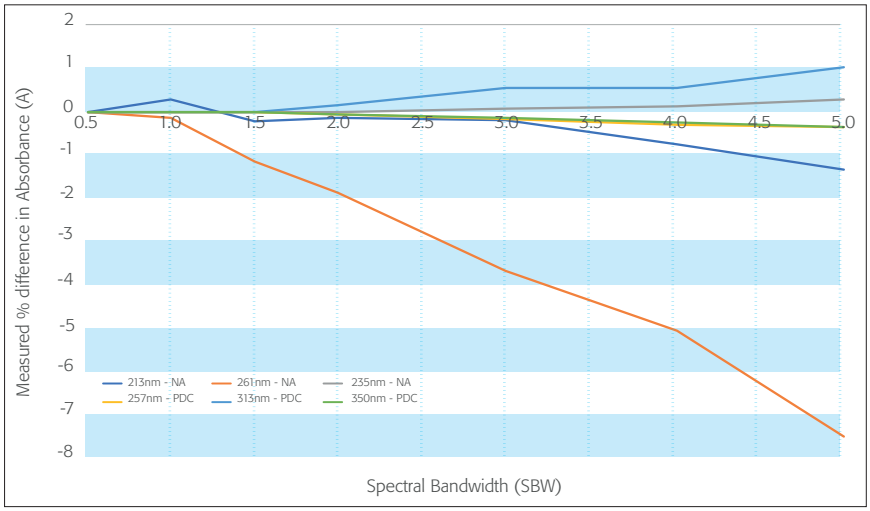


Figure 1. Effect of spectral bandwidth on measured absorbance values of nicotinic acid and potassium dichromate solutions. Nicotinic acid (NA) @ 213 nm and 261 nm vs acidic potassium dichromate (PDC) solution @ 235 nm, 257 nm, 313 nm and 350 nm.

gives a procedure for the preparation of reference solutions from *"nicotinic acid for equipment qualification CRS"*. This material is available as a solid from EDQM. Having prepared the solutions as directed, the user then calculates the reference absorbance values from the "specific absorbance" given in the accompanying certificate. Unfortunately, the variation allowed in the weight of solid to be used will lead to an inexact concentration of the final solution and hence an incorrect calculated absorbance. Furthermore, the certificate gives no indication of the bandwidth used to determine the specific absorbance, so the certified value is fairly meaningless. An instrument could fail to achieve compliance simply because the qualification measurements were unknowingly made using a spectral bandwidth different from that used to determine the certificate value. No guidance is given on the stability or validity period of the solutions once prepared. Use of this material, as described in *10.0*, is, therefore, unlikely to be valid as an absorbance reference. Fortunately, commercial nicotinic acid CRMs are available and can usually be certified at any bandwidth requested by the customer. Used correctly, nicotinic acid is a useful absorbance reference in the far UV but cannot replace potassium dichromate at higher wavelengths.

For compliance, the allowed difference between the measured absorbance and the actual absorbance of the reference material is ±0.010 A or ±1 %, whichever is greater, and "values at approximately the two limits of the expected absorbance range should be verified". This tolerance applies to absorbance values up to 2A, and it is suggested that higher absorbances are dealt with "on the basis of a risk assessment", for which no further details are provided. In this context, both nicotinic acid and potassium dichromate CRMs are available with traceable certified values up to 2.5A and 3.5A, respectively, so direct qualification can be carried out with confidence at these higher levels (Figures 2 and 3).

Control of photometric accuracy and/or linearity in the visible region can be achieved using solid glass filter CRMs, but unlike the previous version (*9.2*) no specific guidance is given with respect to standards for the visible region other than to say that "suitable solid or liquid filters" can be used. The comments made above for wavelength also apply here, so CRMs other than those suggested may be used if they better match the operating conditions used for analysis.

Control of photometric linearity

This is a new requirement in *10.0*. The references used to qualify absorbance accuracy can be used to qualify linearity provided they are compatible with the analytical wavelength and absorbance ranges. Nicotinic acid is cited as an example over the range 5-40 mg L^{-1}. The number of references to be measured over the required absorbance range is not stated, but the coefficient of determination (R^2) is given as 0.999 for compliance. How this requirement is met is left for the laboratory to decide. Fortunately, there is a definitive, internationally recognised ISO standard, ISO 11095,[7] "Linearity Calibration using Reference Materials", which states that the number of references used to assess a calibration function should be at least three. Similarly, the latest USP Chapter <857> simply states that at least three references bracketing the required absorbance range should meet the required absorbance accuracy criteria. Three will probably suffice for a limited absorbance range, say up to 1A, but users may decide to use more when using higher absorbances. When using CRMs, users should remember to compare measured values with certified values and not with concentrations when assessing linearity.

Control of stray light

The standard says: "Stray light is determined at an appropriate wavelength using suitable solid or liquid filters or solutions prepared in-house". The previous Edition (*9.2*) named just one stray light reference, namely 12 g L^{-1} potassium chloride solution, a cut-off filter that indicated stray light at 198 nm. Now, four different aqueous solutions are identified that can allow stray light to be detected over a wavelength range from 198 nm to 370 nm (Table 3).

The test is to be conducted using a water blank cell, and it is observed that "the instrument parameters used for the test, such as slit-width and type of light source (e.g. deuterium or tungsten lamp), must be the same as those intended for the actual measurements". All these reference materials are available as CRMs.

Control of resolution (spectral bandwidth)

This test remains the same as in the previous Edition. Where prescribed in a

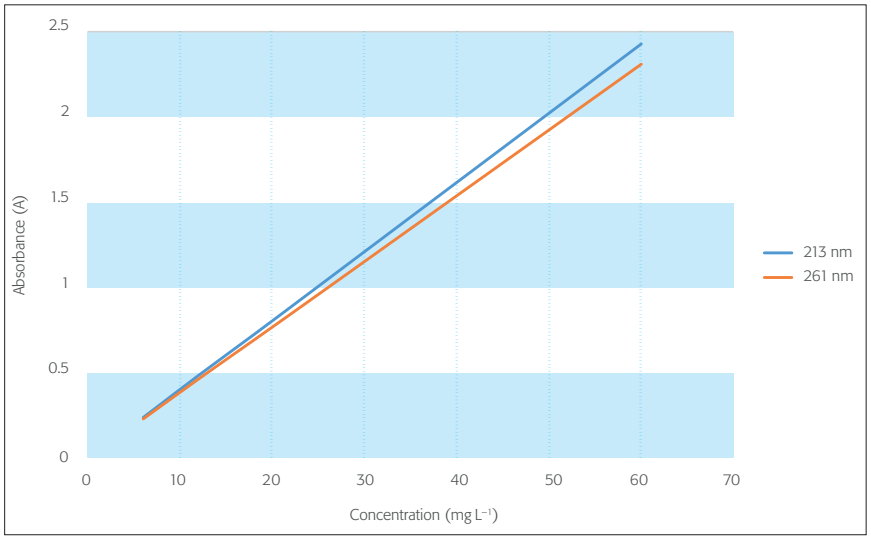


Figure 2. Nicotinic acid linearity.

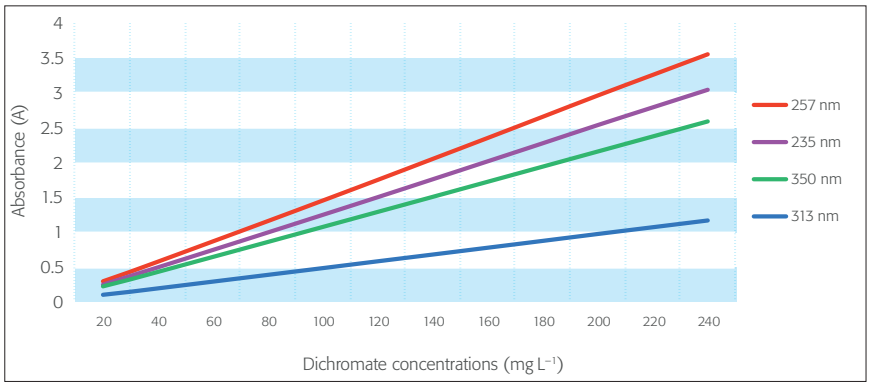


Figure 3. Potassium dichromate linearity.

Table 3

Material	Concentration	Absorbance at wavelength
Potassium chloride	12 g L^{-1}	≮2.0 A at 198 nm
Sodium Iodide	10 g L^{-1}	≮3.0 A at 220 nm
Potassium iodide	10 g L^{-1}	≮3.0 A at 250 nm
Sodium nitrite	50 g L^{-1}	≮3.0 A at 340 nm ≮3.0 A at 370 nm

monograph, the resolution of the instrument can be determined by recording the spectrum of 0.02% v/v toluene in hexane (or heptane), which produces a spectrum with an absorbance maximum at 269 nm and a minimum at 267 nm. The ratio of the maximum at 269 nm to the minimum at 267 nm should be as stated in the monograph. For general guidance, however, Figure 4 shows typical spectra obtained at different bandwidths—a useful guide to instrument bandwidth is shown in Table 4.

Heptane, with lower toxicity than hexane, is proposed as an alternative solvent. This is not an issue, however, if the test material is purchased as a sealed-cell CRM.

The resolution test recommended for derivative spectroscopy is no longer included in the standard.

System suitability

This new section states that:

"System suitability tests may be required prior to sample measurement to verify critical parameters which may have an impact on the result. These tests may cover wavelength accuracy, absorbance accuracy, stray light and photometric linearity. System functionality tests, for example those performed as part of equipment auto testing, may be considered part of the system suitability tests."

Several spectrophotometer models incorporate some degree of automatic self-test facility. A typical example is to use the source lamp (deuterium or xenon) emission lines to provide a wavelength check. As indicated above, however, these checks are only in the visible region, and users will have to decide whether such tests can "verify critical parameters" to the degree required. If not, the implication is that some or all the qualification tests previously described may also need to be performed along with the analysis.

Reference materials: CRM or prepared in-house?

Until the 1970s most laboratories used in-house prepared solutions or proprietary test materials to check the performance of their instrumentation or relied on the manufacturer to calibrate their instruments as part of routine maintenance. Now, the international nature of regulation requires that calibrations must have international validity, which means using universally recognised standards for calibration purposes. CRMs, prepared by accredited suppliers according to international norms, have that validity. It is still perfectly possible for instrument users to prepare their own reference solutions, and instructions are given in this standard, but compared with the use of CRMs this can be a complex process with many pitfalls. Clearly the accuracy of the reference value will depend on the purity of the materials used and the accuracy of preparation processes such as weighing and dilution. It is, therefore, normal to establish an "uncertainty budget" for the

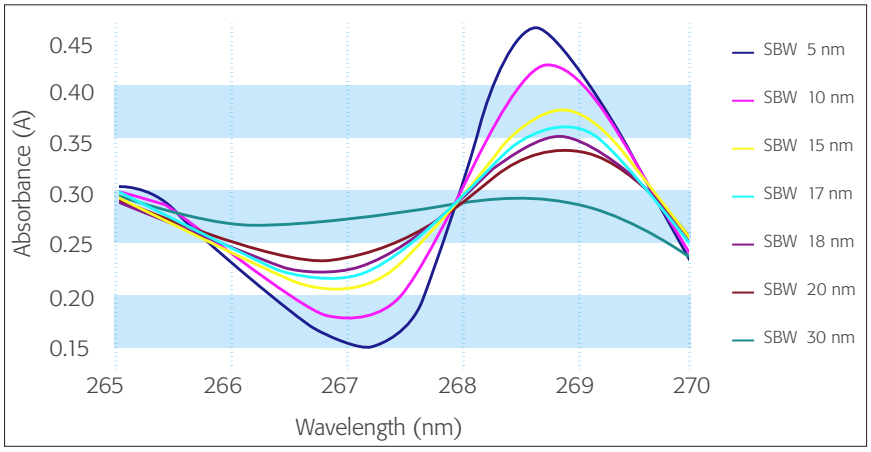


Figure 4. Spectra of 0.02% v/v toluene in hexane at different spectral bandwidths.

