Article

Multivariate Curve Resolution and Carbon Balance Constraint to Unravel FTIR Spectra from Fed-Batch Fermentation Samples

Dennis Vier [1],*, Stefan Wambach [2], Volker Schünemann [1] and Klaus-Uwe Gollmer [3]

[1] AG Biophysik und Medizinische Physik, Technische Universität Kaiserslautern, Kaiserslautern 67663, Germany; schuene@physik.uni-kl.de
[2] Fachbereich Bioverfahrenstechnik, Hochschule Trier, Umwelt-Campus Birkenfeld, Birkenfeld 55761, Germany; s11a81@umwelt-campus.de
[3] Fachbereich Angewandte Informatik, Hochschule Trier, Umwelt-Campus Birkenfeld, Birkenfeld 55761, Germany; k.gollmer@umwelt-campus.de
* Correspondence: dennisvier@gmail.com; Tel.: +49-170-2056148

Abstract: The current work investigates the capability of a tailored multivariate curve resolution–alternating least squares (MCR-ALS) algorithm to analyse glucose, phosphate, ammonium and acetate dynamics simultaneously in an *E. coli BL21* fed-batch fermentation. The high-cell-density (HCDC) process is monitored by ex situ online attenuated total reflection (ATR) Fourier transform infrared (FTIR) spectroscopy and several in situ online process sensors. This approach efficiently utilises automatically generated process data to reduce the time and cost consuming reference measurement effort for multivariate calibration. To determine metabolite concentrations with accuracies between ± 0.19 and $\pm 0.96 \cdot gL^{-1}$, the presented utilisation needs primarily—besides online sensor measurements—single FTIR measurements for each of the components of interest. The ambiguities in alternating least squares solutions for concentration estimation are reduced by the insertion of analytical process knowledge primarily in the form of elementary carbon mass balances. Thus, in this way, the established idea of mass balance constraints in MCR combines with the consistency check of measured data by carbon balances, as commonly applied in bioprocess engineering. The constraints are calculated based on online process data and theoretical assumptions. This increased calculation effort is able to replace, to a large extent, the need for manually conducted quantitative chemical analysis, leads to good estimations of concentration profiles and a better process understanding.

Keywords: multivariate curve resolution; *E. coli*; fed-batch; fermentation; carbon mass balance constraint; soft constraints; alternating least squares; hybrid modelling

1. Introduction

Multivariate curve resolution (MCR) with constrained alternating least squares (ALS), as described by Tauler et al. [1], is a powerful method to deconvolve overlapping spectral signals from chemical and biological reaction systems. The intended purpose is commonly the estimation of concentrations of individual components **C** or the identification of unknown spectral profiles **S** in complex aqueous solutions; generally, MCR has the ability to estimate both simultaneously from a data matrix **X**. The specific feature of MCR is the decomposition of **X** in a physically or chemically meaningful way. Besides MCR and ALS, other bilinear modelling methods and different algorithms can be utilised for the decomposition of **X**, with various resolution performances and limitations [2]. Because of its flexibility and popularity [3], MCR with the ALS algorithm is used in this work. A tutorial for the

application of MCR to analyse multicomponent systems, with special focus on the ALS algorithm, is given in [4]. The main aspects to be considered—such as data set configurations, initial estimates and applicable constraints—are described. A central issue of bilinear decomposition is the impact which particular constraints, initial estimates and the applied algorithm may have on the uniqueness of solutions in the presence of rotational ambiguities [5]. In the present work, the initial MCR settings and constraints for the analysis of FTIR-spectra from an *E. coli* fed-batch bioprocess are described in detail.

Compared with other established chemometric analysis methods, MCR has the potential of simultaneous resolution and quantitation of all mixture components without their chemical or physical separation [5,6]. Besides the recovery of qualitative and quantitative information about analytes, the identification of unknown interferents is possible [7]. In comparison to other multivariate calibration methods, the calibration effort of MCR can be decreased significantly by application of appropriate constraints [8]. Given a suitable set of constraints, this paper demonstrates that single measurements of pure solved analytes suffice to perform quantitative MCR analysis of respective fermentation data.

MCR-ALS has been employed for years in different research fields, especially in chemical reaction processes monitored by different spectroscopic techniques such as X-ray absorption [9], fluorescence, nuclear magnetic resonance, Raman, Near-Infrared and FTIR [10]. In biochemical and biophysical processes, MCR was used to analyse protein and nucleic acids systems concerning denaturation processes, protonation equilibria or complexation processes [11]. Among other chemometric methods, MCR was utilised in single biological cell analysis to unmix information from hyperspectral images [12]. In reference to fermentation processes, several applications of MCR-ALS have been published, e.g., in monitoring alcoholic fermentations with *S. cerevisiae* [13,14], milk lactic acid fermentations with *Streptococcus* and *Lactobacillus* strains [15] and the quantification of penicillin V in bioprocesses with *Pencillium chrysogenum* [16]. In this work, MCR-ALS with tailored constraints is applied to estimate metabolism-relevant concentrations in high cell density cultivation (HCDC) fed-batch processes with *E. coli* BL21 (DE3) pET28a. The cultivated organism produces the recombinant and pharmaceutically utilisable enzyme cytochrome p450 after induction. For evaluating process kinetics and optimising the growth of microorganisms, it is useful to obtain estimations about the quantitative changes of carbon, nitrogen and phosphate sources as well as of metabolic products such as acetic acid in the fermentation broth over process runtime. The aim of this study is the resolution of these substances by a tailored MCR-ALS algorithm.

To monitor the composition of fermentation media, ATR-FTIR spectroscopy is employed as an in-line ex situ analyser. Spectral information from the fermentation process is provided online by automatic cell-free sampling of fermentation broth through an ATR flow-cell. Because of the continuous sterile sampling, the constraint of invariance of the total concentration (closure) is applicable, as described below. An automated in-line flow system can cause the problem of CO_2 and air bubbles, as well as biofilms on the ATR surface [17]. Therefore, biofilms are inhibited by employing cell-free sampling and initial ethanol-cleaning of the flow-system. As technical gas bubble prevention, just gas-tight polytetrafluoroethylene (PTFE) tubes are implemented. However, principally, the problem of gas bubbles can be handled mathematically by the MCR algorithm as shown in this study.

FTIR spectroscopy is an established technique in bioprocess monitoring and, as other IR techniques such as near-infrared (NIR) and Raman, it combines the advantages of non-invasiveness and fast simultaneous measurement of multiple solved substances [18]. An overview of advantages and disadvantages of different spectroscopy techniques in bioreactor monitoring has been described [19]. In the mid-infrared region, covered by FTIR, most excitations of fundamental molecular vibrations can be found. Especially the fingerprint area (1500–500 cm^{-1}) exhibits specific patterns of media compounds [20]. By contrast, the peaks in the NIR spectrum consist of overtones and combinations from primary MIR signals and are less distinctive. In comparison to NIR, the MIR region exhibits a

higher selectivity thus allowing for a better detection of overlapping component spectra in complex aqueous mixtures. Raman spectroscopy is not sensitive to water and the small peak widths of solved components are main advantages of this technique in bioprocess monitoring [21]. However, the Raman scattering is generally weaker than the FTIR signal, while higher concentrations of target analytes are required. A comparison of FTIR, NIR and FT-Raman spectroscopic techniques referring to a lactic acid fermentation shows the best prediction performance for FTIR [22].

Some studies described the analysis of glucose, acetate, ammonium and phosphate concentrations in bioprocesses, using ATR-FTIR spectroscopy. Among other substances, ammonium and glucose are analysed in a complex antibiotic fermentation by at-line measurements on a horizontal attenuated total reflectance (HATR) crystal using a partial least squares (PLS) calibration model [23]. Gluconacetobacter xylinus fed-batch cultures were monitored by an in situ ATR probe aimed at the online PLS analysis of acetate, phosphate and ammonium [24]. The results of these PLS predictions—in contrast to MCR-ALS predictions of glucose, acetate, ammonium and phosphate—are discussed below.

As shown by references, FTIR monitoring of bioprocesses and MCR analysis of complex mixtures such as fermentation broths promise many advantages in simultaneous process information collection. The proposed effective usage of in situ and ex situ online sensor data to calculate carbon mass balance-constrained MCR predictions of several analytes underlines the relevance of ATR-FTIR/MCR-ALS combinations in fermentation analysis.

