Journal of Pharmaceutical and Biomedical Analysis 48 (2008) 772–779

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

Raman spectroscopy as a process analytical technology (PAT) tool for the *in-line* monitoring and understanding of a powder blending process

T.R.M. De Beer[a,*], C. Bodson[b], B. Dejaegher[c], B. Walczak[d], P. Vercruysse[a], A. Burggraeve[a], A. Lemos[a], L. Delattre[b], Y. Vander Heyden[c], J.P. Remon[e], C. Vervaet[e], W.R.G. Baeyens[a]

[a] *Laboratory of Pharmaceutical Chemistry and Drug Analysis, Department of Pharmaceutical Analysis, Ghent University, Harelbekestraat 72, B-9000 Ghent, Belgium*
[b] *Laboratory of Pharmaceutical Technology, Department of Pharmacy, University of Liège, 1 l'Avenue de l'Hôpital, 4000 Liège, Belgium*
[c] *Department of Analytical Chemistry and Pharmaceutical Technology, Pharmaceutical Institute, Vrije Universiteit Brussel–VUB, Laarbeeklaan 103, B-1090 Brussels, Belgium*
[d] *Department of Chemometrics, The University of Silesia, 9 Szkolna Street, 40-006 Katowice, Poland*
[e] *Laboratory of Pharmaceutical Technology, Department of Pharmaceutics, Ghent University, Harelbekestraat 72, B-9000 Ghent, Belgium*

ARTICLE INFO

Article history:
Received 22 April 2008
Received in revised form 21 July 2008
Accepted 25 July 2008
Available online 7 August 2008

Keywords:
Raman spectroscopy
NIR spectroscopy
PAT
Blending
Experimental design
In-line monitoring
Process understanding
MD
Orthogonal distance
SIMCA

ABSTRACT

The aim of this study is to propose a strategy to implement a PAT system in the blending step of pharmaceutical production processes. It was examined whether Raman spectroscopy can be used as PAT tool for the *in-line* and *real-time* endpoint monitoring and understanding of a powder blending process.

A screening design was used to identify and understand the significant effects of two process variables (blending speed and loading of the blender) and of a formulation variable (concentration of active pharmaceutical ingredient (API): diltiazem hydrochloride) upon the required blending time (response variable). Interactions between the variables were investigated as well. A Soft Independent Modelling of Class Analogy (SIMCA) model was developed to determine the homogeneity of the blends *in-line* and *real-time* using Raman spectroscopy in combination with a fiber optical immersion probe. One blending experiment was monitored using Raman and NIR spectroscopy simultaneously. This was done to verify whether two independent monitoring tools can confirm each other's endpoint conclusions.

The analysis of the experimental design results showed that the measured endpoints were excessively rounded due to the large measurement intervals relative to the first blending times. This resulted in effects and critical effects which cannot be interpreted properly. To be able to study the effects properly, the ratio between the blending times and the measurement intervals should be sufficiently high.

In this study, it anyway was demonstrated that Raman spectroscopy is a suitable PAT tool for the endpoint control of a powder blending process. Raman spectroscopy not only allowed *in-line* and *real-time* monitoring of the blend homogeneity, but also helped to understand the process better in combination with experimental design. Furthermore, the correctness of the Raman endpoint conclusions was demonstrated for one process by using a second independent endpoint monitoring tool (NIR spectroscopy). Hence, the use of two independent techniques for the control of one response variable not only means a mutual confirmation of both methods, but also provides a higher certainty in the determined endpoint.

1. Introduction

The Food and Drug Administration's (FDA) Process Analytical Technology (PAT) initiative (http://www.fda.gov/Cder/OPS/PAT.html) forms the basis of the pharmaceutical Good Manufacturing Practice (GMP) rules for the 21st century [1–4]. By means of scientific, risk-based PAT frameworks, it is aimed to design and develop continuously controlled (by timely *in-line*, *on-line* or *at-line* measurements of the critical intermediate steps and endpoints during the process), well understood and efficient processes that will consistently ensure a predefined quality at the end of the manufacturing process. Hence, PAT can be considered as an unbreakable connection between the scientific domains of analytical chemistry and pharmaceutical technology. To fulfil the PAT objectives in a process, it is necessary to use an appropriate combination of PAT tools, such as chemometric tools, process analyzers, endpoint monitoring tools and knowledge management tools.

Homogeneity of a powder blend is essential to guarantee the correct amounts of APIs and additives in every dosage unit at the end of several pharmaceutical production processes (production of tablets, capsules, ...). The importance of *in-line* and *real-time*

* Corresponding author. Tel.: +32 9 2648099; fax: +32 9 2648196.
E-mail address: Thomas.DeBeer@UGent.be (T.R.M. De Beer).

0731-7085/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2008.07.023

endpoint monitoring methods during blending processes with respect to the classical sampling method using a thief probe (*off-line* after-process analysis), is well-documented in numerous publications [5–11]. Till now, especially the ability of spectroscopic techniques (mostly NIR and, to a lesser extent, Raman spectroscopy) for *in-line* and *real-time* powder blending endpoint determination is described [12,13]. However, the implementation of a PAT system in a mixing process is not only involved with endpoint monitoring, but also with process understanding and optimizing process efficiency [14].

This study proposes a strategy to implement a PAT system in a pharmaceutical powder blending process. In a first instance, Raman spectroscopy in combination with chemometric tools are examined as PAT tools for the *in-line* and *real-time* endpoint determination of several different powder blending processes. The processes were determined according to an experimental design (see further). The powder mixture consisted of diltiazem hydrochloride (API), lactose DCL 21, Avicel PH 102 and silicium dioxide (additives).

We described earlier the ability and importance of quantitative *in-line* monitoring during the homogenization process of an aqueous pharmaceutical suspension [14]. Berntsson et al., El-Hagrasy et al., Li et al. described several methods for the quantitative process monitoring of blending processes using NIR [15–17]. However, quantitative *in-line* monitoring was not aimed at in this study, as a blending process is a *closed* process. Once the desired and controlled amounts of powders are added to the blender, a homogeneous powder blend will automatically consist of the correct component proportions. Furthermore, several Raman and NIR spectroscopic methods for the fast (seconds) and non-destructive quantitative analysis of solids (tablets, capsules, ...) are described [18,19]. These methods can easily be implemented at the end of a production process for immediate analysis, allowing *real-time* release [20–24].