Table 4

Ratio	2.4:2.5	2.0:2.1	1.6:1.7	1.3:1.4	1.0:1.1
Spectral bandwidth (nm)	0.5	1.0	1.5	2.0	3.0

preparation of the standard and hence the overall uncertainty in the reference value, but this can also be complicated.[8] It is perhaps not surprising that most laboratories decide to use commercial CRMs, where all this has already been done and the uncertainty is stated on the certificate.

What is a CRM?

As defined by ISO/REMCO (the International Standards Organisation Committee on Reference Materials), a CRM is a "Reference Material, characterised by a metrologically valid procedure for one or more specified properties, accompanied by a reference material certificate that provides the value of the specified property, its associated *uncertainty*, and a statement of metrological *traceability*."[9]

Originally, the only available references for spectrophotometer calibration with internationally accepted property values were those from National Metrology Institutes (NMIs) such as the National Institute of Standards and Technology (NIST) in the USA, whose products were trademarked as Standard Reference Materials (SRMs). In any case, the advent of Good Laboratory Practice and similar quality schemes led to an increase in the demand for SRMs that exceeded the production capacity production capacity of the NMIs. Commercially produced reference materials were available but not necessarily accepted by regulatory authorities, so some producers collaborated with the regulators to develop reference materials that would be recognised as equivalent to SRMs for calibration purposes. Such materials would be known as CRMs and would be recognised by national and international regulators or accreditation bodies. These CRMs can be produced as solutions, supplied permanently sealed into UV quality cells for direct qualification measurements. Not only does this free the user from the task of preparing the reference solutions, but virtually eliminates any hazards that might arise from directly handling the reference materials.

Furthermore, unlike in-house reference materials, the certified value of a CRM does not rely on the accuracy with which the reference material has been prepared, but on a calibration performed on a reference instrument that has itself been calibrated against primary physical standards or SRMs. The certificate values are of course subject to any variability of the calibration instrument, but this can be established by the producer and stated on the certificate that accompanies the CRM. The "expanded uncertainty budget" normally given in the calibration certificate is the uncertainty to be expected in the measured parameter and is conventionally stated with a 95% confidence level.

Armed with this information, instrument qualification becomes very straightforward. When a CRM is used to qualify an instrument, the total allowed tolerance is the sum of the certificate uncertainty and the instrument manufacturer's specified accuracy of the instrument, Table 5.

If the difference between the measured value and the certified value is less than the total tolerance, the instrument can be judged to be operating correctly. The difference should, of course, also be less than the error permitted by the pharmacopoeia or the analytical monograph in use.

Nowadays, most instrument qualification in the pharmaceutical industry is performed using CRMs. Indeed, the United States Pharmacopeia states in its Chapter <857> that "Wherever possible… CRMs are to be used in preference to laboratory-prepared solutions". Sets of CRMs are available tailored to the new regulations, an added convenience of this approach.

Traceability is very important as it lends to the CRM the authority of the internationally recognised references to which its calibration can ultimately be traced. It is defined in ISO/IEC Guide 99:2007[10] as the "property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contrib-

Table 5

	Wavelength	Absorbance
Certificate uncertainty budget	±0.10 nm	±0.0049 A
Instrument specification	±0.30 nm	±0.0050 A
Total tolerance	±0.40 nm	±0.0099 A

uting to the measurement uncertainty". The reference spectrophotometers used by CRM suppliers to establish the certified values must, therefore, be qualified against suitable SRMs or against primary physical references such as elemental emission lines. The references used should be identified on the certificates accompanying the CRM.

The stability of the reference material is also very important, and the validity of the calibration should be stated on the CRM certificate. This is typically two years, but may be less depending on the laboratory's quality protocols. Recertification should be performed periodically to maintain the validity of the certification.

For users to have confidence in purchased CRMs, their suppliers should be properly accredited to ISO 17034:2016 "General requirements for the competence of reference material producers".[11] This is the minimum requirement and covers quality and administration systems and technical and manufacturing operations. This standard includes normative references to another standard: ISO/IEC 17025:2017 "General requirements for the competence of testing and calibration laboratories".[12] ISO 17025 specifies the procedures for reporting and evaluating measurement uncertainty and any competent producer should be accredited to this standard also. ISO 17025 accreditation includes a statement of its "scope", listing the reference materials the laboratory is competent to calibrate. Intending purchasers should check that their proposed supplier's accreditation scope includes the material in question: accreditation to ISO 17025 could be claimed on the strength of just one material or calibration process, which might not cover the item to be purchased.

Conclusions

Like the new USP Chapter <857>, Edition *10.0* of EP 2.2.25 has been considerably expanded to put more emphasis on the "fitness for purpose" of UV/vis instrumentation. Instruments must now be shown to have the necessary performance to function adequately under the operating parameters to be used for analysis. To this end, examples of suitable reference materials are given, but the suggested materials will not cover all situations. There are also uncertainties in the interpretation of the standard, notably in the sections dealing with absorbance accuracy and linearity. Nicotinic acid is suggested as an absorbance reference, but the data given for its preparation is flawed as it is inexact and does not acknowledge the effect of spectral bandwidth. One of the new specifications (cell path length) is unachievable in many instances in practice. Fortunately, however, the standard does allow the use of the very wide range of CRMs now commercially available for instrument qualification. Judicious choice of these materials will sometimes provide a better alignment to the analytical method in use than the references cited in the standard and thus better demonstrate "fitness for purpose", providing a more straightforward and convenient route to achieving compliance.

References

1. European Directorate for the Quality of Medicines & HealthCare (EDQM), "2.2.25 Absorption Spectrophotometry, Ultraviolet and Visible", *European Pharmacopoeia (Ph. Eur.)* **Edition 10.0,** 42 (2020). https://www.edqm.eu/en/european_pharmacopoeia_10th_edition
2. United States Pharmacopeia, "<857> ultraviolet-visible spectroscopy", *USP 42-NF37 Second Supplement (2S)* (2019). https://www.uspnf.com/
3. European Directorate for the Quality of Medicines & HealthCare (EDQM), "2.2.25 Absorption Spectrophotometry, Ultraviolet and Visible", *European Pharmacopoeia (Ph. Eur.)* **Edition 9.2** (2008).
4. Starna Scientific Ltd Hainault Business Park, 52/54, Fowler Rd, Ilford IG6 3UT, UK. https://www.starna.com/reference-materials/uv-visible-spectroscopy
5. Regulation (EC) No. 1907/2006 of The European Parliament and of the Council, 18 December 2006 on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) (2006).
6. The use of Potassium Dichromate in the European Pharmacopoeia; Starna Position Paper. Technical Communication, June 2018
7. International Organization for Standardization, "Linearity calibration using reference materials", *ISO 11095* (1996). https://www.iso.org/standard/1060.html
8. C. Burgess, "Measurement uncertainty without the math", *Pharmaceut. Technol.* **Feb** (2016).
9. International Organization for Standardization, "Reference materials—selected terms and definitions", *ISO Guide 30:2015*. https://www.iso.org/standard/46209.html
10. International Organization for Standardization, "International vocabulary of metrology—basic and general concepts and associated terms (VIM)", *ISO/IEC Guide 99:2007*. https://www.iso.org/standard/45324.html
11. International Organization for Standardization, "General requirements for the competence of reference material producers", *ISO 17034* (2016). https://www.iso.org/standard/29357.html
12. International Organization for Standardization, "General requirements for the competence of testing and calibration laboratories", *ISO/IEC 17025* (2017). https://www.iso.org/standard/66912.html

TONY DAVIES COLUMN

Pre-processing spectroscopic data: for good or ill?

Antony N. Davies[a,b] Jan Gerretzen[a] and Henk-Jan van Manen[a]
[a]Expert Capability Group – Measurement and Analytical Science, Nouryon Chemicals BV, Deventer, the Netherlands
[b]SERC, Sustainable Environment Research Centre, Faculty of Computing, Engineering and Science, University of South Wales, UK

At a recent international conference, I attended a good lecture by a scientist using Ion Mobility Spectrometry (IMS) in a food analysis application. During the talk, one slide mentioned that they had used Savitzky–Golay smoothing on the IMS data and that started me wondering. I asked why they had decided that they needed to smooth the IMS data and was told that as they did it routinely for infrared spectra they just applied it to the IMS data as well.

I thought a better approach might have been to decide what data processing was really required and be able to justify the additional data manipulation steps in terms of improving on an analytical figure of merit, for example. You really need to start by accepting that the spectroscopic data you have just measured isn't fit-for-purpose. Now measuring data of insufficient quality for the role it must play can have as many good (for "good" read unavoidable) reasons as bad.

Why is my raw data not fit-for-purpose?

One common reason is that you do not have enough sample. This may be unavoidable if there simply isn't more available, but can also arise by failure to prepare enough during sub-sampling. Surprisingly often it is worth going back to the source of the sample and simply asking if you can have a specific amount required to carry out your analysis. This can sometimes lead to 5 kg sacks of material requiring disposal at the end of the work, but remember in many settings the people carrying out the sampling normally work in tonnes not in milligrams. Lack of sample amount can also make the answer to the analytical question less reliable if you do not have enough to carry out a number of full-method replicates of the analysis to deliver a good estimate of the error in your result. For a fuller discussion on sampling and errors, see the Sampling Column in this issue.

Another can arise by not paying enough attention to the resolution settings on the spectrometer or method being run on the instrument. Be aware of the settings on instruments which are automatically averaging several scans for each data point they are recording as well as the actual number of data points being recorded across the width of the narrowest peak in the spectrum. Depending on the type of spectrometer being used, taking a setting which records too high a resolution can mean the scan time for each spectrum becomes long if a reasonable signal-to-noise ratio is required. This can also cause issues if the spectrometer is liable to drift, meaning there is not an infinite amount of time available for each of the independent measurements.

For hyphenated methods, such as gas chromatography/ion mobility spectrometry (GC/IMS) data which triggered this article, this resolution consideration will also include the time axis for the sample separation step (Figure 1).

With the introduction of the much more rapid ultra-high-performance liquid chromatography (UPLC® or UHPLC) systems, much effort was spent in increasing the speed at which the attached spectrometers were capable of scanning. This was so that sufficient data points could be obtained to properly define each peak, since analytes were eluting off the columns an order

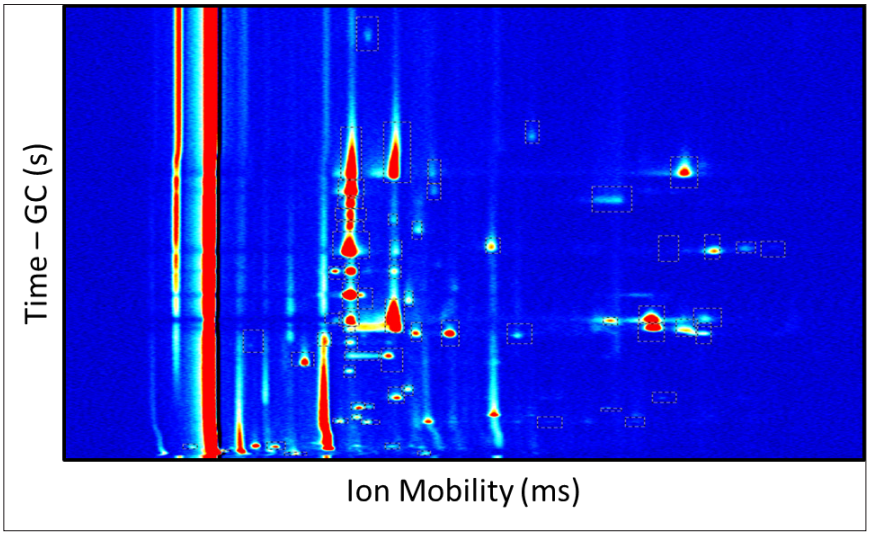


Figure 1. A somewhat typical GC/IMS analytical run showing relatively complex peak shapes compared to infrared spectroscopy.

of magnitude faster, delivering much narrower, more intense peaks.

