European Journal of Pharmaceutics and Biopharmaceutics 134 (2019) 20–28

Contents lists available at ScienceDirect

European Journal of Pharmaceutics and Biopharmaceutics

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

Research paper

Evaluation of chemometric approaches for polymorphs quantification in tablets using near-infrared hyperspectral images

Vitor H. da Silva[a], José L. Soares-Sobrinho[b], Claudete F. Pereira[a], Åsmund Rinnan[c,*]

[a] Department of Fundamental Chemistry, Federal University of Pernambuco, Av. Prof. Moraes Rego, 1235, Cidade Universitária, Recife, Brazil
[b] Department of Pharmaceutical Science, Federal University of Pernambuco, Av. Prof. Moraes Rego, 1235, Cidade Universitária, Recife, Brazil
[c] Department of Food Science, University of Copenhagen, Rolighedsvej 26, 1958 Frederiksberg C, Copenhagen, Denmark

ARTICLE INFO

Keywords:
Hyperspectral images
PLS
MCR-ALS
Polymorphs
Mebendazol
Multivariate calibration

ABSTRACT

Near-Infrared hyperspectral imaging (HSI-NIR) is a useful technique for pharmaceutical research and industry alike. It can provide important surface information such as the polymorphs quantification and its distribution over the tablet. Several chemometric tools are applied for this purpose, with MCR-ALS and PLS regression being the most common approaches. In this work, a detailed comparison between these two approaches is performed. Beyond a "simple" regression comparison, a comparison of the score images (local quantification) was also evaluated. The system under study is tablets with ternary mixtures of Mebendazol (MBZ) polymorphs, microcrystalline cellulose and magnesium stearate. PLS models, in general, gave lower RMSEP (below 1.7% w/w for the three MBZ polymorphs) than the corresponding MCR-ALS predictions. Analyzing the distributions of the scores in the images of each sample shows clear differences between the PLS and MCR-ALS models. The MCR-ALS gave more chemical meaningful distribution maps for all polymorphs, even though the PLS accurately predicts the average concentration across the image. The problem is that the PLS models used the main spectral regions to quantify each MBZ polymorph, but at the same time undermines the minor spectroscopic changes caused by the different polymorphs. Although this may seem as a minor deviation from the truth, the results clearly show that this deviation is detrimental for the analysis of the spatial distribution of the analytes. These results indicate that the optimal multivariate model for multivariate images depend on the goal of the analysis: global quantification or a distribution analysis.

1. Introduction

Near infrared spectroscopy (NIRS) is widely and successfully used for pharmaceutical applications due to its inherent characteristics such as non-destructivity, no sample preparation, fast analysis and easy applicability for industrial process monitoring [1,2]. This vibrational spectroscopy enables quantitative and qualitative analysis and is also a common method for pharmaceutical solid-state assessment such as for polymorphs analysis [3–5].

Polymorphism means the multiple crystalline forms of a substance. Several active pharmaceutical ingredients (API) are characterized by the actual polymorph. Each polymorph displays individual chemical, physical and process properties without undergoing any change in their chemical composition [6–8]. Moreover, these different properties may affect the drug therapeutic efficiency and product manufacturing such as the powder flow, compressibility, tablet stability, etc [9,10]. Therefore, it is important to detect and quantify pharmaceutical polymorphs in bulk and final product to ensure the presence of a correct or acceptable solid-state form and also to avoid health and manufacture problems, such as the famous Ritonavir case, which several manufactured lots were botched due to the appearance of a new polymorph [11]. Mebendazole (MBZ) is a anthelmintic drug that needs careful attention as it displays three polymorphs (A, B and C) with different and critical properties [12,13]. For instance, polymorph A does not present the desired effect when it exceeds 30% in a formulation due to its low solubility. Polymorph B is the more toxic of the polymorphs [14,15], while polymorph C is the pharmaceutically preferred. Nonetheless, all three MBZ polymorphs are found in commercial products [16–18]. Vibrational spectroscopy such as NIRS has been used successfully for the identification and quantification of MBZ polymorphs in raw material [16,19,20]. Ayala et al. characterized MBZ polymorphs using vibrational spectroscopy [16]. Silva et al. described analytical methods using NIRS for MBZ polymorphs quantification using benchtop and portable instruments [19,20].

* Corresponding author.
E-mail address: aar@food.ku.dk (Å. Rinnan).

https://doi.org/10.1016/j.ejpb.2018.11.007
Received 31 July 2018; Received in revised form 5 November 2018; Accepted 6 November 2018
Available online 07 November 2018
0939-6411/

Nowadays, the association of conventional NIRS instruments with cameras makes the acquisition of chemical images in this electromagnetic spectral region possible - so-called Near Infrared-Hyperspectral Images (HSI-NIR). In a HSI-NIR systems, a NIR spectrum is measured for each image pixel and a data cube (3D) is generated (height by width by wavelength). Beyond chemical information (spectral information) obtained with the image samples, spatial information, such as the API distribution, becomes possible. The HSI-NIR provides a large amount of information with several thousands of pixels per image, each measured at many wavelengths, making the subsequent data analysis complex. At the same time, the possible informational output is drastically increased. Multivariate data analysis have shown to be an efficient tool to extract useful information from this type of data [21–24]. Several applications of HSI-NIR coupled with chemometrics for different pharmaceutical application such as determination of API content [25] and its degradation product [26] have been shown. They unveil multiple solid-state transitions [27], and analysis of pharmaceutical pallets [28]. In terms of pharmaceutical quantitative analysis, the most common chemometrics approaches for hyperspectral images are Multivariate Curve Resolution – Alternative Least Squares (MCR-ALS) and Partial Least Squares (PLS) regression [29].

MCR-ALS and PLS are chemometrics approaches well established and widespread with several references describing them in detail [30–33]. Described very briefly, PLS regression aims to correlate spectral information with a dependent variable such as pharmaceutical bulk content. Thus, a model is built and its regression coefficients are used to predict the property of interest from new sample spectra [32,33]. On the other hand, MCR is a curve resolution method that decomposes the data matrix into its pure response profiles (e.g. spectra) and their relative concentrations. MCR works in an iterative way (Alternating Least Squares – ALS) to achieve the best data decomposition. The ALS algorithm allows for several types of constraints (e.g. non-negativity in concentration/ spectral profile) to improve and reach chemical reasonable MCR solutions [30,31].

For polymorph determination, both MCR-ALS and PLS regression have been utilized together with data from HSI-NIR. Rocha et al. [34] described an analytical method to quantify two piroxicam polymorphs in tablets using NIRS images and PLS regression. The authors determine local and global polymorphs concentration in the tablets with prediction errors below 4% w/w. Siddiqui et al. [35] developed models based on PLS regression to predict the nimodipine polymorph concentration in powder mixtures using hyperspectral images. Schönbichler et al. [36] evaluated HSI-NIR for quantifying furosemide crystal polymorphs in ternary powder mixtures using PLS regression. The quantitative results obtained with the images were comparable with the ones obtained using conventional NIRS, Raman and IR spectroscopies. Brondi et al. [37] evaluated fenofenadine polymorph distribution in tablets by HSI-NIR. PLS and MCR-ALS were evaluated and compared for polymorphs quantification. The author concluded that MCR-ALS efficiently quantified the polymorphs, but the PLS model recovered the concentration better. In that work, the models' regression parameters were evaluated to compare the prediction accuracy.