The monitored substances are related to bacterial metabolism. Glucose, ammonia and phosphate are substrates whereas acetate is a by-product of overflow metabolism [25]. The recombinant cytochrome p450 remains unconsidered for its being an intracellular metabolite and therefore not being obtainable in fermentation broth. To predict concentrations of the observed analytes, only four calibration measurements of pure components are required, provided adequate constraints are applied during the alternating least squares procedure. In addition to the required analyte spectra, the implementation of estimated artefact spectra and additionally known fermentation media components is useful. As mentioned above, collecting samples from fermentation broth by a peristaltic pump through a tube system can present the problem of air bubbles in ATR flow-cell with impact on the measured spectra. If the water spectrum is removed from each mixture spectrum prior to that, the air bubble disturbance has its own spectral signature and can be handled like any pure component in multivariate curve resolution. The shape of this artefact is easy to be determined and its implementation in MCR improves the resolution of primary signals, as shown below. All initial estimations for pure spectral components **S** expected in the mixture are also implemented as soft-constraints during alternating least squares. In the following, soft-constraints means the presence of an allowed solution area in a range set by inequality constraints, as in the optimisation problem. In the case of physical spectral shifting in the mixture, a certain flexibility during iterative identification of pure components helps to avoid over-restriction.

In the concentration estimation step, besides the non-negativity constraint, online process data such as input and exhaust gas-flow, fermenter mass, liquid supply from feed reservoir and pH control as well as turbidity are utilized to calculate elementary carbon mass balance constraints. Mass balance constraints (closure) applied to reaction systems have been described [1,26,27]. Closure constraints require invariance of total concentration, granted by the sterile online sampling system and by including total reactor in- and output mass flow in the constraint calculation. So far, as known, the presented application of the carbon mass balance constraint for MCR to analyse a fed-batch fermentation process is a new utilisation of the popular closure constraint. The carbon mass balance constraint for a fed-batch fermentation requires extensive prior calculations, such as different conversion steps and soft sensor approaches. In addition, dynamic in- and output carbon mass-flow in gas and liquid phase as well as continuous and discrete sampling need to be taken into account. The presented algorithm is able to deal with these requirements. Calculating carbon mass balances or recovery rates is an established approach in bioprocess engineering to check the integrity of observed process data [28]. The referenced literature has already shown the application of carbon mass

balances for *Escherichia coli* high-cell-density fed-batch culture and recombinant protein production. During process runtime, carbon recovery rates should take on values of about 1. In the present study, this condition is utilised as a MCR-ALS constraint for the estimation of carbon sources and metabolites such as glucose and acetate in fermentation media. In so doing, the explorative decomposition of measured mixture spectra is coupled with analytical knowledge in order to form a new hybrid multivariate modelling approach.

In summary, the objectives of this paper are

(1)　the interpretation of the MCR-ALS closure constraint as a carbon mass balance constraint for fed-batch fermentation processes;

(2)　to demonstrate that moderate gas bubble disturbances on the ATR crystal can be handled computationally, without any need for technical preventions;

(3)　to show that MCR-ALS with carbon mass constraint is capable of simultaneously predicting four analyte concentrations from FTIR spectra of fermentation media samples, with minor calibration effort.

2. Material and Methods

2.1. Spectra Acquisition, Sampling and Spectra Processing

The MIR spectra are scanned with a Thermo Scientific Nicolet™ iS™ 10 and the extension unit Nicolet iZ™ 10. The Specac's Gateway™ ATR Accessory Kit and a ZnSe ATR crystal with six reflections are mounded as a flow cell in that unit. The flow cell is connected to the bioreactor via PTFE tubes (id 1.1 mm) and a Flownamics® FISP® probe with rapid flow membrane for cell-free sampling. A peristaltic pump, controlled by an Arduino microcontroller and a driver board, delivers the sample liquid continuously to the FTIR flow cell. A background spectrum with pure water in the flow cell is scanned before using the FTIR for bioprocess analysis. During a running sampling process, spectra are scanned in cycles of 10 min. During spectrum acquisition by the Thermo Scientific™ OMNIC™ software [29], the sampling pump remains inactive. The spectra acquisition time for scanning 32 spectra and releasing the mean spectrum for the current sample is about 1 min. OMNIC and microcontroller are both triggered by a C# program that observes and synchronises the sample supply and measurement steps. Before each start of a fermentation trial, tubing and flow cell are treated with 70% ethanol solution to minimise the risk of microbial activity in the sampling section. The initial spectra of known substances are standardised to unit concentration. No further pre-processing steps such as normalisation or differentiation are applied to the mixture-spectra in order to preserve the natural physical properties of the spectra. After the fermentation run, the MCR-ALS analysis of the ex situ online-monitored FTIR spectra is performed for all collected spectra.

2.2. Reference Analysis

To validate MCR-ALS results, reference values for glucose, acetate, ammonia and total phosphate concentrations are measured in the cell-free sample drain after passing the FTIR flow-cell. Glucose analysis was conducted by the YSI 2700 SELECT Biochemistry Analyzer (Yellow Springs, OH, USA). Acetate was determined by HPLC (high performance liquid chromatography) using chromatography column Reprogel H+ (Dr. Maisch GmbH, Ammerbuch, Germany). Total phosphate and ammonia were determined using photometric methods by procedures described in DIN EN 1189, DVGW W 504 and DIN 38406 E5.

2.3. Bioreactor System and Online Measurement Equipment

Fermentations are conducted in a prototype of Bioengineering's 5l rounded-bottom autoclavable laboratory fermenter (RALF), controlled and observed with the Software BioSCADA Lab (Bioengineering AG, Wald, Switzerland). A supply tower with intelligent front modules (IFM) directs

in- and output of control and measurement values. All data interchanged between IFMs and SCADA pass a structured query language (SQL) data base, the central data hub. From there, the needed data for calculating MCR constraints or advanced measurement and control strategies can be acquired by MySQL and MATLAB.

The current work utilises the following bioprocess online measurement instrumentation: Turbidity probe ASD19-N and optek-converter FC10 (optek-Danulat GmbH, Essen, Germany); exhaust gas analyser BlueInOne Ferm (BlueSens GmbH, Herten, Germany); thermal mass flow controller Red-Y Smart for inflow oxygen (0.01, ... , 5 lpm) and air/nitrogen (0.1, ... , 10 lpm) control (Vögtlin Instruments AG, Aesch, Switzerland); balances for online weight/volume observation of fermenter (DE 35K5D, Kern & Sohn GmbH, Balingen,Germany), acid/base (EW6000-1M, Kern & Sohn GmbH, Balingen, Germany) and feed (BL6100, Sartorius, Göttingen, Germany) reservoir.

2.4. Fermentation Strategy

The HCDC process is conducted in three phases: an initial batch phase, a feeding phase for biomass growth and an induction phase for product expression. The substrate and inductor feed is performed by exponential feeding strategy to control the cell specific growth rates µ similar to [30]. Because of the risk of overflow metabolism und protein folding errors at high growth rates, µ is controlled to defensive setpoints of 0.1 h^{-1} (feed phase/biomass production) and 0.05 h^{-1} (induction phase).

2.5. Strain and Fermentation Medium

E. coli BL21 (DE3) pET28a was stored as glycerol cryo-culture at -76 °C. The pre-culture is incubated as overnight culture in 500 mL baffled flasks at 37 °C in a shaker rotating 200 rpm. An amount of 300 mL pre-culture is portioned in equal shares on two shaking flasks. After 24 h pre-culture incubation, 2.7 L sterilised batch medium in the reactor is inoculated with the culture, thus amounting to a total start volume of 3 L.

The media are modified mineral media based on [31]. The pre-culture and batch medium contain per litre: Glucose*H_2O, 16.5 g; KH_2PO_4, 13.3 g; $(NH_4)_2HPO_4$, 4 g; citric acid, 1.7 g; $MgSO_4$*7H_2O, 0.72 g; Fe(II)SO_4*7 H_2O, 113.5 mg; $CoCl_2$*6H_2O, 10.5 mg; $MnCl_2$*4 H_2O, 15 mg; $CuCl_2$, 1.2 mg; H_3BO_3, 3 mg; Na2MoO4*2 H2O, 2.5 mg; thiamine*HCl, 4.5 mg; trisodium citrate dihydrate, 75 mg; Na_2-EDTA, 9.6 mg.

The feeding solution is composed of Glucose*H_2O, 544.4 g; $MgSO_4$*7H_2O, 12 g; Fe(II)SO_4*7H_2O, 43.3 mg; $CoCl_2$*6H_2O, 21.4 mg; $MnCl_2$*4H_2O, 23.5 mg; $CuCl_2$ 2.5 mg; H_3BO_3, 5 mg; Na_2MoO_4*2H_2O, 4 mg; trisodium citrate dihydrate, 116 mg; Na_2-EDTA, 14.8 mg.