In a second part of this study, the influence of two process variables (mixing speed and loading of the blender) and one formulation variable (concentration of API) upon the time required to obtain a homogeneous blend (response variable) was examined using experimental design (2-level full factorial design). This allowed identifying significant effects of factors and factor interactions upon the response variable blending time (process understanding). The results from the different design experiments in which the blending endpoints (blending times) were determined (part 1 of this study), were used to estimate the effects. Within a PAT framework, a process endpoint is not a fixed time (replicate processes do not result in unique endpoints); rather it is the achievement of the desired material attribute. However, this does not mean that blending time is not to be considered. A range of acceptable process times is likely to be achieved during the manufacturing phase and should be evaluated. Considerations for addressing significant deviations from acceptable process times should be developed [25]. Once the effects of different process and formulation variables upon the endpoint are known, it is possible to select their optimum combination resulting in increased process efficiency. Furthermore, it is possible to predict how the process endpoint will be influenced by varying a certain factor. The use of experimental design for process understanding of blending or mixing processes is scarcely described in literature [14,26].

Finally, one blending process was monitored using simultaneously Raman and NIR spectroscopy to check if the two techniques proposed the same process endpoints. Similar process endpoints determined by two independent monitoring techniques would significantly increase the credibility and certainty of the process analysis conclusions.

2. Materials and methods

2.1. Materials

Diltiazem hydrochloride (API) was purchased from Roig Farma (Barcelona, Spain). Avicel PH 102, lactose DCL 21 and silicium dioxide (additives) were obtained from FMC Europe (Little Island, Cork, Ireland), De Melkindustrie (Veghel, The Netherlands) and Alpha Pharma (Nazareth, Belgium), respectively.

2.2. Process description

The blending experiments were performed in a Gral™ 10 high shear mixing system from GEA-Collette (Wommelgem, Belgium). After weighing the correct amounts, all powders were transferred into the blender, each time in the same order (Avicel PH 102, lactose DCL 21, diltiazem hydrochloride and silicium dioxide, respectively), followed by the start of the blending process. For the *in-line* Raman spectroscopic process monitoring, the jacketed bowl (10 L) was perforated to allow the introduction of a Raman probe (Fig. 1).

2.3. Spectroscopic conditions

2.3.1. Raman experiments

A RamanRxn1 spectrometer (Kaiser Optical Systems, Ann Arbor, USA), equipped with an air-cooled CCD detector (back-illuminated deep depletion design) was used in combination with a fiber optic immersion probe (length: 5 m) to monitor the homogenization process *in-line* and non-invasively (Fig. 1). The laser wavelength during the experiments was the 785 nm line from a 785 nm Invictus NIR diode laser. All spectra were recorded at a resolution of 4 cm^{-1} using a laser power of 400 mW. Data collection and data transfer were automated using the HoloGRAMS™ data collection software package, the HoloREACT™ reaction analysis and profiling software package, the Matlab® software package (version 6.5), the Grams/AI—PLSplusIQ software package (version 7.02) and Excel®. Five-second exposures were used for the *in-line* monitoring of the mixing process. Spectra were collected every 10 s. The *in-line* spectral data acquisition using the fiber optic probe was operated in a non-stop mode, hence spectral information was continuously monitored during the blending process. The Raman probe was always positioned in exactly the same way through a tailored-made hole in the blender, the end of the probe (sapphire window) being flushed with the inner wall of the mixing bowl. A constant flow of water (15 °C) through the jacket of the mixing bowl assured a constant product temperature during all mixing experiments since high shear mixing is known to increase product temperature which would affect the intensity of the Raman bands significantly [27].

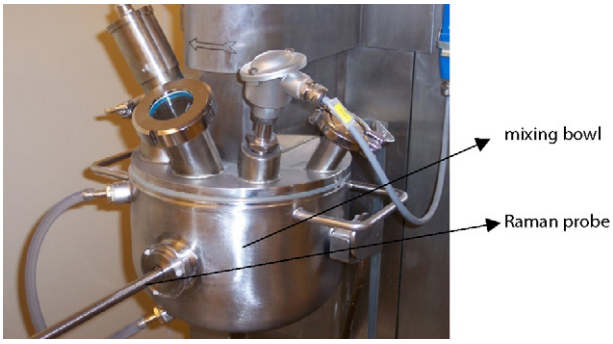

Fig. 1. Experimental setup.

2.3.2. Raman-NIR comparison experiment

One blending process was monitored using Raman and NIR spectroscopy simultaneously. Therefore, an UltimaGral™ 75 high shear mixing system from GEA-Collette (Wommelgem, Belgium) was used. The Raman immersion probe was placed in a hole on the side of the blender, while the NIR probe was positioned on top of the blender. The blender was loaded with 25 kg powder (containing Avicel PH 102, lactose DCL, 21 and silicium dioxide; without diltiazem·HCl) and a mixing speed of 90 rpm was used. Monitoring the same response variable (mixing time) by two independent process analyzers at different places in the blender clearly increases the certainty of the endpoint conclusions. NIR spectra were obtained using a Multi-Purpose Analyzer FT-NIR spectrometer (Bruker Optics, Belgium) in combination with a fiber optic immersion reflectance probe (Series 400 Diffuse Reflectance Probe from Precision Sensing Devices, Massachusetts, USA). The NIR spectrometer was equipped with an InGaAs detector and a tungsten light source. The NIR spectra were collected every 20 s with the Opus software 5.0 (Bruker Optics, Belgium). Each spectrum was the average of 16 scans and the resolution was 8 cm^{-1} over the range from 12,500 to 4000 cm^{-1}. The Raman measurement conditions were identical to those specified higher (cfr. 2.3.1), except that an exposure time of 15 s was used and spectra were collected every 20 s. Blending endpoint determination was based on the conformity index (CI) method, as presented in Ref. [28]. As no diltiazem·HCl was used for this experiment (too expensive), the developed SIMCA model to predict blend homogeneity for the pure Raman experiments of Section 2.3.1 (cfr. 2.5) could not be applied.