According to the described works, PLS regression is the most common approach. Whenever it is compared with MCR-ALS, the focus is on comparing the regressions parameters. However, it is necessary to evaluate the multivariate approaches also by the distribution maps, i.e. analyzing the local concentration prediction within the tablet. In the case of polymorphs analysis, the vibrational spectral profiles between the polymorphs are often similar and it complicates the analysis. Detailed comparison between the two most common chemometrics tools (MCR-ALS and PLS) for polymorphs quantification using HSI-NIR is necessary. Therefore, this work aims to evaluate the analytical performance of MCR-ALS and PLS approaches to quantify global and local concentration of MBZ polymorphs in pharmaceutical tablets.

2. Material and methods

2.1. Samples

Mebendazole polymorphs A and C were supplied by Formil Química (São Paulo, Brazil) and polymorph B were obtained by the recrystallization method described by Kachrimanis et al. and Silva et al. [18,19]. The polymorphs were characterized by IR spectroscopy and Powder X-ray diffraction (PXRD). IR spectra were recorded using a FTIR Spectrum 400 spectrometer (Perkin Elmer) using a universal attenuated total reflectance (U-ATR) accessory in the 4000–650 cm^{-1} range, with a 4 cm^{-1} resolution and an average of 32 scans. PXRD diffractograms were recorded using a D8 advanced diffractometer (Bruker) using Kα radiation of copper (1.54 Å). The samples were analyzed with a scan rate of 0.02° 2θ/s in the angular range of 3–35° 2θ at room temperature. The MBZ patterns were compared with those described in the literature [13,14,18] and after positive agreement, the samples were sieved into a 355 μm (45 mesh screen).

A mixture design, based on the Simplex-Lattice design [38], was used to prepare 25 ternary mixtures of MBZ polymorphs, varying polymorph A and C from 0 to 100% and polymorph B in the 0–30% range. Simplified pharmaceutical tablets containing the polymorph mixtures and the excipients (microcrystalline cellulose – MCC, and magnesium stearate – MgSt) were prepared. All tablets with polymorph mixtures, MCC and MgSt presented 50:49:1 of mass fraction range with total weight of 100 mg. All the samples were weighted in analytical balance (CP225D, Sartorius) with an accuracy of 0.01 mg. Each sample powder mixture was mixed by vortexing for 3 min in 15 mL recipient to ensure homogeneity in the tablets before pressing. The samples were pressed in a mechanical press under to 2 ton/cm^2 to form the MBZ tablets with a diameter of 8.5 mm. Table 1 shows the amount of MBZ polymorphs in each tablet.

2.2. NIR images acquisition

Near infrared hyperspectral images were collected using a SisuCHEMA (Specim) hyperspectral camera in the spectral range of 900–2500 nm and a spectral resolution of 10 nm. This camera measures

Table 1
Amount of MBZ polymorphs in each tablet.

	Polymorph A (% w/w)	Polymorph B (% w/w)	Polymorph C (% w/w)
M1	42.5	7.5	0
M2	0	15	35
M3	35	15	0
M4	21.5	7.5	21.5
M5	50	0	0
M6	25	0	25
M7	0	0	50
M8	0	7.5	42.5
M9	17.5	15	17.5
M10	37.5	4	8.5
M11	30	1.5	18.5
M12	9.5	0.5	40
M13	15	2.5	32.5
M14	40	6	4
M15	12.5	9	28.5
M16	20	12.5	17.5
M17	30	7.5	12.5
M18	10	10	30
M19	45	1	4
M20	16.5	2	31.5
M21	13.5	3	33.5
M22	7.5	5	37.5
M23	3.5	9.5	37
M24	2.5	11	36.5
M25	1.5	13.5	35

cross-sections of the sample, and the distance between the cross-sections is 6.3 nm. The images were measured using a macro lens with 10 mm of field of view and pixel size of 30 × 30 µm, resulting in an average image size of 300 × 300 pixels.

2.3. Data analysis and software

The first step in the data analysis, is to select the Region of Interest (ROI) in each image to eliminate the background. The spatial pre-processing was made by Principal Component Analysis (PCA) frequency histograms obtained from the score images [39]. Afterwards, the images were unfolded from three-dimensional to two-dimensional matrices by unfolding in the spatial direction.

Different pre-processing strategies were applied on the NIR spectra to minimize the instrumental and physical artefacts that are not related with the chemical behavior of the samples and pixels, such as noise and light scattering. Standard Normal Variate (SNV), Multiplicative Signal Correction (MSC) and Savitzky-Golay 1st and 2nd derivative were all investigated [40]. On the pre-processed spectra, MCR-ALS and PLS models were developed to quantify the MBZ polymorphs in the tablet and visualize their surface distribution. The optimal pre-processing method was decided based on a set of figures of merit. The MCR-ALS models were evaluated by the Lack of Fit (LOF), explained variance, residuals and correlation coefficient between the experimental and the recovered data, while the PLS models were evaluated by the RMSEC, RMSECV and the correlation coefficient between the actual and predicted values. Through this analysis, the optimal pre-processing method was found to be the Savitzky-Golay routine using a window size of 15, a polynomial degree of 2 and the estimated 2nd derivative.

The samples were partitioned into a calibration and test set containing 15 samples and 10 samples, respectively. The same calibration and test set was used for both the PLS and the MCR-ALS models.

MCR-ALS models were developed using the hyperspectral images compressed by data binning. This compression technique can be applied in both spectral or spatial dimension and it replaces the original data by the mean value within a determined pixel or wavelength window. This data compression was used to decrease the computational time and it was applied in the spatial dimension in order to reduce the number of pixels by a factor of two. After data binning, each image sample included around 14,000 pixels. In addition, row-wise augmented matrices with the calibration subset samples were made to perform the models; effectively positioning the images below each other in the data matrix. The tablet components (MBZ polymorphs, MCC and MgSt) spectra were provide as initial estimation and non-negativity constraints were applied on the concentration profiles. Finally, the mean resolved concentration value for each image was regressed against the global concentration for each polymorph in the image. In this way, the relative concentration is turned into absolute concentration by the regression line. The recovered MCR-ALS spectra were used to determine the concentration profile for the prediction samples and its global concentration was determined by the calibration model described above. The quality of the calibration model was evaluated by the root mean square error of calibration (RMSEC) and the regression parameters. More details about quantification based on MCR-ALS can be found in a review made by Piqueiras et al. [31].

Partial Least Squares (PLS) regression models were developed using the mean pre-processed spectra of each image. Full cross validation was carried out to select the optimal number of latent variables (LV) in the regression models. The number of LVs was determined from the minimum root mean squared error of cross validation (RMSECV). It should be noted that both the average image and the content of the polymorph where mean-centered.

The predictive ability of the regression models developed in this work was evaluated by the root mean square error of prediction (RMSEP) of the test set and by the regression parameters. The Limit of Detection (LOD) values for each models were determined according to

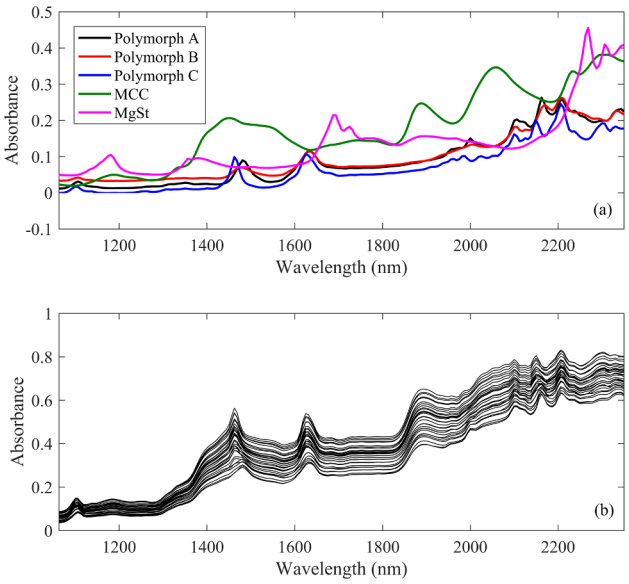

Fig. 1. (a) NIR spectra for each pure tablet component. (b) Average NIR spectra for each tablet.

