Under the influence of light
New chromatographic tools for elucidating photodegradation mechanisms
den Uijl, M.J.

Publication date
2022

Citation for published version (APA):
CHAPTER 4

Modelling SPAM

Abstract

Two-dimensional liquid chromatography (2DLC) offers great separation power for complex mixtures. The frequently encountered incompatibility of two orthogonal separation systems, however, makes its application complicated. Active-modulation strategies can reduce such incompatibility issues considerably. Stationary-phase-assisted modulation (SPAM) is the most-common of these techniques, but also the least robust due to the major disadvantage that analytes may elute prematurely. The range of liquid chromatography (LC) applications continues to expand towards ever more complex mixtures. Retention modelling is increasingly indispensable to comprehend and develop LC separations. In this research, a tool was designed to assess the feasibility of applying SPAM in 2DLC. Several parameters were investigated to accurately predict isocratic retention of analytes on trap columns under dilution-flow conditions. Model parameters were derived from scanning-gradient experiments performed on analytical columns. The trap-to-trap repeatability was found to be similar to the prediction error. Dead volumes for the trap columns could not be accurately determined through direct experimentation. Instead, they were extrapolated from dead-volume measurements on analytical columns. Several known retention models were evaluated. Better predictions were found using the quadratic model than with the log-linear (“linear-solvent-strength”) model. Steep scanning gradients were found to result in inaccurate predictions. The impact of the dilution flow on the retention of analytes proved less straightforward than anticipated. Under certain conditions dilution with a weaker eluent was found to be counterproductive. A tool was developed to quantify the effect of the dilution flow and to predict whether SPAM could be applied in specific situations. For nine different analytes under 36 different sets of conditions and with three different modulation times, the SPAM tool yielded a correct assessment in more than 95% of all cases (less than 5% false positives plus false negatives).

Publication

Assessing the Feasibility of Stationary-Phase-Assisted Modulation for Two-Dimensional Liquid-Chromatography Separations
Mimi J. den Uijl, Tim Roeland, Tijmen S. Bos, Peter J. Schoenmakers, Maarten R. van Bommel, and Bob W.J. Pirok

J. Chromatogr. A, 2022, 1679, 463388, DOI: 10.1016/j.chroma.2022.463388
4.1. Introduction

Liquid chromatography (LC) is one of the most important analytical techniques because of its wide range of applications, the choice of selectivity, and the possibility to couple it to many detection techniques [1]. In an LC separation, the maximum number of peaks that can be resolved is the theoretical peak capacity \( n_c \). While this number has been increasing in recent years by using long columns under high pressure [2], core–shell particles [3–5], monolithic stationary-phases [6], or elevated temperatures [7], it has to be accounted for that the number of peaks that are separated is far lower than the \( n_c \) [8]. Next to that, complex mixtures, like the samples from cultural heritage [9], blood analysis [10], food industry [11], and waste-water effluent [12] require unattainable peak capacities for one-dimensional liquid chromatography. In two-dimensional liquid chromatography (2DLC), more resolving power and higher peak capacities can be obtained [13]. In 2DLC, the effluent of the first-dimension column (1D) is transferred to a second-dimension column (2D), where the analytes undergo another separation. This transfer can be either performed offline, or in an online fashion by means of a valve. For comprehensive 2DLC (LC×LC), all 1D effluent is transferred to the 2D.

The power of two-dimensional liquid chromatography (2DLC) lies in the combination of two different retention mechanisms, which are carefully chosen by i) the sample dimensionality, ii) the compatibility with the detector, iii) their duration (i.e. one needs to be fast), and, of course, iv) the compatibility of the two dimensions with each other. Although 2DLC has been developing in the last years, with 160 applications from 2016 to 2018 [14], this latter compatibility issue significantly complicates its implementation [15].

The employed modulation strategy plays a key role to improve the compatibility of the two dimensions within a 2DLC method. Unmodified transfer of the 1D effluent to the 2D through an 8- or 10-port valve equipped with empty loops is referred to as passive modulation [14]. For this to succeed, the two separation techniques should be compatible to prevent miscibility and adsorption issues such as breakthrough, peak deformation, and peak splitting [16,17]. Another issue is the loss of sensitivity due to the additional dilution induced by a second dimension [18–20]. Driven to combine organic-based separations with aqueous based, these compatibility challenges defined the fundamental work of many groups worldwide in the past decades. These efforts culminated into the development of active modulation techniques: active-solvent modulation (ASM) [21,22], stationary-phase-assisted modulation (SPAM) [9,23–25], thermal modulation [26–29], evaporative membrane modulation [30], vacuum evaporation modulation [31,32], and cold trapping [33]. While the latter four techniques have seen relatively few applications thus far, the former two are considered established in academia [14].
In SPAM, the analytes in the 1D effluent are trapped on a small guard column, which is connected similarly like the loops in passive modulation. In this method, the total volume of the SPAM column, and thus the injection volume of the 2D, is small. Another advantage of this modulation technique is that the modulation volume is not restricted to a loop volume. However, where ASM has been demonstrated to be a robust technique, for SPAM i) guard columns do not feature the same quality and robustness standards as normal columns, rendering reproducibility of the method difficult, ii) pressure pulses on the guard column are expected to reduce its lifetime, and iii) when the analyte retention is too low in the effluent of the 1D, the compounds will elute prematurely from the guard column and be lost. While the first two challenges can be addressed technically, the latter describes a discriminatory attribute that is intrinsically fatal to a method. While premature elution can be detected by installing a 1D detector post-valve to monitor the 1D waste, this is not practical for complex samples of unknown composition.