It is often the case that a system being studied is changing as it is being measured and this dynamic change is what you are studying. Clearly the time available for each independent measurement is constrained by the rate at which the system is changing, so it may not be possible to acquire many scans for each time point in order to achieve excellent signal-to-noise ratios.

A review by Engels and co-workers sums up some of the issues which lead to a demand for spectroscopic data pre-processing to remove unwanted artefacts in data sets under the headings of noise, baseline offset and slope light scatter temporal and spectral misalignment, normalisation, scaling and element-wise transformations, supervised pre-processing methods and finally artefacts in hyphenated techniques.[1] This is an excellent starting point if you wish to go deeper into the subject than this column's space allows. The authors acknowledge how extremely difficult it can be to determine which method or pre-processing methods can successfully be applied. It is important to take into account the specific data set characteristics emphasising that the identification of which artefacts are present among which properties of the spectroscopic data is of considerable importance that cannot be ignored in this choice of pre-processing strategies.

Approaches to spectroscopic data pre-processing: or "my boss told me to do it" syndrome

In some laboratories there are preferences for carrying out certain types of pre-processing as standard, and this includes standard ordering of the pre-processing steps. These have often been handed down over the years and the original reasons for these workflows are no longer known by the current laboratory staff.

Jan Gerretzen and co-workers at the University of Nijmegen working under the Dutch COAST initiative carried out some work to try and eliminate the "black magic" around the selection of the data pre-processing steps and the order in which they should be carried out. They adopted a systematic Design of Experiments approach to varying baseline, scatter, smoothing and scaling pre-processing steps for reference data sets in Latex monitoring (quantifying butyl acrylate and styrene) as well as corn data sets for their moisture content.[2] In a separate report the approach was tested on data from a near infrared (NIR) spectrometer monitoring NaOH, NaOCl and Na_2CO_3 concentrations in a waste treatment system of a chlorine gas (Cl_2) production facility. The gaseous waste effluent of this facility contains chlorine, which is removed by a caustic scrubber where the waste gases are led through a solution containing NaOH.[3]

Selection of pre-processing strategies

Quite often text books or spectroscopic data processing packages will describe the effect of individual pre-processing algorithms. However, there is little support around the consequences of applying multiple pre-processing steps during data analysis. Even the order that the pre-processing steps are applied can have a drastic effect on the quality of the analysis, let alone how the parameterisation of each step impacts subsequent steps or the final result.

Table 1 shows an experimental design used in this approach. A full factorial design was selected to evaluate the influence of each pre-processing step. The response variable measuring the model improvements from the pre-processing steps was the root-mean-square error of prediction figures.

Figure 2 shows how close the rapid Design of Experiments approach came to determining the best sequence and parameterisation of various pre-processing strategies, compared to identifying the absolute best strategy determined by Brute Force number crunching of every possible variable (over 5000 solutions required to be calculated).

Most authors highlight the fact that their work can really only be deemed applicable to the types of data and particular types of samples they are analysing. In Reference 1, the application of variable selection and data pre-processing were only observed to improve the model performance when they were carried out simultaneously[2] and the conclusion was that although the specific "best-case" data pre-processing solutions were found, the more general applicability of this work was in defining a successful generic approach to scientifically decide on the best spectroscopic data pre-processing methodology to use.

Peter Lasch looked at spectral pre-processing for infrared and Raman spectroscopic techniques used in the field of biomedical vibrational spectroscopy and microspectroscopic imaging.[4] Here

Table 1. Data preprocessing Design of Experiments derived from Reference 1.

Experiment	Baseline	Scatter	Smoothing	Scaling
1	Yes	Yes	Yes	Yes
2	Yes	Yes	Yes	No
3	Yes	Yes	No	Yes
4	Yes	Yes	No	No
5	Yes	No	Yes	Yes
6	Yes	No	Yes	No
7	Yes	No	No	Yes
8	Yes	No	No	No
9	No	Yes	Yes	Yes
10	No	Yes	Yes	No
11	No	Yes	No	Yes
12	No	Yes	No	No
13	No	No	Yes	Yes
14	No	No	Yes	No
15	No	No	No	Yes
16	No	No	No	No

TONY DAVIES COLUMN

Data smoothing

Often used to reduce random noise where further data accumulation is not possible. Depending on the data set data smoothing can damage the data set leading to distorted picture of the results. Some typical data smoothing methods include Moving Average across a number of data points, the number of points averaged is adjustable and Savitsky–Golay smoothing which fits a polynomial to segment of the data set. In Savitsky–Golay smoothing the order of the polynomial can be changed (first-order = Moving Average) as well as the range of data to be fitted.

ATR correction

Correction of mid-IR spectra sampled using the Attenuated Total Reflectance (ATR) technique for the penetration depth dependence related to the frequency in the spectrum. It does not attempt to correct for the refractive index differences between the sample and the crystal that can lead to "derivative-like" spectra.

Multiplicative Scatter Correction (MSC)

Rinnan and co-workers took a critical look at a range of pre-processing methods in NIR spectroscopy chemometric modelling including a group of scatter-corrective pre-processing methods includes Multiplicative Scatter Correction using a reference data sets. They also looked at how different pre-processing methodologies impacted on the quality of prediction results for six different spectrometers using filter, dispersive and Fourier transform technologies. In whichever combination they applied pre-processing they could only achieve at best a 25% improvement in the prediction error—and the concluded with a warning about the risks associated with incorrectly setting the parameters for the window size or smoothing functions.[5]

Derivative filters

Quite a popular pre-processing strategy to enhance the resolution of complex spectra assisting in identifying overlapping peaks and also assists in minimising the influence of baseline effects. For instruments that acquire signals in the time domain such as Fourier transform infrared spectrometers several techniques exist to apply filters to enhance resolution and reduce noise in the time domain before the data is transformed to the frequency domain.

techniques including cleaning the datasets (outlier detection), normalisation, filtering, detrending, transformations like ATR correction and "feature" selection are discussed. The article contains some interesting explanatory graphics and longer discussions on water vapour correction, different strategies for normalisation, baseline correction and data filtering for noise removal or spectral resolution enhancement (use of derivative filters). Raman-specific spectroscopic data pre-processing is also addressed, covering topics such as the removal of cosmic ray artefacts and fluorescence background signals. The author acknowledges that a combination of pre-processing steps is usually required to obtain the best results and bemoans the sparsity of systematic investigations in which the effectiveness of different ways of applying pre-processing workflows to the specific needs of subsequent quantitative or classification analytical procedures is investigated. The author acknowledges that it is one of the main data analysis tasks to adapt and optimise these workflows, but this is still more an art rather than a science!

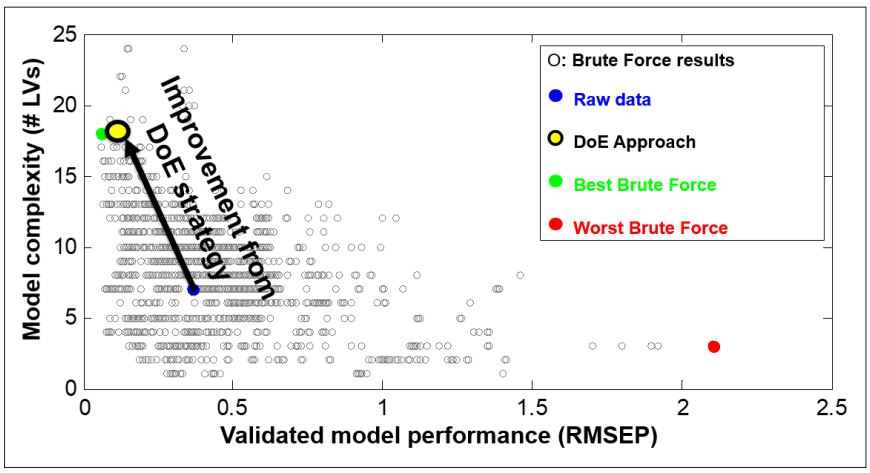


Figure 2. Successful application of a Design of Experiments approach to spectroscopic data pre-processing for model optimisation (data taken from the work reported in Reference 3).

Conclusion

I think it is clear that we are often constrained from measuring the ideal spectra for our tasks and that data pre-processing can eliminate or mitigate some of the problems arising from having to handle sub-optimal measurements. However, it is also clear that these pre-processing steps need to be carried out with our eyes wide open and after giving the problem some thought. The computing power now commonly available allows us to also use the Design of Experiments approach to find the best pre-processing strategy for our specific data sets—and that this pre-processing strategy needs to be re-assessed for each individual problem and not blindly copied across from one spectroscopic field to another.

References

1. J. Engel, J. Gerretzen, E. Szymańska, J.J. Jansen, G. Downey, L. Blanchet and L.M.C. Buydens, *Trends Anal. Chem.* **50,** 96–106 (2013). https://doi.org/10.1016/j.trac.2013.04.015
2. J. Gerretzen, E. Szymańska, J. Bart, A.N. Davies, H.-J. van Manen, E.R. van den Heuvel, J.J. Jansen and L.M.C. Buydens, *Anal. Chim. Acta* **938,** 44e52 (2016). https://doi.org/10.1016/j.aca.2016.08.022
3. J. Gerretzen, E. Szyman´ska, J.J. Jansen, J. Bart, H.-J. van Manen, E.R. van den Heuvel and L.M.C. Buydens, *Anal. Chem.* **87,** 12096−12103 (2015). https://doi.org/10.1021/acs.analchem.5b02832
4. P. Lasch, *Chemometr. Intell. Lab. Syst.* **117,** 100–114 (2012) https://doi.org/10.1016/j.chemolab.2012.03.011
5. Å. Rinnan, F. van den Berg and S.B. Engelsen, *Trends Anal. Chem.* **28(10),** 1201–1222 (2009). https://doi.org/10.1016/j.trac.2009.07.007

SAMPLING COLUMN

Sampling for spectroscopic analysis: consequences for multivariate calibration

Kim H. Esbensen[a] and Brad Swarbrick[b]
[a]Consultant, independent researcher, Dr (h.c.), www.kheconsult.com. E-mail: khe.consult@gmail.com
[b]QBD Consultancy, www.qbd-consultancy.com

An alternative title for this column could have been: "What's in an analytical datum?" Analytical results, and with them multivariate chemometric models, cannot be validated in isolation; analytical results have a pedigree, a history, which influences the quality of determination just as much as the technicalities of the analytical method itself—in fact, often more so. The relevant issue is representativity **with respect to the primary target material, the lot. Analytical aliquots, or direct analytical signals, are the end-products of a compound "lot-to-analysis" pathway in which** all **preceding sampling/signal acquisition operations must be representative in order for analytical results (data in the chemometric world) to be valid renditions of the original lot material. The incurred sampling, sub-sampling, sample preparation and sample presentation errors, collectively constituting the Total Sampling Error (TSE_{TOT}), are typically 10–25+ times larger than the spectroscopic measurement errors alone (TAE_{SPEC}), i.e. TSE_{TOT} dominates the total measurement uncertainty budget. Focussing on analysis alone (TAE_{SPEC}) is, therefore, a breach of due diligence when seen from the point of view of the** user **of analytical results, which forms the basis for critical decision making in science, technology and industry. This column surveys the proper context for all critical steps** before **spectroscopic analysis and their impact on multivariate modelling of spectroscopic signals, irrespective of whether the TSE_{TOT} contributions are large, intermediate or small. All cases must be treated identically, including sensor-based solutions from the Process Analytical Technology (PAT) realm.**

Introduction

The key issue of "sampling" is material and lot *heterogeneity* and how to **counteract** its adverse influence on sampling/signal acquisition, sub-sampling and sample preparation/presentation processes, all of which demonstrably take place **before** analysis. The Theory of Sampling (TOS) is the guiding framework for meta-analysis of all spectroscopic modalities. The TOS emphasises the Fundamental Sampling Principle (FSP), which states that all potential units from an original material must have an equal probability of being sampled in practice, and that samples are not altered in any way after sampling. Units can be particles, particle fragments or collections-of-units making up the practical sampling unit, termed *increments*. In the realm of quantitative spectroscopic analysis, compliance with the FSP is rather often a hidden elephant in the room; far from always properly acknowledged. In this light, many potential pitfalls exist regarding analysis in the lab as well as in Process Analytical Technology (PAT) applications, which **must** be avoided, lest unnecessary Total Sampling Error (TSE_{TOT}) will be produced. These errors will uncontrollably inflate the total Sampling-and-Measurement Uncertainty (SMU).

