



Predicting sample injection profiles in liquid chromatography: A modelling approach based on residence time distributions

Monica Tirapelle^{a,*}, Maximilian O. Besenhard^a, Luca Mazzei^a, Jinsheng Zhou^b, Scott A. Hartzell^b, Eva Sorensen^a

^a Department of Chemical Engineering, University College London, Torrington Place, London, WC1E 7JE, UK

^b Eli Lilly and Company, 893 Delaware St, Indianapolis, 46225, USA

ARTICLE INFO

Keywords:

Injection profile
Liquid chromatography
Analytical chromatography
Residence time distribution theory
In-silico optimization

ABSTRACT

The pharmaceutical and bio-pharmaceutical industries rely on simulations of liquid chromatographic processes for method development and to reduce experimental cost. The use of incorrect injection profiles as inlet boundary condition for these simulations may, however, lead to inaccurate results. This study presents a novel modelling approach for accurate prediction of injection profiles for liquid chromatographic columns. The model uses the residence time distribution theory and accounts for the residence time of the sample through the injection loop, connecting tubes and heat exchangers that exist upstream of the actual chromatographic column, between the injection point and the column inlet. To validate the model, we compare simulation results with experimental injection profiles taken from the literature for 20 operating conditions. The average errors in the predictions of the mean and variance of the injection profiles result to be 8.98% and 8.52%, respectively. The model, which is based on fundamental equations and actual hardware details, accurately predicts the injection profile for a range of sample volumes and sample loop-filling levels without the need of calibration. The proposed modelling approach can help to improve the quality of in-silico simulation and optimization for analytical chromatography.

1. Introduction

A significant proportion of drug manufacturing costs is incurred during downstream purification, where reversed-phase liquid chromatography is a major process [25]. To reduce costs and experimental effort, pharmaceutical companies employ modelling and simulation of liquid chromatography to support method development. This is done by solving a mathematical transport model of the chromatographic separation, such as the Equilibrium-Dispersive Model [13] or the General Rate Model [16], together with an equilibrium isotherm describing the adsorption of solutes onto the stationary phase. In order to solve the model equations, suitable expressions for the initial conditions and the boundary conditions are required. The initial conditions, describing the starting point of the simulation (e.g., an empty column), are normally well established, as is the outlet boundary condition for the exit from the column. The selection of an appropriate sample injection profile as the inlet boundary condition to the column, however, is not trivial, although is of fundamental importance for obtaining accurate chromatographic peak shape predictions. Accurate prediction of the peak

shapes is critical, especially when these are used to derive information about equilibrium isotherm parameters and kinetic data that are later used in method development.

The most commonly employed inlet boundary condition is the simple rectangular boundary condition, whereby the sample injection profile is modelled as a rectangular pulse with height C_0 and width t_p , where C_0 and t_p represent the concentration of the analyte in the sample vial and the feeding time, respectively [7,16,26,14]. To prevent numerical instabilities that might arise from the vertical boundaries of a rectangular shape, the rectangular profile can be flanked by two semi-Gaussian curves [23]. Rectangular injection profiles are usually employed because of their simplicity, however, they are not realistic. Because of axial dispersion of the solutes in the elements or units upstream of the column, for instance, tubing or a heat exchanger (see Fig. 1), true injection profiles actually deviate strongly from the rectangular shape, in fact, they are asymmetric, with a sharp front followed by a tailing decay [16,26,11,18].

To improve the accuracy of simulations, some studies have employed experimentally determined injection profiles. These are mea-

* Corresponding author.

E-mail address: m.tirapelle@ucl.ac.uk (M. Tirapelle).

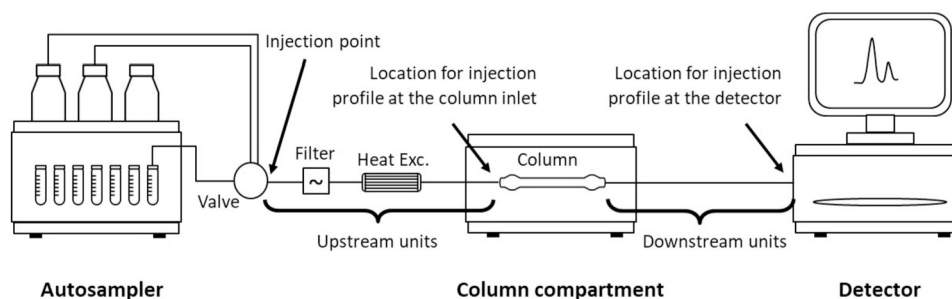


Fig. 1. Schematic representation of a liquid chromatographic system. Our modelling approach for predicting the injection profile into a liquid chromatographic column considers the residence time distributions (RTDs) of the units placed upstream of the column (e.g., connecting tubes and heat exchangers) from the injection point to the column inlet. However, when a direct comparison with experimental injection profile is required, the RTDs of downstream units must also be included.

sured by injecting a tracer into the chromatographic system with the column replaced by a zero-dead volume connector [29]. This procedure is, however, impractical for in-silico optimization since the injection profiles vary with flow rate, injection volume, mobile phase viscosity, solute molecular weight and extra-column volumes [26,11], hence the injection experiments would need to be repeated if any of these parameters are to be changed.

As alternatives to experiments, the literature reports several computational fluid dynamics simulations [8,9], as well as empirical and semi-empirical models for predicting injection profiles. These models vary in terms of accuracy, number of fitting parameters and experimental effort required. In the case of small sample volumes, the simplest and most extensively used models are the half-Gaussian [24], the quasi-Gaussian [17] and the exponentially modified Gaussian (EMG) models [6]. The latter results from the convolution of a Gaussian peak with an exponential decay function, where the Gaussian contribution accounts for the Gaussian-type band broadening in connecting tubes caused by axial dispersion, and the exponential decay function models the exponential tailing caused either by mixer-type extra volumes or by strong axial dispersion [20]. None of these models can, however, predict the concentration plateau that is known to arise at large sample volumes [7,11].

To account for sample volume, Felinger et al. [7] proposed combining the EMG function with a rectangular pulse of width t_p , this being the feeding time of the equivalent rectangular shape (i.e., the ratio between sample volume and flow rate). This model can predict the sample injection profile under different operational conditions, however, it relies on several fitting parameters that have to be calibrated against experimental results. Note that this approach is empirical, it does not offer any physical insight, and the parameters therefore do not have any physical interpretation.

A similar model was developed by Forssén et al. [11]. In their work, a Gaussian function was combined with a square pulse followed by an exponential decay. They also derived linear relationships to express the model parameters as functions of injection volume and flow rate. Even though this model reduces the experimental effort required, it is still empirical and the parameters again have no physical meaning.

