



A systematic investigation of algorithm impact in preparative chromatography with experimental verifications

Martin Enmark^a, Robert Arnell^b, Patrik Forssén^a, Jörgen Samuelsson^a, Krzysztof Kaczmarek^c, Torgny Fornstedt^{a,*}

^a Analytical Chemistry, Department of Chemistry and Biomedical Sciences, Karlstad University, SE-651 88 Karlstad, Sweden

^b AstraZeneca Pharmaceutical Development, SE-151 85 Södertälje, Sweden

^c Department of Chemical and Process Engineering, Rzeszow University of Technology, ul. W. Pola 2, 35-959 Rzeszow, Poland

ARTICLE INFO

Article history:

Received 29 June 2010

Received in revised form 19 October 2010

Accepted 16 November 2010

Available online 24 November 2010

Keywords:

Equilibrium-Dispersive model

Inverse method

IM

Simulation

Experimental validation

Quaternary separation

Selectivity

Efficiency

Process chromatography

Orthogonal collocation on finite elements

OCFE

Rouchon finite difference method

RFD

ABSTRACT

Computer-assisted optimization of chromatographic separations requires finding the numerical solution of the Equilibrium-Dispersive (ED) mass balance equation. Furthermore, the competitive adsorption isotherms needed for optimization are often estimated numerically using the inverse method that also solves the ED equations. This means that the accuracy of the estimated adsorption isotherm parameters explicitly depends on the numerical accuracy of the algorithm that is used to solve the ED equations. The fast and commonly used algorithm for this purpose, the Rouchon Finite Difference (RFD) algorithm, has often been reported not to be able to accurately solve the ED equations for all practical preparative experimental conditions, but its limitations has never been completely and systematically investigated. In this study, we thoroughly investigate three different algorithms used to solve the ED equations: the RFD algorithm, the Orthogonal Collocation on Finite Elements (OCFE) method and a Central Difference Method (CDM) algorithm, both for increased theoretical understanding and for real cases of industrial interest. We identified discrepancies between the conventional RFD algorithm and the more accurate OCFE and CDM algorithms for several conditions, such as low efficiency, increasing number of simulated components and components present at different concentrations. Given high enough efficiency, we experimentally demonstrate good prediction of experimental data of a quaternary separation problem using either algorithm, but better prediction using OCFE/CDM for a binary low efficiency separation problem or separations when the compounds have different efficiency. Our conclusion is to use the RFD algorithm with caution when such conditions are present and that the rule of thumb that the number of theoretical plates should be greater than 1000 for application of the RFD algorithm is underestimated in many cases.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Today there is a growing interest to use computer-assisted optimization of preparative chromatographic separations for industrial applications. This trend is due to the development of modern computing power combined with the recent development of rapid methods for parameter determination; e.g. the Inverse Method (IM). The concept is in line with the FDA guidance (Pharmaceutical CGMPs, SEP 2004) pointing out the importance of (i) understanding and (ii) developing mathematical relationships for the process.

Chiral preparative chromatography is of special importance for the pharmaceutical industry; regulatory authorities today require that the pharmaceutical industry investigate the optical

isomers of each candidate drug regarding pharmacological and, especially, toxicological and pharmacokinetic properties (Pharmaceutical CGMPs, SEP 2004). For each new drug application, there are thousands of candidate drugs, where each optical isomer must be isolated in mg – g amounts for evaluation. By using numerical simulation, we can predict the optimal experimental conditions for maximal throughput of the desired isomer(s) under given restrictions, e.g. yield and recovery. The most important information required for the numerical optimization is the competitive adsorption isotherms of the product and impurities.

The inverse method is a rapid method that has been developed for obtaining the competitive adsorption isotherm input data directly from the overloaded elution profiles and has been successfully applied for many years [1].

For many experimental situations, the column model considered in the IM is the Equilibrium-Dispersive (ED) mass balance equations [1]. The question of which mass balance equation is valid,

* Corresponding author. Tel.: +46 54 700 19 60.

E-mail address: Torgny.Fornstedt@kau.se (T. Fornstedt).

for the particular experiment, is of another nature, and has been discussed in detail earlier [2,3]. In this work, we will use the ED model to describe the propagation of components in a chromatographic separation.

Solving the ED mass balance equation accurately is of great importance in the IM, since errors in the solution will decrease the model's ability to predict elution profiles and hence to accurately estimate the adsorption isotherm parameters. There has been a lot of research regarding this problem, but there still exist unresolved questions. The reason to this is mainly that there are no analytical solutions even for the typical binary, non-linear separation problem. Furthermore, the error analysis of any applied numerical scheme is complicated. In the case of different finite difference schemes, exact conditions on the time and space discretization are only available for linear conditions. However, accumulated experience has shown that the approximate guidelines used are sufficiently good in most cases and this has been verified by the degree of overlap between experimental and simulated elution profiles for different experimental systems and adsorption isotherm models [1].

The accuracy of the commonly used Rouchon Finite Difference (RFD) numerical scheme is only first order. To reduce numerical error, it is possible to state finite difference schemes with error terms of higher order and let the grid of time and space approach infinite discretization, or to use finite element schemes [1,4]. The recommendation has been not to use other finite difference schemes but instead the method of Orthogonal Collocation on Finite Elements (OCFE) [1].

Because of the simplicity of implementation of a finite difference scheme over OCFE, we will systematically compare RFD with another higher order finite difference scheme and the OCFE algorithm. We will not only present numeric studies, but also validate our findings with experimental studies when the inverse method is applied for high and low efficiency chromatographic separation processes.

Although a huge number of successful applications of RFD [5] a complete consensus has not yet been reached concerning when the RFD should be used and when it is needed to use a more time consuming alternative such as OCFE. Therefore there is a need for a more systematic investigation. Moreover, in some of the reports published, some contradictory conclusions appear concerning the limitations of RFD. Earlier, it has been concluded that there exists serious problems with RFD when applied to binary separations with low column efficiency, $N < 1000$, regardless of adsorption isotherm, but negligible differences for $N > 1000$ when considering Langmuir like adsorption isotherms [1]. This is in clear contrast to a recent publication where loss of accuracy was reported in a binary separation with an additive at such high column efficiency as $N > 10,000$ [5]. Loss of accuracy was also reported for binary separations at average column efficiency where the components are present at markedly different concentrations ratios [1] and serious error or even no computability was reported for some adsorption isotherms with different concavity than Langmuir [5]. Common for all these reports are that they demonstrate problems with RFD under some specific conditions and that they indicate that the rule of thumb to require at least 1000 theoretical plates for RFD may be greatly underestimated.

Because low column efficiency, typically less than 1000 and as low as 100 theoretical plates, still are frequently encountered in process scale chromatography, it is also important to investigate the performance of RFD and more accurate algorithms in more detail. In addition, in these often chiral separations of optical isomers, the second eluted isomer regularly has less column efficiency than the first eluted one. Because the RFD algorithm uses average column efficiency, we need to investigate the case of mixed low column efficiency.

To summarize, this work aims at giving more complete guidelines on which algorithm to use under different experimental conditions, more specifically to:

- (i) Estimate adsorption isotherms with the IM and compare the goodness of profile fit for a quaternary high efficiency separation problem, as well as for a low efficiency binary separation problem.
- (ii) Compare the solutions obtained with the different algorithms for a quaternary separation problem for a large set of combinations of column efficiency and load.
- (iii) Compare the solutions obtained with the different algorithms for a binary separation problem when the two components are present at different concentrations and column efficiencies.
- (iv) Compare the solutions obtained with the different algorithms for a binary separation problem when the two components have different column efficiency.
- (v) Compare the solution obtained by different algorithms as a function of the number of components and combination of retention factors.

In contrast to earlier studies, showing isolated cases, we will compare over 1000 simulated eluted profiles in order to investigate trends.