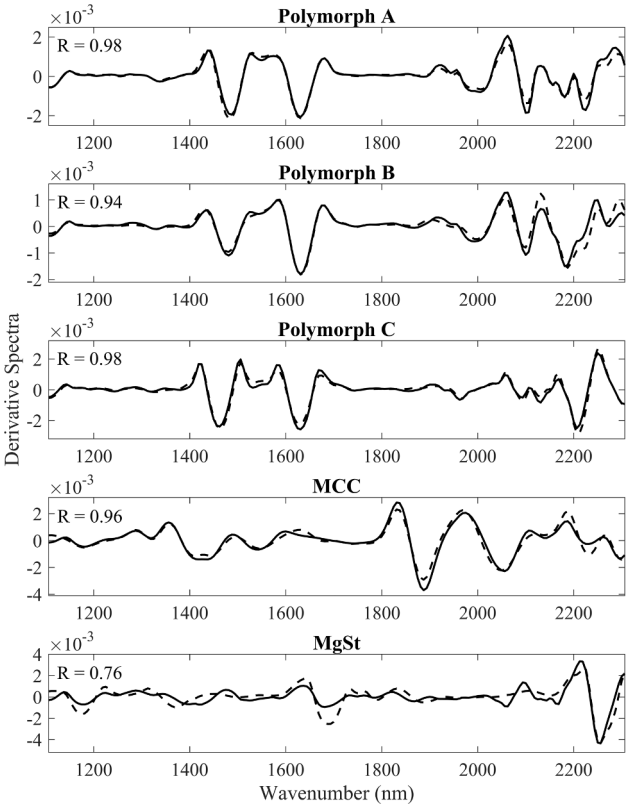

Fig. 2. Pre-processed NIR spectra (line) and recovered spectra by MCR-ALS (dashed) with their respective correlation coefficient.

Booksh and Kowalski [41]. Distribution maps for the test set were obtained by refolding the predicted concentration at each pixel of the HSI.

All chemometrics calculations were performed using Matlab software (Mathworks®), PLS Toolbox (Eigenvector Research Inc.), Hypertools Toolbox [42] and MCR-ALS interface (available at www.mcrals.info) [30].

Table 2
Overall results for PLS and MCR-ALS models for quantifying MBZ polymorph in tablets using HSI-NIR. The values in parentheses refers the number of latent variables used in the PLS regression models.

	Polymorph A				Polymorph B					Polymorph C				
	RMSEC (% w/w)	R^2	RMSEP (% w/w)	R^2	LOD (% w/w)	RMSEC (% w/w)	R^2	RMSEP (% w/w)	R^2	LOD (% w/w)	RMSEC (% w/w)	R2	RMSEP (% w/w)	R^2	LOD (% w/w)
PLS	1.65 (4)	0.99	1.71	0.98	2.06	1.09 (4)	0.95	1.00	0.96	2.15	1.28 (3)	0.99	1.12	0.99	2.29
MCR-ALS	2.87	0.97	4.56	0.95	6.65	1.59	0.91	2.87	0.88	5.10	1.64	0.99	2.22	0.96	4.62

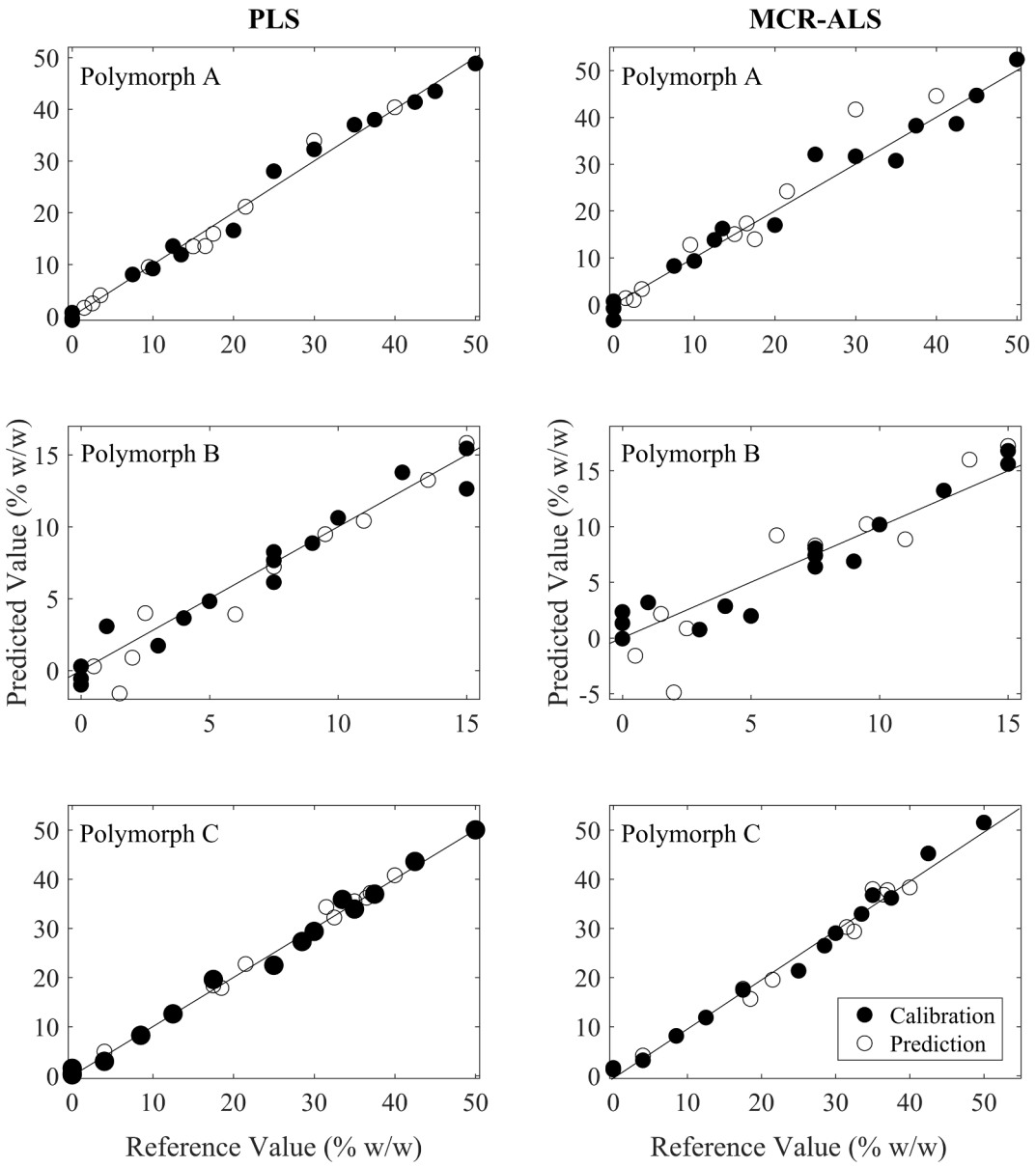

Fig. 3. Plots of reference values versus predicted values obtained by the PLS and MCR-ALS models for quantifying MBZ polymorphs in tablets.

