Chemical Engineering Research and Design 192 (2023) 510–520

Available online at www.sciencedirect.com

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherd

Development of a spatially offset Raman spectroscopy probe for monitoring pharmaceutical drying

Mais Al-Attili[a], Carla Ferreira[a], Chris Price[a,b], Karen Faulds[c], Yi-Chieh Chen[a,*]

[a] Department of Chemical and Process Engineering, University of Strathclyde, Glasgow, UK
[b] ESPRC Centre for Continuous Manufacturing and Advanced Crystallisation, University of Strathclyde, Glasgow, UK
[c] Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow, UK

ARTICLE INFO

Article history:
Received 24 October 2022
Received in revised form 16 February 2023
Accepted 23 February 2023
Available online 26 February 2023

Keywords:
In-line monitoring
PAT
Chemometrics
Critical quality attributes

ABSTRACT

Spatially offset Raman spectroscopy (SORS) is a subset of Raman spectroscopy devised for probing subsurface compositions in non-homogenous media. An example of such media is the wet filter cake during pharmaceutical drying. This non-homogeneity poses a challenge for process monitoring as it could render the determined solvent content during drying, and the end point, inaccurate. In this study, a SORS probe was developed for the monitoring of pharmaceutical drying. The probe includes a 45° illumination point and 0–5 mm equidistant collection offsets. Using the SORS probe, solvent signal detection through variable thicknesses of dry paracetamol was examined. The solvent content during the drying of paracetamol in anisole was then monitored using the SORS probe. Partial least squares regression (PLSR) analysis was applied to evaluate the performance of offset configurations in monitoring the solvent content during drying. The results showed that the solvent signal is detected through thicknesses beyond 6 mm of paracetamol from the larger offsets of 4–5 mm. A more accurate prediction of the solvent content was obtained from larger offsets. PLSR models using offset spectra showed a decrease in estimation error up to 50 % compared to backscattering spectra, with a further decrease upon using standard normal variate pre-processing. This suggests that SORS could offer improved monitoring of pharmaceutical drying processes.

1. Introduction

Spatially Offset Raman Spectroscopy (SORS) is a novel technique that is effective in probing the subsurface of heterogeneous diffusely scattering media (Matousek et al., 2005a). This technique is applied by collecting Raman measurements from positions laterally offset from the illumination point. This is in contrast to conventional backscattering Raman, where the scattered light is collected at the incidence point and contains features mostly of the surface layer. SORS spectra typically exhibit lower intensity than conventional backscattered Raman spectra as the light diffuses through the sample. However, this diffusion could increase the contribution in SORS signals from the sublayers, enabling the probing of larger sample volumes (Matousek et al., 2005b).

Abbreviations: SORS, Spatially offset Raman Spectroscopy; PLSR, Partial least square regression; LV, latent variable; PAT, Process Analytical Technology

* Corresponding author.
E-mail address: yichieh.chen@strath.ac.uk (Y.-C. Chen).

https://doi.org/10.1016/j.cherd.2023.02.041

Since its emergence, SORS has been studied in diverse fields, such as security, medical, and pharmaceutical applications, (Mosca et al., 2021) for authentication of products through packaging, (Nicolson et al., 2017; Eliasson et al., 2007) early diagnosis of breast cancer, (Stone et al., 2007; Ghita et al., 2018) and monitoring of changes in collagen concentration during bone healing (Dooley et al., 2020). In pharmaceutical applications, the technique was investigated to identify pharmaceutical tablet and capsule components, (Eliasson and Matousek, 2007; Matousek and Parker, 2007; Chao et al., 2017) and to provide quantitative analysis of pharmaceutical formulations (Johansson et al., 2007). The potential for SORS to achieve spatial and depth-resolved diagnosis from complex tissue structures has led to early efforts to develop SORS probes (Keller et al., 2011; Matousek, 2006; Schulmerich et al., 2007). However, those studies focus on the in-vivo biomedical application and, to the best of our knowledge, a systematic study on the development, evaluation, and deployment of a SORS probe for process monitoring has not been reported.

Pharmaceutical drying is a process in which a dry solid is obtained from an initial solid-liquid mixture, referred to as the wet cake, by removing the liquid part. The amount of residual solvent allowable in the dry active pharmaceutical product may be determined based on regulatory requirements, related to safety or stability considerations, and is a critical quality attribute of the pharmaceutical product (Aulton and Somavarapu, 2018; Murugesan et al., 2010; ICH, 2020).

During the drying process, the wet cake goes through multiple stages that could affect the physical and chemical stability of the dried active pharmaceutical ingredient (API) and by the end of the process, the initial solid-liquid mixture has changed in solvent content, packing density, and potentially particle size. Those possible effects on the product pose a challenge for the development and optimisation of pharmaceutical drying processes and necessitate the development of innovative techniques for the real-time in-line monitoring of the product throughout the process (Conder et al., 2017). Therefore, the development of an in-line method for the monitoring of the non-homogeneous medium during pharmaceutical drying is crucial.

Monitoring of the drying process was conventionally implemented through loss on drying methods (LOD) that include thermogravimetric analysis and Karl Fischer titration. Those methods involve compromising the process conditions in order to acquire a sample for the determination of the solvent content. More advanced process analytical technology (PAT) methods include gas chromatography, mass spectroscopy, nuclear magnetic resonance (NMR), and high performance liquid chromatography (HPLC), which are time-consuming and are not readily implemented in-line (Chanda et al., 2015).

Near Infrared (NIR) and Raman Spectroscopy, combined with chemometric analysis, have been used for monitoring the drying of pharmaceuticals (Nieuwmeyer et al., 2007; Kogermann et al., 2008; Märk et al., 2010; Peinado et al., 2011; Tewari et al., 2010; Hamilton et al., 2011; Maltesen et al., 2012; Kona et al., 2013; Fonteyne et al., 2014; Reddy et al., 2018; Peters et al., 2018; Gagnon et al., 2021). However, those studies use commercial probes that do not account for the non-uniformity of the monitored system in their signal collection geometry and signal extraction. Measurements using those probes either yield representations of homogeneous systems or an averaged representation of non-homogeneous systems. However, due to the non-homogeneity of the drying process as described earlier, it is clear that the process might start as a homogeneous system, but non-uniformity will naturally occur within the system. This will cause conventional measurements to fall short of describing the system. This phenomenon is embedded in the mechanism of drying as a result of the concentration gradient required to move the solvent to the surface of the powder bed and evaporate it. Due to the demonstrated advantages of probing larger volumes of non-homogeneous media, SORS presents a clear advantage for the monitoring of pharmaceutical drying.

The use of spectroscopic techniques for process monitoring is often coupled with applying chemometric analysis to draw out qualitative and quantitative information from the collected data. For the extraction of quantitative information from spectral measurements, regression may be applied to relate the response, the quantitative value, to the predictor, which is the spectral measurement. Partial least squares regression (PLSR) analysis is a widely used regression method. In PLSR, a set of latent variables are constructed based on both the spectral intensities obtained from spectroscopic measurements and the solvent content values in a calibration set for the prediction of the solvent content (Biancolillo and Marini, 2018).

Here, a SORS probe was developed for the in-line monitoring of pharmaceutical drying. First, a series of experiments were conducted to determine a suitable optical configuration to include in the SORS probe. The development process considered the intensity variations with the increase in offset distance, and the final design incorporated the use of additional optical fibres for collecting the signal from larger offset distances to maintain the signal quality at a comparable level. Then, the developed SORS probe performance was evaluated, and it was used to determine the depth of particulate samples through which solvent signal can be detected. Finally, the SORS probe performance was tested in the monitoring of solvent content during the drying of paracetamol, an active pharmaceutical ingredient, following washing with the wash solvent anisole. To assess the performance of SORS for monitoring solvent content, the collected SORS spectra were then analysed, and PLSR was used to develop models for estimating solvent content.