2. Theory

2.1. Column model

The Equilibrium-Dispersive (ED) model is a generally accepted and frequently used model for many applications of LC [1]. Despite its many simplifications of the chromatographic process, it has been successfully used in modeling the separation of many substances [6],

$$\begin{cases} \frac{\partial C_i(x, t)}{\partial t} + F \frac{\partial q_i(x, t)}{\partial t} + u \frac{\partial C_i(x, t)}{\partial x} = D_{a,i} \frac{\partial^2 C_i(x, t)}{\partial x^2}, \\ 0 \leq x \leq L, \quad t > 0, \quad i = 1, \dots, n, \\ C_i(x, 0) = 0, \\ C_i(0, t) = \varphi_i(t). \end{cases} \quad (1)$$

Eq. (1) is the differential mass balance equation for component i of n . $C_i(x, t)$ and $q_i(x, t)$ are the stationary and mobile concentrations of component i at time t in and column position x . F is the phase ratio, L is the column length and u is the linear velocity. $D_{a,i} = Lu/2N_{ap,i}$ is the lumped mass transfer and dispersion coefficient. $C_i(0, t) = \varphi_i(t)$ is the left boundary condition, i.e., $\varphi_i(t)$ is a measured or idealized injection profile. $C_i(x, 0)$ is the initial conditions of the column.

Existing methods to solve Eq. (1) range from using different kinds of finite difference approximations to finite element approximations [1,7]. In this paper, we will focus on the application of three numerical schemes to solve Eq. (1), two finite difference methods and one finite element method. The first finite difference scheme is the well studied and successfully applied finite difference scheme is the Rouchon Finite Difference method (RFD) [8]. Whenever it is applied in this work, it is the implementation of Forssén et al. [6].

To further investigate the finite difference scheme approach in solving Eq. (1), we have implemented a higher order scheme. Because of the use of a central difference approximation of the spatial derivative, we call this scheme the Central Finite Difference Method (CDM).

The solution of partial differential equations using the orthogonal collocation method was widely discussed in [9], and this method has successfully been applied to the modeling of different separation- and reactor processes in chemical engineering. The OCFE method, initially applied to the solution of Eq. (1) by Ma and

Guiochon [4], will be used here with the refined implementation of Kaczmarzski et al. [10].

While the RFD method is one of the fastest numerical schemes, its order of accuracy, use of averaged column efficiency and component traveling speed will not guarantee exact solution of Eq. (1). Obtaining solutions of the ED model, using the method of OCFE, is considered to be the most accurate and robust strategy available [5].

2.2. Numerical solutions to column model

The RFD numerical scheme used in this paper is an apparent solution of the ideal model of chromatography where the physical dispersion is approximated by the numerical dispersion that is adjusted by the discretization of the ED model,

$$\begin{aligned} \frac{C_{i,j}^{n+1} - C_{i,j}^n}{\Delta t} + F \frac{q_{i,j}^{n+1} - q_{i,j}^n}{\Delta t} + u \frac{C_{i,j+1}^{n+1} - C_{i,j}^{n+1}}{\Delta x} &= 0, \\ i = 1, \dots, N_c, \quad j = 0, 1, \dots, N_{\text{spacesteps}}, \\ n = 0, 1, \dots, N_{\text{timesteps}}, \\ C_{i,j}^0 &= C_{0,i}, \\ C_{i,0}^n &= \varphi_i(t), \quad 0 \leq n \leq N_{\text{injectionsteps}} + 1, \\ \Delta x &= \frac{L}{N_{\text{avg}}}, \\ \Delta t &= \frac{2L}{N_{\text{avg}} u_{\text{avg}}^{\text{linear}}}. \end{aligned} \quad (2)$$

Here N_c is the number of components and N_{avg} is the average column efficiency for all simulated components.

To determine the concentration of a species in the unknown time and space step, the following update formula is used.

$$\begin{pmatrix} C_{1,j+1}^{n+1} \\ C_{2,j+1}^{n+1} \\ \vdots \\ C_{N_c,j+1}^{n+1} \end{pmatrix} = \begin{pmatrix} C_{1,j}^{n+1} - \frac{\Delta x}{u \Delta t} (C_{1,j}^{n+1} - C_{1,j}^n + F(q_{1,j}^{n+1} - q_{1,j}^n)) \\ C_{2,j}^{n+1} - \frac{\Delta x}{u \Delta t} (C_{2,j}^{n+1} - C_{2,j}^n + F(q_{2,j}^{n+1} - q_{2,j}^n)) \\ \vdots \\ C_{N_c,j}^{n+1} - \frac{\Delta x}{u \Delta t} (C_{N_c,j}^{n+1} - C_{N_c,j}^n + F(q_{N_c,j}^{n+1} - q_{N_c,j}^n)) \end{pmatrix} \quad (3)$$

The CDM numerical scheme can be explicitly formulated as follows,

$$\begin{aligned} \frac{C_{i,j}^{n+1} - C_{i,j}^n}{\Delta t} + F \sum_{z=1}^{N_c} \frac{\partial q_i C_{z,j}^{n+1} - C_{z,j}^n}{\Delta t} + u \frac{C_{i,j+1}^{n+1} - C_{i,j}^{n+1}}{2 \Delta x} \\ = D_{a_i} \frac{C_{i,j+1}^{n+1} - 2C_{i,j}^{n+1} + C_{i,j-1}^{n+1}}{\Delta x^2}, \\ i = 1, \dots, N_c, \quad j = 0, 1, \dots, N_{\text{spacesteps}}, \\ n = 0, 1, \dots, N_{\text{timesteps}}, \\ C_{i,j}^0 &= C_{0,i} \\ C_{i,0}^n &= \varphi_i(t), \quad 0 \leq n \leq N_{\text{injectionsteps}}, \\ \frac{\partial C_i(x, t)}{\partial x} \Big|_{x=L} &= 0, \\ \Delta x &= \frac{L}{\max(N_i)}, \\ \Delta t &= \varepsilon \frac{\Delta x}{\max(L/t_R)}, \end{aligned} \quad (4)$$

where ε is a small number, dynamically adjusted to ensure algorithm stability that is imposed by requiring that mass is preserved throughout simulation. If there are steep concentration gradients, e.g., due to high column efficiency or large ratios of component concentrations, the algorithm might be unstable and ε is decreased. The

parameter ε is therefore an implicit function of both q , column efficiency N and C . The parameter t_R is the retention times of simulated components at linear conditions.

The explicit formulation of the update formula is written,

$$\begin{aligned} \begin{pmatrix} C_{1,j}^{n+1} \\ C_{2,j}^{n+1} \\ \vdots \\ C_{N_c,j}^{n+1} \end{pmatrix} &= \frac{1}{\Delta t} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N_c} \\ a_{21} & a_{22} & \dots & a_{2N_c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N_c1} & a_{N_c2} & \dots & a_{N_cN_c} \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_c} \end{pmatrix} + \begin{pmatrix} C_{1,j}^n \\ C_{2,j}^n \\ \vdots \\ C_{N_c,j}^n \end{pmatrix} \\ a_{\alpha\beta} &= \begin{cases} 1 + F \frac{\partial q_\alpha}{\partial C_\beta}, & \alpha = \beta, \\ F \frac{\partial q_\alpha}{\partial C_\beta}, & \alpha \neq \beta, \end{cases} \\ x_k &= D_{a_k} \frac{C_{k,j+1}^n - 2C_{k,j}^n + C_{k,j-1}^n}{\Delta x^2} - u \frac{C_{k,j+1}^n - C_{k,j-1}^n}{2 \Delta x} \end{aligned} \quad (5)$$

RFD requires no external computational routines when compiled using Fortran code, but CDM uses the DGESV and ZGESV routines in the LAPACK library [11].

The explicit statement of the OCFE method used in this paper is much more complex [9]. The idea of the method is to divide the normalized space coordinate in the interval [0,1] into NS subdomains (elements). In each k th element, the $N(k)$ internal collocation points are defined and the solution is approximated by Lagrange polynomial of degree $(N(k)+2)$. The overall solution is obtained by joining the solutions in each element. The details of the discretization of the spatial derivatives following the orthogonal collocation method are presented in [10]. The set of ordinary differential equations obtained through the discretization process was solved using the VODE solver [12]. The VODE solver automatically controls the integration time interval to fulfill the requirement of accuracy of the calculation. However, the number of the sub domains, NS , and the number of internal collocation points, $N(k)$, have to be chosen individually for the analyzed problem.

The OCFE method produces solutions with oscillatory behavior. The oscillations vanish when NS and $N(k)$ increase. The obtained solution can be regarded as “real” when: (i) increasing the number of sub domains or internal collocation points, above the certain values has practically no effect on the solution obtained, (ii) the mass balance is fulfilled.