The present column focuses on the adverse influences that may crop up in the chemometric data modelling "on the other side" of production of analytical results, if the basic representativity demands from TOS are not heeded. The TOS needs only minimal presentation is this column.

Theory of Sampling, TOS

The FSP is the first of six Governing Principles (GP) and four Sampling Unit Operations (SUO), which must be honoured in order to guarantee sampling and analysis representativity. In previous *Spectroscopy Europe* columns, and within the chemometric and spectroscopic communities (the NIR realm in particular), the TOS has been presented extensively to any depth desired, as a unified, systematic framework for all principles and practical operations needed before analysis. While it is often argued that the analyst is only responsible for TAE_{SPEC}, someone else must then be responsible for controlling TSE_{TOT}. This is a most unfortunate division, however, that positively invites a serious sin-of-omission: who is really in charge of guaranteeing representativity of the analytical result, if/when most of the uncertainty is incurred outside the complacent four walls of the analytical laboratory? We here argue that it is better to view the "lot-to-analysis" pathway as a unified whole, as a common responsibility, which **includes** the quantitative analyst (of any spectral modality) as well as the data analyst, whether of chemometric or statistical inclination.

This column is a reasoned call for a *holistic* view of sampling, analysis and data modelling as an integrated whole. The relevant literature is numerous, and presents the minimum TOS competence

SAMPLING COLUMN

necessary to be scientifically, technically and managerially responsible for guaranteeing relevant, meaningful and reliable analytical results—and for multivariate calibration models in all contexts, under all circumstances and for all types of materials that can be subjected to spectroscopic analysis. The comprehensive core reference list is all one needs.[1–8]

Reference 1 is a treatise answering the question: "WHY we need the Theory of Sampling (TOS)". Reference 2 is the most recent presentation of the TOS, and Reference 3 is of particular relevance for quantitative spectroscopic analysis. Reference 4 is the *de facto* international standard on the general principles for representative sampling, often accompanied by Reference 5 for full coverage of the relationship between the TOS and Measurement Uncertainty (MU). Reference 6 introduces readers to the key process technology interplay between the TOS and PAT. Reference 7 is the most recent chemometric textbook, in which the critical pre-sampling realm is fully integrated, including its bearings on proper model validation (issues not covered by any other chemometric textbook). Reference 8 is a feature on the first order issues related to the application of on-line NIR to predict pharmaceutical powder composition, as an example of a study following the holistic call. This column takes up this scope and will in particular deal with the consequences of TOS non-compliance for the chemometric community, for which "data" are usually considered sacred entities—in the sense that nobody cares much about the pre-analytical realm: "Chemometricians analyse and model the data—basta!"

Only two TOS elements are needed for the present purpose:

Fundamental Sampling Principle (FSP): "TOS—the missing link in PAT",[6] amplified by Reference 8, explained the difference between physical extraction of **samples** (representative) or **specimens** (non-representative) relative to TOS-compliant spectral interaction with a stream of matter (representative "PAT process sampling"), and how this difference results from failure to comply with the FSP as applied to flowing streams of matter (all explained more fully below).

Sampling Bias: Failure to eliminate the complement of Incorrect Sampling Errors (ISE), wholly or partly, will unavoidably lead to a sampling bias, of unknown magnitude, which cannot be corrected for as presented in many of this column's References. It will appear that within the PAT approach there are several major pitfalls if/when the pertinent TOS principles are not heeded (or are perhaps unknown).

In medias res

For the purpose of chemometric multivariate calibration/validation/prediction, we are at first interested in the relationship between:

- "From-lot-to-aliquot" (sampling + analysis, i.e. traditional physical sampling), and
- "From-lot-to-spectrum" (sampling via *in-* or *on-line* application of spectral analysis, PAT)

The traditional domain sampling + analysis needs only little comment. Seen from the point of view of the professional analytical laboratory, "samples" arrive in the lab, which is hired to produce the requested analytical results. Preferably representative samples, of course, but it is no mystery why many professional laboratories declare that the relevance, validity, quality and representativity status of primary samples is solely the responsibility of the client who supplies them, as this conveniently saves the day w.r.t. decisions made on the basis of the analytical results produced. Not surprisingly, we see in many cases that results supplied by Quality Control (QC) laboratories are understood as the absolute truth. And apparently with good reason, the analysis **is** representative of the sample delivered, but all issues about whether this means that it is representative of the entire lot from which it was extracted have disappeared. The demarcation between QC and production is never more pronounced than in this situation.

But it **is** also fair to say that such critical pre-analysis issues have begun to appear on the agenda, at least for some laboratories: "We know about the potential for gross sampling errors, that may very well jeopardise the objective of the client. Shall we tell him, or not?" There is (very) much more to discuss concerning the complex relationship between client, in this case production, and laboratory, and this was recently subject to an extensive analysis earlier in the Sampling Column.[9,10]

The in-, on-line realm (analysing *while* sampling) is of particular interest to the current column. Historically there has been a trend within PAC (Process Analytical Chemistry) and PAT to consider installing PAT *sensors* into a pipeline as synonymous with: "No sampling needed—spectra are acquired directly", but this is a mistake of the highest order! Reference 6 was the first to deal inten-

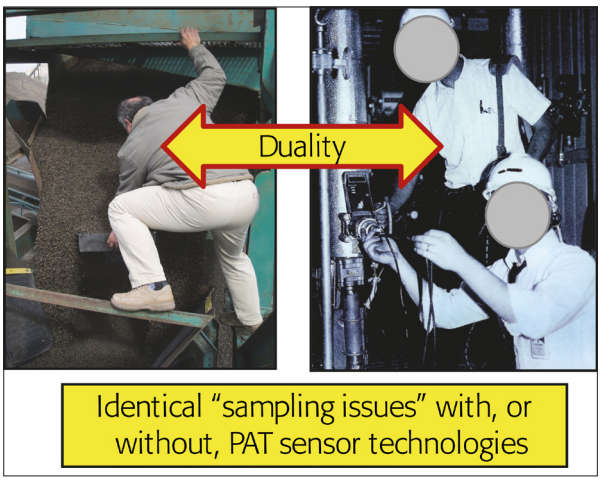


Figure 1. The fundamental sampling **duality**. Physical **or** optical grab sampling incur identical sampling error effects (ISE, CSE). Illustration copyright KHE Consulting, reproduced with permission.

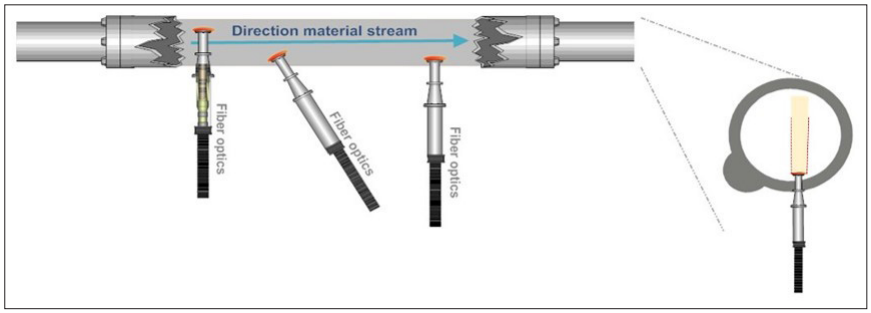


Figure 2. The "PAT sensor application solves all sampling issues" fallacy, which follows because PAT sensors are not "seeing" a volume corresponding to a full cross-sectional **slice** of a moving stream of matter, see also Figure 3. Illustration copyright KHE Consulting and Martin Lischka, reproduced with permission.

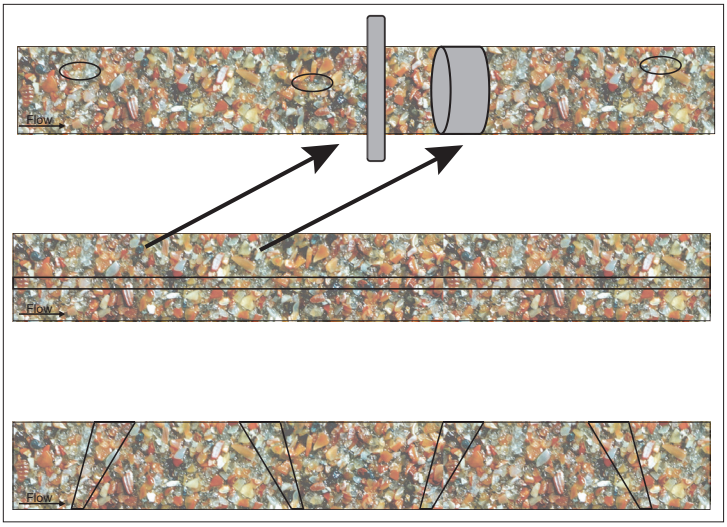


Figure 3. Massive ISE as a result of **non**-compliance with TOS' principles. This figure can alternatively be understood as looking down on the top of moving conveyor belts, or as representing a longitudinal section of a ducted flow in a pipeline. "TOS-correct" delineation and extraction of cross-sectional increments, or sliced stream volumes, are shown in grey in the top panel. All other indicated increments give rise to a significant sampling bias. Illustration copyright KHE Consulting and Martin Lischka, reproduced with permission.

sively with what is a sampling *duality*, Figure 1.

This "no sampling" fallacy optimism is illustrated by showing how "direct" application of a PAT sensor does **not** eliminate the occurrence of massive ISE, Figures 2 and 3.

It seems difficult to understand why this fallacy has originated, and why it has been propagated during at least two decades in the PAT realm. The strict truth is that **only** a full slice of the stream of matter qualifies as the proper volume/mass *support* for a representative increment/signal.[a]

The above leads directly to a fundamental distinction in this context between:

1) A *sample* cell (sample: noun) and,
2) A *sampling* cell… (sampling: verb)

Upon reflection, it is the act of *simultaneous* sampling-and-analysis that distinguishes the sampling cell—and which in a sense may appear as making (physical) sampling superfluous. But this latter is critically dependent on representative sampling, a proviso of overwhelming importance. If sampling is **not** representative, all manner of unknown, inconstant sampling bias will still be part of the equation, totally destroying the "no sampling" claim. On the other hand, if/when a sampling-and-analysis cell complies with TOS' demand for representativity, conditions are right for reaping the powerful advantages of the PAT revolution, but only then. Reference 6 treats these issues in detail.