2.4. Experimental design methodology

The effects of one formulation variable (concentration of diltiazem hydrochloride) and two process variables (mixing speed and loading) and their interactions on the time required to obtain a homogeneous blend (response variable) were studied. These variables are known to influence the required blending times. From an industrial perspective, it is important to characterize the effect of concentration on mixing because a manufacturer's product line may frequently include products of different strengths. Furthermore, it was shown that similar mixtures with different API concentrations display different endpoints of mixing [26]. Typically, industrial scale blenders mix at a lower blending speed compared to lab-scale blenders. The question raised is whether the time required to achieve blend homogeneity at a slower speed is equivalent to the one needed for mixing the same powder blend at a higher speed. Finally, it is evident that the amount of powder in the blender (loading) can affect the required blending time. Table 1 provides an overview of the variables (factors) with their lowest and highest levels.

A 2-level full factorial design was performed to determine the influence of the factors upon the response within the design space.

Table 1
Overview of the studied variables and their levels

Variable	Unit	Levels	
		–	+
Diltiazem·HCl concentration	% (w/w)	9.5	21.0
Blending speed	rpm	100	400
Loading	kg	2.5	3.0

In this design, $2^3 = 8$ experiments are required. It was chosen to perform 2 additional experiments at the center of the design space, resulting in a total of 10 experiments (Table 2).

2.5. In-line and real-time homogeneity determination

As the time required to obtain a homogeneous blend was the response variable, a mathematical method that allows *in-line* and *real-time* homogeneity detection by extracting useful information from the spectral data was essential. In the literature, several methods have been described for this purpose. These methods can be divided into two types.

The first type does not use any historical measurements or information: for instance, plotting the peak area(s) or height(s) of the Raman bands known to be produced by the compound of interest as a function of blending time [29], the mean square of differences (MSD) method between two consecutive spectra [11,30,31], the moving block standard deviation (MBSD) method [32–34], multivariate methods using dendrograms derived from hierarchical cluster analysis [35], score plots obtained from principal component analysis (PCA) [36], etc. For these methods, it is assumed that, once the mixture is homogeneous, the spectra will not change anymore.

The second type of methods uses a training set of spectra taken from one or more homogeneous mixtures. By comparing each new measurement with the training set spectra via dissimilarities [33,34,36,37], PCA [34,36,37], partial least squares (PLS) [15], CI [28,38], Euclidean distance (ED), Mahalanobis distance (MD) [29,38] or by using SIMCA models or Principal Component-Modified Bootstrap Error-Adjusted Single-Sample Technique (PC-MBEST) as pattern recognition techniques [39], it is examined when mixture homogeneity is reached. The ED and MD compare the distance of the observed spectra to the spectra that comprise the model space. This model space is set when a training set is established.

SIMCA was used in this study. SIMCA [40] is a very popular supervised classification method. Since an object can belong to one, to any, or to several groups at the same time, it is called a soft-classification method. The theory of SIMCA and some adaptations to the original SIMCA method are discussed in several papers [40–45].

Table 2
Overview of the performed experiments and obtained results

Experiment number	Diltiazem·HCl concentration (%, w/w)	Blending speed (rpm)	Loading (kg)	Results (s)
1	21	400	3	140
2	21	100	3	70
3	9.5	100	3	100
4	9.5	400	3	50
5	9.5	400	2.5	50
6	21	100	2.5	70
7	15.25	250	2.75	60
8	15.25	250	2.75	60
9	9.5	100	2.5	90
10	21	400	2.5	50

After pre-processing, SIMCA builds PCA models for each group (or class) of objects individually. The model set **X** is divided into subgroups, $\mathbf{X_k}$, with *m* objects of class K and *n* variables. For class K, the PCA model with *f* factors is obtained from singular value decomposition (SVD) of $\mathbf{X_k}$ after column centring.

$$\mathbf{X_k} - \overline{\mathbf{X_k}} = \mathbf{S_k} \cdot \mathbf{V_k} \cdot \mathbf{D_k^T} = \mathbf{Sc_k} \cdot \mathbf{D_k^T} + \mathbf{E_k} \quad (1)$$

In each column, the matrix $\overline{\mathbf{X_k}}$ contains the corresponding means of the $\mathbf{X_k}$ columns, and the score matrix $\mathbf{S_k}$ the f normalized score vectors. $\mathbf{V_k}$ is a diagonal matrix with singular values, $\mathbf{D_k}$ the matrix with f loadings, $\mathbf{Sc_k}$ the unnormalized score matrix of PC, containing f (unnormalized) score vectors, and $\mathbf{E_k}$ the matrix of residuals. To determine the optimal number of components in PCA, a cross-validation procedure is applied [40–43,45].

Then, class boundaries are defined, using model objects for each class. This can be done in several ways [40–45]. In our study, the scores of each object in a class are predicted using LOO cross-validation, and these cross-validated scores are then used to calculate the cut-off values. The MDs for the objects in the score space, describing the distances to the center of the PCA model, and the Orthogonal Distances (OD) from the PCA model, describing the deviations to the model or the residuals, are calculated for all objects in the model set and used to determine their cut-off values. For this purpose, the MD and OD distances (*d*) of the objects from class K are centered according to their corresponding means (mean (d_K)), and the MD and OD cut-off values are defined as three times the standard deviations of the corresponding vector elements [45].

$$\left| d - \text{mean}\,(d_K) \right| \geq 3 \cdot \sigma(d_K) \quad (2)$$

In other words, when assuming a normal distribution, 99.90% of the centered distances should fall within the interval of three times the standard deviations of these distances.

To verify whether a new test object belongs to class K, it is projected in the space defined by the selected factors (PC's) of the corresponding model set, $\mathbf{X_k}$. Then, the MD and OD are calculated for this object of the test set, and the distances are centered according to the corresponding class means.

