ELSEVIER

Journal of Pharmaceutical and Biomedical Analysis
17 (1998) 141–152

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

Influence and correction of temperature perturbations on NIR spectra during the monitoring of a polymorph conversion process prior to self-modelling mixture analysis

K. DeBraekeleer [a,*], F.Cuesta Sánchez [a], P.A. Hailey [b], D.C.A. Sharp [b], A.J. Pettman [b], D.L. Massart [a]

[a] *Vrije Universiteit Brussel, ChemoAC, Laarbeeklaan 103, B-1090 Brussels, Belgium*
[b] *Pfizer Central Research, Sandwich, Kent CT13 9NJ, UK*

Received 9 April 1997; accepted 15 July 1997

Abstract

The influence of temperature variations on the rank of a NIR dataset, has been investigated by comparing the results of principal component analysis (PCA) and evolving factor analysis (EFA), applied to two datasets measured at constant temperature and varying temperature. After temperature correction, the concentration profiles and spectra were obtained with PCA, SIMPLISMA and the orthogonal projection approach (OPA). The same resolution methods were used on the dataset measured at constant temperature.

Keywords: Reaction monitoring; NIR spectroscopy; Temperature influence; Temperature correction; Polymorph conversion; Self-modelling mixture analysis; Chemometrics

1. Introduction

The understanding of how a reaction proceeds is at the core of chemistry. Typically, analytical techniques used to 'follow' the reaction progress involve sampling and to some extend perturbation of the reaction medium. A technique that lends itself to on-line non-invasive monitoring is NIR spectroscopy. NIR spectroscopy is an ideal process technique as it requires no sample preparation and using suitable sample interfaces measurements from heterogeneous and homogeneous reactions can be made. The NIR spectra that result from monitoring the reaction process contain both physical and chemical information and often require the application of statistical approaches to extract meaningful information from the spectroscopic data.

Different computational methods have been developed for analysing spectral data obtained from evolutionary processes such as processes monitored as a function of time. The objective of these resolution techniques is to resolve the data into the concentration profiles and the spectra of the components present. In general they can be classified into two groups: hard-modelling and

* Corresponding author.

0731-7085/98/$19.00
PII S0731-7085(97)00177-5

soft or self-modelling methods. The advantage of self-modelling in comparison with hard-modelling is that no prior information about the shape of the pure spectra and/or concentration profiles is required. The self-modelling approaches encompass methods based on principal component analysis (PCA) and methods which compare rows or columns of the data matrix with a reference. Methods such as evolving factor analysis (EFA) [1], fixed size moving window evolving factor analysis (FSW-EFA) [2], eigenstructure tracking analysis (ETA) [3], heuristic evolving latent projections (HELP) [4–6] and iterative target transformation factor analysis (ITTFA) [7] belong to the first group, while SIMPLISMA [8–10] and the orthogonal projection approach (OPA) [11] can be classified in the second group. These self-modelling approaches require a bilinear data structure, which means that the effects due to different compounds are additive and linear with concentration.

In this paper two sets of data are considered. The first concerns the NIR spectra obtained as a function of time during the polymorph conversion of species P_1 to P_2 at constant temperature. The second describes the same process, carried out with varying temperature. The aim is to be able to obtain the concentration profiles in both cases. This is made more difficult in the second case because temperature has an influence on the NIR spectra and this must be corrected for to achieve the objective of the analysis.

2. Theory

2.1. Data pretreatment

Variation within individual near-infrared spectra is the result of the physical characteristics and the chemical composition of the sample. Therefore the raw spectra are mathematically pretreated in order to eliminate the differences between them due to physical sources. Several mathematical transformations such as the standard normal variate transformation (SNV) [12], detrending [12], linear baseline correction and second derivation are often used for this purpose. The advantage of the second derivation method in comparison with SNV, detrending and linear baseline correction is that it enhances small differences between similar spectra. This can be useful to detect selective wavelengths in the spectral data. Therefore, the convolution method of Savitzky and Golay [13] with a second order polynomial and a window size 11 was used in this paper.

2.2. Analysis of the data

2.2.1. Principal component analysis

The NIR spectra were collected on a NIRSystems 6500 spectrophotometer using a reflectance fibre-optic probe. The typical wavelength range is measured at 700 wavelengths (1100–2500 nm). The spectra can be represented as points in the wavelength space. PCA is performed to visually display the spectra in a two dimensional latent variable space defined by principal component 1 (PC_1) and principal component 2 (PC_2).

2.2.2. Resolution techniques

2.2.2.1. Evolving factor analysis [1]. The principle of EFA is to follow the change of the rank of a data matrix as a function of the ordered variable, which for evolutionary processes is typically time. For this purpose PCA is performed on an increasing data matrix, starting with the first or last row and adding one row at the time. The logarithms of the singular values (SV) are plotted as a function of the ordered variable. To make the graph clearer, all points calculated for the first SV are connected with a line; analogously, lines are also drawn to connect the second SV and so on. If one starts with the first row the method is called forward EFA. The graph in theory, indicates the time were each compound appears. Starting with the last row the method is called backward EFA. The plot must now be read in reverse order, from right to left, to obtain information about the disappearence of the components. The number of components, or the rank of the data matrix, can be determined by visual inspection of the EFA graphs. The number of SV's rising out of the noise level indicate the number of components present in the data matrix.

2.2.2.2. SIMPLISMA. SIMPLISMA [8–10], a self-modelling approach, is based on the selection of what are called pure variables. They can be found either in the wavelength or in the time direction of the measured spectral data matrix and are defined as variables of which the intensities are due to only one of the compounds present in the data set.

If the SIMPLISMA approach is applied in the wavelength direction, the intensities measured for each selected pure wavelength represent the relative amounts of the unknown compound, for which the wavelength is selected, in the mixture spectra of the data matrix. Performing SIMPLISMA in the time direction results in the selection of regions with pure spectra.

