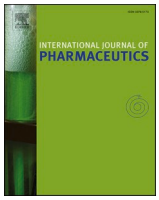
International Journal of Pharmaceutics 664 (2024) 124650

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

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



In-line indirect concentration measurement of ultralow dose API during twin-screw wet granulation based on NIR and Raman spectroscopy

Dorián László Galata [a], András Domokos [a], Balázs Démuth [b], Petra Záhonyi [a], Gergő Fülöp [c], Zsombor Kristóf Nagy [a,*], Brigitta Nagy [a]

[a] Department of Organic Chemistry and Technology, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary
[b] SeraNovo B.V., J.H. Oortweg 21, 2333 CH Leiden, the Netherlands
[c] Gedeon Richter Plc., Formulation R&D, Gyömrői u. 19-21, H-1103 Budapest, Hungary

ARTICLE INFO

Keywords:
Twin-screw wet granulation
Process analytical technology
In-line concentration measurement
NIR spectroscopy
Raman spectroscopy
Ultralow dose
Chemometrics

ABSTRACT

Twin-screw wet granulation (TWSG) is a promising continuous alternative of pharmaceutical wet granulation. One of its benefits is that the components dissolved in the granulation liquid are distributed homogeneously in the granules. This provides an elegant way to manufacture products with ultralow drug doses. Near-infrared (NIR) and Raman spectroscopy are well-established process analytical technology (PAT) tools that can be used for the in-line monitoring of TSWG. However, their detection limit does not enable the measurement of components in the ultralow (*i.e.*, ppm) range. In this paper, an indirect approach is presented that enables the real-time determination of the concentration of a drug in concentrations between 40 and 100 ppm by using the signal of an excipient, in this case, the polyvinylpyrrolidone (PVP). This component is also dissolved in the granulation liquid; therefore, it is distributed in the same way as the active ingredient. Results of HPLC measurements have proved that the models trained to quantify the concentration of PVP in real-time gave an accurate determination for the active ingredient as well (root mean squared error was 7.07 ppm for Raman and 5.31 ppm for NIR spectroscopy, respectively). These findings imply that it is possible to indirectly predict the concentration of ultralow dose drugs with in-line analytical techniques based on the concentration of an excipient.

1. Introduction

Wet granulation is a vital step of pharmaceutical downstream production. This process usually relies on a binder compound to form larger granules from the particles of a powder. As a result, the flowability and compressibility of the material are improved, which enables the production of tablets with more consistent quality. The uniformity of tablet mass and dosage becomes better, and the crushing strength and friability of the tablets are also improved when granules are used for tablet production (Dhenge et al., 2010). There are two dominant batch technologies for wet granulation, namely fluidized bed granulation (Petrović et al., 2011) and high shear granulation (Liu et al., 2021), while the most desirable continuous wet granulation method is twin-screw wet granulation (TSWG) (Bandari et al., 2020). This process uses a modified twin-screw extruder, where the powder and the granulation liquid are thoroughly mixed with two counter-rotating or parallel rotating screws.

Thanks to the intensive and efficient mixing inside the twin-screw granulator, the distribution of the components in the granules can be very homogeneous. This leads to another great advantage of TSWG: the thorough mixing enables the effective homogenization of products with the active pharmaceutical ingredients (APIs) at an ultralow dosage. Démuth *et al.* showed that it is possible to manufacture granules with an API concentration of approximately 0.035 % w/w with a relative standard deviation (RSD) of less than 2.5 % (Démuth et al., 2020).

In a continuous process, only a small amount of product leaves the equipment at any given time. Therefore, by using a real-time analytical tool, it is feasible to characterize the whole mass of the product. This is highly beneficial compared to the traditional quality inspection of the pharmaceutical industry, where each batch is evaluated by small samples that are measured with slow, destructive methods. There are many process analytical technology (PAT) methods that can be used for the real-time monitoring of TSWG. Typically, the following characteristics can be predicted based on PAT measurements. Moisture content is a vital attribute of granules, that must be monitored to evaluate whether drying

* Corresponding author.
E-mail address: zsknagy@oct.bme.hu (Z.K. Nagy).

https://doi.org/10.1016/j.ijpharm.2024.124650
Received 4 July 2024; Received in revised form 14 August 2024; Accepted 27 August 2024
Available online 28 August 2024

was successful. Microwave spectroscopy (Peters et al., 2017) and near-infrared (NIR) spectroscopy (Domokos et al., 2021; Palmer et al., 2020) are the best tools for this purpose. The particle size of the granules can be monitored in real-time with machine vision-based tools (Madarász et al., 2022; Meng et al., 2019). The physical properties of the granules, such as bulk density, tapped density, and flowability can also be predicted based on NIR spectroscopy, Raman spectroscopy, and machine vision (Fonteyne et al., 2012). Raman spectroscopy is sensitive to changes in the polymorphism of APIs, therefore several authors have used this method to detect changes in the crystalline state during granulation (Fonteyne et al., 2013; Nomura et al., 2020).

The API concentration of the granules must also be evaluated to ensure that the material used for tableting is homogeneous so that the tablets will have an appropriate content uniformity (CU). Martinetz *et al.* used NIR spectroscopy for API concentration measurement during dry granulation. They placed integrating spheres after the blender and the roller compactor, and their model was used for granules with 10 % w/w and 20 % w/w API (Martinetz et al., 2018). Román-Ospino *et al.* utilized a NIR line scan spectrometer to measure 9–13.5 % w/w API concentration of granules on a conveyor belt right after the granules were discharged from the twin-screw granulator (Román-Ospino et al., 2020). Meng *et al.* applied a Raman spectrometer above a conveyor transporting granules produced on a TWSG appliance, where they could measure API concentrations around 8 % w/w (Meng et al., 2019). Harting *et al.* used a Raman spectroscope placed at the outlet of a twin-screw granulator to measure the API concentration of granules between 5–50 % w/w (Harting and Kleinebudde, 2018, 2019). These works show that both NIR and Raman spectroscopy are capable of monitoring the API concentration of granules in real-time. However, so far, these methods have not been used in the case of API concentrations below 1 % w/w. This is not surprising considering the fact that the sensitivity of these methods is limited, therefore, the limit of quantification (LoQ) of vibrational spectroscopy techniques does not allow for the accurate measurement of small concentrations. Consequently, when TSWG is used to produce low-dose granules, the real-time measurement of API concentration may no longer be possible. In this case, the API concentration can still be predicted if a component is present in larger concentration and which concentration is proportionate to the API. Typically, when low-dose granules are manufactured, the API is dissolved in the granulation liquid, which enables the thorough homogenization of the API. In this case, the other components dissolved in the granulation liquid are expected to be distributed in the granules similarly to the API. Ficzere *et al.* exploited this phenomenon by including a colored tracer material in the granulation liquid (Ficzere et al., 2021). The color of the granules depended on the L/S ratio, which also determined the API concentration of the granules. They placed a digital camera after the TSWG that acquired images of the granules leaving themachine, and the average color of the granules was correlated with the API concentration. This way, they could accurately determine the concentration of the API based on the images. The API concentration was around 0.05 % w/w, and the relative error of the prediction was 2.62 %. As an alternative method, the average diameter of the granules was also calculated, which also correlated with the API concentration with a relative prediction error of 2.30 %. It is also possible to determine the residence time distribution (RTD) of the system and predict the concentration of the output material by convolving the input mass flows with the probability distribution function. Gyürkés *et al.* demonstrated that this soft sensor-based approach can predict small changes in the API concentration of the granules caused by input mass flow changes (Gyürkés et al., 2022). However, the simplest tracer material for the API is the binder of the granulation liquid. This component is usually present in concentrations of at least 5 % w/w, and is expected to be distributed similarly to the API. To the author's knowledge, so far no work was published where the binder is used to indirectly predict the concentration of a low-dose API based on real-time measurements.

