



openaccess

Study of component distribution in pharmaceutical binary powder mixtures by near infrared chemical imaging

Manel Bautista,^{*a} Jordi Cruz^b and Marcelo Blanco^a

^aUnidad de Química Analítica, Departamento de Química, Facultad de Ciencias, Universidad Autónoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. E-mail: bautista.manel@googlemail.com

^bEscola Universitària Salesiana de Sarrià, Passeig Sant Joan Bosco, 74, 08017 Barcelona, Spain

Near infrared chemical imaging (NIR-CI) has recently emerged as an effective technique for extracting spatial information from pharmaceutical products in an expeditious, non-destructive and non-invasive manner. These features have turned it into a useful tool for controlling various steps in drug production processes. Imaging techniques provide a vast amount of both spatial and spectral information that can be acquired in a very short time. Such a huge amount of data requires the use of efficient and fast methods to extract the relevant information. Chemometric methods have proved especially useful for this purpose. In this study, we assessed the usefulness of the correlation coefficient (CC) between the spectra of samples, the pure spectra of the active pharmaceutical ingredient (API) and we assessed the excipients to check for correct ingredient distribution in pharmaceutical binary preparations blended in the laboratory. Visual information about pharmaceutical component distribution can be obtained by using the CC. The performance of this model construction methodology for binary samples was compared with other various common multivariate methods including partial least squares, multivariate curve resolution and classical least squares. Based on the results, correlation coefficients are a powerful tool for the rapid assessment of correct component distribution and for quantitative analysis of pharmaceutical binary formulations. For samples of three or more components it has been shown that if the objective is only to determine uniformity of blending, then the CC image map is very good for this, and easy and fast to compute.

Keywords: correlation coefficient, NIR chemical imaging, homogeneity, concentration maps, pharmaceutical samples

Introduction

In recent years, chemical imaging techniques have emerged as effective methods for surface analysis which not only circumvent some shortcomings of conventional spectroscopic techniques, but also afford more comprehensive characterisation of the measured surface. These imaging techniques are especially useful in this context because of their ability to extract vast spectral and spatial information (thousands of spectra) from a single sample in a short time. The robust and reliable combination of chemical (molecular spectroscopy) and physical (digital imaging) techniques have aroused an increasing interest in a number of scientific areas such

as medical science, food science, and pharmaceutical and cultural heritage research.¹⁻⁶

The use of near infrared chemical imaging (NIR-CI) as a working methodology in pharmaceutical analysis is on the rise, with applications including tablet uniformity assessment,⁷ coating characterisation and particle/domain size determination,⁸ spectral matching,^{9,10} and the homogeneity of pharmaceutical powder blends characterisation.¹¹⁻¹⁴ All these features make NIR-CI techniques a useful tool for reducing the time and resources required for manufacturing while still improving quality control.

Such expansion has resulted mainly from the above-mentioned ability of NIR-CI to measure whole surfaces of samples, which is not the case with conventional NIR spectroscopy. NIR-CI information consists of a three-dimensional data cube where two dimensions are spatial directions (pixels) and a third dimension contains spectral information for each pixel (spectra). The body of NIR-CI data constitutes a hyperspectral cube consisting of thousands of spectra from which useful information can be extracted by using an appropriate mathematical tool.

One of the most critical steps in any pharmaceutical tableting process is the blending of the pure constituents to obtain a homogeneous final blend. Ensuring correct blending in terms component distribution and the absence of segregation (i.e. substantial differences in size between particles) is essential to comply with the strict regulations required of the pharmaceutical industry. The use of NIR-CI in combination with suitable data analysis methods has proved useful for this purpose. However, assessing powder blending uniformity requires using a fast, reliable methodology in order to obtain a picture of the distribution of the active pharmaceutical ingredient (API) and its excipients at any time during the production process, as well as quantifying the pharmaceutical components on the measured surface of the final blend.

Multivariate analysis has been used in combination with NIR-CI to overcome the serious overlap problems of NIR radiation ever since the latter technique emerged. Thus, principal component analysis (PCA) and cluster analysis have been used to obtain a preliminary view of the target surface or to segment previously obtained images into their constituent parts.⁷ Also, multivariate calibration models based on partial least squares (PLS) and classical least squares (CLS) methodologies^{15,16} and resolution models based on multivariate curve resolution (MCR)¹⁷ have been used to quantify components in various pharmaceutical preparations.

In this study, the usefulness of correlation coefficients between the spectrum for a pure analyte and its NIR chemical image have been tested for rapidly determining the distribution of major pharmaceutical components to be blended into binary mixtures; this was assessed by using laboratory-made binary mixtures of acetylsalicylic acid (ASA) and sucrose (SUC). Using the proposed methodology facilitated construction of calibration models to quantify the ASA and SUC contents in the measured surface at any time during the blending operation.