Once the pure variables (wavelengths or times) have been determined, the data set can be resolved into the pure spectra and the concentration profiles using a least squares procedure. The data matrix $\mathbf{X}$ $(n \times m)$ can be expressed by the following equation:

$$\mathbf{X} = \mathbf{CA} \tag{1}$$

in which the columns of the matrix $\mathbf{C}$ $(n \times k)$ contain the relative concentrations of the k pure components in the mixture spectra and the rows of the matrix $\mathbf{A}$ $(k \times m)$ contain the pure spectra. If SIMPLISMA is performed in the wavelength direction, the matrices $\mathbf{X}$ and $\mathbf{C}$ are known, and the pure spectra $\mathbf{A}$ can be obtained by using the pseudo-inverse $(\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'$ of the matrix $\mathbf{C}$:

$$\mathbf{A} = (\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\mathbf{X} \tag{2}$$

The concentration profiles are then recalculated using the pseudo-inverse $\mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}$ of the matrix $\mathbf{A}$:

$$\mathbf{C} = \mathbf{X}\mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1} \tag{3}$$

The selection of the pure variables and more specifically of pure spectra has been extensively explained in [8], therefore, only a brief description is given here. The purity of each spectrum, p_i , is calculated with the following equation:

$$p_i = w_i \times \frac{\sigma_i}{\mu_i + \alpha} \quad i = 1, ..., n \tag{4}$$

where μ_i is the mean [8] and σ_i the standard deviation [8] of each spectrum x_i. The offset α is a percentage of the maximum mean value (μ_i). It is included to avoid that spectra with low mean intensity obtain high purity values. The weight of each spectrum, w_i , is a measure for the dissimilarity of the spectrum with the selected ones and it is calculated as the determinant of the dispersion matrix $\mathbf{Y}_i \cdot \mathbf{Y}_i'$ of $\mathbf{Y}_i$, which contains the pure spectra selected and the normalised spectrum z_i. The spectra in the matrix $\mathbf{Y}_i$ are normalised according to the equation [8]:

$$z_{ij} = \frac{x_{ij}}{\sqrt{m(\sigma_i^2 + (\mu_i + \alpha)^2)}} \text{ for } i = 1, \ldots, n \text{ and } j = 1, \ldots, m \tag{5}$$

Initially, when no spectrum has been selected the matrix $\mathbf{Y}_i$ contains only the normalised spectrum. The weight, w_i, is then equal to the length of the normalised spectrum. The purity of each spectrum is calculated and plotted as function of time. The spectrum with the highest purity is selected and included in the matrix $\mathbf{Y}_i$ after normalisation. The matrix $\mathbf{Y}_i$ now contains two spectra, the first selected one and spectrum z_i. The determinant of the dispersion matrix of $\mathbf{Y}_i$ measures the area of the parallellogram determined by spectrum z_i and the first selected spectrum. The higher the area, the higher the weight and purity value for spectrum z_i. The weight factors thus avoid the selection of variables close to the selected ones.

To confirm that all the information has been substracted, the total signal spectrum can be calculated [8]. The total signal spectrum, t, is a linear combination of the pure spectra of the compounds present in the mixture:

$$t = C_1S_1 + C_2S_2 + \cdots + C_nS_n \tag{6}$$

where C_1, C_2,..., C_n are the total concentrations of the compounds and S_1, S_2,..., S_n are the molar absortivities at the different wavelengths for each compound. When the proper number of pure variables has been selected, the approximated total intensity spectrum should be equal to t. The criterion used to measure the difference is the square root of the relative sum of squares differ-

ences between the total signal spectrum and the approximated total signal spectrum.

In our context the NIR spectra of P_1 and P_2 are expected to be very similar, so that no pure zones in the spectra will be found. According to the literature, SIMPLISMA can be applied in the wavelength direction after second derivation. The selection of pure (or at least the purest possible) wavelengths is performed on the positive part of the inverted second derivative spectra. The inversion (i.e. change of sign) is useful because in this way maxima in the second derivative spectra coincide with maxima in the original spectra [9,10]. For the resolution the original dataset is used in combination with the selected pure variable intensities from the positive part of the inverted second derivative data.

2.2.2.3. Orthogonal projection approach [11]. OPA is based on the determination of the dissimilarity. The dissimilarity of each spectrum is defined as the determinant of the dispersion matrix $\mathbf{Y}_i \cdot \mathbf{Y}_i'$ of $\mathbf{Y}_i$, which contains initially the mean spectrum of the data matrix, normalised to length = 1, as a reference and the unchanged measured spectrum i. The determinant of the dispersion matrix of $\mathbf{Y}_i$ measures the area of the parallellogram determined by spectrum i and the mean spectrum. The first spectrum selected is thus the one most dissimilar with respect to the mean spectrum. In the next step, the first selected spectrum, normalised to length = 1, is taken as a reference in $\mathbf{Y}_i$. The second dissimilarity plot represents the dissimilarity of each spectrum with respect to the first selected one. The process is repeated including the first and the second spectrum selected, normalised to length = 1, as references in the matrix $\mathbf{Y}_i$. The procedure continues adding new reference spectra, until the dissimilarity plot represents only noise. The number of selected spectra equals the number of absorbing compounds in the system. Those selected spectra are then used for the decomposition of the matrix into the concentration and pure spectra matrices, using an alternating least-squares approach.

3. Data

Two NIR datasets, measured on-line in a reaction vessel, were available. The spectra of both matrices were measured at 700 different wavelengths (from 1100 to 2500 nm). Signal attenuation due to the fibre optic necessitates the truncating of spectral data from 2200 to 2500 nm. The matrix obtained at a constant temperature of 30°C contains 80 spectra. The matrix obtained with variation in temperature contains 127 spectra. The first 15 spectra of this matrix were measured at an ambient temperature of 21°C. The temperature rises slowly from then on and reaches 60°C at the time that spectrum 38 is measured. This temperature is maintained until spectrum 115 is measured. From this time on the cooling process starts which ends at a temperature of 25°C.