3. Theory and Calculation

3.1. Nomenclature

Matrices: Uppercase fat letters
Vectors: Lowercase fat letters
Scalars: Lowercase letters

3.2. Multivariate Curve Resolution and Its Physical Interpretation

The bilinear model of multivariate curve resolution [1] for FTIR data can be deduced from the Lambert–Beer law which describes the attenuation of light travelling through material. The absorbance x of a material is given as

$$x = \lg\left(\frac{I_0}{I_1}\right) = c\ \varepsilon\ d$$

The logarithm of incident radiant intensity (I_0) divided by transmitted radiant intensity (I_1) is equal to the product of substance concentration (c), the molar attenuation coefficient (ε) and the pathlength (d). In this work, the technique of attenuated total reflection is used, so d is the penetration depth of an evanescent wave into the sample on the ATR crystal. The material and wavelength dependent factors ε and d can be pooled to s which consolidates the optical properties of a substance:

$$x = c\,s$$

For mixtures of several substances $k = 1, \dots, \Omega$, each absorbance value x_{ij} related to its wavelength in a spectrum $j = 1, \dots, n$ for a particular concentration profile $i = 1, \dots, m$ is calculated as

$$x_{ij} = \sum_{k=1}^{\Omega} c_{ik}\,\varepsilon_{kj}d_{kj} = \sum_{k=1}^{\Omega} c_{ik}s_{kj}$$

In chemometrics, it is usual to term $i = 1, \dots, m$ as the objects or samples of a dataset, whereby j counts the n features or variables. Here, the m objects are samples of fermentation broth over process runtime and the n features are absorbance values over the wavenumbers of FTIR spectra.

According to the previous sum equation, the decomposition of absorbance values over sample and wavenumber can be organised in matrices

$$
\begin{pmatrix}
x_{11} & x_{12} & \cdots & x_{1n} \\
x_{21} & x_{22} & \cdots & x_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
x_{m1} & x_{m2} & \cdots & x_{mn}
\end{pmatrix}
=
\begin{pmatrix}
c_{11} & c_{12} & \cdots & c_{1\Omega} \\
c_{21} & c_{22} & \cdots & c_{2\Omega} \\
\vdots & \vdots & \ddots & \vdots \\
c_{m1} & c_{m2} & \cdots & c_{m\Omega}
\end{pmatrix}
\begin{pmatrix}
s_{11} & s_{12} & \cdots & s_{1n} \\
s_{21} & s_{22} & \cdots & s_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
s_{\Omega 1} & s_{\Omega 2} & \cdots & s_{\Omega n}
\end{pmatrix}
$$

In matrix representation, we get the simplified description:

$$\mathbf{X} = \mathbf{C}\mathbf{S}^{\mathrm{T}}$$

That is the decomposition of absorbance spectra indicated by multivariate curve resolution assuming the data matrix $\mathbf{X}$ is bilinear.

3.3. An Implementation of the Alternating Least Squares Algorithm

With an initial estimation for concentration matrix $\hat{\mathbf{C}}_0$ or pure components $\hat{\mathbf{S}}_0$ and existing data $\mathbf{X}$, the ALS algorithm can run and perform multivariate curve resolution iteratively [1]. Assuming the chemical rank of the observed data matrix is estimated and one assumption per each expected spectral independent component is available, the ALS procedure can start with an initial pure component matrix. Thus, in the first iteration, the estimated unconstrained concentration matrix $\hat{\mathbf{C}}$ is obtained by

$$\hat{\mathbf{C}} = \mathbf{X}\hat{\mathbf{S}}_0\left(\hat{\mathbf{S}}_0^{\mathrm{T}}\hat{\mathbf{S}}_0\right)^{-1} = \mathbf{X}\left(\hat{\mathbf{S}}_0^{\mathrm{T}}\right)^{+}$$

whereby the superscripted + indicates the pseudoinverse.

In the next step, $\hat{\mathbf{S}}^{\mathrm{T}}$ is estimated in an unconstrained way by

$$\hat{\mathbf{S}}^{\mathrm{T}} = \left(\hat{\mathbf{C}}^{\mathrm{T}}\hat{\mathbf{C}}\right)^{-1}\hat{\mathbf{C}}^{\mathrm{T}}\mathbf{X} = \hat{\mathbf{C}}^{+}\mathbf{X}$$

With that pure component estimation, a new concentration matrix calculation can be performed. That loop is repeated until a termination criterion is achieved.

Because of rotational and intensity ambiguities, it is necessary to constrain the solutions for $\hat{\mathbf{C}}$ and $\hat{\mathbf{S}}^{\mathrm{T}}$ to obtain a physically meaningful separation of mixture components.

To calculate constrained linear least-squares solutions in this work, the `lsqlin` function with the *active-set* algorithm from MATLAB and the "Optimization Toolbox" is applied [32]. `lsqlin` makes use of mathematically rigorous methods of applying equality and inequality constraints with a better numerical stability than approximate methods commonly used in chemometrics. The approximate methods are easy to use and code, but they exhibit poor least squares behaviours and in some cases they result in an increase in the magnitude of residuals [33].

`lsqlin` solves linear least-squares curve fitting problems of the form

$$\min_{\mathbf{c}^T} \| \mathbf{x}^T - \mathbf{S}\,\mathbf{c}^T \|_2^2 \text{ such that } \begin{cases} \mathbf{A}\mathbf{c}^T \leq \mathbf{b} \\ \mathbf{A}_{eq}\mathbf{c}^T = \mathbf{b}_{eq} \\ \mathbf{l} \leq \mathbf{c}^T \leq \mathbf{u} \end{cases}$$

Hence, the MCR-ALS algorithm using `lsqlin` is implemented as shown in Figure 1 to solve the present problem of resolving **X** in a hybrid modelling way with the target of reducing ambiguities of least squares solutions. To bring a priori knowledge about pure spectra and the bioprocess into ALS solutions, linear inequality constraint vectors (e.g., **b**) and matrices (e.g., **A**) are applied. Further, the non-negativity constraint for concentrations is set by using lower bounds (**l**).
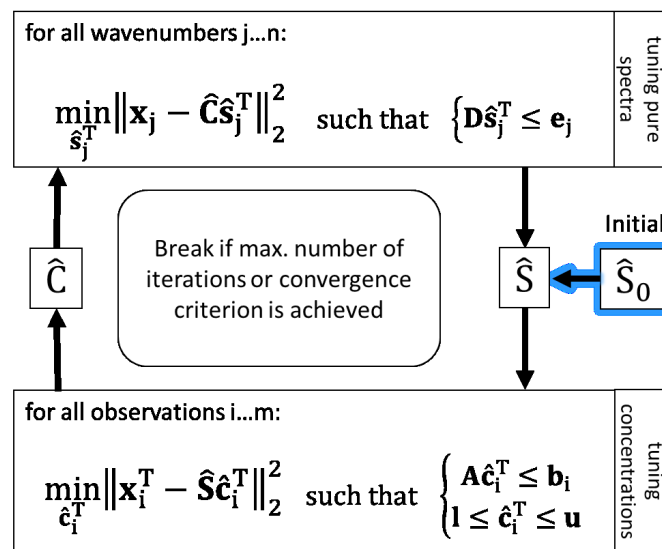
Figure 1. Scheme of the MCR-ALS algorithm with inequality constraints applying MATLAB `lsqlin` function.

3.4. Constraints for Pure Spectral Component Estimation

Pure spectra of components which are known and expected in mixture are constrained based on measured und normalised spectra of respective pure substances. Therefore, the same spectra used as initial estimations $\hat{\mathbf{S}}_0$ are also basis values of inequality constraints to calculate $\hat{\mathbf{S}}$. Assuming the shapes of pure spectra in mixture closely resemble the pure measured spectra, in each iteration the associated pure component estimations may only vary inside the defined ranges relative to the measured $\hat{\mathbf{S}}_0$. Depending on the amount of expected deviation in mixture, the range for the respective component can be adapted. In so doing, over-restriction can be avoided e.g., in the case of smaller rates of band shifting or in the case of differences in signal-to-noise ratios between high concentrated pure substance measurement and lower concentrations in mixture.

Regarding the inequation constraint for tuning $\hat{\mathbf{S}}$ in Figure 1 (upper box), **D** is composed vertically of the positive and negative identity matrices **I** and $-\mathbf{I}$ both of dimension (Ω, Ω).

$$\mathbf{D} = \begin{pmatrix} +\mathbf{I} \\ -\mathbf{I} \end{pmatrix}$$

The positive part is associated with the upper bounds $\mathbf{e}_j^u$, the negative with the lower bounds $\mathbf{e}_j^l$ represented in $\mathbf{e}_j$ for all components on each wavenumber. The allowed upper deviation $\mathbf{u}$ and lower deviation $\mathbf{l}$ are relative to the total ranges of the minimal and maximal values of the pure initial spectra for each component $(\hat{\mathbf{s}}_0)_{k=1,...,\Omega}$.

$$\mathbf{\Delta}_s = \begin{pmatrix} \max((\hat{\mathbf{s}}_0)_{k=1}) \\ \vdots \\ \max((\hat{\mathbf{s}}_0)_{k=\Omega}) \end{pmatrix} - \begin{pmatrix} \min((\hat{\mathbf{s}}_0)_{k=1}) \\ \vdots \\ \min((\hat{\mathbf{s}}_0)_{k=\Omega}) \end{pmatrix}$$

$$\mathbf{e}_j = \begin{pmatrix} +\mathbf{e}_j^u \\ -\mathbf{e}_j^l \end{pmatrix} = \begin{pmatrix} (\hat{\mathbf{s}}_0)_j^T + \mathbf{u} \circ \mathbf{\Delta_s} \\ (\hat{\mathbf{s}}_0)_j^T - \mathbf{l} \circ \mathbf{\Delta_s} \end{pmatrix}$$

In our application, the chemical rank of the mixtures $\mathbf{X}$ was estimated at 12 significant spectroscopically independent components by principal component analysis (PCA). The loadings of PCA were manually evaluated for the presence of spectra-like structure, which is strongly present on the first principal components and decreases on higher factors. Among the above mentioned, significant spectroscopically independent components were spectra of known media components, expected metabolic products, artefacts (like air bubbles) and unknown components. Only components evaluated as certainly present in the spectral mixture $\mathbf{X}$ are constrained, notably the pure spectra of wanted substances: glucose, ammonia, total phosphate ($H_2PO_4^- + HPO_4^{2-}$) and acetate. The estimations for those pure components may take on values at an interval of $\pm 10\%$ in the range of each pure component spectrum starting on the initial spectrum (see Table 1).