2. Materials and methods

2.1. *Materials*

The model system used in this study consisted of the API paracetamol (granular, 265 μm median particle size, Mallinckrodt Inc., Raleigh, N.C.) along with anisole (99 %, Alfa Aesar, Lancashire, UK), as the wash solvent, in which paracetamol displays minimal solubility (Ottoboni et al., 2020). The other material used in the development and testing of the SORS configurations was polyethylene terephthalate (PET) sheets (24 ×0.9 mm thickness, Covestro Ltd., UK).

2.2. *Benchtop SORS system for determining the SORS probe configurations*

2.2.1. *Optical setup of the benchtop SORS system*

The setup of the benchtop SORS system, shown in Fig. 1, includes a non-contact laser excitation source and a detector in a point-based geometry, as reported previously (Hopkins

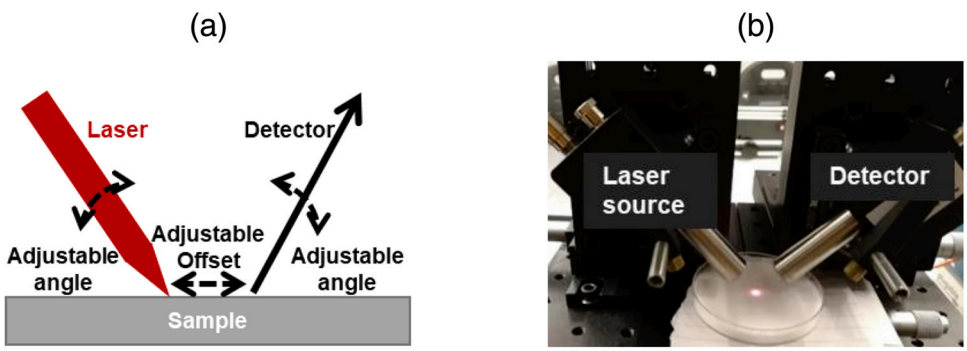

Fig. 1 – (a) Illustration of SORS with point-like illumination/collection geometry and (b) non-contact benchtop SORS setup.

Table 1 – Non-contact SORS combination of source/detector configurations.

Source/Detector notation	Source angle	Detector angle	Offsets tested /mm
SD 10°/0°	10°	0°	13–14
SD 30°/0°	30°	0°	6–14
SD 45°/0°	45°	0°	1–12
SD 45°/45°	45°	45°	1–12

and Pelfrey, 2012; Asiala et al., 2017). The setup offers variations in the incident angle of the laser and collecting angle for the Raman signal, as well as variations in the offset distance. A 785 nm laser (I0785MM0350MF, Innovative Photonic Solutions, USA) with a power of 100 mW, measured with a handheld power meter (PM100D/SC130C, Thorlabs GmbH, Germany), was used for illumination with an acquisition time of 1 s. A Raman spectrometer (WP 785, Wasatch Photonics, USA) was used to collect the signal. Four combinations of the source/detector (SD) angular arrangements were investigated for offset distances between 1 and 14 mm with 1 mm increments. A summary of the angular and spatial configurations is given in Table 1. The smaller offset ranges reported for 30°/0° and 10°/0° are due to the physical constraints of the incidence and collection probes.

SORS spectra of granular paracetamol were collected to evaluate the signal performance from different SD configurations and offsets. To assess the quality of SORS spectra from paracetamol beneath a barrier and to determine the maximum offset distance, a set of measurements was collected by placing a 21.6 mm layer of PET on top of granular paracetamol.

2.3. *Experimental setup for SORS probe*

2.3.1. *Optical configurations of the SORS probe*

A custom-made probe (FiberTech Optica, Ontario, Canada) was built based on the configurations determined using the benchtop SORS system. The probe consists of a single excitation fibre of 100 µm core diameter, which illuminates at 45° to the sample surface and produces a laser spot size of approximately 800 µm across. As illustrated in Fig. 2(a), a total of 19 fibres of 50 µm core diameter are used to collect the spectra from a range of offset distances between 0 and 5 mm (D1-D6) to the incident light with 1 mm distance between different offset configurations, where the 0 mm offset corresponds to conventional backscattering. Higher numbers of fibres are used to collect signals from the larger offsets to increase the signal throughput.

The probe was connected to a Raman spectrometer (RXN1, Kaiser Optical Systems Inc., USA) via a custom ferrule, where spacer fibres of 50 µm are inserted in the array of the collection fibres to separate the collection fibres for each of the offset distance groups, as shown in Fig. 2(b), to mitigate signal crosstalk. The 785 nm laser source used for the benchtop system was used to produce 100 mW of laser power at the illumination spot. The optical alignment of the SORS probe connection to the spectrometer was verified using an argon spectral calibration lamp (3060AR, 10 mA, Newport, USA) to obtain argon emission lines. Calibrating the SORS probe with the argon lamp is conducted by placing the probe

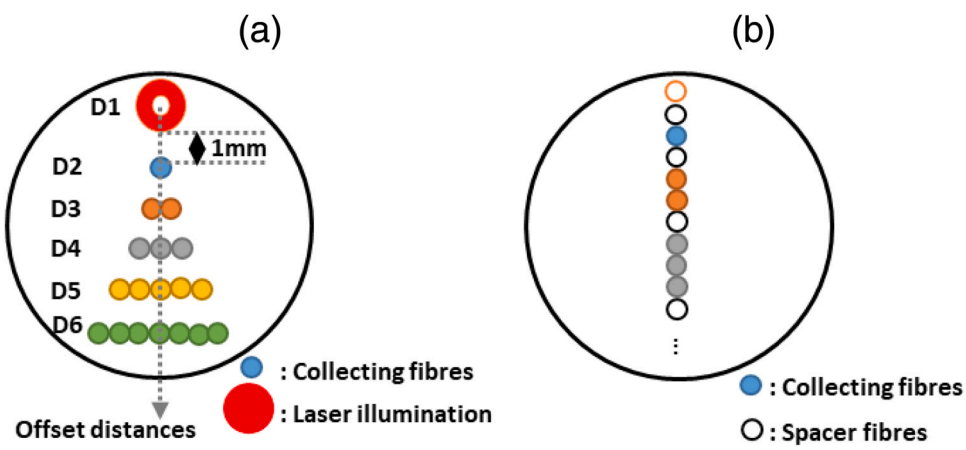

Fig. 2 – Schematic illustration of (a) the SORS probe field of view and (b) the custom ferrule end that connects to the spectrometer. Fibres correspond to different offset distances and are separated by a spacer fibre of 50 µm.

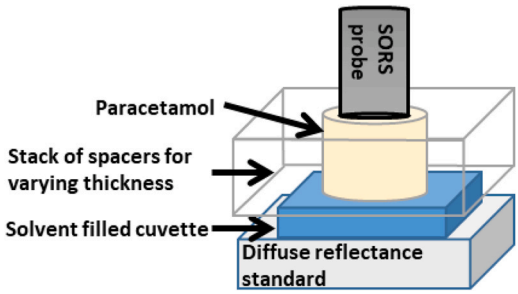

Fig. 3 – Depth of solvent signal detection setup using diffuse reflectance target underneath a cuvette filled with anisole and separated by multiple layers of equal thickness (1.2 mm) of granular paracetamol.

at the port of an integrating sphere (7N6322A, Newport, US) with the argon lamp placed in the sphere. The highly diffusely reflective internal coating of the sphere produces a uniform distribution of argon light to be collected by the probe. The Raman spectrometer consists of a two-dimensional 1024 pixel x 256 pixel detector. The full spectral range provided is between 780 and 1080 nm, with the grating of the system splitting them into two smaller regions of 780–920 and 920–1080 nm, corresponding to Raman shift in 0–1870 cm^{-1} and 1870–3480 cm^{-1}, respectively, for a 785 nm excitation wavelength. The signal was collected through this system as a two-dimensional (2D) spectral-spatial image. Using the detector control software (Andor Solis, Oxford Instruments, UK), the Raman signals from all 19 collection fibres were acquired using 20 s as the acquisition time in one spectral image (single scan), (Qin et al., 2016) as opposed to the sequential collection for each offset with the benchtop SORS system. In this study, the focus was on the lower Raman wavenumber region, as it covers most of the Raman peaks of the materials examined. The data collected through this setup was then processed using a script developed in MATLAB (R2020b, Mathworks, USA) to subtract the dark signal, sum the intensity corresponding to each offset configuration, and express the signal as Raman spectra.