For the solution of a chromatographic separation process, when the isotherm model is described by convex upward isotherm, $N(k)$ can be assumed to equal 3 for each sub domain. In such case, the number of sub domains should be close to a tenth of N , where N is the number of theoretical plates. In this work we have assumed $N(k)=3$, $NS=0.1N$ and checked that further increasing NS has no effect on the solution obtained.

2.3. Adsorption isotherms and their determination

The adsorption isotherm applied in this work is the competitive bi-Langmuir,

$$q_i = \frac{a_{1,i} C_i}{1 + \sum_{j=1}^{N_c} b_{1,j} C_j} + \frac{a_{2,i} C_i}{1 + \sum_{j=1}^{N_c} b_{2,j} C_j} \quad (6)$$

where $a_{j,i}$ and $b_{j,i}$ are the distribution coefficient and the association equilibrium constant for the j th site and i th compound.

There are numerous methods of determining adsorption isotherm parameters and the reader is referred to the detailed reviews of Guiochon et al. [1] and Morgenstern [13]. In this study, the inverse method is used, it has been discussed and applied numerous times in the literature [1,6,14]. The steepest descent procedure used in this work does not ensure that a global optimum is found, i.e., the best possible adsorption isotherm parameters are

found, once the convergence criterion is reached in the inverse method. Using implementations involving genetic search algorithms decreases this risk [15,16].

3. Experimental

3.1. Separation of ethyl- and methyl mandelate on AGP column

A 10×0.46 cm AGP column (ChromTech, Hägersten, Sweden), consisting of $5 \mu\text{m}$ silica beads with immobilized α_1 -acid glycoprotein (AGP) as chiral selector was installed in a Hewlett-Packard, Agilent 1100 system (Agilent Technologies, Palo Alto, CA, USA). Eluent was prepared by dissolving acetic acid and sodium acetate in MilliQ water, making a 75 mM acetate buffer at pH 5, ionic strength 50 mM. Additionally, methanol was added to a concentration of 1.4% (v/v). The buffer was filtered through a $0.22 \mu\text{m}$ filter (Millipore) before use.

The column was equilibrated with eluent at a constant, calibrated flow at 0.7 mL/min and kept at 20°C in water bath during all experiments. UV-detection was made by a DAD detector and chromatograms were exported at 254 nm.

4.8 mM stock solutions of S(+)/R(–)-ethyl/methyl mandelate were prepared by dissolving S(+)/R(–)-ethyl/methyl mandelate (99% purity, Sigma-Aldrich, Stockholm, Sweden) directly in the mobile phase. To measure the injection profile of the system, the capillary tube connecting the injector and column was connected directly into the UV-detector whereupon a triplicate of $60 \mu\text{L}$ injections of 0.5 mM S(+)-ethyl/methyl mandelate were made.

Initial analytical injections to determine efficiency (half height method) of the two enantiomers were made in triplicate by injecting $7 \mu\text{L}$ of $40 \mu\text{M}$ racemate and single injections of each enantiomer. The column efficiencies were found to be 2740, 2690, 2270 and 2360 theoretical plates for the enantiomers S(+)/R(–) methyl-mandelate and S(+)/R(–)-ethyl-mandelate, respectively.

Four overloaded injections were made with injection volume $60 \mu\text{L}$ and with injection concentrations were 0.6, 0.8, 0.96 and 1.2 mM per component, respectively. Maximum elution concentrations, determined by converting the detector signal, for these injections were 0.04, 0.08, 0.18 and 0.23 mM for the enantiomers S(+)/R(–) methyl-mandelate and S(+)/R(–)-ethyl-mandelate, respectively. Uracil (99%, Sigma, Stockholm, Sweden) was used as unretained void-volume marker. The elution peak of uracil was also compared to the minor disturbance of pure eluent injection peak and was found to coincide at t_0 at 1.5 min (triplicate injections of $7 \mu\text{L}$). Component response curves were measured by pumping different ratios of mobile phase and 4.8 mM stock solution directly into the detector, where the detector response was found to be linear.

3.2. Binary low efficiency separation

A small racemic drug intermediate molecule from AstraZeneca was used as a model substance. No details of the molecular properties can be presented since this information is classified. It will be referred to as the “AstraZeneca substance”.

A 25×0.46 cm ChiralCel-OD column (Daicel Chemical Industries Ltd., Tokyo, Japan), consisting of $20 \mu\text{m}$ silica beads with modified cellulose as chiral selector was installed in a Waters 2695 Separation system. Eluent was prepared by mixing heptane (Fischer Scientific, Loughborough, Leicestershire, UK) and EtOH (Kemetyl AB, Haninge, Sweden) to a 7:3 (v/v) proportion.

The column was equilibrated with eluent at flow at 1.0 mL/min. and kept at a controlled temperature of 25.0°C in column oven. UV-detection was made by a DAD detector and a wavelength of 270 nm where the detector response was found to be non-linear.

A 47 g/L racemate (AZ substance) stock solution was prepared by dissolving substance directly in the eluent. The injection profile of the system was measured by connecting the capillaries from the injector directly into the detector. Also, using this setup, 20 of 1 mL pulses ranging from concentrations of 47 g/L to 0.3 g/L were injected to calibrate the detector response. The non-linear calibration curve was then interpolated by a cubic spline.

Initial analytical injections to determine the column porosity and efficiency (half height method) of the two enantiomers were made by injecting $10 \mu\text{L}$ stock solution with additional 0.25 mM of 1,3,5-tri-tert-butylbenzene (TTBB) as unretained void-volume marker. Void times of 3.05, 2.0 and 1.55 were recorded at 1, 1.5 and 2 mL/min, respectively. The column efficiency, for the first and second eluting component of the AstraZeneca substance, respectively, was found to be 410 and 380 theoretical plates at 2 mL/min, 550 and 570 theoretical plates at 1.5 mL/min and 820 and 760 theoretical plates at 1 mL/min.

Overloaded injections were made using 100, 300, 450 and $600 \mu\text{L}$ injections of 47 g/L stock solution at the flow 2 mL/min. Injections were duplicated to ensure experimental stability. The maximum eluted concentration was 19.9 and 15.3 g/L for the first and second eluting component of the AstraZeneca substance, respectively.

3.3. Determination of adsorption isotherm parameters

For the quaternary case, all overloaded elution profiles recorded in Section 3.1 were simultaneously used to estimate the adsorption isotherm parameters with the inverse method. This particular system has previously been studied using the perturbation peak method and the bi-Langmuir model was found to describe the raw adsorption data excellently [17] and was therefore used here. This estimation was repeated using the RFD, CDM and OCFE algorithm to solve the ED equations in the inverse method, see results in Table 1. For RFD the average column efficiency 2515 theoretical plates were used. The rounded parameters estimated using the RFD algorithm was then used in the synthetic study in Sections 4.1–4.4.

In the low efficiency, binary separation, all overloaded elution profiles recorded in Section 3.2 were simultaneously used to estimate bi-Langmuir adsorption isotherm parameters with the inverse method in the same manner as above, see results in Table 1. For RFD the average column efficiency 395 theoretical plates were used.

3.4. Simulations

All simulations use an measured injection profile $\varphi_i(t)$, which was determined from quaternary or binary experiment in Sections 3.1 and 3.2. The measured experimental injection profile was interpolated using a piecewise cubic Hermite interpolating polynomial. The interpolated profile $\varphi_i(t)$ could then be used for all injection concentrations when keeping the injection volume and flow rate fixed. For OCFE and CDM, $\varphi_i(t)$ is identical for all values of N but due to the coarser space-time discretization of RFD, $\varphi_i(t)$ cannot be identical to that of OCFE/CDM at low column efficiencies.

All simulated injection volumes in Sections 4.1–4.4 was fixed at $60 \mu\text{L}$ and the flow rate fixed at 0.7 mL/min. Basic model data in these sections, such as column length, phase ratio and void time were all taken from Section 3.1. The RFD bi-Langmuir adsorption isotherm parameters for S(+)/R(–) methyl/ethyl mandelate obtained in Section 3.3 were used and the components are referred to as “Components 1–4”. Whenever CDM or OCFE is used, each component has its unique defined apparent dispersion coefficient $D_{a,i}$ but for RFD the average value, D_a is used.

Table 1
The estimated bi-Langmuir adsorption isotherm parameters using the IM with three different algorithms. The top set of parameters describes the adsorption of methyl mandelate (MM) and ethyl mandelate (EM) isomers S(+) and R(-), they are used for numerical simulation studies in chapter 4, and are there referred to as components 1–4. The bottom set of parameters describes the adsorption of the "AstraZeneca substance".

