

Application of Multivariate Curve Resolution in Pharmaceutical Process Understanding

Dongsheng Bu
Martin Kermit

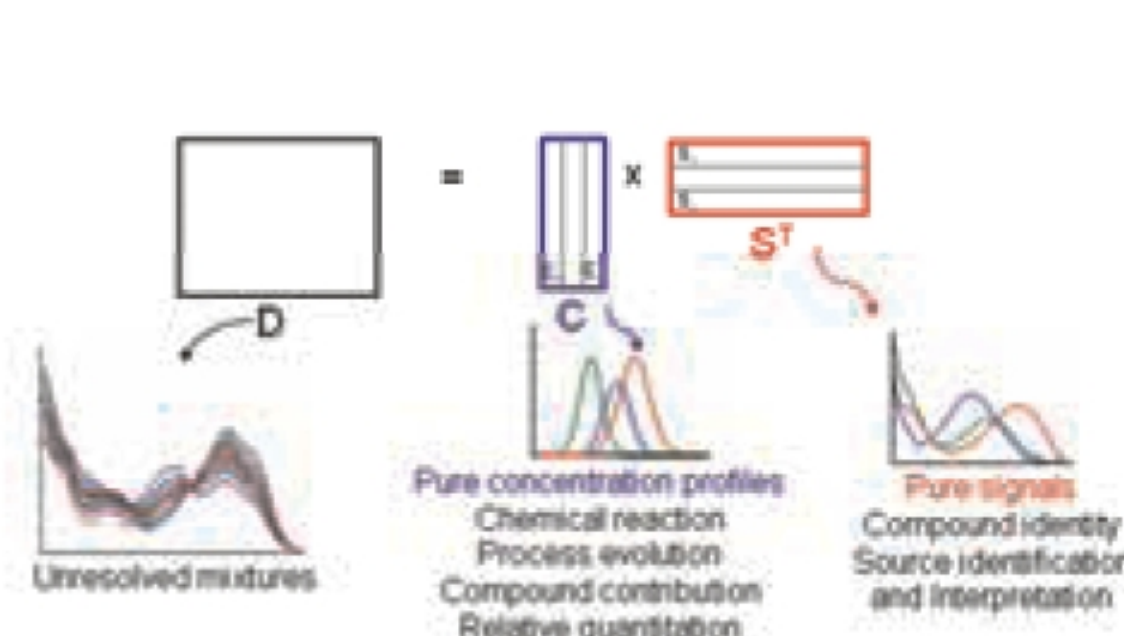
CAMO Software

ABSTRACT

We are currently applying Multivariate Curve Resolution (MCR) to understand and monitor manufacturing processes. One of the major applications is real-time detection of reaction end-point by FT-IR/Raman spectroscopy. In this case of an active ingredient synthesis, reference sampling was difficult, and MCR was applied to obtain relative concentrations of a major reagent. The estimated concentrations were then post-processed by Relative Standard Deviation (RSD) to detect a steady state of the synthesis in real-time. The application of MCR in wavelength selection for building quantitative model is discussed. Principal Component Analysis (PCA) score plots of different batches are useful in tracing the synthesis process and to study batch-to-batch variations.

INTRODUCTION

MCR methods may be defined as a group of techniques [1,2,3] which intend the recovery of concentrations and response profiles of the components in unresolved mixtures. It has been increasingly used in chemical reaction monitoring by FT-IR/Raman, particularly while reference sampling is difficult or there are possible intermediates during the reaction process. End-point detection of a synthesis is one of the most common applications in Process Analytical Technology (PAT). This paper is to explore the application of MCR in the understanding of an active pharmaceutical ingredient (API) synthesis process.



Materials and Methods

FT-IR data of three batches was provided by a pharmaceutical company. Each batch contains about 40 spectra. The spectral measurements as shown in Fig 1(a) were acquired in real-time at interval of 60 seconds. The synthesis is under low temperature condition, thus sampling for reference of reagent concentrations is difficult during the process. There are about 5-6 reference measurements of reagent concentration for each batch, which are shown as blue dots in Figure 1(c). The objective is to detect end-point which is defined as 1% of the reagent concentration.

Unscrambler-Online (version 2.0.0.3 CAMO Software AS, Oslo, Norway) was used for real-time prediction and PCA projection. A house-made algorithm for RSD computation was done in Matlab R2007a (Mathworks, Natick, MA). The Unscrambler® (Version 9.7, CAMO Software AS, Oslo, Norway) was used for all other data analysis and computation.