Correlation coefficients provide numerical information about the similarity between spectra at each pixel. Also, correlation maps provide visual information about the distribution of each analyte in a sample; therefore, increasing the proportion of a given analyte in the sample raises its correlation coefficient for the analyte at each pixel. Correlation coefficients have so far been successfully used to determine analytes from NIR hyperspectral images¹⁸ and, in combination with conventional NIR spectroscopy, to determine the API content in different samples from the same production process.¹⁹

Using correlation coefficients allows the development of simple calibration models; this is especially so with pharmaceutical samples where the pure spectrum for each analyte is available. As noted earlier, the pharmaceutical industry has a growing need for simple methods for component distribution analysis. With binary samples, meeting such a need entails the use of an intuitive parameter (e.g. a correlation coefficient) to obtain quantitative information in a fast, simple manner. In this study, we compared the effectiveness of using correlation coefficients in combination with univariate methodology to multivariate methodologies including CLS, PLS and MCR.

This paper demonstrates the advantages of using correlation coefficients in combination with NIR-CI over multivariate methodologies described in the literature^{7-9,11-14} for the assessment of component distribution uniformity in pharmaceutical binary preparations. In fact, correlation coefficients provide a simpler, faster methodology for the qualitative and quantitative analysis of major components in pharmaceutical samples. In addition, a discussion of the potential problems with samples that have more than two components is also included.

Experimental

Reagents and instruments

Binary samples preparation

A total of 15 binary samples containing approximately 10 g of ASA and SUC, and spanning the ASA content range 25–75% (w/w) (Table 1) were prepared. SUC was used in the form of granular powder of ca. 20 μm particle size and ASA was presented as needle-shaped particles 100–250 μm long. Both compounds were used as received and mixed in a Turbula mechanical shaker (Willy A. Bachofen, Basel, Switzerland) for 20 min, which was deemed long enough to ensure uniform distribution. Ten samples, from the above 15 samples spanning the concentration range were manually milled in a mortar into a fine powder and used to obtain the calibration line by the regression of the CC vs the concentration of the ASA in each sample. The other five samples were used to predict the ASA concentrations by using previously constructed regression curves.

Instruments

All spectra were obtained on a Think Spectrally Roda-25 focal plane imaging NIR spectrophotometer (Think Spectrally, Valencia, Spain) equipped with a mercury cadmium telluride (MCT) detector of 320 \times 256 pixels spanning the wavelength range of 1200–2400 nm with a spectral resolution of 7 nm. The total image-acquisition time was 120 s.

Under the selected operating conditions, a spatial resolution of 100 \times 100 μm per pixel image was collected from each sample; therefore, the overall image spanned an area of 32.0 \times 25.6 mm. Samples were irradiated with four halogen lamps placed at 45° angles around them.

Table 1. Standard deviation (*SD*) and relative standard deviation (*RSD*) of correlation coefficient of samples spectra respect to acetyl salicylic acid between pixels.

Sample	ASA concentration (%)	Non-milled samples		Milled samples	
		<i>SD</i> (%)	<i>RSD</i> (%)	<i>SD</i> (%)	<i>RSD</i> (%)
N01	24.8	0.056	12.07	0.024	4.43
N02	30.1	0.028	5.02	0.028	4.63
N03	34.7	0.035	6.09	0.027	4.51
N04	40.3	0.039	6.77	0.024	3.71
N05	45.3	0.023	3.38	0.026	3.71
N06	49.9	0.037	5.16	0.029	3.85
N07	57.7	0.019	2.58	0.017	2.12
N08	65.5	0.027	3.31	0.014	1.64
N09	70.2	0.022	2.60	0.010	1.13
N10	75.2	0.014	1.50	0.009	0.96

Software

Images were acquired by using the software bundled with the TS hyperspectral camera (Think Spectrally, Valencia, Spain) calibrated via a graphical user interface (GUI) for MatLab, TS_Gui, also supplied with the camera.

Spectral treatments and CLS were applied and correlation coefficient maps were constructed by using customised functions developed in MATLAB (The MathWorks, Natick, MA, USA). These routines can be freely downloaded from the web.²⁰ MCR-ALS analysis was performed by using software described elsewhere.²¹ PLS models were constructed with PLS-Toolbox (Eigenvector Research, Wenatchee, WA, USA).

Recording of images

The recording of images was preceded by calibration of the instrument with six AP-0200 NIR standards of 99%, 80%, 40%, 20%, 10% and 0.2% reflectance from Foss (Silver Springs, MD, USA). These standards were used to construct a calibration model used to define the absorbance scale; the software TS_Gui makes a quadratic adjustment from Digital Numbers (DN) to absorbance values, fixing the 100% and 0% reflectance values.

The images were obtained with a 7 nm resolution over the wavelength range of 1200–2000 nm.

The images of the samples were obtained by placing 3 g of each powder sample in a cylindrical glass cell about 3.8 cm of diameter and pressing their surfaces with a metal disc to obtain a flat surface in order to acquire the NIR spectrum.

The final information area used was 150 × 150 pixels, equivalent to a useful surface of 15 × 15 mm. The pure component images were also measured in a similar manner.