The purpose of this work was to create and test an analytical method that quantifies the concentration of an ultralow dose API in granules produced on a twin-screw granulator, where the API is dissolved in the granulation liquid. Granules were produced containing carvedilol (CAR) as model drug in concentrations between 40–100 ppm. NIR and Raman spectroscopy were used to determine the concentration of polyvinylpyrrolidone (PVP) K30, the binder in the granulation liquid. The API being dissolved in the granulation liquid, the concentrations of PVP and CAR correlate in the granules, enabling the indirect quantification of CAR concentration. The aim was to to create robust chemometric models for the quantification of the PVP concentration, and validate the hypothesis that the concentration of CAR can be determined based on the PVP. For validation, a continuous experiment with in-line NIR and Raman monitoring was carried out, where the L/S ratio was changed, resulting in granules with varying PVP and CAR content. The in-line measurement results were validated by using HPLC measurements.

2. Materials and methods

2.1. Materials

Carvedilol (CAR) was used as the model API of the formulation, provided by Merck (Budapest, Hungary), lactose monohydrate (Granulac 230®), used as filler, was obtained from Meggle Pharma (Duisburg, Germany), potato starch (filler) was supplied by Roquette Pharma (Lestrem, France). Polyvinylpyrrolidone (PVP) K30 – used as the binding agent of the granulation – was purchased from BASF (Ludwigshafen, Germany). Ethanol was obtained from Merck (Darmstadt, Germany).

2.2. Methods

2.2.1. Continuous twin-screw wet granulation

Granules containing an ultralow dose of CAR were prepared on a twin-screw granulator (TSG) (Quick 2000 Ltd., Hungary). A powder mixture of lactose monohydrate and potato starch (the exact ratio is confidential) was mixed in a polyethylene bag for 5 min. The granulation liquid was prepared by dissolving CAR and PVP K30 in a 96 % V/V ethanol–water mixture, and stirring the mixture for 24 h to ensure that the components were properly dissolved. The API (CAR) was dissolved in the granulation liquid in order to ensure that it would be distributed evenly in the granules, granules were produced with concentrations between 0–200 ppm. The TSG was equipped with 16 mm, 25 L/D ratio screws, operated at 200 rpm. The powder mixture was fed using a gravimetric feeder (Brabender Technologie, Duisburg, Germany) in gravimetric mode at a feed rate of 1 kg/h. The granulation liquid was dosed using a peristaltic pump (Watson Marlow 120U, Wilmington, MA, USA) at varying rates. The feeding rate of the pump has some oscillations due to its working principle (Démuth et al., 2020). The granules leaving the TSG fell on a conveyor belt (Brabender Technologie, Duisburg, Germany), where in-line NIR and Raman spectroscopic measurements were performed as detailed in Sections 2.2.3 and 2.2.4. The acquired spectra were imported into MATLAB and evaluated in real-time using an in-house developed script. The setup is described in more detail in our earlier work (Domokos et al., 2021).

2.2.2. Sample preparation for chemometric model calibration

The chemometric models that aim to quantify the composition of the granules need to be robust to the changes in the residual solvent concentration, as it can interfere with signals of the solid components of interest. This is because immediately after the material leaves the TSG, the concentration of the liquid and solid components will be proportionate to each other, the ratio corresponding to the composition of the granulation liquid. However, during the drying, the ratio of these components will continuously change. Therefore, depending on the length of drying, as well as the environmental factors (temperature, relative humidity) in the manufacturing site, the spectrum will be different due to the changes in the residual moisture content. For this

reason, the models were created by recording Raman and NIR spectra of the same granules immediately after granulation when the concentration of ethanol was high and also after the granules were dried. In this way, our aim was to develop such chemometric models that are robust to the varying moisture content, and represent the PVP (and API) variations irrespective of the drying stage.

Granules for calibration purposes with known composition were prepared by performing wet granulation with a mortar and pestle. The granulation liquid was added to a known mass of powder mixture in different amounts to produce granules with different PVP (and API) concentrations, and the components were thoroughly mixed, creating approximately 5 g of product in each case. The nominal concentration of PVP in the granules was set to the following values: 0, 1, 2, 3, 4, 5, 5.5, 6, 6.5, 7, 8, 9 and 10 % w/w, which correspond to 0, 20, 40, 60, 80, 100, 110, 120, 130, 140, 160, 180 and 200 ppm API concentration (the API concentration is always the 1/500 of the concentration of PVP). For the off-line validation of the models, independent test samples with 3.5, 6 and 7.5 % w/w PVP were created with the same technique but on different days. These samples represent concentrations levels that are not included in the calibration, this way the reliability of the models can be more objectively tested, as they must yield predictions for previously unseen concentrations.

Two sample sets were created and measured using NIR and Raman spectroscopy. In the first case, the granules were immediately measured after homogenization in the mortar. In the second scenario, the granules were allowed to dry for at least one day at room temperature before being measured. 3 spectra were recorded from each sample.