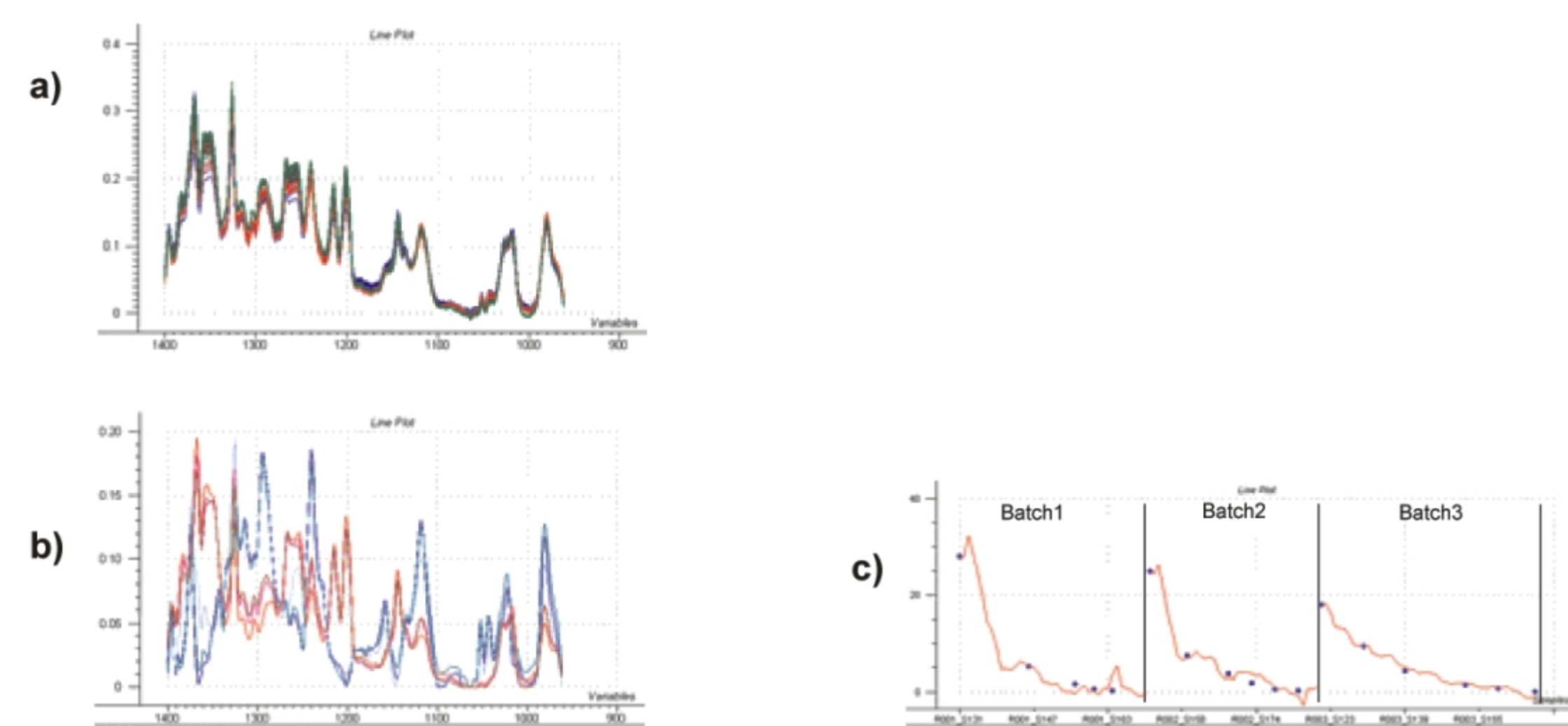


Fig 1. FT-IR spectra and their MCR results: a) All spectra of three batches that are plotted and color grouped by individual batch; b) Estimated spectra of 2 pure components that MCR was applied to individual batch; and c) Estimated concentrations (scaled) of pure component 1 over sample number, and available references in blue dots.

