Supplement January 2021

SPECTROSCOPY europe

Spectroscopy since 1975

Top 20 in 2020

IMPOpen

ISSN: 2634-2561

CONTENTS

Spectroscopy Europe is a digital magazine, published eight times a year and available free-of-charge to qualifying individuals in Europe and the Middle East. Institutions can subscribe at the rate of €152 (Europe), £108 (UK), $208 (ROW, postage included) for the eight issues published in 2021. All paid subscription enquiries should be addressed to: Subscriptions, IM Publications Open, 6 Charlton Mill, Charlton, Chichester, West Sussex PO18 0HY, United Kingdom.

IMPOpen

Publisher
Ian Michael
E-mail: ian@impopen.com

Advertising Sales
Ian Michael
IM Publications Open, 6 Charlton Mill, Charlton, Chichester, West Sussex PO18 0HY, United Kingdom. Tel: +44-1243-811334, Fax: +44-1243-811711,
E-mail: ian@impopen.com

Paid subscriptions
IM Publications Open, 6 Charlton Mill, Charlton, Chichester, West Sussex PO18 0HY, United Kingdom. Tel: +44-1243-811334, Fax: +44-1243-811711,
E-mail: katie@impopen.com

Spectroscopy Europe is published by IM Publications Open LLP, 6 Charlton Mill, Charlton, Chichester, West Sussex PO18 0HY, United Kingdom.

Supplement January 2021

Top 20 in 2020

This is a Supplement to *Spectroscopy Europe*, bring together the Top 20 most viewed articles published online in 2020.

Each piece is organised into the section it originally appeared under and has a prefix of its rank from 20 to 1. All 20 are listed opposite in the order they appear throughout the issue.

If you have not already completed the short Reader Survey, I would be most grateful if you would: https://www.spectroscopyeurope.com/2021-reader-survey

We will start publication of our regular digital issues early in February 2021.

NEWS

20 NIR spectroscopy can detect acrylamide

Acrylamide is a natural neurotoxin often found in starchy foods. It is formed as a chemical reaction when the food is cooked at high temperatures. Examples of foods with higher levels of acrylamide include fries/chips, potato chips/crisps, breads, cereals and coffee. According to the National Cancer Institute, studies in rodent models have shown that exposure to acrylamide can increase the risk of several types of cancer. Furthermore, because the body converts acrylamide into glycidamide, there is also an association between acrylamide intake and mutations in and damage to DNA.

In light of these risk factors, in 2018, the European Union implemented a directive that urges food operators to closely monitor the presence of acrylamide in foodstuffs. Unfortunately, doing so is easier said than done. That's because existing techniques for detecting acrylamide are time-consuming, expensive and overly complex. To help food operators comply with the EU directive, the EU-funded AFREELAMIDE project developed a detection device that uses visible/near infrared (vis/NIR) spectroscopy. The new device allows users to conduct real-time monitoring of individual food pieces on site and at a cost that is five times cheaper than traditional inspection means.

"This is the first time that spectroscopy systems have been used to detect acrylamide levels in regular, everyday foods", explains Efrén García, Founder and CEO of Centaurea, a Spanish tech company and lead partner in the AFREELAMIDE project. "With spectroscopy, we can monitor 100 % of a producer's food products for the presence of acrylamide, saving the company money and ensuring the health and well-being of consumers."

During the course of the project, the AFREELAMIDE system was tested in several real industrial settings, including global agri-food company Siro Group. "These tests demonstrated that by using a near infrared system to measure acrylamide, companies can ensure that their entire output is completely free of this potentially dangerous neurotoxin", explains García.

Based on a market study conducted by project researchers, AFREELAMIDE has the potential to be both scalable and profitable. Sold at a unit price of EUR 20,000, García notes that his company anticipates sales of more than EUR 8 million by 2025. In preparation for its marketisation, Centaurea is currently working to obtain financing through the EIC Accelerator. This funding will be used to develop a large database of all the foods affected by acrylamide.

16 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 coupled plasma-mass spectrometry (ICP-MS) for the characterisation of microplastics (MP). The approach relies on the ultrafast 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).

15 COVID-19 therapy potential targets highlighted by new mass spectrometry technique

A team of biochemists and virologists at Goethe University and the Frankfurt University Hospital were able to observe how human cells change upon infection with SARS-CoV-2, the virus causing COVID-19 in people. The scientists tested a series of compounds in laboratory models and found some which slowed down or stopped virus reproduction. These results now enable the search for an active substance to be narrowed down to a small number of already approved drugs (*Nature*, http://doi.org/dw7s). Based on these findings, a US company reports that it is preparing clinical trials. A Canadian company is also starting a clinical study with a different substance.

Since the start of February, the Medical Virology of the Frankfurt University Hospital has been in possession of a SARS-CoV-2 infection cell culture system. The Frankfurt scientists in Professor Sandra Ciesek's team succeeded in cultivating the virus in colon cells from swabs taken from two infected individuals returning from Wuhan (http://doi.org/ggpw78). Using a mass spectrometry technique developed at the Institute for Biochemistry II at Goethe University Frankfurt, researchers from both institutions were together able to show how a SARS-CoV-2 infection changes the human host cells. The scientists used a particular form of MS called the multiplexed enhanced protein dynamics (mePROD) method, which they had developed only a few months previously. This method makes it possible to determine the amount and synthesis rate of thousands of proteins within a cell.

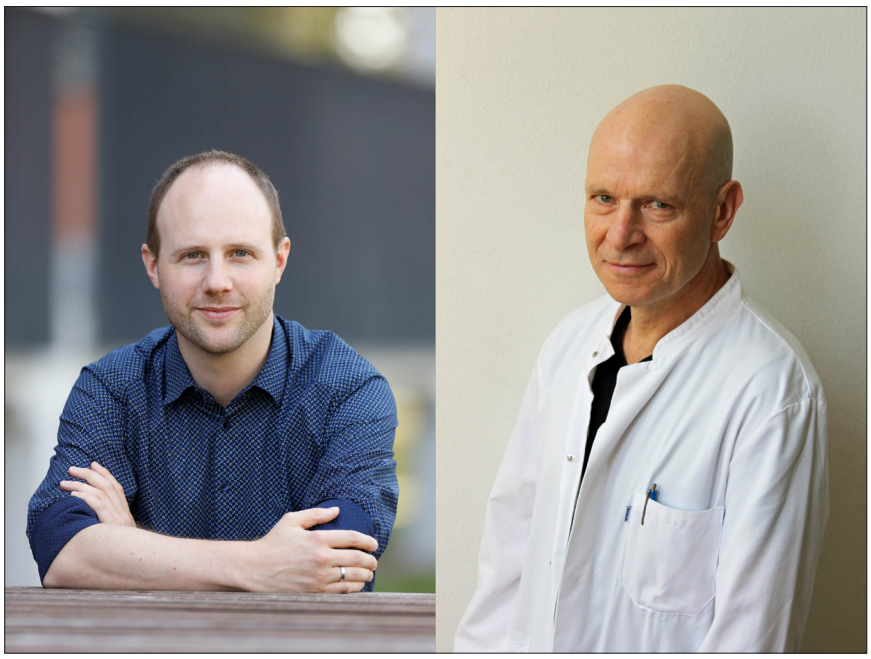
The findings paint a picture of the progression of a SARS-CoV-2 infection: whilst many viruses shut down the host's protein production to the benefit of viral proteins, SARS-CoV-2 only slightly influences the protein production of the host cell, with the viral proteins appearing to be produced in competition to host cell proteins. Instead, a SARS-CoV-2 infection leads to an increased protein synthesis machinery in the cell. The researchers suspected this was a weak spot of the virus and were indeed able to significantly reduce virus reproduction using translation inhibitors, which shut down protein production.

Twenty-four hours after infection, the virus causes distinct changes to the composition of the host proteome: while cholesterol metabolism is reduced, activities in carbohydrate metabolism and in modification of RNA as protein precursors increase. In line with this, the scientists were successful in stopping virus reproduction in cultivated cells by applying inhibitors of these processes. Similar success was achieved by using a substance that inhibits the production of building blocks for the viral genome.

In keeping with common practice since the beginning of the corona crisis, the Frankfurt researchers made these findings immediately available on a preprint server and on the website of the Institute for Biochemistry II (http://pqc.biochem2.de#coronavirus). Professor Ivan Dikic, Director of the Institute, comments: "Both the culture of 'open science', in which we share our scientific findings as quickly as possible, and the interdisciplinary collaboration between biochemists and virologists contributed to this success. This project started not even three months ago, and has already revealed new therapeutic approaches to COVID-19."

Professor Sandra Ciesek, Director of the Institute for Medical Virology at the University Hospital Frankfurt, explains: "In a unique situation like this we also have to take new paths in research. An already existing cooperation between the Cinatl and Münch laboratories made it possible to quickly focus the research on SARS-CoV-2. The findings so far are a wonderful affirmation of this approach of cross-disciplinary collaborations."

Among the substances that stopped viral reproduction in the cell culture system was 2-Deoxy-D-Glucose (2-DG), which interferes directly with the carbohydrate metabolism necessary for viral reproduction. The US company Moleculin Biotech possesses a substance called WP1122, a prodrug similar to 2-DG. Recently, Moleculin Biotech announced that they are preparing a clinical trial with this substance based on the results from



Left: Dr Christian Münch (Credit: Uwe Dettmar for Goethe University Frankfurt). Right: Prof. Dr. rer. nat. Jindrich Cinatl (Credit: University Hospital Frankfurt).

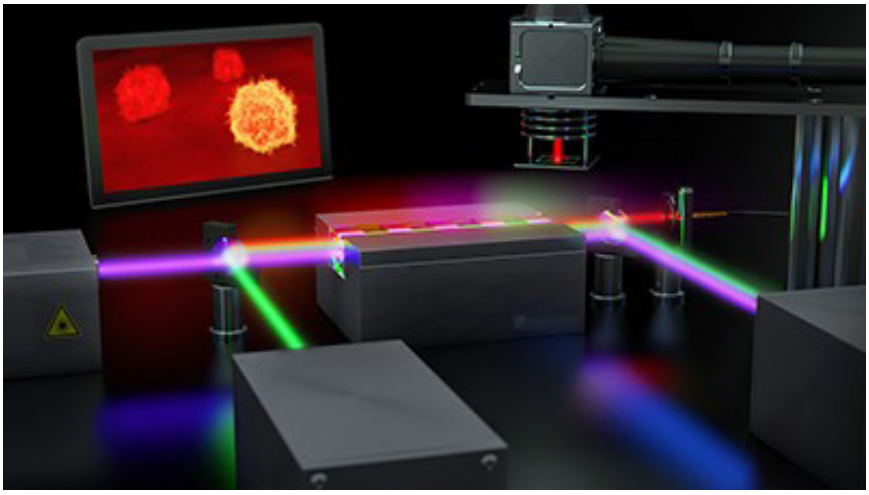
Frankfurt, https://www.moleculin.com/covid-19/.

Based on another one of the substances tested in Frankfurt, Ribavirin, the Canadian company Bausch Health Americas is starting a clinical study with 50 participants: https://clinicaltrials.gov/ct2/show/NCT04356677?term=04356677&draw=2&rank=1

Dr Christian Münch, Head of the Protein Quality Control Group at the Institute for Biochemistry II and lead author, comments: "Thanks to the mePROD-technology we developed, we were for the first time able to trace the cellular changes upon infection over time and with high detail in our laboratory. We were obviously aware of the potential scope of our findings. However, they are based on a cell culture system and require further testing. The fact that our findings may now immediately trigger further *in vivo* studies with the purpose of drug development is definitely a great stroke of luck." Beyond this, there are also other potentially interesting candidates among the inhibitors tested, says Münch, some of which have already been approved for other indications.

Professor Jindrich Cinatl from the Institute of Medical Virology and lead author explains: "The successful use of substances that are components of already approved drugs to combat SARS-CoV-2 is a great opportunity in the fight against the virus. These substances are already well characterised, and we know how they are tolerated by patients. This is why there is currently a global search for these types of substances. In the race against time, our work can now make an important contribution as to which directions promise the fastest success."

14 Using entangled photons to image in the IR and THz



While optical analysis techniques such as microscopy and spectroscopy are extremely efficient in visible wavelength ranges, they quickly reach their limits in the infrared or terahertz range. That, however, is precisely where valuable information is hidden. For example, biosubstances such as proteins, lipids and other biochemical components can be distinguished based on their characteristic molecular vibrations in the mid-infrared to terahertz range and are very difficult to detect with conventional measurement techniques. "If these motions could be captured or induced, it would be possible to see exactly how certain proteins, lipids and other substances are distributed in cell samples. For example, some types of cancer have a characteristic concentration or expression of certain proteins. This would mean that the disease could be detected and treated more efficiently. More precise knowledge of the distribution of bio-substances could bring major advances in drug research, as well", says quantum researcher Dr Markus Gräfe from Fraunhofer IOF.

The quantum mechanical effect of photon entanglement is helping the researchers allowing them to harness twin beams of light with different wavelengths. In an interferometric setup, a laser beam is sent through a non-linear crystal in which it generates two entangled light beams. These two beams can have very different wavelengths depending on the crystal's properties, but they are still connected to each other due to their entanglement.

"So now, while one photon beam in the invisible infrared range is sent to the object for illumination and interaction, its twin beam in the visible spectrum is captured by a camera. Since the entangled light particles carry the same information, an image is generated even though the light that reaches the camera never interacted with the actual object", explains Gräfe. The visible twin essentially provides insight into what is happening with the invisible twin.

The same principle can also be used in the ultraviolet spectral range: UV light easily damages cells, so living samples are extremely sensitive to that light. This significantly limits the time available for investigating, for instance, cell processes that last several hours or more. Since less light and smaller doses of radiation penetrate tissue cells during quantum imaging, they can be observed and analysed at high resolution for longer periods without destroying them.

The researchers are currently working to make the system more compact, shrinking it to the size of a shoebox, and to further enhance its resolution. The next step they hope to achieve is, for example, a quantum scanning microscope. Instead of the image being captured with a wide-field camera, it will be scanned, similar to a laser-scanning microscope. The researchers expect this to yield even higher resolutions of less than 1 µm, enabling the examination of structures within individual cells in even greater detail. In the long term, they want to see quantum imaging integrated into existing microscopy systems as a basic technology, thus lowering the barriers for industry users.

8 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, says Peter So, a professor of biological and mechanical engineering at MIT.

"Today, diabetes is a global epidemic", says So, who is one of the senior authors of the study and the director of MIT's Laser Biomedical Research Center. "If there were a good method for continuous glucose monitoring, one could potentially think about developing better management of the disease."

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.

The researchers tested the system in pigs and found that after 10–15 min of calibration, they could get accurate glucose readings for up to an hour. They verified the readings by comparing them to glucose measurements taken from blood samples.

"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.

7 Imaging spectroscopy to track plant pathogens through the atmosphere

The multidisciplinary team of scientists has been selected for a $750,000 NASA grant to combine their expertise in remote sensing, climate and earth system computer modelling, plant pathology and genomics to study the effects of soilborne plant pathogens which can travel in dust clouds from Africa to the Western Hemisphere. They will also use Earth system modelling to predict how regions will change over time and how that may influence plant disease dispersal with dust.

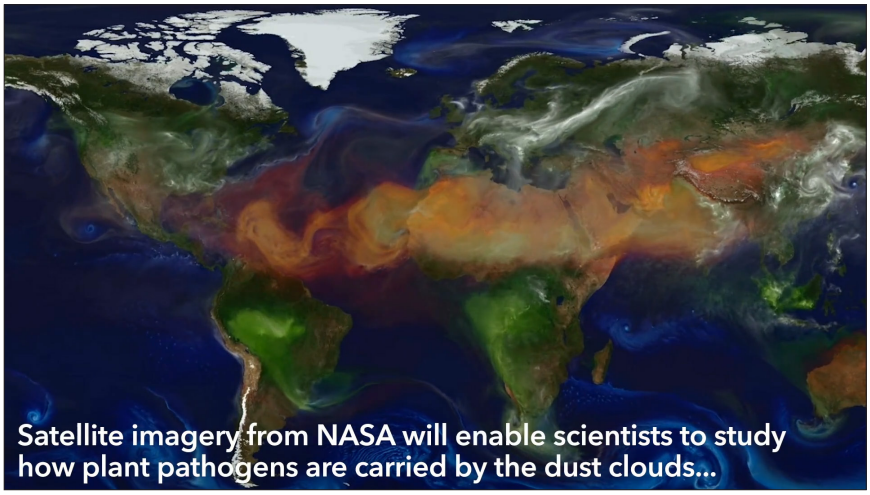
If the origins and landing spots of specific pathogens can be better predicted, farmers can be advised on how to avoid practices that would increase its spread, such as those that kick up dust from farm fields, and perhaps grow less susceptible crops where such dust falls.

"We lose anywhere from 15% to 30% of the global harvest to plant diseases annually; here in 2020, people still die because they don't have access to food, because of losses due to plant disease", said principal investigator Katie Gold, assistant professor of plant pathology. "Remote sensing can do a lot help mitigate the impacts of plant disease on the global food supply."

"It's just a fascinating combination of cross-disciplinary work that's going to allow us to address things that no one has been able to address before", said co-investigator Natalie Mahowald, professor of Earth and atmospheric sciences and an expert in atmospheric modelling.

NASA's Release of Research Opportunities in Space and Earth Science Interdisciplinary Science grant is for three years, which will allow the team to lay the foundation for a global surveillance system to assess risk and track and potentially prevent the global spread of plant diseases.

Other co-investigators include Ryan Pavlick, an imaging spectroscopy technologist at NASA's Jet Propulsion Lab, and Sharifa Crandall, assistant professor of soilborne disease dynamics and management at Pennsylvania State University.



NASA remote sensing satellites will help Cornell University faculty members Katie Gold, assistant professor of plant pathology, and Natalie Mahowald, professor of Earth and atmospheric sciences, study the effects of soil-borne plant pathogens which can travel in dust clouds from Africa to the Western Hemisphere. Credit: John Munson/Cornell University

6 Atomic spectroscopy of Stonehenge stones reveals their origin

The ancient stone circle at Stonehenge in the UK has been of fascination for us and our ancestors for around 5000 years. Now, we wonder at how the vast stones were transported to the site and erected. In the past, the fascination was with the movement of the Sun and the passing of the seasons. It has been known for some years that the smaller "bluestones" came from the Preseli Hills in Wales. However, larger (up to 7 m) and heavier (20 tonnes) stones, known as sarsens, are also used at Stonehenge, and the origin of these—more homogeneous in composition—has been impossible to identify until now.