3. Results and discussion

Fig. 1 shows the raw NIR spectra from the tablets used in this works and their components. Although the hyperspectral camera operates from 900 to 2500 nm, the NIR spectra were reduced to the spectral range of 1060–2350 nm to eliminate the high frequency noise observed at each end of the spectra. Fig. 1a shows the NIR spectra profile for each component in the tablets and their characteristic absorption bands. It is possible to identify the main characteristic band for the MBZ polymorphs, the first overtone of N-H stretching vibrations around 1450 nm. This band indicates differences in the intensity and width among the polymorphs due to their packing difference in the crystal lattice. The 2nd and 1st overtones of the C-H stretching vibration for the polymorphs arises around 1130 nm and 1650 nm, respectively. The polymorph bands above 2000 nm are assigned to combinations of C-H and N-H stretching and bending vibration. This spectral region is important to distinguish the polymorph A and B due to their overlapped spectral profile.

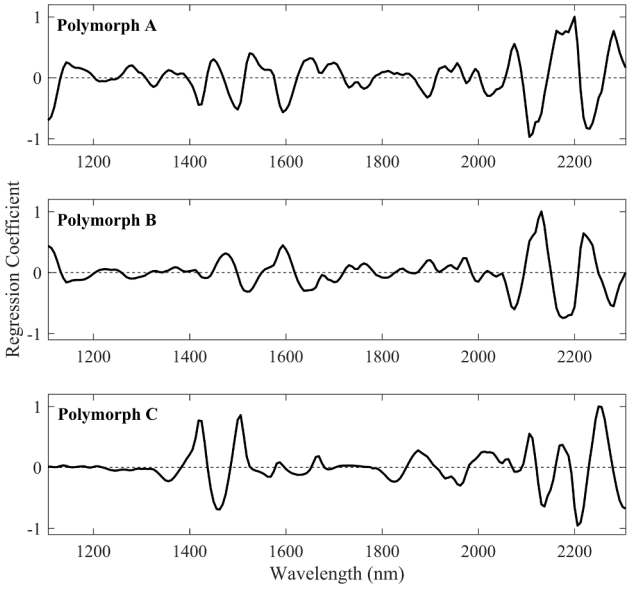

Fig. 4. PLS regression coefficients for MBZ polymorphs.

MCC has broad bands around 1500 nm, 1900 nm and 2100 nm related with the 1st and 2nd overtone of O-H stretching (the two first) and its deformation (at 2100 nm). As can be seen, the MMC spectrum overlaps the major bands used to distinguish the MBZ polymorphs. It is worth noting that the global MCC content in the images is the same (50% w/w) and its information on the mean spectra do not change. However, spectral variability for MCC is observed in the pixels' spectra due to its distribution across the tablet. Finally, the magnesium stearate spectral profile showed bands around 1150 nm and 1700 nm related with the 1st and 2nd overtones of C-H stretching, as well as characteristic combination bands (1st C-H stretching + 1st C-H deformation) around 2280 nm.

Fig. 1b shows the mean raw NIR spectra for all tablets. As can be seen, the main information observed in the mean spectra is related with the MBZ polymorphs and MCC. Spectral information about the magnesium stearate is not clear due to its low content (1% w/w) in the tablets. Scattering effect is observed in these measurements and this effect was removed prior to the analysis by pre-processing in the spectral mode.

Different pre-processing strategies were evaluated for MBZ quantification using both MCR-ALS and PLS regression. However, the second derivative with Savitzky-Golay smoothing filter with second order polynomial and 15 points window size provided the best results based on PLS and MCR-ALS predictive ability.

MCR-ALS models were built using two different strategies, (1) all the pre-processed NIR spectra after data binning and (2) their mean spectra for calibration samples. Between these models, the first strategy provided best results since more spectral variability were included and equality constrains can be used to obtain good recovered spectra. Fig. 2 shows the recovered and pre-processed MBZ spectra for the best MCR-ALS model and the respective correlation coefficient between the recovered MCR-ALS and the pure spectra.

Table 2 summarizes the results obtained for quantification of MBZ polymorphs using both MCR-ALS and PLS regression. For all polymorphs, the PLS models exhibit better results for global tablet concentration with lower calibration and prediction errors when compared with the MCR-ALS results. Probably the better analytical performance for global determinations using PLS is because the model was developed with the mean spectra, in other words just general tablet information was included in the model. On the other hand, the MCR-ALS models were based on the local information and the global concentration was obtained through an extra regression step. This means that it is likely that MCR-ALS is more accurate regarding local distribution of the chemical components across the whole image. In addition, for local quantification (distribution maps), any variability not reflected in the average image spectrum may affect the PLS distribution maps. Therefore, only the regression parameters (as shown in Table 2) are not enough to evaluate the analytical performance of these approaches for local determination. In order to include this a pixel by pixel evaluation is needed, as the models developed indicate that they could behave differently with regards to the local concentration evaluation.

Fig. 3 shows the reference versus predicted values plots for the models summarized in the Table 2. All samples are distributed around the target line demonstrating agreement between the predicted and nominal values. The higher random prediction error for the MCR-ALS models was expected due to how the model has been generated. Among

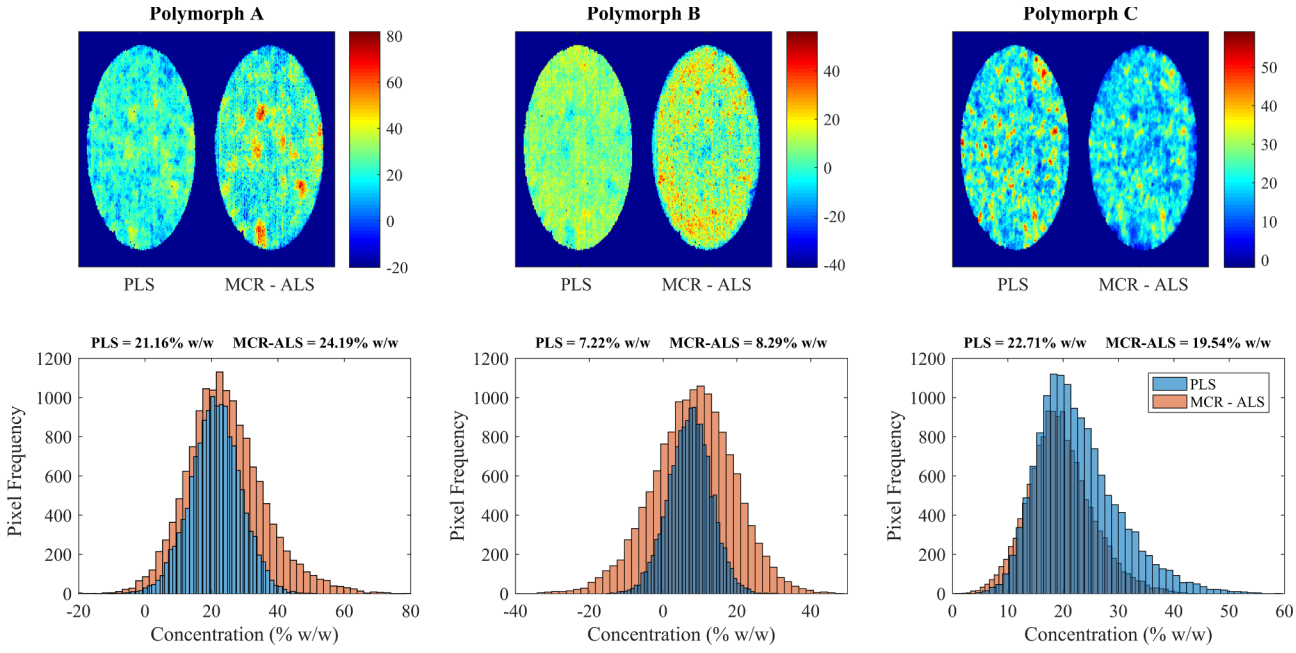

Fig. 5. Distribution maps and its frequency histograms for MBZ polymorphs on the sample M4.