Pre-treatment of hyperspectral images

The three-dimensional nature of the data cube obtained (Figure 1) required unfolding for adaptation to the two-dimensional matrix typically used by the most common data processing

algorithms.⁷ The resulting 2D matrix was subjected to various spectral pre-treatments for testing and evaluation using an appropriate spectral treatment and wavelength that maximised differences between spectra according to ASA and SUC composition, thereby providing a wide range of correlation coefficients in terms of component concentrations. Also, it resulted in improved predictions when using univariate calibration models to obtain concentration maps.

Data analysis

Assessment of component distribution via correlation coefficients

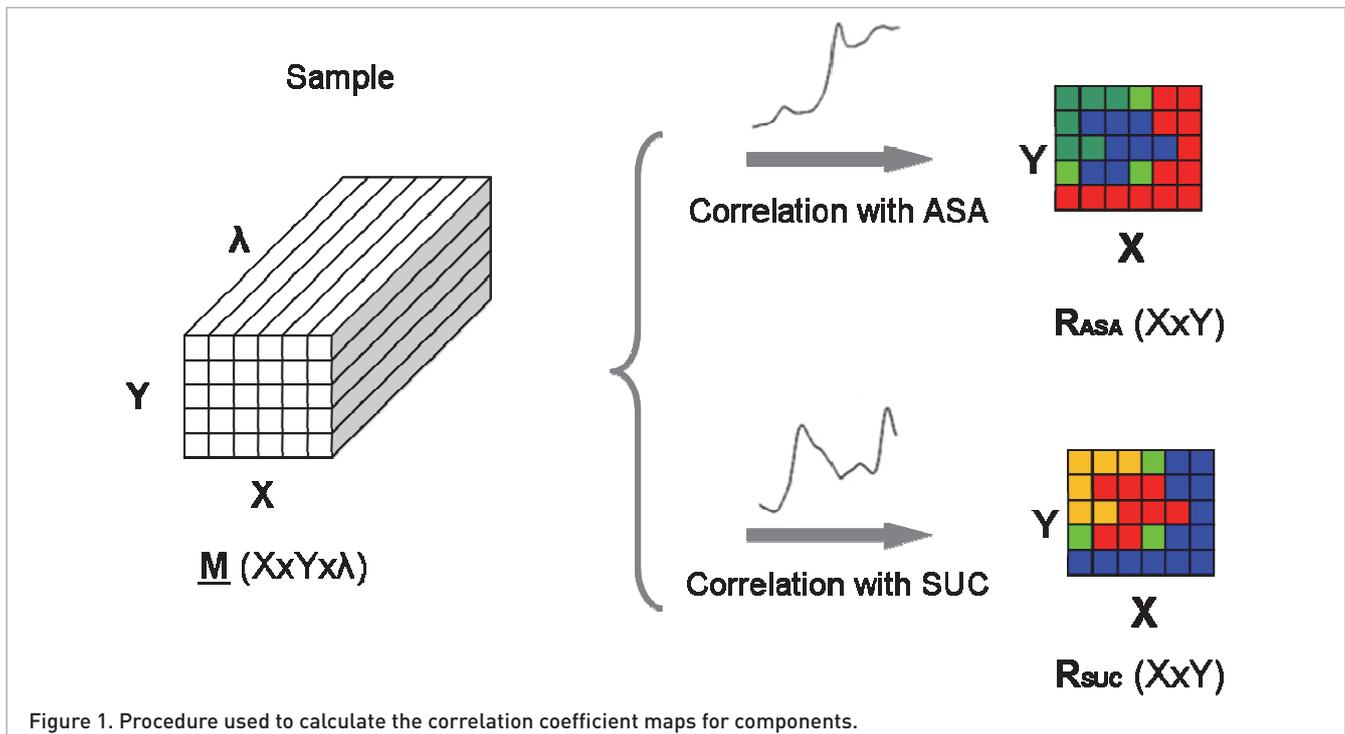
Correlation maps

The correlation coefficient measures the similarity between two spectra and is defined as the ratio between the covariance of spectra and the standard deviation for each spectrum. So correlations can be calculated to each pixel. Mathematically, the correlation coefficient between the spectrum for each pixel and for each pure component (ASA or SUC) is calculated as Pearson's correlation coefficient:¹⁸

$$\rho_{xy} = \frac{\sum_{\lambda=1}^{\Lambda} (a_{xy} - \bar{a})(b - \bar{b})}{(\lambda - 1)\sigma_{a_{xy}}\sigma_b} \quad (1)$$

where a_{xy} and b represent the spectrum obtained for each x - y pixel and the pure spectrum respectively as measured at λ wavelengths and $\sigma_{a_{xy}}$ and σ_b denote the standard deviation of a_{xy} and b , respectively.