A paper in *Science Advances* (doi.org/d5d7) reports a study that has

pinpointed the source of the sarsens to an area around 15 miles north of the stone circle site. The breakthrough came when a core—drilled from Stonehenge's "Stone 58" during repair work in the 1950s—was returned to English Heritage from Florida last year. This was at the request of one of those involved at the time, Mr Robert Phillips. This presented a unique opportunity to analyse the interior of one of the sarsens with destructive techniques.

The team first carried out non-destructive testing of all the remaining sarsens at Stonehenge using portable x-ray fluorescence spectrometry (pXRF). This revealed that most, including Stone 58, shared a similar chemistry and came from the same area. They then analysed sarsen outcrops from across the south of the UK using inductively coupled plasma mass spectrometry (ICP-MS) and ICP-atomic emission spectrometry (ICP-AES). The return of the core drilled from Stone 58 in the 1950s offered the opportunity to perform destructive analysis as well using ICP-MS and ICP-AES, and to compare this with the compositions of a number of stones from possible sources around the country. This showed that the composition of Stone 58 matched the chemistry of sarsens at West Woods, just south of Marlborough: about 40 minutes drive from Stonehenge today.

English Heritage Senior Properties Historian Susan Greaney said: "To be able to pinpoint the area that Stonehenge's builders used to source their materials around 2500 BC is a real thrill. Now we can start to understand the route they might have travelled and add another piece to the puzzle. While we had our suspicions that Stonehenge's sarsens came from the Marlborough Downs, we didn't know for sure, and with areas of sarsens across Wiltshire, the stones could have come from anywhere. We can now say, when sourcing the sarsens, the over-riding objective was size—they wanted the biggest, most substantial stones they could find and it made sense to get them from as nearby as possible. This is in stark contrast to the source of the bluestones, where something quite different—a sacred connection to these mountains perhaps—was at play. Yet again this evidence highlights just how carefully considered and deliberate the building of this phase of Stonehenge was."

Professor David Nash, University of Brighton, said: "It has been really exciting to harness 21st century science to understand the Neolithic past, and finally answer a question that archaeologists have been debating for centuries. We were able to investigate the chemistry of the sarsens at Stonehenge using x-ray fluorescence, a non-destructive technique. This showed that most of the stones shared a similar chemistry and likely came from a similar source. We then applied mass spectroscopy to samples from sarsen outcrops across southern England and to tiny pieces of the Phillips' Core from Stonehenge. Each outcrop was found to have a different geochemical signature, but it was the chance to test the returned core that enabled us to determine the source area for the Stonehenge sarsens. We're incredibly grateful to the Phillips family for returning the core to us."

ARTICLE

19 Identification of dental materials in food

Johannes Hesper
Shimadzu Europa GmbH, 47269 Duisburg, Germany

Nuts, bread crusts, toffee, apples—a hearty bite into food can be the trigger that breaks a tooth. However, the real reason may not be the food or contaminants or particles in it, but the tooth's natural age, structure and chemical make-up, or even caries. In case of complaints, how can food producers trace what has happened and begin an analysis of the cause? It starts with the materials…

There are many different types of dental products and their characteristics vary according to their intended purpose. Examples include temporary dressings, dental restorations (fillings, crowns, bridges), endodontic products (used in root canal therapy), impression materials, prosthetic materials (dentures) or dental implants. They all differ from the natural substances of a tooth.

In order to analyse the substance of materials, Fourier transform infrared (FT-IR) spectroscopy and energy-dispersive x-ray fluorescence (ED-XRF) are the methods of choice. Here, we report the analysis of a human tooth and various dental restorative materials by an FT-IR spectroscopy instrument coupled to an FT-IR microscope for molecular analysis and an ED-XRF spectrometer for elemental analysis. Due to the diversity of dental materials, which include hybrid materials containing organic and inorganic components, metal materials and others, a combined analysis approach was chosen.

Dental materials

Dental materials, such as metals, porcelain, ceramics and composite resins, are used in restorations, for example to repair carious lesions or other defects on single teeth. Fillings may be necessary in both the crown part and the root part of a tooth. A classic example is the amalgam filling, which has been used for centuries.[1]

Fillings can break off the tooth completely or only in pieces, and they may then be found as supposedly unknown foreign bodies in food. In some cases, this contamination is wrongly perceived as food contamination and consumers send it to the food manufacturer, complaining that the contamination originates from the production process. This can be clarified quickly by a combined analysis with FT-IR and ED-XRF; the complementary nature of these techniques can identify the source of contamination unambiguously.

Measurement methods and instruments

ED-XRF is a method used in material analysis and is one of the most frequently used methods for the qualitative and quantitative determination of the elemental composition of a sample, since the samples are not destroyed by the measurement and no digestions are required. It is widely used in the metal-working industry, in the analysis of glass, ceramics and building materials, as well as in the analysis of lubricants and mineral oil products.

The detection limit of ED-XRF depends on the sample matrix and the element itself. It can reach ppm levels for most elements. Here, an ED-XRF system with a Rhodium tube was used, leading to a visible Rh signal in the ED-XRF energy spectrum. These x-ray target signals are subtracted automatically from the result table, since they are not part of the sample composition. Nevertheless, Rh can be analysed by ED-XRF via the use of primary filters.

FT-IR spectroscopy can be used for the analysis of a solid, liquid or gas, simultaneously collecting data with high spectral resolution over a wide spectral range. With appropriate accessories, such as attenuated total reflection (ATR) or an

FT-IR microscope, even small samples (down to few µm) can be characterised quickly.

The analyses were conducted using an ED-XRF spectrometer (Shimadzu EDX-8000) and an AIM-9000 infrared microscope attached to an IRAffinity-1S FT-IR spectrophotometer (Shimadzu). Compared to other methods such as atomic absorption spectroscopy (AAS), inductively coupled plasma optical emission spectroscopy (ICP-OES) or inductively coupled plasma mass spectrometry (ICP-MS), both measurement techniques examine the sample destruction-free in its original state. For larger samples, it may be necessary to use a smaller piece to make it fit into the measuring instrument or accessory. If required, further tests can be performed with other analytical instruments to determine additional details of the sample.

Table 1 shows the analysis conditions of the ED-XRF and FT-IR instruments.

Human tooth—strontium and calcium phosphate make the difference

In a first step, a human tooth was analysed. Figure 1 shows a photograph of the contaminant, and Figure 2 shows a qualitative and quantitative analysis by ED-XRF. Calcium (Ca) and phosphorus (P), which are the main components of teeth, could be confirmed. Although the elemental composition is similar to that of bone, the lower content of strontium (Sr) is a characteristic feature of teeth. Like calcium, strontium is an alkaline earth metal and, therefore, can also be used to build teeth.

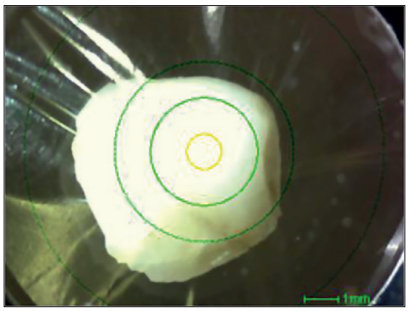
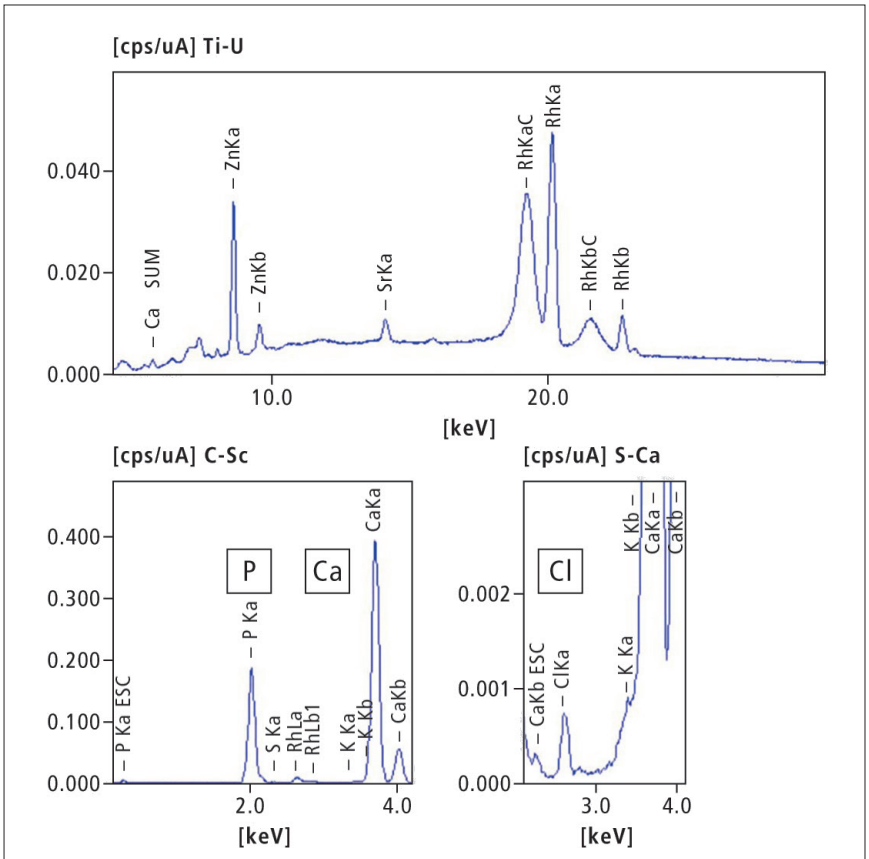


Figure 1. Photograph of contaminant (the yellow circle is the x-ray irradiation range: 1 mm ∅).

Table 1. ED-XRF and FT-IR analysis conditions.

Instrument: EDX-8000/EDX-8000P	
X-ray tube target	Rh target
Tube voltage/Tube current	15 kV (C-Sc, S-Ca), 50 kV (Ti-U)/Auto
Atmosphere	Vacuum
Collimator	1 mm ∅
Integration time	90 s
Sample container (film used)	PP film, 5 µm
Instrument: IRAffinity-1S + AIM-9000	
Resolution	8 cm^{-1}
Accumulation	40 times
Apodisation function	Sqr-triangle
Detector	MCT (mercury cadmium telluride)



Element	Ca	P	Cl	Zn
Quantitative value (%)	70.99	27.23	1.21	0.21
Element	**K**	**S**	**Sr**	
Quantitative value (%)	0.18	0.16	0.019	

Figure 2. ED-XRF qualitative and quantitative analysis results of human tooth.

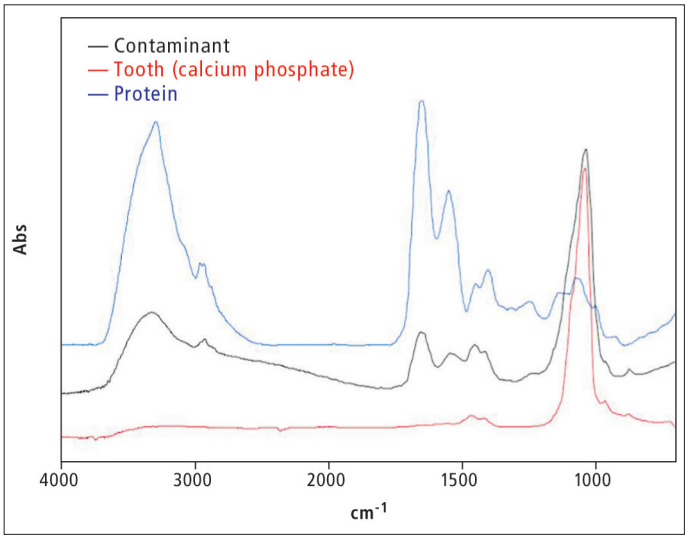
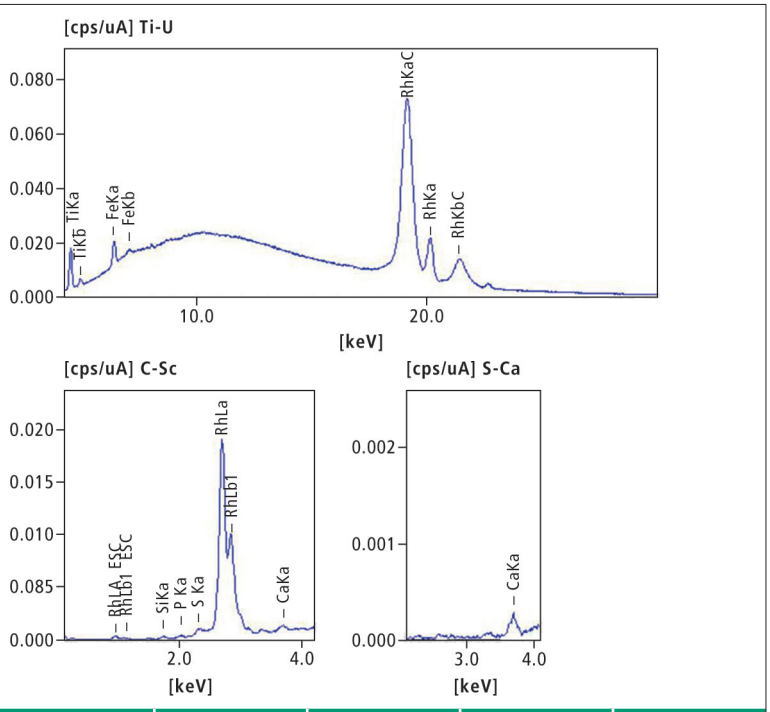


Figure 3. FT-IR absorbance spectrum of human tooth, contaminant and protein.



Element	Ti	Si	S	Ca
Quantitative value (%)	0.032	0.023	0.017	0.013
Element	**P**	**Fe**	**CH_2O**	
Quantitative value (%)	0.011	0.006	99.90	

Figure 5. EDX qualitative and quantitative analysis results.

Figure 3 shows the results of a qualitative analysis by FT-IR spectroscopy. Because the contaminant had a distorted shape and was extremely hard, a small amount of the sample was scraped off and a microscopic transmission measurement was conducted while holding the sample in a diamond cell.

The main component is calcium phosphate, and peaks originating from proteins can also be confirmed. The proteins are considered to be deposits on the tooth surface. When an adhering deposit exists, the sample is cleaned with water or ethanol, dried and then measured. However, care is required so that the contaminant is not lost before it can be analysed.

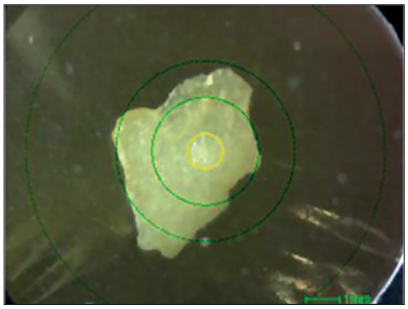


Figure 4. Photograph of contaminant (the yellow circle is the x-ray irradiation range: 1 mm ∅).

Artificial tooth (composite resin): PMMA and PMEA

In the second part of the experiment, a piece of an artificial tooth was analysed. A photograph (Figure 4) shows the contaminant, and Figure 5 shows the results of a qualitative and quantitative analysis by ED-XRF. Because the main component is organic material, a quantitative calculation was carried out assuming the balance is CH_2O. This is necessary since ED-XRF systems cannot analyse organic material, but can take the matrix effect of the organic composition into account for the elemental concentration calculation. A simplified description of the organic matrix, here CH_2O, leads to good quantitative results.

Figure 6 shows the results of a qualitative analysis by FT-IR spectroscopy. A small amount of the sample was scraped off, and a microscopic transmission measurement was conducted while holding the sample in a diamond cell. An FT-IR library search of the spectra was performed, and hits were obtained for polymethyl methacrylate (PMMA) and polyethyl methacrylate (PEMA). Based on these results, the contaminant is considered to be a piece of artificial tooth (composite resin).

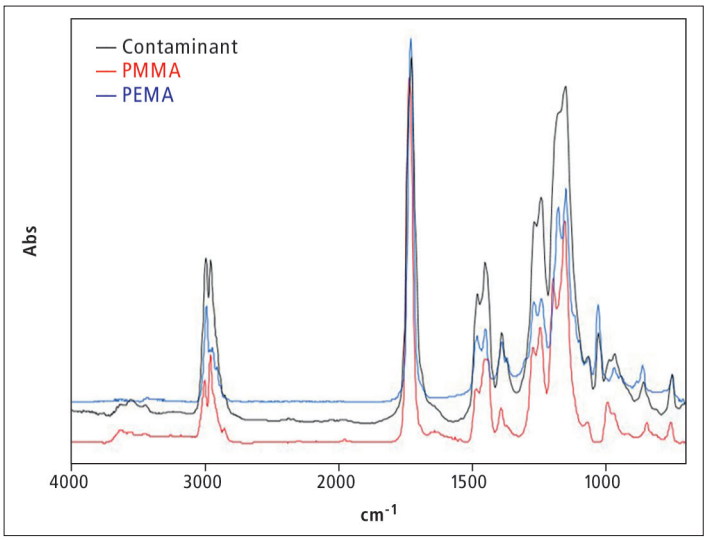
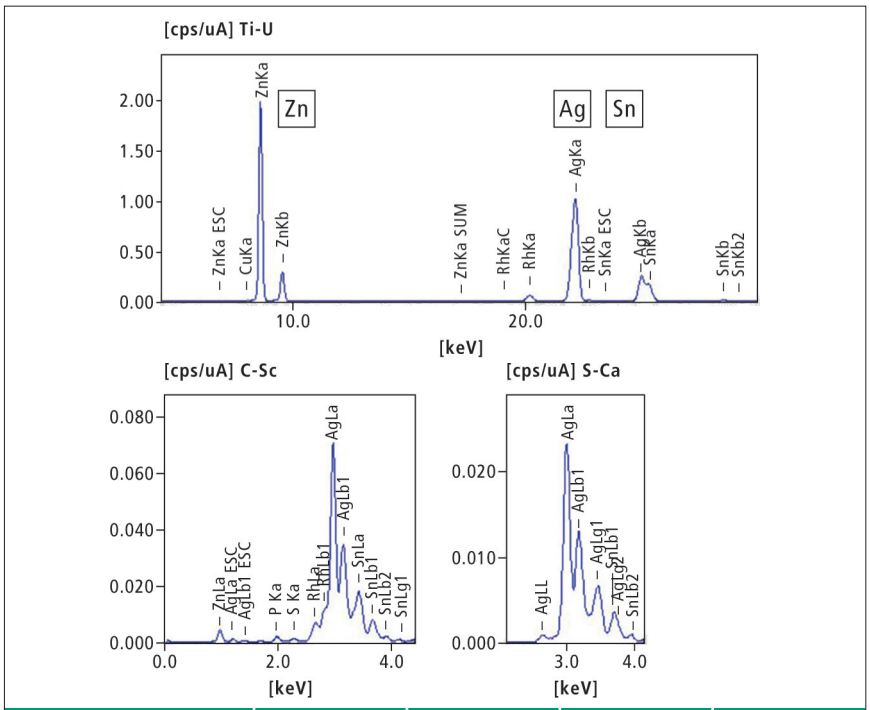


Figure 6. FT-IR absorbance spectrum of contaminant, PMMA and PEMA.