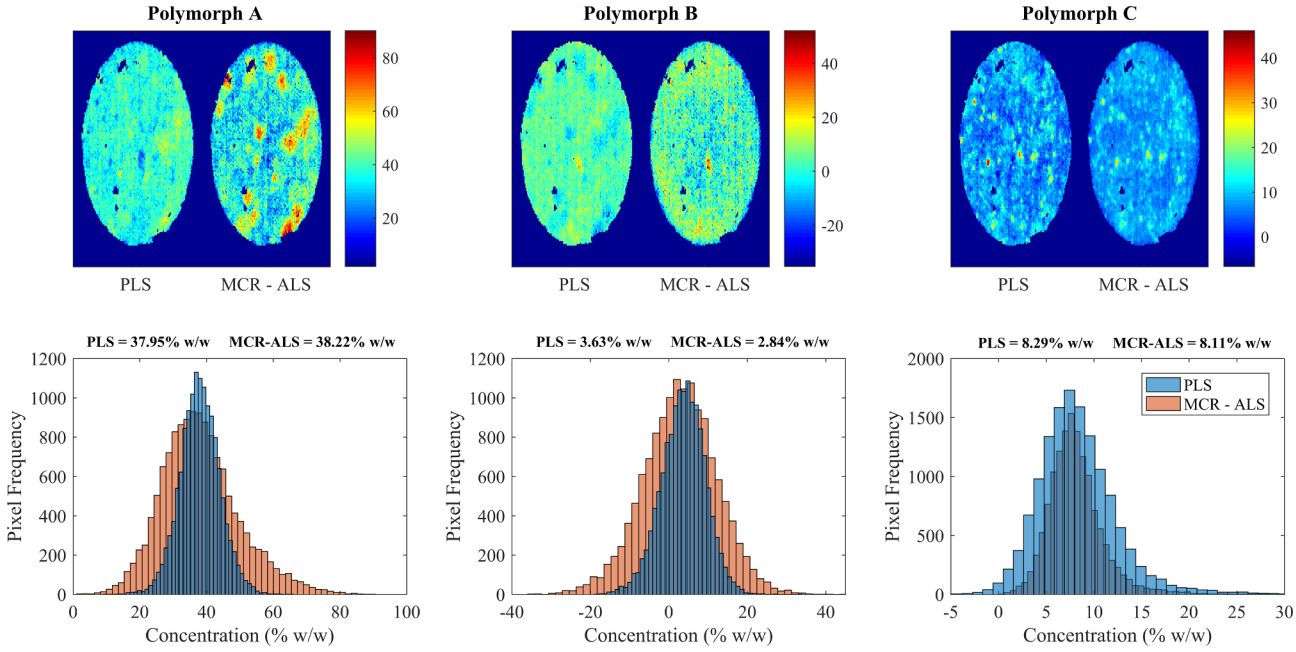

Fig. 6. Distribution maps and its frequency histrograms for MBZ polymorphs on the sample M10.

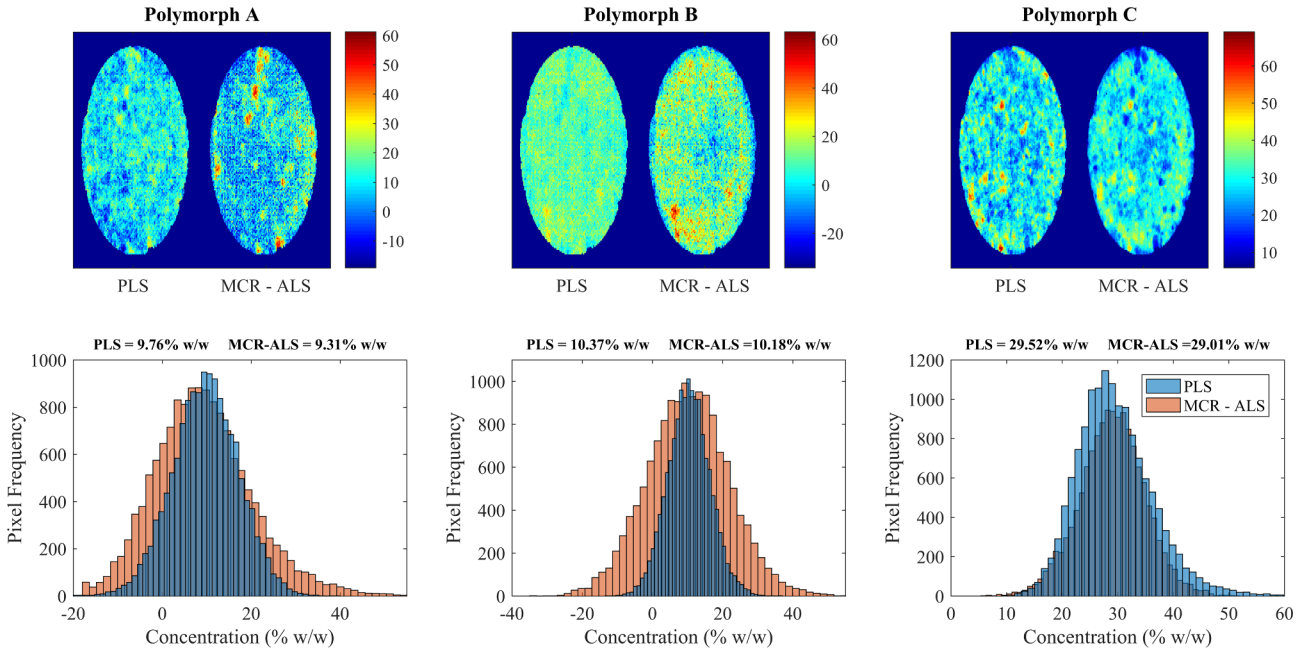

Fig. 7. Distribution maps and its frequency histrograms for MBZ polymorphs on the sample M18.

the polymorphs, the polymorph B showed the highest deviations from the target line. This is not a surprise as the polymorph A overlapped the polymorph B spectra making its quantification more difficult. To ensure that the PLS models were using information associated with the respective polymorphs, the Net Analytical Signal (NAS) from the MCR-ALS recovered spectra were calculated and compared with the PLS regression coefficient. The correlation coefficient between the NAS and the PLS regression coefficient were 78% for polymorphs A and B, and 59% for polymorph C. The largest differences between the NAS and PLS regression coefficient for each polymorph are related with the smaller peaks utilized during the quantification in the PLS models.

Considering the LOD (Table 2), the PLS models achieved values lower than those obtained with the MCR-ALS models. These results also showed the better sensitivity of the PLS models for global polymorph quantification in tablets. In addition, the PLS results obtained in this work were comparable with the analytical methods previously published using conventional NIR, MIR and Raman spectroscopies for quantify MBZ polymorphs in raw materials [19,20,43].