Weatherbee et al. [30] used the Forssén model as a starting point for developing what they referred to as a *global* model. In their work, they considered a two-dimensional liquid chromatography (2D-LC) system and focused on predicting the shape of the injection profile at the inlet of the second column by estimating five parameters over a number of experimental conditions. Then, they developed empirical functions that related the parameter values to flow rate and loop volume. Unlike in the work by Forssén et al. [11], Weatherbee et al. [30] found non-linear dependencies.

A slightly different approach was developed by Gritti et al. [14], and later implemented by Pepermans et al. [23]. In their work, they described the injection profile as an asymmetric double sigmoidal function with two empirical time parameters that have to be tuned against experimental data [14].

All the models presented so far, and summarized in Table 1, are empirical or semi-empirical and do not consider the hardware details of the system. Their applicability will therefore depend on the specific chromatographic system in question. To the authors' knowledge, the only fundamental investigation of the injection profile in liquid chromatography so far was made by Samuelsson et al. [26]. In their work, they employed a 2D-convection-diffusion differential equation with cylindrical coordinates and the equation for the parabolic velocity profile to investigate the dependence of the injection profile on different experimental parameters. The same model was successively employed by Baran et al. [1] to predict band broadening and band deformation of slow diffusing macromolecules. The model, which enables accounting for the parabolic velocity profile and the radial diffusion that arises in the injection loop and in all the capillaries connecting the injector to the column, led to highly accurate predictions and fundamental knowledge acquisition on the dispersion mechanisms involved. However, every new flow rate, sample volume, filling level, changes in tubing etc. will require a new simulation. As argued by Forssén et al. [11], this modelling approach is hard to implement, computationally expensive, and thus inaccessible to practitioners.

This work presents a new versatile modelling approach for predicting the injection profile into a liquid chromatography column that requires minimal experimental effort and is easy to implement. The approach is based on the convolution of the residence time distribution (RTD) functions of the sample through the elementary units of a chromatographic system that are located between the sample injection point and the column inlet, such as connecting tubes and heat exchangers (see Fig. 1), together with the concentration profile of the sample at the injection point. The model, which is based on fundamental understanding of the material transport phenomena in a continuous flow system, accounts for the system geometry and enables predicting the injection profiles over a wide range of operating conditions (i.e., sample loop volumes and loop-filling levels). This is essential for proper simulation accuracy as required by in-silico optimization of analytical applications in order to reduce the time and cost of method development.

This paper is organized as follows. Section 2 introduces the RTD theory and the convolution integral. Section 3 describes the proposed modelling approach in detail. In Section 4, the only unknown parameter of the model is estimated, and the model is validated against experimental findings from the literature. Further discussions are provided in Section 5, with conclusions drawn in Section 6.

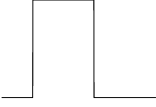
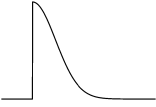


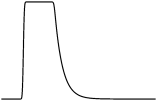

2. Residence time distribution (RTD) theory

Since the early work of Peter Danckwerts in 1953, the RTD theory has been a standard tool for understanding and analysing flow systems in a variety of physical sciences [22]. In this work, the RTD theory is employed, for the first time, to describe the injection profile in liquid chromatography.

By definition, the RTD function is the probability density function (PDF) of the time that different fluid elements, atoms or molecules, are

Table 1

Summary of empirical injection profiles reported in the literature. In these functions, C_0 is the analyte concentration in the sample, $t_{dwell,inj}$ is the dwell time from the injection point to the column inlet, t_p is the feeding time of the (equivalent) rectangular pulse, m and σ are the mean and the standard deviation of a Gaussian function, τ is the time constant of an exponential decay function, A and a are fitting parameters, while a_1 and a_2 are empirical time parameters.

Most used name	Shape	Equation	References
Rectangular pulse		$C(0,t) = \begin{cases} C_0 & \text{if } t_{dwell,inj} \leq t \leq t_{dwell,inj} + t_p \\ 0 & \text{if } t < t_{dwell,inj} \text{ or } t > t_{dwell,inj} + t_p \end{cases}$	Felinger et al. [7], Guiochon et al. [16], Gritti et al. [14]
Half Gaussian		$C(0,t) = \frac{2A}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right) \text{ for } t \geq m$	James et al. [17]
Exponentially modified Gaussian (EMG)		$C(0,t) = \frac{A}{2\tau} \exp\left(\frac{\sigma^2}{2\tau^2} + \frac{m-t}{\tau}\right) \left[1 - \operatorname{erf}\left(\frac{\sigma}{\sqrt{2\tau}} - \frac{t-m}{\sqrt{2\sigma}}\right)\right]$	Felinger [6]
EMG convoluted with a rectangular pulse		$C(0,t) = \frac{1}{2a} \left\{ \operatorname{erfc}\left(\frac{m-t}{\sqrt{2\sigma}}\right) - \operatorname{erfc}\left(\frac{m-(t-t_p)}{\sqrt{2\sigma}}\right) + \exp\left(\frac{\sigma^2}{2\tau^2} + \frac{m-t}{\tau}\right) \times \left[\exp\left(\frac{t_p}{\tau}\right) \operatorname{erfc}\left(\frac{\sigma}{\sqrt{2\tau}} + \frac{m-(t-t_p)}{\sqrt{2\sigma}}\right) - \operatorname{erfc}\left(\frac{\sigma}{\sqrt{2\tau}} + \frac{m-t}{\sqrt{2\sigma}}\right) \right] \right\}$	Felinger et al. [7]
Gaussian function convoluted with an exponentially decaying pulse characterized by an initial plateau of width θ		$C(0,t) = \frac{A}{2} \left[\operatorname{erf}\left(\frac{2m-2t+\theta}{\sqrt{2\sigma}}\right) + \operatorname{erf}\left(\frac{2t-2m+\theta}{\sqrt{2\sigma}}\right) + \exp\left(\frac{2m-2t+\theta}{\tau} + \frac{\sigma^2}{2\tau^2}\right) \operatorname{erfc}\left(\frac{2m-2t+\theta}{\sqrt{2\sigma}} + \frac{\sigma}{\sqrt{2\tau}}\right) \right]$	Forssén et al. [11], Weatherbee et al. [30]
Asymmetric double sigmoidal function		$C(0,t) = C_0 \left[1 + \exp\left(-\frac{t-t_{dwell,inj}+0.5t_p}{a_1}\right) \right]^{-1} \times \left\{ 1 - \left[1 + \exp\left(-\frac{t-t_{dwell,inj}+0.5t_p}{a_2}\right) \right]^{-1} \right\}$	Gritti et al. [14], Pepermans et al. [23]

likely to spend in one or more elements or units; and the function is often denoted as $E(t)$ [4,20]. Thus, the quantity $E(t)dt$ yields the fraction of material whose residence time in the system lies in the differential range dt around the time t . The RTD function of a system can be determined experimentally by performing either pulse injection or step change experiments of an inert tracer at the inlet of the system and recording the tracer concentration in the outlet stream as a function of time [10]. In the case of pulse injection, and for constant flow rates, $E(t)$ is directly determined by normalizing the concentration profile of the tracer in the outlet stream by its underlying area:

$$E(t) = \frac{C_{out}(t)}{\int_0^\infty C_{out}(t)dt} \quad (1)$$

In the case of a step change of the type $C_{in}(t) = C_0 H(t)$, with $H(t)$ being the Heaviside step function, one can directly determine the cumulative RTD function $F(t)$ from the concentration profile of the tracer in the outlet stream:

$$F(t) = \frac{C_{out}(t)}{C_0} \quad (2)$$

whose derivative coincides with $E(t)$:

$$E(t) = \frac{dF(t)}{dt} \quad (3)$$

For the case where the inlet concentration of the tracer is a generic function, $E(t)$ can be obtained from the outlet concentration only via deconvolution.