2.3.2. *Depth of solvent signal detection*

To assess the depth through which the signal of a solvent can be detected, the probe was tested using a setup comprising a layer of dry paracetamol of variable thickness placed over a solvent-filled cuvette, as illustrated in Fig. 3. The anisole-filled 10 mm path length cuvette (Quartz SUPRASIL® 300, Hellma, Germany) was placed on a diffuse reflectance target (99 % reflective Zenith Polymer®, SphereOptics, Germany). A barrier layer of granular paracetamol with variable thicknesses was built on the cuvette using a stack of 10 spacer sheets, each of 1.2 mm in thickness. These spacer sheets have a hollow centre, which was filled with paracetamol crystals. An initial thickness of 12 mm of paracetamol was prepared to achieve uniform packing within different barrier thicknesses. Then, the thickness was reduced by removing the top spacer sheet with the excess paracetamol. SORS spectra were collected at each thickness with the probe in contact with the surface of the paracetamol. After extracting the signal collected from each thickness, baseline subtraction was carried out using the built-in 'msbackadj' function in MATLAB.

2.3.3. *Drying experiment setup*

The drying of the paracetamol and anisole wet cake was carried out using a 2-litre agitated Nutsche Glass Filter-Dryer (GFD® Lab 050 Series, Powder Systems Limited, UK). The jacket temperature of the dryer was set to 65°C and regulated through a heater/chiller. The powder surface temperature was monitored with an infrared sensor. Additionally, the cake temperature was monitored using a thermocouple sensor placed in the powder bed. Vacuum filtration was performed by means of a diaphragm vacuum pump, and the pressure within the vessel was monitored using a pressure sensor. This setup is shown in Fig. S1. (Supplementary Information).

Paracetamol was loaded into the dryer, followed by anisole, and the mixture was agitated at a speed of 5 revolutions per minute for 10 min. Next, the agitation was paused for de-liquoring with the aid of the vacuum pump. The de-liquoring process was stopped as the solvent reached the surface level of the solid of a fully saturated filter cake, which is the starting point of drying. The agitation and vacuum were maintained during the drying process and only paused for extracting samples from the mixture. The extracted samples were placed in a beaker for at-line SORS measurement collection. At-line measurements were collected as opposed to in-line measurements due to constraints of the port size of the laboratory-scale dryer. Those samples were then used for reference solvent content determination using a moisture analysing balance (MA160, Sartorius AG, Germany), which provides a reading of the solvent content as a percentage of the total mass of the wet sample. Then, the samples were taken for particle size measurements by laser diffraction (Mastersizer 3000 with a dry dispersion unit, Malvern Instruments Ltd, UK) to confirm that no change occurred to the particle size distribution (Chen et al., 2012). The parameters and protocol set out for the drying process aim to mitigate any particle breakage or agglomeration during the drying process and allow for the focus to assess the SORS probe performance in monitoring the solvent content.

2.3.3.1. *Description of datasets collected during drying.*

A total of 7 drying runs with various sampling intervals were performed to create a dataset with a range of solvent content. It was noted that the samples extracted at the start of drying contained excessive amounts of solvent, resulting in large variations in estimating the solvent content. Therefore, samples of a solvent content lower than 20 %, which are 69 samples, were used for the following quantitative analysis. This is of limited consequence because the samples collected later with a lower solvent content are more important in the quantitative analysis focused on the drying endpoint (Kjeldahl and Bro, 2010). Since each SORS measurement produces 6 spectra, corresponding to the offsets from 0 to 5 mm, the final dataset consisted of 414 spectra.

2.3.3.2. *Multivariate regression analysis.*

To quantitatively evaluate the performance of SORS measurements in predicting the solvent content, the collected dataset was analysed using partial least squares regression (PLSR) analysis. The PLSR analysis was performed in MATLAB, where the algorithm incorporates spectral range selection and signal pre-processing for optimisation of the analysis, as previously reported (Steponavičius and Thennadil, 2009). For PLSR analysis, the dataset was randomly split into a

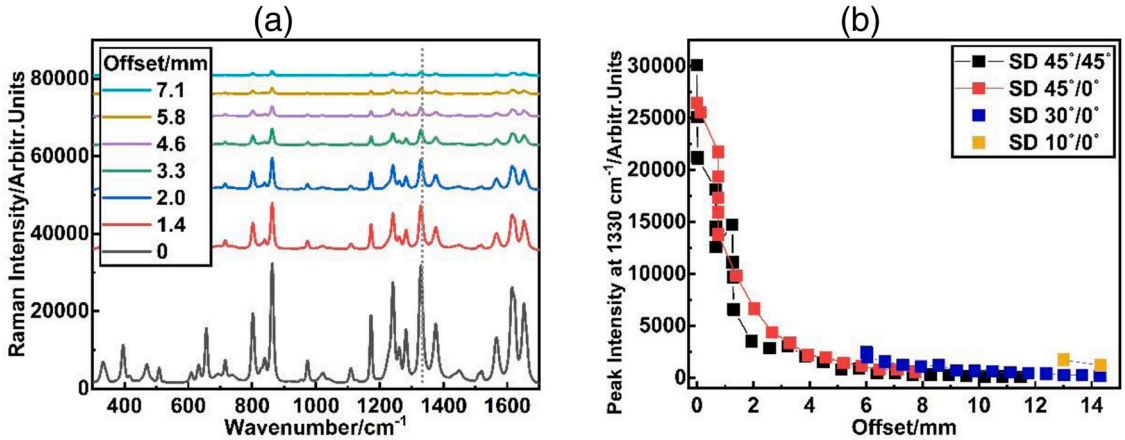

Fig. 4 – (a) Spectra of granular paracetamol from multiple offsets from SD 45°/0°, where the dashed line shows the peak at 1330 cm^{-1}; spectra intensity is offset on the y-axis for clarity. (b) Change in Raman signal intensity per offset distance using the different SD configurations.

calibration and a test set, consisting of 52 and 17 samples, respectively. Standard normal variate (SNV) pre-processing method was employed to remove the variation in the SORS spectra. For the calibration dataset, 5-fold cross-validation was employed to avoid overfitting the calibration models (Næs et al., 2002). The suitable number of latent variables (LVs) for the calibration model was determined by examining the root mean square error of cross-validation (RMSECV) curves, the coefficient of determination (R^2), and the prediction and residuals plots. The chosen models were then applied to the test dataset to obtain the root mean square error of prediction (RMSEP). The overall merit of the model is described with these performance metrics and the consistency between RMSECV and RMSEP.

3. Results and discussion

The results of this study are organised into three main parts. The first part describes the determination of a suitable source/detector (SD) configuration and offset distances to build in the SORS probe. This is followed by the characterisation of the developed SORS probe. The final part includes the analysis of the performance of the SORS probe in monitoring the solvent content during the drying of granular paracetamol in anisole.

3.1. *Development of the SORS probe from benchtop SORS setup*

Assessing the effect of SD configuration (Table 1) on the SORS signal was essential to determine a suitable configuration to be built into the probe, as well as determining the strategy to overcome the challenges of the weaker signal typically associated with signals from larger offsets/depths.