Element	Ag	Sn	Zn	P
Quantitative value (%)	61.81	22.08	15.59	0.38
Element	**S**	**Cu**		
Quantitative value (%)	0.088	0.054		

Figure 8. EDX qualitative and quantitative analysis results.

Artificial tooth (metal alloy) made of Ag, Sn and Zn

An artificial tooth piece was analysed in the last part of the study. Figure 7 shows a photograph of the contaminant, and Figure 8 shows the results of the qualitative and quantitative analysis by ED-XRF. The main components are silver, tin and zinc. No significant peaks were detected in the FT-IR analysis. Based on these results, the contaminant is an artificial metal tooth filling alloy. Figure 7 shows a piece of a silver crown.

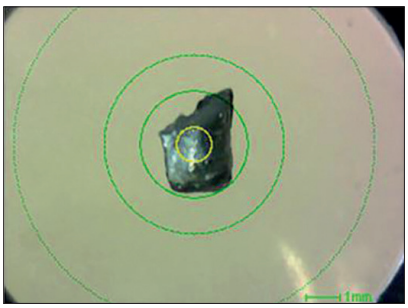


Figure 7. Photograph of contaminant (the yellow circle is the x-ray irradiation range: 1 mm ∅).

Conclusion

Contaminants in the form of a human tooth and various dental restorative materials were analysed by ED-XRF and FT-IR microspectroscopy. Due to the diversity of dental materials, which include hybrid materials containing organic and inorganic components, metal materials and others, a combined analysis technique using ED-XRF and FT-IR spectroscopy is very effective. Together, both techniques clarify the molecular and elemental composition of the food containment. Stainless steel, small stones, glass or rubber pieces can be analysed and differentiated quickly.

These techniques support food producers in detecting the cause of food contamination when they receive pieces found in food which are wrongly supposed to originate from the production process, enabling possible customer complaints to be clarified quickly. Since the methods are non-destructive, further analytical steps can be taken afterwards such as AAS, ICP-OES, ICP-MS and others. They provide more in-depth results, if required even down to concentrations in sub-ppm levels.

References

1. W. Bengel, "Amalgam-werkstoff und klinik", *Die Quintessenz* **20,** 1129–1144 (1990).

ARTICLE

13 Hyperspectral mid-infrared ellipsometric measurements in the twinkling of an eye

Karsten Hinrichs, Christoph Kratz and Andreas Furchner
Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V., Department Berlin, Schwarzschildstr. 8, 12489 Berlin, Germany
Contacts
karsten.hinrichs@isas.de (0000-0002-6580-7791)
andreas.furchner@isas.de (0000-0002-5448-7956)

Introduction

Fast, contact-less and destruction-free hyperspectral infrared (IR) techniques that enable large-area mapping within short measurement times are highly relevant for research and industry in environmental, biomedical, material and space applications. Laser-based methods provide high optical throughput as well as high spectral, spatial and temporal resolution, and are thus of particular interest for analytical, process, laboratory and field applications.

In this article, we focus on a recently introduced, rapid, laser-based hyperspectral method for thin-film analysis in the mid-IR fingerprint range. Featuring a polarimetric single-shot design, both phase and amplitude information related to refractive and absorption indices are simultaneously recorded. A tuneable, pulsed quantum cascade laser (QCL) enables spectrally highly resolved ($<0.5\,cm^{-1}$) ellipsometric mapping at lateral resolutions of ≤120 µm. High QCL tuning speeds provide access to time-dependent measurements of single spots in the µs/ms range, and fast hyperspectral mapping of large sample areas ($50 \times 50\,mm^2$). The single-shot approach ensures robustness with respect to pulse-to-pulse intensity variations and changes in environmental conditions like humidity. Hyperspectral and time-dependent phase and amplitude images facilitate the chemical identification of a specific material via its vibrational fingerprint, but also interpretation with respect to film thickness, molecular structure, composition and homogeneity of these parameters.

Laser-based IR imaging

Mid-infrared (MIR) imaging beyond classical Fourier-transform infrared (FT-IR) techniques recently showed promising developments for material science,[1–8] medical diagnostics and the study of biological samples.[9–11] IR spectroscopic applications are significantly broadened and strengthened by modern technical advances such as upconversion-based hyperspectral imaging,[4,9] hyperspectral nanospectroscopy,[1–3,5–7,10] frequency-comb techniques,[12] visible-diffraction limited optical photothermal IR multispectral imaging,[11] and QCL-based single-shot polarisation dependent hyperspectral concepts.[7,8,13]

IR imaging as a linear optical technique can reveal vibrational and structural sample information. Being contact-less, label-free and non-destructive, IR imaging can also be performed under varying environmental conditions. The latest advances in detection systems, optical set-ups and the use of brilliant MIR light sources open the door for new IR spectroscopic possibilities with respect to lateral and time resolution, structure analyses, as well as imaging capabilities in general. Particularly the involvement of QCL sources allows one to overcome the limits of classical FT-IR spectroscopy both in far-field and near-field IR-related techniques. Typical applications are time-resolved[13] and hyperspectral IR ellipsometry[8] (Sentech Instruments), photothermal AFM-IR[2] (e.g., from Bruker/Anasys), visible-diffraction limited spectroscopy[11] [e.g., optical photothermal infrared (O-PTIR) spectroscopy from Photothermal Spectroscopy Corp.], and scattering near-field IR spectroscopy[1,5] (e.g., from Neaspec and Bruker).

In this article, we focus on MIR laser-based polarimetric imaging (i.e., hyperspectral ellipsometry) as a new method to simultaneously reveal vibrational, structural and thickness information.[8,13]

Method of hyperspectral single-shot ellipsometry

The developed hyperspectral IR laser polarimeter incorporates a four-channel division-of-amplitude concept[14] that enables simultaneous single-shot amplitude and phase measurements at the laser repetition rate. Time-dependent measurements of individual spots (pixels) are possible with sub-decisecond resolution spectrally and sub-millisecond resolution down to 10 µs at a single wavelength. Fast hyperspectral amplitude and phase imaging is achieved by lateral sample mapping. Compared to classical IR ellipsometric imaging using FT-IR, measurement times are considerably reduced with this new laser-based ellipsometric technique.

The polarimeter (Sentech Instruments) consists of a tilt- and height-adjustable sample mapping stage (50 mm × 50 mm) and an auto-collimation unit for defined sample alignment. A pulsed external-cavity QCL (MIRcat 2100, Daylight Solutions) is employed as a brilliant, broadband-tuneable radiation source. A custom-built beamsplitter optics divides and directs the laser beam into four parallel detection channels, each equipped with a polarising unit (KRS-5 wire-grid polariser, Specac) and a photovoltaic InAsSb detector (P13894-211, Hamamatsu). A custom-built gated-integrator electronics synchronises the polarisation-state analysers of the four channels, thus enabling single-shot ellipsometric measurements.

By selecting specific polariser settings and additional optical elements like retarders, it is possible to tailor the four detection channels. This flexible approach allows one to measure various polarimetric parameters, for example, ellipsometric amplitudes ($\tan\Psi$), phases (Δ) or specific Mueller-Matrix elements related to the sample's polarimetric properties (e.g., circular/linear dichroism and birefringence).[15,7]

In this contribution, we focus on single-shot measurements of the ellipsometric angles Ψ and Δ. Using a different polariser setting α in each of the four detection channels, the single-shot ellipsometer measures Ψ and Δ from the intensities I_α according to:

$$\cos 2\Psi = \frac{I_{90°} - I_{0°}}{I_{90°} + I_{0°}},$$
$$\sin 2\Psi \cos\Delta = \frac{I_{45°} - I_{135°}}{I_{45°} + I_{135°}} \qquad (1)$$

Further technical details regarding the device and measurement protocol can be found in References 8 and 13.

The ellipsometric angles Ψ and Δ are functions of the incidence angle φ_0, the wavelength λ, the optical constants of substrate (N_S), ambient medium (N_0) and layers (N_j), as well as of the individual layer thicknesses d_j,

$$\tan\Psi \cdot e^{i\Delta} = \frac{r_p}{r_s} = F\left(\varphi_0, \lambda, N_S, N_0, N_j, d_j\right) \qquad (2)$$

j=0, 1, 2,… number of layers, with r_p and r_s being the p- and s-polarised complex reflection coefficients. The optical constants $N=n+ik$ (N: complex refractive index; n: refractive index; k: absorption index) are related to the complex dielectric function $\varepsilon=\varepsilon_1+i\varepsilon_2$ via $N=\sqrt{\varepsilon}$.

Measuring both Ψ and Δ yields complementary information regarding the sample's optical and structural properties, thus providing a means to in-depth sample analysis. Various routes are possible for evaluating the structure-related baselines and material-specific vibrational bands of hyperspectral IR ellipsometric data, including direct spectral interpretation, optical modelling and multivariate analysis.

Hyperspectral IR ellipsometry

Single-shot hyperspectral IR polarimetric imaging of a sample provides ellipsometric parameters dependent on wavelength and up to three spacetime coordinates. Regarding data interpretation, characteristic vibrational bands such as carboxyl, CH_x and amide bands[16] of lipids, proteins and polymers yield detailed insights into the film's properties.[13] Such data, therefore, enable unprecedentedly comprehensive, laterally resolved analysis with respect to molecular structure, chemistry, interactions, optical anisotropy, composition, morphology, film thickness and variations thereof.

In the following, we present two examples: one on the simultaneous phase and amplitude mapping of a heterogeneous surface and one on a time-dependent investigation of individual sample spots under external temperature stimuli.

Hyperspectral imaging of heterogeneous thin films

The first sample is a drop-cast thin film of myristic acid (MyA) [$CH_3(CH_2)_{12}COOH$] on a gold substrate.[13] The homogeneous, 150 nm thick film was partially chemically modified using droplets of NaOH dissolved in ethanol, resulting in areas of pure MyA, intermediate regions and areas of complexed-like MyA. Hyperspectral measurements of the film are presented in Figure 1, showing exemplary images and spectra obtained from the multidimensional data cube.

Figure 1a shows complementary phase and amplitude contrast images at different IR wavelengths. These images readily identify areas with varying thicknesses (mainly in cos Δ) and/or a chemical structure (mainly in tan Ψ). The chemical contrast image in Figure 1b shows the difference in band amplitudes of two vibrational markers at 1706 cm^{-1} (associated with C=O carboxyl vibrations of pure MyA) and at 1379 cm^{-1} (associated with complexed MyA), thus enabling the direct visual domain separation of pure and complexed MyA.

A spectral tan Ψ linescan along the red line in Figure 1b is presented in Figure 1c, highlighting the distinct vibrational characteristics of the different sample areas. Comparing the variations in spectral intensities reveals that regions of pure MyA are rather homogeneous, whereas those with complexed MyA exhibit pronounced heterogeneity in surface coverage. Pure and complexed domains are separated by an area with low coverage.

A detailed vibrational analysis of, for example, the ν(C=O) band can elucidate molecular interactions and orientations in such complex heterogeneous samples. Beyond such a spectral interpretation, a polarimetric approach can (as stated in Reference 15) "provide a number of contrast mechanisms besides traditional unpolarized radiation intensity, including linear depolarization, circular depolarization, cross-polarization, directional birefringence and dichroism." The high contrast resulting from the ellipsometric measurement could have tremendous application potential for the imaging and characterisation of complex surfaces (e.g., tissue classification for cancer identification, quality and morphology control of materials). In this regard, hyperspectral data analysis based upon neural networks, multivariate analysis and/or optical modelling could be a promising tool for gaining a detailed understanding of the structural and chemical specificities of the sample.

Together with the high stability concerning disturbing environmental

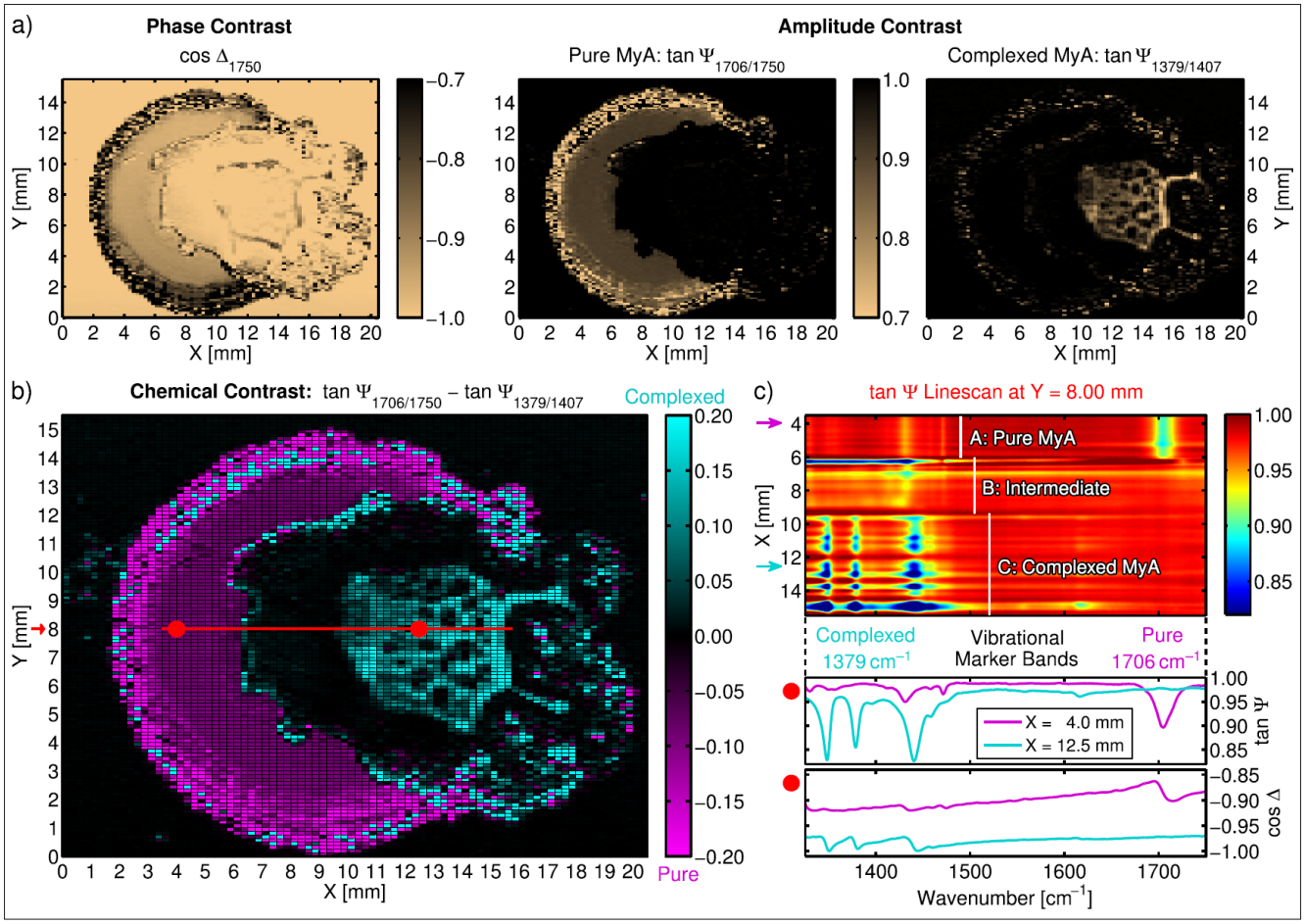


Figure 1. Hyperspectral IR ellipsometry images of a partially chemically modified, drop-cast MyA film. a) Phase and amplitude contrast at different wavelengths. b) Chemical contrast between pure and complexed MyA. c) tan Ψ linescan across the heterogeneous sample surface and exemplary ellipsometric spectra. Reprinted with permission from Reference 8 (©2019, Optical Society of America).

absorptions,[7] the capabilities of the IR laser ellipsometer make accessible multiple new IR spectroscopic imaging applications regarding the analysis of structured, anisotropic films and biological samples. Interesting systems to be investigated with the device range from functionalised surfaces and coatings, to polymer and protein materials, minerals, as well as solar cells, OLEDs and other optoelectronic devices and sensors.

Time-resolved studies of thin-film phase transitions

Spectra of single spots can be monitored within about 100 ms, and single wavelengths even as rapidly as 10 µs. The IR laser polarimeter, therefore, pushes new possibilities for time-resolved measurements of non-cyclic processes via flexible measurement durations from the µs to hour range. This point is illustrated in Figure 2, which displays the temperature-resolved investigation of the thermal phase transition of a 150 nm thick MyA film around 55 °C.[13] The amplitude and phase images in Figure 2b show the ν(C=O) band progression (cf. spectra in Figure 1c). Characteristic band components due to differently interacting C=O groups can be identified. A strongly interacting component (mode 1) is found in the solid phase at 50 °C, and a weakly interacting one (mode 2) occurs at higher wavenumbers mainly in the liquid phase at 60 °C.

Figure 2c shows the time-dependent development of the two ν(C=O) band components obtained from Figure 2b during slow heating. Figure 2d displays the corresponding single-wavelength monitoring of mode 1 during rapid cooling with a time resolution of 200 µs. While both mode 1 and 2 are present below the phase transition, only mode 2 is found above it. These observed changes reveal a tight interrelationship between molecular interactions and thermo-induced phase transition.

As these presented measurements demonstrate, the new IR polarimeter covers multiple time scales ranging from µs to minutes. The results highlight the applicability and sensitivity of the technique for time-dependent analyses of non-cyclic, irreversible processes and reactions. We see powerful applications in process and quality control, but also in rheology, relaxation and related studies. Because of the ellipsometer's small spot size (0.03 mm^2), low sample volumes and amounts can also be studied.