Results and Discussion

1. End-point detection by MCR

An approach similar to leave-one-out cross validation was used in MCR model building and applying. Estimated spectra of pure components were calculated by MCR with spectral data of 2 batches, and those spectra were used as Initial Guess for the concentrations computation of the left-out batch. The MCR calculated (relative in scale) concentrations were subjected to relative standard deviation computing to fit real-time end-point detection purpose. As shown in Figure 2, end-points were detected successfully in all 3 batches if an end-point is identified at 4th point of consequent %RSD values equal or smaller than 2.

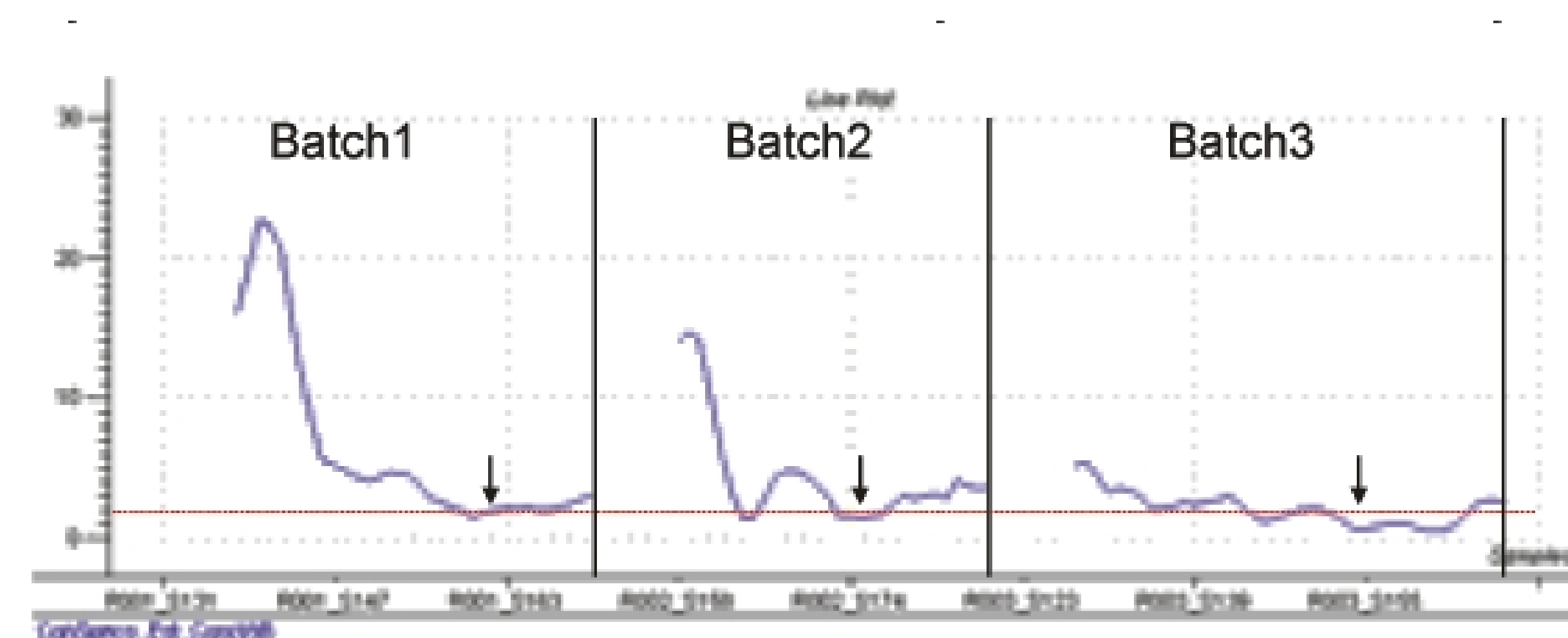


Fig 2. %RSD of MCR estimated concentrations over sample number (time) of all 3 batches. Leave-one-batch-out approach was used. Arrows indicate individual batches' end-point, which is defined as 4th point of consequent RSD values <= 2% (red dashed line: 2 %RSD).

2. End-point detection by Partial Least Square (PLS) regression and prediction

As a straight-forward approach, PLS directly predicts the concentrations of targeted reagent. MCR estimated spectra of pure components were used to select wavelength regions in order to reduce interference from other components. Figure 3 (a) shows the regression coefficients of a PLS1 model with MCR selected wavelengths. The model's RMSECV (not showed) decreases to 0.569 from 0.886 of a full region model. Such improvement is critical for this application. As shown in Figure 3 (b), end-points were detected successfully in all 3 batches by predicting 1% of a target reagent concentration.