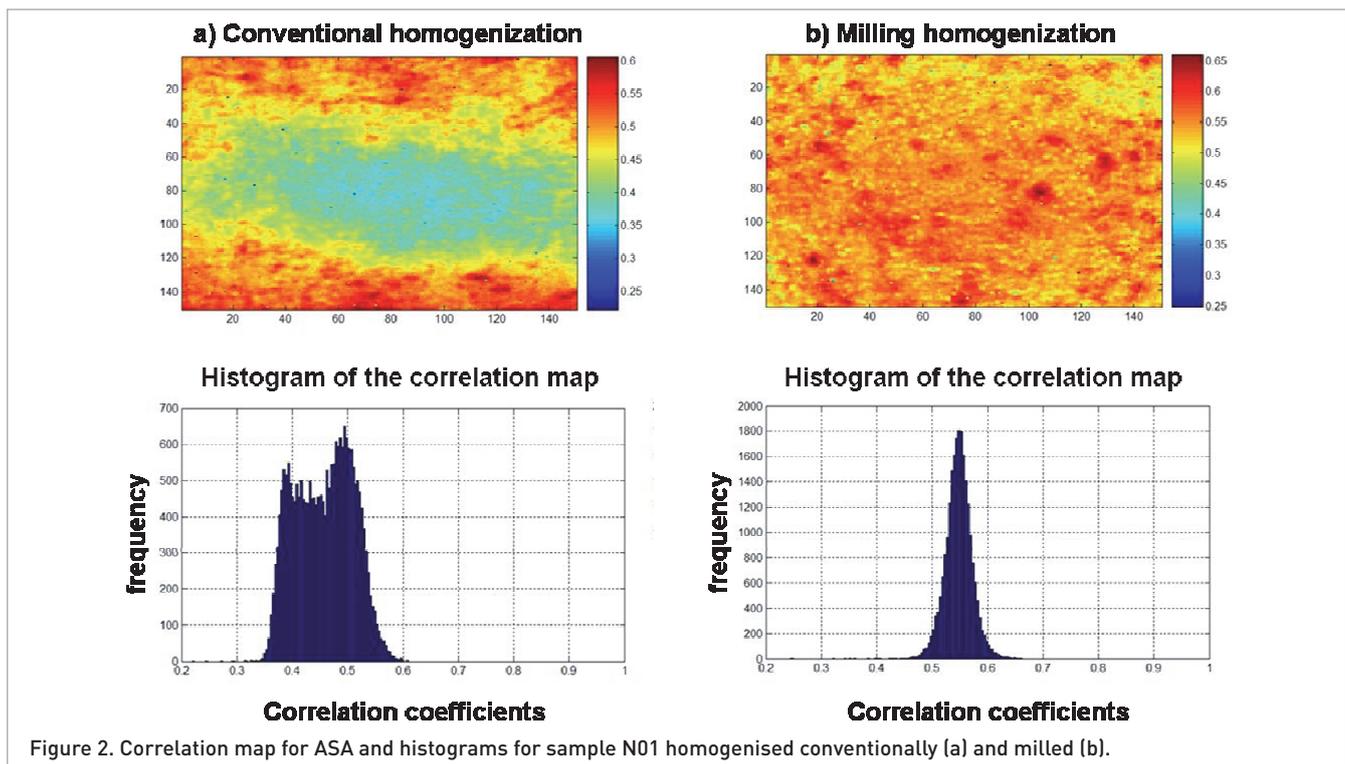
Calculating the correlation coefficient (CC) for each pixel of the image, correlation map $\mathbf{R}(x \times y)$ for each component can be obtained in order to identify the predominance areas for each analyte (Figure 2) and the plot of the CC histograms for each analyte allows us to examine the distribution of the correlation coefficients for both analytes in each sample.



The correlation maps obtained were used to calculate the corresponding mean values (\bar{X}), the standard deviation (SD) and the relative standard deviations (RSD) in order to derive useful information regarding the distribution of each component at the sample surface. A high RSD for a component was indicative of poor uniform distribution in the corresponding sample.

Quantitative analysis

Calibration models were constructed by univariate regression between the mean CC value obtained from the correlation maps for each component and their nominal concentration as described elsewhere.^{19,22} Jovanovic *et al.* illustrated the quadratic fit between concentration and coefficient correlation in the whole range (from 0% to 100%). Nevertheless, a linear



performance is observed when using short concentration intervals.

For this reason, both linear and quadratic models were developed in order to find the calibration with best prediction performance.

The models were applied pixel wise to prediction samples in order to construct a concentration map for each component in all samples and the map was used to obtain an average concentration value for each analyte.

Comparison with calibration and resolution methodologies

The quantitative results obtained were compared with those provided by other calibration models in order to assess the usefulness of correlation coefficients for quantification purposes. Therefore, quantitation methods, including PLS and CLS,¹⁵ and multivariate curve resolution–alternating least squares (MCR-ALS),^{6,17,22} were also used to compare the performance of these different algorithms.

The predictive ability of their models were assessed in terms of precision and accuracy via the root mean square error of prediction (*RMSEP*) and standard error of prediction (*SEP*) [Equations 2 and 3] with provision for bias (average value of residuals) in each model (Equation 4):

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (\hat{c}_i - c_i)^2}{n}} \quad (2)$$

$$SEP = \sqrt{\frac{\sum_{i=1}^n (\hat{c}_i - c_i - bias)^2}{n-1}} \quad (3)$$

$$bias = \frac{\sum_{i=1}^n (\hat{c}_i - c_i)}{n} \quad (4)$$

In these expressions, c_i is the reference concentration, \hat{c}_i the calculated concentration and n the number of samples studied.