Table 1. Columns in $\hat{\mathbf{S}}_0$ with initial spectra and constraint settings.

	$(\hat{\mathbf{s}}_0)_1$	$(\hat{\mathbf{s}}_0)_2$	$(\hat{\mathbf{s}}_0)_3$	$(\hat{\mathbf{s}}_0)_4$	$(\hat{\mathbf{s}}_0)_5$	$(\hat{\mathbf{s}}_0)_6$	$(\hat{\mathbf{s}}_0)_7$	$(\hat{\mathbf{s}}_0)_8$	$(\hat{\mathbf{s}}_0)_9$	$(\hat{\mathbf{s}}_0)_{10}$	$(\hat{\mathbf{s}}_0)_{11}$	$(\hat{\mathbf{s}}_0)_{12}$
pure	glc	ace	NH_4^+	$H_2PO_4^-$ HPO_4^{2-}	IPTG	$MgSO_4$	air	H_2O	citric acid	unifrnd	unifrnd	unifrnd
u	0.1	0.1	0.1	0.1	0.1	0.1	inf	inf	0.1	inf	inf	inf
l	−0.1	−0.1	−0.1	−0.1	−0.1	−0.1	−inf	−inf	−0.1	−inf	−inf	−inf

Initial pure components: glucose (glc); acetate (ace); isopropyl β-D-1-thiogalactopyranoside (IPTG); air bubbles estimation (air); uniformly distributed random values (unifrnd); infinity (inf).

Because of water background subtraction on each taken spectrum $\mathbf{x}_i$, air bubbles in the flow cell have the shape of inverted water spectra. Moreover, a pure water spectrum is also initialised for the case of air bubble presence during the background recording. Both estimations are not constrained and can vary depending on the actual mixture content.

To demonstrate the validity of the assumption for the spectral air bubble model, a simple aqueous solution containing glucose ($15 \cdot gL^{-1}$), ammonium ($0.7\ gL^{-1}$) and phosphate ($8 \cdot gL^{-1}$) was compounded. From this solution, a first FTIR spectrum was acquired from a mixture covering the entire ATR crystal whereas a second spectrum resulted from the same mixture covering only about half of the crystal surface. In this way, a part of the IR beam reflections interacts with the aqueous solution on the ATR crystal, while another part interacts just with the air on the crystal surface. The latter liquid-free surface part simulates a large air bubble on the crystal in ATR flow cell. In the case without ALS iterations, the mixture matrix of the known solution is multiplied once with a simple pseudo inverse of the estimated initial pure components matrix. In a first measurement, S_0 contains just pure measurements of glucose, ammonium and phosphate, respectively standardised to unit concentration. Next, S_0 additionally contains an inverted water spectrum. The differences in concentration estimations are shown in Figure 2 whereas the actual concentration values are listed in

Table 2. Obviously, the integration of the air bubble model brings an improvement of the prediction results in the case of air bubble presence.
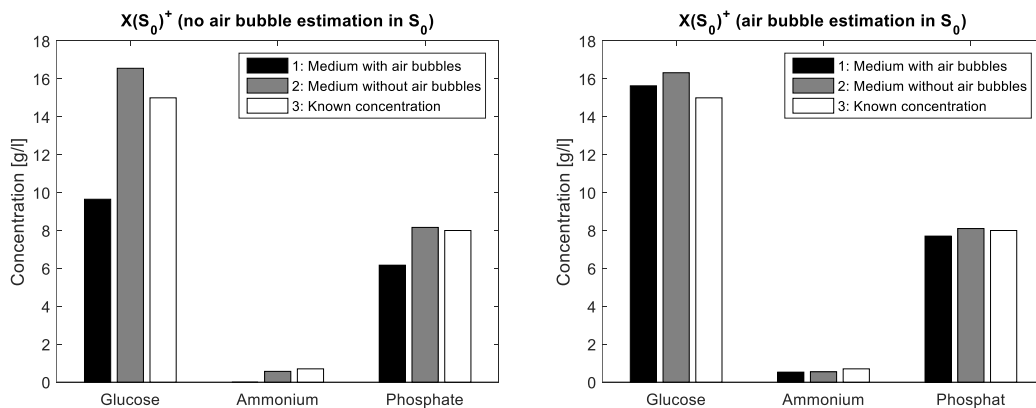

Figure 2. Proof of the initial estimation for the spectral air bubbles signature: in a known aqueous solution of glucose, ammonium and phosphate, air bubbles are present in one case and absent in another. On the left, the known spectral components S_0 do not consider the air bubble model. On the right, the air bubble model is integrated and the differences between concentrations estimated in medium with and without air bubbles are similarly close to the known concentrations.

Table 2. Concentration estimation by multiplying known mixture spectra with pseudo inverses of pure spectral component matrices with and without integration of an air bubble model.

	Air Bubble Model in S_0	Glucose $[gL^{-1}]$	Ammonium $[gL^{-1}]$	Phosphate $[gL^{-1}]$
Medium with air bubbles	No	9.65	0	6.17
	Yes	15.63	0.52	7.70
Medium without air bubbles	No	16.56	0.57	8.16
	Yes	16.32	0.55	8.10
Known concentration	-	15	0.7	8

3.5. Constraints for Concentration Estimation

In bioprocess engineering, the carbon balance and recovery rate of a fermentation process are commonly used as a check for the integrity of process monitoring and sensors as well as the assessment of the release of outer membrane components. Carbon balances in a fed-batch culture are based on the mass of carbon in the total fermenter volume. Thereby, the recovery rate is the relation between the recovered carbon $m^{C,rec}(t)$ and the carbon brought into the bioreactor $m^{C,in}(t)$ over process runtime t.

$$r^C(t) = \frac{m^{C,rec}(t)}{m^{C,in}(t)}$$

Suppose all carbon compounds are determinable and measurement errors are negligible, $r^C(t)$ is equal to 1 for all t. Because of the presence of measurement errors and not identified soluble organic carbon compounds, a tolerance range must be assumed. The carbon recovery considering biomass, CO_2, glucose and acetate is assumed as being about 90% [28].

The recovered carbon is the sum of carbon mass in the reactor liquid phase L, gas phase g, sample liquid phase divided in cell-free sampling scf and cell containing sampling scc.

$$m^{C,rec}(t) = m^{C,L}(t) + m^{C,g}(t) + m^{C,scc}(t) + m^{C,scf}(t)$$

The brought-in carbon is the sum of initial carbon mass in fermentation medium at process start time (t = 0) and the supplied carbon mass m^r from the feed reservoir r.

$$m^{C,in}(t) = m^{C,L}(t = 0) + m^{C,r}(t)$$

If the integrity of measurement equipment and data observation is already proved, the carbon mass balance can be applied as a MCR-ALS constraint for glucose and acetate estimation from the spectra on each observation i over process runtime. For that, several non-spectroscopic measurements and assumptions must be applied to calculate carbon balances on each FTIR measurement.

The carbon in the reactor liquid phase is located in biomass in fractions of $\alpha^{C,cell}$ as well as in dissolved CO_2 in fractions of α^{C,CO_2}. The fraction of carbon in biomass is an assumption based on the analysis of elemental biomass composition of *E. coli* with an elemental analyser taken from literature [34]. Further carbon, of course, is located in glucose (glc) and acetate (ace), for which the concentrations c in reactor volume V^L are to be determined by FTIR/MCR-ALS. For MCR execution, $m_i^{C,L}$ must be split into one term containing the required concentrations from FTIR ex situ online measurement (ex) and into another term containing carbon compounds with concentrations accessible by in situ online measurement (on). For calculation of the online term, biomass concentration is observed by turbidity measurement and a calibrated exponential model. Dissolved carbon dioxide concentration is estimated by a soft-sensor based on Henry's law and the CO_2 mole fraction measured by a gas sensor in exhaust gas flow [35].

$$m_i^{C,L} = m_i^{C,L,on} + m_i^{C,L,at}$$

$$m_i^{C,L,on} = \left(\alpha^{C,cell} c_i^{cell,L} + \alpha^{C,CO_2} c_i^{CO_2,L} \right) V_i^L$$

$$m_i^{C,L,ex} = (\alpha^{C,glc} c_i^{glc,L} + \alpha^{C,ace} c_i^{ace,L}) V_i^L$$

The ex situ online term ex, containing the concentrations to estimate by MCR, must be converted concerning the left side of inequality constraint $\mathbf{A}c_i^T \leq \mathbf{b}_i$. Thus, the $(2,\Omega)$-matrix $\mathbf{A}$ contains in the first row, on positions associated with concentrations of glucose and acetate, the fractions $\alpha^{C,glc}$ and $\alpha^{C,ac}$ multiplied with reactor volumes V_i^L. The first row is associated with the upper bounds b^u of the constraint. The second row is the negative of the first row and is associated with the lower bounds b^l.