2.1. RTDs of common process units

The RTDs of common process units have been derived and extensively used in the literature. Below, we review the RTDs of the ideal

Plug Flow Reactor (PFR), Laminar Flow Reactor (LFR) and ideal Continuous Stirred Tank Reactor (CSTR), as we will use these expressions to model the different units of a chromatography system (e.g., connection tubes and heat exchangers).

The ideal PFR is characterized by a piston flow. No mixing takes place in the axial direction and the residence time is the same for all elements of fluid [20]. This means that a PFR delays, but does not reshape, the input signal. The RTD function of an ideal PFR is a spike of infinite height, zero width and unit area, mathematically represented by [4,10]:

$$E(t) = \delta(t - \tau) \quad (4)$$

where δ denotes the Dirac delta function and τ represents the residence time of all the fluid elements. Perfect plug flow is often assumed for describing the flow through catalytic reactors, heat exchangers and packed towers [4].

Unlike the ideal PFR, the ideal CSTR is characterized by perfect mixing. As a consequence, the concentration within the vessel volume is uniform and identical to the concentration in the outlet stream [4,10,28]. This leads to an exponentially decaying RTD function:

$$E(t) = \frac{1}{\tau} \exp\left[-\frac{t}{\tau}\right] \quad (5)$$

where τ represents the mean residence time of the fluid elements.

Experimentally measured RTDs deviate from ideality and they are usually in between the RTDs of the ideal CSTR and of the ideal PFR. This happens because, in turbulent flow, velocity fluctuations lead to axial dispersion, whereas in laminar flow, the velocity profile is parabolic according to the Hagen–Poiseuille law, with the fraction of fluid closer to the wall spending more time inside the pipe [10]. The RTD for a purely

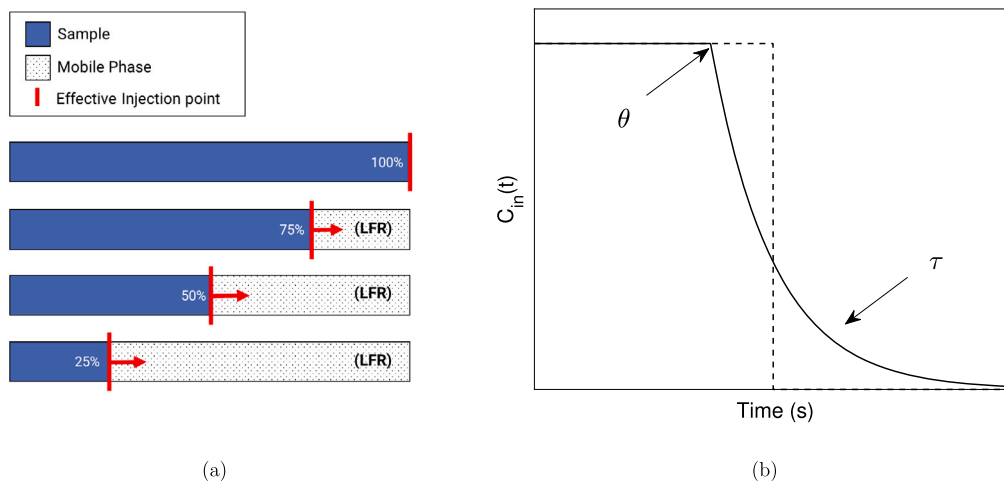


Fig. 2. (a) Schematic representation of sample loops when considering 100, 75, 50 and 25% of loop filling. When the sample loop is operated in the First-In/First-out (FIFO) mode (i.e., the sample loop is loaded and unloaded in the same direction), the sample has to flow through the dotted volume before reaching the first capillary tube, and thus, it is convenient to define an effective injection point (i.e., the thick bar). (b) Input concentration at the effective injection point $C_{in}(t)$.

convective laminar flow reactor (LFR) was derived by Danckwerts [4] and reads:

$$E(t) = \begin{cases} 0, & t < \tau/2 \\ \frac{t^2}{2\tau^3}, & t \geq \tau/2 \end{cases} \quad (6)$$

where $\tau/2$ is the minimum time the tracer stays within the reactor.

2.2. Convolution

To model the injection profile, we consider the elements or units upstream of the chromatographic column as an assembly of unit operations in series. The RTD of the overall process $E(t)$ can thus be obtained by combining the RTDs characterizing all the individual unit operations:

$$E(t) = E_1(t) * E_2(t) * \dots * E_n(t) \quad (7)$$

where $*$ denotes the convolution operator in the time domain. By definition, the convolution of two functions $E_1(t)$ and $E_2(t)$ defined in $[0, +\infty) \rightarrow \mathbb{R}$ is:

$$E(t) = (E_1 * E_2)(t) = \int_0^t E_1(s)E_2(t-s)ds \quad (8)$$

The convolution integration can also be computed as a multiplication in the frequency domain using the Fast Fourier transform (FFT) and its inverse [12]:

$$E(t) = FFT^{-1} [FFT[E_1(t)] \times FFT[E_2(t)]] \quad (9)$$

Since the convolution operation is commutative, the order of the units does not matter. Additionally, convolution can also be used to relate the output concentration of a given system of interest to the input concentration [10]:

$$C_{out}(t) = C_{in}(t) * E(t) = \int_0^t C_{in}(s)E(t-s)ds \quad (10)$$

3. Modelling approach

In this work, we propose a new modelling approach for predicting the injection profile in liquid chromatography under various operating conditions based on the RTD theory described above. If we consider $C_{in}(t)$ to be the input concentration at the injection point, the corresponding output concentration $C_{out}(t)$ can be computed by introducing Eq. (7) into Eq. (10):

$$C_{out}(t) = C_{in}(t) * E_1(t) * E_2(t) * \dots * E_n(t) \quad (11)$$

where $E_1(t), E_2(t) \dots E_n(t)$ are the RTDs of all the units composing the system chain. As we need the sample injection profile as an inlet boundary condition to solve the mass balance equation of the liquid chromatographic column, we have to consider the RTDs of all the units placed upstream of the column, from the injection point to the column inlet. We will refer to this as *injection profile at the column inlet*. If, on the other hand, we want to validate predicted injection profiles against experimental injection profiles as part of an investigation, we have to consider the RTDs of all the units placed upstream *and downstream* of the chromatographic column, all the way to the detector. This is because the injection profiles are experimentally measured at the detector having the column replaced by a zero-dead volume connector. We will refer to this as *injection profile at the detector* (see Fig. 1).