Results and discussion

Correlation maps

The distribution of components in 10 binary samples containing variable proportions of ASA from 25% to 75% w/w and SUC was studied. The two components were mixed in a conventional solid shaker for 20 min, which was assumed to be long enough for adequate homogenisation. Different spectral ranges and various spectral treatments were used to identify the most informative correlation map for every analyte in each sample. This maximised the differences in correlation coefficients between samples, and also those between pixels in each sample, thereby increasing the sensitivity of the ensuing calibration model (Table 2). Using the wavelength range of 1200–2000 nm in combination with SNV and smoothing with a seven-point moving window provided the widest possible range of correlation coefficients.

For example, Figure 2(a) shows the correlation map for ASA in sample N01. As can be seen, the surface image exposed a clearly different distribution of the areas highly correlated with ASA (in red) and additionally where correlation was lower than expected (in green). Furthermore, the associated histogram exhibited duplicated variability in the correlation coefficients. These two results are complementary inasmuch as a correlation map shows the surface distribution of correlation coefficients, whereas a histogram is more closely associated with the values themselves.

Blending processes depend not only on mixing time but also on the physical properties of the compounds that allow or do not allow a better and faster mixing when morphology and particle size is quite similar. In our study, a visual inspection with a stereoscopic magnifying glass shows a very high aggregation of ASA particles and a reduced mixture with

Table 2. Calibration curves for ASA and SUC by using correlation coefficients.

ASA calibration curve				
Model	Ten milled samples in range 25–75%			
	Linear		Quadratic	
Quadratic term	—		1.8	
Slope	127.2		124.6	
Offset	-44.1		-43.1	
R^2	0.994		0.994	
SUC calibration curves				
Model	Ten milled samples in range 25–75%		Six milled samples in range 45–75%	
	Linear	Quadratic	Linear	Quadratic
Quadratic term	—	74.8	—	36.1
Slope	103.7	2.3	90.6	47
Offset	-22.4	9.9	-15.1	-2.5
R^2	0.989	0.998	0.998	0.999

sucrose particles. The needle form of ASA particles greatly limited the mixture with the rounded sucrose when mixed with a mechanical shaker, and only after an intense manual milling, which produces a significant reduction in the size of them, can a homogeneous distribution of the components be achieved. Only thorough milling of the two joint mixture components ensured homogeneous blending.

In conclusion, conventional homogenisation of the mixture failed to ensure homogeneity in the resulting sample as well as shaking for a prolonged time. Cohesive forces between the particles dictate their fluency and the ease with which they can be mixed homogeneously as a result. Only thorough joint milling of the two mixture components ensured homogeneous blending.

As can be seen from Figure 2(b), the studied zone exhibited a very soft gradation suggestive of improved distribution of the sample components. The associated histogram exhibited a homogeneous (Gaussian) distribution around a correlation coefficient value suggestive of the absence of clustering in the coefficients

Table 1 gives the *SD* and *RSD* of the correlation coefficient maps for the 10 samples in the calibration set. *SD* and *RSD* decreased with increasing API (ASA) content; the small differences in correlation coefficients between pixels are suggestive of increased homogeneity in the mixture.

The Cochran test, a test which allows *K* variances to be compared, between normal distributions but limited to cases where all classes are of equal size, especially useful for detecting whether one variance is much larger than the other. A Cochran test applied to samples exposed differences in *SD* between milled and unmilled samples. This result can be ascribed to the ease of homogenisation increasing with API content due to the effect of increasing cohesion forces between ASA particles hindering their movement and homogenisation of the mixture.^{23,24} Based on this information, simply milling the samples to a very small particle size (<25 µm) can be expected to facilitate mutual mixing of the two components.

Quantitative analysis

Ten finely powdered binary samples containing 25%–75% ASA were used to construct calibration curves for ASA and SUC by univariate regression of the average correlation coefficients for each sample and their respective nominal concentrations (Table 3). Both linear and quadratic models were tested and found to provide similar results for ASA. Therefore, the linear model was adopted in order to facilitate interpretation of the models. On the other hand, the quadratic model for SUC

Table 4. Comparison of the results obtained by correlation coefficients (CC) with classical least squares (CLS), multivariate curve resolution-alternating least squares (MCR-ALS) and partial least squares (PLS).

	CC	CLS	MCR-ALS	PLS
<i>RMSEP</i>	2.38	23.06	3.33	2.35
Bias	−0.90	23.01	−2.65	−0.04
SEP	3.18	51.49	6.33	2.62

performed better than its linear counterpart. For simplicity, however, the linear model was chosen for SUC as well. This model was applied over a restricted SUC content range (45%–75%), which provided acceptable values for quality-related parameters.

Five new mixtures of ASA and SUC containing 50%–70% API were prepared and used to assess the predictive ability of the calibration models. Preliminary tests revealed that the distribution of both components in these samples was similar to that in the calibration samples and that their degree of homogeneity was acceptable. A concentration map was constructed by plotting the concentration value obtained in order to show the concentration distribution for both analytes at the sample surface.