An example of spectra of granular paracetamol from different offset distances using SD 45°/0° is shown in Fig. 4(a). To assess the change in Raman intensity as the collection offset distance increases for each SD configuration, the intensity of the paracetamol peak at 1330 cm^{-1} was plotted against the detector offset distance, as shown in Fig. 4(b). The backscattering spectra exhibit the strongest intensity. For spectra from SD 45°/45° and 45°/0°, a sharp decrease in intensity was observed as the offset increased to around 4 mm. Measurements using SD 30°/0° and 10°/0° were constrained by the dimensions of the Raman probes of the benchtop system, which limit the offset ranges that can be investigated. Nevertheless, the 4 SD configurations suggest a consistent decrease in intensity, as seen in Fig. 4(b), where the intensity levels are comparable at larger offset distances. This can be attributed to the reduced photon density as a result of the diffusion of photons as they travel deeper into the sample (Matousek et al., 2005a). The incident excitation photons from the focused laser are almost unidirectional. However, the travelling direction of these photons changes due to the light scattering effect caused by the paracetamol particles leading to the diffusion of the photons. SORS spectra obtained at larger offset distances correspond to photons that travelled in various directions after having been scattered multiple times by solid particles. Due to the constraint in the setup of SD 30°/0° and 10°/0° restricting the use of shorter offsets distances, those two configurations were discarded from further evaluation.

To further differentiate the influence of the collection angle in SORS measurements, spectra of paracetamol under a 21.6 mm layer of PET were collected using SD 45°/45° and 45°/0°; these are shown in Fig. 5(a) and (c), respectively. Overall, a decrease in the PET peak intensity as the offset increases can be seen in spectra collected from both SD configurations. To assess the details of the changes in paracetamol and PET peak intensity with the offset distances, the intensity of the paracetamol peak at 1676 cm^{-1} and PET peak at 1640 cm^{-1}, highlighted by arrows in Fig. 5(a) and (c), were plotted in Fig. 5(b) and (d) and for SD 45°/45° and 45°/0°, respectively. As the light scattering effect of a particle is dependent on the frequency of the incident photons, (Larkin and Larkin, 2011) the proximity of the selected paracetamol and PET peaks can mitigate the difference in the light scattering conditions, reducing the complexity in analysing the spectral intensity.

Fig. 5(b) and (d) show the contrast in intensity between PET and paracetamol in their response to the increase in offset distances. All SORS spectra exhibit strong PET features, where the intensity decreases gradually as the offset distance increases, while the paracetamol peak intensity gradually increases as the offset is increased. For SD 45°/0°, a maximum of the paracetamol peak intensity is reached at an offset of 8–10 mm. For SD 45°/45°, the paracetamol peak intensity continues to increase with the offset distance up to

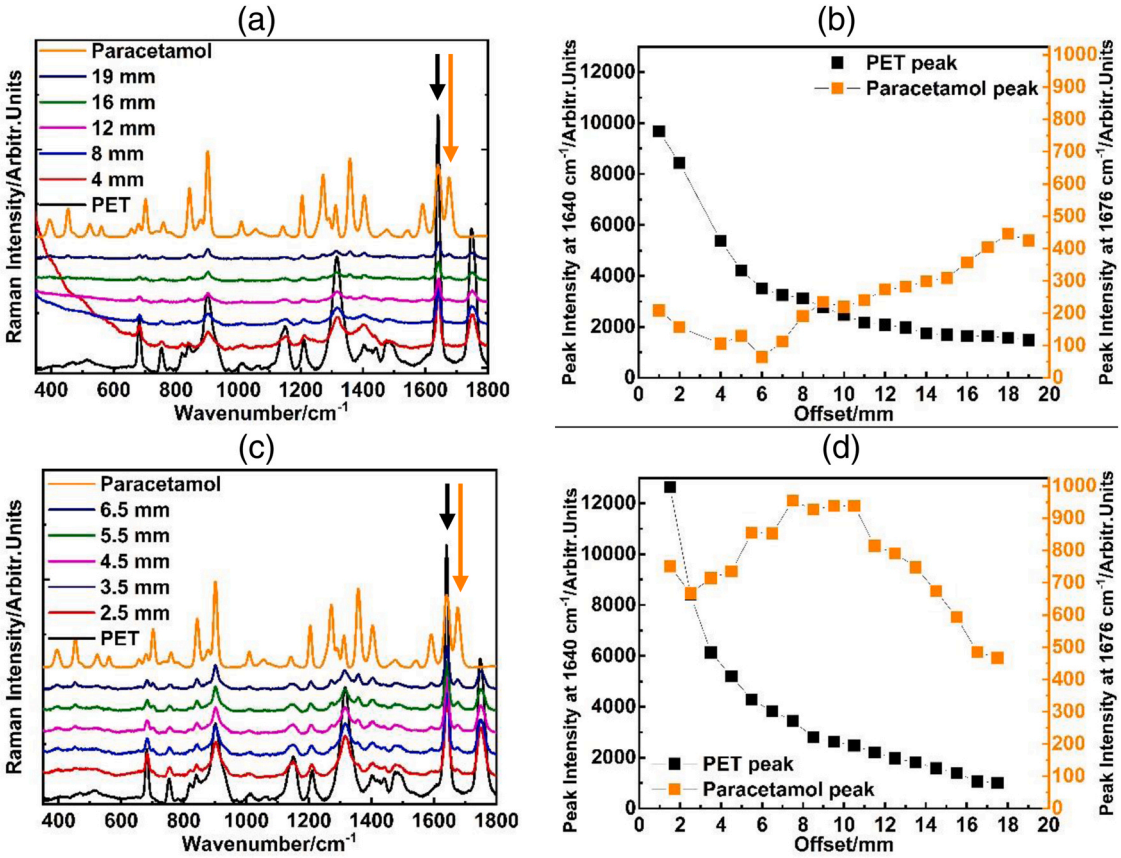

Fig. 5 – SORS spectra of paracetamol under a PET barrier of 21.6 mm thickness using (a) 45°/45° and (c) 45°/0° SD configurations. The PET peak at 1640 cm^{-1} and paracetamol peak at 1676 cm^{-1} are highlighted with the arrows in (a) and (c), and the change in their intensity as the offset distance increases is plotted in (b) and (d), respectively.

18 mm. The observations in Fig. 5 could be explained by the optical geometry of the measurement setup. By using angular illumination, the incident laser travels through the non-scattering and absorbing PET layer and reaches the paracetamol layer at a position shifted horizontally from where it encounters the PET layer. Similarly, collecting the Raman signal at an angle would lead to a horizontal shift in the measurement position. Therefore, Fig. 5(d) suggests that the illumination laser reaches the paracetamol layer at an offset of around 8–10 mm, and the offset required to observe the maximum intensity will be increased if an angular collection configuration is used. One could also anticipate that, by changing the thickness of the top layer, the intensity maximum might be achieved at a different offset distance.

The findings in Figs. 4 and 5 suggest that selecting an optimal configuration to use with the in-line probe is not straightforward and depends on the desirable sample depth to reach and the optical characteristics of the sample. Furthermore, practical constraints, such as the allowable dimension, the mechanical and optical components required, and the size of the access port of the process reactor, must be considered. Nevertheless, the observations from Fig. 5 suggests that the use of SD 45°/0° configuration allows the collection of photons that have travelled to deeper layers using smaller offset distances. Fig. 4 suggests that the maximum offset distance, after which the signal intensity decreases considerably, is around 5 mm. Therefore, offset distances employed within the SORS probe prototype were increments of 1 mm with 5 mm as the maximum offset distance. Moreover, additional collection fibres were incorporated for the larger offset distances, as described in Sec. 2.3.1 and illustrated in Fig. 2, to increase the total signal throughput, improving the signal quality of the resulting spectra. The total number of collection fibres is limited, considering the utilisation of the available 2D detecting area of the spectrometer.