Later, the models were also augmented with samples produced on the TSG with various L/S ratios. In order to obtain a concentration value for these samples, the concentration of API was measured using HPLC.

2.2.3. Near-infrared spectroscopy

NIR spectra of the granules were recorded using a Bruker MPA FT-NIR spectrometer and OPUS 7.5 software (Bruker Optik Gmbh, Ettlingen, Germany). In-line and off-line diffuse reflection measurements were realized with a Solvias fiber optic probe. Spectra were recorded using a high-intensity Tungsten NIR source and an InGaAs detector in the 4000–12500 cm^{-1} wavenumber range with a resolution of 32 cm^{-1}, and 32 scans were averaged for each spectrum. The measurement time was approximately 10 s. Double-sided, forward–backward acquisition and 10 kHz scanner velocity were used. For the in-line measurement during the continuous process, the probe was positioned in a way that it touched the surface of the material on the conveyor. The off-line spectra were recorded with the same fiber optic probe and the same settings by inserting the probe into the surface of the samples.

2.2.4. Raman spectroscopy

Raman spectra were acquired with a Kaiser Raman Rxn2^{TM} spectrometer (Kaiser Optical Systems, Ann Arbor, MI, USA) and iC Raman 4.1 (Mettler-169 Toledo AutoChem Inc., USA) software. A Pharmaceutical Area Testing (PhAT) probe was used in reflection mode by illuminating the samples with a 400 mW diode laser at 785 nm wavelength in a 6 mm diameter circular area. Reflection spectra were recorded in the 200–1890 cm^{-1} Raman shift range with a resolution of 4 cm^{-1}. Spectra were recorded with an exposure time of 10 s and two scans. For the in-line measurements, the probe was placed above the material moving on the conveyor belt. Off-line spectra were acquired with the same settings by placing the samples below the PhAT probe.

2.2.5. HPLC measurements

The CAR concentration of the granules was determined using an Agilent 1200 HPLC apparatus (Agilent Technologies, Santa Clara, CA, USA). For each sample, 3 replicate HPLC measurements were performed the following way. 80 mg of powder was dissolved in 10.0 mL volumetric flasks in a 1:1 mixture of methanol and phosphorous acid solution (1:200 phosphorous acid: water solution). The dissolution was facilitated by placing the flasks inside an ultrasound bath for 10 min. Before injecting them into the HPLC system, the samples were filtered through a 0.45 μm PTFE syringe filter. The HPLC measurements were performed on a reverse-phase Gemini® NX-C18 3 μm 100 × 4.6 mm column (Phenomenex Inc, Torrance, CA, USA). The temperature was set to 25 °C, gradient elution was applied, and the flow rate was 1.0 mL/min. The UV absorption at 285 nm was used to determine the concentration of CAR in the solutions.

2.2.6. Creation of the chemometric models

Data analysis and chemometric model building were performed in MATLAB version 9.8 (Mathworks, Natick, MA, USA), using PLS_Toolbox 8.8.1 (Eigenvector Research, Manson, WA, USA).

The models determining the PVP concentration based on NIR or Raman spectra were created using the partial least squares (PLS) method (Geladi and Kowalski, 1986). The CAR concentration of the samples was determined by dividing the predicted PVP values by 500.

Three types of models were developed for both NIR and Raman spectroscopy. The first model – referred to as 'wet' model – contained only the spectra of material that was collected immediately after granulation and thus had a high ethanol content. The second model used only the spectra of the dried samples ('dry' models). Lastly, models were also created that included both the spectra of the wet and dry samples, referred to as 'combined' models. Consequently, a total of 6 models were obtained.

The performance of these models was evaluated based on the R^2 (Eq (1).) (Renaud and Victoria-Feser, 2010), and root mean squared error (RMSE) metrics (Eq (2).) (Porep et al., 2015).

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_{measured} - y_{predicted})^2}{\sum_{i=1}^{n} (y_{measured} - y_{predicted})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(y_{predicted} - y_{measured}\right)^2}{n}} \quad (2)$$

In Eqs. (1) and (2), n is the number of samples, $y_{measured}$ is the actual value of the variable and $y_{predicted}$ is the value of the variable predicted by the model. Both metrics are calculated for calibration, cross-validation and test samples separately, resulting in RMSEC/R^2(calibration), RMSECV/R^2(cross-validation), and RMSEP/R^2(prediction) values, respectively.

The models were also characterized by calculating the limit of detection (LoD) (Eq. (3)) and limit of quantification (LoQ) (Eq. (4)) values according to the ICH Q2 guideline (European Medicines Agency, 1995).

$$LoD = \frac{3.3\sigma}{S} \quad (3)$$

$$LoQ = \frac{10\sigma}{S} \quad (4)$$

where σ is the standard deviation of the y-intercept of the calibration line, and S is the slope of the calibration curve.

3. Results and discussion

3.1. Comparison of the NIR spectroscopy models

The suitability of NIR spectroscopy for the measurement of PVP concentration can be qualitatively evaluated by comparing the spectrum of PVP to the signals of ethanol and the wet granules. The NIR spectra of these components are shown in Fig. 1. PVP has several characteristic peaks in the region below 6000 cm^{-1}. Ethanol has very strong absorbance in the region below 9000 cm^{-1}, therefore a spectrum comparable to the other components could only be recorded by diluting ethanol with

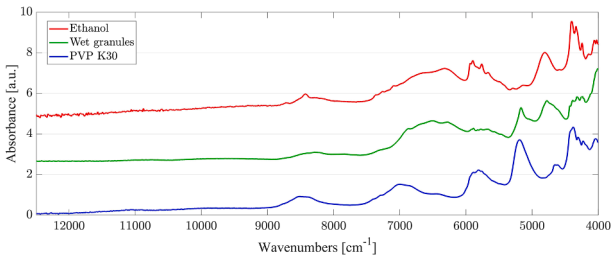


Fig. 1. The NIR spectra of ethanol, PVP K30 and a sample of wet granules. The spectra were preprocessed using SNV. The baselines of the spectra were shifted for better visibility.

KBr. On this diluted spectrum, a peak can be observed between 9000 and 8000 cm^{-1}. Between 7000 and 4000 cm^{-1}, ethanol has four additional peaks, some of these overlaping with the spectrum of the PVP. Several of these peaks also appear in the spectrum of the wet granules. However, the peak of PVP at 5200 cm^{-1} does not overlap with the peaks of ethanol, therefore, it gives an opportunity to measure PVP concentration without disturbance from ethanol.