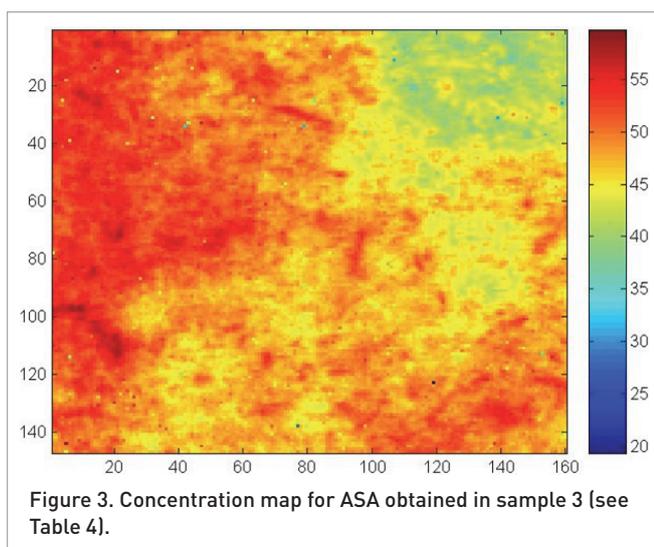
Table 4 gives the average ASA concentration obtained for each sample as calculated by using the calibration model with correlation coefficients. As can be seen, differences between the model-predicted and reference values were acceptable, which testifies to the quality of the proposed model. The largest difference was for sample 3 (Figure 3), which was exhibiting the highest *SD* (i.e. the highest heterogeneity).

Comparison between calibration models

Five freshly prepared samples were used in order to test the quality of the CC-built model over other quantitative algorithms for calculating the concentration commonly used in chemometric treatment for NIR chemical images. Recorded images of these samples were quantified using three calibration models: CLS, MCR-ALS and partial least regression (PLSR). The results are shown in Table 4. CLS directly calculates the concentration of the compounds by using the pure spectra of the compounds and all of the spectra of the image. As can be observed in Table 4, CLS does not perform a correct prediction with both concentration and distribution parameters.

Table 3. Comparison between nominal and predicted values of ASA concentration by using correlation coefficients. *SD* (%) and *RSD* (%) account for the standard deviation and relative standard deviation of the concentration, respectively.

Sample	<i>SD</i> (%)	<i>RSD</i> (%)	Reference (% ASA)	Prediction (% ASA)	Residual
1	0.0279	3.58	59.1	56.5	−2.6
2	0.0125	1.44	68.4	67.9	−0.5
3	0.0282	3.95	52.3	48.3	−4.0
4	0.00954	1.14	61.7	64.0	2.3
5	0.0073	0.83	70.3	70.6	0.3



MCR-ALS also uses a direct quantification but with the difference that it works iteratively; optimising matrices for concentration and spectra. Using pure spectra for initial estimations allows significantly better results than CLS to be obtained. These two methods do not require a calibration set (only the spectra of components) and their results can be obtained faster than PLS and CC.

A calibration data set is a prerequisite and is always available for PLS and CC. This is both an advantage and a disadvantage in generating the most accurate chemical images because it is time consuming to produce a robust calibration.

CC, PLS and MCR-ALS provided similar predictions and consequently any of these three methodologies can be used for exploratory analysis of quantitative and qualitative levels in binary samples within the pharmaceutical industry.

The use of CC for samples of three or more components

This study concentrated on samples of two components; to extend it for more complex samples, some simulations were carried out and these showed us that this method could be extended to complex samples of three or more components if the objective is only to test the uniformity of blending.

Assuming the ideal situation, which is perfect blending of the sample, then the CC image of component 1 and a mixture containing component 1 and one (component 2), or further other components (components 2, 3, 4), should be a very uniform image of a constant CC. It does not matter if it is a binary mixture or a mixture of three or more components, for example.

In this case the mixture of three components (components 2, 3 and 4) is a mixture, but since it is an ideal blend, the composition does not vary, and in fact it can be thought of as a "single" component (but which is in fact a mixture) and therefore the entire question returns to the CC of a binary mixture. And this is the same case as in pharmaceutical samples, which are composed of an active pharmaceutical ingredient and other components, and can also be considered as binary samples if they have a constant composition: API

on one side and other components on the other. In these conditions the other components can be treated as a single component.

For non-uniform blending, the composition of the samples is heterogeneous, the CC image for component 1 (or any other component) has to vary and this variation will spatially represent the non-uniformity of the material.

This non-uniformity can be caused in three ways:

- The uniformity of component 1 does in fact vary, which you can see in the CC map.
- The uniformity of component 1 is very good, but the uniformity of component 2 changes: component 2 is a pure component and the relative concentration to 1 changes.
- Component 2 is in fact a mixture of more than one ingredient, and the relative proportions change.

The problem is that if the blending is not uniform, this can be detected with this methodology, but it is impossible to say which of the three possible causes is the correct one.

Conclusions

Chemical imaging techniques can produce a very large amount of data and it is therefore essential to find simple tools that allow extraction of relevant information.

