Journal of Chromatography A, 1218 (2011) 6009–6014

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Probabilistic model for immiscible separations and extractions (ProMISE)

Joost de Folter*, Ian A. Sutherland

Brunel Institute for Bioengineering, Brunel University, Kingston Lane, Uxbridge, Middlesex UB8 3PH, UK

ARTICLE INFO

Article history:
Available online 23 December 2010

Keywords:
Counter-Current Chromatography
CCC
Modelling
Simulation
Probabilistic
Bayes
ProMISE

ABSTRACT

Chromatography models, liquid–liquid models and specifically Counter-Current Chromatography (CCC) models are usually either iterative, or provide a final solution for peak elution. This paper describes providing a better model by finding a more elemental solution. A completely new model has been developed based on simulating probabilistic units. This model has been labelled ProMISE (probabilistic model for immiscible phase separations and extractions), and has been realised in the form of a computer application, interactively visualising the behaviour of the units in the CCC process. It does not use compartments or cells like in the Craig based models, nor is it based on diffusion theory. With this new model, all the CCC flow modes can be accurately predicted. The main advantage over the previously developed model, is that it does not require a somewhat arbitrary number of steps or theoretical plates, and instead uses an efficiency factor. Furthermore, since this model is not based on compartments or cells like the Craig model, and is therefore not limited to a compartment or cell nature, it allows for an even greater flexibility.

1. Introduction

Chromatography models, liquid–liquid models and specifically Counter-Current Chromatography (CCC) models are usually either iterative [1–5], or provide a final solution for peak elution [3,6,7]. In the former, the column is typically divided into discrete compartments, resembling the test tube based counter-current distribution (CCD) process [8]. Sample components are distributed between the phases, mixed, and then transferred to the next compartment, each according to its K value (distribution coefficient) and so on. This is an iterative process until the peaks have moved beyond the column (eluted out). Under certain conditions, a single equation describing the eluted peaks can be formulated, but always includes a factorial, which in modelling terms, is an iterative operation. The advantage of an iterative model is that the complete chromatography process is described including the time spent inside the column. Other models are usually based on mass transfer or diffusion theory, also referred to as rate models. These models consist of solving one or more differential equations, and providing a solution only describing the final outcome.

However CCD, as a discreet process, remains fundamentally different from CCC which is a continuous process. Having a single mathematical solution is not necessary, and in fact an iterative solution gives additional advantage.

The aim of this research is finding an iterative model that better describes the CCC process. This paper describes providing a better model by finding a more elemental solution.

2. Theory

In 1941, Martin and Synge [1] showed that CCD can be described using a binomial solution. This is also known as the probability mass function which describes the probability of getting exactly r successes in n trials:

$$\binom{n}{r} p^r (1-p)^{n-r} \tag{1}$$

p is the probability of each trial (being 0 or 1), and $r = 0, 1, 2, \ldots, n$, where the binomial coefficient is defined as

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \tag{2}$$

The probability mass function is based on Bayes' theorem, which shows the relation between two conditional probabilities which are the reverse of each other [9]. Using a more elemental approach, a simple model can be developed.

3. Model

3.1. Concept

This model has been labelled ProMISE (probabilistic model for immiscible phase separations and extractions). Considering a

---

* Corresponding author. Tel.: +44 01895 266856; fax: +44 01895 274608.
E-mail address: joost.defolter@brunel.ac.uk (J. de Folter).

0021-9673/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2010.12.079

molecule of a particular compound in a two phase system, assuming it is located somewhere in either phase, its behaviour can be described by the probability of it moving to the other phase. This probability is then simply

$$\frac{1}{KX+1} \quad \text{or} \quad \frac{KX}{KX+1} \tag{3a,b}$$

depending on which phase the unit is in, where K is the distribution coefficient:

$$K = \frac{C_U}{C_L} \tag{4}$$

with C_U and C_L the sample concentrations of the upper and lower phase respectively, and for the current definition of K (Eq. (4)). X is defined as the phase distribution:

$$X = \frac{U_F}{L_F} \tag{5}$$

where U_F is the proportion of the column volume occupied by the upper phase and L_F is the proportion of the column occupied by the lower phase. Furthermore an efficiency factor is introduced here by simply multiplying by the efficiency factor for the final probability value.