All other brought-in and recovery terms that are directly or indirectly accessible by online process sensors and soft-sensors, but not by FTIR/MCR-ALS, are used to form the $\mathbf{b}_i$ vector.

The carbon in the exhaust gas phase is calculated by the CO_2 removal rate Q^{CO2}, which in turn is calculated based on measurements of CO_2 mass flow at gas phase entry and of inert gas balance to estimate exit mass flow.

$$m_i^{C,g} = \alpha^{C,CO_2} \int_1^i V_i^L(t) Q_i^{CO_2}(t) dt$$

The calculation of carbon in samples of fermentation media starts at the first FTIR observation i = 1 with known initial media concentrations c_0 and sample volumes ΔV_0^{scc} and ΔV_0^{scf} taken before the first FTIR measurement is observed. A certain error in sample carbon mass calculation must be accepted since the respective current values of $c_i^{glc,L}$ and $c_i^{ace,L}$ are unknown at the time of constraint calculation. Hence, at i > 1, the results of the last MCR step i-1 are utilised. Considering comparative slow bioprocess kinetics and a higher sampling frequency, this is a reasonable approximation.

$$m_i^{C,scc} = \sum_{i=1}^m \left(\alpha^{C,glc} c_{i-1}^{glc,L} + \alpha^{C,ace} c_{i-1}^{ace,L} + \alpha^{C,cell} c_{i-1}^{cell,L} + \alpha^{C,CO_2} c_{i-1}^{CO_2,L} \right) \Delta V_{i-1}^{scc}$$

$$m_i^{C,scf} = \sum_{i=1}^m \left(\alpha^{C,glc} c_{i-1}^{glc,L} + \alpha^{C,ace} c_{i-1}^{ace,L} + \alpha^{C,CO_2} c_{i-1}^{CO_2,L} \right) \Delta V_{i-1}^{scf}$$

The brought-in carbon is the sum of the carbon fractions of glucose, acetate and cell mass in the initial medium as well as the supplied glucose from the feed reservoir.

$$m_i^{C,in} = m_{i=1}^{C,L} + m_i^{C,r} = \left(\alpha^{C,glc}\, c_{i=1}^{glc,L} + \alpha^{C,ace}\, c_{i=1}^{ace,L} + \alpha^{C,cell}\, c_{i=1}^{cell,L} \right) V_{i=1}^L + \alpha^{C,glc}\, c_i^{glc,r} V_i^r$$

By that information, $\mathbf{b}_i$ can be calculated as

$$\mathbf{b}_i = \left(\begin{array}{c} b^u \cdot m_i^{C,in} - \left(m_i^{C,L,on} + m_i^{C,g} + m_i^{C,scc} + m_i^{C,scf} \right) \\ -b^l \cdot m_i^{C,in} + \left(m_i^{C,L,on} + m_i^{C,g} + m_i^{C,scc} + m_i^{C,scf} \right) \end{array} \right)$$

The settings of the upper and lower tolerance bounds b^u and b^l of the carbon balance constraint are based on different considerations. Recovery rates higher than 1 are only caused by measurement errors while values below 1 are caused by both, measurement errors and not identified by-products. Therefore, the upper bound can be set tighter than the lower, with values in an interval of $b^u = (0.9, 1.1)$, depending on the process phase. An upper bound lower than 1 may be suitable if it is evident that the carbon compounds which are considered in the constraint calculation but which lie outside the optimisation of $\hat{\mathbf{C}}$ are underestimated (e.g., biomass). An upper bound higher than 1 is indicated if external carbon compounds seem to be overestimated. Thus, by tuning b^u, it therefore is possible to compensate measurement errors in online sensor equipment. As mentioned above, the recovery rate can reach approximately 90% at the end of the *E. coli* process, although just based on glucose, acetate, CO_2 and biomass. In order to take the formation of not considered carbon compounds into account, the lower bound is set in a defensive way to $b^l = 0.8$ to avoid over-restriction.

Furthermore, the non-negativity constraint is set for all concentration values, and at each new curve resolution step, the start value for lsqlin optimisation is set to the last estimation result. The initial concentration values for glucose, acetate, total phosphate and ammonia at i = 1 are set to the known batch medium concentration and may vary in a range of $\pm 10\%$.

Some online measurements such as fermenter weight, gas analysis and turbidity have a higher noise level and have to be filtered before further processing. Biomass estimated by turbidity is smoothed by application of an exponential smoothing filter with a smoothing factor alpha set to 0.05. The online signal of fermenter weight is prone to disturbances in form of high needle peaks, often caused by manual contact with the reactor e.g., while taking an offline sample. Those disturbances can easily be removed automatically by a threshold filter detecting differences between one measurement point to the next, higher than a threshold value e.g., >1.5 L, since offline samples usually have values below 1.5 and since the actual fermenter volume changing rate is much inferior. These few values in the sequel above the threshold are overwritten by the last value lower than the threshold. In this way, measurement errors can be significantly reduced since all concentration values are depending on the reactor volume. Outliers in the online gas analysis are treated correspondingly.

4. Results and Discussion

The carbon balance constraint algorithm with appropriate initial pure spectra estimations results in physically reasonable MCR solutions. By setting suitable start values and tolerance bounds, the rotatory and intensity ambiguities are reduced significantly. As a consequence, the concentration profiles of the substrates glucose, ammonia, total phosphate and the expected metabolic by-product acetate can be unfolded from the spectral mixture matrix $\mathbf{X}$ with minor manual measurement effort. An overview of the entire process spectra is displayed in Figure 3.
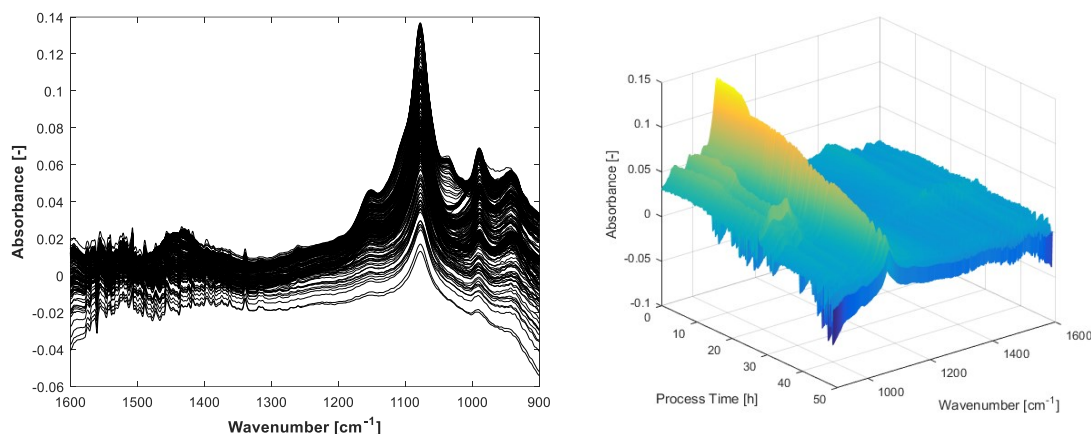
Figure 3. The 264 infrared spectra acquired from approximately 50 h fed-batch fermentation.

The FTIR spectra show negative values because of the water background subtraction. The inflexions downwards on the left and right borders of the display are caused by air bubbles in the flow cell. These artefacts can be handled by MCR. Before integrating the spectral air bubble model in the MCR-ALS algorithm, the assumption for the spectral air bubble model was ascertained by concentration prediction of a simple aqueous solution containing glucose, acetate and phosphate. The solution was measured by FTIR with and without air on the ATR crystal surface. The prediction was executed by multiplying the measured spectra **X** with the pseudoinverse of $\mathbf{S}_0$, whereby $\mathbf{S}_0$ is a composition of pure spectra of known mixture components. In one experiment, $\mathbf{S}_0$ involves an estimated spectral model for air bubbles, in the other just the pure spectra of the solved components are compounded. As evident from Figure 2 and Table 2, the integration of the estimated air bubble signature results in a significant prediction improvement.