4. Results and discussion

4.1. Influence of thermal effects upon rank analysis

The reaction in which the unwanted polymorph (P_1) is converted into the prefered one (P_2) is followed by measuring NIR-spectra at regular time intervals. The score plot obtained for the second derivative spectral data, obtained when the conversion is performed at a constant temperature of 30°C, is shown in Fig. 1a. Monitoring the reaction performed with variation in temperature results in the score plot presented in Fig. 1b.

The score plot obtained after PCA on the second derivative spectra of the combined matrices is shown in Fig. 1c. The spectra measured with varying temperature are distinguished from those measured at constant temperature, by adding the value 80 to each spectrum number. This means that the measured spectra 1, 13, 39, 115, 121,..., 127 correspond with the numbers 81, 93, 119, 195, 201,..., 207 in the score plots.

Because P_1 is converted in P_2, the sum of the concentrations at each time is constant. The spectra measured in the selective chromatographic regions of P_1 and P_2 will therefore be plotted as two clusters. The mixture spectra of P_1 and P_2 will be situated between these two clusters.

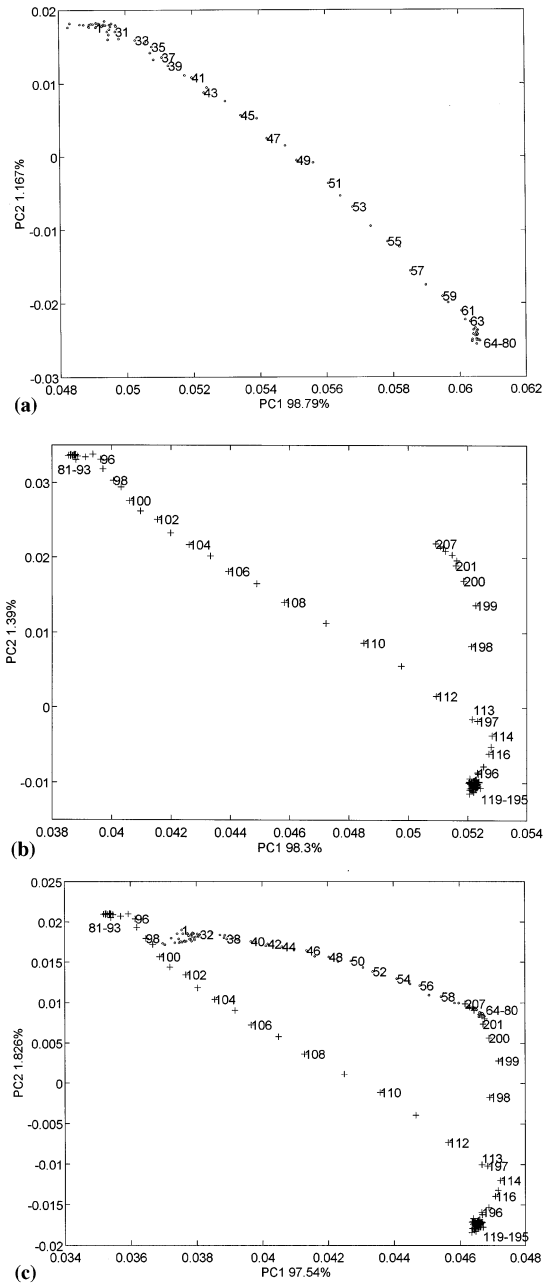

Fig. 1. (a) Score plot of second derivative spectra measured at constant temperature. (b) Score plot of second derivative spectra measured at varying temperature. (c) Score plot obtained after PCA on the second derivative spectra of the combined matrices (constant (..) and varying (+) temperature).

When the conversion is performed at 30°C two clusters 1–32 and 64–80 corresponding with P_1 and P_2 are respectively obtained. The mixture spectra of P_1 and P_2 (33–63) are situated between these two clusters on a straight line. The variation in the data is almost completely explained by PC_1, which indicates that the variation due to composition changes is mainly explained by this principal component. The conversion with varying temperature shows three clusters (81–93, 119–195 and 201–207). The small cluster (81–93) represents spectra of P_1 taken at a temperature of 21°C. The mixture spectra 94–113 measured at different temperatures varying from 21 to 51.5°C are also situated on a straight line. The big cluster containing the scores 119–195 corresponds with the spectra obtained at a temperature of 60°C. Spectra 196–200 are taken during the gradual cooling of the sample. The scores 201–207, representing the end product P_2 at 25°C, coincide with the cluster 64–80 corresponding with P_2 at 30°C. Due to the variation between the scores 119–200 mainly explained by PC_2, one can conclude that these spectra represent the change of the spectra of P_2 with temperature. Thus, the spectra 114–118 also represents P_2 at 53–58°C. The clusters 1–32 and 81–93 indicating P_1 at 30 and 21°C are found in each others neighbourhoods but do not coincide, which must be due to the temperature difference of 9°C.