3.2. *Characterisation of the optical performance of the SORS probe*

3.2.1. *Signal calibration*

A series of calibrations and analyses were conducted to characterise the optical performance of the SORS probe. This included the evaluation of the optical throughput of each of the individual collection fibres and the total intensity for each of the offset distances as well as the ability to detect the solvent signal through a barrier of paracetamol. Fig. 6(a) shows the typical 2D spectral image captured by the spectrometer with the argon lamp illuminated at the probe tip. The pixels in the x-axis to the Raman spectral range of 0–1870 cm^{-1}, while the y-axis corresponds to the spatial arrangement of the collecting fibres. The 19 horizontal dashed lines correspond to the 19 collection fibres in the probe; those 19 fibres form six groups corresponding to the 0–5 mm offset distances as described in Sec. 2.3.1. The vertical dashed lines correspond to the argon emission lines, where the intensity (brightness) of the lines corresponds to the intensity variations of the characteristic argon peaks. The position of the characteristic Argon peaks were used to calibrate the wavenumber position of the spectra.

Fig. 6(b) shows the intensity from each of the individual fibres as extracted from Fig. 6(a), with the argon characteristic peak at 416.45 cm^{-1}, corresponding to the emission line

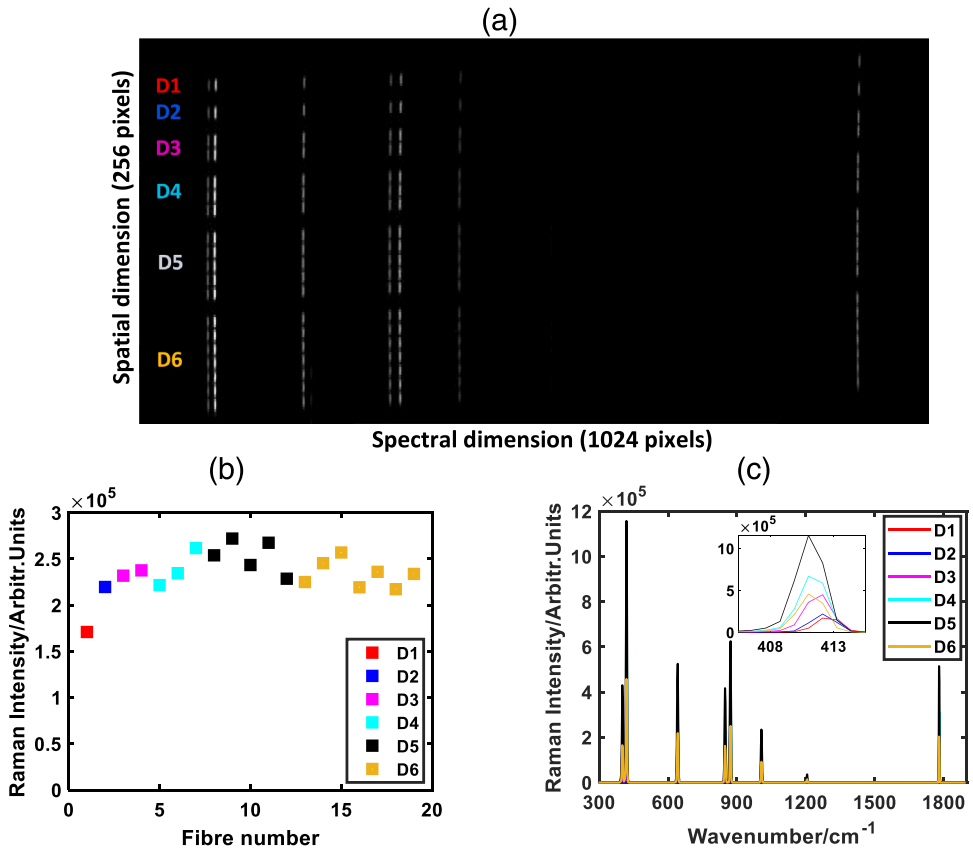

Fig. 6 – (a) Spectral image of argon showing the signal corresponding to 19 collection fibres from backscattering and 5 offsets, D1-D6, (b) maximum intensity of the argon peak at 416.45 cm^{-1} (811.54 nm) for the 19 individual collection fibres, and (c) argon signal combined as one spectrum per offset distance. The inset shows a zoomed-in view of the argon peak at 416.45 cm^{-1}.

at 811.54 nm. Small variation across all fibres, exhibited in Fig. 6(b), is consistently observed from SORS calibration. Due to the use of an integrating sphere in the calibration, a uniform intensity from all collecting fibres would be anticipated as a result of the uniform argon intensity provided by the sphere. The observation in Fig. 6(b) could therefore be attributed to the variation in the optical throughput of the fibres. It can be further explained by the design of the optical probe, which centred and optimised the optical alignment of the fibres corresponding to the 4 mm offset to the lens and filter in the probe. This design enhances the optical throughput for configurations that typically exhibit weak SORS signals while balancing the use of additional fibres in the signal collection.

The signal from fibres of the same offset group can be combined to form the spectra for each offset, as shown in Fig. 6(c). A slight systematic shift across all argon peaks, shown in the inset in Fig. 6(c) and Fig. S3, was observed. Further analysis suggests that this may be due to chromatic aberration of the optics in the spectrometer, with the small deviation observed (± 0.2 nm) considered negligible. The use of additional fibres increases the total signal throughput for the larger offset configurations. Spectra collected using the SORS probe exhibit comparable intensity across the offset distances, as shown in Fig. S2; this is in contrast to Fig. 4, where the decrease in SORS intensity with the increase in offset distance occurs. The increase in the optical throughput would improve the quality of the spectra, which could be critical to the analysis of the performance of the offset configuration in estimating the solvent content.

3.2.2. *Depth of solvent signal detection*

To determine the depth through which the signal of the solvent anisole can be detected, spectra of anisole placed under a layer of paracetamol of variable thicknesses were investigated. Spectra obtained from samples of various thicknesses of paracetamol were collected from all offset distances, and the anisole peak at 1002 cm^{-1} was selected for the analysis due to its intense and sharp spectral feature, in addition to being resolvable from the other major paracetamol peaks, as shown in Fig. S4.

An example of the results from the 3 mm offset (D4) is given in Fig. 7(a), where an increase in anisole peak intensity as the thickness of the paracetamol barrier layer decreases can be observed. Results from all offsets for all paracetamol thickness are provided in Fig. S5. At a depth of 1.2 mm of paracetamol, Fig. 7(a) and Fig. S5 show an increase in solvent peak intensity as the collection offset distance increases. Moreover, the configurations of larger offsets show that the solvent peak continues to be resolvable at a larger depth. This finding suggests a larger contribution from the subsurface to the collected spectra as the offset distance increases. In order to examine and compare the magnitude of the solvent signal from each of the offsets, the intensity of the anisole peak at 1002 cm^{-1} was divided by the paracetamol peak intensity at 1020 cm^{-1} and expressed as a function of depth, as seen in Fig. 7(b). A higher ratio indicates a stronger intensity of the solvent peak, allowing estimation of the depth through which the solvent peak is detectable. The curve is flat beyond 1.2 mm for D1. This is increased to 3.6 mm for D2 and D3, 4.8 mm for D4, and exceeds 6 mm for D5 and D6. The