Another important consideration relates to the sample loading [3]. In the case that the sample loop is operated in the First-In/Last-out (FILO) mode (i.e., the sample loop is loaded with the sample, and then unloaded into the chromatographic system in the direction opposite from which it was loaded), the first unit to consider in the system chain is the first capillary tube. However, in the case that the sample loop is operated in the First-In/First-Out (FIFO) mode (i.e., the sample loop is loaded and unloaded in the same flow direction), and for partial loop filling, the RTD function of a further unit has to be considered. This is because when injection starts, the sample has to flow through the remaining portion of the loop before reaching the first capillary tube. To account for this, the *effective injection point* is defined at the interface between sample and mobile phase (see Fig. 2-a). Both loading modes are used without one being more commonly employed than the other [21].

Eq. (11) can be solved as long as $E_i(t)$ for $i = 1, \dots, n$ and $C_{in}(t)$ are all known (n is the number of units). With this aim, we imposed two simplifying assumptions. As a first approximation, each unit is assumed to behave either like an ideal PFR, or like an ideal CSTR, or like a LFR, and thus its characteristic RTD is given by either Eq. (4), or Eq. (5) or Eq. (6), respectively. This approximation is enough to successfully model the main physical features of the hardware system in analytical chromatography. Furthermore, it is convenient since the equations are already available and easy to solve. However, in the case of LFR, it is worth verifying that 1) the flow regime is indeed laminar (by evaluating the Reynolds number of the flow), 2) mass transfer is mainly driven by convection (so that diffusion is negligible) and 3) the ratio L/d , where L and d are the unit length and diameter, respectively, is sufficiently small (for more information, refer to Chapter 15 of Levenspiel [20]). Note that the RTD function for LFRs cannot be used to model laminar flow

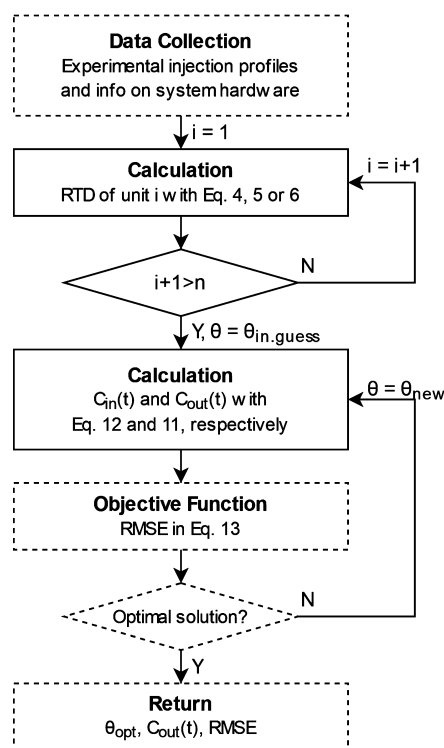


Fig. 3. Block diagram of the modelling approach developed for predicting the injection profile in liquid chromatography. The main steps are: data collection, recovery of the residence time contributions (RTDs) for each unit or element in the experimental setup, and parameter estimation. Note that, if the parameter θ is known, the modelling approach reduces to the blocks with solid lines.

systems in the Taylor-Aris regime, because radial molecular diffusion significantly influences mass transfer and in turn the RTD function. For more information, refer to the Supplementary Information section. The second assumption concerns the input concentration at the effective injection point. To allow for a certain unknown amount of backward flux to be considered when the sample is pumped through the loop, $C_{in}(t)$ is assumed to be a sharp front followed by an exponential decay:

$$C_{in}(t) = \begin{cases} 0 & \text{if } t < 0 \\ C_0 & \text{if } 0 \leq t < \theta \\ C_0 \exp[-(t-\theta)/\tau] & \text{if } t \geq \theta \end{cases} \quad (12)$$

where θ is the width of the plateau and τ is the length (in time-space) of the exponential decay (see Fig. 2-b). Because the area beneath $C_{in}(t)$ is known and reflects the injected amount of substance, only one parameter, either θ or τ , has to be estimated.

Let us establish θ as the only unknown parameter of our model, which consists of Eqs. (11) and (12). The optimal value for $\theta \in [0, t_p]$ can be estimated by minimizing the root-mean-square error (RMSE) between predicted and experimental injection profiles:

$$\min_{\theta \in [0, t_p]} \text{RMSE}(\theta) = \sqrt{\frac{1}{j} \sum_{i=1}^j (C_{out,i} - C_{out,i}^{exp})^2} \quad (13)$$

where j is the number of data points considered.

Fig. 3 shows our modelling approach in step-by-step instructions. It begins with the collection of experimental injection profiles, the identification of the hardware details of the system, the estimates of RTDs for each unit, and the estimate of an initial guess for θ . Then, it involves the calculation of θ , which is performed by minimizing the objective function as shown in Eq. (13). With the optimal value of θ , it is then possible to predict the injection profile at the column inlet. In the case where θ is known, the modelling approach reduces to just estimating

the RTDs for each unit operation and calculating $C_{in,i}$ and $C_{out,i}$, i.e. no experimental investigation is required.

In this work, the proposed modelling approach is implemented and solved in the Matlab workspace using the built-in *conv* function to compute convolution. As minimization method, we employed the routine *fminsearchbnd* [5], which uses the Nelder-Mead simplex algorithm [19] and supports variable bounds. To qualitatively compare experimental and predicted injection profiles, we also calculated the mean residence time (the first raw moment of the overall RTD):

$$t_m = \mu_1 \equiv \int_0^{\infty} t E(t) dt \quad (14)$$

and the variance of the curve, namely the band broadening contribution due to extra column volumes (the second central moment):

$$\sigma^2 = \int_0^{\infty} (t - t_m)^2 E(t) dt \quad (15)$$

4. Results

4.1. Case study: experimental and model setup

To validate our model, we employ the injection profiles experimentally obtained by Stoll et al. [27] in a customized Agilent HPLC system (1290 Infinity I line). (Their study is the only one in the literature considering injection profiles that report system geometry and dimensions.) Their injection profiles refer to the sample injected from the interface valve of a two-dimensional liquid chromatography (2D-LC) into the second dimension column. Note that, being able to accurately predict the injection profile of the ¹D effluent into the ²D column is particularly important when the ¹D effluent has a higher solvent strength than the mobile phase in the second dimension (i.e., in the case of solvent mismatch).