Component	a_1			b_1 [m ⁻¹]			a_2			b_2 [m ⁻¹]		
	RFD	CDM	OCFE	RFD	CDM	OCFE	RFD	CDM	OCFE	RFD	CDM	OCFE
N > 2000												
MM S(+), (1)	1.37	1.25	1.25	1340	1350	1790	2.35	2.55	2.59	2260	2258	1980
MM R(-), (2)	1.33	1.26	1.28	483	487	456	3.47	3.56	3.55	3640	3636	3480
EM S(+), (3)	2.14	2.00	1.97	492	493	486	6.19	6.39	6.45	5720	5765	5720
EM R(-), (4)	2.17	2.43	2.32	0.52	0.610	0.24	11.3	10.82	11.0	8830	8467	8410
N < 500												
AZ subst. comp. 1	0.30	0.40	0.40	14.1	11.5	6.24	0.23	0.09	0.13	9.30	0.05	8.74
AZ subst. comp 2	1.20	1.25	1.22	0.01	0.00	0.00	0.22	0.15	0.10	139	115	47.6

To quantify the overlap of experimental and simulated elution profiles for each algorithm, the total area overlap formula was used,

$$\text{Overlap} = 100 \cdot \frac{\int_0^\infty \min(C_{\text{sim}}(t), C_{\text{exp}}(t)) dt}{\int_0^\infty C_{\text{sim}}(t) dt} \quad (7)$$

This formula can be considered to measure how much two 2D-objects with equal area intersects, i.e., how large portion of each object is in the intersection of them. An overlap approaching 100% indicates complete overlap, while 0% indicates that they are completely separated.

Because the time discretization of the RFD algorithm is much coarser than that used in CDM, the computation of overlap will introduce a small error because $\min(C_{\text{CDM}}(t), C_{\text{RFD}}(t))$ cannot be evaluated at the exact same point in time. It would be possible to interpolate the RFD profile to avoid this but this would also add a small error.

The binary model system described in Section 3.2, with adsorption isotherm parameters listed in Table 1, was used to maximize production rate in Section 4.6. Parameters t_0 and N were interpolated for flow rates of 1–2 mL/min. The injection profile was interpolated for flow rates of 1–2 mL/min and injection volumes of 100–600 μ l. Besides measured N values, we made investigations on a quarter and double that value. The constraints on maximum production rate were an excess of 95% purity and 75% yield of the first eluting component.

4. Results and discussion

The purpose of the study is to investigate, more carefully and completely, the previous claims in the literature of the insufficiency of the RFD algorithm. More specifically binary separations where one has low column efficiency and different injection concentration ratios. To do this, we have made a large number of simulations to determine trends in the difference between the RFD solution, the OCFE/CDM solution. Additionally, we have investigated areas not yet studied: the case of the different solutions obtained for a quaternary simulation with different column loads and efficiencies and a binary separation where the two components have different column efficiencies. To experimentally validate this, we used two different experimental systems: one quaternary and one binary with low column efficiency. To limit the scope of our investigation, we focus on the most commonly encountered adsorption isotherm, type I, e.g. Langmuir. Because the implemented finite difference scheme of higher order (CDM) gives identical solutions to OCFE, which in turn is considered as the reference solution to the ED model, whenever the RFD solutions deviates we will interpret this as an error in the RFD algorithm.

4.1. Quaternary simulations: impact of column efficiency and load

Simulations of a four component separation were performed at equal injection concentrations ranging from 0.1 to 1.5 mM and column efficiencies ranging from 200 to 3000 in a grid of 225 equally spaced points. The simulated elutions profiles using the RFD and CDM algorithm, each corresponding to a point in the grid, were then compared and the overlap was calculated component wise. A smaller subset of 36 evenly distributed simulations across the entire set of profiles, using CDM, were compared to solutions acquired by OCFE. These profiles were found to be identical.

The column load and efficiency correlation to degree of overlap between the RFD and CDM algorithm are different when considering co-eluting and base-line separated species as they are different when one considers the least and most retained components. For co-eluting species, corresponding to components 1 and 2, we have found that a higher column load is necessary to reach a stable and high overlap of the different solutions even for higher column efficiencies, see Fig. 1a and b. However, when considering base-line separated species, components 3 and 4, the degree of overlap is less sensitive to changes in column load and or efficiency, see Fig. 1c and d. Also, there seems to be a correlation between the overlap calculated for the earliest and latest eluting component, components 1 and 4, in Fig. 1a and d, respectively, which in these quaternary simulations might be a manifestation of the simulated average effective component traveling velocity using the RFD algorithm, see details in Section 2.2. It can clearly be seen that for low column load, an efficiency exceeding 2500 is not sufficient to guarantee the accuracy of the RFD solution.

4.2. Binary simulations: impact of column efficiency

Because the RFD algorithm uses the average column efficiency when simulating an experiment, we investigated, if and where, a correlation of algorithm overlap and component efficiency ratio becomes noticeable for a binary simulation. Simulations were performed at injection concentrations of 1 mM per component. This value was chosen from the results shown in Fig. 1, where 1 mM gives good overlap for most efficiencies. Components 1+2 and 1+3 in the quaternary example, were simulated separately, using combinations of efficiencies ranging from 200 to 3000 theoretical plates in a grid of 225 equally spaced points. The simulated profiles obtained by RFD and CDM for each set of binary simulations were compared to each other and the overlap calculated.

Again, different results were obtained depending on if the considered simulation has co-eluting or base-line separated components. For co-eluting components, simulations describing the individual components differ if one leaves a 1:1 ratio of column

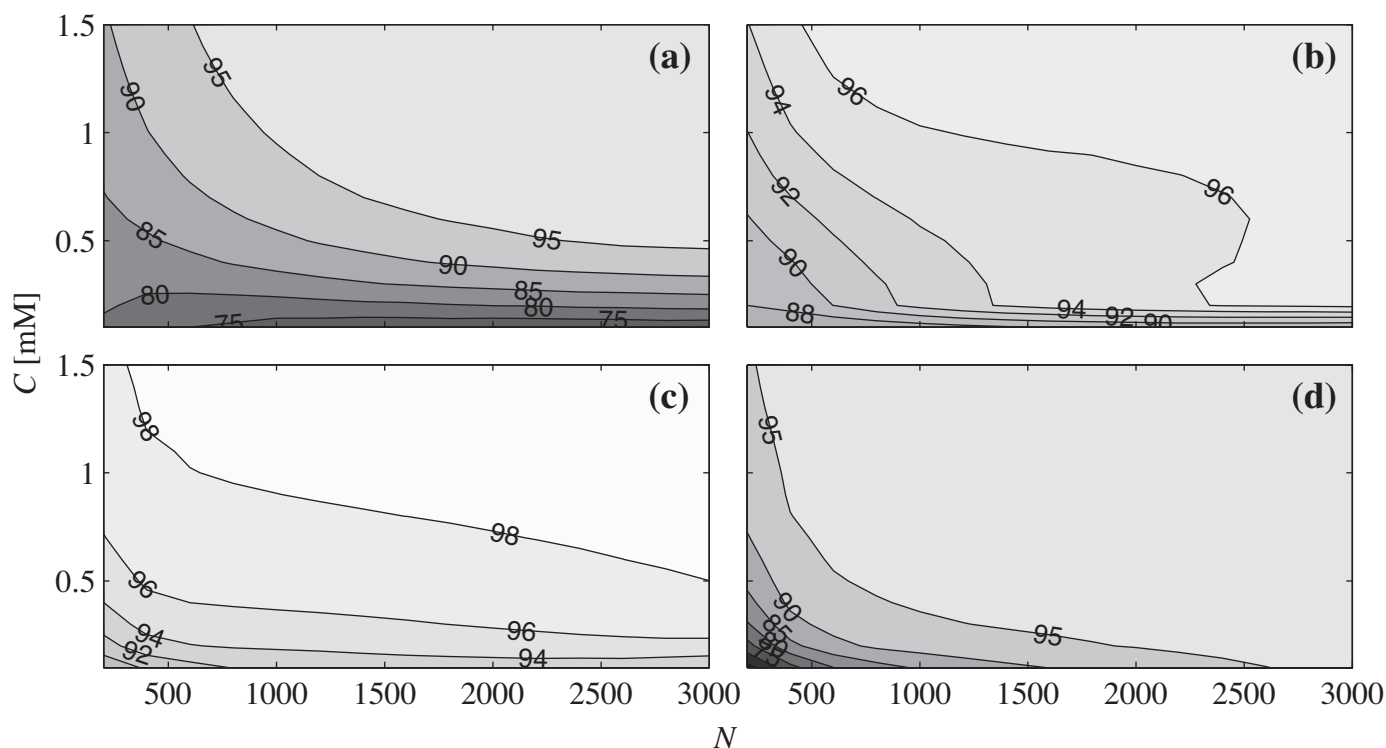


Fig. 1. Overlap contour surface of the RFD and CDM algorithm used for quaternary simulations of components 1–4 with the adsorption isotherm parameters listed in Table 1. Each overlap is calculated for the individual components 1 (a) to 4 (d) and presented from 0 to 100%. Identical column efficiency is used for all components regardless of algorithm and 60 μ L of the sample is injected. The overlap contour surface consists of a grid of 225 points where the injection concentration is between 0.1 mM and 1.5 mM per component and column efficiency is between 200 and 3000 theoretical plates.