An object with centered MD and/or OD larger than the corresponding cut-off value(s) is considered an outlier. Here, three situations are possible, which can be seen when plotting the OD versus the MD for each object (Fig. 2). The (- - -) lines on this plot represent their cut-off values. Objects situated in region I are called high residual objects or vertical outliers (high-residuals from PCA model, high OD), those in region II good leverage objects (far from majority of data, fit PCA model, high MD), and those in region III bad leverage objects (both high MD and OD).

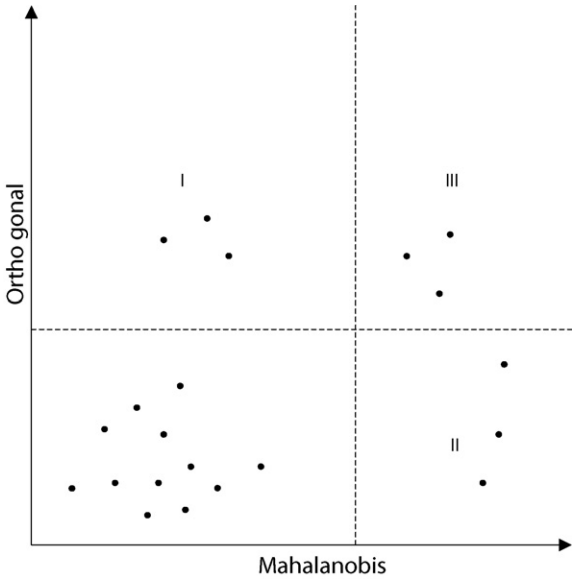

Fig. 2. Types of outlying objects when plotting the OD vs. the MD. The (- - -) lines represent their cut-off values; I = high-residual objects; II = good leverage objects; III = bad leverage objects.

3. Results and discussion

3.1. In-line and real-time endpoint determination

In Table 3, the number of spectra for each batch (=each design experiment) is given. Raman spectra were collected every 10 s during each blending experiment. The peak intensity as a function of Raman shift (cm^{-1}) for the 506 Raman spectra of the 10 batches is shown in Fig. 3.

The goal was to determine the time at which a blend is homogeneous. Therefore, the last seven spectra of each batch were chosen as reference spectra, since these Raman spectra certainly origin from a homogeneous blend. In our study, only one class of objects is used. These last seven reference spectra from each batch constitute the model set (70 spectra), and thus are used to build the SIMCA model and define the boundaries of this class. The other 436 spectra are considered as test set, and it is verified whether they belong

Table 3
Number of Raman spectra for each batch (=each design experiment)

Batch	Number of Raman spectra
1	45
2	35
3	48
4	46
5	51
6	74
7	61
8	33
9	57
10	56

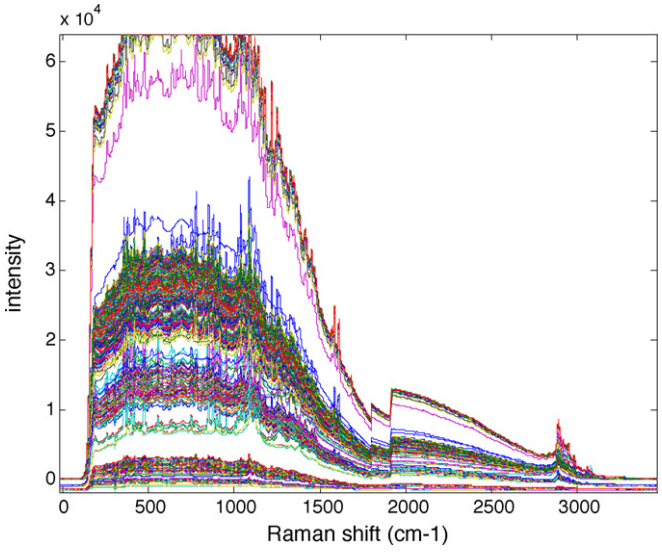

Fig. 3. Peak intensity as a function of Raman shift (cm^{-1}) for the 506 Raman spectra of the 10 batches. The range of Raman shifts from 1554.6 till 1629.6 cm^{-1}, is the range in which most differences are found between the spectra of different batches.

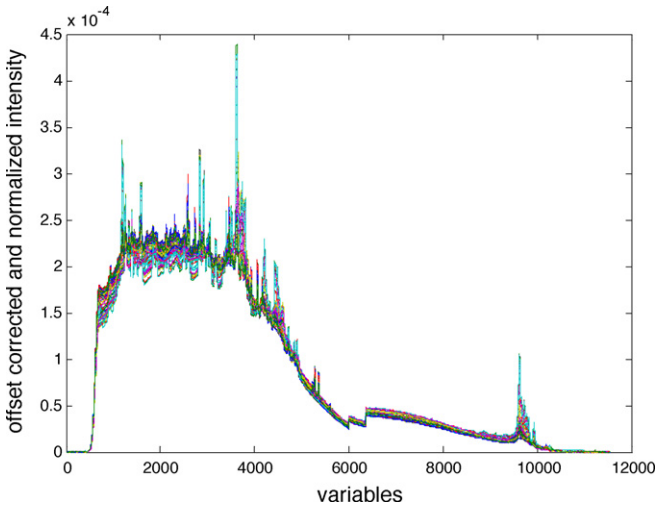

Fig. 4. Offset corrected and normalized peak intensity as a function of the variable number for the 506 Raman spectra of the 10 batches.