To experimentally measure the injection profile, Stoll et al. [27] replaced the chromatographic column with a zero-dead volume connector. As Fig. 4 shows, their system was equipped with an 8-port/2-position interface valve with two injection loops operated in the FIFO configuration. The valve was connected to a pre-column filter, removing any particulate matter, via capillary tube *a* characterized by 0.12 mm i.d. and 50 mm length. The thermostated column compartment was equipped with a 1.0 μ L heat exchanger and a zero-dead volume connector. The injected concentration profiles were measured in a UV detector equipped with a 0.6 μ L low dispersion flow cell. The connecting tubes *b*, *c* and *d* were, in order, 220 mm x 0.075 mm i.d., 100 mm x 0.075 mm i.d. and 220 mm x 0.075 mm i.d. stainless steel capillaries. The total overall upstream and downstream volumes were estimated to be about 3.0 mm³ and 1.6 mm³, respectively. The experimental injection profiles were detected by UV absorbance at 254 nm with five sample loops differing in geometry and volume, and considering four different sample loop filling levels (25, 50, 75 and 100%), for a total of 20 operating conditions tested. A 10 μ g/mL solution of uracil in water was used as tracer and the flow rate was set at 2.5 mL/min. More information about the experimental setup and the experimental procedure can be found in Stoll et al. [27].

In deriving the predicted injection profiles at the detector, we implemented the procedure described in Fig. 3. We modelled each capillary tube as a LFR because, in each tube, laminar flow occurs (i.e. the Reynolds number $Re \equiv \rho u d / \mu$ is much lower than 2100) and convection transport dominates over diffusive transport (i.e. the Péclet number $Pe \equiv u d / D$ is much greater than L/d). For more information, refer to the Supplementary Information section. In terms of the heat exchanger, we assumed this to behave like an ideal PFR as suggested by Danckwerts [4]. For the case of 100% loop filling, the effective injection point coincides with the end of the sample loop, and thus capillary tube *a* was the first unit considered in determining the RTD of the overall system.

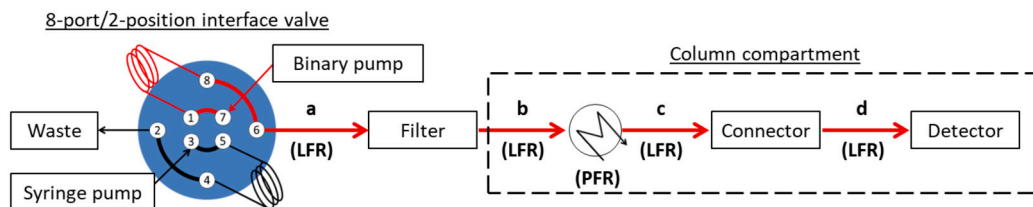


Fig. 4. Schematic representation of the customised Agilent HPLC system used in Stoll et al. [27]. The figure provides details about the flow path upstream and downstream of the chromatographic column (the column is replaced with a zero-dead volume connector for the investigation). The elements *a*, *b*, *c* and *d* are capillary tubes.

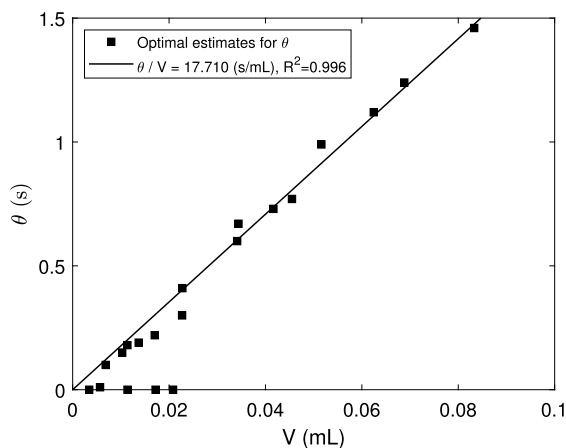


Fig. 5. The parameter θ as a function of the sample volume V .

However, since the sample loop is operated in FIFO mode, in the case of partial loop filling, the sample is first pushed through the remaining portion of the loop before entering tube *a*. To account for this, we located the effective injection point within the sample loop, at the interface between the sample and the mobile phase, and we considered the rest of the sample loop as a further element in the system chain. As a first approximation, the latter was assumed to behave like a LFR. Note that, in this case study, we considered neither the pre-column filter, nor the connecting valves, nor the low-dispersion flow cell. However, their volumes are small compared to the volume of the other units and, as a consequence, their contribution in shaping the overall RTD is negligible. All simulations were performed considering the loop volumes measured by Stoll et al. [27] (13.7, 22.7, 45.5, 68.8, 83.3 μL) rather than the nominal loop volumes (13.5, 20, 40, 60 and 80 μL).

4.2. The optimal estimate for parameter θ

Our modelling approach relies on an unknown parameter, θ , that is obtained by minimizing the RMSE between experimental and predicted injection profiles (see Eq. (13)). Fig. 5 shows the optimal estimate for θ as a function of the sample volume for all the 20 conditions tested. Interestingly, we found that there exists a linear relationship between θ and the sample volume V , and that the five data that do not follow the trend, i.e., the five markers whose value is close to zero, correspond to cases where only 25% of the loop is filled with sample. If we disregard these outliers (there must be some unaccounted effects that are significant when it comes to 25% of loop filling), the data are well fitted by $\theta/V = 17.71$ (s/mL), with an R-squared value equal to 0.996. As a consequence, if the loop filling ranges between 50% and 100% and the chromatographic system is similar to that considered in this work, then the value of θ can be predicted *a priori* and our modelling approach becomes parameter-free, i.e. not requiring any experimental investigation. The model can then be used to predict injection profiles for a wide range of operative conditions and hardware settings.

For loop filling below 50%, on the other hand, the trend is unclear. We believe that the difference between simulation and experiment is

due to unaccounted effects that become significant when the sample volume is small compared to the injection loop volume. Examples are fluxes due to valve movements when injecting the sample into the mobile phase, different viscosities of the sample and the mobile phase, or longitudinal diffusion over time.

The relationship between θ and V must be re-calibrated for chromatographic systems and operating conditions that deviate from those considered here. In particular, lower values of θ/V (i.e., C_{in} characterized by a narrower plateau and a longer tail) might result from solutes with larger molecular weights and solvents with higher viscosity. Theoretically, θ/V should be re-calibrated also when the flow rate changes. However, since the influence of the flow rate in shaping the injection profile is minor [26,11,30], almost negligible at high sample volumes, predicting θ/V only once (e.g., in the middle of the range considered) is enough in most practical situations.

In cases where re-calibration is really needed, only one experiment is required (preferably performed with a loop filling greater than 50%). The fine-tuned θ/V can then be used to model a range of injection loop volumes and loop-filling levels.