Currently, we are working on coupling microfluidic flow cells[17] to the instrument in order to image and analyse processes

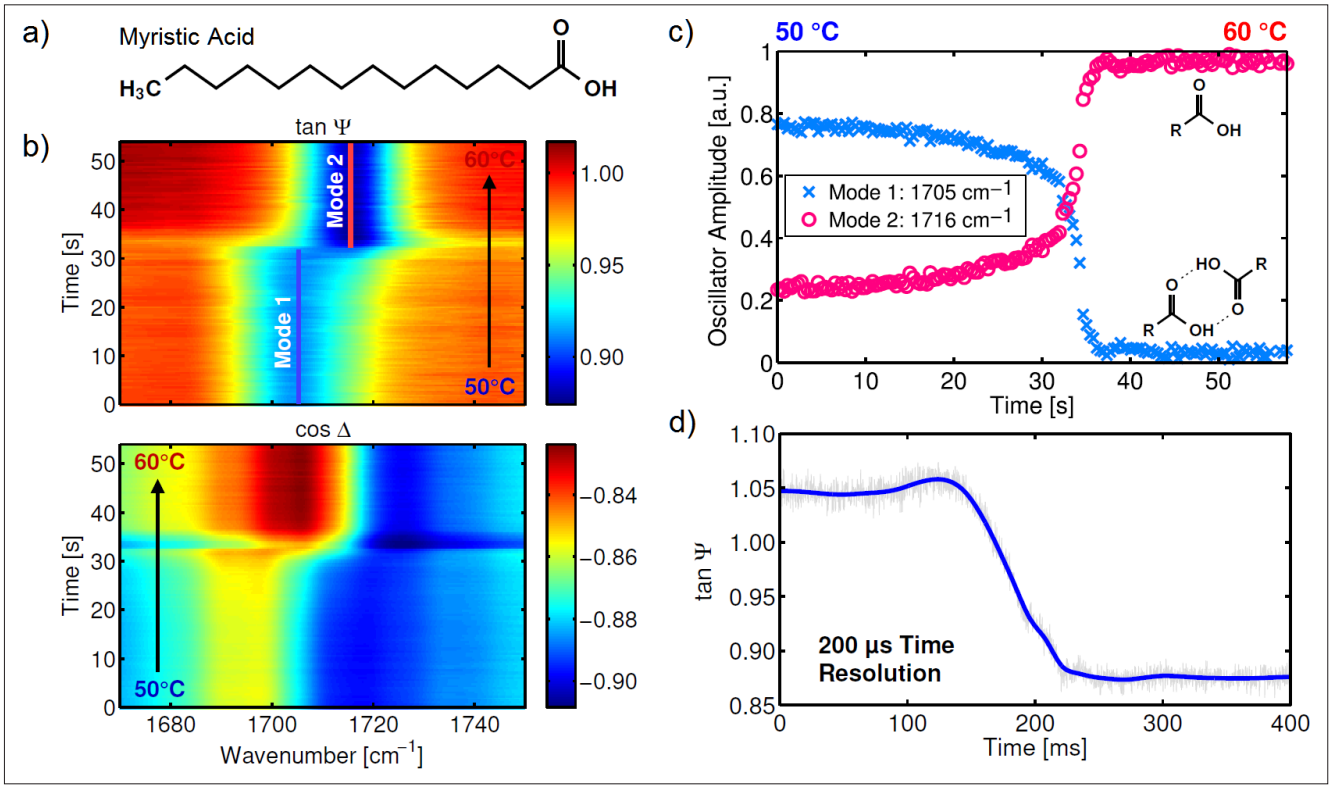


Figure 2. Time-resolved IR laser ellipsometry of the solid-to-liquid phase transition of a fatty-acid thin film upon heating/cooling. a) Chemical structure of MyA. b) ν(C=O) band amplitude and phase during slow heating. c) Oscillator amplitudes of mode 1 and 2 fitted from the spectral images. Reprinted with permission from Reference 13 (©2019, Optical Society of America). d) Sub-ms single-wavelength monitoring of mode 1 during rapid cooling.

of functional, sensor and biocompatible surfaces at solid–liquid interfaces. The set-up will then allow for sensing of molecular adsorptions, structural transitions and the study of intra- and intermolecular interactions of nL to mL samples.

A laser application laboratory will be opened at ISAS Berlin (anticipated for the end of 2020) operating and making available a hyperspectral IR laser ellipsometer.

Funding

Europäischer Fonds für regionale Entwicklung (EFRE) (1.8/13); Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen; Regierender Bürgermeister von Berlin – Senatskanzlei Wissenschaft und Forschung; Bundesministerium für Bildung und Forschung.

References

1. J.M. Atkin, S. Berweger, A.C. Jones and M.B. Raschke, "Nano-optical imaging and spectroscopy of order, phases, and domains in complex solids", *Adv. Phys.* **61,** 745–842 (2012). https://doi.org/10.1080/00018732.2012.737982
2. A. Dazzi and C.B. Prater, "AFM-IR: technology and applications in nanoscale infrared spectroscopy and chemical imaging", *Chem. Rev.* **117,** 5146–5173 (2017). https://doi.org/10.1021/acs.chemrev.6b00448
3. A. Huber, D. Kazantsev, F. Keilmann, J. Wittborn and R. Hillenbrand, "Simultaneous IR material recognition and conductivity mapping by nanoscale near-field microscopy", *Adv. Mater.* **19,** 2209–2212 (2007). https://doi.org/10.1002/adma.200602303
4. S. Junaid, J. Tomko, M.P. Semtsiv, J. Kischkat, W.T. Masselink, C. Pedersen and P. Tidemand-Lichtenberg, "Mid-infrared upconversion based hyperspectral imaging", *Opt. Exp.* **26,** 2203–2211 (2018). https://doi.org/10.1364/OE.26.002203
5. I. Amenabar, S. Poly, M. Goikoetxea, W. Nuansing, P. Lasch and R. Hillenbrand, "Mid-infrared upconversion based hyperspectral imaging", *Nat. Commun.* **8,** 14402 (2017). https://doi.org/10.1038/ncomms14402
6. K. Hinrichs and T. Shaykhutdinov, "Polarization-dependent atomic force microscopy–infrared spectroscopy (AFM-IR): infrared nanopolarimetric analysis of structure and anisotropy of thin films and surfaces", *Appl. Spectrosc.* **72,** 817–832 (2018). https://doi.org/10.1177/0003702818763604
7. K. Hinrichs, T. Shaykhutdinov, C. Kratz and A. Furchner, "Brilliant mid-infrared ellipsometry and polarimetry of thin films: toward laboratory applications with laser based techniques", *J. Vac. Sci. Technol. B* **37,** 060801 (2019). https://doi.org/10.1116/1.5122869
8. A. Furchner, C. Kratz, J. Rappich and K. Hinrichs, "Hyperspectral

infrared laser polarimetry for single-shot phase–amplitude imaging of thin films", *Opt. Lett.* **44,** 4893–4896 (2019). https://doi.org/10.1364/OL.44.004893

9. M. Hermes, R. Brandstrup Morrish, L. Huot, L. Meng, S. Junaid, J. Tomko, G.R. Lloyd, W.T. Masselink, P. Tidemand-Lichtenberg, C. Pedersen, F. Palombo and N. Stone, "Mid-IR hyperspectral imaging for label-free histopathology and cytology", *J. Opt.* **20,** 023002 (2018). https://doi.org/10.1088/2040-8986/aaa36b
10. D. Perez-Guaita, K. Kochan, M. Batty, C. Doerig, J. Garcia-Bustos, S. Espinoza, D. McNaughton, P. Heraud and B.R. Wood, "Multispectral atomic force microscopy-infrared nano-imaging of malaria infected red blood cells", *Anal. Chem.* **90,** 3140–3148 (2018). https://doi.org/10.1021/acs.analchem.7b04318
11. D. Zhang, C. Li, C. Zhang, M.N. Slipchenko, G. Eakins and J.-X. Cheng, "Depth-resolved mid-infrared photothermal imaging of living cells and organisms with submicrometer spatial resolution", *Sci. Adv.* **2,** e1600521 (2016). https://doi.org/10.1126/sciadv.1600521
12. K.L. Vodopyanov, E. Sorokin, I.T. Sorokina and P.G. Schunemann, "Mid-IR frequency comb source spanning 4.4–5.4 μm based on subharmonic GaAs optical parametric oscillator", *Opt. Lett.* **36,** 2275–2277 (2011). https://doi.org/10.1364/OL.36.002275
13. A. Furchner, C. Kratz and K. Hinrichs, "Sub-second infrared broadband-laser single-shot phase–amplitude polarimetry of thin films", *Opt. Lett.* **44,** 4387–4390 (2019). https://doi.org/10.1364/OL.44.004387
14. R. Azzam, "Multichannel polarization state detectors for time-resolved ellipsometry", *Thin Solid Films* **234,** 371–374 (1993). https://doi.org/10.1016/0040-6090(93)90288-Z
15. J. Qi and D.S. Elson, "A high definition Mueller polarimetric endoscope for tissue characterisation", *Sci. Rep.* **6,** 25853 (2016). https://doi.org/10.1038/srep25953
16. A. Barth, "Infrared spectroscopy of proteins", *Biochim. Biophys. Acta* **1767,** 1073–1101 (2007). https://doi.org/10.1016/j.bbabio.2007.06.004
17. C. Kratz, A. Furchner, T.W.H. Oates, D. Janasek and K. Hinrichs, "Nanoliter sensing for infrared bioanalytics", *ACS Sensors* **3,** 299–303 (2018). https://doi.org/10.1021/acssensors.7b00902

ARTICLE

3 The UV/Vis+ photochemistry database

Andreas Noelle* and Gerd. K. Hartmann
science-softCon, Auf der Burg 4, 63477 Maintal, Germany. E-mail: andreas.noelle@science-softcon.de

The *science-softCon UV/Vis+ Photochemistry Database* (www.photochemistry.org) is a large and comprehensive collection of extended ultraviolet, vacuum ultraviolet, ultraviolet, visible and near infrared spectral data and other photochemical information assembled from published peer-reviewed papers. The database contains photochemical data including absorption, fluorescence, photoelectron, and circular and linear dichroism spectra, as well as quantum yields and photolysis-related data that are critically needed in many scientific disciplines. This article gives an outline regarding the structure and content of the *science-softCon UV/Vis+ Photochemistry Database*. The accurate and reliable molecular level information provided in this database are fundamental in nature and help in proceeding further to understand photon-, electron- and ion-induced chemistry of molecules of interest, not only in spectroscopy, astrochemistry, astrophysics, Earth and planetary sciences, environmental chemistry, plasma physics, combustion chemistry, but also in applied fields such as analytical chemistry, medical diagnostics, pharmaceutical sciences, biochemistry, agriculture and catalysis.

Introduction

Photochemical data and information such as absorption spectra, fluorescence spectra, photoelectron spectra, circular and linear dichroism spectra, quantum yields etc. are important parameters needed in many scientific disciplines. Back in 1999, there was deemed to be a need for publicly accessible on-line databases containing such data and information in digital format (machine-readable). A first *UV/Vis Spectra of Atmospheric Constituents* CD-ROM[1] was published which contained, at that time, the largest collection of ultraviolet/visible (UV/vis) spectral data available free-of-charge. Based on this CD and the motivation to provide spectral data and information in digital format to the scientific community via the World Wide Web, the *UV/Vis Spectra Database* went on-line in August 2000 as a non-profit project.

In the beginning, the on-line database contained about 1200 spectra/datasheets for 120 substances and the compiled data extended beyond atmospheric research to allow for interdisciplinary application. To enable platform independent usability, both the spectral data as well as the datasheets (metadata such as publication, authors, source, wavelength range, temperature, pressure, phase etc.) are available as plain ASCII text. To guarantee the high quality of the fast growing *UV/Vis+ Spectra Database*, an international "Scientific Advisory Group" (SAG) was established in 2004, and the database was operated in accordance with the "open access" definitions and regulations of the CSPR Assessment Panel on Scientific Data and Information (International Council for Science, 2004, ICSU Report of the Committee on Scientific Planning and Review Assessment Panel on Data and Information).[2]

Since 2004, in addition to publishing the on-line database, every two years a mirror of the on-line database has been published on CD-ROM. The latest edition in the *science-softCon UV/Vis+ Spectra Database* series was published in 2019.[3] The on-line database currently (as of July 2020) contains about 14,200 spectra/datasheets as well as 5300 graphical representations for about 3000 substances, and is sub-divided into 28 substance groups (e.g. hydrocarbons, pharmaceuticals, pesticides, polycyclic aromatic hydrocarbons etc.). The database is updated weekly. In addition to the inclusion of new data, a main focus of the database is the preservation of data from older publications.

A more detailed description of the database and its applications has been published recently.[4]

As mentioned by the CSPR Assessment Panel on Scientific Data and Information, database maintenance and management are costly.[2] Collection of data, preparation of metadata and provision of professional data management expertise and institutional support for data dissemination and permanent archiving will add to the overall expense of specific research projects and maintaining the larger research infrastructure.

"Full and open access" to data implies equitable, non-discriminatory access to all data that are of value for science. It does not necessarily equate to "free of cost" at the point of delivery. There are several economic models for providing scientists with access to data for research and education.[2] The *UV/Vis+ Photochemistry Database* allows free and open access to all metadata, and cost-recovery pricing for data (or data licenses) in order to support the full data infrastructure. A choice of charged subscriptions giving full access to the data are available: for example an annual campus-wide licence provides full access to all data and information for less than USD 1 per day (for universities, governmental

organisations, non-profit organisations) and a "One-Time Registration" licence allows perpetual access to all data and information. Both licenses include a copy of the 12th edition of the *UV/Vis*+ *Spectra Database* CD-ROM.[3] In addition, those colleagues who support us in maintaining the database through the provision of new or missing data and information can get personal free-of-charge access to all data and information. More information is available at www.photochemistry.org.

Database structure and content

The database contains spectral information (gas, liquid and solid phase) from the extreme ultraviolet to the near infrared spectral regions (EUV-VUV-UV-Vis-NIR) and related data (e.g. information concerning publications on quantum yield studies or photolysis studies) from published peer-reviewed papers. Besides absorption spectra, which comprise most of the available data, fluorescence spectra, photoelectron spectra, circular and linear dichroism spectra, quantum yields etc. are available. The database is structured into 28 categories, which just provide a rough classification.

The datasheets provide metadata (substance name, formula and CAS number, data source, full reference, including title, authors, journal and DOI when available, spectral range and resolution, temperature, pressure, phase etc.), as well as data in various forms obtained and presented in the literature. This includes, for example, absorption data measured over a specific wavelength/energy range in tabulated form. In many applications (e.g. quantum yield studies or photolysis studies), the absorption cross-section (σ) or the molar extinction coefficient (ε) at a specific wavelength (λ) are determined, and these single wavelength data are also included in the database. For many substances temperature dependent data are available.

Most of the available data are from published peer-reviewed papers (>99%), data presented at scientific meetings and conferences are also available (<0.5%), as well as data from PhD theses, reports and unpublished material (<0.5%). As an example of the database structure and contents, absorption data of carbonyl fluoride (COF_2) from three different sources are presented in Figure 1. The data sets are as provided by the authors or listed in the relevant publications. To enable platform independent usability, all data are provided as plain ASCII text.

More recently, almost 3000 graphical representations, mostly from older publications have been digitised and added to the database. In addition, we have converted more than 100 datasets from "floppy discs" and hence prevented these data from being lost as technology has evolved.

Since 2019 the database has been extended to include circular and linear dichroism spectral data (Figure 2). The absorbance curves were recorded with the electric vector of the sample beam parallel and perpendicular to the stretching direction of the polyethylene polymer.

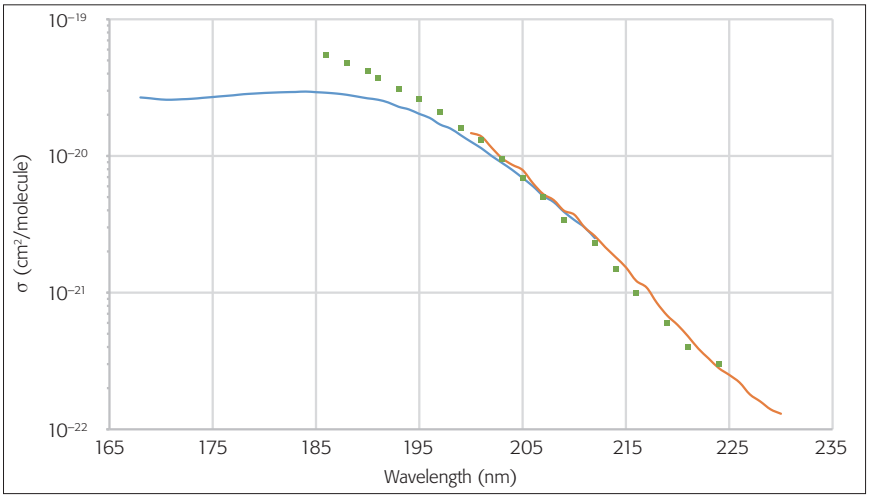


Figure 1. UV absorption spectrum of COF_2 obtained by Noelle[5] (blue curve), Noelle *et al.*[6] (red curve) and Molina and Molina[7] (green squares).

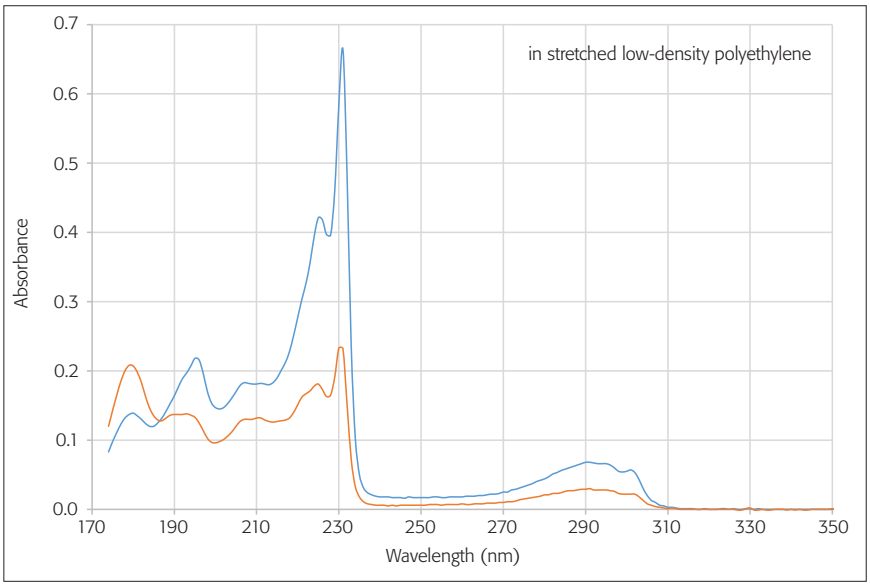


Figure 2. Linear dichroism absorbance spectrum of dibenzo-*p*-dioxin.[8] The absorbance curves were recorded with the electric vector of the sample beam parallel (blue line) and perpendicular (orange) to the stretching direction of the polyethylene polymer.