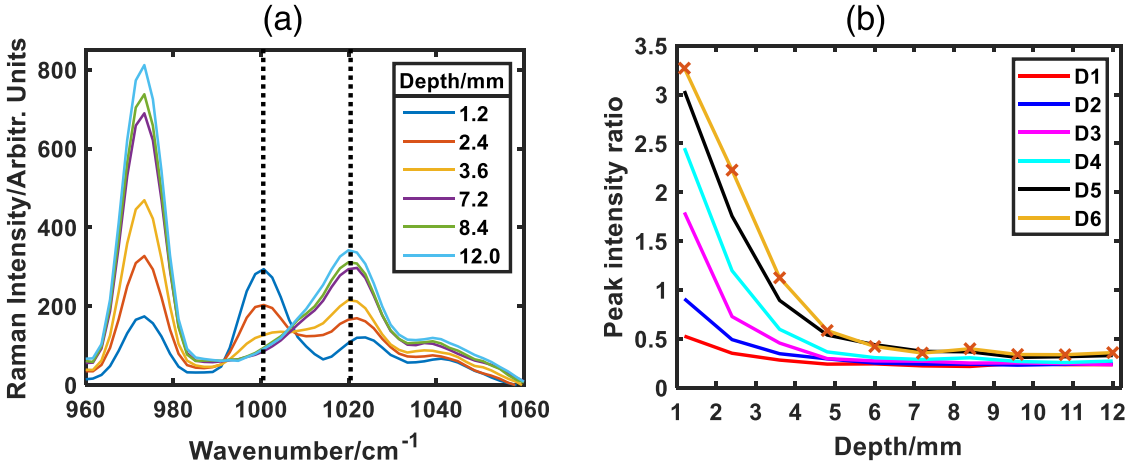

Fig. 7 – (a) SORS spectra from D4, 3 mm offset, of anisole beneath paracetamol of the depths shown in the legend. The anisole and paracetamol peaks at 1002 and 1020 cm^{-1} are highlighted by the dashed lines. (b) Change of the intensity ratio of anisole peak to the paracetamol peak, highlighted in (a), with the paracetamol depth for each offset. Ticks indicate the depths at which the measurements were taken; the lines are plotted to guide the eye.

larger depth reached by the larger offset distances observed in Fig. 7 illustrates the advantage of using SORS for characterising the subsurface of non-homogeneous systems.

3.3. *Monitoring of pharmaceutical drying using the developed SORS probe*

In this part of the study, SORS spectra were collected intermittently from multiple drying runs to build a sufficiently large dataset of spectra representing variable solvent content values. Fig. 8(a) shows the drying curves for each of the runs with the solvent content determined using loss on drying measurements. Those curves display some variation encountered within the drying runs despite the use of the same process parameters in all runs. This is due to the change in the vacuum pressure during the process as a combined result of the imperfect sealing around the agitator shaft and a fluctuation in the function of the vacuum pump. The non-uniformity of the wet cake during the drying is another contributing factor to the observed variation in the drying curve. The figure illustrates the significant, and practical, challenges in process control for ensuring critical quality attributes are met and in determining the end point of the drying.

Fig. 8(b) shows spectra collected from D3, 2 mm offset, during one of the drying runs. Due to the variation in the baseline of the spectra, the changes in the anisole peak intensity with the solvent content are not clearly evident looking at those raw spectra. Similar spectral responses from other offset distances were also observed across the runs. To remove the variation in the baseline, standard normal variate (SNV) was selected as the pre-processing method. Unlike many other pre-processing methods, such as the multiplicative scatter correction (MSC), the SNV method normalises a spectrum using the mean and standard deviation of the spectrum, eliminating any possible effects from other spectra in the dataset (Rinnan et al., 2009). Fig. 8(c) shows the SNV pre-processed spectra from Fig. 8(b). The decrease in anisole peak intensity following the decrease in solvent content can be seen and is readily distinguishable from the paracetamol peaks by the removal of the baseline effect. This suggests that those spectra could be used for the quantification of solvent content using PLSR analysis.

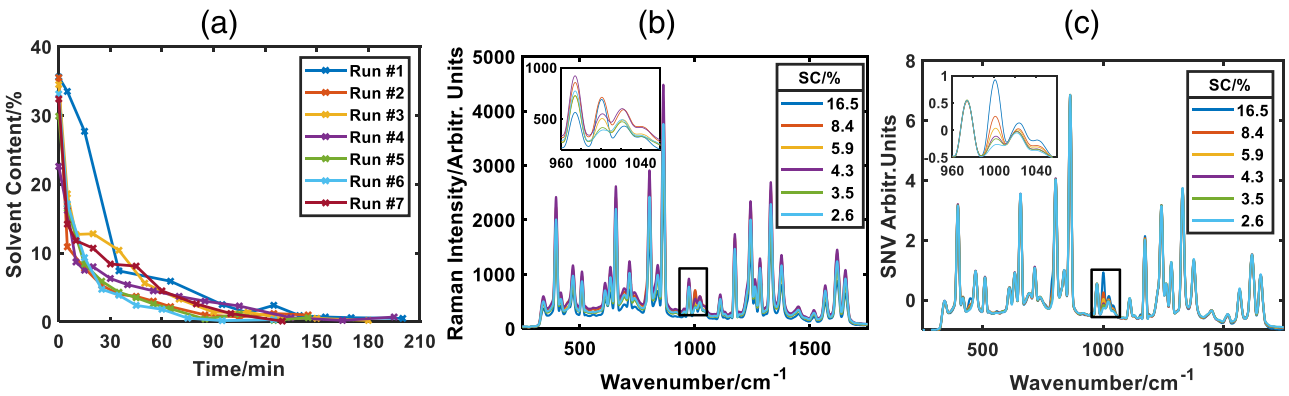

Fig. 8 – (a) Drying curves of all runs of granular paracetamol in anisole. (b) Raw spectra from D3 of Run 5. (c) SNV pre-processed spectra from (b). The legend shows the solvent content percentage (SC %) in the sample of each spectrum. Insets are the zoom-in view on the anisole peak at 1002 cm^{-1}.

Table 2 – Summary of PLSR model performance of raw and SNV pre-processed spectra of each of the detector offsets.

Detector	Raw				SNV Pre-processed			
	#LVs	RMSECV/%	RMSEP/%	R^2	#LVs	RMSECV/%	RMSEP/%	R^2
D1	3	2.22	1.47	0.78	3	1.07	0.76	0.95
D2	3	1.93	2.32	0.83	2	0.98	0.93	0.96
D3	4	1.95	2.54	0.83	3	0.85	0.76	0.97
D4	5	1.57	2.02	0.89	4	0.71	0.63	0.98
D5	5	1.14	1.28	0.94	4	0.65	0.58	0.98
D6	4	1.38	2.23	0.91	5	0.64	0.55	0.98

PLSR analysis was applied using spectra collected from all runs from the different offset distances, where results from D1, the backscattering spectra, were used as a benchmark. This analysis aimed to establish a systematic investigation of the model performance using the SORS configurations, the advantage of pre-processing of SORS spectra, and the source of improvement where observed. The full spectral range of 250–1750 cm^{-1} was used for the analysis.

The suitable number of latent variables (LVs) for the PLSR model from each configuration is determined by examining the RMSECV curve, where the latent variable at which the curves reach the lowest RMSECV value or a plateau is selected. The model with a suitable number of LVs is then used to predict the solvent content in the samples of the independent test set. The models using raw and SNV pre-processed spectra are summarised in Table 2. This table shows the RMSECV and coefficient of determination of the calibration model in addition to the RMSEP obtained for the test set. Detailed comparison of the RMSECV curves of the models built using both raw and SNV pre-processed spectra from each of the offset distances are shown in Fig. S5. All RMSECV curves show a sharp decrease in the cross-validation error after incorporating the first two latent variables in the model. In addition, models using the pre-processed spectra exhibit a greater decrease than those of the raw spectra, with Fig. S5(a) showing more pronounced variation among RMSECV curves for different offset configurations.