4.3. Predicted and experimental injection profiles: a comparison

Fig. 6 compares the experimental injection profiles measured at the detector (markers) with the injection profiles predicted considering $\theta/V = 17.71$ (s/mL) (solid lines) for all the 20 conditions tested. All profiles are normalized to a maximum peak height of 1. To get a quantitative measure of the goodness of the comparison, the RMSE is provided above each picture. The agreement between experimental and predicted injection profiles is good for most of the injection volumes and filling levels tested. The largest deviations are observed when the sample is injected through a sample loop filled for 25% (first column of Fig. 6) as previously discussed, and when there are larger volumes of the injection loop not filled with the sample (bottom right corner of Fig. 6). This indicates that assuming laminar flow through the remaining portion of the sample loop might be inaccurate (Stoll et al. [27] do not provide any information about the length and diameter of the injection loops; therefore, we could not verify the applicability of the pure convective model, i.e., whether the LFR assumption was suitable in all cases). The model accounts for the time delay the sample requires to reach the injection port and the parabolic velocity profile, but it does not account for phenomena such as mixing and Taylor-Aris dispersion which may be present. Even if our model is inaccurate at low loop-filling levels, it works well for loop-filling levels equal to or greater than 50%. Furthermore, the model should always work when the sample loop is operated in the FILO mode (in this case, there is no need to model a portion of the sample loop and thus, we expect the same accuracy obtained in the FIFO mode reported here for 100% loop filling).

Table 2 reports the mean residence time and the time variance of the experimental and predicted injection profiles, together with their relative errors. In both cases, the median relative error of the predictions is $\sim 7.5\%$. Note that, the time variance of the injection profile at the detector represents the band broadening associated with the migration of the solute in the extra-column volumes, and thus it is useful to determine the influence of extra-column contribution to peak width under both isocratic and gradient conditions [15,29].

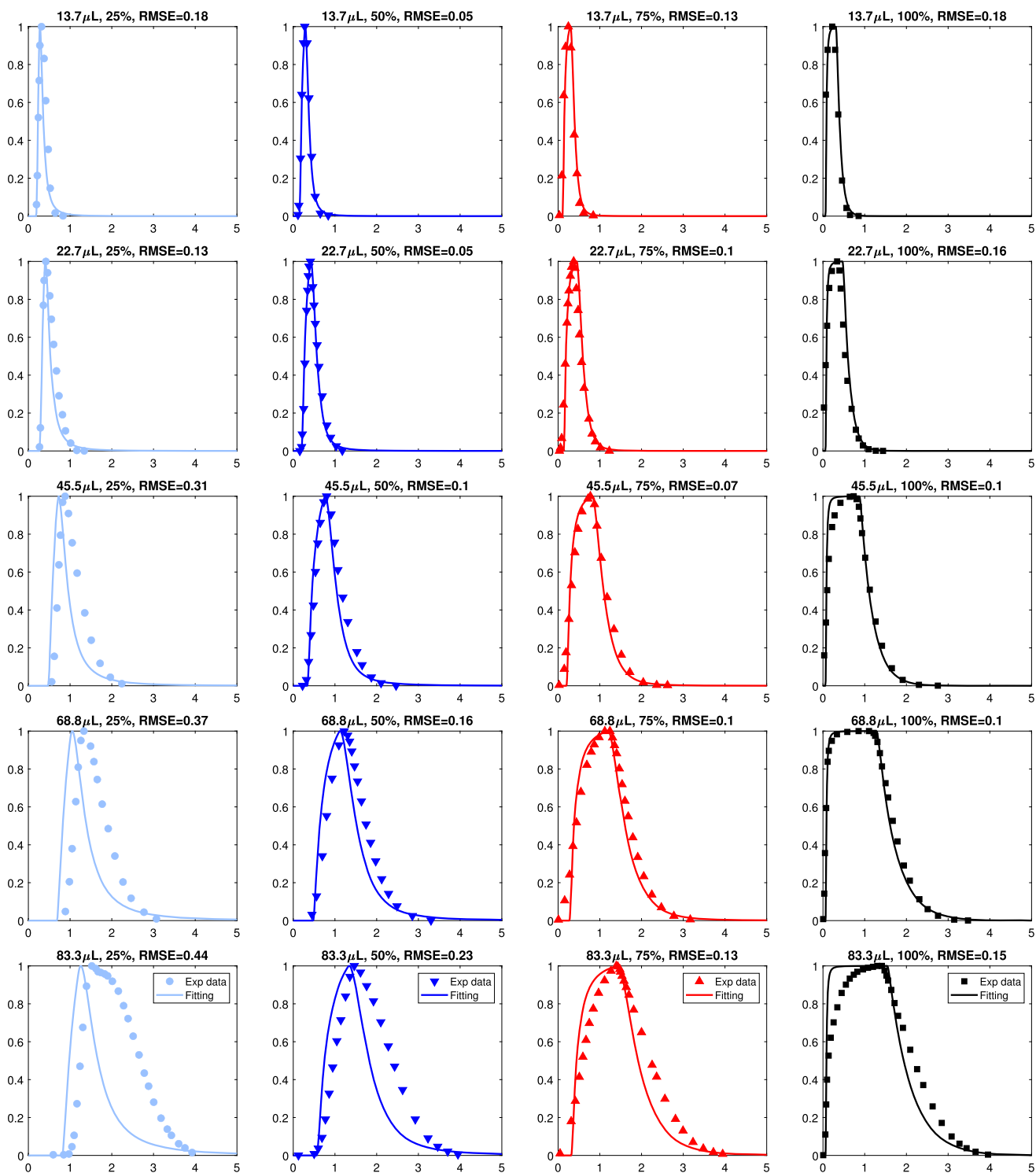


Fig. 6. Comparison between the experimental injection profiles provided by Stoll et al. [27] (markers) and the injection profiles predicted considering $\theta = 17.71$ (s/mL) V (solid lines). The effect of varying the injection loop volumes (i.e. 13.7, 22.7, 45.5, 68.8 and 83.3 μL) and the loop-filling levels (i.e. 25, 50, 75 and 100%) are reported by rows and by column, respectively. All the curves are normalized, and the flow rate is 2.5 mL/min.

Table 2

Mean residence time and variance of experimental and predicted injection profiles with relative errors for the 20 conditions tested. The experimental injection profiles are the one provided by Stoll et al. [27].