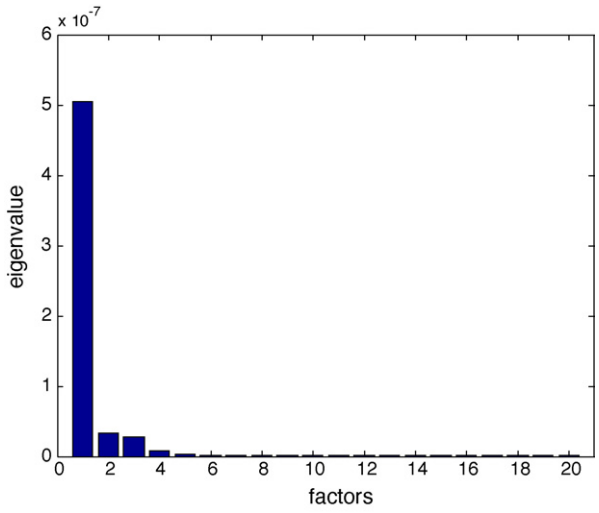

Fig. 6. Eigenvalues plotted as a function of the first 20 factors or PC's.

to the class of reference spectra or, in other words, whether or not the blend can be considered homogeneous.

Most differences between the spectra of different batches were found in the range of Raman shifts from 1554.6 till 1629.6 cm^{-1}. SIMCA models were built using both the entire spectrum and only this selected range of Raman shifts. Since a better model was obtained when using the entire spectrum, only these results are discussed below.

To pre-process the data, first an offset correction was applied, followed by normalization of the spectra. Since negative peaks occur in the range 294.6 till 315.6 cm^{-1}, the corresponding variable range was removed from the data. The spectra obtained after the above preprocessing are shown in Fig. 4.

Then, the data were centered according to three classes, i.e., the three different API concentrations (9.5, 15.25 and 21% (w/w) diltiazem·HCl concentration, see Table 2). The latter was necessary, since otherwise, when performing PCA, different groups are distinguished, depending on the API concentration.

Next, the 70 reference spectra were used to build a SIMCA model. Model complexity was determined by means of LOO cross-validation. In Fig. 5, the RMSECV is given as a function of the first 20 factors or PC's. Optimal complexity was found when including four factors. This is confirmed by Fig. 6, where the eigenvalues are plotted as a function of the first 20 factors or PC's.

The plot of the OD versus the MD for the 70 spectra of the model set is given in Fig. 7. The cut-off values for MD and OD then are calculated as $MD_{model,centered} \pm 3 \times$ standard deviations and $OD_{model,centered} \pm 3 \times$ standard deviations.

The calculated MD and OD for the 436 spectra of the test set are represented in Fig. 8. The plot of the OD versus the MD for the spectra of the test set is given in Fig. 9. On both figures, the cut-off values for both the OD and MD are represented by blue lines. In Fig. 9, points representing spectra situated in region I are called high residual objects, those in region II good leverage objects, and those in region III bad leverage objects. It can be seen that at the beginning of the blending time for each batch, the MD and OD are large (see Fig. 8), which on Fig. 9 represent points representing spectra situated outside domain determined by the MD and OD cut-off values.

Finally, the Raman peak area (P; 1554.6–1629.6 cm^{-1}) is plotted as a function of blending time for each batch (Fig. 10; only batch 1 is shown). Outlying spectra, i.e., situated in regions I, II or III of

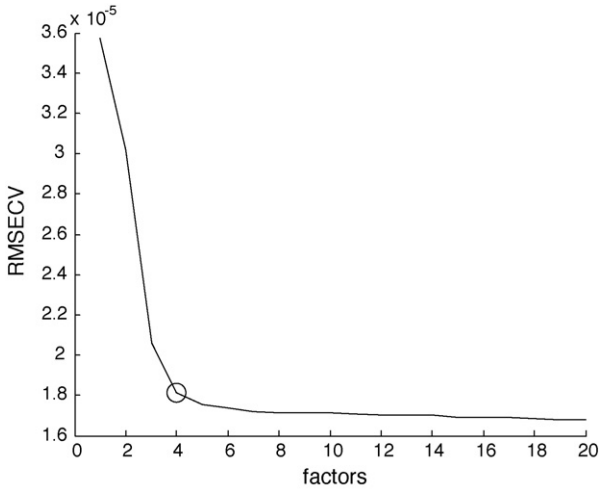

Fig. 5. RMSECV plotted as a function of the first 20 factors or PC's.

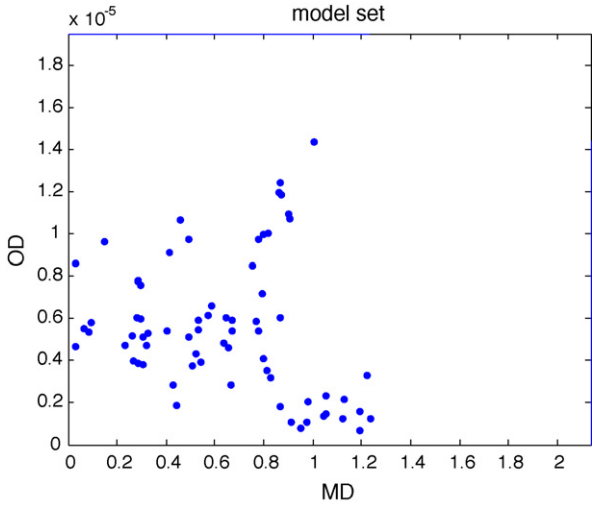

Fig. 7. SIMCA: OD plotted as a function of the MD for the 70 spectra of the model set.

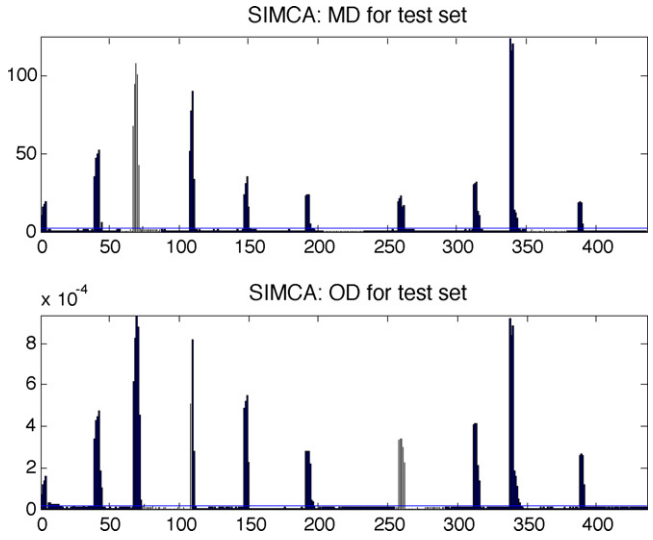

Fig. 8. SIMCA: OD and MD for the 436 spectra of the test set. The horizontal lines represent the cut-off values.