Compared to the models built with the raw spectra, the models built with the pre-processed spectra exhibit lower errors in estimating the solvent content. By simply rescaling the spectra, a notable reduction in the estimation error by 50 % can be achieved. The larger values of RMSEP compared to RMSECV of the raw spectra indicate some limitations in the calibration set for describing the variation in the test set. This limitation could be related to the random variation in spectral baseline, such as the example shown in Fig. 8(b), which may not be fully described by the calibration models. Another interesting observation in Table 2 is the improvement in model performance as the offset distance increases. This was also observed in the model built from spectra from D6, despite the lower optical throughput seen in Fig. S2. The model built from spectra from D5 shows an approximately 50 % reduction in RMSECV, with a consistent level of RMSEP, compared to those of the benchmark configuration, D1. Further comparison of the prediction and residual plots are shown in Fig. 9 for the models built from the SNV pre-processed spectra from D1 and D5. While the linearity of the prediction is clear in both models and also indicated by the R^2 in Table 2, the PLSR model of D5 spectra showed more consistent error over the solvent content range compared to the benchmark configuration. Despite the decrease in error for models of pre-processed spectra from all configurations, R^2 values display a slight increase, suggesting strong linearity in the models.

To investigate the source of the improvement in models from D5, further analysis on the scores and loadings of each LV used for the model was performed. The percentage of variance in the spectra described by each LV is summarised in Table 3 for the model built from the spectra from D1 and D5. Detailed comparison of the loading curves to the pure component spectra of paracetamol and anisole, as well as the associated scores, are given in Fig. S6 and Fig. S7 for D1 and D5, respectively. The second latent variable, LV2, is found to capture most of the features related to anisole, as the loading curve closely resembles the anisole spectra. Nevertheless, the main anisole features, the peak at 1002 cm^{-1}, also appear in the loading curve for LV3 in the model built from the spectra from D5, although it also contains features resembling paracetamol spectra. In comparison, these anisole features are less resolvable in LV3 for the model built using the spectra from D1 than those using the spectra from D5. The loading for LV4 in Fig. S7 only resembles paracetamol spectra, particularly for the spectral range over 1000 cm^{-1}. Therefore, a combined variance of 47.63 % and 48.87 % can attribute to the solvent-related information captured by the model using spectra from D1 and D5, respectively.

The findings from Table 3, Fig. S6, and Fig. S7, suggest that the spectra from the offset configurations contain: (1) more solvent-related information due to the longer optical path length involved, (2) information more consistent with the solvent content estimated by the reference loss on drying analysis, or the combination of both (1) and (2), compared to those from the backscattering configuration. The longer optical path length travelled and deeper sample area reached, as illustrated in Fig. 7 and Fig. S5, signify that the offset configuration could better describe the bulk properties of the sample and mitigate issues related to localised sample non-uniformity. Furthermore, the reference analysis quantitates the solvent content in the bulk of the sample, which may not be representative of the condition at the surface due to the different physical environments that the sample is exposed to, introducing errors in the model from the measurements from D1. Considering the fundamental mechanism and dynamic nature of the drying process, it would be challenging to quantitate the differences in solvent content analysis for samples at the surface and in bulk, or to mitigate such differences.

The study presented here highlights the advantage of applying SORS measurements in-line over traditional in-line (backscattering) Raman spectroscopy for providing a real-time, in-situ assessment of solvent content with compatible sampling volume to the off-line analysis used as a common

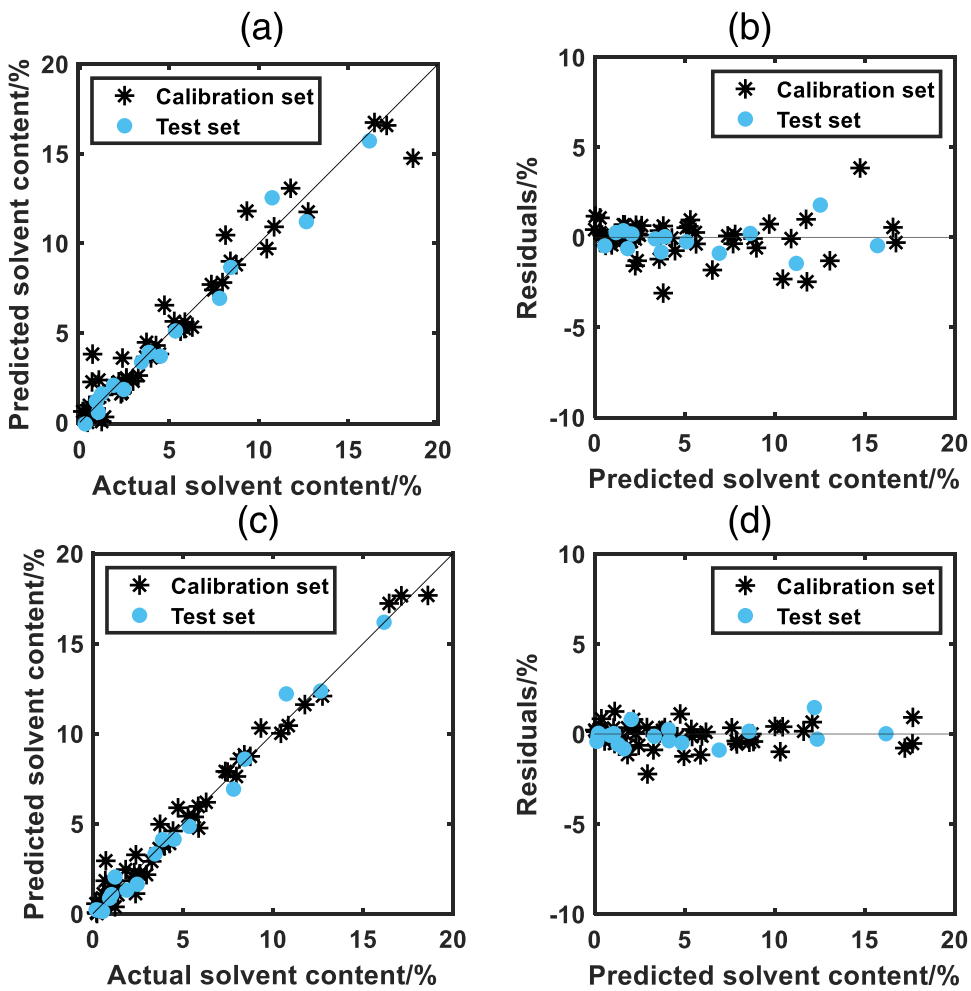

Fig. 9 – (a) Regression and (b) residuals plot of PLSR models constructed from D1 spectra and (c) regression and (d) residuals from D5 SNV pre-processed spectra of granular paracetamol in anisole.

Table 3 – Summary of model variance captured by each LV of PLSR models of SNV pre-processed spectra from D1 and D5 of granular paracetamol in anisole.

LV	D1		D5	
	Variance in LV/%	Sum of variance in model/%	Variance in LV/%	Sum of variance in model/%
1	50.44	50.44	50.40	50.40
2	47.17	97.61	47.33	97.74
3	0.46	98.07	1.54	99.27
4	–	–	0.08	99.36

practice. While this study focuses on the development of an in-line SORS probe and the evaluation of the performance obtained using SORS configurations, the setup and the probe design enable simultaneous analysis of multiple offset configurations, providing great potential for analysing sample uniformity and investigating and monitoring non-uniform drying and its mechanism.

4. Conclusions

A spatially offset Raman spectroscopy probe was developed for in-line process monitoring. The optical configurations, source and collection angles and collection offset distances, were determined through the evaluation of the signal obtained from potential configurations using a benchtop SORS setup to collect granular paracetamol spectra. The performance of those configurations and the increased contribution of the sublayer signal for larger offsets were further explored by analysing paracetamol spectra through a barrier of PET. The optical configuration of the SORS probe consisted of a 45° illumination angle and five equidistant collection offsets of 1–5 mm in addition to a backscattering configuration. This study also explores the depth through which the anisole signal under layers of dry paracetamol can be detected. The solvent signal from larger depths was detected using larger collection offsets, whereas the backscattering signal had the lowest detection depth. This provided an initial demonstration of the advantage of using SORS to probe non-homogeneous particulate media.