Vol (μL)	Filling (%)	$t_{m,exp}$	$t_{m,pre}$	ER_{t_m}	σ_{exp}^2	σ_{pre}^2	ER_{σ^2}
13.7	25	0.33	0.30	6.82	0.01	0.01	3.68
13.7	50	0.29	0.30	1.21	0.01	0.01	10.36
13.7	75	0.25	0.29	13.68	0.01	0.01	2.61
13.7	100	0.22	0.26	17.53	0.01	0.01	3.65
22.7	25	0.50	0.48	5.18	0.02	0.02	3.46
22.7	50	0.44	0.43	1.75	0.02	0.02	5.19
22.7	75	0.37	0.39	6.45	0.02	0.02	0.67
22.7	100	0.33	0.38	12.67	0.04	0.03	6.84
45.5	25	0.95	0.82	13.45	0.06	0.06	7.62
45.5	50	0.83	0.77	6.70	0.08	0.07	9.36
45.5	75	0.73	0.72	2.00	0.10	0.09	12.47
45.5	100	0.65	0.63	2.85	0.14	0.13	3.04
68.8	25	1.49	1.25	15.72	0.11	0.13	16.11
68.8	50	1.36	1.25	8.04	0.14	0.15	12.34
68.8	75	1.17	1.13	2.71	0.21	0.19	7.73
68.8	100	0.95	0.97	2.60	0.37	0.35	6.39
83.3	25	1.97	1.51	23.42	0.30	0.25	15.36
83.3	50	1.63	1.35	17.07	0.31	0.25	19.53
83.3	75	1.41	1.30	8.08	0.31	0.26	16.36
83.3	100	1.16	1.03	11.52	0.46	0.43	7.69

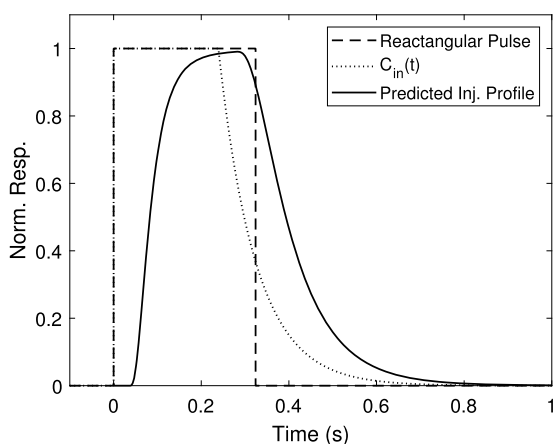


Fig. 7. Rectangular pulse (dashed line), $C_{in}(t)$ (dotted line) and predicted injection profile (solid line) for 13.5 μL sample volume and 100% loop filling. The total area overlap between the ideal rectangular pulse and the predicted injection profile is 70.78%.

5. Discussion

In the previous sections, we have shown that the RTD theory can be used to accurately predict the injection profiles and to estimate the extra-column band broadening for sample loops of different sizes and different loop-filling levels. Furthermore, for loops filled between 50% and 100%, the parameter θ scales with the sample volume, and the model becomes parameter-free and is therefore valid for many chromatographic settings and combinations.

5.1. Deviation from the ideal rectangular pulse

Fig. 7 compares the predicted injection profile (solid line) with the ideal rectangular pulse (dashed line) for the case of 13.5 μL sample volume and 100% loop filling. The total area overlap between the two curves is only 70.78%. In other cases considered, the overlap is even lower (not shown). This demonstrates that true injection profiles might significantly deviate from ideal rectangular pulses, and therefore one

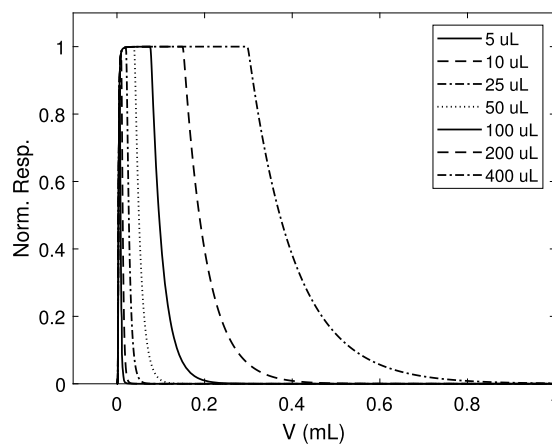


Fig. 8. Normalized injection profile for 5, 10, 25, 50, 100, 200 and 400 μL at a flow rate of 2.5 mL/min, considering $\theta/V = 17.71$ (s/mL), 100% loop filling, and the same system represented in Fig. 4.

must impose as inlet boundary condition a profile that is as close to the real one as possible. In doing so, however, one must accept a trade-off between accuracy and experimental effort. Our model leads in most cases to reasonably accurate inlet profiles, while significantly reducing the experimental effort required to obtain the values of the model parameters. Fig. 7 also shows how the input concentration $C_{in}(t)$ (dotted line) looks like for this particular case study.

5.2. Dependence on injection volume

In their work, Forssén et al. [11] demonstrated that the injection volume is the main parameter that needs to be accounted for when modelling the injection profile. Thus, we report in Fig. 8 the injection profiles predicted for varying injection volumes, at 100% loop filling and with $\theta/V = 17.71$ (s/mL). The profiles are characterized by a steep front and a significant tail that arises because of sample dispersion through the different elements or units. Furthermore, the dispersion of the injection profiles increases with the injection volume and, for the higher injection volumes, the concentration plateau emerges. The shapes of these profiles are qualitatively similar to those reported by Samuelsson et al. [26] and Forssén et al. [11], respectively, in their Figs. 1 and 2, however, their models require either more parameters or more computational effort than our model. We can conclude that our modelling approach can accurately and easily predict the shape of the injection profile across a set of injection volumes and with only one parameter (θ) that is easy to estimate.

6. Conclusions

Actual injection profiles in liquid chromatographic systems generally deviate from the ideal rectangular shape often assumed, and one of the main challenges that researchers face when modelling liquid chromatography is implementing an injection profile that mimics the actual one. Previous literature addressed this problem by using either experimentally determined injection profiles or surrogate models derived from the combination of Gaussian, square, and exponential residence time profiles [2]. However, these models require a relatively large amount of experimental data, and performing experiments is both expensive and time-consuming.

In this work, we have proposed a new modelling approach for in-silico estimation of injection profiles for liquid chromatography that is based on the residence time distribution (RTD) theory. The approach consists in combining, or convoluting, the RTD functions of all the different units or elements composing the chromatographic system that are placed between the injection point and the column inlet (e.g., tubing, heat exchanger etc). To obtain the injection profile, the overall RTD is

then convolved with the inlet concentration of the sample. The model, which has been validated against independent experimental data, relies on a single easily fitting parameter whose value depends on the injection volume.

The modelling approach requires either one experimental injection profile, or no experiments at all. The main advantages of the model are that it is applicable over a range of operating conditions, there is no need to solve mass balance equations, and the experimental work is either minimized or not required. Furthermore, the approach is cheap, fast and easy to use, and is therefore accessible to most practitioners, with the only requirement being the knowledge of the size and geometry of the different chromatographic system elements. Future work includes testing and adapting the proposed approach to preparative scale chromatography where different assumptions may be required.

Funding

This work was funded by the Engineering and Physical Sciences Research Council (EPSRC), grant code EP/T005556/1.