According to the PLS regression coefficient (Fig. 4), the main spectral region for quantifying polymorphs A and B is the region of combination bands. In addition to this region, polymorph C presented high regression coefficient values at the first overtone of N-H stretching vibration for its quantification. These spectral regions verified in the PLS regression coefficient showed agreement with the spectral bands used to characterize each MBZ polymorphs.

To evaluate the local concentration for both PLS and MCR-ALS models, several tablet images with different polymorphs concentrations were recovered using the models. Figs. 5–7 show the distribution maps

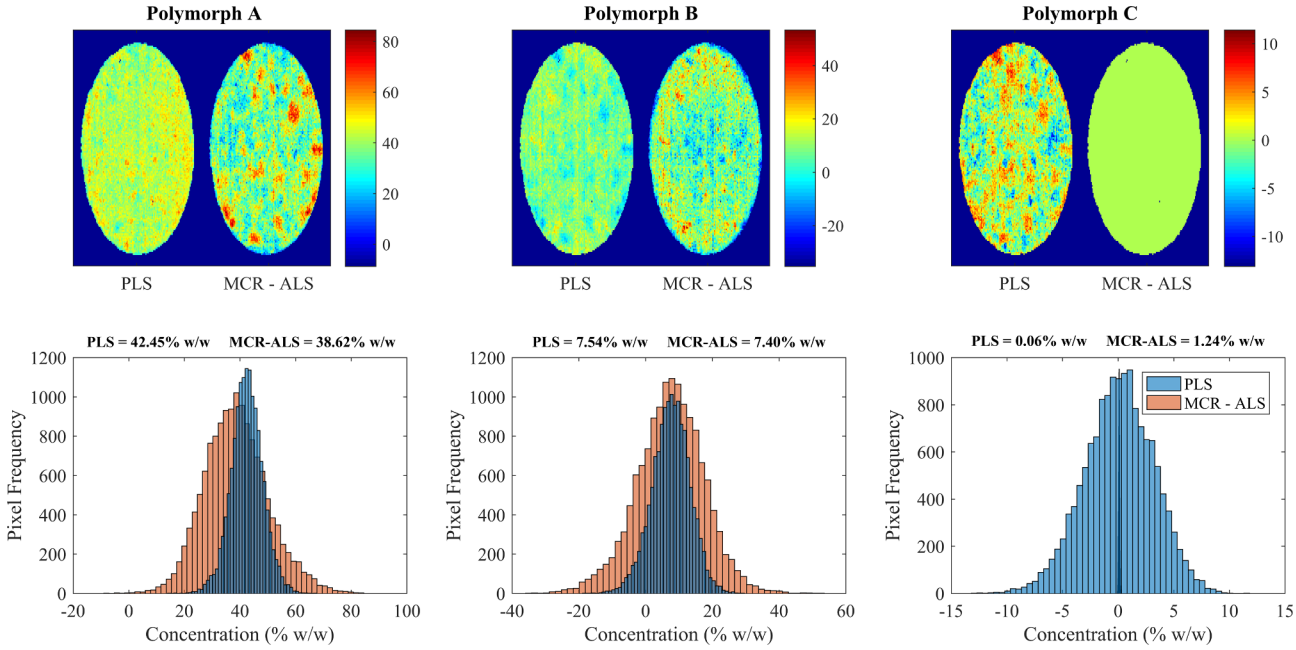

Fig. 8. Distribution maps and its frequency histograms for MBZ polymorphs on the sample M1.

for the MBZ polymorphs and its frequency histograms for the tablets M4, M16 and M18 (Table 1).

The different powder features among the MBZ polymorphs are very important for these evaluations, as this comparison cannot be performed quantitatively, only qualitatively. All the MBZ polymorphs are cohesive and easily agglomerates, with polymorph B with the lowest degree of agglomeration [44]. Since the tablets used in this work were powder blends, it is expected that the samples carry on the natural features of the powders. It means that the more aggregating powders (polymorph A and C) should indicate clusters on the tablet. While the others components are expected to be more homogenous.

The distribution maps and their histograms reveal some difference between the two strategies. For polymorphs A and B, the MCR-ALS gives a wider histogram compared with the PLS models, while the PLS models accurately predict the global quantity. However, with the MCR-ALS models it is possible to indicate the powder clusters in the distribution maps. These clusters indicate high content in that local region demonstrating the sample heterogeneity. However, the distribution maps calculated by PLS demonstrated more homogeneous tablet surfaces. The PLS regression coefficient (Fig. 4) showed that the combination bands are the most important region for quantifying polymorphs A and B. This spectral region (2000–2200 nm) is overlapped by the broad MCC band (Fig. 1a). As the PLS model uses the mean NIR spectra, the MCC information does not change and any local variability in the MCC is not taken into account. However, the MCR-ALS clearly shows that there is a variability in the local MCC concentration making the prediction by PLS more difficulty and less realistic as this variability is not taken into account in the PLS model. Furthermore, the local quantification by PLS for polymorph B is even more difficult due to this polymorph presented the lowest content in the tablet ($\leq$ 30% w/w) and its absorption bands are readily overlapped by the MCC.

For polymorph C the results showed opposite behavior than these obtained for the other two MBZ polymorphs on the recovered images. The PLS models show wider histograms and clear cluster visualization over the images. It is worth noting that the two main regions used for this quantification according to PLS regression coefficients are also highly overlapped by MCC spectra. It means that when the polymorph C content decreases across the tablet more MCC information is observed in the same spectral region. It is thus evident that the model used for understanding the local variability can cause misinterpretation of the distribution maps, as observed in the binary mixture sample M1 (Fig. 8).

Fig. 8 shows a tablet sample without the polymorph C. As can be seen, the distribution map for this polymorph by PLS evidenced some cluster over the image, although this global concentration (0.06% w/w) was accurate. The distribution map for polymorph C is similar to that observed for polymorph A. However, the PLS regression coefficient was not looking for polymorph A. The PLS model observed the lack of MCC in the sample, which for that sample is similar with the distribution maps for polymorph A due to its high content in the tablet. The regression coefficient for polymorph C has a band around 1900 nm related with the MCC and that region presented a negative contribution in the model.

For polymorph C, the MCR-ALS (Figs. 5–7) showed the most homogenous distribution maps among the polymorphs. As the MCR-ALS decomposed the data, its distribution maps for polymorph C are more realistic for the concentration range used and for that complexity data.

It is worth to highlight that the same information and behavior for all MBZ polymorphs were observed over the images. Even though the focus of this work is related with the quantification of the polymorphs, the tablet excipient distributions are only possible with the MCR-ALS strategy. The recovered images for the excipients were analyzed and all of them showed an agreement with the other compounds; and areas with high content of other compounds show low content of the major excipients.

As stated before, the regression parameters are not sufficient to evaluate analytical performance in different quantification approaches for images analysis. That kind of evaluation is of less importance when distribution maps also are of interest, which is the case of many applications of HSI. Sabin et al. [45] showed in their work that there is no difference between the PLS and MCR-ALS results for both local and global quantification in cabemezepine tablets. An agreement between the approaches used for global and local quantification is related with the data complexity, for instance the number of APIs, as well as the spectral features among the chemical components of the tablet. Therefore, quantification analysis in both global and local distribution are necessary and needs to be done carefully in quantitative analysis using hyperspectral images.

4. Conclusion

In this work a detailed comparison between MCR-ALS and PLS models to quantify MBZ polymorphs in tablets using HSI-NIR was presented. For global tablet determination, the PLS showed better prediction performance with lower calibration and prediction errors when compared with the MCR-ALS results. The superior performance for PLS for global quantification is because these models are developed with the average image spectra and only general information about the tablet was included.