The results of MCR-ALS concentration prediction based on the 264 measured process spectra are shown in Figure 4. Elapsed calculation time for 300 ALS iterations was about 10 min on an Intel Core i7-4790 @3.6 GHz (4 Cores). It should be noted that, besides the constraints described above, just single manual measured spectra for each estimated pure component are utilised to achieve the resolution. Likewise, some of the pure component start values are just vectors of uniformly distributed random numbers. After 300 alternating least squares iterations, satisfactory approximations of the process dynamics are obtained. Glucose and phosphate are present in higher concentrations, so the resolution succeeds nearly without artefacts. At concentrations close to zero, a higher presence of artefacts and noise is expectedly obtained. Accordingly, the lower concentrated ammonium and acetate show a higher ratio of disturbances.

The error evaluation takes place by comparing the FTIR/MCR-ALS concentration measurements with reference measurements. Concerning the residuals, the root mean squared errors (RMSE) are calculated and shown in Table 3.

The prediction results of the proposed MCR-ALS algorithm can be compared with prediction performances of PLS models. Acetate, ammonium and phosphate concentrations of a *Gluconacetobacter xylinus* fed-batch culture were predicted from spectra of in situ ATR-FTIR measurements by a PLS model with accuracies of 0.2, 0.17 and 0.24 gL^{-1}, respectively [24]. The validation errors for offline samples of the same process were 0.22 gL^{-1} (acetate), 0.24 gL^{-1} (ammonium) and 0.18 gL^{-1} (phosphate). The applied PLS regression model is based on 56 mixture solutions, used as calibration standards. The accuracies of MCR-ALS estimation for ammonium and acetate are similar to the PLS errors of the referenced paper. The absolute error of phosphate prediction is higher for the MCR-ALS approach than for the described PLS method, the measurement range being about two times higher, too. In consideration of the minor calibration effort of the proposed MCR-ALS approach, the results are impressive. Furthermore, the PLS glucose prediction accuracy by at-line ATR-FTIR monitoring of

an antibiotic fermentation process is with $0.56 \, \text{gL}^{-1}$ similar to the present prediction by MCR-ALS [23]. The PLS calibration model for glucose is based on 70 filtrated fermentation samples. Here, too, the reduction of calibration effort by effective online sensor data usage is evident when compared to PLS.

The estimations of pure component spectra are displayed in Figure 5. In addition to the notice concerning the associated concentrations, the higher noise level of the lower concentrated ammonium and acetate is also apparent in the pure spectral components.
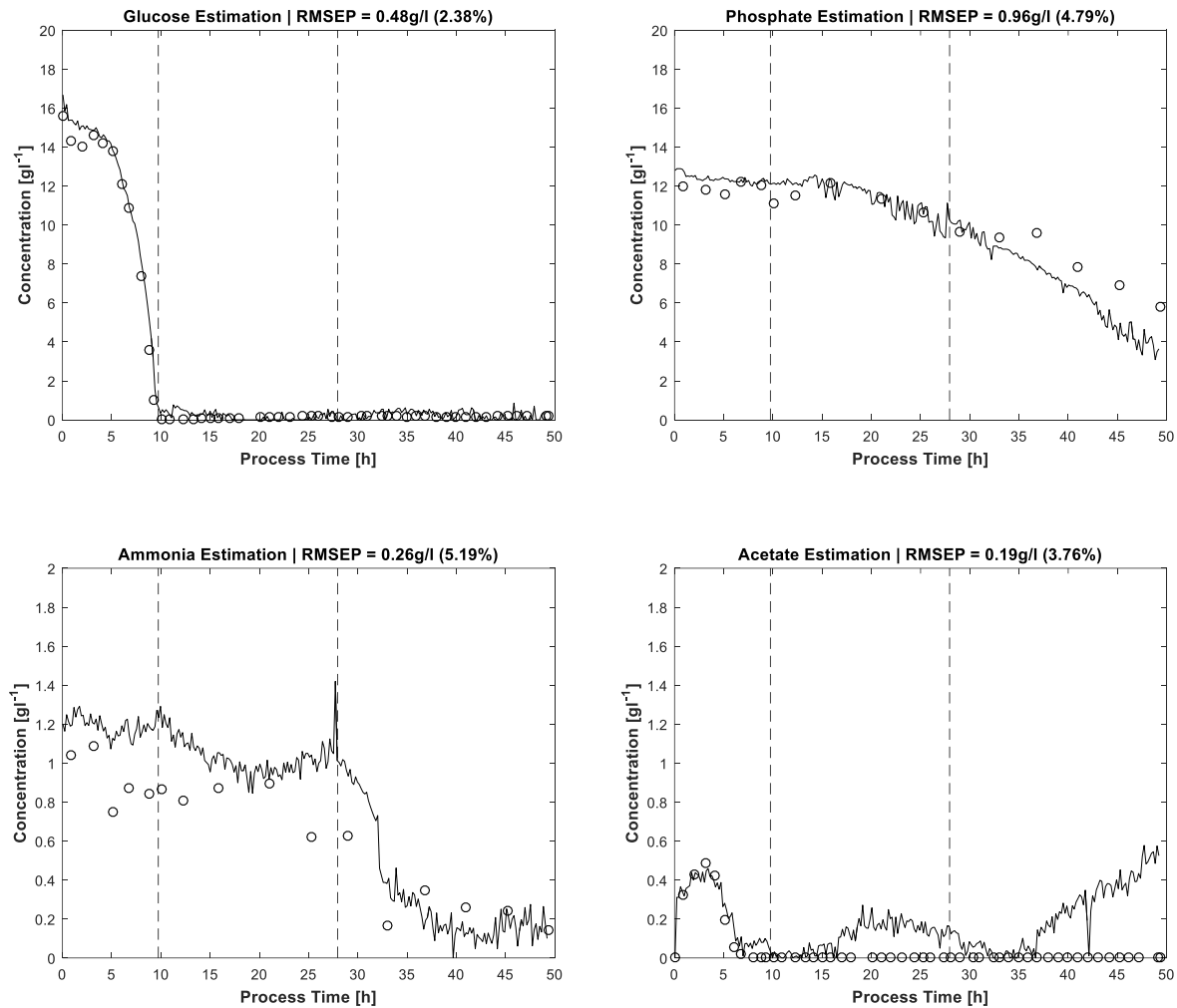

Figure 4. Concentration profiles estimated by FTIR/MCR-ALS (solid lines) and reference measurements (dots). The two dashed vertical lines in each plot distinguish the three process phases: batch phase (left), feeding phase/biomass production (middle), induction phase (right).

Table 3. Prediction performance: error estimation quantified by root mean squared distances between FTIR-MCR-ALS predictions and reference measurements.

| Substance	RMSE $[\text{gL}^{-1}]$	rel. RMSE [%]	Expected Range $[\text{gL}^{-1}]$
Glucose	0.48	2.38	(0, 20)
Phosphate	0.96	4.79	(0, 20)
Ammonium	0.26	5.19	(0, 5)
Acetate	0.19	3.76	(0, 5)

Root mean squared error (RMSE); relative RMSE related to expected range (rel. RMSE).
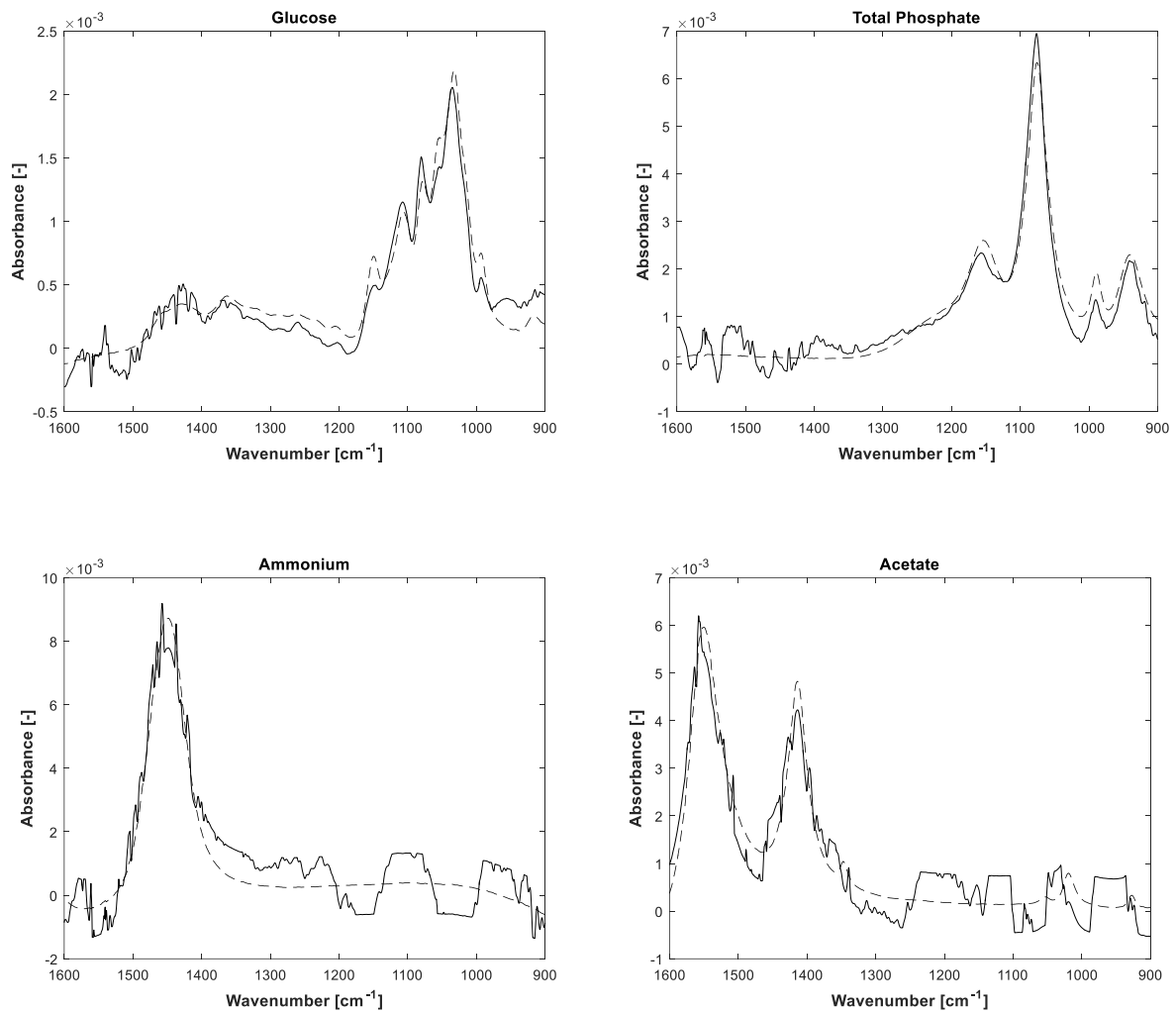
Figure 5. Initial measurements of pure substances (dashed lines) and identified pure spectral components after 300 ALS iterations (solid lines).