This behaviour is followed regardless of other molecules of the same compound, e.g. (local) compound concentration in the phases. The model consists of simulating many of these representative units, following the probabilistic rules described here, where each unit represents a very small amount of sample compound. Each unit is then moved according to the movement of the phase it is located in.

Because the model is based on compound units, a density function with an adaptive Gaussian filter is used to convert the separate unit values into a chromatogram. The nature of this model allows its internal values to be volume or time.

3.2. Output

The model output consists of a number of units, each having a position value. The main peak values can be directly obtained from the model output. The peak position is equal to the mathematical average of the weighted units:

$$\mu = \frac{1}{m_{tot}} \sum_{i=0}^{n} m_i x_i \tag{6}$$

where x_i is the position value of each unit index i, out of n total units. m_i represents the (relative) weight of each unit, and m_{tot} the total weight of all units. Note that the peak average does not necessarily coincide with the peak maximum (in case of asymmetrical peaks).

The peak width is subsequently obtained by taking the standard deviation from the units:

$$\sigma = \sqrt{\frac{1}{m_{tot}} \sum_{i=0}^{n} m_i (x_i - \mu)^2} \tag{7}$$

Using this theory, the compounds naturally distribute according to their K value.

The retention times can therefore be predicted using standard theory. Expanding the equations determined in previous research [4,5], the peak width for (normal flow mode) can be calculated as follows:

$$\sigma = \sqrt{t_R \frac{KX}{\omega}} \tag{8}$$

where ω is the rotational speed giving the number of mixing/settling steps per unit time, and t_R the retention time. The peak width is equal to 4σ.

The mixing/settling efficiency is taken into account in the model by multiplying the probabilities with the efficiency factor f. Accordingly the peak width can be calculated by modifying Eq. (8):

$$\sigma = \sqrt{t_R \frac{KX}{\omega f'}} \tag{9}$$

where f' represents the efficiency effect.

However in reality this is a simplified equation for normal flow mode. In the same way, a more general equation incorporating different flow modes can be expanded from previous research [5]:

$$\sigma = \frac{\sqrt{t_R (KX/\omega f')}}{X_{FL} + KXX_{FU}} \tag{10}$$

where X_{FU} and X_{FL} are the normalised upper and lower phase flow rates:

$$X_{FU} = \frac{F_U}{U_F} \quad \text{and} \quad X_{FL} = \frac{F_L}{L_F} \tag{11a,b}$$

where F_U and F_L are the upper and lower phase flow rates. X_{FU} and X_{FL} are further normalised dividing by the maximum of (X_{FU}, X_{FL}) to give dimensionless values between 0 and 1. This set of equations for peak width appears to work well in most cases, though a general equation that perfectly satisfies all flow modes correctly has not yet been found.

3.3. Efficiency

The efficiency of the mixing/settling in the model is determined by an efficiency factor f. However, it was found that this efficiency factor could not be directly incorporated into the equations for the peak shape (Eq. (9)). The relationship between the model mixing/settling efficiency and resulting efficiency effect was obtained by using Eq. (9) to determine f' (Fig. 1a).

Using curve fitting techniques, the relationship between the initial efficiency factor f and the resulting efficiency effect f' was empirically found to approximate (Fig. 1b) as follows:

$$f' = \frac{f}{2^{1-f}} \tag{12}$$

This efficiency factor actually represents a composite factor, not only for the mixing/settling efficiency but also, for the CCC system in general.

4. Results

This new model has been realised in the form of a computer application, interactively visualising the behaviour of the units in the CCC process (see Fig. 2).

Results from this new model show good correlation with the currently tested operation modes: conventional, co-current and dual flow. Theoretical results were obtained by applying the standard predictive equations for peak retention [10], and peak width (Eqs. (9) and (10)). All result values are in volume units, hence instead of T_R retention is labelled V_R.

For each of the flow modes described in the following sections, the model was set up and calibrated according to the experimental set up. This includes the coil volume, the rotational speed of the CCC apparatus, the stationary phase volume retention, the mobile phase flow rates and the K-values of the components. Furthermore the model was calibrated for each experimental condition, finding an effective efficiency.

Because of its nature, the model can be set up more accurately than the CCD based model [5]. Because of the limitation of the number of cells of the CCD based model, an effective smaller number of cells had to be used that incorporated the efficiency. So in each mixing/settling step, an efficiency of 1 (100%) was used. The new