To confirm the above mentioned temperature influences on the spectra of P_1 and P_2 a comparison was made between the raw mean spectra, normalised to maximum intensity, of the clusters observed in Fig. 1c (Fig. 2). In Fig. 2a the mean spectra of P_1 (..) and P_2 (–) both at 30°C are shown. The spectral regions 1600–1700 and 2100–2200 nm show some minor bands which will be used to further distinguish the spectrum of P_1 from that of P_2. The mean spectra of the clusters 1–32 (–) and 81–93 (..) are given in Fig. 2b. The shape of both spectra is similar. The small differences in absorbance values are due to a temperature difference of 9°C. The same conclusions can be drawn from Fig. 2c which represents P_2 at 30°C (–) and at 25°C (..). In Fig. 2d the mean spectrum of the cluster 119–195 (..) is plotted together with the mean spectrum of the cluster

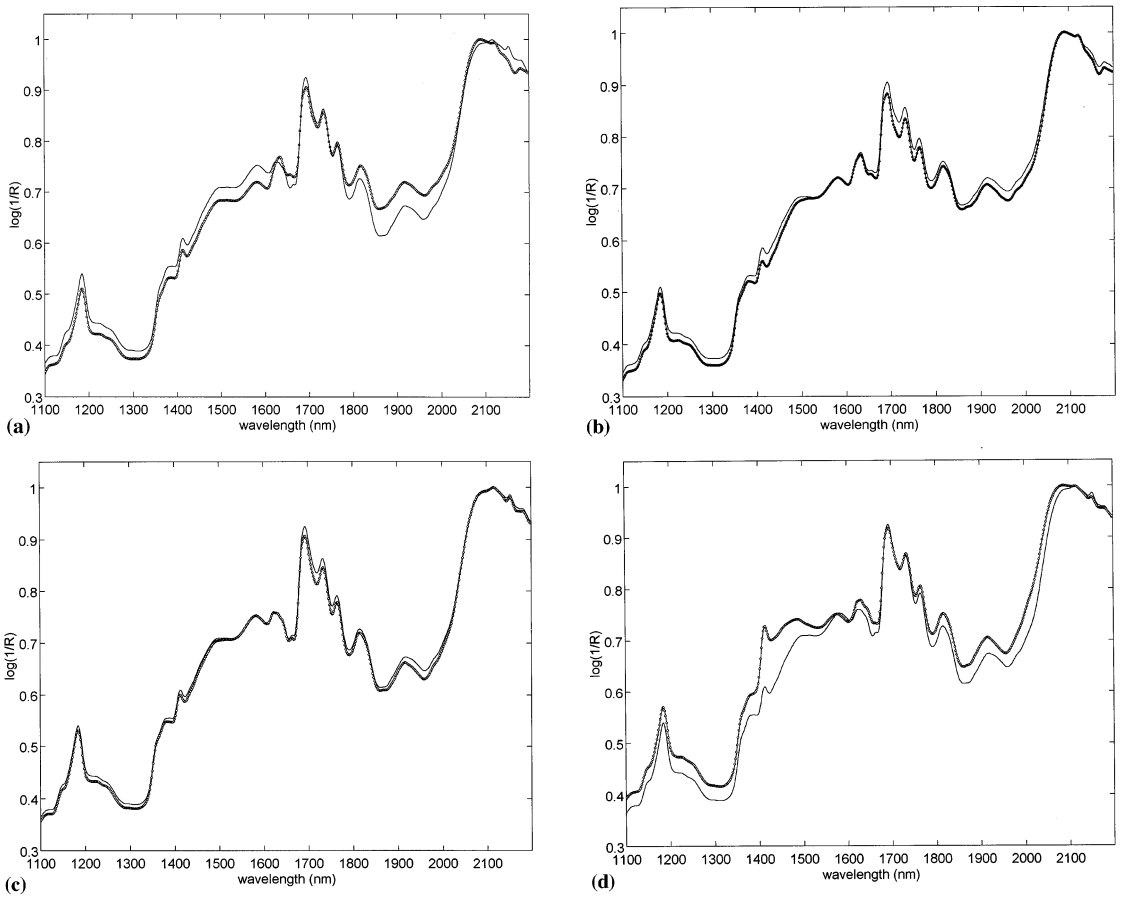

Fig. 2. Comparison of the raw mean spectra, normalised to maximum intensity, of the clusters observed in Fig. 1. (a) Mean spectra of the cluters 1–32 (..) and 64–80 (–). (b) Mean spectra of the clusters 1–32 (–) and 81–93 (..). (c) Mean spectra of the clusters 64–80 (–) and 201–207 (..). (d) Mean spectra of the clusters 64–80 (–) and 119–195 (..).

64–80 (–). The spectral regions 1600–1700 and 2100–2200 nm indicate that the mean spectra both correspond with P_2. Nevertheless it might be argued that spectral changes of this type are due to a temperature dependent equilibrium of distinct chemical species. The cluster 119–195 could thus be due to an unexpected third chemical component giving a true chemical rank of three. In this case the simplest and most probable interpretation is to attribute the third source of variability to the change of the spectrum of P_2 due to temperature. The difference between the spectra is the temperature effect on the spectrum of P_2. The process of conversion from P_1 to P_2 results from microsolution of the polymorph in the organic solvent.

In Fig. 3 the forward EFA graphs are shown for both matrices after second derivation. The number of chemical compounds can be determined visually from these plots as the number of SVs with a sharp increase from the noise. The noise level is estimated as the largest SV which is caused by noise and obtained when PCA is performed on the whole data matrix. For the matrix measured at constant temperature the noise level is calculated as the logarithm of the third SV. The noise level for the matrix with temperature effects is considered to be the logarithm of the fourth SV. Thus, the number of chemical compounds or the chemical rank of the matrix, measured at constant

temperature, is = 2. For the matrix obtained with varying temperature the observed rank is three.

Variation in the temperature during the polymorph conversion thus results in an additional source of variation. Therefore, it is necessary to correct the spectral data for temperature influences before further conclusions can be drawn.