First, models were created by using only the spectra of the off-line prepared samples, testing several preprocessing methods in various combinations. The first derivative was applied to remove the baseline and increase the sharpness of the typically broad peaks in NIR spectroscopy. As a normalization technique, multiplicative signal correction (MSC) and standard normal variate (SNV) were tested. The leave-one-out cross-validation method was used. Variable selection was also utilized to improve the performance of the models. The interval PLS (iPLS) was chosen with a window width of 100 variables. The iPLS technique improved the models in the case of the dry and the combined models, while the wet model had the best performance when the whole spectrum was included. The parameters of the dry, wet and combined NIR spectroscopy models are shown in Table 1.

The three NIR models and the predictions obtained for the test samples are shown in Fig. 2. Comparing the dry and wet models, it can be observed that both perform well in terms of calibration and cross validation metrics, resulting in a 0.2–0.3 % w/w RMSE. However, when the independent test samples (including both wet and dry samples) are predicted, there is a large difference between the performance of the two methods. Both methods yield accurate predictions for test samples that were prepared similarly to their respective (wet or dry) calibration samples. The wet model (Fig. 2/a) gives poor predictions for the dry test samples, underestimating the PVP concentration (test samples with a PVP concentration of 3.5, 6 and 7.5 % w/w are estimated to be approximately 1.5, 2.5 and 3 % w/w by the model). This points out that the calibration model fitted mainly accounting for the covariance of the ethanol NIR signals and the PVP concentration in the wet calibration samples. However, in the dry samples, the ethanol is not present anymore, therefore a reduced PVP concentration is erroneously predicted. The predictions of the test samples still correlate with the PVP concentration, meaning that the NIR signal of PVP also contributes to the predictions, but without ethanol, the values are underestimated. The dry model (Fig. 2/b) performs much better, indicating that the features of ethanol have a much smaller influence on the predictions. The PVP concentration of wet test samples is slightly overestimated, but the dry technique appears to be more robust to changes in ethanol concentration. Combining the two calibration datasets results in significantly worse calibration and cross-validation errors (Fig. 2/c). However, the predictions of the test samples are more accurate compared to the dry and wet models. This shows that it is worth sacrificing calibration and cross-validation metrics for a more robust overall model. The LoD and LoQ values of the wet and dry are lower than the combined model. The combined model contains data with larger variability, resulting in higher detection and quantification limits. The robustness of the combined model is also reflected in the number of latent variables. The wet and dry models yielded the best predictions for the test samples with 5 latent variables, but the combined yielded the best performance with only 3. It should be noted that the spectral region chosen by the iPLS method (5450.09–4597.67) for the combined model contains the peak of PVP that does not overlap with the peaks of ethanol. This shows that the variable selection technique was capable of finding the part of the spectrum that enables the prediction of PVP concentration without disturbances from ethanol. This region was also included in the dry model.

3.2. Comparison of the Raman spectroscopy models

The Raman spectra of ethanol, PVP and a sample of wet granules are displayed in Fig. 3. PVP has many characteristic peaks in the spectrum that can potentially be used to quantify its concentration in the granules. Ethanol also has a few peaks, but these do not overlap with the signal of PVP. This can facilitate the model building, as it is possible to develop a model that ignores the regions where ethanol appears and focuses on the signal of PVP by suitable variable selection.

During the creation of Raman calibration models, the tested preprocessing methods included smoothing with the Savitzky-Golay method to remove noise from the spectra, baseline removal with the automatic Whittaker filter method, and normalization via MSC, SNV, and area under the curve. The leave-one-out cross-validation technique was used. Variable selection was performed using the iPLS algorithm with a window width of 25. The application of the iPLS method enhanced the performance of the dry model, but it could not improve the wet and combined models. The parameters of the dry, wet and combined NIR spectroscopy models are shown in Table 2.

The three Raman models and the predictions obtained for the test samples are shown in Fig. 4.

Similar to NIR spectroscopy, the wet and dry models do not give

Table 1
Properties of wet, dry, and combined NIR spectroscopy models for PVP concentration prediction.

Model type	Wet model	Dry model	Combined model
Variable selection	**none**	**iPLS**	**iPLS**
Spectral range (cm^{-1})	12493.2–3999.82	9403.62–7498.21; 5453.95–4597.67	9403.62–7498.21; 5450.09–4597.67
Preprocessing	SNV, mncn*	SNV, mncn*	1stDer, SNV, mncn*
Cross validation	leave-one-out	leave-one-out	leave-one-out
Number of latent variables	5	5	3
RMSEC [% w/w]	0.258	0.272	0.593
RMSECV [% w/w]	0.294	0.320	0.619
RMSEP [% w/w]	2.479	0.691	0.486
R^2 (calibration)	0.994	0.991	0.970
R^2 (cross validation)	0.993	0.987	0.968
R^2 (prediction)	0.565	0.942	0.976
LoQ [% w/w]	2.62	2.62	6.03
LoD [% w/w]	0.86	0.87	1.99

* 1stDer: first derivative; SNV: standard normal variate; mncn: mean centering.