efficiency. The worst case scenario is when the first component has over 1000–3000 theoretical plates and the second is below 1000. For the cases of 1000 to 3000 theoretical plates, the solutions for both components are very similar. Base line separated species show very little sensitivity to either combination, but shows similar trends as co-eluting species.

4.3. Binary simulations: impact of injection concentrations

We have made some more detailed observations regarding needed precaution when simulating different injection concentration ratios for a binary separation, i.e., effectively simulating potentially steep concentration gradients between components. We have investigated combinations of injection concentrations ranging from 0.5 to 1.5 mM in an equally spaced grid of 121 points, at fixed column efficiencies of 400 and 2000 theoretical plates. Each pair of simulated elution profiles obtained by the RFD and CDM algorithm were then compared.

Yet again, we observe different behavior of the earliest eluting component versus the later eluting component, as well as minor differences between co-eluting and base-line separated components. When the efficiency is set to 2000, the solutions obtained for two co-eluting components using either the RFD or CDM algorithm are practically identical for even the extremes of ratios, see Fig. 2a and b. For an efficiency of 400, the overlap is poor for the fastest eluting component, see Fig. 2c. However the overlap is rather uniformly dependent on concentration ratios, indicating that low efficiency is the main cause for poor overlap. Interestingly, the difference in solutions for the second component are practically the same as for 2000 theoretical plates, compare Fig. 2b and d.

For base-line separated components, the average of the overlap of the second eluting component is higher than for the first

eluting component and almost independent of injection ratio at either low or high efficiency, see Fig. 4. At high efficiency, the first eluting component is clearly dependent on the injection concentration ratio, see Fig. 3a, but still relatively insensitive. For low efficiencies, we find similarities with the simulations performed with components 1 and 2 Fig. 2c and d, but overall the overlap is lower.

4.4. Binary simulations: impact of the number of components and combination of retention factors

Different combinations of simulated unary, binary, ternary and quaternary separations were performed to investigate how the combination of relative retention between the simulated components effects the difference in obtained solutions by the RFD and CDM algorithm.

To do this, we calculated the overlap of the first eluting component when simulated alone and then in the presence of additional components. The best overlap given a certain column efficiency and at a fixed column load is obtained, as expected, for single component simulation. Adding additional components with continuously increasing retention factors continuously decreases the overlap. This confirms the correlation of decreasing accuracy of solutions to increasing numbers of simulated components. However, adding additional components with mixed retention factors to the simulation do not lead to a continuous decrease in overlap, see Fig. 4. This is not surprising, considering that the RFD algorithm discretization is based on average effective traveling velocity of the simulated components, the larger the relative difference in retention factors, the greater the error. The point to be made is that it is the relative difference of retention factors of the simulated components that is important, beside that of the actual

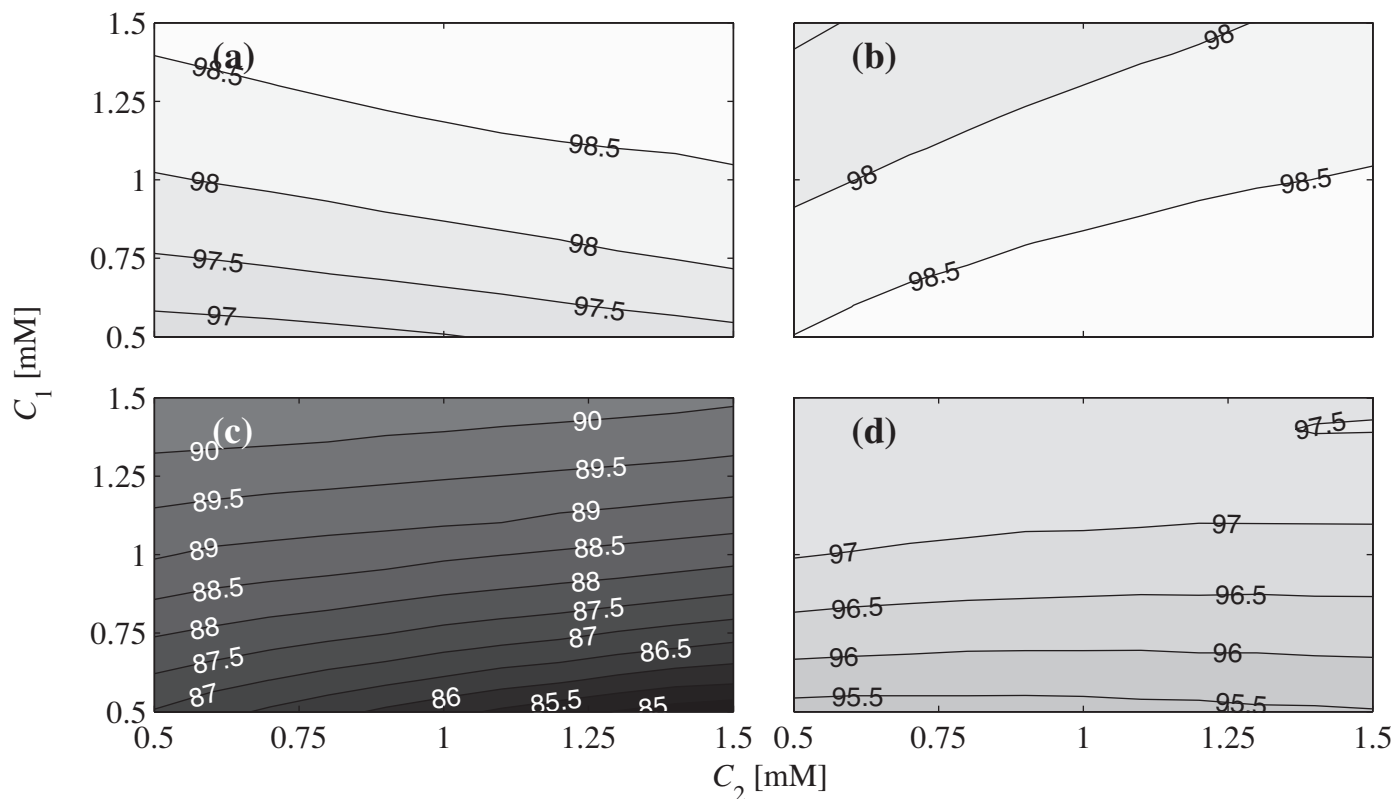


Fig. 2. Overlap contour surface of the RFD and CDM algorithm used for binary simulations of components 1 and 2 with adsorption isotherm parameters listed in Table 1 and 60 μ L of the sample is injected. Each overlap is calculated for the individual components. The overlap contour surface consists of a grid of 121 points where C_1 and C_2 vary from 0.5 to 1.5 mM, i.e., we investigate component ratios between 1:3 and 3:1. (a, b) Components 1 and 2 at $N=2000$ and (c, d) components 1 and 2 at $N=400$.