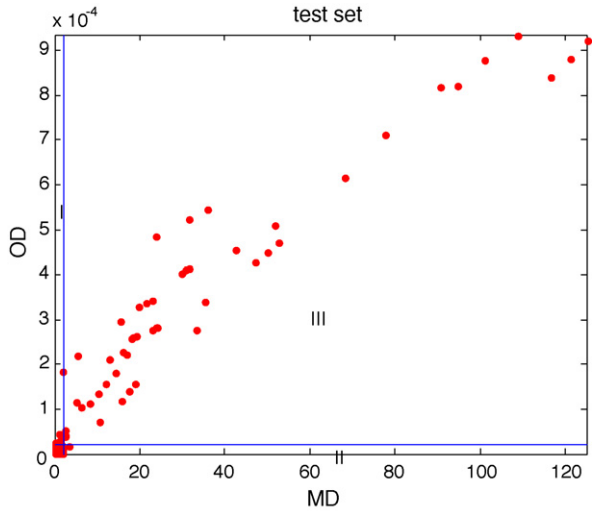

Fig. 9. SIMCA: OD plotted as function of the MD for the 436 spectra of the test set. The lines represent the cut-off values. I = high residual objects. II = good leverage objects. III = bad leverage objects.

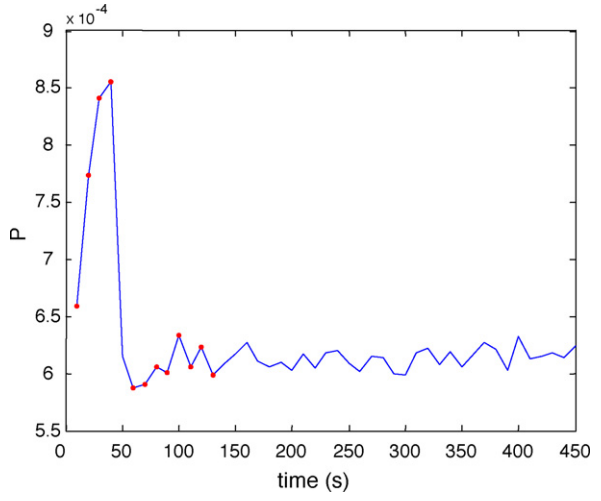

Fig. 10. Raman peak area (P) plotted as a function of blending time for batch 1 (experiment 1). Outlying spectra are shown with dots (regions I, II or III of Fig. 9).

Fig. 9, are shown with dots on Fig. 10. The time points at which homogeneity is reached are included in Table 2.

3.2. Screening design

From the experimental results (Table 2), a regression model for the blending time was calculated (i.e., the effects and interactions were estimated) by Statgraphics® *Plus*. Since two-level full factorial designs take into account all linear terms, and all possible k way interactions, the number of different types of terms can be predicted by the binomial theorem: $k!/(k-m)!m!$ for mth-order interactions and k factors. As the third order interaction is considered negligible, only second order ($m = 2$) interactions were estimated [46,47]. Hence, the following equation was built:

$$y = b_0 + b_1A + b_2B + b_3C + b_{12}AB + b_{13}AC + b_{23}BC \quad (3)$$

where A is the concentration of diltiazem (%, w/w), B the mixing speed (rpm), C the loading (kg), $b_0, b_1, b_2, b_3, b_{12}, b_{13}, b_{23}$ the regression coefficients, and y the response variable (required blending time in s).

The calculated model based on the experimental results is $y = 74.0 + 5.0A - 5.0B + 12.5\,C + 17.5AB + 10.0AC + 10.0BC$.

For the calculation of this model, each factor was put on an equal scale, with the highest level coded +1 and the lowest −1. The effects of all variables (A, B, C) and variable interactions (AB, AC, BC) are the double of their corresponding coefficients in the equation (i.e., A: 10, B: 10, C: 25, AB: 35, AC: 20 and BC: 20) [46].

Next, the significance of the effects (and hence the coefficients) from the factors (A, B, C) and interactions (AB, AC, BC) was examined using a standardized Pareto chart. A standardized Pareto chart is a visualization of the Student's t-test used to evaluate the effects significance. Fig. 11 shows the standardized Pareto chart for $\alpha = 0.05$. It indicates each effect in decreasing order of magnitude. The length of the bar is equal to the standardized effect, calculated as the estimated effect divided by its standard error. The vertical line on the plot is the critical t-value. Bars extending beyond the line correspond to effects that are statistically significant at the 95% confidence level. From Fig. 11, it is concluded that none of the factors and interactions have a significant influence on the response.

The standardized effects were calculated by dividing the effects by its standard error s_D.

$$s_D = \frac{\sqrt{2s^2}}{\sqrt{n}} \quad (4)$$

where n = number of factors or interactions

$$s^2 = \frac{\sum(\text{measured value} - \text{predicted value})^2}{\text{remaining degrees of freedom}} \quad (5)$$

Hence, s_D was calculated on the basis of the residuals from the calculated model, which means that the error estimate is dependent from the complexity of the model. Including a third order

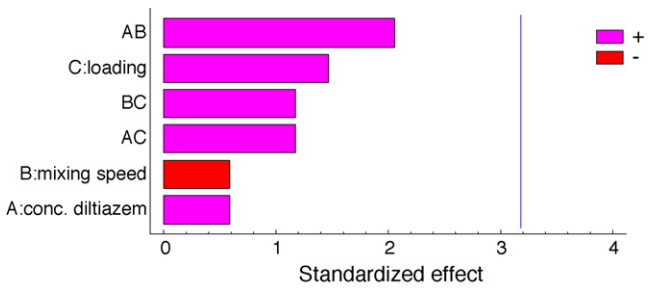

Fig. 11. Standardized Pareto chart for $\alpha = 0.05$.

interaction factor in the model would decrease the residuals (as the fit increases) and hence also the s_D, which is not logic. An effect should be considered as significant or non-significant in all situations. A good estimate of the error should always lead to the same conclusion. Therefore, adequate error estimates using dummies, intermediate precision estimates or the Dong algorithm are advised [48]. However, evaluation of the significance of the effects using the Dong algorithm also indicated that none of the effects is significant.