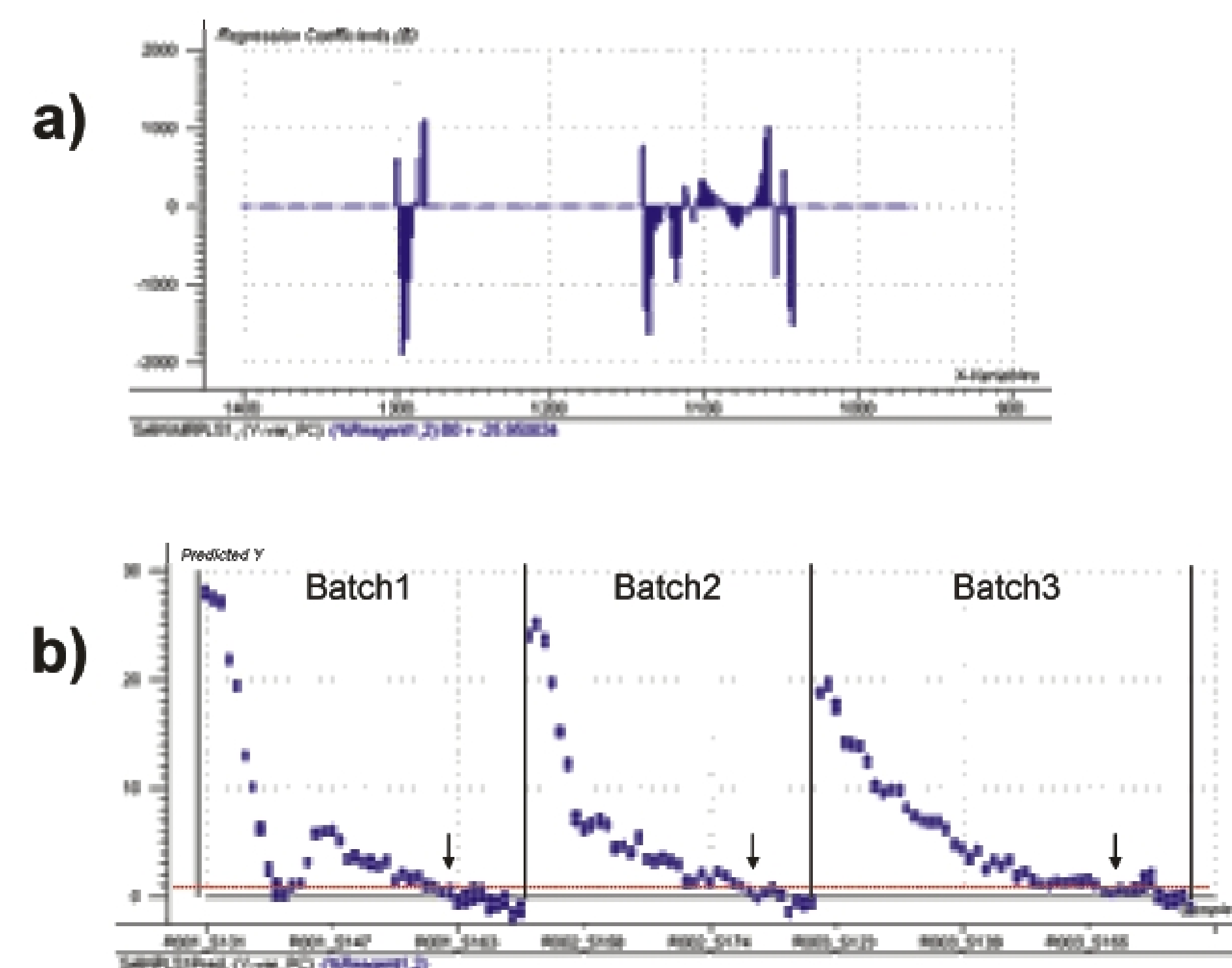


Fig 3. Regression model and prediction results of the targeted reagent: a) Regression coefficients of the PLS1 model that utilizes wavelengths selected by MCR, includes 17 samples of available references in the model and 2nd Savitzky-Golay derivative data preprocessing; b) Prediction of targeted reagent. Arrows indicate 4th point of consequent concentration <= 1% (red dashed line: 1 %).