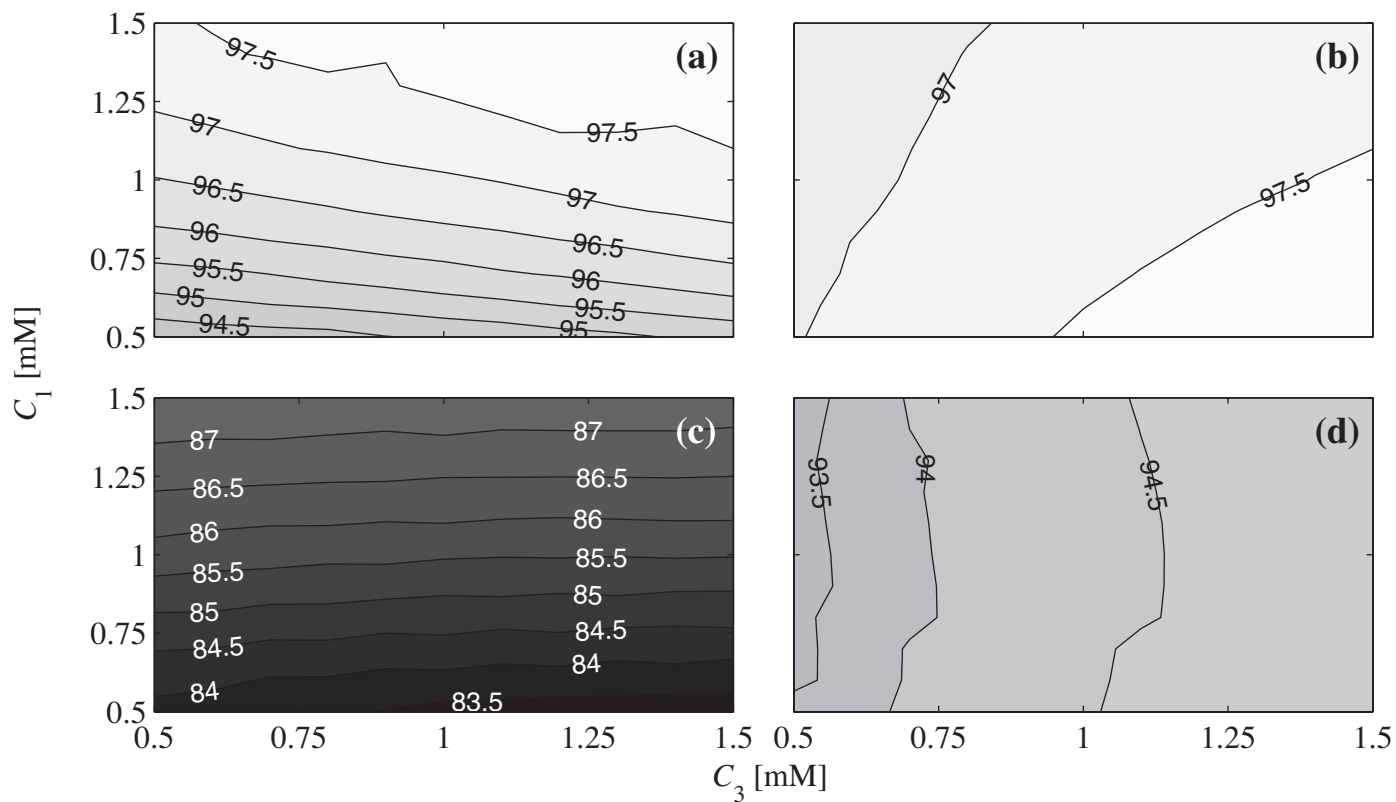


Fig. 3. Overlap contour surface of the RFD and CDM algorithm used for binary simulations of components 1 and 3 with adsorption isotherm parameters listed in Table 1 and 60 μ L of the sample is injected. Each overlap is calculated for the individual components. The overlap contour surface consists of a grid of 121 points where C_1 and C_3 vary from 0.5 to 1.5 mM, i.e., we investigate component ratios between 1:3 and 3:1. (a, b) shows components 1 and 3 at $N=2000$ and (c, d) components 1 and 3 at $N=400$.

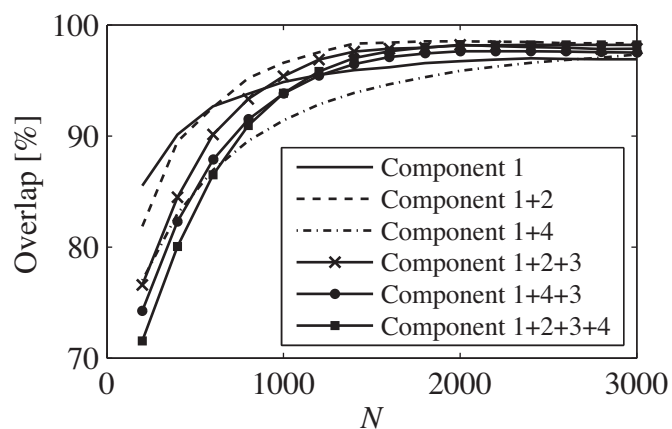


Fig. 4. Overlap between the RFD and the CDM solutions for single, binary, ternary and quaternary simulation of components 1 and 4 using the adsorption isotherm parameters listed in Table 1. Simulated injection concentration is 1 mM per component and injection volume of 60 μ l. Each overlap is calculated for the earliest eluting component (1). The simulations are performed in the following order: component 1 (solid line), components 1 and 2 (dashed line), components 1 and 4 (dashed-dotted line), components 1–3 (solid-crossed line), components 1, 4 and 2 (solid-circled line) and components 1–4 (solid-squared line).

number of simulated components as well as column efficiency and load. We have also showed that considerably higher efficiency than 1000 is required when simulating two components with markedly different retention.

4.5. Experimental validation of findings

Experimental comparison of the simulated and experimental profiles for the quaternary, high efficiency in Section 3.1 show almost identical results when uses either the RFD, CDM or OCFE algorithm, see Fig. 5. Parameters found describing the bi-Langmuir adsorption isotherm are, within error margins, almost similar for the three algorithms, see Table 1. Correlating this finding to the sim-

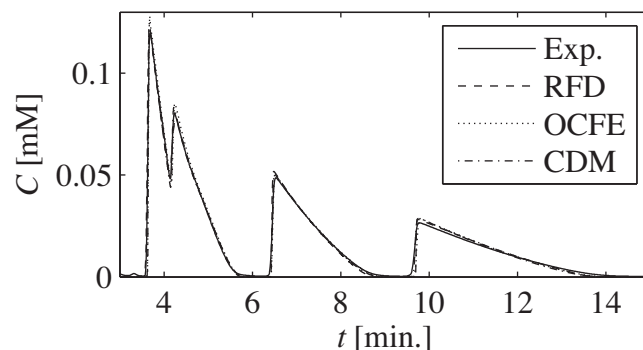


Fig. 5. Experimental and simulated chromatograms of quaternary racemic methyl- and ethylmandelate mixture using the adsorption isotherm parameters listed in Table 1 and the RFD, OCFE or CDM algorithm. The experimental chromatogram shows a quaternary injection of 60 μ l 0.6 mM *S*(+)/*R*(-)-methyl-mandelate and *S*(+)/*R*(-)-ethyl mandelate. Experimental data is solid line, RFD simulated (dashed), OCFE (dotted), CDM (dash-dotted). Notice, that these components are referred to as components 1–4 in the synthetic study.

ulations at high efficiency, we expect the RFD and CDM algorithm to produce accurate, but not necessarily identical, solutions of Eq. (1), provided sufficient column load and efficiency, and this is what we observe, see Fig. 5. This experimental validation can be compared to the overlap contour surface in numeric study in Fig. 1 at $N = 2500$ and $C = 0.6$ mM. Here the degree of overlap of RFD to CDM is around 95–97%, which is also what we find comparing the three simulated chromatograms to the experimental elution profiles.