The use of correlation coefficients calculated from pure NIR spectra allow the construction of correlation maps which facilitate the expeditious observation of the distribution of analytes at the sample surface and therefore the rapid and accurate assessment of component distribution in mixtures.

Correlation maps, *SD* and *RSD* values, and histograms of the correlation maps provide useful information to assess mixture homogeneity. Moreover, correlation coefficients allow the easy development of quantitative models for predicting API surface contents as supplementary information to that obtained by visual inspection of the distribution of components in a powder mixture. Predictions and quality-related parameters are as good as those obtained with more complex multivariate calibration methodologies.

Correct application of spectral pre-treatments, appropriate selection of spectral bands and the resolution of the imaging system all influence the results obtained with the proposed methodology.

The need for pure spectra of the analytes necessitates that this technique cannot be used with natural products which do not have pure spectra. But in the pharmaceutical industry the pure spectrum of the compounds is easy to obtain, so this technique can be widely used. The CC itself can be used to perform an exploratory and initial analysis on the distribution, and even quantification. The main disadvantage of CC use is the degree of spectral similarity between components; high spectral differences between components improve the range of CC and the differences in concentration maps.

Despite these drawbacks, the ease with which CC can be calculated and processed by using univariate regression with direct interpolation to obtain the final result makes it an attractive choice for rapidly checking correct distribution

and extracting quantitative information about pharmaceutical blending processes among others.

Although only binary samples were studied, correlation coefficients can also be useful with mixtures in which one compound varies its concentration, and the others remain constant; since only two of the compounds vary their concentration, the concentration of the rest of the components remain constant.

If the objective is only to determine uniformity of blending, then the CC image map is very good for this, and easy and fast to compute, but it is not possible to know the reason why a mixture appears non-uniform.

Acknowledgements

The authors gratefully acknowledge funding of this work by Spain's Ministry of Science and Technology (Project CTQ2009-08312).