A cursory survey of relevant industrial process technology and dedicated PAT literature from the last 10 years or so does not impress. Unwitting neglect of the "full slice" dictum can be found in abundance (but there are moments of satisfaction as well). The mission here is **not** to identify which are which, but only to direct attention to the critical need for a certain minimum TOS competence in the PAT realm.

Proper application of TOS' relevant GP and SUO in the pre-analysis realm is a mandatory requirement in order to **guarantee** that samples, or the spectral acquisitions from matter streams, can be proven to be representative. Failure to live up to this demand will result in *compromised* analytical samples/signals with which to begin a subsequent chemometric data analysis or modelling.

As a prominent contemporary example, consider the rapidly expanding case of continuous manufacturing (CM) in the pharmaceutical industry sector. Figure 5 shows the many locations in the CM pathway where NIR spectroscopic characterisation finds very good use. Figure 5 also shows where one would easily lose one's way were not a modicum of

[a]The present column presents the strict demands for representative increment extraction/signal acquisition. The reader will realise that often there are severe practical difficulties involved when trying to comply herewith, for example that the effective NIR path length is ~30 mm while the effective duct diameter can be larger (much larger), say 100 mm or more? "Smartly" implemented reflectance probes *may* go a certain way to remedy this shortfall, but are essentially bracketed by the same path length maximum. A bypass duct will quickly become of significant interest a.o. The issues raised in the present column will be addressed in the form of **solutions** in the next Columns in this series.

SAMPLING COLUMN

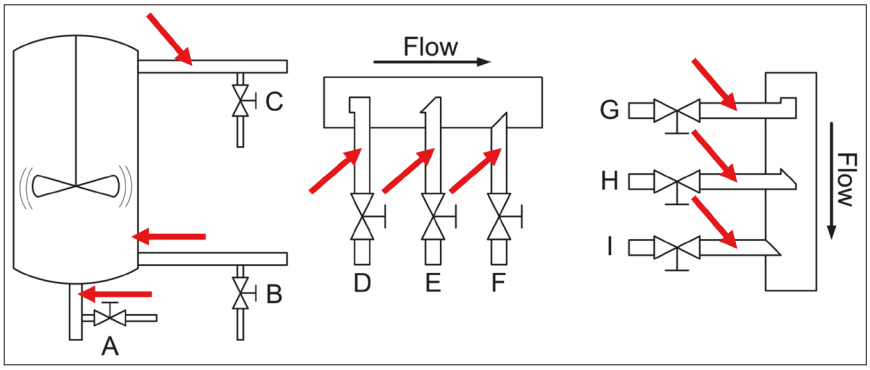


Figure 4. Overview of the many possibilities for installing "sample extraction valves" and "PAT sensors" in traditional process industry. Note that all configurations shown here will lead to incorrect, i.e. biased samples or spectral signals as regards representativity because their support volumes do not correspond to full stream slices. Reference 6 treats these process sampling issues in depth w.r.t. **solutions** to the problems emphasised. Illustration copyright KHE Consulting, reproduced with permission.

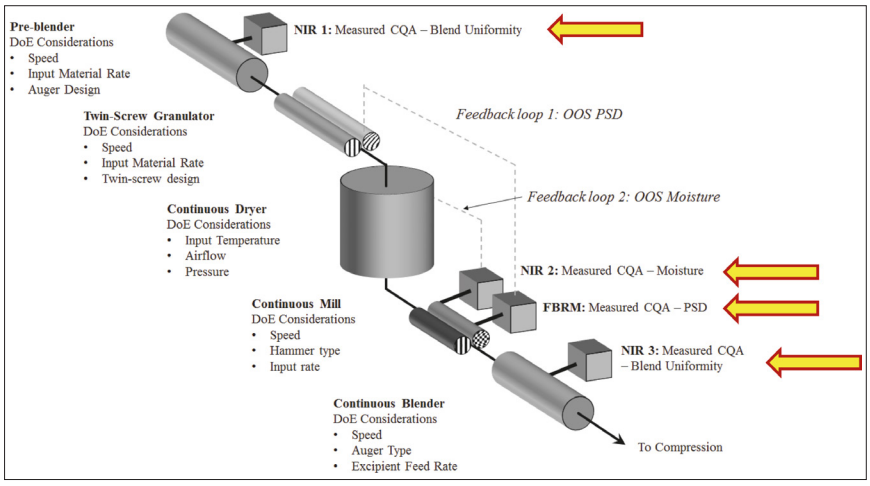


Figure 5. Four process analytical locations involved in CM, based on in-line sampling-and-NIR spectroscopic prediction of pharmaceutical API composition, moisture, PSD and blend uniformity. In addition to ISE associated with incorrect support volumes, which will affect the quality of spectral X-data in multivariate calibration, there is also the equally important issue of the quality of corresponding Y-data, for which representative reference samples must be obtained. But **where** exactly, and **how**, should these be extracted? This is a classical sample extraction issue, far from always properly acknowledged and far less satisfactorily solved, primarily because TOS continues to be a partly (largely) little acknowledged critical success factor in process technology, spectroscopy and chemometrics. Illustration copyright QbD Consultancy, reproduced with permission.

TOS competence involved in the design, implementation, validation and operation phases of starting up using this manufacturing approach.

While it may often *seem* to be relatively easy to obtain "direct spectral data" *via* inserted PAT sensors, as indicated in Figures 4 and 5 (X-data in regression), these are nevertheless critically *dependent* on whether the ISE issues outlined above have been successfully eliminated, or not. Failure to comply with this requirement is the by far the most often met with deficiency within a trigger-happy chemometric community; the literature is full of illustrative examples, but we shall here refrain from identifying journals, papers, authors—the task here is to sound a warning against continuing to be unaware (or to wilfully neglect) the critical support volume dictum. Fortunately, a lot of work has been performed by the pioneers of CM systems to address such sampling issues and the fact is that a CM system essentially reduces a traditional 3-D sampling plan (traditional approach to manufacturing) with a 1-D sampling situation. This is the optimal, TOS-correct understanding from which to begin to look for solutions to the sampling issues warned about.

In addition, in order to perform *proper* multivariate calibrations for one or more y-variables, it is necessary to extract relevant, valid and representative reference samples (Y-data in regression). Indeed, this also applies for proper test set samples to be used for *validation* of the desired multivariate calibration models with which to carry out on-line prediction of blend uniformities (real-time compositional variation), moisture, PSD… .[7] A documented facility for representative acquisition of **both** sensor signals (X-spectra) and reference samples (reference data) **must** be present, or multivariate calibration/validation models for prediction will forever continue to suffer "impossible to reduce" prediction errors etc. These issues are often described in suspiciously murky fashions in the literature; one is tempted to interpret this as if authors actually do understand the fundamental ISE issue here, but are at a complete loss to come up with solutions that work. In fact, "sweeping the problem under the carpet" has led to many process failure investigations and "incomprehensible results", which in reality is fighting fires that simply are just not there in the first place.

Many skills needed

This column has the purpose to introduce all elements from the diverse disciplines of i) the TOS, ii) process engineering, iii) spectroscopic analysis, iv) sensor technology, v) PAT and vi) chemometric data analysis. All need to acknowledge that analytical results pertaining to heterogeneous materials and systems have a **history** in which some degree of

[b] i.e. sub-sampling/splitting in several stages; these are *bona fide* sampling processes in their own right.

[c] Exceptions, for example uniform materials, de Beer dilutions etc. cannot establish a basis for solving the infinitely more complex issues surrounding the kind of significantly heterogeneous materials treated here.

sampling (primary, secondary, tertiary[b]) is always present.[c] For this fundamental reason "What's in an analytical datum?" will always include a significant complement of error effects and uncertainties stemming from this pre-analysis realm (smaller or larger, but never absent). The point is that this state-of-affairs must be acknowledged by analysts and data analysts; this point has been forcefully argued in Reference 5. It is no longer appropriate to pass the responsibility onto someone else. If PAT is going to be implemented successfully by an organisation, all involved in quality, production, regulatory affairs, as well as management, must be on the same page.

Chemometric data modelling

"Compromised samples or signals"—in what sense? Compromised, because extraction of physical samples or acquisition of spectral signals will be associated with a significant sampling bias. The nature of a sampling bias is that it cannot be corrected for by any means, data analytical or statistical, as distinct from an analytical bias, which can be subjected to a statistical bias-correction. This, perhaps surprising, distinction is treated in full in References 2–7.

The effect of this on chemometric data analysis and modelling is like the proverbial elephant in the room, generally unnoticed. Put simply, no manner of data analysis, data modelling etc. from the chemometric and the statistical domains will be able to correct for a sampling bias; see References 6 and 7 for a full argumentat for *why* this is not possible; also see Figure 6.

What is the specific effect on multivariate data analysis, modelling and calibration?

First: There will be an *inflated* total MU associated with every analytical result, very often significantly larger than the specific analytical error itself (which *may* occasionally also be significant of course, but only as a result of an analytical method not in proper control; such an issue will eventually be brought under control, GLP a.o.).

Second: There **will** always be a component of the multivariate data errors so effectively screened away by the powerful bi-linear data modelling approaches in chemometrics that must be taken into account when appropriate. Figure 7 shows a principal illustration of chemometric decomposition of multivariate data into systematic data structures (principal- and PLS-components) and decoupled multivariate data errors (εs in chemometric parlance).

A tacit understanding within chemometrics has been that bi-linear errors (εs) would turn out to include TOS-errors stemming from sampling deficiencies, in which case it would actually be possible to correct for ISE after all. Indeed, multivariate data analysis would then appear on the scientific scene with an unbelievable power, not even known or foreseen in the genesis and development of chemometrics. Alas, this is not so!

The effect of sampling bias inflation of the total sampling + analysis uncertainty level is such that both data analytical *components* as well as their complementary errors (εs) are affected by the inconstant bias effect. Because TOS-errors are expressed for single variables in turn, the bias will affect each individual variable differently. The sum-effect of an unresolved sampling bias is such that TSE_{TOT} will vary every time a new analytical determination is attempted on a new sample. This means that every new sample added to an already existing data matrix, think of a training data set, will each add its own, varying contribution to the total data variance—and thus also to the total data set covariance. Thus, both components and errors will be affected. These issues are described in more detail in Reference 7.

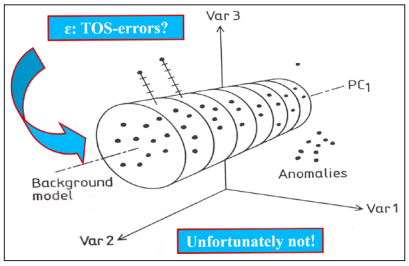


Figure 7. Chemometric bi-linear data model of pervasive data structures, PCA or PLS-components (schematic sketch). The multivariate data model errors (εs) are often optimistically thought of as "TOS-errors" (TSE_{TOT} + TAE), but they are not, see text for clarification.

For completion, the complement of Correct Sampling Errors (CSE) will also affect each variable individually, after elimination or maximal reduction of ISE.