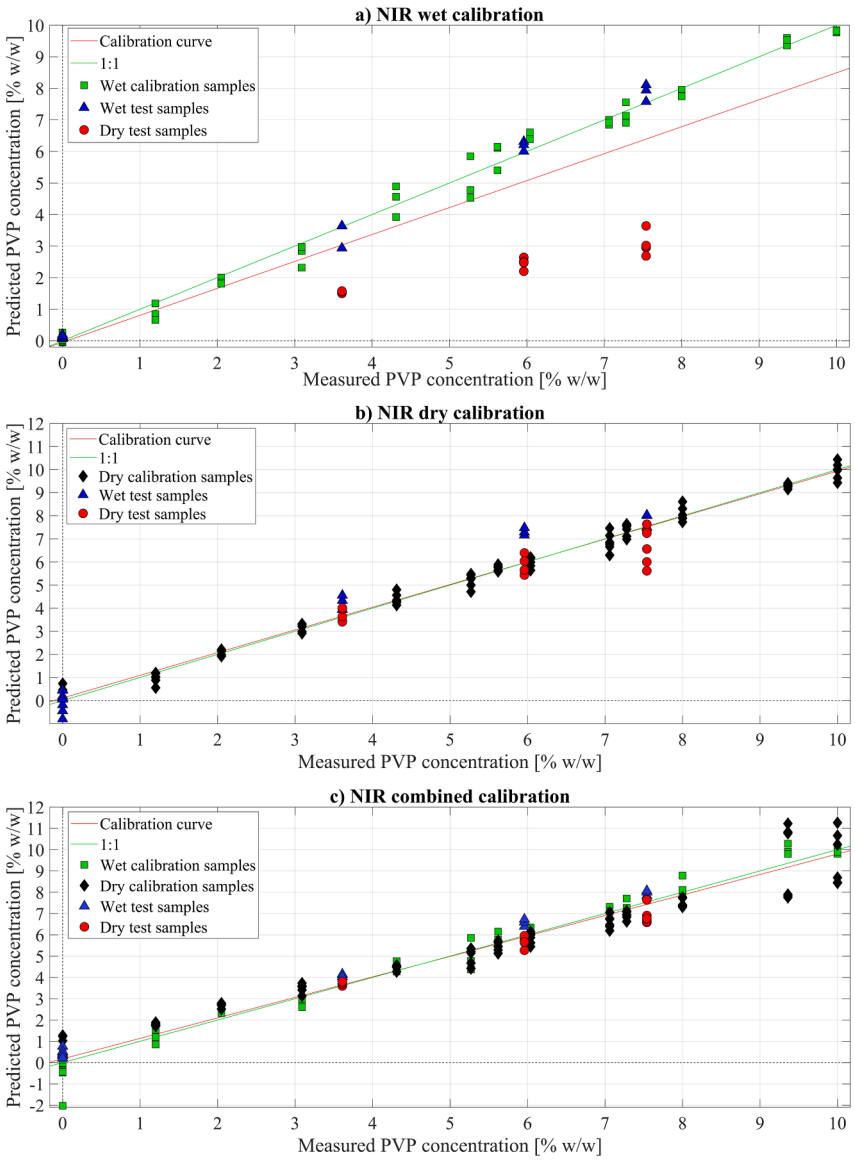


Fig. 2. NIR models for the prediction of PVP concentration.

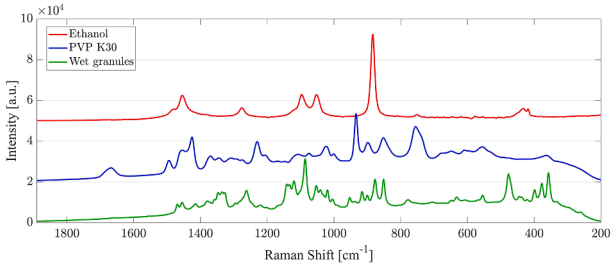


Fig. 3. The Raman spectrum of ethanol, PVP K30 and a sample of wet granules The baselines of the spectra were shifted for better visibility.

reliable predictions for the test samples. The worst performance belongs again to the wet model despite its lowest RMSEC and RMSECV values. The wet Raman model relies on the signal of ethanol more heavily than its NIR counterpart, thus while it yields accurate result for wet test samples, its predictions for the dry test samples are even worse. The trend of increasing PVP concentration barely appears in these predictions (all test samples are predicted between 0 and 1 % w/w PVP). In the case of Raman spectroscopy, the dry model does not perform substantially better either. It slightly underestimates the concentration of the dry test samples, and in the case of wet test samples, the predictions are much lower than the measured values. This is possibly because the Raman spectra consist of much sharper peaks, therefore when ethanol is absent, the shape of the spectrum deviates from the one expected by the model. Combining the wet and dry datasets yields more spectacular results in the case of Raman spectroscopy. The combined model performs excellently in the case of calibration, cross-validation and testing. The combined Raman model has an RMSEP value of 0.338 % w/w compared to 0.486 % w/w in the case of NIR spectroscopy. The LoD and LoQ values of the combined Raman model (1.05 % w/w and 3.17 % w/w) are also much lower than that of NIR spectroscopy (1.99 % w/w and 6.03 % w/w). Interestingly, the iPLS variable selection did not yield a more accurate combined model. Based on these results, Raman spectroscopy appears to be a more promising candidate for the monitoring of PVP concentration. The deviations between samples at the same concentration level can originate from noise caused by changes in ambient light, the physical structure of the sample, or the contact between the probe and the measured sample can also vary, causing slight changes in the spectra.

Table 2
Properties of wet, dry, and combined Raman spectroscopy models for PVP concentration prediction.

Model type	Wet model	Dry model	Combined model
Variable selection	**none**	**iPLS**	**none**
Spectral range (cm^{-1})	1890–200	940–916; 765–741; 690–666	1890–200
Preprocessing	bl, SNV, mncn*	bl, nm (1-norm), mncn*	bl, nm (1-norm), mncn*
Cross-validation	leave-one-out	leave-one-out	leave-one-out
Number of latent variables	2	2	5
RMSEC [% w/w]	0.195	0.404	0.271
RMSECV [% w/w]	0.213	0.436	0.309
RMSEP [% w/w]	3.692	2.750	0.338
R^2 (calibration)	0.997	0.981	0.994
R^2 (cross-validation)	0.996	0.978	0.992
R^2 (prediction)	0.145	0.398	0.981
LoQ [% w/w]	1.98	4.19	3.17
LoD [% w/w]	0.65	1.38	1.05

* bl: baseline correction; SNV: standard normal variate; nm (1-norm): normalization for area under curve; mncn: mean centering.

3.3. Testing the method with a real-time experiment

The performance of the combined models described in Sections 3.1 and 3.2 was tested by recording in-line Raman and NIR spectra during an independent continuous granulation experiment. The purpose of this experiment was to evaluate how the models can track changes in the granule composition. This was realized by adjusting the L/S ratio by changing the rotational speed of the peristaltic pump responsible for dosing the granulation liquid. Manufacturing was operated for two hours, during this in-line Raman and NIR spectra were recorded at every 30 s and 12 s, respectively, as well as 24 samples were taken for off-line analysis. The CAR concentration of these samples was measured using HPLC. This way, the predictions of the spectroscopy methods could be compared to actual measured values.

Initially, it was found that the combined models based solely on the off-line calibration samples described in Sections 3.1 and 3.2 did not yield accurate predictions when they were deployed to monitor the real-time experiments. Therefore, the models were augmented with new samples collected during in-line monitoring. These samples were measured in-line with the spectrometers, then they were collected from the conveyor belt and HPLC was used to determine their true CAR concentration. The spectra of these samples were then included in the training dataset of the PLS models. The samples were collected from two separate experiments, meaning 24 samples from the first and 17 from the second experiment. Table 3 shows the parameters of the augmented models.