References

1. E.N. Lewis, L.H. Kidder, E. Lee and K.S. Haber, *Spectrochemical Analysis using Infrared Multichannel Detectors*. Blackwell Publishing, Oxford, UK (2005).
2. V.W. Smail, A.K. Fritz and D.L. Wetzel, "Chemical imaging of intact seeds with NIR focal plane array assists plant breeding", *Vibr. Spectrosc.* **42**, 215(2006). doi: [10.1016/j.vibspec.2006.02.004](https://doi.org/10.1016/j.vibspec.2006.02.004)
3. M. Politi and C.D. Tran, "Investigation of inhomogeneity in chemical compositions and kinetics of sol-gel by near-infrared multispectral imaging technique", *J. Non-Cryst. Solids* **304**, 64 (2002). doi: [10.1016/S0022-3093\(02\)01005-0](https://doi.org/10.1016/S0022-3093(02)01005-0)
4. E. Lee, L.H. Kidder, V.F. Kalasinsky, J.W. Schoppelrei and E.N. Lewis, "Forensic visualization of foreign matter in human tissue by near-infrared spectral imaging: methodology and data mining strategies", *Cytometry* **69A**, 888 (2006). doi: [10.1002/cyto.a.20277](https://doi.org/10.1002/cyto.a.20277)
5. M.E. Klein, B.J. Aalderink, R. Padoan, G. de Bruin and T.A. Steemers, "Quantitative hyperspectral reflectance imaging", *Sensors* **8**, 5576 (2008). doi: [10.3390/s8095576](https://doi.org/10.3390/s8095576)
6. A. de Juan, R. Tauler, R. Dyson, C. Marcolli, M. Rault and M. Maeder, "Spectroscopic imaging and chemometrics: a powerful combination for global and local sample analysis", *Trac-Trend. Anal. Chem.* **23**, 70 (2004). doi: [10.1016/S0165-9936\(04\)00101-3](https://doi.org/10.1016/S0165-9936(04)00101-3)
7. C. Gendrin, Y. Roggo and C. Collet, "Content uniformity of pharmaceutical solid dosage forms by near infrared hyperspectral imaging: a feasibility study", *Talanta* **73**, 733 (2007). doi: [10.1016/j.talanta.2007.04.054](https://doi.org/10.1016/j.talanta.2007.04.054)
8. A.A. Gowen, C.P. O'Donnell, P.J. Cullen and S.E.J. Bell, "Recent applications of chemical imaging to pharmaceutical process monitoring and quality control", *Eur. J. Pharm. Biopharm.* **69**, 10 (2008). doi: [10.1016/j.ejpb.2007.10.013](https://doi.org/10.1016/j.ejpb.2007.10.013)
9. P.A. Hailey, P. Doherty, P. Tapsell, T. Oliver and P.K. Aldridge, "Automated system for the on-line monitoring of powder blending processes using near-infrared spectroscopy Part I. System development and control", *J. Pharm. Biomed. Anal.* **14**, 551 (1996). doi: [10.1016/0731-7085\(95\)01674-0](https://doi.org/10.1016/0731-7085(95)01674-0)
10. S.S. Sekulic, J. Wakeman, P. Doherty and P.A. Hailey, "Automated system for the on-line monitoring of powder blending processes using near-infrared spectroscopy Part II. Qualitative approaches to blend evaluation", *J. Pharm. Biomed. Anal.* **17**, 1285 (1998). doi: [10.1016/S0731-7085\(98\)00025-9](https://doi.org/10.1016/S0731-7085(98)00025-9)
11. (a) J.M. Amigo, J. Cruz, M. Bautista, S. Maspocho, J. Coello and M. Blanco, "Study of pharmaceutical samples by NIR chemical image and multivariate analysis", *Trac-Trend. Anal. Chem.* **27**, 696 (2008). doi: [10.1016/j.trac.2008.05.010](https://doi.org/10.1016/j.trac.2008.05.010); (b) E.N. Lewis, J. Schoppelrei and E. Lee, "Molecular spectroscopy workbench near-infrared chemical imaging and the PAT initiative", *Spectroscopy* **19**, 26 (2004).
12. H. Ma and C.A. Anderson, "Characterization of pharmaceutical powder blends by NIR chemical imaging", *J. Pharm. Sci.* **97**, 3305 (2008). doi: [10.1002/jps.21230](https://doi.org/10.1002/jps.21230)
13. T. Puchert, D. Lochmann, J.C. Menezes and G. Reich, "A multivariate approach for the statistical evaluation of near-infrared chemical images using Symmetry Parameter Image Analysis (SPIA)", *Eur. J. Pharm. Biopharm.* **78**, 117 (2011). doi: [10.1016/j.ejpb.2010.12.035](https://doi.org/10.1016/j.ejpb.2010.12.035)
14. R.C. Lyon, D.S. Lester, E.N. Lewis, E. Lee, L.X. Yu, E.H. Jefferson and A.S. Hussain, "Near-infrared spectral imaging for quality assurance of pharmaceutical products: analysis of tablets to assess powder blend homogeneity", *AAPS PharmSciTech* **3(3)**, 1 (2002). doi: [10.1007/BF02830615](https://doi.org/10.1007/BF02830615)
15. C. Ravn, R. Bro and E. Skibsted, "Near-infrared chemical imaging (NIR-CI) on pharmaceutical solid dosage forms—Comparing common calibration approaches", *J. Pharm. Biomed. Anal.* **48**, 554 (2008). doi: [10.1016/j.jpba.2008.07.019](https://doi.org/10.1016/j.jpba.2008.07.019)
16. J.M. Amigo, C. Ravn, N.B. Gallagher and R. Bro, "A comparison of a common approach to partial least squares-discriminant analysis and classical least squares in hyperspectral imaging", *Int. J. Pharm.* **373**, 179 (2009). doi: [10.1016/j.ijpharm.2009.02.014](https://doi.org/10.1016/j.ijpharm.2009.02.014)
17. J. Cruz, M. Bautista, J.M. Amigo and M. Blanco, "NIR-chemical imaging study acetylsalicylic acid in commercial tablets", *Talanta* **80**, 473–478 (2009). doi: [10.1016/j.talanta.2009.07.008](https://doi.org/10.1016/j.talanta.2009.07.008)
18. A.G. Asuero, A. Sayago and A.G. Gonzalez, "The correlation coefficient: an overview", *Crit. Rev. Anal. Chem.* **36**, 41 (2006). doi: [10.1080/10408340500526766](https://doi.org/10.1080/10408340500526766)
19. N. Jovanovic, A. Gerich, A. Bouchard and W. Jiskoot, "Near-infrared imaging for studying homogeneity of pro-

tein-sugar mixtures”, *Pharmaceut. Res.* **23**, 2002 (2006).
doi: [10.1007/s11095-006-9037-y](https://doi.org/10.1007/s11095-006-9037-y)

20. <http://www.models.life.ku.dk/~jose/home.html>

21. J. Jaumot, R. Gargallo, A. de Juan and R. Tauler, “A graphical user-friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB”, *Chemometr. Intell. Lab. Syst.* **76**, 101 (2005). doi: [10.1016/j.chemolab.2004.12.007](https://doi.org/10.1016/j.chemolab.2004.12.007)

22. M. Blanco, J. Cruz and M. Bautista, “Development of a univariate calibration model for pharmaceutical analysis based on NIR spectra”, *Anal. Bioanal. Chem.* **392**, 1367 (2008). doi: [10.1007/s00216-008-2426-9](https://doi.org/10.1007/s00216-008-2426-9)

23. L.J. Bellamy, A. Nordon and D. Littlejohn, “Effects of particle size and cohesive properties on mixing studied by non-contact NIR”, *Int. J. Pharm.* **361**, 87 (2008). doi: [10.1016/j.ijpharm.2008.05.030](https://doi.org/10.1016/j.ijpharm.2008.05.030)

24. L.J. Bellamy, A. Nordon and D. Littlejohn, “Real-time monitoring of powder mixing in a convective blender using non-invasive reflectance NIR spectrometry”, *Analyst* **133**, 58 (2008). doi: [10.1039/b713919e](https://doi.org/10.1039/b713919e)