For local quantification, analyzing the distribution maps, the MCR-ALS provides more meaningful results for all polymorphs, especially if one base the evaluation on earlier reported observations. The MCR-ALS results provided clearer and more realistic information about the polymorphs powder clustering on the tablet. According to the PLS regression coefficients, the important regions for quantification of MBZ polymorphs are related with the spectral region used to characterize them. However, these spectral regions are highly overlapped by the MCC spectral information and it makes the PLS results vulnerable to local variations in MCC.

These results demonstrate the efficiency of HSI-NIR to provide global polymorph content in a tablet and its distribution. Moreover, the prediction performance themselves are not enough to describe which chemometric model provides most meaningful distribution maps, since the recovered image quality depend of the spectral information about the polymorphs and their variability. PLS models provided better results for global quantification probably because the model is based on the average image spectra and any variability over this is not considered in the model. On the other hand, MCR-ALS models include more spectral information about the samples and can better separate the polymorphs information for local distribution. Therefore, distribution maps should be evaluated for different chemometrics approaches even though a specific model provided good global prediction results. Thus, the optimal model depends on the goal of the analysis of HSI-NIR. If the global quantification is important, PLS should be chosen. On the other hand, if one is interested in accurate distribution maps, MCR-ALS should be the preferred model.

Acknowledgments

The authors are grateful to FACEPE/NUQAAPE (APQ-0346-1.06/14) and INCTAA (Proc. CNPq No. 465768/2014-8, FAPESP Proc. No. 2014/50951-4). Scholarship granted by CAPES/PDSE (Process n°.: 88881.134155/2016-01) and the University of Copenhagen we are also gratefully acknowledging. We also thank Dr. Everaldo P. de Medeiros at Embrapa Algodão (PB – Brazil) for allowing us to use the hyperspectral camera.

References

[1] M. Jamrógiewicz, Application of the near-infrared spectroscopy in the pharmaceutical technology, J. Pharm. Biomed. Anal. 66 (2012) 1–10, https://doi.org/10.1016/j.jpba.2012.03.009.
[2] J. Luypaert, D.L. Massart, Y. Vander Heyden, Near-infrared spectroscopy applications in pharmaceutical analysis, Talanta 72 (2007) 865–883, https://doi.org/10.1016/j.talanta.2006.12.023.
[3] A.A. Bunaciu, H.Y. Aboul-enein, V. Dang, Vibrational spectroscopy used in polymorphic analysis, Trends Anal. Chem. 69 (2015) 14–22, https://doi.org/10.1016/j.trac.2015.02.006.
[4] R.B. Chavan, N. Bhargavi, A. Lodagekar, N.R. Shastri, Near infrared spectroscopy: a tool for solid state characterization, Drug Discov. Today 22 (2017) 1835–1843, https://doi.org/10.1016/J.DRUDIS.2017.09.002.
[5] N. Chieng, T. Rades, J. Aaltonen, An overview of recent studies on the analysis of pharmaceutical polymorphs, J. Pharm. Biomed. Anal. 55 (2011) 618–644, https://doi.org/10.1016/j.jpba.2010.12.020.
[6] R. Purohit, P. Venugopalan, Polymorphism: an overview, Resonance. 14 (2009) 882–893.
[7] P.H. Karpinski, Polymorphism of active pharmaceutical ingredients, Chem. Eng. Technol. 29 (2006) 233–238, https://doi.org/10.1002/ceat.200500397.
[8] H.G. Brittain, Theory and principles of polymorphic systems, in: H.G. Brittain (Ed.), Polymorph. Pharm. Solids., vol. 192. Drugs Pharm. Sci., 2nd ed., Informa Helthcare, New York, 2009, pp. 1–23.
[9] J. Bauer, Polymorphism—a critical consideration in pharmaceutical development, manufacturing, and stability, J. Valid. Technol. 5 (2008) 15–23 http://www.ivtnetwork.com/sites/default/files/Polymorphism_01.pdf.
[10] N.L. Calvo, R.M. Maggio, T.S. Kaufman, Chemometrics-assisted solid-state characterization of pharmaceutically relevant materials. Polymorphic substances, J. Pharm. Biomed. Anal. 147 (2018) 518–537, https://doi.org/10.1016/J.JPBA.2017.06.018.
[11] J. Bauer, S. Spanton, R. Henry, J. Quick, W. Dziki, W. Porter, J. Morris, Ritonavir: an extraordinary example of conformational polymorphism, Pharm. Res. 18 (2001) 859–866.
[12] S. Kumar, G. Chawla, M.E. Sobhia, A.K. Bansal, Characterization of solid-state forms of mebendazole, Pharmazie 63 (2008) 136–143, https://doi.org/10.1691/ph.2008.7168.
[13] M. Himmelreich, B.J. Rawson, T.R. Watson, Polymorphic forms of mebendazole, Aust. J. Pharm. Sci. 6 (1977) 123–125.
[14] J. Costa, M. Freno, L. Guzmán, A. Igual, J. Oliva, P. Vidal, A. Pérez, M. Pujol, Polymorphic forms of mebendazole: analytical aspectcs and toxicity, Circ. Farm. 49 (1991) 415–424.
[15] P. Charoenlarp, J. Waikagul, C. Muennoo, S. Srinophakun, D. Kitayaporn, Efficacy of single-dose mebendazole, polymorphic forms A and C, in the treatment of Hookworm and Trichuris infections, Southeast Asian, J. Trop. Med. Public Heal. 24 (1993) 712–716.
[16] A.P. Ayala, H.W. Siesler, S.L. Cuffini, Polymorphism incidence in commercial tablets of mebendazole: a vibrational spectroscopy investigation, J. Raman Spectrosc. 39 (2008) 1150–1157, https://doi.org/10.1002/jrs.
[17] S.B. Honorato, S. Farfan, A. Viana, J.M. Filho, G.C. Camarão, F.V. Fechine, M.E.A. Moraes, M.O. Moraes, M. Ferro, V. Dabbene, S.L. Cuffini, A.P. Ayala, Polymorphism evaluation in generic tablets containing mebendazole by dissolution tests, J. Brazilian Chem. Soc. 23 (2012) 220–227.
[18] K. Kachrimanis, M. Rontogianni, S. Malamataris, Simultaneous quantitative analysis of mebendazole polymorphs A-C in powder mixtures by DRIFTS spectroscopy and ANN modeling, J. Pharm. Biomed. Anal. 51 (2010) 512–520, https://doi.org/10.1016/j.jpba.2009.09.001.
[19] V.H. Silva, J.L. Gonçalves, F.V.C. Vasconcelos, M.F. Pimentel, C.F. Pereira, Quantitative analysis of mebendazole polymorphs in pharmaceutical raw materials using near-infrared spectroscopy, J. Pharm. Biomed. Anal. 115 (2015) 587–593, https://doi.org/10.1016/j.jpba.2015.08.018.
[20] V.H. da Silva, J.J. da Silva, C.F. Pereira, Portable near-infrared instruments: application for quality control of polymorphs in pharmaceutical raw materials and calibration transfer, J. Pharm. Biomed. Anal. 134 (2017) 287–294, https://doi.org/10.1016/j.jpba.2016.11.036.
[21] C. Gendrin, Y. Roggo, C. Collet, Pharmaceutical applications of vibrational chemical imaging and chemometrics: a review, J. Pharm. Biomed. Anal. 48 (2008) 533–553, https://doi.org/10.1016/J.JPBA.2008.08.014.
[22] J.M. Prats-Montalbán, A. de Juan, A. Ferrer, Multivariate image analysis: a review with applications, Chemom. Intell. Lab. Syst. 107 (2011) 1–23, https://doi.org/10.1016/j.chemolab.2011.03.002.
[23] J.M. Amigo, H. Babamoradi, S. Elcoroaristizabal, Hyperspectral image analysis. A tutorial, Anal. Chim. Acta. 896 (2015) 34–51, https://doi.org/10.1016/j.aca.2015.09.030.
[24] P.-Y. Sacré, C. De Bleye, P.-F. Chavez, L. Netchacovitch, P. Hubert, E. Ziemons, Data processing of vibrational chemical imaging for pharmaceutical applications, J. Pharm. Biomed. Anal. 101 (2014) 123–140, https://doi.org/10.1016/J.JPBA.2014.04.012.
[25] J. Cruz, M. Bautista, J.M. Amigo, M. Blanco, Nir-chemical imaging study of acetylsalicylic acid in commercial tablets, Talanta. 80 (2009) 473–478, https://doi.org/10.1016/j.talanta.2009.07.008.
[26] L. de M. França, M.F. Pimentel, S. da S. Simões, S. Grangeiro, J.M. Prats-Montalbán, A. Ferrer, NIR hyperspectral imaging to evaluate degradation in captopril commercial tablets, Eur. J. Pharm. Biopharm. 104 (2016) 180–188, https://doi.org/10.1016/J.EJPB.2016.05.005.
[27] G.L. Alexandrino, J.M. Amigo, M.R. Khorasani, J. Rantanen, A.V. Friderichsen, R.J. Poppi, Unveiling multiple solid-state transitions in pharmaceutical solid dosage forms using multi-series hyperspectral imaging and different curve resolution approaches, Chemom. Intell. Lab. Syst. 161 (2017) 136–146, https://doi.org/10.1016/J.CHEMOLAB.2016.11.004.
[28] G.P. Sabin, M.C. Breitkreitz, A.M. de Souza, P. da Fonseca, L. Calefe, M. Moffa, R.J. Poppi, Analysis of pharmaceutical pellets: an approach using near-infrared chemical imaging, Anal. Chim. Acta. 706 (2011) 113–119, https://doi.org/10.1016/J.ACA.2011.08.029.
[29] N.L. Calvo, R.M. Maggio, T.S. Kaufman, Characterization of pharmaceutically relevant materials at the solid state employing chemometrics methods, J. Pharm. Biomed. Anal. 147 (2018) 538–564, https://doi.org/10.1016/J.JPBA.2017.06.017.
[30] J. Jaumot, A. de Juan, R. Tauler, MCR-ALS GUI 2.0: new features and applications, Chemom. Intell. Lab. Syst. 140 (2015) 1–12, https://doi.org/10.1016/j.chemolab.2014.10.003.
[31] S. Piqueras, J. Burger, R. Tauler, A. de Juan, Relevant aspects of quantification and sample heterogeneity in hyperspectral image resolution, Chemom. Intell. Lab. Syst. 117 (2012) 169–182, https://doi.org/10.1016/J.CHEMOLAB.2011.12.004.
[32] T. Naes, T. Iraksson, T. Fearn, T. Davies, A User-Friendly Guide to Multivariate Calibration and Classification, NIR Publications, Chichester-UK, 2004.
[33] R.G. Brereton, Introduction to multivariate calibration in analytical chemistry, Analyst 125 (2000) 2125–2154, https://doi.org/10.1039/b003805i.
[34] W.F. de C. Rocha, G.P. Sabin, P.H. Março, R.J. Poppi, Quantitative analysis of