For the binary, low efficiency, enantiomeric separation, defined in Section 3.2, the results are different. If we use the inverse method with the RFD, CDM or OCFE algorithms, they will have almost the same degree of overlap with the experimental profiles at higher column load, see Fig. 6c and d, but not at lower load, see Fig. 6a and b. This is what could be expected from the synthetic studies. Despite the low efficiency at a certain high enough column load, the overlap will be almost identical, but if we decrease the column load we will encounter conditions where the simulations differ. To our

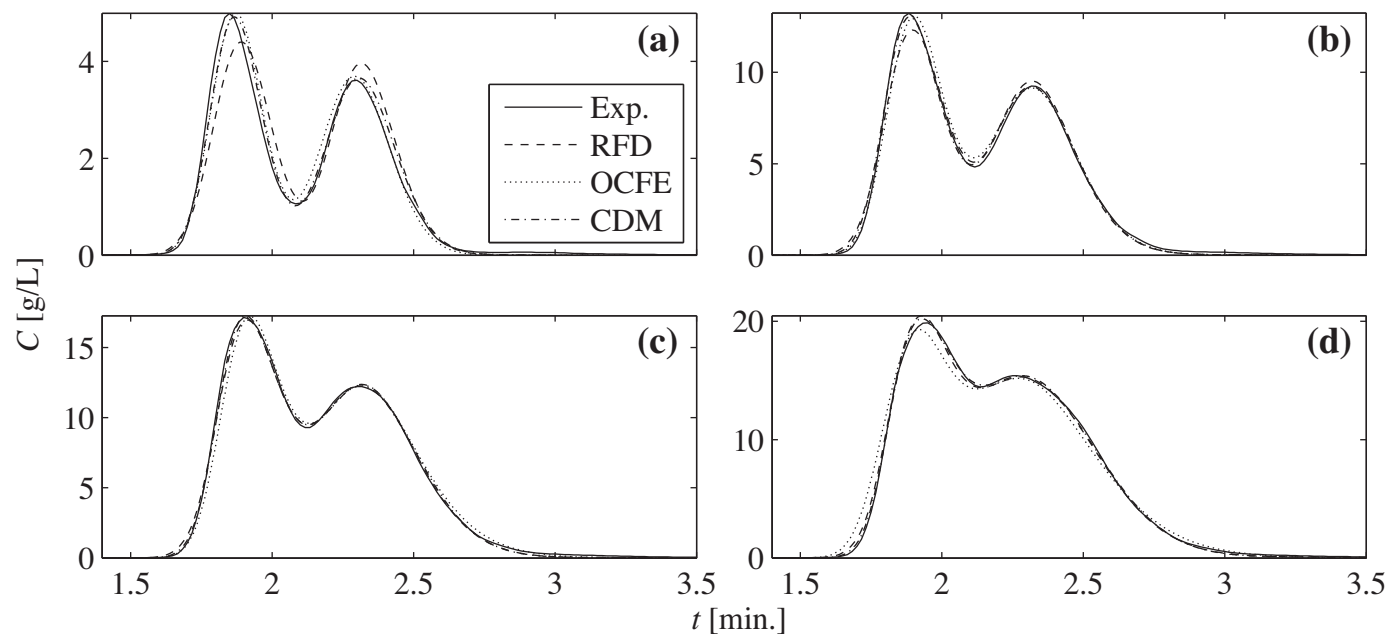


Fig. 6. The chromatogram shows overlay of experimental and simulated chromatograms of the “AstraZeneca” binary racemic drug intermediate. The inverse method was used with RFD, OCFE and CDM. The experimental chromatogram shows 100–600 μ l of 47 g/L racemic injections. Experimental data is solid line, RFD simulated (dashed), OCFE (dotted), CDM (dash-dotted). Binary isotherm parameters are listed in Table 1.

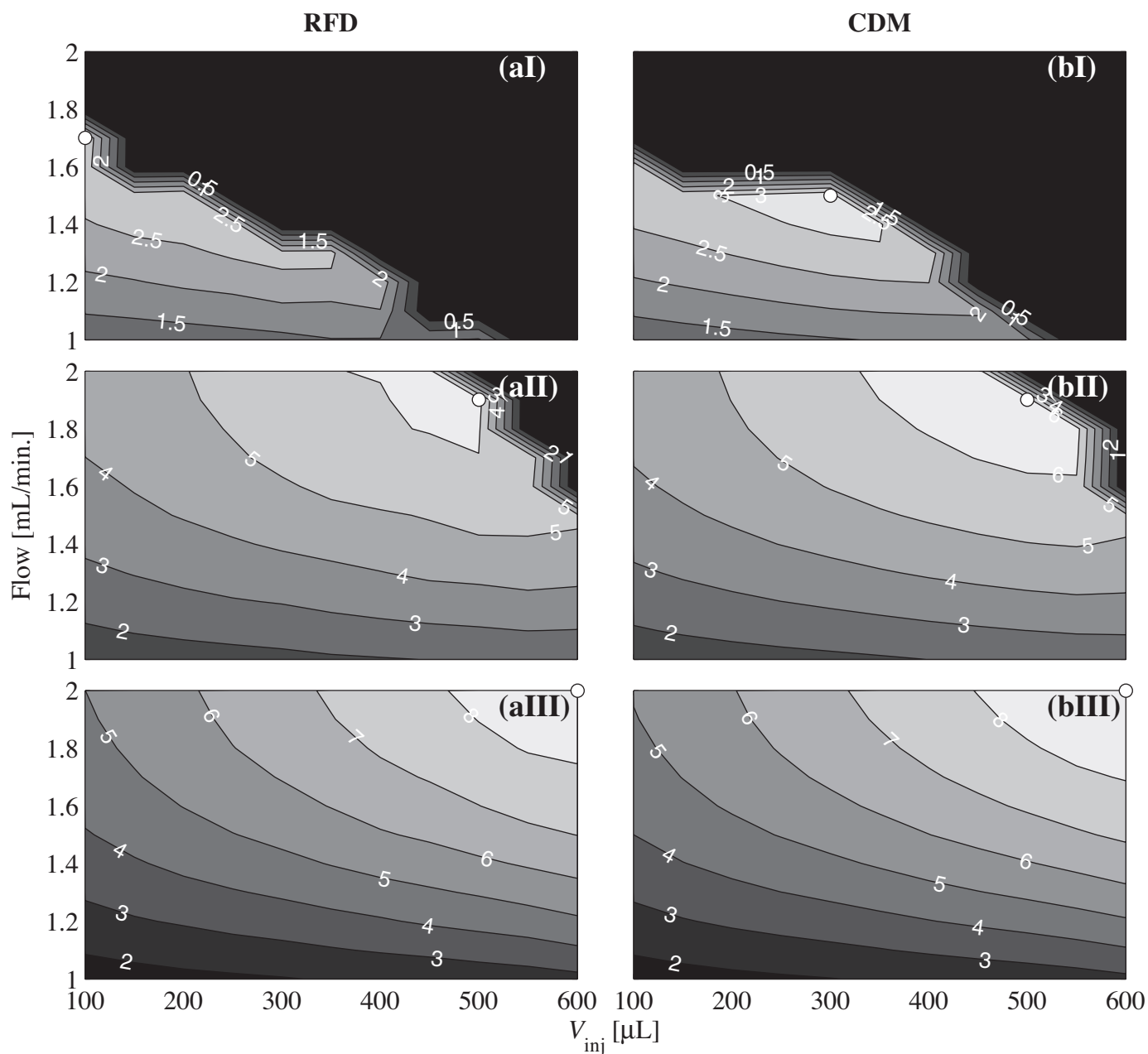


Fig. 7. Contour surfaces for maximum production rate of the earliest eluting component in the binary experiment described in Section 3.2. Production rates [mg/min] are calculated for injections between 100 and 600 μL at flows between 1 and 2 mL/min on a 11×11 grid. Constraints of >95% purity and >75% were used. Black color indicates an area where either purity or yield requirements could not be met for any cut-point. (a) Production rate calculated by RFD and (b) production rate calculated by CDM. In (I) for $100 < N < 200$, in (II) for $400 < N < 800$ and in (III) for $800 < N < 1600$.

knowledge, for the first time, these have been shown with actual experimental data.

Differences between the three algorithms are also shown by the obtained adsorption isotherm parameters. It seems as the error in RFD algorithm to degree of overlap, as we have observed in the synthetic study, is compensated by error in adsorption isotherm parameters. Using the acquired parameters for RFD or CDM/OCFE and simulating using the other algorithm, gives lower degree of overlap, confirming this suspicion.

4.6. Calculation of optimal production rate for a binary separation

Here we optimize the flow rate and injection volume to maximize the production rate for a binary separation based on the

experimental parameters in Section 3.2. The common objectives are that the first eluted component should be collected in excess of 95% purity and the yield should be at least 75%. Using the measured efficiency we found negligible differences between the predictions of maximum productivity of RFD and CDM, see Fig. 7a and b. At half the efficiency, the predictions are also quite similar. At a quarter of the measured efficiency, RFD and CDM both predict markedly different maximum production rate and also different regions where the constraints are fulfilled. These findings are in line with our earlier results.