After augmentation with the in-line samples, both models showed increased RMSEC and RMSECV values. However, despite this, their predictions improved in the case of new in-line samples. Fig. 5 shows their performance during an independent in-line experiment by comparing the predicted PVP and CAR concentration with the results of HPLC measurements.

Generally, it can be observed that the predicted concentration values correlate well with both the target values and the concentrations measured with HPLC. It is worth noting that during the intervals when the target concentration is the same, the HPLC measurements still show fluctuations, which is also detected by both spectroscopic methods. In certain data points, the spectra predict unrealistically large or small values. These spectra could be associated with large Hotelling T^2 values. The Hotelling T^2 gives a numeric representation of the variation in the spectra that is not explained by the PLS model. Large values imply that the spectrum deviates from the spectra that were used to create the model, therefore predictions based on this spectrum are not reliable. An acceptance limit for the Hotelling T^2 value was created by calculating the average and standard deviation of the T^2 values obtained for the calibration samples, and the limit was the average plus 10 times the standard deviation. This limit was 39.32 for NIR and 28.87 for Raman spectra, above this spectra were marked as outliers. These faulty spectra most likely occurred because during the measurement, the amount of material on the conveyor was not sufficient. Furthermore, the Raman spectrometer is also sensitive to ambient light, thus noise caused by external factors can also be responsible for outlier results. Based on the Hotelling T^2 acceptance values, it is possible to exclude these spectra from the analysis.

To quantitatively compare the in-line performance of the spectroscopic methods, the 3 spectroscopy predictions closest to the time point of the off-line sample collection were averaged and the RMSEP values were calculated by comparing these values to the HPLC measurements. The RMSEP of PVP concentration determination were 0.35 % w/w for Raman spectroscopy and 0.27 % w/w for NIR spectroscopy. In the case of CAR concentration, Raman spectroscopy has an RMSEP of 7.07 ppm and NIR spectroscopy has a value of 5.31 ppm. With a target concentration of 60 ppm or higher (here the relative error of Raman is 11.7 % and NIR is 8.9 %), these results are enough to reliably recognize outlier samples that are not within 85–115 % of the target concentration. Consequently, this method can be used for the real-time content uniformity of samples prepared on a TSWG. These results are interesting considering the LoQ values in Table 3. Both spectroscopic techniques have a relatively high LoQ (6.03 % w/w for NIR and 3.17 % w/w for Raman spectroscopy). However, despite the fact that almost all measurements happened below the LoQ of NIR spectroscopy, it yielded more accurate results than Raman spectroscopy. Most likely this can be attributed to the fact that the models were created by combining spectra recorded under various conditions, therefore there is a rather high variability between samples belonging to the same concentration level. Eventually this manifests in worse calibration performance metrics including LoD and LoQ values. This shows that the inclusion of diverse spectra resulted in more robust models that do not yield good values during calibration, but perform well with independent test samples.

These results extend the lower limit of drug concentration measurement in pharmaceutical applications compared to previously published work. Bostijn et al. applied in-line UV spectroscopy to measure the concentration of a drug in a suspension of 0.09 % w/w (Bostijn et al., 2018). Griffen et al. used transmission Raman spectroscopy to measure the concentration of warfarin salts in tablets in concentrations between 0.25 and 0.7 % w/w. They achieved an LoQ value of 0.19 % w/w with this off-line measurement technique (Griffen et al., 2018). Igne et al. utilized light-induced fluorescence spectroscopy to measure the drug concentration of granules inside the feed frame of a tablet press (Igne et al., 2021). The target concentration was 0.89 % w/w. With the proposed indirect measurement technique, it is possible to predict concentrations one or two magnitudes lower compared to the previously mentioned works.

These results show that the models are capable of accurately quantifying the API concentration even in such a low dose. Furthermore, the HPLC results validate that the concentration of PVP and CAR change together in the granules. The results indicate that the indirect concentration measurement of an ultralow dose API is feasible based on NIR and Raman spectroscopy. However, before using this technique with other systems, it would be necessary to check whether this correlation is true in the case of that particular product. The application of the proposed technique would enable manufacturers to obtain real-time API concentration information during the manufacturing of ultralow dose products. This would eliminate the need for frequent sample taking from highly potent powders and the high organic solvent and labour cost of measuring numerous samples with HPLC. Using the technique to completely replace HPLC measurements would require a more thorough validation process according to pharmacopoeial recommendations. This procedure could also define acceptance levels by considering the