Consequences for chemometrics

Because there are many influential agents involved for each sample extracted, or for each signal acquired by a PAT instrument, it may easily be an unhelpful simplification to understand all

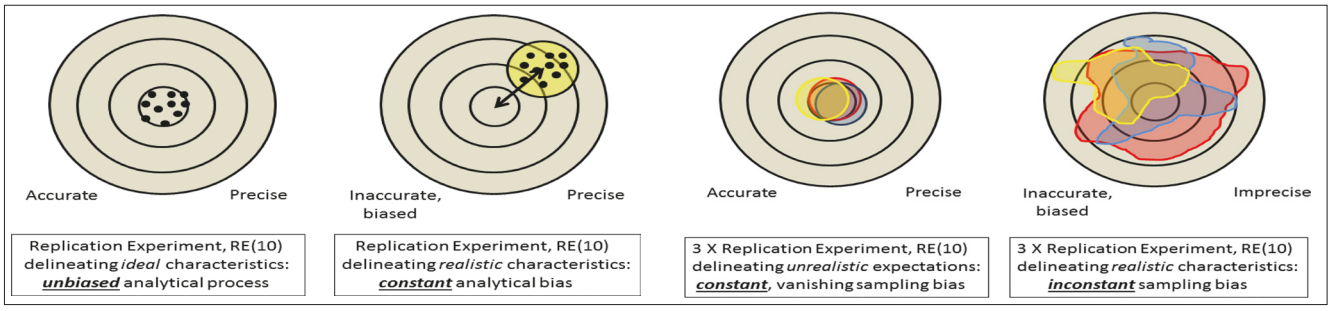


Figure 6. Analytical bias (left) vs sampling bias (right). Bias effects are shown as manifested by Replication Experiments (RE).

SAMPLING COLUMN

"data" as but identical realisations of variables, each with a systematic information content to which is added a stochastic error complement. Within chemometrics the former can be successfully modelled by data analytical "components" and the latter can be conveniently identified, quantified—and then discharged, so this would be "all one needs to know" **if** data are always, *ipso facto*, representative and reliable. This is not so, however!

This column has argued that the background for each individual datum, and for analytical data collected as matrices must be appreciated in a more fully developed setting in which significant parts of the traditional "measurement error" also contain contributions, often large contributions, related to the specific history of each sample, aliquot or spectral signal support. The data analyst must be mindful of this intricate relationship, lest glib, simplistic interpretations of "measurement error" will run a grave risk of not reflecting the more complex reality.

The archetypal manifestation of these relationships is shown in Figure 8, which highlights the fact that any chemometric prediction model falling short of sufficient performance, for example as evidenced by a "too high" $RMSEP_{validation}$, can **only** be improved upon by caring about the TSE incurred for all data, critically based on the full understanding that the error complement is overwhelmingly made up of contributions by the TSE_{TOT}.

Thus, it is not a guaranteed successful strategy to care only for the "data" as such, with an aim of optimal data analytical modelling systematics (chemometric components), *perhaps* acknowledging a minor measure of accidental analytical error in addition. This will, therefore, **not** include the major determinants stemming from unrecognised, or deliberately overlooked, sampling errors and their incurred uncertainties. Interpretation of standard regression-prediction figures-of-merit, e.g. $RMSEP_{validation}$ must be based on a modicum of TOS knowledge and competence in order to be able to *improve* on unsatisfactorily "imprecise" prediction performance statistics. In our collective experience many data analysts skip straight to the R^2 value of a

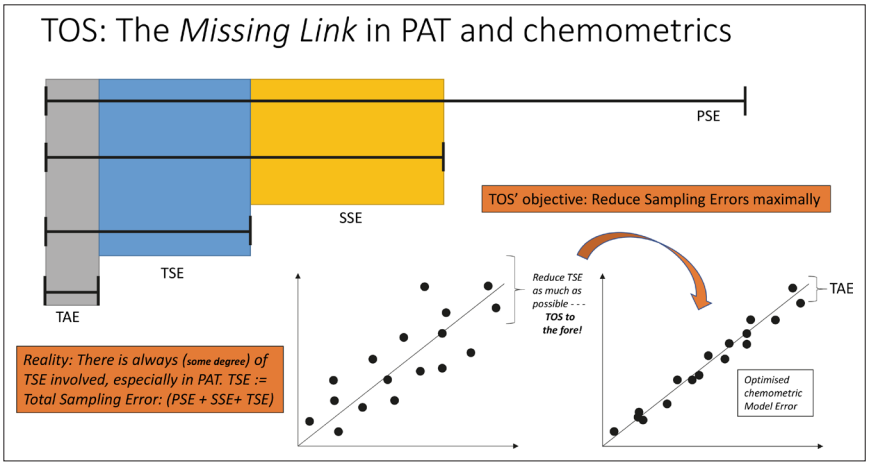


Figure 8. Theory of Sampling (TOS)—the missing link in PAT and chemometrics. Key effects shown here concern decoupling of TSE_{TOT} into contributions from primary sampling errors (PSE), secondary and tertiary sampling (SSE, TSE) and the analytical error s.s. (TAE). The lower panels illustrate how an unsatisfactory $RMSEP_{VALIDATION}$ of a multivariate prediction model manifests itself on a "predicted vs reference" plot. The TOS is the only approach that outlines solutions that lead to reduction of TSE_{TOT}.

fitted regression line **as if** this was some form of magical truth statistic, but it only relates to model fitting.

Here is the most relevant criterion for evaluation: for an optimal bi-linear model, the RMSEP errors should be statistically *comparable* to the reference chemistry errors and this situation should be achieved with a relatively small number of model components. The implicit modelling of random artefacts in the spectral data to overcome material heterogeneity and non-representative sampling **only** relates to the miniscule test volumes involved. If a large number of model components is required to achieve your error target (if this is even at all possible), this is only modelling a *mirage*. If a chemometric model cannot be brought below your *a priori* established RMSEP error threshold, it is telling you there is a fundamental problem *outside* of the analytical realm.

This is true information—for which no amount of calling for "more samples", more spectra (to average), more model components will ever help. The unpleasant situation, Figure 8, simply means that you must focus on improving your sampling practices. TOS to the fore!

Conclusions

The one sure way **not** to be able to reduce the uncertainty elements behind data analytical models that does not comply with desired prediction performance goals, is the traditional call for more data (an approach very often cited in the literature and observed in practice). More data, meaning more samples for analysis, will always display the same TSE_{TOT} characteristics as the samples already included in the contemporary training data set, see Reference 7. The number of times this futile call has been heard in practice is overwhelming, and is usually preached to those with little experience in the PAT/chemometrics fields in order to avoid the more difficult problems revealed here. Focussing on the root cause, i.e. *why* the samples and their analytical result do not match with reality is a simple sampling issue, however, and must be treated as such.

In fact, most of the initial efforts in PAT implementation and data modelling should be focused on improving and optimising sampling—way before analysis and data analysis. As the saying goes: "if the data already contain the information, then the chemometrics will succeed". However, if the data are swamped by sampling noise, even applying implicit or explicit "correction functions" will still not improve the accuracy of the analytical results, because this inaccuracy can never be modelled away.

SAMPLING COLUMN

Chemometrics is not a black box, "push button" approach where the modelling will automatically do the rest! Automated chemometrics routines in software packages should be outlawed and never used! How dare a vendor provide software to replace the many years of competent chemometricians' experience around the world by reducing our collective practice down to a single automated routine! And when this approach does not work, the same vendors tell their clients to collect "more samples" to avoid the issue at hand. Chemometrics is not a supermarket of models, but a scientific expertise area where all sources of variation must be understood such that the model is interpretable and validateable.

Barring trivial, accidental TAE mishaps (which are always special cases, only of interest to themselves), the only way to reduce unsatisfactory (TSE + TAE) levels is by reducing TSE_{TOT}. Thus, the **only** way to be able reduce the "troublesome", apparently incompressible uncertainty contributions behind unsatisfactory multivariate data analytical models, Figure 8, is to master the necessary basics of the TOS.

Chemometrians are not exempt from these scientific insights. There is no longer an excuse to hide behind "I don't need to learn chemometrics, the superior software will sort it out for me". Like with CGMPs for the 21st Century, we also need to take a 21st Century approach to the full sampling–analysis–data analysis pathway, otherwise we will be travelling the same merry go round, always chasing our own tail and never progressing.

The promise

We shall address the many issues pointed to in this column from the point of view of solutions in the next columns in this series.

References

1. K.H. Esbensen and C. Wagner, "Why we need the Theory of Sampling", *The Analytical Scientist* **21 (October),** 30–38 (2014).
2. R.C.A. Minnitt and K.H. Esbensen, "Pierre Gy's development of the Theory of Sampling: a retrospective summary with a didactic tutorial on quantitative sampling of one-dimensional lots", *TOS Forum* **7,** 7–19 (2017). https://doi.org/10.1255/tosf.96
3. K.H. Esbensen, R.J. Romañach and A.D. Román-Ospino, "Theory of Sampling (TOS): a necessary and sufficient guarantee for reliable multivariate data analysis in pharmaceutical manufacturing", in *Multivariate Analysis in the Pharmaceutical Industry*, Ed by A.P. Ferreira, J.C. Menezes and M. Tobyn. Academic Press, Ch. 4, pp. 53–91 (2018). https://doi.org/10.1016/B978-0-12-811065-2.00005-9
4. *DS 3077, Representative sampling—Horizontal Standard*. Danish Standards (2013). http://www.ds.dk
5. K.H. Esbensen and C. Wagner, "Theory of Sampling (TOS) versus Measurement Uncertainty (MU)—a call for integration", *Trends Anal. Chem. (TrAC)* **57,** 93–106 (2014). https://doi.org/10.1016/j.trac.2014.02.007
6. K.H. Esbensen and P. Mortensen, "Process Sampling (Theory of Sampling, TOS) – the Missing Link in Process Analytical Technology (PAT)", in *Process Analytical Technology*, 2nd Edn, Ed by K.A. Bakeev. Wiley, pp. 37–80 (2010). https://doi.org/10.1002/9780470689592.ch3
7. K.H. Esbensen and B. Swarbrick, *Multivariate Date Analysis – An introduction to Multivariate Data Analysis, Process Analytical Technology and Quality by Design*. CAMO (2018).
8. R.J. Romañach, R. Méndez and K.H. Esbensen, "Application of Theory of Sampling principles for real-time monitoring of pharmaceutical powder blends by near infrared spectroscopy", *Spectrosc. Europe* **31(5),** 22–27 (2019). https://www.spectroscopyeurope.com/sampling/application-theory-sampling-principles-real-time-monitoring-pharmaceutical-powder-blends
9. K.H. Esbensen, "A tale of two laboratories I: the challenge", *Spectrosc. Europe* **30(5),** 23–28 (2018). https://www.spectroscopyeurope.com/sampling/tale-two-laboratories-i-challenge
10. K.H. Esbensen, "A tale of two laboratories II: resolution", *Spectrosc. Europe* **30(6),** 23–28 (2018). https://www.spectroscopyeurope.com/sampling/tale-two-laboratories-ii-resolution

The authors

Kim H. Esbensen, Ph.D, Dr. (hon), has been research professor in geoscience data analysis and sampling at three universities (1990–2015), after which he moved to be an independent researcher and consultant: www.kheconsult.com. He is a member of several scientific societies and has published over 260 peer-reviewed papers. Together with Brad Swarbrick, he authored a widely used textbook in Multivariate Data Analysis, published in 2018 (http://bit.ly/chemometrics). He is a co-founder of the International Pierre Gy Sampling Association www.intsamp.org and is editor of the magazine *TOS forum*, impopen.com/tos-forum, and of this Sampling Column.

Brad Swarbrick is owner of Quality by Design Consultancy (www.qbd-consultancy.com) and Co-founder of KAX Group (www.kaxgrp.com). He is a world recognised expert in the application of chemometrics and design of experiment methodology to Process Analytical Technology (PAT) applications in Quality by Design (QbD) environments.