Outlook

The *science-softCon UV/Vis⁺ Photochemistry Database* is continually evolving and growing. As of July 2020, it contains about 14,200 spectra/datasheets as well as 5300 graphical representations for about 3000 substances and is sub-divided into 28 substance groups (e.g. hydrocarbons, pharmaceuticals, pesticides, polycyclic aromatic hydrocarbons etc.) This is a tremendous effort and requires a lot of manpower, not to mention technical infrastructure. We hope that the database proves useful to the scientific community and will facilitate their day-to-day work.

Since the support by the scientific community is crucial for such a photochemistry database, we would like to encourage all colleagues to assist us in maintaining the database and join our initiative "Photochemical Data and Information Sharing Platform—Share Photochemical Data & Information, Find Answers".

This initiative should develop the photochemical database towards a photochemical data sharing platform. The advantage of such a photochemical data sharing platform is that the more scientists provide their data for inclusion in the database, the better is the chance for all users to find specific photochemical data within the database. In addition, the platform becomes increasingly beneficial for use across multiple disciplines.

Database examples (datasheet, data, graph) are available at www.photochemistry.org.

Since the "UV/Vis⁺ Photochemistry Database" is operated as a non-profit "open access" database, any support from both sides, academic and commercial, would be highly appreciated.

Acknowledgements

The authors would like to thank the members of the SAG and the colleagues who support us in maintaining the database.

References

1. A. Noelle, F. Pätzold, S. Pätzold, R. Meller, G. Moortgat, E.P. Röth, R. Ruhnke and H. Keller-Rudek, *UV/Vis Spectra of Atmospheric Constituents.* DLR Publication (1998). ISBN: 3-89100-030-8
2. *Scientific Data and Information*. A Report of the CSPR Assessment Panel on Scientific Data and Information, International Council for Science (ICSU) (2004). ISBN: 0-930357-60-4, https://council.science/wp-content/uploads/2017/05/PAA_Data_and_Information_report.pdf
3. A. Noelle, G.K. Hartmann, A. Fahr, D. Lary, Y.-P. Lee, P. Limao-Vieira, R. Locht, F.J. Martin-Torres, K. McNeill, J.J. Orlando, F. Salama, A.C. Vandaele, R.P. Wayne and B.N. Rajasekhar, *UV/Vis⁺ Spectra Database*, 12th Edn. science-softCon (2019). ISBN: 978-3-00-063188-7
4. A. Noelle, A.C. Vandaele, J. Martin-Torres, C. Yuan, B.N. Rajasekhar, A. Fahr, G.K. Hartmann, D. Lary, Y.P. Lee, P. Limao-Vieira, R. Locht, K. McNeill, J.J. Orlando, F. Salama and R. Wayne, "UV/Vis⁺ photochemistry database: Structure, content and applications", *J. Quant. Spectr. Rad. Transfer* **253,** 107056 (2020). https://doi.org/10.1016/j.jqsrt.2020.107056
5. A. Noelle, *Bestimmung von Quantenausbeuten und Absorptionsquerschnitten von COF_2 und COFCl im UV-Bereich*. PhD Thesis, University Frankfurt, Germany (1996).
6. A. Noelle, H. Heydtmann, R. Meller, W. Schneider and G. Moortgat, "Absorption spectrum and absorption cross sections of COF_2 at 296 K in the range 200–230 nm", *Geophys. Res. Lett.* **19(3),** 281–284 (1992). https://doi.org/10.1029/91GL02316
7. L.T. Molina and M.J. Molina, "Chemistry of fluorine in the stratosphere", in *182nd American Chemical Society National Meeting*, New York (1982).
8. D.D. Nguyen, N.C. Jones, S.V. Hoffmann and J. Spanget-Larsen, "Electronic states of dibenzo-*p*-dioxin. A synchrotron radiation linear dichroism investigation", *Chem. Phys.* **519,** 64–68 (2019). https://doi.org/10.1016/j.chemphys.2018.12.003

ARTICLE

1 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*

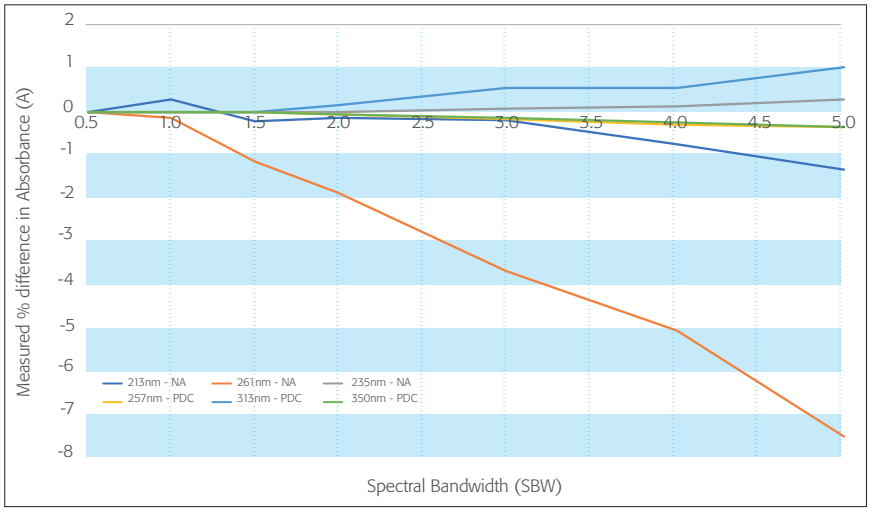


Figure 1. Effect of spectral bandwidth on measured absorbance values of nicotinic acid and potassium dichromate solutions. Nicotinic acid (NA) @ 213 nm and 261 nm vs acidic potassium dichromate (PDC) solution @ 235 nm, 257 nm, 313 nm and 350 nm.

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 2.5 A and 3.5 A, 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 1 A, 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

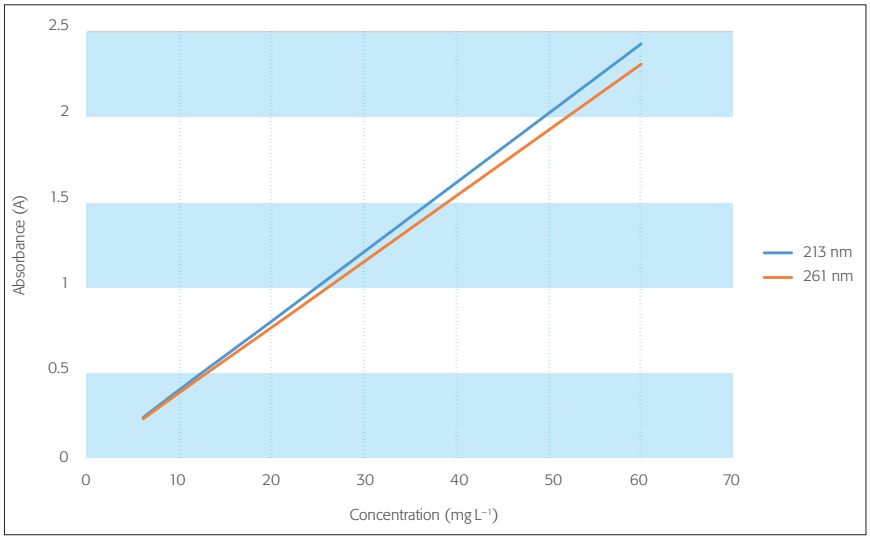


Figure 2. Nicotinic acid linearity.

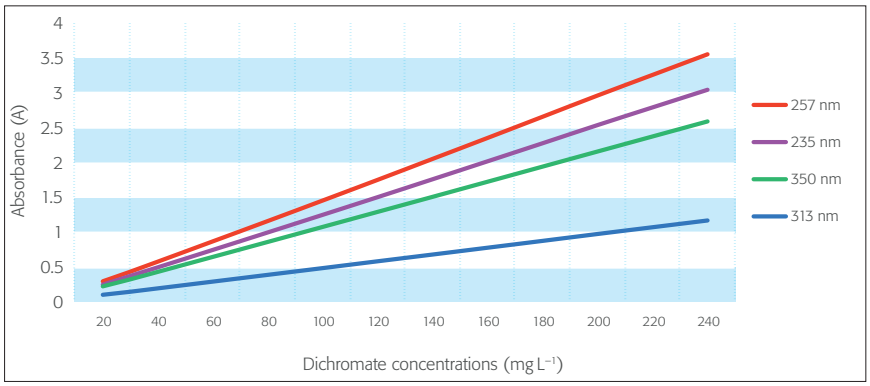


Figure 3. Potassium dichromate linearity.

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

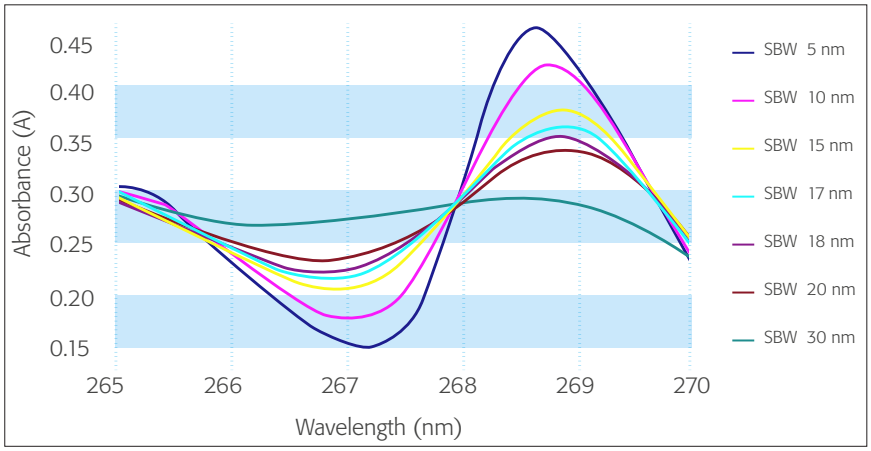


Figure 4. Spectra of 0.02% v/v toluene in hexane at different spectral bandwidths.

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,

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

each contributing 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

18 COVID-19: Lock-down and up-skill

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

With a significant proportion of our regular readership probably under home lock-down, we were wondering if we could help you at this difficult time by pointing out some useful online resources. So, when we finally come out of this pandemic, you could do so better skilled and more up-to-date than when we went in to it. Indeed, it is extremely rare in our modern lives that we get the chance to sit back and reflect on what we can be doing better, more efficiently and more collaboratively. Take the opportunity to download and test out some of the free versions of innovative data processing and reporting software that is available. There are also numerous very informative video tutorials available around new instruments and accessories, so go explore. This column will point you at a few resources you might like to use to start off. If you find some great resource yourself or feel we have overlooked something you yourself have put online, please let us know (and/or use the online commenting facility) and we will look to publicise the resource.

As many of us are home… let us start there

At https://www.spectroscopyeurope.com, Ian and the team have brought together articles, news, product information and events together by technique area.

While it can never be definitive, you might find some interesting leads to follow by looking at activities in techniques you currently don't have in your own laboratory. The content is a useful mix of high-level introductory content down to detailed discussions of specific research results in targeted problem areas—so something for most tastes.

With the demise of other sites integrating information across multiple vendors, this is a good starting point. The majority of instrument and software vendor websites unfortunately are now focussing more and more on promoting their specific products rather than helping potential new customers in understanding the various techniques in general.

It is worth looking for general interest areas within the offerings of the scientific publishers. One example from Wiley is the Essential Knowledge Briefings—a resource I must admit I had never heard of before, and these are quite nice relatively short summaries of different techniques with examples of their application which can be downloaded not only as PDFs but also in Kindle and Apple device formats for different eReaders (https://www.essentialknowledgebriefings.com). One little warning: if you are searching for articles you might like to read, be aware that the pre-defined filters under the box "Specialties" are a little arbitrary (Figure 1). Raman Microscopy is under the categories Imaging, Microscopy and Spectroscopy, which is fine, but the Dynamic Secondary Ion Mass Spectrometry briefing I wanted to read was only in Microscopy! What is very nice to see is that the reference lists for further reading with each briefing are not restricted to Wiley publications.

Instrument vendor resources

One of the instrument vendor websites which includes non-product specific training videos from basics to more advanced applications is hosted by Thermo Fisher Scientific (https://www.thermofisher.com/uk/en/home/industrial/spectroscopy-elemental-isotope-analysis/spectroscopy-elemental-isotope-analysis-learning-center.html). It also does not require you register to access the content.

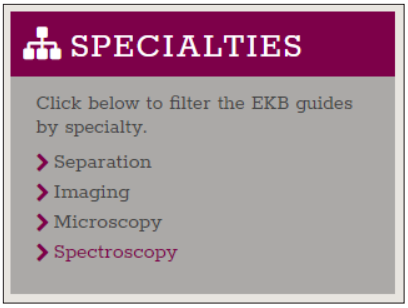


Figure 1. Preconfigured "Specialties" filters on the Wiley Knowledge Briefings site—make sure you use with care.

Their Spectroscopy, Elemental & Isotope Analysis Learning Center includes their Academy sections of introductory videos and webinars going right back to the first principles and explanation of the terms used without overloading the viewer with sales pitches for their products. Obviously as you move to more detailed webinars, such as the webinar on the Near Infrared Analysis of Biofuels for example, you will encounter more product-focused content. However, this is well mixed in with a good discussion of the benefits of the technique as a whole in this application area against other analytical approaches and as such is quite educational. The lessons you will learn are then equally applicable to a range of product solutions from other vendors.

TONY DAVIES COLUMN

For educators, there are also resources available in the Chemistry in the Classroom Resource Center with several pre-planned lessons and lab experiments. The individual links there lead to content ranging from simple single experiments to full multiple-experiment course plans or quite involved sample preparation, extraction and measurement practicals.

Agilent also hosts a vast amount of content within the Agilent University (https://inter.viewcentral.com/events/cust/catalog2.aspx?cid=agilent&pid=1&lid=1). Unfortunately, Agilent still require a lengthy personal registration process be completed to access any of this content and it appears that some of the content hides behind a pay-wall. If you persist and register yourself, then there is some very interesting educational content, although it does take some finding. Don't forget to tick on the "remember me" box or you will be forced to fill in all your details for every single content access.

Agilent's approach is different to Thermo's in the way the content is built with almost a complete dump of all their resources: generic information presentations as well as very targeted highly specific instrument operation training in one location. Before you give up in despair, you need to use the filters on the left-hand side of the screen, which are very useful in firstly filtering on "free" content and then drilling down to those subject areas you are actually interest in, rather than scrolling through pages and pages of links, see Figure 2.

Once you have mastered their search strategy, there is some excellent content, such as a long and detailed webinar on troubleshooting GC separation: very helpful for the large number of us with GC-MS systems delivering huge numbers of analyses. There is also content on the main Agilent website (https://www.agilent.com/en/training-events/eseminars). This is a different domain and your stored login information from the "University" won't be recognized, so you will have to repeat the process here. The content is made available in different ways, so for some content it is not possible to run the training in full-screen mode… for others the installation of, for example, the Cisco WebEx browser extension plug-in will be required. Try also https://www.agilent.com/en/training-events/eseminars as a route in, where you will get a list of future live webinars you can sign up for as well as access to the large library of recordings.

Bruker as you would expect also makes a lot of information available that you can benefit from looking at during your enforced absence from normal working practices (for example, https://www.bruker.com/se/service/education-training/webinars/optical-spectroscopy.html). Here I was rather confused by the "register to order" the webinar link for the recording I wanted to view. I then received an email confirming that I had requested the webinar. The reason appears to be the age of the webinar I was interested in viewing. Bruker makes use of YouTube to host their more recent webinars and if you follow the links they provide, this content is (obviously) available without the need to register. There are also some webinars hosted directly on Bruker.com. One drawback I found was that accessing their site, similarly to Agilent, initially presents content in a multitude of languages not making use of my browser's language settings. Automatically filtering in this way would be useful to control the amount of content initially presented to the user.

Now please don't just stick to the big multi-national vendors, there is plenty of very good educational content made available by the more specialist instrument manufacturers and software vendors; these smaller vendors have a tendency to keep their content more up to date.

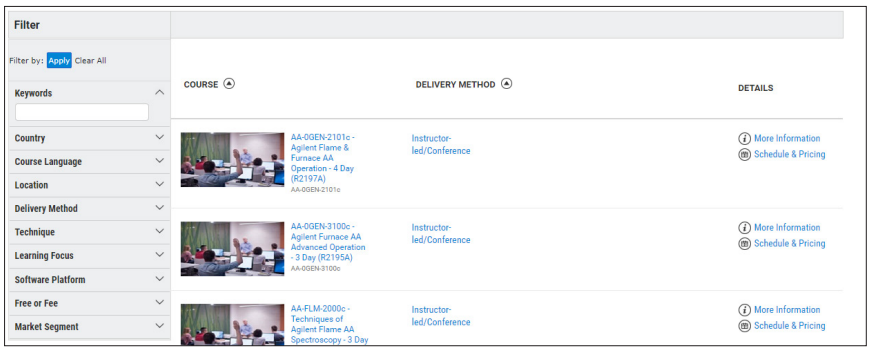


Figure 2. Agilent University webinars: use the Filter function on the left side to simplify identifying interesting content.

Download and try-out

So now you are overwhelmed with on-line training, let's move to an area where you can take what you learn and potentially improve the working environment for you and your colleagues when the world returns to normal. There are several vendors who allow you access to fully functioning copies of their software on a trial basis. This is a very generous offer, which I think you probably now have time to make the most of. I am sure that many of you have had the experience of being given a fantastic demonstration of a software product by an experienced sales person or software engineer, only later finding out that the operations you were shown only really worked on the test data provided by the vendor or the specific operations demonstrated in a predefined order they were shown.

Mestrelab Research, for example, will allow you to install a fully functioning version of their Mnova data processing suite (https://mestrelab.com/software/mnova/) for you to test as you wish. There are some very helpful videos supporting the functionality you will be trying out. Finally, are you bored of up-skilling yourself on your own? Have you been forced by COVID-19 to realise that your collaborative environments are well short of what they need to be? Are you still having to use that old USB stick to get data off your spectrometer… oh dear! You could look at another Mestrelabs fully functioning trial, of their cloud-based electronic laboratory notebook solution Mbook. This has

a free 45-day license and the capability of adding your co-workers into your test environment for a more realistic exploration. Once you have completed the registration, you will have a short wait for the environment to be set up after which you will be ready to enroll your colleagues for a truly collaborative trial—maybe role-playing different positions within your organisation.

Now there are many other vendors with trial or cut-down versions of their software available for free for you to install and test now, but this column certainly doesn't have the space to list them all. So, as mentioned above, if you find something you would like to recommend to others or have content that fits the criteria we are trying to use in this column then please get in touch.