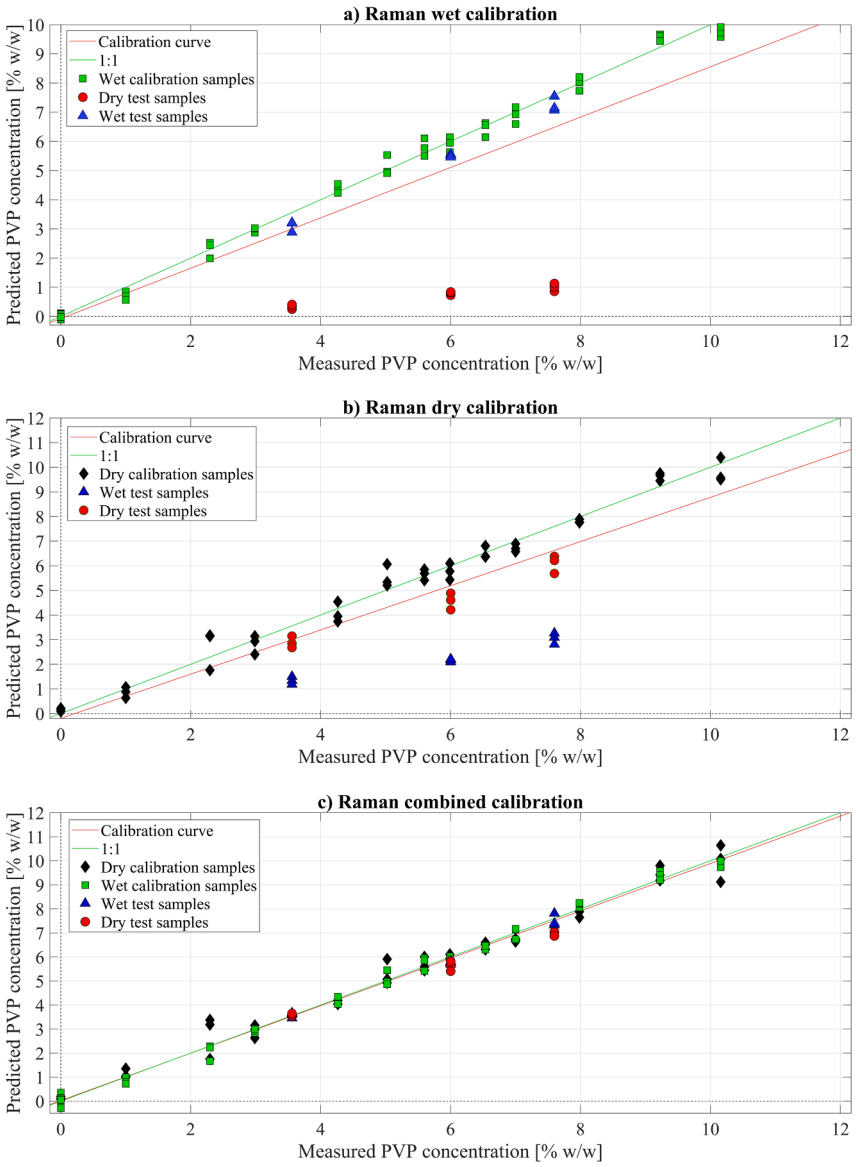


Fig. 4. Raman models for the prediction of PVP concentration.

Table 3
The parameters of the PLS models augmented with in-line samples.

Model type	NIR combined augmented	Raman combined augmented
Variable selection	**iPLS**	**none**
Spectral range (cm^{-1})	9403.62–7498.21; 5450.09–4597.67	1599–600
Preprocessing	1stDer, SNV, mncn*	bl, nm (1-norm), mncn*
Cross validation	contiguous block with 10 splits	contiguous block with 10 splits
Number of latent variables	3	4
RMSEC [% w/w]	0.708	0.329
RMSECV [% w/w]	0.777	0.375
R^2 (calibration)	0.941	0.986
R^2 (cross validation)	0.929	0.982
LoQ [% w/w]	6.03	3.17
LoD [% w/w]	1.99	1.05

measurement error of the technique.

The presented methodology could be used for several drug products with ultralow API content, with the prerequisite that the drug substance is in the liquid phase of the TSWG process, – which is a suggested approach for ultralow dose formulations for homogeneity purposes. This would give operators of the technique real-time information about the dose of the product leaving the granulator. Even if the technique does not eliminate the need of HPLC measurements, its value lies in the ability to alert operators about large deviations in the concentration, therefore they have the chance to react accordingly and compensate for the deviation. The method is especially promising in the case of high-potent drugs (*e.g.*, hormones, cytostatics) that are often present in low concentrations and can have strong side effects if not used in the recommended dosage.

4. Conclusion

The proposed indirect API concentration measurement technique was successfully applied for the real-time monitoring of TSWG. By monitoring the PVP concentration of the granules, the presented methods could predict the concentration of CAR. This could be realized because both materials were dissolved in the granulation liquid, therefore their concentrations were proportionate in the granules. During the real-time monitoring of an experiment validated with HPLC, the NIR

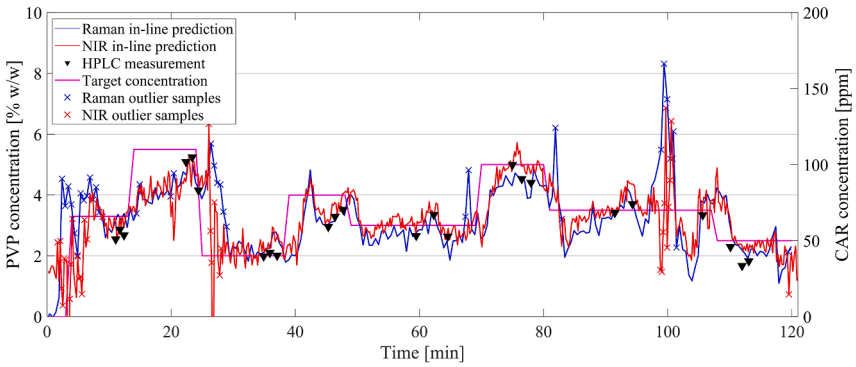


Fig. 5. Concentrations predicted using Raman and NIR spectroscopy and measured using HPLC.

spectroscopy-based method predicted the API concentration with an RMSE value of 5.31 ppm, while Raman spectroscopy had an RMSE of 7.07 ppm. Concentrations in this range can not be measured directly with these spectroscopic methods. This accuracy enables a real-time, non-destructive content uniformity and assay measurement of granules produced using TSWG. The technique is expected to work similarly in the case of other low-dose products where the API is dissolved in the granulation liquid.

The research also provided insights about the efficient chemometric model building. The moisture content of the granules needs to be considered because without a carefully constructed training dataset, the created chemometric models can confuse the signals of the binder material with the solvent, resulting in poor prediction for samples with different moisture content. This issue could be circumvented by including both wet and dry granule samples in the training dataset. The models built this way were found to be robust to changes in moisture content. Our results also show during model creation, chasing the lowest RMSE values of calibration and cross validation is not always the best way, because ultimately, a more robust model with modest parameters can perform better when unknown samples are measured.

CRediT authorship contribution statement

Dorián László Galata: Writing – review & editing, Writing – original draft, Visualization, Data curation. **András Domokos:** Validation, Methodology, Investigation. **Balázs Démuth:** Methodology, Investigation, Conceptualization. **Petra Záhonyi:** Methodology, Investigation. **Gergő Fülöp:** Methodology, Investigation. **Zsombor Kristóf Nagy:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization. **Brigitta Nagy:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Investigation, Formal analysis, Data curation.

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.

Data availability

Data will be made available on request.

Acknowledgement

Project no. RRF-2.3.1-21-2022-00015 has been implemented with the support provided by the European Union. This work was supported by OTKA grants K-143039 and PD-142970 and ÚNKP-23-5-BME-443 New National Excellence Program of the Ministry for Culture and Innovation from the Source of the National Research, Development and Innovation Fund. This paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Science.

References

Bandari, S., Nyavanandi, D., Kallakunta, V.R., Janga, K.Y., Sarabu, S., Butreddy, A., Repka, M.A., 2020. Continuous twin screw granulation–An advanced alternative granulation technology for use in the pharmaceutical industry. Int. J. Pharm. 580, 119215.