NEW PRODUCTS

IMAGING

Compact MEMS FT-NIR spectrometer

Ocean Insight has introduced a compact spectral sensor with a wavelength range of 1350–2500 nm. The NanoQuest is a MEMS-based FT-NIR spectrometer in a small and more accessible package. The NanoQuest uses patented micro-electro-mechanical systems (MEMS) technology that allow a continuous-wave Michelson interferometer to be created monolithically on a MEMS chip. This enables detection of all wavelengths simultaneously across the 1350–2500 nm spectral range, using a single-photodetector design to reduce instrument footprint and maintain low-noise, high-stability performance. Each NanoQuest comes with an optical fibre and operating software, and can be coupled to Ocean Insight light sources and accessories to configure systems for absorbance/transmission or reflectance measurements. Typical NanoQuest NIR applications include authentication of counterfeit products; characterisation and quantification of food, soil nutrients and industrial materials; and compositional analysis of bodily fluids and other biological specimens. For industrial applications, NanoQuest offers the advantages of scalability, low power needs and tolerance to vibration and other motion effects.

Ocean Insight

http://link.spectroscopyasia.com/32-002

Simplified deployment of hyperspectral imaging for inspection applications

Pleora Technologies and perClass BV have announced a technology partnership to simplify the deployment of machine learning hyperspectral imaging for inspection applications. The Pleora AI Gateway and perClass AI plug-in allow end-users and integrators to deploy machine learning hyperspectral capabilities without any additional programming knowledge. Images and data are uploaded to perClass Mira® "no code" training software on a host PC, which automatically generates AI models that are deployed on the Pleora AI Gateway in a production environment. Pleora's AI Gateway works with any standards-compliant hyperspectral sensor, bridging the gap between applications and existing machine vision software by automatically handling image acquisition from the hyperspectral imaging source and sending out the processed data over GigE Vision to inspection and analysis platforms. Pleora's AI Gateway provides additional plug-in AI skills for classification, sorting and detecting, with the processing flexibility of an NVIDIA GPU to train and deploy open source or custom algorithms developed in popular frameworks like TensorFlow and OpenCV. Lead customers are now evaluating the AI Gateway in inspection applications to help reduce costly inspection errors, false-positives and secondary screenings.

Pleora Technologies

http://link.spectroscopyasia.com/32-006

SWIR imaging camera

Teledyne Princeton Instruments have introduced the NIRvana HS, adding to their NIRvana SWIR camera portfolio. The NIRvana HS runs at 250 frames per second in 16-bit mode and offers both integrate-then-read (ITR) and integrate-while-read (IWR) modes for low noise and high duty cycle. The thermal design includes deep cooling to –55 °C and incorporates a vacuum sealed chamber to provide a lifetime of maintenance-free operation. There is advanced image correction and LightField® software provides an intuitive interface and analytical functions, eliminating the need for any third-party hardware or software.

Teledyne Princeton Instruments

http://link.spectroscopyasia.com/32-012

NEW PRODUCTS

Rapid spectroradiometer for LED production monitoring

The CAS 125 spectroradiometer from Instrument Systems has a CMOS sensor that is linked to a specially developed electronic readout circuit. This combination enables very low measurement times of 0.01 ms while also optimising long-term stability. The spectrograph design is based on the existing high-end CAS 140D device. This gives the CAS 125 a level of optical performance comparable to that of the CAS 140D in terms of both stray light suppression and optical throughput. The device-specific electronic readout circuit enables time-optimised control of the spectrometer through parameterisation of successive measurements in Recipe mode on the CAS 125. This eliminates the time-consuming step of communicating with the PC to initialise each subsequent stage of the measurement process. Spectral range is 200–1100 nm.

Another feature of the CAS 125 sensor is built-in temperature stabilisation. This results in dark current behaviour that is independent of the ambient conditions, enabling the CAS 125 to ensure optimum long-term stability even in environments where temperatures fluctuate. The flash trigger can also be parameterised, which helps users synchronise the spectrometer with other system components, for example by triggering a photodiode measurement.

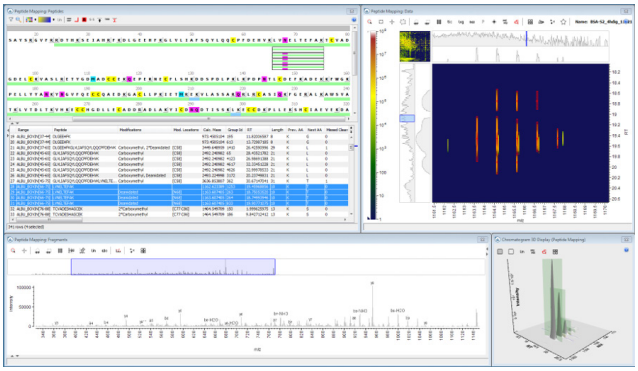
Instrument Systems

http://link.spectroscopyasia.com/32-011

MASS SPEC

New version of Genedata Expressionist for characterisation of biopharmaceuticals

Genedata has released Genedata Expressionist® 13.5, which has been developed in collaboration with partners across the biopharmaceutical industry, and provides unique workflow solutions for specific customer requirements. Out-of-the box data processing and analysis activities are configured to automate complex mass spectrometry-based workflows enabling users to scale up their experimental throughput while improving the quality and consistency of results. The new release includes the following highlights. New peptide mapping functionalities that enable searches for unexpected "wildcard" modifications and facilitate thorough review of results. Novel intact mass analysis providing automated deconvolution over broad mass ranges up to 1 MDa with artefact suppression and a selection of full-range or zoomed views; this improves the resolution and relative quantitation of coeluting species in analysis of intact proteins or subunits.



This release also includes integration with the newest SCIEX mass spectrometers and data formats—such as wiff and wiff2.

Genedata

http://link.spectroscopyasia.com/32-014

Increased performance for new PTR-TOF 1000 Ultra

Ionicon have introduced new and improved hardware components for the PTR-TOF 1000 ultra, which will increase the instrument's sensitivity significantly. Moreover, the PTR-TOF 1000 series now also features a higher mass resolution of >2000. The Ion-Booster has been revised and new PTR-TOF 1000 ultra instruments have 3× higher sensitivity of around 1500–3000 cps/ppbv as the new standard configuration.

The new X2 option is a combination of the Ion-Booster and the hexapole Ion-Guide. It increases the instrument's sensitivity to 10,000–20,000 cps/ppbv. Existing customers will be able upgrade.

Ionicon

http://link.spectroscopyasia.com/32-008

NEW PRODUCTS

Software tool for processing non-targeted analysis GC/MS data

The Southwest Research Institute has introduced Floodlight™, a software tool for processing non-targeted analysis (NTA) data from gas chromatography mass spectrometry and other instrument data. This cheminformatics machine learning tool integrates algorithms with analytical chemistry software to provide deep analysis of the instrument data. NTA data can require careful examination by a chemist to identify and exclude data artefacts caused by equipment, techniques or conditions. This data quality review can be automated with Floodlight.

Southwest Research Institute

http://link.spectroscopyasia.com/32-010

Portable quadrupole gas analyser

The Hiden *p*QA portable gas analyser is a versatile mass spectrometer with a range of interchangeable sampling inlets to suit a broad application range. MIMS inlets are offered for analysis of dissolved species in ground water, fermentation cultures, soil samples and general applications where analysis of dissolved species in liquid sample is required. The system is suited to gas analysis applications, where sample volume is small, and for environmental applications where detection of low concentration levels is required. The *p*QA system has a mass range of 200 amu and sub-ppb detection levels. Extended mass range to 300 amu is optional.

The system is supplied in a Pelican® case and can be powered by a 12V supply for field use, battery and/or solar powered or a 220V supply for laboratory use.

Hiden Analytical

http://link.spectroscopyasia.com/32-001

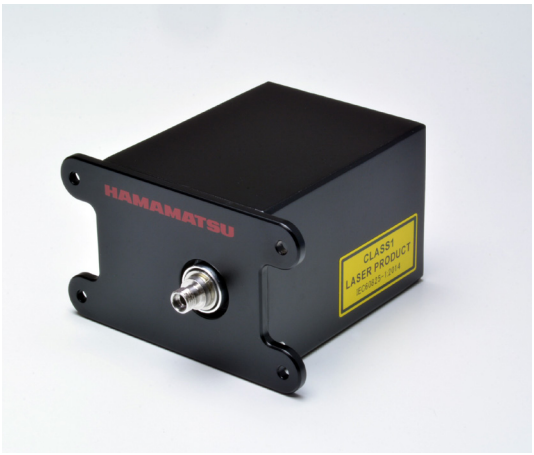
NEAR INFRARED

Palm-sized MEMS FT-NIR engine

Hamamatsu Photonics has developed a palm-sized FT-NIR engine with a wavelength range from 1100 nm to 2500 nm. The model C15511-01 is built using Hamamatsu's own microelectromechanical systems (MEMS) technology. Optical components such as the moving and fixed mirrors have been in a compact orientation that also minimises the error in the relative angle between the mirrors. The moving mirror is only 3 mm in diameter. The FT-NIR engine can help in creating handheld FT-NIR spectrophotometers for analytical applications including real-time monitoring of chemicals on production lines and ingredient analysis of agricultural products in the field.

Hamamatsu Photonics

http://link.spectroscopyasia.com/32-005



UV/VIS

New QC software for the Thermo Scientific NanoDrop OneC

Thermo Fisher Scientific has released the Thermo Scientific NanoDrop QC Software for the UV/vis microvolume NanoDrop OneC spectrophotometer. The NanoDrop QC Software natively runs chemometric methods, allowing materials scientists to obtain results of chemometric analysis in as little as 10 s. Once a spectroscopist develops the method, technicians can gather the data and complete the analyses. The NanoDrop QC Software is hoped to open applications in a wide range of industries such as petrochemical companies, adhesive and lubricant manufacturers, and food dye producers that need a fast and accurate way to test sample quality.

Thermo Fisher Scientific

http://link.spectroscopyasia.com/32-009

VACUUM

Agilent introduces new smart-connected turbomolecular pumps

Agilent Technologies have introduced two new models of turbomolecular pumps to their TwisTorr turbo pump range, both with a more compact design and smart capabilities. The TwisTorr 305 FS and TwisTorr 305 IC pumps both come with smart connectivity, a new feature for Agilent turbomolecular pumps. An app called Vacuum Link, which can be installed on Apple or Android phones, enables users to communicate remotely with the pump, so they can type commands and modify parameters to control the pump. The TwisTorr 305 FS pump is a standalone unit, featuring an external remote controller. The TwisTorr 305 IC pump features an integrated controller, and a small footprint making it of interest to original equipment manufacturers and other companies that want to integrate the pump in an instrument. An advanced function enables users to extract log files so they can share pump operating data easily, saving time. It also

enables quick communication with Agilent service and support teams, speeding up the company's response time.

Agilent Technologies

http://link.spectroscopyasia.com/32-007

X-RAY

Portable XRF MARPOL analyser for IMO 2020 low sulfur fuel oil requirements

Bruker has introduced a complete solution to test and verify adherence to the International Maritime Organization (IMO) Low Sulfur Fuel Oil Standard Requirement in response to the significant reduction of the maximum permissible levels of sulfur in marine fuels from 3.5% to 0.5%, as being enforced from the beginning of 2020. The new MARPOL package is based on the portable CTX™ 500S XRF analyser, and includes a ready-to-go MARPOL calibration set-up, a quality control (QC) kit with sample cups, XRF safety foil and QC standards. The calibration set-up, specifically developed for MARPOL applications, enables the instrument to provide detection limits of 30 ppm (3σ) for sulfur, making it suitable for marine fuel testing at service labs, supply stations, on docks, in ports and aboard ships, even for the ultra-low 0.1% sulfur limit which continues to be the standard in Emission Control Areas (ECA). Fuel samples can be placed inside the analysis chamber using sample cups or other liquid containers. All user operation is through an easy-to-use front panel touchscreen display or an optional PC via Wi-Fi or USB.