4.2. Resolution of the data measured at constant temperature

The spectra for P_1 and P_2 are shown in Fig. 2a. The concentration profiles (Fig. 4) are obtained from the score plot of the second derivative spectra (Fig. 1a). The scores 1 and 80 representing the first and last spectrum measured define the vector on which the scores of the spectra 1–80 are projected. Plotting the distances between the score of spectrum 80 and the orthogonal projections of the scores 1–80 in function of time results in the concentration profile for P_1. The profile for P_2 is obtained by plotting the distances between score 1 and the orthogonal projections of the scores 1–80. The distances for P_1 and P_2 are calculated according to the Eq. (7) and Eq. (8), respectively.

$$\text{distance}_i = \frac{((\boldsymbol{s}_i - \boldsymbol{s}_{80})^T \cdot (\boldsymbol{s}_1 - \boldsymbol{s}_{80}))}{\|\boldsymbol{s}_1 - \boldsymbol{s}_{80}\|} \quad i = 1, 2, \ldots, 80 \tag{7}$$

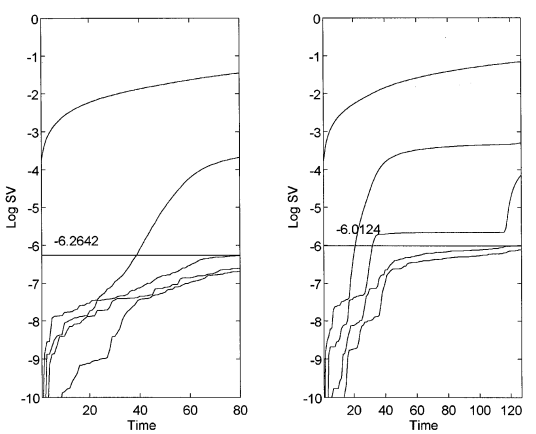

Fig. 3. (a) EFA graph obtained for forward calculation, for the second derivative matrix measured at constant temperature. (b) EFA graph obtained for forward calculation, for the second derivative matrix measured with varying temperature.

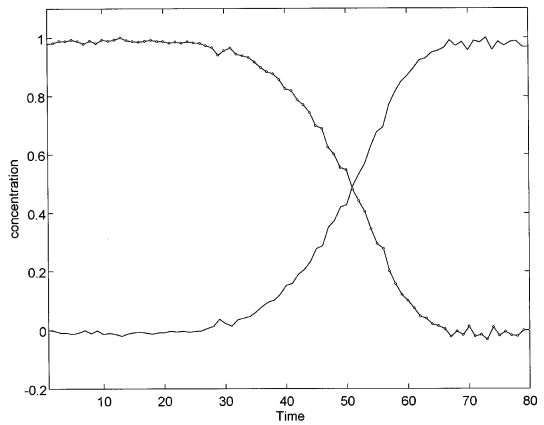

Fig. 4. Concentration profiles, normalised to maximum intensity, for P_1 (..) and P_2 (–) obtained from the score plot of the second derivative spectra.

$$\text{distance}_i = \frac{((\boldsymbol{s}_i - \boldsymbol{s}_1)^T \cdot (\boldsymbol{s}_{80} - \boldsymbol{s}_1))}{\|\boldsymbol{s}_{80} - \boldsymbol{s}_1\|} \quad i = 1, 2, \ldots, 80 \tag{8}$$

where $\boldsymbol{s}_1$, $\boldsymbol{s}_i$ and $\boldsymbol{s}_{80}$ are the vectors of the spectra 1, i and 80 in the two dimensional latent variable space defined by PC_1 and PC_2.

Applying the SIMPLISMA approach on the positive part of the inverted second derivative spectra results in the selection of the wavelengths 1644 and 2126 nm (The offset used is 1% of the maximum mean value). The selected wavelengths are situated in the spectral regions where the minor absorption bands in the original spectra of P_1 and P_2 were observed. The square root of the relative sum of squares of the residuals (rssq) between the total intensity spectrum and its least squares approximation calculated with the selected wavelengths 1644 and 2126 nm is 0.1915. Due to the selection of a third wavelength 2196 nm that does not reduce the rssq significantly (rssq = 0.1691), we can confirm that only two compounds are present. The significance of the change in the rssq values is determined by visual inspection of these values. The SIMPLISMA approach is an interactive process, which means that the selection of pure variables depends on the judgement of the spectroscopist. The spectra are obtained using Eq. (2). The matrix **C** contains the selected pure variable intensities from the positive

part of the inverted second derivative spectra in its columns. The matrix **X** contains the original spectra. Using the resulting pure spectra in combination with the original data matrix **X**, we can recalculate the concentration profiles using Eq. (3). The pure spectra are very well comparable with the ones shown in Fig. 2a. The concentration profiles, shown in Fig. 5, are not comparable with the profiles of Fig. 4. In the concentration profiles obtained with SIMPLISMA negative concentration values are observed for P_2. This artefact shows that the resolution with SIMPLISMA applied on the positive part of the inverted second derivative spectra do not give good results.

OPA was also used on the second derivative spectra. The results are shown in the Fig. 6a–c. The dissimilarity of each spectrum with respect to the mean spectrum is plotted in Fig. 6a. The first spectrum selected is the one at time 73. Each spectrum is then compared with the spectrum at time 73 and the dissimilarity is plotted as a function of time in Fig. 6b. The spectrum at time 13 has the highest dissimilarity value and because of this, it is the second spectrum selected. Each spectrum's dissimilarity with respect to the already selected spectra (times 73 and 13) is determined and plotted in Fig. 6c. This plot shows the highest dissimilarity value at time 49. Due to the much lower dissimilarity values in this plot, compared with the ones obtained in the plot before, spectrum 49 does not represent a third component. The spectra and concentration profiles are shown in the Fig. 7a,b respectively. The differences between the spectra are situated in the same spectral regions, 1600–1700 and 2100–2200 nm, as in the spectra shown in Fig. 2a. The concentration profiles are comparable with those of Fig. 4.