Bostijn, N., Hellings, M., Van Der Veen, M., Vervaet, C., De Beer, T., 2018. In-line UV spectroscopy for the quantification of low-dose active ingredients during the manufacturing of pharmaceutical semi-solid and liquid formulations. Analytica Chimica Acta 1013, 54–62.

Démuth, B., Fülöp, G., Kovács, M., Madarász, L., Ficzere, M., Köte, Á., Szabó, B., Nagy, B., Balogh, A., Csorba, K., 2020. Continuous manufacturing of homogeneous ultralow-dose granules by twin-screw wet granulation. Period. Polytechn. Chem. Eng. 64, 391–400.

Dhenge, R.M., Fyles, R.S., Cartwright, J.J., Doughty, D.G., Hounslow, M.J., Salman, A.D., 2010. Twin screw wet granulation: Granule properties. Chem. Eng. J. 164, 322–329.

Domokos, A., Pusztai, É., Madarász, L., Nagy, B., Gyürkés, M., Farkas, A., Fülöp, G., Casian, T., Szilágyi, B., Nagy, Z.K., 2021. Combination of PAT and mechanistic modeling tools in a fully continuous powder to granule line: Rapid and deep process understanding. Powder Technol. 388, 70–81.

European Medicines Agency, 1995. ICH Topic Q2 (R1) Validation of Analytical Procedures: Text and Methodology.

Ficzere, M., Mészáros, L.A., Madarász, L., Novák, M., Nagy, Z.K., Galata, D.L., 2021. Indirect monitoring of ultralow dose API content in continuous wet granulation and tableting by machine vision. Int. J. Pharm. 607, 121008.

Fonteyne, M., Soares, S., Vercruysse, J., Peeters, E., Burggraeve, A., Vervaet, C., Remon, J.P., Sandler, N., De Beer, T., 2012. Prediction of quality attributes of continuously produced granules using complementary pat tools. Eur. J. Pharm. Biopharm. 82, 429–436.

Fonteyne, M., Vercruysse, J., Díaz, D.C., Gildemyn, D., Vervaet, C., Remon, J.P., Beer, T. D., 2013. Real-time assessment of critical quality attributes of a continuous granulation process. Pharm. Dev. Technol. 18, 85–97.

Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. Analytica Chimica Acta 185, 1–17.

Griffen, J.A., Owen, A.W., Matousek, P., 2018. Quantifying low levels (< 0.5% w/w) of warfarin sodium salts in oral solid dose forms using Transmission Raman spectroscopy. J. Pharm. Biomed. Anal. 155, 276–283.

Gyürkés, M., Madarász, L., Záhonyi, P., Köte, Á., Nagy, B., Pataki, H., Nagy, Z.K., Domokos, A., Farkas, A., 2022. Soft sensor for content prediction in an integrated continuous pharmaceutical formulation line based on the residence time distribution of unit operations. Int. J. Pharm. 624, 121950.

Harting, J., Kleinebudde, P., 2018. Development of an in-line Raman spectroscopic method for continuous API quantification during twin-screw wet granulation. Eur. J. Pharm. Biopharm. 125, 169–181.

Harting, J., Kleinebudde, P., 2019. Optimisation of an in-line Raman spectroscopic method for continuous API quantification during twin-screw wet granulation and its application for process characterisation. Eur. J. Pharm. Biopharm. 137, 77–85.

Igne, B., Baldasano, C.N., Airiau, C., 2021. Feasibility of using light-induced fluorescence spectroscopy for low-dose formulations monitoring and control. J. Pharm. Innov. 16, 391–397.

Liu, B., Wang, J., Zeng, J., Zhao, L., Wang, Y., Feng, Y., Du, R., 2021. A review of high shear wet granulation for better process understanding, control and product development. Powder Technol. 381, 204–223.

Madarász, L., Köte, Á., Hambalkó, B., Csorba, K., Kovács, V., Lengyel, L., Marosi, G., Farkas, A., Nagy, Z.K., Domokos, A., 2022. In-line particle size measurement based on image analysis in a fully continuous granule manufacturing line for rapid process understanding and development. Int. J. Pharm. 612, 121280.

Martinetz, M., Karttunen, A., Sacher, S., Wahl, P., Ketolainen, J., Khinast, J., Korhonen, O., 2018. RTD-based material tracking in a fully-continuous dry granulation tableting line. Int. J. Pharm. 547, 469–479.

Meng, W., Román-Ospino, A.D., Panikar, S.S., O'Callaghan, C., Gilliam, S.J., Ramachandran, R., Muzzio, F.J., 2019. Advanced process design and understanding of continuous twin-screw granulation via implementation of in-line process analytical technologies. Adv. Powder Technol. 30, 879–894.

Nomura, K., Titapiwatanakun, V., Hisada, H., Koide, T., Fukami, T., 2020. In situ monitoring of the crystalline state of active pharmaceutical ingredients during high-shear wet granulation using a low-frequency Raman probe. Eur. J. Pharm. Biopharm. 147, 1–9.

Palmer, J., O'Malley, C., Wade, M., Martin, E.B., Page, T., Montague, G., 2020. Opportunities for process control and quality assurance using online NIR analysis to a continuous wet granulation tableting line. J. Pharm. Innov. 15, 26–40.

Peters, J., Bartscher, K., Döscher, C., Taute, W., Höft, M., Knöchel, R., Breitkreutz, J., 2017. In-line moisture monitoring in fluidized bed granulation using a novel multi-resonance microwave sensor. Talanta 170, 369–376.

Petrović, J., Chansanroj, K., Meier, B., Ibrić, S., Betz, G., 2011. Analysis of fluidized bed granulation process using conventional and novel modeling techniques. Eur. J. Pharm. Sci. 44, 227–234.

Porep, J.U., Kammerer, D.R., Carle, R., 2015. On-line application of near infrared (NIR) spectroscopy in food production. Trends Food Sci. Technol. 46, 211–230.

Renaud, O., Victoria-Feser, M.-P., 2010. A robust coefficient of determination for regression. J. Statist. Plann. Inference 140, 1852–1862.

Román-Ospino, A.D., Tamrakar, A., Igne, B., Dimaso, E.T., Airiau, C., Clancy, D.J., Pereira, G., Muzzio, F.J., Singh, R., Ramachandran, R., 2020. Characterization of NIR interfaces for the feeding and in-line monitoring of a continuous granulation process. Int. J. Pharm. 574, 118848.