By way of comparison, Figure 6 shows the results for glucose and acetate concentrations without application of the carbon balance constraint but including the same constraints as used for pure spectra estimation, see above. Between hour 10 and 15 there is a significant artefact observable in the glucose concentration profile. The concentration estimation is too high, also discernible by carbon recovery rates approaching almost 1.2 in this process phase. A second drift in glucose concentration is located around t = 35 h without obvious reflection in carbon balance because of the lower deviation. In any case, without the carbon balance constraint, the solution space of MCR is enlarged, thereby also increasing the risk of ambiguities which can cause physically nonsensical solutions. For the same reason, the acetate profile in Figure 6 gives the impression of increasing concentrations which actually are not present. Nevertheless, the shapes of the actual existing concentration profiles in the batch phase are more or less recognised, the artefacts increasing mostly in the respective zero-concentration phases.
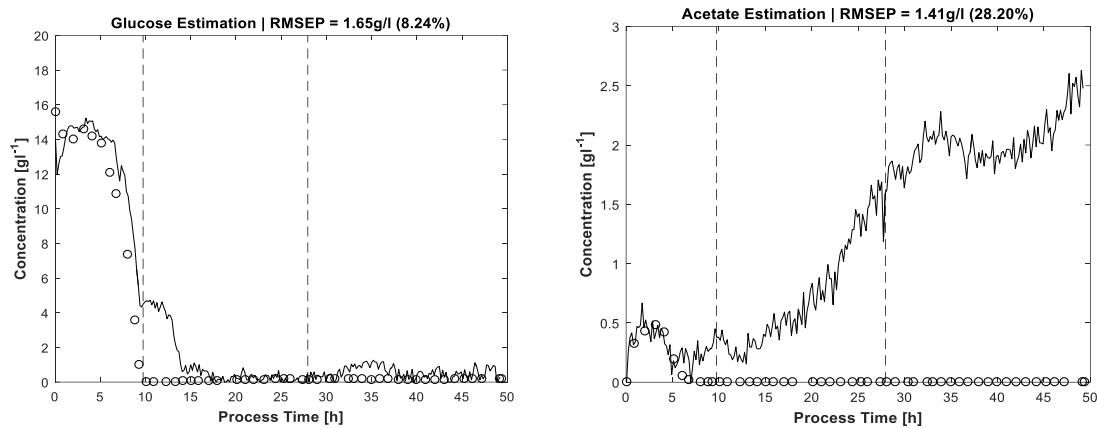
Figure 6. Concentration profiles for glucose (left) and acetate (right), estimated without application of the carbon balance constraint.

The carbon recovery at the end of the MCR-ALS procedure is shown in Figure 7. Without application of the carbon balance constraint, the recovery rate exceeds two times the value of 1.09, once in the beginning and once again at the end of fed-batch phase. Even the lower value of 0.8 is slightly undershot at the beginning of the process. Around t = 25h, near the end of the feeding phase, the recovery rate with the enabled carbon balance constraint touches the highest upper bound of 1.09. As for the final estimated concentrations, the lower bound of 0.8 is not reached.
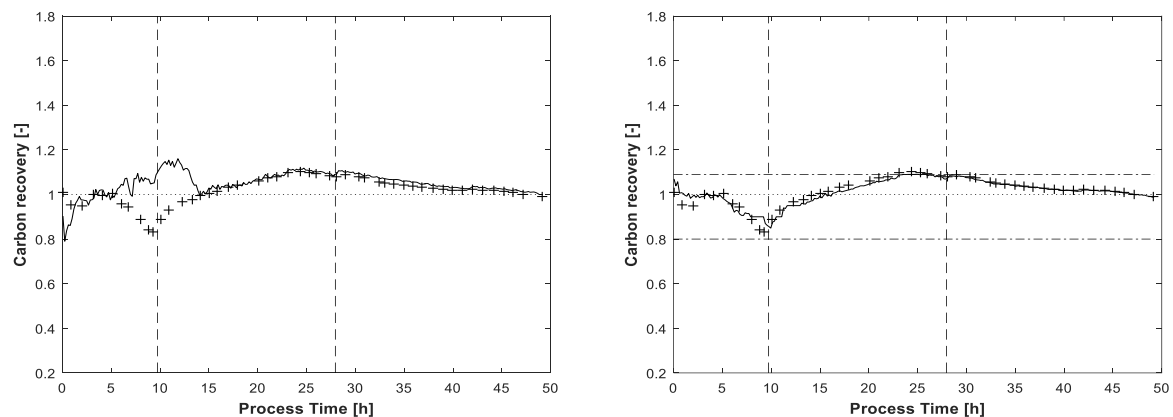
Figure 7. Carbon recovery based on the final glucose and acetate estimations after 300 ALS iterations, without (left) and with (right) application of the carbon balance constraint (solid line). Carbon recovery calculated by offline determined glucose and acetate concentrations (+). The dot-dashed horizontal lines indicate the highest upper and lowest lower bounds, only enabled in the run displayed on the right. The dotted horizontal line is the theoretical recovery rate of 1. The vertical lines indicate the three process phases: batch, feeding and induction.

5. Conclusions

This study has shown that MCR-ALS with tailored constraints is capable of analysing simultaneously the concentrations of glucose, acetate, ammonium and total phosphate from ex situ online recorded FTIR spectra of an *E. coli* HCDC fermentation process. The required concentration information, extracted from 264 FTIR spectra and recorded over 50 h process time, has been estimated in accuracies between 0.19 and 0.96 gL^{-1}. These results are comparable to established concentration estimations by PLS models, but are achieved with less calibration effort. It became apparent that the application of appropriate constraints, in particular the carbon balance constraint, improves the accuracy of concentration estimation in the ALS solution process by avoiding artefacts

caused by rotatory ambiguities. In MCR-ALS concentration estimation, the carbon mass balance constraint, calculated by online sensor data, reduced ambiguities in glucose and acetate concentrations significantly. In pure spectra estimation, initial FTIR measurements of the required analytes as well as a spectral air bubble model led to appropriate MCR solutions. Besides the automatically sampled online FTIR spectra, all applied constraints are calculated broadly based on automated measurements and analytical process knowledge. It is shown that by introducing prior knowledge and processed non-spectroscopic online sensor data into the ALS procedure, better spectra resolution performances as well as efficient fermentation process analysis can be achieved.