Drying of granular paracetamol in anisole was monitored at-line using the developed SORS probe. PLSR analysis of the raw and SNV pre-processed spectra from the offsets showed improvement in performance as the collection offset increased. The RMSECV and RMSEP were reduced in the models of the pre-processed spectra for all offset distances. A 50 % decrease in the error of models using raw D5 spectra compared to the conventional backscattering configuration was seen, in addition to a further 50 % decrease from SNV pre-processed D5 spectra compared to the raw D5 spectra. This was attributed to greater depths probed by the larger offset distances as was illustrated in the results of the depth of solvent signal detection. Further investigation into the

scores and loadings indicated that more solvent-related information was captured in the model of the spectra of the offset compared to the model of the backscattering spectra. This study demonstrates the developed SORS probe potential for addressing current challenges and needs by providing a real-time, in-situ analytical solution in monitoring solvent content during pharmaceutical drying.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to acknowledge the Enabling Technologies Consortium for funding this work and the University of Strathclyde for the studentship funding MA.

The authors acknowledge the ESPRC Centre for Continuous Manufacturing and Advanced Crystallisation (CMAC) for providing the facilities utilised in this work.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cherd.2023.02.041.

References

Asiala, S.M., Shand, N.C., Faulds, K., Graham, D., 2017. ACS Appl. Mater. Interfaces 9, 25488.
Aulton M.E., Somavarapu, S. , 2018. In: Aulton's Pharmaceutics, The Design and Manufacture of Medicines; 5th ed.; Aulton ME, Taylor, K.M.G., Ed.; Elsevier Health Sciences: NP.
Biancolillo, A., Marini, F., 2018. Front. Chem. 6.
Chanda, A., Daly, A.M., Foley, D.A., LaPack, M.A., Mukherjee, S., Orr, J.D., Reid, G.L., Thompson, D.R., Ward, H.W., 2015. Org. Process Res. Dev. 19, 63.
Chao, K., Dhakal, S., Qin, J., Peng, Y., Schmidt, W.F., Kim, M.S., Chan, D.E., 2017. Sensors 17, 618.
Chen, Z.-P., Li, L.-M., Jin, J.-W., Nordon, A., Littlejohn, D., Yang, J., Zhang, J., Yu, R.-Q., 2012. Anal. Chem. 84, 4088.
Conder, E.W., Cosbie, A.S., Gaertner, J., Hicks, W., Huggins, S., MacLeod, C.S., Remy, B., Yang, B.S., Engstrom, J.D., Lamberto, D.J., Papageorgiou, C.D., 2017. Org. Process Res. Dev. 21, 420.
Dooley, M., McLaren, J., Rose, F.R.A.J., Notingher, I., 2020. J. Biophotonics 13, e202000190.
Eliasson, C., Matousek, P., 2007. Anal. Chem. 79, 1696.
Eliasson, C., Macleod, N.A., Matousek, P., 2007. Anal. Chem. 79, 8185.
Fonteyne, M., Gildemyn, D., Peeters, E., Mortier, S.T.F.C., Vercruysse, J., Gernaey, K.V., Vervaet, C., Remon, J.P., Nopens, I., De Beer, T., 2014. Eur. J. Pharm. Biopharm. 87, 616.
Gagnon, F., Desbiens, A., Poulin, É., Bouchard, J., Lapointe-Garant, P.-P., 2021. Chem. Eng. Res. Des. 174, 254.
Ghita, A., Matousek, P., Stone, N., 2018. J. Biophotonics 11, e201600260.
Hamilton, P., Littlejohn, D., Nordon, A., Sefcik, J., Slavin, P., Dallin, P., Andrews, J., 2011. Analyst 136, 2168.
Hopkins, R.J., Pelfrey, S.H., 2012. Shand NC Anal. 137, 4408.
ICH, 2020. In EMA/CHMP/ICH/213867/2020; Use ICFHOTRFPFH, Ed.; European Medicines Agency: Amsterdam. https://www.ema.europa.eu/en/ich-q3c-r8-residual-solvents-scientific-guideline.
Johansson, J., Sparén, A., Svensson, O., Folestad, S., Claybourn, M., 2007. Appl. Spectrosc. 61, 1211.
Keller, M.D., Vargis, E., de Matos Granja, N., Wilson, R.H., Mycek, M.-A., Kelley, M.C., Mahadevan-Jansen, A., 2011. J. Biomed. Opt. 16 (7), 077006.
Kjeldahl, K., Bro, R., 2010. J. Chemom. 24, 558.
Kogermann, K., Aaltonen, J., Strachan, C.J., Pöllänen, K., Heinämäki, J., Yliruusi, J., Rantanen, J., 2008. J. Pharm. Sci. 97, 4983.
Kona, R., Qu, H.-b, Mattes, R., Jancsik, B., Fahmy, R., 2013. Hoag SW Int. J. Pharm. 452 (1–2), 63.
Larkin P.J., Larkin P.J., Ed., 2011; Elsevier Science: NP.
Maltesen, M.J., van de Weert, M., Grohganz, H., 2012. Aaps Pharmscitech 13, 747.
Märk, J., Karner, M., Andre, M., Rueland, J., Huck, C.W., 2010. Anal. Chem. 82, 4209.
Matousek, P., 2006. Appl. Spectrosc. 60, 1341.
Matousek, P., Parker, A.W., 2007. J. Raman Spectrosc. 38, 563.
Matousek, P., Clark, I.P., Draper, E.R.C., Morris, M.D., Goodship, A.E., Everall, N., Towrie, M., Finney, W.F., Parker, A.W., 2005a. Appl. Spectrosc. 59, 393.
Matousek, P., Morris, M.D., Everall, N., Clark, I.P., Towrie, M., Draper, E., Goodship, A., Parker, A.W., 2005b. Appl. Spectrosc. 59, 1485.
Mosca, S., Conti, C., Stone, N., Matousek, P., 2021. Nat. Rev. Methods Prim. 1, 21.
Murugesan S., Sharma, Praveen K., Tabora, Jose E., 2010. In Chemical Engineering in the Pharmaceutical Industry; Ende DJa, Ed., p 315.
Næs T., Isaksson T., Fearn T., Davies T. , 2002. A user-friendly guide to multivariate calibration and classification; NIR Publications.
Nicolson, F., Jamieson, L.E., Mabbott, S., Shand, N.C., Graham, D., Faulds, K., 2017. J. Raman Spectrosc. 48, 1828.
Nieuwmeyer, F.J.S., Damen, M., Gerich, A., Rusmini, F., van der Voort Maarschalk, K., Vromans, H., 2007. Pharm. Res. 24, 1854.
Ottoboni, S., Simurda, M., Wilson, S., Irvine, A., Ramsay, F., Price, C.J., 2020. Powder Technol. 366, 305.
Peinado, A., Hammond, J., Scott, A., 2011. J. Pharm. Biomed. Anal. 54, 13.
Peters, J., Teske, A., Taute, W., Döscher, C., Höft, M., Knöchel, R., Breitkreutz, J., 2018. Int. J. Pharm. 537, 193.
Qin, J., Kim, M.S., Schmidt, W.F., Cho, B.-K., Peng, Y., Chao, K., 2016. J. Raman Spectrosc. 47, 437.
Reddy, J.P., Jones, J.W., Wray, P.S., Dennis, A.B., Brown, J., Timmins, P., 2018. Int. J. Pharm. 541, 253.
Rinnan, Å., Berg, Fvd, Engelsen, S.B., 2009. TrAC Trends Anal. Chem. 28, 1201.
Schulmerich, M.V., Dooley, K.A., Vanasse, T.M., Goldstein, S.A., Morris, M.D., 2007. Appl. Spectrosc. 61, 671.
Steponavičius, R., Thennadil, S.N., 2009. Anal. Chem. 81, 7713.
Stone, N., Baker, R., Rogers, K., Parker, A.W., Matousek, P., 2007. Analyst 132, 899.
Tewari, J., Dixit, V., Malik, K., 2010. Sens. Actuators B: Chem. 144, 104.