3. End-point detection by PCA projection method

PCA based projection is the method that PCA scores of a batch can be obtained by using a PCA model built on historical data. Scores over time of all batches are shown in Figure 4 by the use of a PCA model built on Batch1 and 2. Projected scores were subjected to relative standard deviation computing to fit real-time endpoint detection.

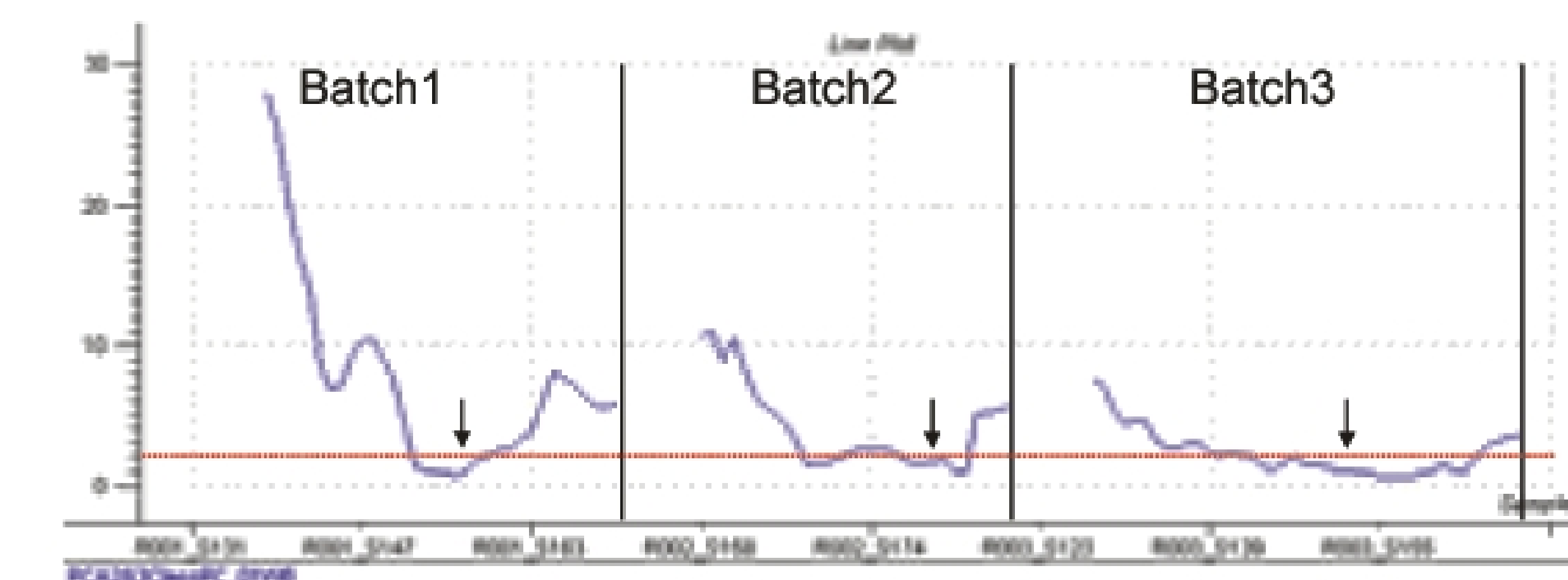


Fig 4. %RSD of projected score over time. PCA model was built on batch 1 and 2 data. Arrows indicate individual batches' end-point, which is defined as 4th point of consequent RSD values <= 2% (red dashed line: 2 %RSD).

Conclusions

- ▶ End-point of an API synthesis process was detected successfully by all three approaches, i.e. MCR estimated concentrations, PLS prediction, and PCA projected scores.
- ▶ MCR provides unique batch analysis tools for FT-IR/Raman process data while reference sampling is difficult.
- ▶ MCR provides practical wavelength selection tool while component interaction or intermediate is possible.
- ▶ Relative Standard Deviation is needed for post-processing MCR relative concentration or PCA scores.
- ▶ Real-time process monitoring needs appropriate Multivariate Data Analysis software tool.

References

1. R. Tauler and B. Kowalski, Anal. Chem. 65, 2040 (1993)
2. W. Windig and J. Guilment, Anal. Chem. 63, 1425 (1991)
3. D. Bu, C. Brown, Appl. Spectrosc. 54, 1214 (2000)