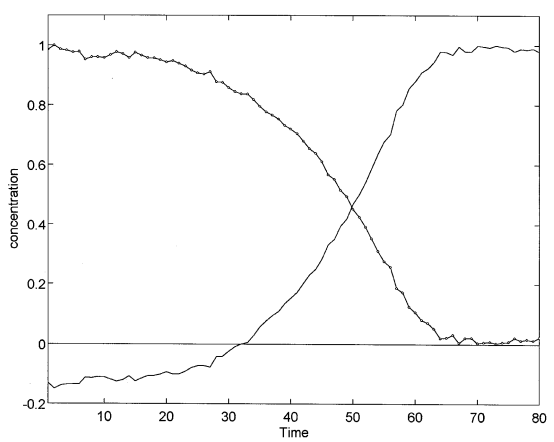

Fig. 5. Normalised concentration profiles for P_1 (..) and P_2 (–) resulting from the resolution (SIMPLISMA) of the data measured at constant temperature.

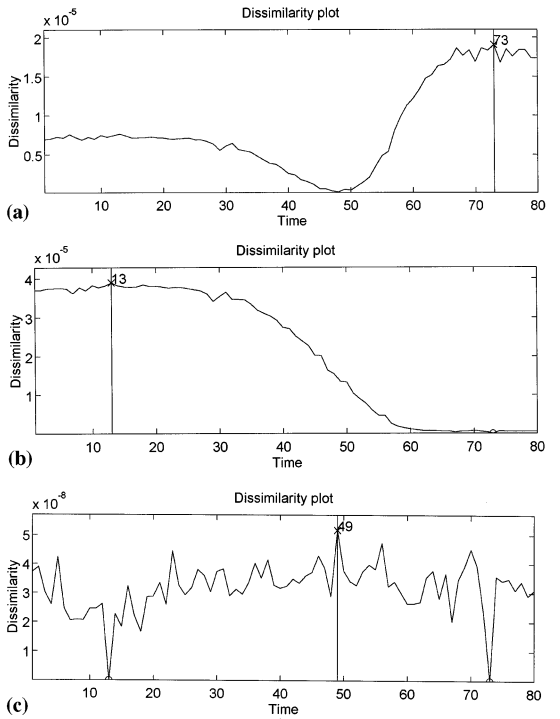

Fig. 6. Results of OPA applied on the second derivative spectra measured at constant temperature.

4.3. Temperature correction

The spectral data matrix measured with variation in temperature, $\mathbf{X}_{\text{measured}}$, can be represented as the sum of a bilinear data matrix, $\mathbf{X}_{\text{corrected}}$, containing the spectra at a constant temperature and a matrix, **T**, containing the temperature influences for each spectrum. The matrix can thus be expressed by the following equation:

$$\mathbf{X}_{\text{measured}} = \mathbf{X}_{\text{corrected}} + \mathbf{T} \quad (9)$$

As discussed in the previous paragraph the chemical rank of the matrix $\mathbf{X}_{corrected}$ is two. The rank of $\mathbf{X}_{measured}$ is three due to thermal effects. The temperature effects will therefore be modelled by a matrix of rank one and substracted from the measured data to obtain the matrix which contains the chemical information at a constant reference temperature.

The temperature correction is performed to obtain the spectra in the matrix $\mathbf{X}_{corrected}$ at a constant reference temperature of 25°C and is based on the following assumptions:

1. The polymorphs P_1 and P_2 are assumed to have similar molecular structures. Due to this their response to temperature change is supposed to be equal.
2. Spectrum 195 represents 100% P_2 measured at 60°C.
3. The thermal effects are linear with temperature.

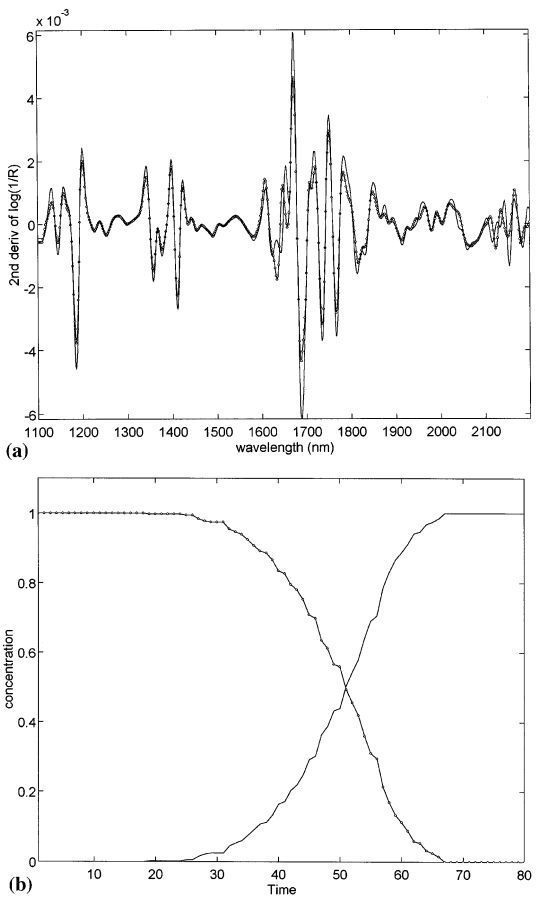

Fig. 7. (a) Pure second derivative spectra for P_1 (..) and P_2 (–) resulting from the application of OPA on the data measured at constant temperature. (b) Concentration profiles for P_1 (..) and P_2 (–) resulting from the resolution (OPA) of the data measured at constant temperature.