Author Contributions: Dennis Vier, Klaus-Uwe Gollmer and Volker Schünemann conceived and designed the experiments; Dennis Vier and Stefan Wambach performed the experiments; Dennis Vier analyzed the data; the University of Applied Sciences Trier–Umwelt Campus Birkenfeld (Departments "Bioprocess Engineering" and "Applied Informatics") contributed reagents/materials/analysis tools; Dennis Vier wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tauler, R.; Kowalski, B.; Fleming, S. Multivariate curve resolution applied to spectral data from multiple runs of an industrial process. *Anal. Chem.* **1993**, *65*, 2040–2047. [CrossRef]
2. Zhang, X.; Tauler, R. Measuring and comparing the resolution performance and the extent of rotation ambiguities of some bilinear modeling methods. *Chemom. Intell. Lab. Syst.* **2015**, *147*, 47–57. [CrossRef]
3. Ruckebusch, C.; Blanchet, L. Multivariate curve resolution: A review of advanced and tailored applications and challenges. *Anal. Chim. Acta* **2013**, *765*, 28–36. [CrossRef] [PubMed]
4. De Juan, A.; Jaumot, J.; Tauler, R. Multivariate Curve Resolution (MCR). Solving the mixture analysis problem. *Anal. Methods* **2014**, *6*, 4964–4967. [CrossRef]
5. Abdollahi, H.; Tauler, R. Uniqueness and rotation ambiguities in Multivariate Curve Resolution methods. *Chemom. Intell. Lab. Syst.* **2011**, *108*, 100–111. [CrossRef]
6. Pomerantsev, A.L.; Rodionova, O.Y. Process analytical technology. A critical view of the chemometricians. *J. Chemom.* **2012**, *26*, 299–310. [CrossRef]
7. Azzouz, T.; Tauler, R. Application of multivariate curve resolution alternating least squares (MCR-ALS) to the quantitative analysis of pharmaceutical and agricultural samples. *Talanta* **2008**, *74*, 1201–1210. [CrossRef] [PubMed]
8. Richards, S.E.; Becker, E.; Tauler, R.; Walmsley, A.D. A novel approach to the quantification of industrial mixtures from the Vinyl Acetate Monomer (VAM) process using Near Infrared spectroscopic data and a Quantitative Self Modeling Curve Resolution (SMCR) methodology. *Chemom. Intell. Lab. Syst.* **2008**, *94*, 9–18. [CrossRef]
9. Conti, P.; Zamponi, S.; Giorgetti, M.; Berrettoni, M.; Smyrl, W.H. Multivariate curve resolution analysis for interpretation of dynamic Cu K-edge X-ray absorption spectroscopy spectra for a Cu doped V(2)O(5) lithium battery. *Anal. Chem.* **2010**, *82*, 3629–3635. [CrossRef] [PubMed]
10. Garrido, M.; Rius, F.X.; Larrechi, M.S. Multivariate curve resolution-alternating least squares (MCR-ALS) applied to spectroscopic data from monitoring chemical reactions processes. *Anal. Bioanal. Chem.* **2008**, *390*, 2059–2066. [CrossRef] [PubMed]
11. Jaumot, J.; Vives, M.; Gargallo, R. Application of multivariate resolution methods to the study of biochemical and biophysical processes. *Anal. Biochem.* **2004**, *327*, 1–13. [CrossRef] [PubMed]
12. Piqueras, S.; Duponchel, L.; Offroy, M.; Jamme, F.; Tauler, R.; de Juan, A. Chemometric strategies to unmix information and increase the spatial description of hyperspectral images: A single-cell case study. *Anal. Chem.* **2013**, *85*, 6303–6311. [CrossRef] [PubMed]
13. Blanco, M.; Peinado, A.C.; Mas, J. Monitoring alcoholic fermentation by joint use of soft and hard modelling methods. *Anal. Chim. Acta* **2006**, *556*, 364–373. [CrossRef]

14. Grassi, S.; Amigo, J.M.; Lyndgaard, C.B.; Foschino, R.; Casiraghi, E. Assessment of the sugars and ethanol development in beer fermentation with FT-IR and multivariate curve resolution models. *Food Res. Int.* **2014**, *62*, 602–608. [CrossRef]

15. Grassi, S.; Alamprese, C.; Bono, V.; Casiraghi, E.; Amigo, J.M. Modelling Milk Lactic Acid Fermentation Using Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS). *Food Bioprocess Technol.* **2014**, *7*, 1819–1829. [CrossRef]

16. Koch, C.; Posch, A.E.; Goicoechea, H.C.; Herwig, C.; Lendl, B. Multi-analyte quantification in bioprocesses by Fourier-transform-infrared spectroscopy by partial least squares regression and multivariate curve resolution. *Anal. Chim. Acta* **2014**, *807*, 103–110. [CrossRef] [PubMed]

17. Mazarevica, G.; Diewok, J.; Baena, J.R.; Rosenberg, E.; Lendl, B. On-line fermentation monitoring by mid-infrared spectroscopy. *Appl. Spectrosc.* **2004**, *58*, 804–810. [CrossRef] [PubMed]

18. Landgrebe, D.; Haake, C.; Höpfner, T.; Beutel, S.; Hitzmann, B.; Scheper, T.; Rhiel, M.; Reardon, K.F. On-line infrared spectroscopy for bioprocess monitoring. *Appl. Microbiol. Biotechnol.* **2010**, *88*, 11–22. [CrossRef] [PubMed]

19. Lourenço, N.D.; Lopes, J.A.; Almeida, C.F.; Sarraguça, M.C.; Pinheiro, H.M. Bioreactor monitoring with spectroscopy and chemometrics: A review. *Anal. Bioanal. Chem.* **2012**, *404*, 1211–1237. [CrossRef] [PubMed]

20. Schmitt, J.; Flemming, H.-C. FTIR-spectroscopy in microbial and material analysis. *Int. Biodeterior. Biodegrad.* **1998**, *41*, 1–11. [CrossRef]

21. Shih, C.-J.; Smith, E.A. Determination of glucose and ethanol after enzymatic hydrolysis and fermentation of biomass using Raman spectroscopy. *Anal. Chim. Acta* **2009**, *653*, 200–206. [CrossRef] [PubMed]

22. Sivakesava, S.; Irudayaraj, J.; Ali, D. Simultaneous determination of multiple components in lactic acid fermentation using FT-MIR, NIR, and FT-Raman spectroscopic techniques. *Process Biochem.* **2001**, *37*, 371–378. [CrossRef]

23. Roychoudhury, P.; Harvey, L.M.; McNeil, B. At-line monitoring of ammonium, glucose, methyl oleate and biomass in a complex antibiotic fermentation process using attenuated total reflectance-mid-infrared (ATR-MIR) spectroscopy. *Anal. Chim. Acta* **2006**, *561*, 218–224. [CrossRef]

24. Kornmann, H.; Valentinotti, S.; Duboc, P.; Marison, I.; von Stockar, U. Monitoring and control of Gluconacetobacter xylinus fed-batch cultures using in situ mid-IR spectroscopy. *J. Biotechnol.* **2004**, *113*, 231–245. [CrossRef] [PubMed]

25. Xu, B.; Jahic, M.; Enfors, S.-O. Modeling of Overflow Metabolism in Batch and Fed-Batch Cultures of *Escherichia coli*. *Biotechnol. Prog.* **1999**, *15*, 81–90. [CrossRef] [PubMed]

26. Tauler, R.; Smilde, A.; Kowalski, B. Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution. *J. Chemom.* **1995**, *9*, 31–58. [CrossRef]

27. Diewok, J.; de Juan, A.; Maeder, M.; Tauler, R.; Lendl, B. Application of a Combination of Hard and Soft Modeling for Equilibrium Systems to the Quantitative Analysis of pH-Modulated Mixture Samples. *Anal. Chem.* **2003**, *75*, 641–647. [CrossRef] [PubMed]

28. Han, L.; Enfors, S.-O.; Häggström, L. *Escherichia coli* high-cell-density culture: Carbon mass balances and release of outer membrane components. *Bioprocess Biosyst. Eng.* **2003**, *25*, 205–212. [CrossRef] [PubMed]

29. Thermo Fisher Scientific Inc. *OMNIC 9. Version 9.1.24*; Thermo Fisher Scientific Inc.: Waltham, MA, USA.

30. Korz, D.J.; Rinas, U.; Hellmuth, K.; Sanders, E.A.; Deckwer, W.-D. Simple fed-batch technique for high cell density cultivation of *Escherichia coli*. *J. Biotechnol.* **1995**, *39*, 59–65. [CrossRef]

31. Riesenberg, D.; Schulz, V.; Knorre, W.A.; Pohl, H.-D.; Korz, D.; Sanders, E.A.; Ross, A.; Deckwer, W.D. High cell density cultivation of *Escherichia coli* at controlled specific growth rate. *J. Biotechnol.* **1991**, *20*, 17–27. [CrossRef]

32. The MathWorks, Inc. *MATLAB and Optimization Toolbox Release*; The MathWorks, Inc.: Natick, MA, USA, 2015.

33. Van Benthem, M.H.; Keenan, M.R.; Haaland, D.M. Application of equality constraints on variables during alternating least squares procedures. *J. Chemom.* **2002**, *16*, 613–622. [CrossRef]

34. Kayser, A. Metabolic flux analysis of *Escherichia coli* in glucose-limited continuous culture. I. Growth-rate-dependent metabolic efficiency at steady state. *Microbiology* **2005**, *151*, 693–706. [CrossRef] [PubMed]

35. Schumpe, A.; Quicker, G.; Deckwer, W.-D. Gas solubilities in microbial culture media. In *Reaction Engineering*; Springer: Berlin/Heidelberg, Germany, 1982; pp. 1–38.