Both the Pareto chart and Dong's approach seem to indicate that the uncontrollable random variation of the blending processes appears to be higher than the variation caused by the studied process and formulation variables. In fact, the apparent finding that this blending process endpoint is not significantly influenced by the examined factors in the experimental design forced us to consider the possible reasons; process thinking and process understanding being two essential features of the PAT strategy.

However, the conclusion of non-significant effects drawn from the statistical analyses should be approached with some cautiousness. As the effect from each parameter upon the response is calculated by subtracting the average of the response values at the low level of the considered parameter from the average of the response values at the high level of the same parameter, the effects from *A*, *B*, *C*, *AB*, *AC* and *BC* are 10, −10, 25, 35, 20 and 20, respectively. This means for the smallest estimated effect (e.g., effect *A* = 10) that the blending time will change 10 s by altering factor *A* from −1 to 1. Hence, the response varies approximately 13% as the average response of the design is 77.5 s. From a practical point of view and without considering any statistical result, such a change would be considered relevant in many cases. A similar effect will only be statistically insignificant when there is a major variability in the blending time of replicated experiments. This is not shown from our study as the replicated experiments (7 and 8) resulted in an identical blending. However, we assume that this was by coincidence as the measurement time intervals (10 s) are large with respect to the short blending times (between 50 and 100 s) for all different processes. We can state that to be able to estimate both effects and critical effects on *t*-values properly, the measurement time interval should be considerably smaller (e.g., 1 s). However, the instrumental limitations might make this impossible. The excessively rounded blending times resulted in insensitive effect estimates from which we consider it inappropriate to interpret them physicochemically. The above learns that the examination of increasing process efficiency is mainly relevant for processes having long process times (i.e., much longer than the possible measurement time intervals). For the present study, where the process endpoints are reached between 1 and 2 min for all experiments, it is of higher importance to have a reliable method allowing the correct detection of the process endpoint.

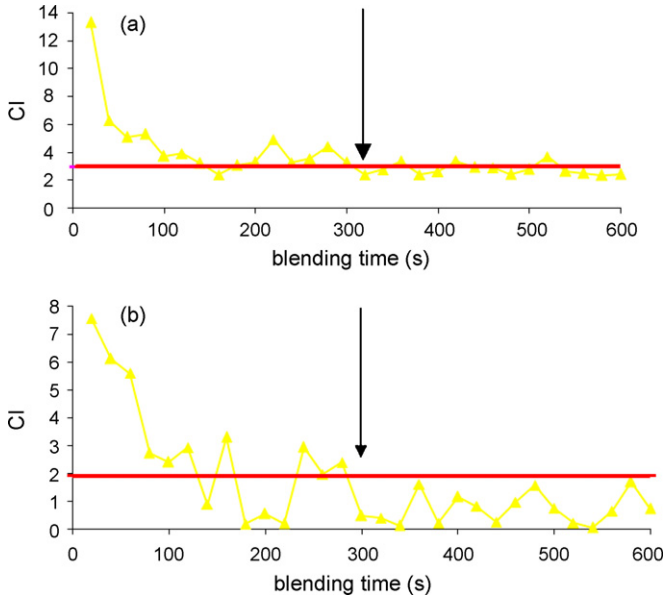

Fig. 12. (a) CI plot for NIR data: homogeneity is reached after about 300 s. (b) CI plot for Raman data: homogeneity is reached after about 300 s.

3.3. Raman-NIR comparison experiment

The CI versus blending time plots for the Raman and NIR measurements are given in Fig. 12(a and b).

The Raman and NIR measurements lead to similar endpoint conclusions (blending time of 300 s). Hence, both independent process-monitoring techniques confirm each other. This clearly increases the credibility and certainty of the process analysis conclusions.

4. Conclusion

A strategy is proposed to implement a PAT system in a powder blending process, hence avoiding sampling errors and making time-consuming and labour intensive *off-line* analyses unnecessary. In a first instance, a Raman spectroscopic method based on a chemometric (SIMCA) model was developed that allows *in-line* and *real-time* monitoring of the endpoint (homogeneity) of the mixing process. Important changes during the process can be detected immediately, hence the necessary adjustments can be undertaken, if necessary, to avoid batch loss.

From an experimental design approach, the effects of process and formulation parameters upon the response can be studied, which makes it possible to understand the process better and to make the process as efficient as possible.

The correctness of the Raman endpoint conclusions can be assured by using a second independent *in-line* endpoint-monitoring tool (NIR spectroscopy) simultaneously. Equal process endpoint determinations provided by two independent monitoring techniques clearly increase the credibility and certainty of the process analysis conclusions.