5. Conclusions

In this study, over 1000 simulated synthetic experiments have been performed and presented to compare the solutions of the

Rouchon finite difference method (RFD) to those of the easily implemented and more accurate algorithm, the central finite difference method (CDM), and the more complicated Orthogonal Collocation on Finite Element (OCFE) method. Overlap contour surfaces have been presented to show where precaution might be needed when choosing between algorithms, but caution must be taken not to draw to general conclusions. OCFE is considered the most accurate algorithm available and was used as a reference solution.

All our studies have shown that the single most important parameter, always leading to different solutions to Eq. (1), when using the RFD or CDM algorithm, is the apparent column efficiency N . Considering that we in most cases find that the CDM and RFD solutions differs when $N < 1000$ theoretical plates, we can also confirm the rule of thumb that not to use RFD in these cases. However, as we also have demonstrated, this should in no way be considered a stringent requirement for accuracy, as we have presented many cases where twice that value is required. Together with the findings in Ref. [5], this concludes us to state that RFD should be avoided and no real certainty of its correctness can be guaranteed unless one compares the solutions with more accurate algorithms.

In more detail, we have found the correlation of column load with efficiency to have a great impact on the obtained solutions of Eq. (1). For certain combinations of column efficiency and load, effectively the degree of column overload, the solutions markedly differ. For a fixed column load, increasing the efficiency generally leads to higher degree of overlap. For a fixed efficiency, increasing the column load generally leads to higher degree of overlap. Considering the simulation of a single component, this behavior is indicated by the Shirazi number [1]. As the number indicates that when either the load or efficiency is increased, the solution of the ED-model approaches that of the ideal model, the RFD algorithm is expected to give good results. However, for multiple components, the Shirazi number is not defined. Our simulations of four components shows that if the column load is small enough but still in the nonlinear region of the adsorption isotherm it is possible to increase the efficiency without noticeably improving overlap at all, see Fig. 1. This is an important new finding.

Our main observations are the following: difference between the RFD and CDM solution was found to increase,

- (i) With decreasing column efficiency and load.
- (ii) When the efficiency difference between the two components increases, as long as one of the components has less than at least 1000 theoretical plates. In particular in the case when one has partially co-eluting components, whereas the solutions for baseline resolved components are less sensitive.
- (iii) As the sample composition deviates from 1:1 ratio, the solution for the first eluting component is more sensitive than for the second eluting component in the case of a binary separation.
- (iv) When the number of components is increased, each with an increased retention factor. The divergence does not seem to increase noticeably if further components with intermediate retention factors are included.

CDM solutions were found to be identical to those obtained by OCFE, however more calculation time was required. Our results show that RFD yields inaccurate solutions under a number of circumstances when compared to CDM/OCFE. We therefore recommend implementation of OCFE for any practical purposes where accurate predictions are needed. Although it is much easier to implement compared to OCFE, CDM is not recommended solely based on its slowness. Our results indicate that RFD cannot be used with confidence for as low as 1000 or even 3000 theoretical plates or not at all regardless of plate number, when considering Langmuir-like adsorption isotherms for a quaternary simulation. This is in line with but does not validate the previous conclu-

sion that $N \sim 10,000$ is required for RFD ternary calculations. The results presented by Kaczmarski [5], regarding simulations where two components are injected and the third is a modifier in the mobile phase is a special, but none the less important case, of the ternary problem. This effect was probably due to the fact that the third component was in the eluent causing so called additive deformations of the binary elution profiles; it has recently been demonstrated how sensitive these effects are to the number of plates [18]. A more thorough study would clarify the problems with additive deformations. Nevertheless, the results in [5] is yet another indication of that the rule of thumb of employing at least 1000 theoretical plates for the RFD algorithm is largely underestimated.

When it comes to using the IM for process optimization, the RFD inaccuracy may not always be an issue. Sufficiently good predictions may be obtained although that the solution is inaccurate, since the IM will converge to the best possible overlap with experimental chromatograms (given the choice of column model, PDE algorithm and data quality) which is not necessarily equivalent with the true adsorption isotherm parameters. In this study the three algorithms resulted in similar but not identical adsorption isotherm parameters when using the IM on a high-efficiency experimental system, $N_{\text{avg}} = 2500$, whereas quite different sets of parameters were obtained for a low-efficiency experimental system $N_{\text{avg}} = 395$, where even the number of Langmuir sites are estimated differently. The IM solutions corresponding to the three algorithms all gave similar solutions at high column load but clearly distinguishable different simulations at low column load. At low column load, solution of the OCFE and CDM algorithm clearly has a higher overlap with experimental data, than then one that is obtained by the RFD algorithm. It is of particular importance that the overlap is satisfactory in the co-elution zone(s), so that the process chromatographer can determine when to collect the pure fraction(s) based on simulations. Provided the experimental data has sufficient quality, the column model and adsorption isotherm model have been correctly chosen, our data indicate that it is reasonable to assume that in general, OCFE and CDM should provide better overlap and thus better process predictions than RFD. We have demonstrated that the prediction of optimal production rate for a binary simulation is clearly dependent on the choice of algorithm at low column efficiency, as predictions are clearly different, but it remains to be explicitly demonstrated by performing experiments. Because of the earlier studies on other type of adsorption isotherms, for example the complete failure of RFD to simulate a BET type adsorption isotherm [5], it also remains to in detail study the type II–IV isotherms. However, in these cases, the particular plate number and column load may be irrelevant to explain the major cause of error.

Acknowledgements

This work was supported by a grant from The Research Council for Environment, Agricultural Sciences and Spatial Planning for the project “High-Value Compounds from Agricultural and Forestry Waste by Sustainable Methods- an Interdisciplinary Approach for Bioresource Utilization” and by a grant from the Swedish Research Council (VR) for the project “Fundamental Studies on Molecular Interactions Aimed at Preparative Separations and Biospecific Measurements.

References

- [1] G. Guiochon, A. Felinger, D.G. Shirazi, A.M. Katti, *Fundamentals of Preparative and Nonlinear Chromatography*, 2nd ed., Elsevier Academic Press, 2006.
- [2] K. Kaczmarski, D. Antos, H. Sajonz, P. Sajonz, G. Guiochon, *J. Chromatogr.*, A 925 (2001) 1.

- [3] D. Antos, K. Kaczmarski, W. Piatkowski, A. Seidel-Morgenstern, J. Chromatogr., A 1006 (2003) 61.
- [4] Z. Ma, G. Guiochon, Comput. Chem. Eng. 15 (1991) 415.
- [5] K. Kaczmarski, J. Chromatogr., A 1176 (2007) 57.
- [6] P. Forssén, R. Arnell, T. Fornstedt, Comput. Chem. Eng. 30 (2006) 1381.
- [7] G. Guiochon, B. Lin, Modeling for Preparative Chromatography, Academic Press Elsevier, 2003.
- [8] P. Rouchon, M. Schonauer, P. Valentin, G. Guiochon, Sep. Sci. Technol. 22 (1987) 1793.
- [9] J. Villadsen, M.L. Michelsen, Solutions of Differential Equation Models by Polynomial Approximation, Prentice-Hall, New York, 1978.
- [10] K. Kaczmarski, M. Mazzotti, G. Storti, M. Morbidelli, Comput. Chem. Eng. 21 (1997) 641.
- [11] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen, LAPACK Users' Guide, 3rd ed., Society for Industrial and Applied Mathematics, Philadelphia, 1999. ISBN 0-89871r-r447-8.
- [12] P.N. Brown, G.D. Byrne, A.C. Hindmarsh, SIAM J. Sci. Stat. Comput. 10 (1989) 1038.
- [13] A. Seidel-Morgenstern, J. Chromatogr., A 1037 (2002) 255.
- [14] Felinger, A. Cavazzini, G. Guiochon, J. Chromatogr., A 986 (2003) 207.
- [15] Y. Zhang, S. Rohani, A.K. Ray, J. Chromatogr., A 1131 (2006) 176.
- [16] Y. Zhang, S. Rohani, A.K. Ray, J. Chromatogr., A 1202 (2008) 34.
- [17] J. Lindholm, P. Forssen, T. Fornstedt, Anal. Chem. 76 (2004) 5472.
- [18] P. Forssén, R. Arnell, T. Fornstedt, J. Chromatogr., A 1216 (2009) 4719.