Learn a new skill

Design of Experiments

Those who know me well will never forgive me if I didn't use this opportunity to suggest that you should also use the available time to look to getting some training in skills that will make you—whether academic or industrial spectroscopist—better at delivering your daily workload when the world gets back to normal. One area I have been preaching about has been the wider use of Experimental Design or Design of Experiments (DoE). You might be working on anything from driving efficiency in a multi-million euro chemical plant or finding the best settings for a new analytical method—adopting the DoE approach will enable you to deliver more robust solutions faster than if you attempt to carry out the work the old-fashioned way optimising one factor at a time. There are several strong vendors in this field, but I have been looking at Design Expert by StatEase, as our University has recently deployed it along with the chemometrics package Unscrambler from Camo AS. With our PhD students in lockdown, I have been suggesting they try out the freely available DoE training available from StatEase. As with other vendors they offer extensive paid training courses, but what caught my eye was some freely available training in what they call the Stat-Ease Academy (eLearning) (https://www.statease.com/training/academy/). Having signed up, the training starts off with a short (if embarrassing) test on your level of basic statistics knowledge. Going on at a nice slow pace you can take the PreDOE: Basic Statistics for Experimenters course which might be a little basic for some of you. However, it does provide some essential explanation of terms that you will use later in the courses.

If you want to try out the concepts as you go along and don't have access to the DesignExpert package, you can download a fully functional version of the package on a one-month temporary license to play with.

Project management

I would also like to take the opportunity to push for better Project Management skills, especially for those who do not primarily see themselves as project managers in the classical sense (i.e. spectroscopists!). For those working in research, development or any environment that involves change, having a basic understand of good project management techniques can certainly make your life simpler and more efficient. An additional benefit of learning to structure your projects well is the removal of any sort of blame culture when issues arise around delivery. Documenting the aims of a project at the beginning and having them signed off by your management against planned costs and timelines with milestones, binds all levels of stakeholders into a transparent team with clear roles and responsibilities and targets on delivery.

We successfully trained over 60 scientists, engineers and support staff in different levels of project management over a three-year period resulting in significant improvements in the planning, execution, documentation, quality and timely closure of their project work. We chose to standardise on the PRINCE2 Project Management qualifications (PRojects IN Controlled Environments https://www.axelos.com/best-practice-solutions/prince2) as PRINCE2 is very flexible and scalable, allowing the methodology to be tailored to meet our differing internal project demands. At the start, we focussed on classroom delivery style, as many of the concepts were new to the teams so the ability to question and discuss in the formal sessions and outside the structured delivery was important. However, more recently, and perhaps more relevant to this column, was the move to online delivery for the initial Foundation level qualification allowing colleagues to study at their own speeds.

Now if you want to obtain the official PRINCE2 certification, as we wanted for all our staff as part of their continuing professional development, you will need to get your boss to agree this is a useful use of your time stuck at home and pay for the course and examination fee. There are several online course providers so I won't go out on a limb to recommend any specific vendor and they will vary depending on where you are located. Regardless of the provider you select, the examinations are all centralised to ensure quality and conformity.

Conclusion

There is far too much content to list everything here, but please think a little out of the box when looking for educational material to make good the time you must spend at home in the coming months. Self-discipline will help in ensuring that when you are finally allowed back into a normal working rhythm you are still useful to your colleagues. Being self-critical or being honest about the strengths and weaknesses of one technique or another is quite rare in our field. One final recommendation could be the Metrohm webinars: https://www.metrohm.com/en-gb/support-and-service/webinar-center/ where (once they came back online) it was refreshing to see the speakers in the NIR spectroscopy webinar I watched actually going into details of when the technique they were "selling" was weak and should not be used. Nice work! Talking of NIR spectroscopy, seven of the series of ICNIRS conference proceedings are freely available at https://www.impopen.com/nir-proceedings ranging from 1995 to 2017.

TONY DAVIES COLUMN

12 COVID-19 #2: Compliant data processing from your home office

Antony N. Davies[a] and Mohan Cashyap[b]
[a]SERC, Sustainable Environment Research Centre, Faculty of Computing, Engineering and Science, University of South Wales, UK
[bc]MASS Informatics, Harpenden, UK

This column continues our theme of supporting working whilst unable to freely or safely access the analytical laboratory. We want to look at what advances have been made in systems allowing spectroscopic data processing from your home office. This has always caused particular problems for those working in highly regulated environments, such as the pharmaceutical industry, and their supplier and support contractors.

Definitions of Open and Closed systems, blockchain

In general, regulated industries have tried to avoid their IT environments falling into the "Open" category due to the increased requirements to ensure data is not capable of being tampered with. Within a well-protected company network this should not be a problem, as they are classical "Closed" environments under the definition. "Closed system means an environment in which system access is controlled by persons who are responsible for the content of electronic records that are on the system." In contrast, cloud provision can fall under the Open system definition. "Open system means an environment in which system access is not controlled by persons who are responsible for the content of electronic records that are on the system."[1]

Now, if you refer to the introductory columns on use of Cloud Computing,[2,3] we discussed various types of Cloud systems which were more or less likely to be able to meet compliance criteria. The FDA Open system adds additional burdensome requirements on the IT infrastructure and software solutions, such as encryption of the data (not just when in transit) and full electronic signatures. But the problems don't stop there, as the requirements for full training records for all systems staff to prove they are GxP compliant and up-to-date doesn't vanish when you outsource your IT infrastructure in an Open system… it just transfers the responsibilities to your external host/provider.

The big cloud hosting organisations claim to have regulatory compliant offerings, but if you approach them you need to know exactly what your strategy will be. For regulatory inspectors, the focus is more and more on data integrity. Who has had access to what and what did they do in your compliant environment is key to demonstrating data integrity and that your security features have been correctly installed and are operating fit-for-purpose. Where some cloud providers have issues is in making their system audit trails open for inspection, so this is a key area you, as customer, need to ensure you have what you need.

Advances in blockchain technologies is one area where we may be able to steal innovations driven by other sectors. Our requirements on automated audit trails cover all actions mandated for audit and signoff by our different regulators. These must remain secure even when data moves outside our immediate internal IT environment, a common critical underlying functionality in blockchain systems. Here the chain of data and the audit trail can be proven to be tamper-proof. But to exploit this we need not only better software systems but also improvements in our hardware environments to ensure no data leakage. Things to consider, for example, when allowing remote working include ensuring no password sharing, IP tracking can easily be spoofed if you want to beat the system, so you need better systems to detect data leakage, maybe something like the anomaly detection capabilities used by banks with similar problems.

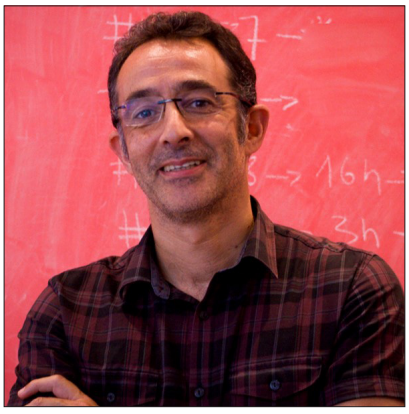
One of the essential tools for ensuring better data integrity is extensive automation of the data transmission and processing functions. This also leads to opportunities to support the work of the Quality Person in a regulated environment through the deployment of some levels of artificial intelligence to support the checking of, for example, study documentation. Scanning documentation for "obvious" or formal errors such as unlikely, incorrect or missing date, out-of-tolerance results, missing fields etc. can easily be automated. This does not replace the work of the Quality Person, but assists them by focusing their work on the anomalies in documentation they would normally have to find themselves. Taking this a step further, you can imagine working on the supporting analytical data and, for example, being able to identify where two spectra were so "identical" as to effectively mean it was impossible that they had come from two

separate QC measurements; thereby flagging possible errors or attempts at fraud.

Deployment examples of compliant cloud solutions

So where are we in terms of moving our systems safely into a cloud environment? Santi Dominguez, CEO of MestreLab, had some interesting comments which also reach back into the previous column[4] on making use of the extra time we have working from home to upskill: they have been running a series of additional free training workshops.[5] They saw that many people taking part were already using their software but were looking for additional training in more advanced spectroscopic data processing in areas they did not currently exploit. After years of this column complaining that the chromatographers are well ahead of the spectroscopists in the advanced level of IT system support with their chromatography data systems (CDS), it was interesting and nice to hear from him a feeling of obligation to develop similar advanced data handling and analytical workflow oriented support for spectroscopists…

"At Mestrelab we have been moving towards allowing our users to work remotely and freeing them from geographical restrictions. We see this access to data anytime from anywhere as being a critical part of the Lab of the Future. The design of our tools and solutions has had this idea at its heart for several years, and we are either there or getting there with most of the tools. The coronavirus pandemic has illustrated this by making remote work compulsory rather than desirable, and the amazing attendance we have had to our COVID19 workshops has shown the interest in the community in the value that this geographical flexibility offers."



Santi Dominguez

With LIMS systems being very much focused on standardised procedures or biased to handling chemical structures at their core, it would be great if spectroscopists would finally have a cloud-based enterprise application to support our work. Santi and his colleagues have taken this on board and are producing a system which can automate large parts of the "request>measure>capture_data>retrieve>process>report>archive" workflows we all use. He commented…

"With the technology available today, there is no reason why you should not be able to continue to progress your research and work just because you are travelling, at a conference or because a global pandemic prevents you from going to your workplace. It is up to us, as solution developers, to allow our users to transcend those geographical limitations, and this is at the core of our philosophy as a company."

We also had a really useful discussion with Heather Longden, a former colleague (of TD) who has a role as Senior Marketing Manager at Waters for Pharmaceutical Regulatory Intelligence, is a specialist in compliance to e-record regulations and an active member of ISPE GAMP Community of Practice, where she is called on as an expert in Data Integrity. In light of the working from home challenges today, Heather acknowledged that the Empower Cloud CDS has been adopted by a number of highly regulated laboratories. As Steve Bird, former Director of Informatics Strategic Marketing included in an Amazon Web Services (AWS) whitepaper...

"Users can sign on to Empower Cloud from any online computer or device, inside or outside of their organisation's network, using the same Empower credentials they would use at their desks or in their laboratories. This change significantly enhances their business continuity and data security capabilities while also ensuring their compliance and validation requirements are met."

Heather was very positive about putting scientific data processing systems into the cloud and had some positive stories where the deployment to the cloud used in an IAAS (Infrastructure as a Service) can actually greatly improve the compliance position of a company. Here I must apologise for citing a CDS system, but it does show what is now possible. Clearly there are additional challenges to solve for SaaS (Software as a Service) applications for regulated laboratories, but the IaaS model allows scientists to run dedicated individual single tenant solution on cloud infrastructure. In Waters' case, they have partnered with AWS as a cloud provider, and leverage automated AWS provided scripts to "install" the application, which is more reliable and consistent than an IT expert deploying applications on inhouse developed, on-premise infrastructure.

Heather did point out that when auditing or verifying your cloud providers understanding and delivery of GxP compliance requirement, be prepared to phrase questions in a way that IT provider's understand, discussing security and authentication, consistent installation etc, rather than IQ, OQ PQ, audit trails and data approval or batch release. Key to this is to ensure that you not only have understood the additional risks, and noting the mitigated risk, but that you also have clear documented agreement laying out who is responsible for what.

"... this is what the cloud provider is responsible for…".

She has been looking at the difference between US and European compliance, which is normally very closely aligned. However, an additional requirement in Annex 11 of the European regulations[6] is to "regularly review" audit trails (and consequently to have documented somewhere that this activity has been carried out).

The annex 11 is not a clear departure from Part 11. It explicitly clarified an expectation of both agencies that ALL critical data and meta data is reviewed.

TONY DAVIES COLUMN

Especially during the pandemic, a compliance trap has opened up with vendors being supportive by making additional licenses of their products available to people working from home. Although it might be obvious, just because the software is the same release version as what you have installed on your desktop computer in your laboratory, it will still need to be a validated installation on a validated computer system. So beware of trying to install your scientific data processing software onto the family's ultra-fast gaming PC… you might produce those 2-million datapoint surface plots really quickly, but you will not be able to use the results in a compliant manner!

I would like to finish off with an example Heather cited of the use of the cloud, not just to reduce costs in your IT environment but to exploit it to produce a much stronger compliant position where companies are working with external third parties such as contract research organisations (CROs) or contract manufacturing organisations (CMOs). Here, the contracting organisation uses an IaaS cloud deployment of their own to be the SaaS provider to their subcontracting CROs and CMOs (Figure 1). This reduces the worries about setting up and ensuring rock-solid Chinese walls with your subcontractors especially around data leakage.

Essentially, the subcontractors are carrying out the work for the contracting company in their own laboratories, but the instrumentation is run through the cloud software deployment of the contracting organisation. Again, a clear case where everything must be very well documented, but does eliminate many of the compliance hurdles associated with out-sourcing much of your new product development activities while maintaining an overall strong compliance position. Waters have a funny short video explaining all this much better than we can, which I would recommend watching if you have a spare four minutes.[7]

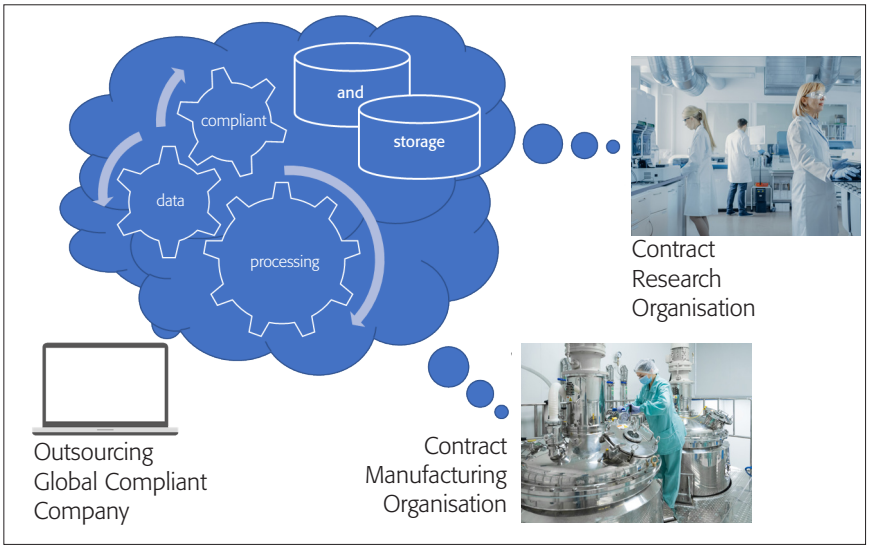


Figure 1. CxO organisations creating data which is acquired directly into the cloud-hosted enterprise application and owned by the outsourcing company.

Conclusions

So, thankfully, it seems that we, as a community, have moved substantially forward since our earlier articles on the introduction of cloud-based solutions. The solutions have addressed the compliance issues and seem to have started to actually deliver more flexible enhanced compliance positions over conventional deployments. If you have any good examples of such innovation yourself, please let us know and we will see if we can feature them in future columns.

Thanks!

Special thanks to Heather Longden at Waters and Santiago Dominguez at MestreLab for some very useful discussions and inspiration when putting this column together!

References

1. *Electronic Code of Federal Regulations, Title 21: Food and Drugs, Part 11—Electronic Records; Electronic Signatures*, e-CFR data is current as of 14 May 2020. https://www.ecfr.gov/cgi-bin/text-idx?SID=10b1b35c3a3ca71cf54d46dde50ab704&mc=true&node=pt21.1.11&rgn=div5
2. A.N. Davies and M. Cashyap, "A head in the clouds?—Part one", *Spectroscopy Europe* **26(3),** 21–22 (2014). https://www.spectroscopyeurope.com/td-column/head-clouds-part-one
3. A.N. Davies and M. Cashyap, "Who's ahead in the clouds?—Part three", *Spectroscopy Europe* **27(1),** 18–21 (2015). https://www.spectroscopyeurope.com/td-column/whos-ahead-cloud-part-three
4. A.N. Davies and H.-J. van Manen, "COVID-19: Lock-down and up-skill", *Spectroscopy Europe* **30(2),** 19–21 (2020). https://www.spectroscopyeurope.com/td-column/covid-19-lock-down-and-skill
5. COVID-19 Mestrelab workshops: https://resources.mestrelab.com/webinars/
6. EudraLex, *The Rules Governing Medicinal Products in the European Union, Volume 4, Good Manufacturing Practice, Medicinal Products for Human and Veterinary Use*. Annex 11: Computerised Systems (2010). https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-4/annex11_01-2011_en.pdf
7. https://www.waters.com/waters/en_US/Empower-Cloud-with-AWS/nav.htm?locale=101&cid=134950285

10 Planning for EuroAnalysis 2021

Antony N. Davies[a,b] and Lutgarde Buydens[c]
[a]SERC, Sustainable Environment Research Centre, Faculty of Computing, Engineering and Science, University of South Wales, UK
[b]Expert Capability Group – Measurement and Analytical Science, Nouryon Chemicals BV, Deventer, the Netherlands
[c]Chair EuroAnalysis2021, Radboud University, IMM Analytical Chemistry & Chemometrics, HG01.715 Heyendaalseweg 135, 6525 AJ Nijmegen, the Netherlands. E-mail: lutgarde.buydens@ru.nl

To run a top international conference that meets with approval from all attendees from industry as well as academia requires a highly skilled and experienced planning and execution team. It also is essential to start early with the planning, so I travelled down to Nijmegen to discuss how preparations for EuroAnalysis 2021 were coming along with Lutgarde Buydens who is chairing the event. One topic was how the community of *Spectroscopy Europe* readers could help in turning this into a very memorable event!

Ensuring EuroAnalysis 2021 addresses real needs

EuroAnalysis 2021 will be held over five days at the end of August at the beautiful conference venue, De Vereeniging, next to the historical town centre of Nijmegen. So apart from enjoying the famous Dutch hospitality, we should be in for some fantastic weather. The conference itself will be structured to give equal time to how analytical science is addressing societal needs, as well as to more developments in particular scientific analytical techniques. So, what areas can we expect to see focussed on?

Current planning sees three afternoon sessions dedicated to Societal Challenges. Table 1 shows the Societal Challenges from the old Horizon 2020 EU funding structure compared against the new (2021–2027 100 billion €) Horizon Europe successor programme "clusters" and "mission areas". More than half the total EU research budget is planned to be spent in the Global Challenges and European Industrial Competitiveness Pillar 2.