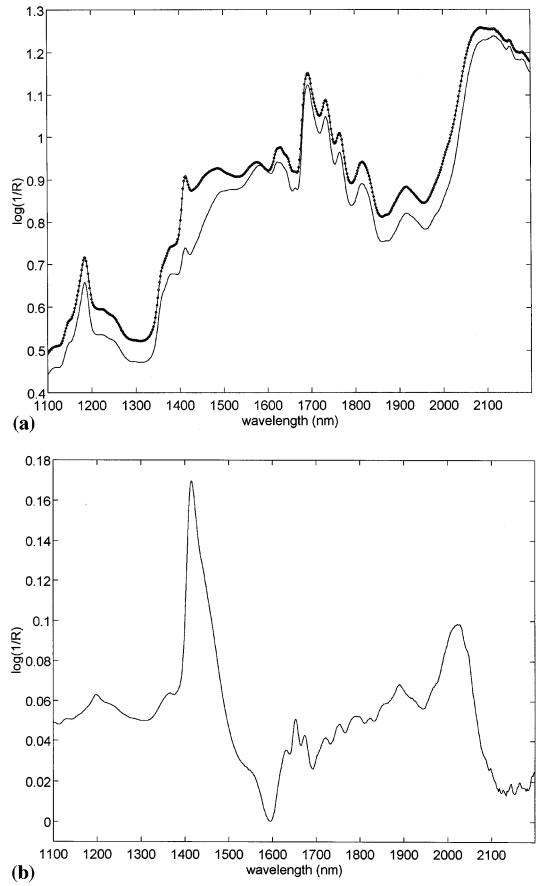

Fig. 8. (a) Spectra of P_2 at 60°C (..) and 25°C (–). (b) Temperature sensitivity spectrum (*ts*).

The temperature sensitivity spectrum (*ts*), shown in Fig. 8b, is the influence on the spectra of both polymorphs when the temperature changes from the highest temperature measured during the conversion to the reference temperature. Taken into account the first two assumptions, it was obtained by calculating the differences in absorbance values between the spectra 195 and 207 (Fig. 8a). Spectrum 195 represents P_2 at the

highest temperature (60°C) and spectrum 207 is P_2 at the reference temperature (25°C).

Because the assumption was made that the thermal effects change linearly with temperature [14], the spectra are corrected according to the following equation:

$$\mathbf{X}_{\text{corrected}} = \mathbf{X}_{\text{measured}} - (((t - 25°\text{C})/(60°\text{C} - 25°\text{C}))'.ts) \quad (10)$$

in which t is an array of dimension (127×1) containing the temperatures for each spectrum.

The score plot obtained after second derivation of the temperature corrected data is shown in Fig. 9. This plot now reveals two clusters 81–93 and 114–207 corresponding with P_1 and P_2, respectively. Furthermore the forward EFA graph of the temperature corrected second derivative data in Fig. 10 shows only two significant SV's instead of three, which means that there are two substances present. These observations permit us to conclude that the temperature correction is succesful.

4.4. Resolution of temperature corrected data

The raw mean spectra, normalised to maximum intensity, for the clusters observed in the score plot of the second derivative spectra (Fig. 9) are comparable with those in Fig. 2a. The concentration profiles in Fig. 11 are obtained from the score plot of the second derivative spectra (Fig. 9), using Eqs. (7) and (8).

The SIMPLISMA approach (offset 1%) used on the positive part of the inverted second derivative spectra results in the selection of the same wavelengths, 2126 and 1644 nm, as for the data measured at constant temperature. The spectra are comparable with those in Fig. 2a. In the concentration profiles (Fig. 12) negative concentration values are again observed for P_2. Furthermore it seems that the values for the zero component region of P_1 are estimated to be high.

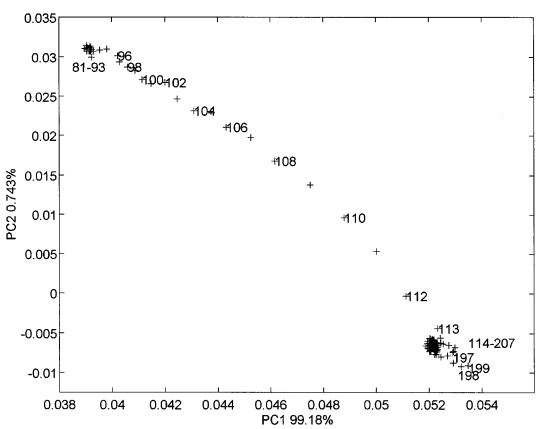

Fig. 9. Score plot after second derivation of the temperature corrected data.

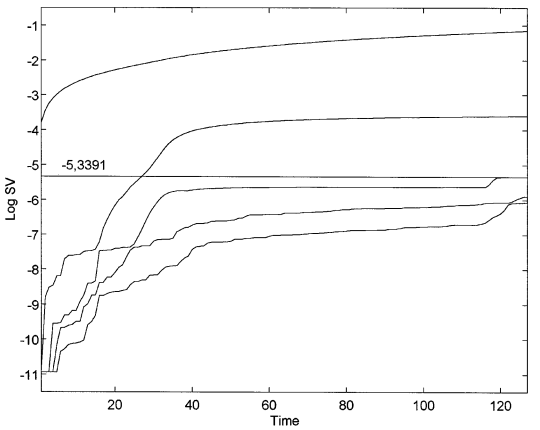

Fig. 10. EFA graph obtained for forward calculation, for the temperature corrected second derivative matrix.

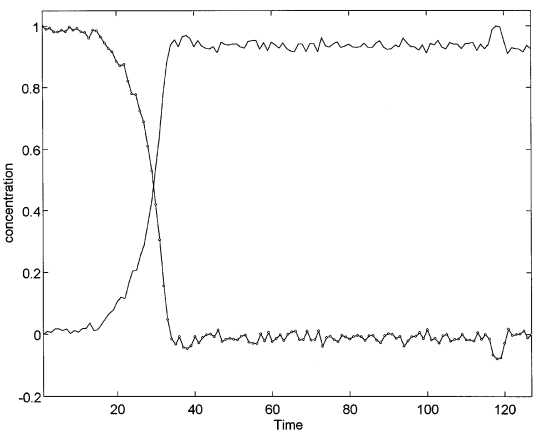

Fig. 11. Concentration profiles, normalised to maximum intensity, for P_1 (..) and P_2 (–) obtained from the score plot of the second derivative spectra.