Bruker

http://link.spectroscopyasia.com/32-004

X-505 handheld XRF from SciAps

SciAps has released the X-505—the second addition to its new product line of small, light handheld X-ray analysers. The X-505 is a high-performance analyser at a competitive price for users who do not need the speed on Al, Mg, Si, P and S of the premium NDT/PMI X-550 model. The SciAps X-505 has the same balanced ergonomic design, narrow nose and light weight as the X-550 released at the end of 2019, but it is slightly slower and less expensive than the premium model. With its state-of-the-art design and excellent heat handling, the X-505 can run all day long, even in the hottest climates, without requiring down time due to overheating. The X-505 weighs 2.8 lbs (1.27 kg) with the battery.

SciAps

http://link.spectroscopyasia.com/32-003

SPECTROSCOPY asia

FREE REGISTRATION

Use this page to register for a free subscription to *Spectroscopy Asia*. You can also register online at the address below.

Online version at **http://www.spectroscopyasia.com/register**

1. Your Details

Title (Professor/Dr etc.):	First Name:	Last Name:
Organisation Name:	Address:	
City:	Postal Code:	State/Region:
Country:	Job Title:	

☐ Please send me/continue to send me a free copy of *Spectroscopy Asia*.

Signed: _____ E-mail: _____

Date: _____

Field of Work:
(please tick ALL that apply)

- ☐ Agriculture
- ☐ Biotechnology
- ☐ Chemicals
- ☐ Cultural Heritage
- ☐ Electronics/Semiconductors
- ☐ Energy and Fuels
- ☐ Environment
- ☐ Food
- ☐ General Analysis
- ☐ Instrumentation
- ☐ Life Science
- ☐ Materials Science
- ☐ Medical Sciences
- ☐ Metals and Minerals
- ☐ Nanotechnology
- ☐ Pharmaceuticals
- ☐ Proteomics and other -omics
- ☐ Polymers and Peptides
- ☐ Security and Forensics
- ☐ Water

Techniques:
(please tick ALL that apply)

- ☐ Atomic Absorption
- ☐ Atomic Emission
- ☐ Inductively Coupled Plasma (ICP)
- ☐ ICP/MS
- ☐ Atomic Fluorescence
- ☐ Chemicals and Reference Materials
- ☐ Data Handling/Chemometrics
- ☐ Gamma-Ray Spectroscopy
- ☐ Ion Mobility
- ☐ Imaging
- ☐ Infrared Spectroscopy
- ☐ Laser Spectroscopy
- ☐ LIBS
- ☐ Luminescence/Fluorescence
- ☐ Mass Spectrometry
- ☐ MRI
- ☐ Mobile Spectrometers
- ☐ Mössbauer Spectroscopy
- ☐ Near Infrared
- ☐ NMR, ESR, EPR
- ☐ Photonics and Optics
- ☐ Polarimetry
- ☐ Process spectroscopy
- ☐ Raman spectroscopy
- ☐ Sampling
- ☐ Separation Science
- ☐ Spectroradiometry
- ☐ Surface Analysis
- ☐ Terahertz Spectroscopy
- ☐ UV/Vis
- ☐ X-Ray Diffraction
- ☐ X-Ray Spectrometry
- ☐ None

Area of employment:
(please tick ONE only)

- ☐ Industry
- ☐ Independent Lab
- ☐ University/Institute
- ☐ Hospital
- ☐ Government
- ☐ Research Inst/Foundation
- ☐ Other_____

Job function: *(please tick ONE only)*

- ☐ Analyst
- ☐ Quality Control/Assurance
- ☐ Research Scientist
- ☐ Lab Manager
- ☐ University Professor
- ☐ Engineering/Design
- ☐ Manufacturing/Processing
- ☐ Teaching
- ☐ Marketing/Sales
- ☐ Other_____

Do you recommend, specify or authorise the purchase of spectroscopic instrumentation or services: ☐ Yes ☐ No

Your Personal Data (we will never pass your personal details to any third party)
John Wiley & Sons Ltd and IM Publications LLP will use the information you have provided to fulfil your request. In addition, please tick to receive the following options:

Digital options

☐ E-mail from Spectroscopy Europe/Asia about new content, products and services
☐ E-mail with information from carefully selected spectroscopy companies (your details are never disclosed)

Postal options

☐ Mailings with information from carefully selected spectroscopy companies (your details are never disclosed)
☐ Mailings from John Wiley & Sons and IM Publications

Now fax this page to **+44-1243-811711**, or return by mail to: *Spectroscopy Asia*, 6 Charlton Mill, Charlton, Chichester, West Sussex PO18 0HY, UK.

Conferences

2020

22–26 March, Philadelphia, United States. **259th American Chemical Society National Meeting.** natimtgs@asc.org, https://www.acs.org/content/acs/en/about/governance/committees/cwd/meetings.html.

4–7 April, San Diego, United States. **Experimental Biology 2020.** eb@faseb.org, https://experimentalbiology.org.

13–14 April, Prague, Czech Republic. **World Congress and Expo Meet on Polymer Science and Chemical Technology.** polymerscience2k20@gmail.com, https://polymerscience.scientificmeeticon.com/.

13–14 April, Prague, Czech Republic. **International Summit on Nanomedicine & Nanotechnology.** nanomednanotech2020@gmail.com, https://nanomedicine.scientificmeeticon.com/.

26–29 April, Oviedo, Spain. **The 5th International Glow Discharge Spectroscopy Symposium.** Peter Robinson, pete@masscare.co.uk, https://www.ew-gds.com/.

24–28 May, Chiba, Japan. **Japan Geoscience Union (JpGU) Meeting 2020.** http://www.jpgu.org/en/articles/20171208meetingplan.html.

24–26 May, Rome, Italy. **8th CMA4CH Meeting, Measurements, Diagnostics, Statistics in Environment and Cultural Heritage Fields.** infocma4ch@uniroma1.it, http://www.cma4ch.org.

24–28 May, Winnipeg, Canada. **103rd Canadian Chemistry Conference.** http://www.ccce2020.ca/.

24–28 May, Chiba City, Japan. **Japan Geoscience Union Meeting 2020.** http://www.jpgu.org/.

31 May–4 June, Houston, Texas, United States. **68th ASMS Conference.** https://www.asms.org/conferences/annual-conference/future-annual-conferences.

15–17 June, Bali, Indonesia. **International Conference on Materials Science and Engineering 2020.** materialsasia@prismscievents.com, https://www.materialsconferenceasia.com/.

17–18 June, Vancouver, Canada. **International Conference on Plant Science.** plantscience@meetingsengage.com, https://plantscience.peersalleyconferences.com/.

21–26 June, Courmayeur, Italy. **18th Chemometrics in Analytical Chemistry Conference (CAC2020).** ludovic.duponchel@univ-lille.fr, https://cac2020.sciencesconf.org/.

21–26 June, Honolulu, Hawali, United States. **2020 Goldschmidt Conference.** helpdesk@goldschmidt.info, https://goldschmidt.info/2020/.

28 June–4 July, Gangwon, South Korea. **AOGS 17th Annual Meeting.** info@asiaoceania.org, http://www.asiaoceania.org/society/public.asp?view=up_coming.

28 June–1 July, Denver, Colorado, United States. **Fluoropolymer 2020.** https://www.polyacs.net/20fluoropolymeprogram.

5–8 July, Skagen, Denmark. **International Association for Spectral Imaging (IASIM) 2020.** 2020@iasim.net, https://2020.iasim.net/.

6–8 July, Rome, Italy. **2nd Global Congress on Material Science & Engineering.** material2020science@gmail.com, https://www.medwideconferences.com/materialscience/.

25–31 July, Chambersburg, United States. **International Diffuse Reflectance Conference (IDRC) 2020.** info@cnirs.org, http://www.cnirs.org/.

17–18 August, Cairns, Queensland, Australia. **19th Australian Near Infrared Spectroscopy Group (ANISG) Conference.** secretary@anisg.com.au, https://anisg.com.au/.

23–28 August, Boston, MA, United States. **XXIX International Conference on Magnetic Resonance in Biological Systems (ICMRBSXXIX).** John Markley, jmarkley@wisc.edu, http://www.icmrbs.org/.

6–10 September, Singapore, Singapore. **SETAC 8th World Congress.** setac@setac.org, https://singapore.setac.org/.

13–16 September, Orlando, United States. **134th AOAC International Annual Meeting & Exposition.** meetings@aoac.org, http://www.aoac.org.

20–26 September, Aachen, Germany. **17th International Symposium of Trace Elements in Man and Animals (TEMA17).** Prof. Dr. Lothar Rink, immunologie@ukaachen.de, https://www.ukaachen.de/kliniken-institute/institut-fuer-immunologie/institut.html.

20–25 September, Kyoto, Japan. **11th International Conference on Laser-Induced Breakdown Spectroscopy (LIBS2020).** Yoshihiro Deguchi, ydeguchi@tokushima-u.ac.jp, http://www.fm.ehcc.kyoto-u.ac.jp/Sakkalab/member/sakka/LIBS2020/index.htm.

30 September–2 October, Amsterdam, Holland. **11th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS).** http://www.spectroexpo.com/whispers/.

6–7 October, Sanur, Bali. **The 4th International Seminar on Photonics, Optics, and its Applications (ISPhOA 2020).** secretariat@isphoa.org, https://isphoa.org/.

Introduction to the Theory and Practice of Sampling

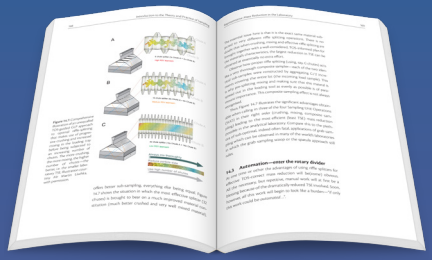
NEW BOOK

Kim H. Esbensen

with contributions from Claas Wagner, Pentti Minkkinen, Claudia Paoletti, Karin Engström, Martin Lischka and Jørgen Riis Pedersen

"Sampling is not gambling". Analytical results forming the basis for decision making in science, technology, industry and society must be relevant, valid and reliable. However, analytical results cannot be detached from the specific conditions under which they originated. Sampling comes to the fore as a critical success factor before analysis, which should only be made on documented representative samples. There is a complex and challenging pathway from heterogeneous materials in "lots" such as satchels, bags, drums, vessels, truck loads, railroad cars, shiploads, stockpiles (in the kg–ton range) to the miniscule laboratory aliquot (in the g–µg range), which is what is actually analysed.

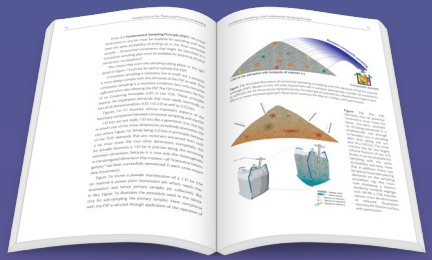
This book presents the Theory and Practice of Sampling (TOS) starting from level zero in a novel didactic framework without excessive mathematics and statistics. The book covers sampling from stationary lots, from moving, dynamic lots (process sampling) and has a vital focus on sampling in the analytical laboratory.



"I recommend this book to all newcomers to TOS"

"This book may well end up being the standard introduction sourcebook for representative sampling."

"One of the book's major advantages is the lavish use of carefully designed didactic diagrams"



impopen.com/sampling

IMPublicationsOpen