References

[1] M.L. Balboni, Pharm. Technol. 27 (2003) 54–66.
[2] R.J. Romañach, J. PAT 1 (2004) 10–12.
[3] A.M. Afnan, J. PAT 1 (2004) 8–9.
[4] A.S. Hussain, J. PAT 2 (2005) 8–13.
[5] J. Berman, Drug Dev. Ind. Pharm. 21 (1995) 1257–1283.
[6] J. Berman, A. Schoeneman, J.T. Shelton, Drug Dev. Ind. Pharm. 22 (1996) 1121–1132.
[7] R.K. Chang, J. Shukla, J. Buehler, Drug Dev. Ind. Pharm. 22 (1996) 1031–1035.
[8] T.P. Garcia, M.K. Taylor, G.S. Pande, Pharm. Dev. Technol. 3 (1998) 7–12.
[9] F.J. Muzzio, M. Roddy, D. Brone, A.W. Alexander, O. Sudah, Pharm. Technol. 23 (1999) 92–110.
[10] A.S. El-Hagrasy, H.R. Morris, F. D'Amico, R.A. Lodder, J.K. Drennen, J. Pharm. Sci. 90 (2001) 1298–1307.
[11] G.J. Vergote, T.R.M. De Beer, C. Vervaet, J.P. Remon, W.R.G. Baeyens, N. Diericx, F. Verpoort, Eur. J. Pharm. Sci. 21 (2004) 479–485.
[12] K.A. Bakeev, Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries, Blackwell Publishing Ltd., Oxford, 2005.
[13] J. Rantanen, J. Pharm. Pharmacol. 59 (2007) 171–177.
[14] T.R.M. De Beer, W.R.G. Baeyens, J. Ouyang, C. Vervaet, J.P. Remon, Analyst 131 (2006) 1137–1144.
[15] O. Berntsson, L.G. Danielsson, B. Lagerholm, S. Folestad, Powder Technol. 123 (2002) 185–193.
[16] A.S. El-Hagrasy, J.K. Drennen, J. Pharm. Sci. 95 (2006) 422–434.
[17] W. Li, M.C. Johnson, R. Bruce, H. Rasmussen, G.D. Worosila, J. Pharm. Biomed. Anal. 43 (2007) 711–717.

[18] C.J. Strachan, T. Rades, K.C. Gordon, J. Rantanen, J. Pharm. Pharmacol. 59 (2007) 261–269.
[19] J. Luypaert, D.L. Massart, Y. Vander Heyden, Talanta 72 (2007) 865–883.
[20] T. Herkert, H. Prinz, K.A. Kovar, Eur. J. Pharm. Biopharm. 51 (2001) 9–16.
[21] R.P. Cogdill, C.A. Anderson, M. Delgado, R. Chisholm, R. Bolton, T. Herkert, A.M. Afnan, J.K. Drennen, AAPS PharmSciTech. 6 (2005) E262–E272.
[22] M. Blanco, M. Alcalá, Anal. Chim. Acta 557 (2006) 353–359.
[23] C. Bodson, E. Rozet, E. Ziemons, B. Evrard, P. Hubert, L. Delattre, J. Pharm. Biomed. Anal. 45 (2007) 356–361.
[24] E.T.S. Skibsted, J.A. Westerhuis, A.K. Smilde, D.T. Witte, J. Pharm. Biomed. Anal. 43 (2007) 1297–1305.
[25] U.S., Department of Health, Human Services, Food, Drug Administration, Center for Drug Evaluation, Research, Center for Veterinary Medicine, Office of Regulatory Affairs. Guidance for Industry: PAT–A Framework for Innovative Pharmaceutical Manufacturing and Quality Assurance, 2004.
[26] A.S. El-Hagrasy, F. D'Amico, J.K. Drennen, J. Pharm. Sci. 95 (2006) 392–406.
[27] R.L. McCreery, Raman Spectroscopy for Chemical Analysis, Wiley-Interscience, New York, 2000.
[28] C. Bodson, W. Dewé, P. Hubert, L. Delattre, J. Pharm. Biomed. Anal. 41 (2006) 783–790.
[29] D.S. Hausman, R.T. Cambron, A. Sakr, Int. J. Pharm. 298 (2005) 80–90.
[30] M. Blanco, R.G. Bano, E. Bertran, Talanta 56 (2002) 203–212.
[31] M. Popo, S. Romero-Torres, C. Conde, R.J. Romañach, AAPS PharmSciTech. 3 (2004) (article 24).
[32] P. Hailey, P. Doherty, P. Tapsell, T. Oliver, P. Aldridge, J. Pharm. Biomed. Anal. 14 (1996) 551–559.
[33] S.S. Sekulic, H.W. Ward, D.R. Brannegan, E.D. Stanley, C.L. Evans, S.T. Sciavolino, P.A. Hailey, P.K. Aldridge, Anal. Chem. 68 (1996) 509–513.
[34] S. Sekulic, J. Wakeman, P. Doherty, P. Hailey, J. Pharm. Biomed. Anal. 17 (1998) 1285–1309.
[35] T. Norris, P.K. Aldridge, Analyst 121 (1996) 1003–1008.
[36] F. Cuesta Sanchez, J. Toft, B. van den Bogaert, D.L. Massart, S.S. Dive, P. Hailey, Fresen. J. Anal. Chem. 352 (1995) 771–778.
[37] R. De Maesschalck, P. Cuesta Sanchez, D.L. Massart, P. Doherty, P. Hailey, Appl. Spectrosc. 52 (1998) 725–731.
[38] G.E. Ritchie, H. Mark, E.W. Ciurczak, AAPS PharmSciTechnol. 4 (2003) (article 24).
[39] A.S. El-Hagrasy, M. Delgado-Lopez, J.K. Drennen, J. Pharm. Sci. 95 (2006) 407–421.
[40] S. Wold, Pattern Recogn. 8 (1976) 127–139.
[41] S. Wold, M. Sjöström, J. Chemom. 1 (1987) 243–245.
[42] R. De Maesschalck, A. Candolfi, D.L. Massart, S. Heuerding, Chemom. Intell. Lab. Syst. 47 (1999) 65–77.
[43] A. Candolfi, R. De Maesschalck, D.L. Massart, P.A. Hailey, A.C.E. Harrington, J. Pharm. Biomed. Anal. 19 (1999) 923–935.
[44] K. Vanden Branden, M. Hubert, Chemom. Intell. Lab. Syst. 79 (2005) 10–21.
[45] M. Daszykowski, K. Kaczmarek, I. Stanimirova, Y. Vander Heyden, B. Walczak, Chemom. Intell. Lab. Syst. 87 (2007) 95–103.
[46] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part A, Elsevier, Amsterdam, 1997.
[47] R.G. Brereton, Chemometrics: Data Analysis for the Laboratory and Chemical Plant, Wiley-Interscience, Chichester, 2003.
[48] Y. Vander Heyden, A. Nijhuis, J. Smeyers-Verbeke, B.G.M. Vandeginste, D.L. Massart, J. Pharm. Biomed. Anal. 24 (2001) 723–753.