Unsurprisingly, Analytical Spectroscopy or advanced multivariate spectroscopic data processing does not seem to feature at the forefront of European thinking. However, I am afraid I am preaching to the converted when I state that without strong research programmes in analytical science, NONE of the Global Challenges or Industrial Competitiveness drives will get very far. As to the Mission Areas, it is clear that strong robust and verifiable analytical data must be at the heart of all scientific advances in all the Mission Areas.

The initial drafting of first Horizon Europe Work Programme based on the Strategic Plan is currently underway, but there is already more information around what will potentially fall under the Cluster headings (see Table 2).[1]

EuroAnalysis 2021 will be one of the first major networking opportunities for analytical scientists to get together after the new Horizon Europe funding stream goes live. If we look at the intervention areas, it is clear that the research we currently undertake can be applicable across many of the intervention areas: work on lab-to-the-sample is as applicable to Cultural Heritage as it is to Personalised Medicine.

We all know that the analytical chemist is today's most obvious example of a modern polymath. If we believe

Table 1. Comparison of Horizon 2020 Societal Challenges with Horizon Europe Global Challenges and Mission Areas

EU Horizon 2020 Societal Challenges	Horizon Europe Preliminary structures
■ Health, demographic change and wellbeing ■ Food security, sustainable agriculture and forestry, marine & maritime and inland water research, and the Bioeconomy ■ Secure, clean and efficient energy ■ Smart, green and integrated transport ■ Climate action, environment, resource efficiency and raw materials ■ Europe in a changing world—inclusive, innovative and reflective societies ■ Secure societies—protecting freedom and security of Europe and its citizens	**Pillar 2 Global Challenges and European Industrial Competitiveness Clusters** ■ Health ■ Culture, creativity and inclusive society ■ Civil security for society ■ Digital, industry and space ■ Climate, energy and mobility ■ Food, bioeconomy, natural resources, agriculture and environment **Mission Areas** ■ Adaptation to climate change, including societal transformation ■ Cancer ■ Climate-neutral and smart cities ■ Healthy oceans, seas, coastal and inland waters ■ Soil health and food

Table 2. Some more details on thinking around the content of Horizon Europe published in 2019.

Clusters	Areas of intervention	
Health	■ Health throughout the life course ■ Non-communicable and rare diseases ■ Tools, technologies and digital solutions for health and care, including personalised medicine	■ Environmental and social health determinants ■ Infectious diseases, including poverty-related and neglected disease ■ Health care systems
Culture, creativity and inclusive society	■ Democracy and Governance ■ Social and economic transformations	■ Culture, cultural heritage and creativity
Civil security for society	■ Disaster-resilient societies ■ Protection and security	■ Cybersecurity
Digital, industry and space	■ Manufacturing technologies ■ Advanced materials ■ Next generation internet ■ Circular industries ■ Space, including earth observation ■ Emerging enabling technologies	■ Key digital technologies, including quantum technologies ■ Artificial intelligence and robotics ■ Advanced computing and big data ■ Low-carbon and clean industry ■ Emerging enabling technologies
Climate, energy and mobility	■ Climate science and solutions ■ Energy systems and grids ■ Communities and cities ■ Industrial competitiveness in transport ■ Smart mobility	■ Energy supply ■ Buildings and industrial facilities in energy transition ■ Clean, safe and accessible transport and mobility ■ Energy storage
Food, bioeconomy, natural resources, agriculture and environment	■ Environmental observation ■ Agriculture, forestry and rural areas ■ Circular systems ■ Food systems	■ Biodiversity and natural resources ■ Seas, oceans and inland waters ■ Bio-based innovation systems in the EU bioeconomy

Wikipedia, then Johann von Wowern apparently defined polymathy in 1603 as "*knowledge of various matters, drawn from all kinds of studies... ranging freely through all the fields of the disciplines, as far as the human mind, with unwearied industry, is able to pursue them*", which seems as applicable to analytical scientists now as then.[2]

So, the conference structure of reserving the mornings for technology-focussed presentations will allow for the scientific developments in general to be presented. This should avoid the danger seen at many conferences of the sessions only being organised by application areas, causing many participants to miss key presentations of a fundamental development which they can apply equally well to their own research in a different field.

The big challenge for the hosts, the International Advisory Committee and the Scientific Committee is to get the mix right![3]

Who are the hosts and how can we help?

As I mentioned above, the conference is being chaired by Lutgarde Buydens, great for readers of this column due to the long-standing innovation from her and her team at Radboud University in the field of chemometrics and advanced analytical data processing! So, the pedigree of the responsible organisers is surely up to the task, but in our discussions it was clear that the organisers were keen to receive any feedback on the global challenges people or communities would like to see covered. Or maybe even groups who would be interested in using the event to get together at a wonderful location and organise their own sessions.

EuroAnalysis 2021 will be the 21st biannual meeting of the EuChemS Division of Analytical Chemistry, hosted by the Section Analytical Chemistry (SAC) of the Royal Dutch Chemical Society (KNCV) and COAST, the Dutch Community of Innovation for Analytical Science and Technology, who have been enormous supporters of innovative data handling and data processing projects in recent years. Many of which we have discussed in this column.

So please, if this short column has sparked any ideas about the topics raised that you would like to share either get in touch with myself, Lutgarde or maybe one of your friends in the organising committees—this is the best way to ensure that EuroAnalysis 2021 will focus on current topics of interest to you all!

References

1. https://ec.europa.eu/info/files/horizon-europe-investing-shape-our-future_en
2. K. Murphy, "Robert Burton and the problems of polymathy", *Renaissance Studies* **28(2),** 279 (2014). https://doi.org/10.1111/rest.12054
3. https://congresscompany.eventsair.com/euroanalysis-2021/commitees

4 When to automate spectroscopic data processing

Gary Sharman,[a] Marcel G. Simons[b] and Antony N. Davies[c]
[a]Mestrelab Research S.L, Feliciano Barrera, 9B-Bajo 15706, Santiago de Compostela, Spain
[b]Expert Capability Group - Measurement and Analytical Science (ECG-MAS), Nouryon, Zutphenseweg 10, 7418 AJ Deventer, Netherlands
[c]SERC, Sustainable Environment Research Centre, Faculty of Computing, Engineering and Science, University of South Wales, UK

I read with interest a recent article in *Chemical Science* originating out of Jonathan Goodman's group at the University of Cambridge. Jonathan is another long-standing IUPAC campaigner for scientific data standardisation and his group has been working on an improved solution to tricky nuclear magnetic resonance (NMR) spectra interpretation.[1] Their approach exploits modern higher processing speeds to enhance their fully automatic molecular structure elucidation software. Their DP4-AI uses the quantum chemical Gauge-Independent Atomic Orbital Density Functional Theory (GIAO-DFT) method calculations starting from chemical structures with undefined stereochemistry. ^{1}H and ^{13}C-NMR peak picking algorithms handle noisy spectra to predict relative stereochemistry. A statistical value is generated for the likelihood that each of the candidate molecules is correct based on the analysed spectra with almost no need for human intervention. This makes it an ideal tool to rapidly solve difficult problems like natural product library validation.

Clearly, there is still strong demand for improved NMR data interpretation and prediction software. I wondered how much such systems were being used on a day-to-day basis in industry, so talked to Gary Sharman, who has enjoyed a 20-year career in analytical science in the pharmaceutical industry and Marcel Simons, a very experienced NMR expert and one of my old colleagues at AkzoNobel/Nouryon.

Why do automation?

Many years ago, I heard a comment that has stuck in my mind and still raises a smile when I have occasion to remember it. One of the pharmaceutical industry customers of Creon·LabControl AG were testing an innovative combined ultraviolet/visible (UV/vis) and mass spectrometry (MS) automated approach for natural product library screening against "known chemistry" to select extracts for further work. After testing for a while, the customer explained to the software developers the reason behind his excitement. Completely ignoring the technological advances and clever programming that had gone into the system being tested, the customer simply pointed out that the automated spectroscopic data processing system effectively eliminated the boring repetitive work. Extracts that were of no interest (known chemistries) were automatically removed allowing him and his team to very rapidly focus on the extracts of interest that were potential new active molecules. "I can finally spend most of my time doing the expert job my company is actually paying me for".

So much for the thoughts that people increasing automation might be responsible for taking jobs away from spectroscopists! Gary Sharman highlighted three areas that can be seen as major drivers for better automation:

- Lost opportunities: problems that we would not even dare to start without automation.
- Free up time for more interesting work. We all became spectroscopists for the tricky, interesting problems, not to churn a handle on routine analysis and be bookkeepers. Let automation take care of the drudgery so you can focus on the fun problems. (Like the UV/vis–MS example above.)
- Less silly mistakes/book-keeping errors. We all like to think we are accurate and precise, but the fact is humans make lots of silly mistakes, particularly in collating data. Computers do not make these kinds of mistakes.

Have realistic expectations

The danger of having so much automation at our fingertips is that we might be setting ourselves up for some spectacular falls when the automation encounters problems it simply cannot master. You often see this in much simpler systems such as gas chromatography (GC)/MS database search results of electron ionisation spectra. We have discussed many innovative solutions in this column in the past, but time and again I see reports where the first database hit is cited as being the compound identified—even if the chemistry of the proposed molecule can have nothing to do with what is actually being worked on. If the scientist/student had taken the time to look further down the hit list they would have found a substance that made much more sense in terms of the experiments being undertaken.

So, as Gary put it… If you want perfectly assigned NMR spectra every time—give up now! A much better aim is

to really ask yourself what level of errors you are prepared to tolerate, and how that trades off against effort. For example, consider the quality control of a large library; without automation you may conclude it cannot be done. With automation perhaps we have 5% false positives. It is not perfect, but surely better than having no data on purity.

So, ask yourself what level or errors you are willing to accept. Be realistic. Everyone says "I want 100% accuracy", but not even an experienced spectroscopist can achieve that. You might make a trivial error like mixing up two samples or simply working on complex chemistries which you are unfamiliar with.

The automation process

Gary described the process in a similar way to Jonathan Goodman's group and this actually applies for different types of spectroscopy (Figure 1).

Although this might be seen as a rather simple schema, it is good to see how automation will benefit us at the various steps in the process.

- Data preparation and metadata extraction. Not to be overlooked—this may be one of the quick wins. For example, automatically finding and opening connected bits of data, looking up a structure and loading it, saving results—all parts that take time and are tedious bookkeeping, but every process needs them.
- Data processing such as peak picking and categorisation. This can be a very crucial part of the process. Many automated structure validation "mistakes" that are just down to poor peak picking of the data.
- Prediction—unless we are looking up a known thing in a database, we typically must predict the expected result to allow comparison. This could be quite simple (what is the expected ion for MS) or complex (a prediction of NMR by *ab initio* methods).
- Matching predicted to experimental. For some applications, this may be trivial: is the biggest peak in the mass spectrum the same as the *m/z* I expect. For proton NMR, with the complexities of coupling, overlap and higher order effects, it is exceedingly difficult.
- Scoring and output—we need to return a useful value that can be used to set actions. We might also want to return "quality factors" that indicate if the result is to be believed or if manual review is a good idea: these two things may well be orthogonal. A fail in the test may not mean the data needs review, and a pass may not mean it is a valid result.

Review by exception strategy

Although you may regard this as an oversimplification, manual analysis is "slow and accurate". Automation is often seen as "fast but error prone". By flagging samples for review where there is a reason to believe the automated result may be suspect, we can get the best of both worlds (Figure 2).

We do not work alone!

One of the critical questions which we are always asking is exactly how does some new wonder-software fit into our daily working practices and processes?

- The automation steps are only half the problem—how are you going to link your process to other processes in your organisation? This can make or break the automation. Workflow tools like the Swiss KNIME, the Konstanz Information Miner (a free and open-source data analytics, reporting, and integration platform)[2] or Biovia's Pipeline Pilot[3] can be valuable here. Also, having information exposed through APIs or web services makes integration easier.
- Constraints. You may have to work with legacy systems, other software with particular requirements or unhelpful interfaces to other data. This can be a major part of the problem that impacts design and implementation.
- The soft part—no one likes to be told by a computer they made a mistake. To get acceptance for a system, it may need thought about how people are informed of failures. For example, an e-mail saying you did something wrong with your boss copied in is probably a bad move. Flagging an error to an expert who reviews it and has a quiet word might be more accepted.
- New problems. Real world data is not perfect. Low signal-to-noise, poorly prepared samples and other components like residual solvents may lead to failures that a person would deal with as part of accepted normal practice.
- Edge cases. Software is built and validated on limited sets of test data. You can guarantee that over time edge cases will be detected that it does not handle well. Hopefully over time, more and more edge cases are dealt with and they become less and less frequent.

So, sticking with the world's COVID-19 theme, an Automated Structural Verification (ASV) software package like Mestrelab's "Verify" module can do an excellent job of assigning a molecule, such as a pharmaceutical active ingredient in a clean sample. Expecting a perfect assignment every time may be setting our sights too high. Imperfections do not stop a system being useful.

Enabling non-spectroscopist colleagues

Marcel Simons and colleagues have been working hard to help support colleagues from other disciplines in a speciality chemicals research and manufacturing area in a way that embodies many of the advantages listed above,

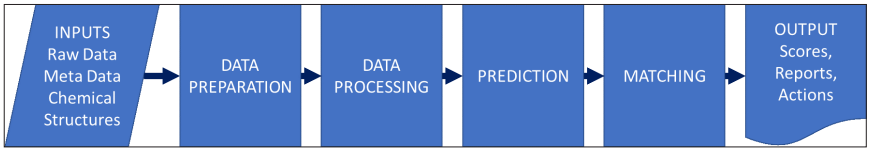


Figure 1. Parts of an automation process; not all processes have all parts. As well as the steps themselves, the inputs and outputs and their interfaces to other systems may be key to success.

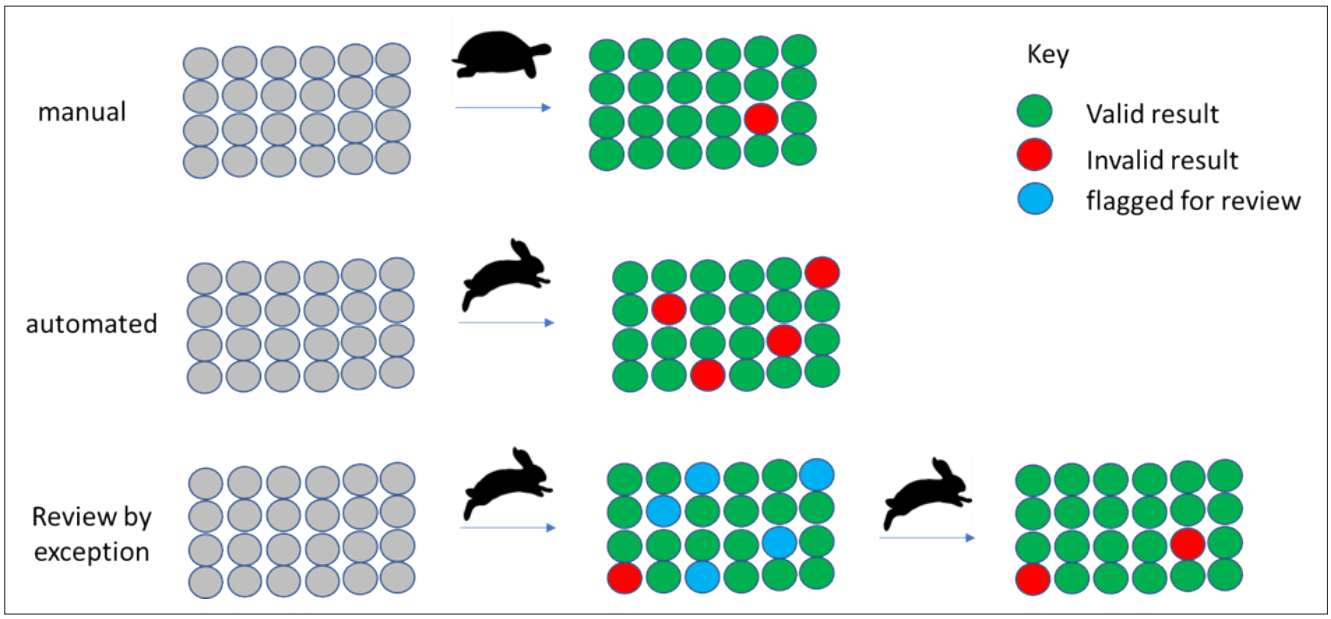


Figure 2. Automation supporting and reducing manual data analysis by focussing on the suspect results.

but in a quite different environment. Their challenges are far more to do with quantitative analysis by NMR rather than purely structure elucidation. The Expert Capability Group's open shop NMR has to cope with a very high workload of business- and time-critical samples—often being generated out of normal laboratory hours. They started configuring automated spectroscopic data analysis back in 2006. With instrument vendor support, they have developed and deployed over 30 automated methods that do tasks such as data processing for manufacturing plant support. These methods go well beyond the out-of-the box tools, and are designed to work using simple sampling strategies on all liquid samples with usable signals even without the use of deuterated solvents.

The automation results are basically processed spectra and a dedicated Excel file with the desired integrals and calculated molar ratios and/or calculated and normalised weight percentages. Depending on the targeted recipient of the automated processing and the demands of the specific business customers, conditional formatting is applied highlighting the results in green if the processing has delivered the expected result and red if the data is not what was expected and additional actions are potentially required (Figure 4).

Conclusions

So, it looks like there is a good clear case for continuing to develop faster and less error prone automated spectroscopic data processing. Jonathan's group have made their new software available under the Open Source MIT license, so if you feel like trying it out while you sit at home worrying about a second COVID-19 wave it can be downloaded from GitHub.[4]

Gary was one of the authors on a recent paper that pulled together many of the topics discussed here.[5] The paper discusses an automated system to verify new compound registrations. At its core was Mestrelab's Verify engine which automatically verified registered structures against their NMR and liquid chromatography-MS data. This was wrapped in a web service to make access by external processes simple. Bookkeeping tasks, scheduling and interfaces to other systems were taken care of by a KNIME server, and a streamlined review process was put in place to ensure there was a human face put on dealing with any problem samples.

Compound	m/m%	mol%	A/C
A	76.2	61.3	3.92
B	19.4	35.2	
C	4.4	3.4	

Compound	m/m%	mol%	A/C
A	66.0	49.1	2.38
B	27.7	46.4	
C	6.3	4.5	

Figure 4. At the end of a complex automated NMR data processing method, the customers question may boil down to "is the ratio of the concentration of two compounds within specific target boundaries to the quality criteria". In this figure, the results show a pass and the lower a fail.