CRediT authorship contribution statement

Monica Tirapelle: Conceptualization, Investigation, Methodology, Software, Writing – review & editing. **Maximilian O. Besenhard:** Conceptualization, Writing – review & editing. **Luca Mazzei:** Conceptualization, Supervision, Writing – review & editing. **Jinsheng Zhou:** Supervision. **Scott A. Hartzell:** Supervision. **Eva Sorensen:** Project administration, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Funding from Eli Lilly and Company and the UK EPSRC, through the PharmaSEL-Prosperty Programme (EP/T005556/1), is gratefully acknowledged.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.chroma.2023.464363>.

References

- [1] K. Baran, W.K. Marek, W. Piatkowski, D. Antos, Effect of flow behavior in extra-column volumes on the retention pattern of proteins in a small column, *J. Chromatogr. A* 1598 (2019) 154–162.
- [2] M.O. Besenhard, A. Tsatse, L. Mazzei, E. Sorensen, Recent advances in modelling and control of liquid chromatography, *Curr. Opin. Chem. Eng.* 32 (2021) 100685.
- [3] P. Carr, D. Stoll, *Two-Dimensional Liquid Chromatography—Principles, Practical Implementation and Applications*, Agilent Technologies Inc., Germany, 2015, p. 182.
- [4] P.V. Danckwerts, Continuous flow systems: distribution of residence times, *Chem. Eng. Sci.* 2 (1953) 1–13.
- [5] J. D'Errico, `fminsearchbnd`, `fminsearchcon`, MATLAB Central File Exchange, <https://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd-fminsearchcon>, 2022.
- [6] A. Felinger, *Data Analysis and Signal Processing in Chromatography*, Elsevier, 1998.
- [7] A. Felinger, D. Zhou, G. Guiochon, Determination of the single component and competitive adsorption isotherms of the 1-indanol enantiomers by the inverse method, *J. Chromatogr. A* 1005 (2003) 35–49.
- [8] B. Filip, R. Bochenek, K. Baran, D. Strzałka, D. Antos, Influence of the geometry of extra column volumes on band broadening in a chromatographic system. Predictions by computational fluid dynamics, *J. Chromatogr. A* 1653 (2021) 462410.
- [9] B. Filip, R. Bochenek, W.K. Marek, D. Antos, Flow behavior of protein solutions in a lab-scale chromatographic system, *J. Chromatogr. A* 464178 (2023).
- [10] H.S. Fogler, *Elements of Chemical Reaction Engineering*, 5th edition, Pearson Education, 2016.
- [11] P. Forssén, L. Edström, J. Samuelsson, T. Fornstedt, Injection profiles in liquid chromatography II: predicting accurate injection-profiles for computer-assisted preparative optimizations, *J. Chromatogr. A* 1218 (2011) 5794–5800.
- [12] Y. Gao, F.J. Muzzio, M.G. Ierapetritou, A review of the residence time distribution (RTD) applications in solid unit operations, *Powder Technol.* 228 (2012) 416–423.
- [13] S. Golshan-Shirazi, G. Guiochon, The equilibrium-dispersive model of chromatography, in: *Theoretical Advancement in Chromatography and Related Separation Techniques*, vol. 35, 1992.
- [14] F. Gritti, M. Gilar, J. Hill, Mismatch between sample diluent and eluent: maintaining integrity of gradient peaks using in silico approaches, *J. Chromatogr. A* 1608 (2019) 460414.
- [15] F. Gritti, G. Guiochon, Band broadening in fast gradient high-performance liquid chromatography: application to the second generation of 4.6 mm id silica monolithic columns, *J. Chromatogr. A* 1238 (2012) 77–90.
- [16] G. Guiochon, A. Felinger, D.G. Shirazi, A.M. Katti, *Fundamentals of Preparative Chromatography*, 2006.
- [17] F. James, M. Sepúlveda, F. Charton, I. Quinones, G. Guiochon, Determination of binary competitive equilibrium isotherms from the individual chromatographic band profiles, *Chem. Eng. Sci.* 54 (1999) 1677–1696.
- [18] L.N. Jeong, R. Sajulga, S.G. Forte, D.R. Stoll, S.C. Rutan, Simulation of elution profiles in liquid chromatography I: gradient elution conditions, and with mismatched injection and mobile phase solvents, *J. Chromatogr. A* 1457 (2016) 41–49.
- [19] J.C. Lagarias, J.A. Reeds, M.H. Wright, P.E. Wright, Convergence properties of the Nelder–Mead simplex method in low dimensions, *SIAM J. Optim.* 9 (1998) 112–147.
- [20] O. Levenspiel, *Chemical Reaction Engineering*, John Wiley & Sons, 1998.
- [21] A. Moussa, T. Lauer, D. Stoll, G. Desmet, K. Broeckhoven, Modelling of analyte profiles and band broadening generated by interface loops used in multi-dimensional liquid chromatography, *J. Chromatogr. A* 1659 (2021) 462578.
- [22] E.B. Nauman, Residence time theory, *Ind. Eng. Chem. Res.* 47 (2008) 3752–3766.
- [23] V. Pepermans, S. Chapel, S. Heinisch, G. Desmet, Detailed numerical study of the peak shapes of neutral analytes injected at high solvent strength in short reversed-phase liquid chromatography columns and comparison with experimental observations, *J. Chromatogr. A* 1643 (2021) 462078.
- [24] I. Quinones, J.C. Ford, G. Guiochon, High-concentration band profiles and system peaks for a ternary solute system, *Anal. Chem.* 72 (2000) 1495–1502.
- [25] A.C.A. Roque, C.R. Lowe, M.Á. Taipa, Antibodies and genetically engineered related molecules: production and purification, *Biotechnol. Prog.* 20 (2004) 639–654.
- [26] J. Samuelsson, L. Edström, P. Forssén, T. Fornstedt, Injection profiles in liquid chromatography. I. A fundamental investigation, *J. Chromatogr. A* 1217 (2010) 4306–4312.
- [27] D.R. Stoll, R.W. Sajulga, B.N. Voigt, E.J. Larson, L.N. Jeong, S.C. Rutan, Simulation of elution profiles in liquid chromatography-II: investigation of injection volume overload under gradient elution conditions applied to second dimension separations in two-dimensional liquid chromatography, *J. Chromatogr. A* 1523 (2017) 162–172.
- [28] P. Toson, P. Doshi, D. Jajcevic, Explicit residence time distribution of a generalised cascade of continuous stirred tank reactors for a description of short recirculation time (bypassing), *Processes* 7 (2019) 615.
- [29] K. Vanderlinden, K. Broeckhoven, Y. Vanderheyden, G. Desmet, Effect of pre- and post-column band broadening on the performance of high-speed chromatography columns under isocratic and gradient conditions, *J. Chromatogr. A* 1442 (2016) 73–82.
- [30] S.L. Weatherbee, T. Brau, D.R. Stoll, S.C. Rutan, M.M. Collinson, Simulation of elution profiles in liquid chromatography – IV: experimental characterization and modeling of solute injection profiles from a modulation valve used in two-dimensional liquid chromatography, *J. Chromatogr. A* 1626 (2020) 1–10.