piroxicam polymorphs pharmaceutical mixtures by hyperspectral imaging and chemometrics, Chemom. Intell. Lab. Syst. 106 (2011) 198–204, https://doi.org/10.1016/j.chemolab.2010.04.015.
[35] A. Siddiqui, Z. Rahman, V.A. Sayeed, M.A. Khan, Chemometric evaluation of near infrared, fourier transform infrared, and raman spectroscopic models for the prediction of nimodipine polymorphs, J. Pharm. Sci. 102 (2013) 4024–4035, https://doi.org/10.1002/jps.23712.
[36] S.A. Schönbichler, L.K.H. Bittner, A.K.H. Weiss, U.J. Griesser, J.D. Pallua, C.W. Huck, Comparison of NIR chemical imaging with conventional NIR, Raman and ATR-IR spectroscopy for quantification of furosemide crystal polymorphs in ternary powder mixtures, Eur. J. Pharm. Biopharm. 84 (2013) 616–625, https://doi.org/10.1016/j.ejpb.2013.01.006.
[37] A.M. Brondi, L.A. Terra, G.P. Sabin, J.S. Garcia, R.J. Poppi, M.G. Trevisan, Mapping the polymorphic forms of fexofenadine in pharmaceutical tablets using near infrared chemical imaging, J. Near Infrared Spectrosc. 22 (2014) 211–220, https://doi.org/10.1255/jnirs.1094.
[38] R. Huisman, H.V. Van Kamp, J.W. Weyland, D.A. Doornbos, G.K. Bolhuis, C.F. Lerk, Development and optimization of pharmaceutical formulations using a simplex lattice design, Pharm. Weekbl. Sci. Ed. 6 (1984) 185–194, https://doi.org/10.1007/BF01999941.
[39] M. Vidal, J.M. Amigo, Pre-processing of hyperspectral images. Essential steps before image analysis, Chemom. Intell. Lab. Syst. 117 (2012) 138–148, https://doi.org/10.1016/j.chemolab.2012.05.009.
[40] Å. Rinnan, F. van den Berg, S.B. Engelsen, Review of the most common pre-processing techniques for near-infrared spectra, TrAC – Trends Anal. Chem. 28 (2009) 1201–1222, https://doi.org/10.1016/j.trac.2009.07.007.
[41] K.S. Booksh, B.R. Kowalski, Theory of analytical chemistry, Anal. Chem. 66 (1994) 782–791, https://doi.org/10.1021/ac00087a001.
[42] N. Mobaraki, J.M. Amigo, HYPER-Tools, A graphical user-friendly interface for hyperspectral image analysis, Chemom. Intell. Lab. Syst. 172 (2018) 174–187, https://doi.org/10.1016/J.CHEMOLAB.2017.11.003.
[43] E.M. Paiva, V.H. da Silva, R.J. Poppi, C.F. Pereira, J.J.R. Rohwedder, Comparison of macro and micro Raman measurement for reliable quantitative analysis of pharmaceutical polymorphs, J. Pharm. Biomed. Anal. 157 (2018) 107–115, https://doi.org/10.1016/j.jpba.2018.05.010.
[44] E. Swanepoel, W. Liebenberg, M.M. de Villiers, Quality evaluation of generic drugs by dissolution test: changing the USP dissolution medium to distinguish between active and non-active mebendazole polymorphs, Eur. J. Pharm. Biopharm. 55 (2003) 345–349, https://doi.org/10.1016/S0939-6411(03)00004-3.
[45] G.P. Sabin, W.F. de Carvalho Rocha, R.J. Poppi, Study of the similarity between distribution maps of concentration in near-infrared spectroscopy chemical imaging obtained by different multivariate calibration approaches, Microchem. J. 99 (2011) 542–547, https://doi.org/10.1016/j.microc.2011.07.007.