References

1. A. Howarth, K. Ermanis and J.M. Goodman, "DP4-AI automated NMR data analysis: straight from spectrometer to structure", *Chem. Sci.* **11,** 4351–4359 (2020). https://doi.org/10.1039/d0sc00442a
2. *KNIME*. https://www.knime.com
3. *Pipeline Pilot*. https://www.3ds.com/products-services/biovia/products/data-science/pipeline-pilot/
4. https://github.com/KristapsE/DP4-AI
5. J.A. Lumley, G. Sharman, T. Wilkin, M. Hirst, C. Cobas and M. Goebel, "A KNIME workflow for automated structure verification", *SLAS Discovery* (2020). https://doi.org/10.1177/2472555220907091

SAMPLING COLUMN

5 Chemical analysis of contaminated soil for sound environmental site assessment. Part 1: the critical role of proper sampling

Jean-Sébastien Dubé, ing. PhD and François Duhaime, ing. PhD
Laboratoire de géotechnique et génie géoenvironnemental, Département de génie de la construction, École de technologie supérieure, 1100 Notre-Dame ouest, Montréal (QC), Canada H3C 1K3

Proper sampling of particulate matter for instrumental analysis is a common task in many applied scientific, technology and engineering fields. It is a crucial task for ensuring that measurements made on a given set of samples are representative estimate of the parameters of interest in the original sampling target. Unfortunately, sampling particulate matter is in many fields performed without a scientific basis, mostly because its critical role is ignored, or at best, misunderstood, and because of an unawareness of, sometimes a disregard for, the Theory of Sampling. This two-part column illustrates this important point using experience in the field of geo-environmental engineering.

Environmental site assessment guidelines require representative sampling, but do not define how: a recipe for decision-making disaster

A noteworthy example of how sampling is performed without a proper scientific basis is the sampling involved in environmental site assessment of *contaminated soil*. In this context, soil samples are analysed for their content of contaminants (chemical, physical). For chemical contaminants, analytical protocols generally require a few grams of soil for analysis only, and specify that this small quantity **must** be representative of the field parcel from which it is derived. This implies that a few grams of soil **must** represent a volume up to several hundred cubic metres of particulate matter in the field. This implies a mass reduction of nothing less than six to nine orders of magnitude, while ensuring that at each stage of the mass reduction process the resulting sub-sampled quantity of matter still represents the entire original soil parcel.

With the current state-of-affairs in this field (guidelines, standards, tradition, ignorance) this is a well-nigh impossible task. We find it incumbent upon us to sound a serious alarm within the field of geo-environmental engineering—but the examples and lessons described below have a much wider impact in many applications fields with *similar* heterogeneity issues.

The representativeness of an analytical measurement, i.e. the degree to which it represents the real contaminant content in the soil, compositionally as well as spatially, is directly related to the representativeness of the *sampling process*. This means the degree to which the proportion of each type of constitutive element of the soil, particles and contaminant(s) is preserved during the "from-field-to-analysis" sampling/subsampling process.

However, in the vast majority of current cases, the degree of representativeness is not assessed, far less even mentioned. In most guidelines for sampling of contaminated soils, representativeness is a vague concept, mostly owing to some form of wishful thinking. Without a formal **definition** of representativeness and guidelines on **how to obtain** a desired degree of sampling representativeness (called "fit-for-purpose reperesentativity"), sampling is performed more or less intuitively, haphazardly or based on subjective judgement. This approach is called grab sampling in the Theory of Sampling (TOS). It most commonly involves taking the desired mass of soil ("not-too-much") from some accessible part of the soil in *one increment*. In today's practice in the field, this would result in a grab sample of a few hundred grams which is sent to the laboratory, where a grab sub-sample of a few grams is then taken for analysis.

Below are two realistic, real-world examples of how this approach to sampling can produce extremely poor results.[a]

Assessment of zinc contamination

The first example is typical of a common situation in the practice of environmental assessment. A field sample from a site contaminated with zinc (Zn) was sent to an analytical laboratory by a geo-environmental consultant charged with the environmental assessment study. Field and laboratory sampling were performed by grab sampling. As per common practice, the laboratory was charged with providing a single analytical result from the material in the container delivered. This measurement resulted in a Zn concentration of 1900 mg kg^{-1}, thus indicating a contamination well above the regulation threshold for the current usage of the site (see further below). This result would lead to the demand that the soil from the parcel must be removed.

However, several *in situ* semi-quantitative measurements were also made by the consultant on the soil parcel with the use of a portable X-ray fluorescence spectrometer, and these had indicated a possibly smaller concentration.

Therefore, the consultant asked the laboratory for "a second measurement" based on the same sample container. This time the results came in at 79 mg kg^{-1}. Such a major discrepancy "naturally" prompted a third measurement, which, however, failed to detect **any** Zn in the soil! At the end of a very confusing day, a total of seven individual measurements were made based on the same 300 g soil sample as shown in Table 1.

Que faire?

As a way of trying to shift the burden of explaining these wildly varying results to the consultant, the laboratory concluded that the sample received was not homogeneous. Although this conclusion is correct, such a conclusion is profoundly naïve as **all** soils are heterogeneous, it is only a matter of to which degree (TOS).

Typical test pits in geo-environmental engineering site soil characterisation. One attribute rules the day: "significant heterogeneity". It is obvious that any single field sample (a grab sample in the TOS parlance) will not be able to represent the entire site. For this job, diligent compliance with the TOS' principle of *composite sampling* is necessary (see part 2).

This self-evident truth was exacerbated in the present case by severely "incorrect" sub-sampling in the laboratory (grab sampling from the *same* field sample container). So, whatever heterogeneity was revealed only pertained to the scale of the volume of the field container. Whether this is the same heterogeneity characterising the significantly larger site volume under investigation is still a completely open question: how well does the field container represent the entire site?

The applicable regulatory thresholds were 140 mg kg^{-1} (I), 500 mg kg^{-1} (II) and 1500 mg kg^{-1} (III), each value representing the maximum allowed Zn concentration in soil for specific usages of the site, or specific means of disposal of the excavated soil. Table 1

Table 1. "Autopsy" of a single 300 g field soil sample, and the resulting soil remediation status (categorisation).

Measurement	Concentration (mg kg^{-1})	Categorisation based on measurement
1	1900	>III
2	79	<I
3	<4	<I
4	<4	<I
5	<4	<I
6	700	II–III
7	25	<I

I, II and III represent regulatory thresholds of 140, 500 and 1500 mg kg^{-1}, respectively.

[a]For the record: the examples and procedures discussed here pertain to significantly heterogeneous materials that can**not** be subject to mixing before sampling. If a significantly heterogeneous lot to be sampled *happens* to be so small that it is economically feasible to mix it thoroughly in its entirety, the rules of the game have been altered because mixing leads to a significantly reduced distributional heterogeneity. However, the resultant lot is still compositionally heterogenous and still needs to be treated as such. Such cases are exceedingly rare, and consequently of overwhelmingly little interest within geo-environmental engineering.

SAMPLING COLUMN

also shows the categorisation of the soil with respect to these thresholds based on each of the seven "replicated" measurements.

It comes as no surprise that the consultant was now confronted with the confounding problem of correctly *categorising* the soil parcel represented by one field sample, but seven analytical results. If the categorisation decision had been made based on a single measurement, as is the usual practice, a highly significant error would have been introduced. This would have transferred unwarranted significant uncertainty to the site remediation process. The key issue is, of course, that under "normal practices" this would not even have been known to any of the stakeholders involved.

It would be hazardous to fit a statistical distribution to such a small dataset in which 43% of the data are left censored. However, it is possible to roughly *estimate* the categorisation probabilities based on proportions as shown in Table 2.

If the consultant had used the first measurement, as in current practice, he would have categorised the soil as larger than criterion III and, therefore, in need of disposal off site (A). But the probability that this decision would have been correct is only 14.3% (Table 2).

The consultant was, therefore, well advised to ask the laboratory for supplemental measurements. While these vary widely, a Kaplan–Meier (KM) estimate of the mean Zn concentration, 388 mg kg^{-1},[b] indicates that the soil could be categorised as being lower than criterion III, and thus kept on the site. This decision would have had an 85.7% probability of being correct. The problem of categorising the soil becomes more acute when the soil must be excavated and disposed of off site, since the disposal cost is related to the contamination level category. In the present case, based on the singular initial measurement, the soil would have been categorised as larger than criterion III, and disposed of at a larger cost, most probably incurring unnecessary expenditures from the site owner. However, based on the KM mean of 388 mg kg^{-1}, the soil would have been categorised as between criteria I and II, and thus disposed of at a much smaller cost or even reused as fill material in some jurisdictions. This example illustrates well how much uncertainty can be introduced in the decision-making process if based on a single 300 g field soil sample.

It can come as no surprise then that the documented uncertainty points to the highly likely situation that the target lot from which this single field sample originated must be significantly heterogeneous itself. The key issue is: is the single field sample representative of this target lot? To answer that, attention must be directed elsewhere: **how** was the primary sample (the field container) sampled in the field? Were the principles and rules in the TOS complied with, or not?

The situation depicted is common and typical, but it is not acceptable. The only

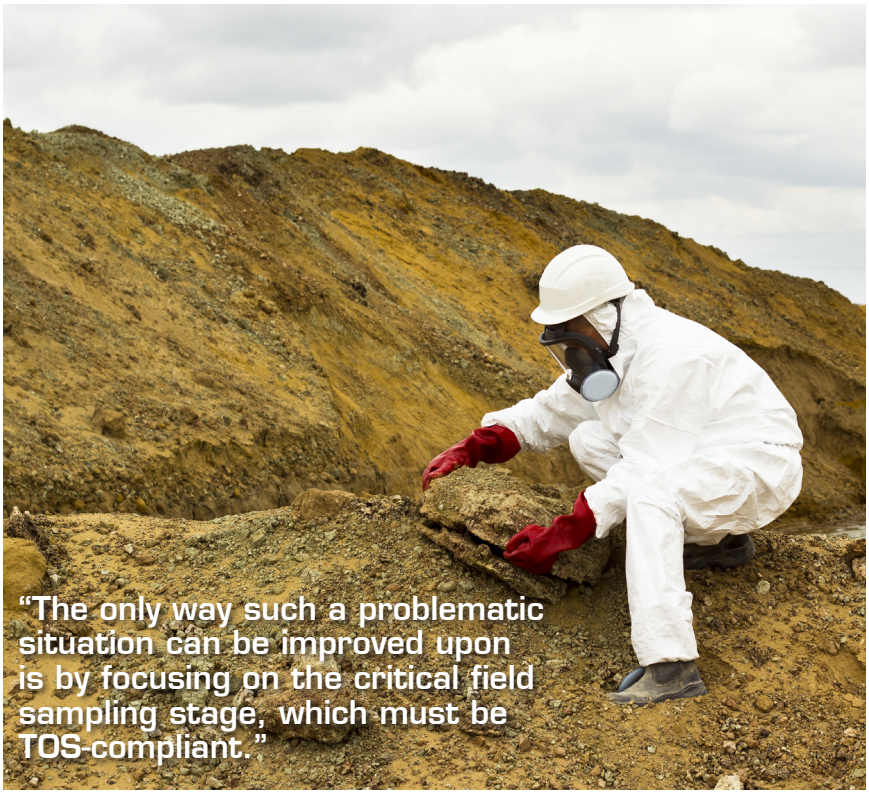


Table 2. Estimates of categorisation probabilities (categories A, B, C are identical to categories I, II, III in Table 1).

Category	Probability
x<A	0.714
A<x<B	0
B<x<C	0.143
x>C	0.143

[b]Note that calculating the mean Zn concentration by arbitrarily substituting the censored concentration measurements, i.e. <4 mg kg^{-1}, by 0 **or** 4, we obtain a mean Zn concentration ranging from 383 mg kg^{-1} to 388 mg kg^{-1}. While these estimates of the mean are close to the KM estimate in this case, arbitrary substitution in environmental datasets can lead to unreliable and biased estimates of descriptive parameters). Dennis Helsel (doi.org/fdmnj8) comments on arbitrary substitution: "There is an incredibly strong pull for doing something simple and cheap". This statement can just as aptly also be applied to grab sampling at all stages from field to aliquot.

Testimony

Understanding what sampling variation is, and how it is estimated, has been a "light-bulb" moment for our analysts after having been introduced to the Theory of Sampling (TOS) principles. So often we have had a situation where analytical work and results can be verified, but our customer still insists it doesn't meet expectations. Short of driving the poor analyst crazy with re-work tasks, which usually only produces the same "incorrect result", I now have an avenue of action that allows us to guide the customer and analysts to the path on how to focus on only taking representative samples. This is decidedly more welcome than always having to hear: "Take the sample back to the lab—repeat the analysis".

Much time is spent determining the combined total uncertainty for specific analytical methods under validation, however, very little attention is given to the preceding sampling errors and the challenges heterogeneity poses to this issue. I now know that sampling errors dominate over their analytical cousins. Also, using variographic characterisation as a quality control tool for process and measurement system monitoring is a very powerful technique that could help process controllers explain the sources of real process variations that occur on their product lines instead of simply following through by blaming the analytical lab. I found that the new international standard DS 3077 (2013) and in particular its use of illustrations and industrial examples captured the true complexity of the principal types of sampling errors and helped to conceptualise the TOS principles in a strikingly visual way, making it easier for a typical chemical analyst to relate to the scenarios involved before analysis. After all, we have to isolate the absolutely smallest aliquot for analysis—as demanded by highly sophisticated analytical instrumentation. It is, therefore, highly surprising that the one area of greatest error affecting analysts' results is the same topic largely ignored in Analytical Chemistry/Science Training programmes, again the sampling errors. This gives rise to "brilliant" analytical results, i.e. extremely precise results, but for non-representative samples for which accuracy with respect to the lot is not accounted for. In fact the accuracy of the analytical results with reference to the original lot is completely without control—and one cannot even estimate the magnitude of the sampling bias incurred (because it is inconstant, as is another insight provided by TOS). This makes for a very unsure analytical laboratory. After this course I wonder how many questionable results have been released by laboratories all over the world over many, many decades—and the revelations brought about by TOS are still not known!

Dr Melissa C. Gouws, InnoVenton Analytical, Port Elizabeth, South Africa

way such a problematic situation can be improved upon is by invoking a stronger focus on the characteristics of the full sampling process, notably the primary field sampling stage.

This case is also "representative" of the ill-informed practice of pouring more money into the analysis stage, i.e. making a larger number of measurements from each primary sample. Instead, more care should be taken in reducing variability at the primary sampling stage. It should not be difficult to understand that the debilitating heterogeneity revealed in Table 1 is only a reflection of the state-of-affairs in the singular field sample upon arrival at the analytical laboratory. No manner of repeated analysis based on this sample alone, can produce any information as to the real-world heterogeneity of the **entire** soil parcel, which **must** be larger but to an unknowable degree. The obvious solution is an appropriate deployment of *composite sampling* covering the entire 3-D parcel site.

Preliminary conclusion, Part 1

The first part of the *full* sampling-and-analysis process occur in the field and is often performed by the consultant's field technician. This gap in the "chain of custody" of the sampling process between the consultant and the laboratory is particularly problematic, especially as much as the current incorrect sampling practices are left without a clear responsibility. No one takes full responsibility for the representativeness of the complete sampling process in such circumstances.

Fix your sampling, not your results

In part 2 we will further illustrate how measurement variability can be controlled at the sampling stage with a second real-world example from a recent study conducted at École de technologie supérieure, Montréal, in partnership with the same consultant involved in the first example presented here. In this second study, we compare the uncertainty derived from grab sampling to that derived from a TOS-compliant composite sampling process.

References

A complete list of References will be included in part 2.

NEW PRODUCTS

17 Handheld FT-IR and SERS analyser measures of low concentrations of illegal substances

The Thermo Scientific Gemini analyser now comes with LowDoseID, specifically to addresses the rising trend of low concentrations of illicit substances. The Gemini combines Fourier transform infrared (FT-IR) and surface enhanced raman scattering (SERS) into one device to facilitate comprehensive and confirmatory chemical identification for substances with a concentration between 1% and 10%. The ability to detect small amounts of substances enhances accuracy and efficiency for first responders, particularly when identifying the presence of Fentanyl, heroin and cocaine.

In addition to new functionality through the introduction of LowDoseID, the Gemini has also undergone a software update. The version 1.8 software enhances ease of use with on-screen colour-coded alerts based on analysis results and an expanded factory chemical library, as well as smoother operation.

Thermo Fisher Scientific

http://link.spectroscopyeurope.com/32-082

11 Spectrum 3 FT-IR spectrometer

PerkinElmer's new Spectrum™ 3 FT-IR spectrometer covers the near, mid and far-IR ranges (11,000–30 cm^{-1}). There is automatic switching between beamsplitters without manual user intervention and the sources can be switched from IR to NIR at the touch of a button. Two sampling accessories can be installed simultaneously, avoiding the need to switch accessories between measurements. A fully integrated TG-IR (EGA4000) solution can be added, and the spectrometer is upgradeable to IR microscopy and imaging with automated switching of the beam to any PerkinElmer microscopy or imaging system. Experiments can be run directly from the SmartPanel on the instrument without returning to the PC, and IR data can be accessed from anywhere with cloud connectivity.

PerkinElmer

http://link.spectroscopyeurope.com/32-042

9 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.spectroscopyeurope.com/32-002

2 Quantification of Ethanol and Isopropanol in Alcohol-Based Hand Sanitizers

The most important parameter to consider in compounding hand sanitiser is the alcohol content. It has also been determined that hand sanitiser with alcohol concentration below 60% (v/v) is not effective and could leave the user at higher risk of infection. This application note describes the use of FT-IR spectroscopy to determine ethanol and isopropanol in hand sanitiser.

▸ Download the Full-Text PDF

Accelerate Your Publication

Letters published in 3 weeks subject to satisfactory review

JSI—Journal of Spectral Imaging is an Open Access, peer-reviewed journal dedicated to the rapid publication of leading research in spectral, hyperspectral and chemical imaging, as well as in related areas such as remote sensing, and chemometrics and data handling for spectral image data.

- *JSI* Letters offer extremely fast publication whilst retaining thorough peer review, with publication within three weeks of submission, subject to satisfactory review.
- *JSI* Letters can either express the view of the author on a particular aspect of spectral imaging that they wish to comment on, or can very briefly report an interesting aspect of the author's work: perhaps details of a chemometric routine or a sampling procedure. The conclusions should represent a significant advance in the understanding of an important issue and carry major implication for spectral imaging as a discipline.
- Authors are expected to submit revised manuscripts within 48 hours of receiving the referees' reports.
- Given the accelerated treatment accorded to Letters, it is particularly important that they are written to a standard of English to allow rapid refereeing.

impopen.com/jsi-author-centre