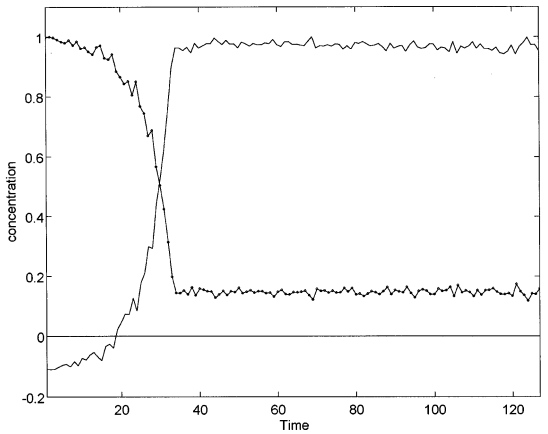

Fig. 12. Concentration profiles for P_1 (..) and P_2 (–) resulting from the SIMPLISMA approach applied on the temperature corrected data.

These artefacts can be the result of the fact that the selected wavelength 2126 nm is not completely pure or because the negative part of the inverted second derivative spectra was not taken into account.

The results of OPA applied on the second derivative spectra are shown in the Fig. 13a,c. Two spectra were selected: the first one at time 1 and the second one at time 118. The second derivative spectra obtained after resolution of the data matrix are comparable with the ones shown in Fig. 7a. The differences between the spectra are situated in the same spectral regions. The concentration profiles, shown in Fig. 14, are comparable with the profiles of Fig. 11.

5. Conclusion

The pure spectra and concentration profiles of the two polymorph substances can be obtained. The temperature correction proposed in this paper can be used when the spectra of the compounds present in the mixture are very similar. The resolution in this case can be performed by applying PCA and OPA (in the time direction). The SIMPLISMA approach performed in the wavelength direction on the positive part of the inverted second derivative spectra does not appear

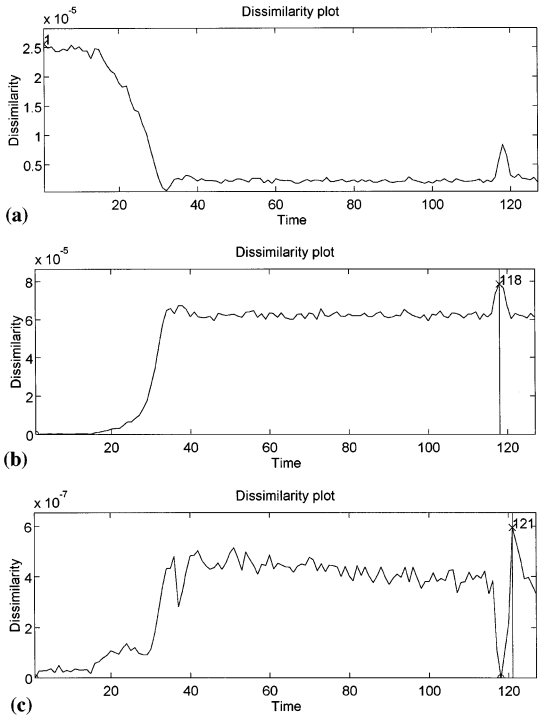

Fig. 13. Results of OPA applied on the temperature corrected second derivation spectra.

to provide good results. Alternative SIMPLISMA procedures are needed when second derivative data are applied.

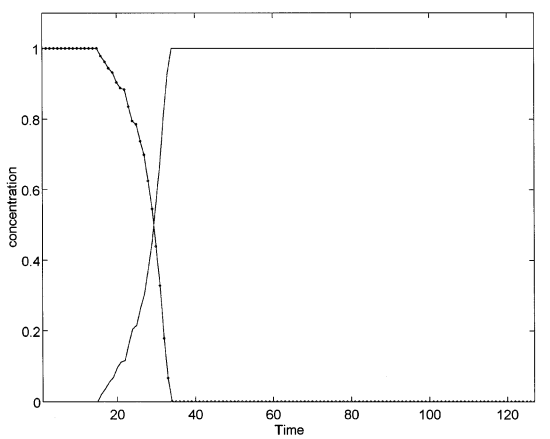

Fig. 14. Concentration profiles for P_1 (..) and P_2 (–) resulting from the resolution (OPA) of the temperature corrected data.

References

[1] H.R. Keller, D.L. Massart, Chemometr. Intell. Lab. Syst. 12 (1992) 209–224.
[2] H.R. Keller, D.L. Massart, Anal. Chim. Acta 246 (1991) 379–390.
[3] F.C. Sánchez, J. Toft, D. Kvalheim, D.L. Massart, Anal. Chim. Acta 314 (1995) 131–139.
[4] O.M. Kvalheim, Yi-Zeng Liang, Anal. Chem. 64 (1992) 936–946.
[5] Yi-Zeng Liang, O.M. Kvalheim, J. Chemometr. 7 (1993) 15–43.
[6] Yi-Zeng Liang, O.M. Kvalheim, H.R. Keller, D.L. Massart, P. Kiechle, F. Erni, Anal. Chem. 64 (1992) 946–953.
[7] B.G.M. Vandegiste, W. Derks, G. Kateman, Anal. Chim. Acta 173 (1985) 253–264.
[8] F.C. Sánchez, D.L. Massart, Anal. Chim. Acta 298 (1994) 331–339.
[9] W. Windig, D.A. Stephenson, Anal. Chem. 64 (1992) 2735–2742.
[10] W. Windig, J. Guilment, Anal. Chem. 63 (1991) 1425–1432.
[11] F.C. Sánchez, J. Toft, B. van den Bogaert, D.L. Massart, Anal. Chem. 68 (1996) 79–85.
[12] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Appl. Spectrosc. 43 (5) (1989).
[13] P.A. Gorry, Anal. Chem. 62 (1990) 570–573.
[14] R. Bourne, C. Burgess, Analyst 120 (1995) 2075–2080.