To prevent premature elution, methods that use SPAM often feature an active dilution flow to the 1D effluent to reduce the elution strength, however these dilution flows are often chosen rather randomly [9]. Moreover, the improved retention factor by the weaker elution strength is countered by the additional elution volume induced by the additional flow. For users it is thus difficult to gauge whether SPAM could be applied for a sample of interest and, if so, what conditions should be used. Unknown loss of analytes to the lack of retention renders SPAM unreliable, yet some applications do rely on complete removal of the 1D eluent and thus this problem must be addressed.

One potential solution may be found in empirical retention modelling. Here, retention parameters are established that relate analyte retention as a function of mobile-phase composition [34]. These parameters are obtained by fitting an empirical model to retention data for each analyte. This data is obtained by measuring analyte retention at several mobile-phase compositions which is often used to develop optimal gradient conditions [35]. For SPAM, there is an opportunity to employ retention modelling since the guard columns used typically feature an identical or similar stationary phase relative to the analytical 2D column. This suggests that retention parameters obtained from analytical 2D columns can be extrapolated to the shorter guard columns to predict retention on SPAM columns, a concept that has recently also been used [36], is particularly attractive because the 2D is often reversed-phase liquid chromatography (RPLC) [37] and often a 1DLC starting point for any 2DLC method. If successful, retention modelling could be used to estimate the success rate of the implementation of SPAM for all analytes, without trial-and-error 2DLC experiments.
In this work we aim to develop a tool to quickly assess the feasibility of using SPAM in 2DLC. We evaluate retention prediction on SPAM (trap) columns with sufficient accuracy to decide on i) the feasibility of SPAM and ii) the desired dilution flow rate. We aim to use empirical retention models constructed using data obtained from scanning gradients on analytical columns. Various factors need to be controlled to achieve this goal. System parameters, such as the extra-column residence times and column dead time, and the trap-to-trap repeatability must be controlled. We aim to establish an optimal set of gradient-scanning experiments and an optimal retention model to predict isocratic retention factors on SPAM columns. The obtained models will be used to predict the minimal dilution flow needed to achieve sufficient retention on the trap columns. We finally aim to use all this information to decide whether a SPAM process can be used successfully within certain boundaries, such as (minimum and maximum) modulation time, dilution flow, 1D flow rate, and 1D modifier composition.

4.2. Experimental

4.2.1. Chemicals
Milli-Q water (R = 18.2 MΩ cm) was obtained from a purification system (Arium 611UV, Sartorius, Germany). Acetonitrile (ACN, HPLC grade), Acetone (HPLC grade) and Toluene (LC-MS grade) were purchased from Biosolve Chemie (Valkenswaard, The Netherlands). Formic acid (98%) was purchased from Fluka (Buchs, Switzerland). Riboflavin (≥99%), crystal violet (≥90%), Ammonium formate (≥99%), phenol (~99%), orange G, 4-hydroxybenzoic acid (≥99%), propranolol (≥99%), trimethoprim (≥99%), uracil (≥99%), and acetaminophen (≥99%) were purchased from Sigma Aldrich (Darmstadt, Germany).

Stock solutions were made in H₂O/ACN (50/50, [v/v]) of 500 ppm (riboflavin, toluene, phenol, uracil, acetaminophen, propranolol, trimethoprim, and 4-hydroxy benzoic acid) and 1000 ppm (crystal violet and orange G). Two solutions were made from these separate stock solutions, each containing five compounds, which were further diluted with H₂O/ACN (50:50, [v/v]). The final concentrations can be found in Appendix B-1, Table B-1.

4.2.2. Instrumental
All experiments were carried out on an Agilent 1290 series Infinity 2DLC system (Agilent, Waldbronn, Germany), which was configured for one-dimensional operation. This system was equipped with a binary pump (G7120A) equipped with a 35-μL JetWeaver mixer, an autosampler (G7129B), column oven (G7116B) and a diode-array detector (DAD, G7117B).
For all measurements on the analytical column, a Zorbax RRHD Eclipse Plus C18 column with dimensions 50 × 2.1 mm, 1.8 µm (Agilent) was used. For the measurements on the trapping column the Zorbax RRHD Eclipse plus C18 guard column with dimensions 5 × 2.1 mm, 1.8 µm UHPLC guard (Agilent) was used. The temperature was not controlled. A DAD detector was equipped at several wavelengths, depending on the absorption spectrum of the compound. The detection wavelength for phenol, trimethoprim, propranolol, and toluene was 214 nm, for uracil, 4-hydroxy benzoic acid and acetaminophen it was 254 nm, 590 nm was used for crystal violet, 492 nm for orange G, and 450 nm for riboflavin. The slit size was set to 4 nm and the sampling rate to 240 Hz.

4.2.3. Analytical methods

The following mobile phases were used for all experiments in this project unless stated otherwise. Mobile phase A consisted of a 5 mM ammonium formate buffer targeted to a pH of 3. To prepare 1 L of buffer 0.0476 grams of ammonium formate and 0.195 grams (± 160 µL) formic acid were mixed with 1 L of water. The pH of the buffer was verified using a pH-meter (Metrohm, Herisau, Switzerland) and adjusted to pH = 3.00 using additional formic acid and a glass Pasteur pipette. Mobile phase B consisted of ACN.

The dwell volume of the system was determined to be about 0.1775 mL, which was experimentally determined using acetone as modifier. Solvent A was water and solvent B was water with 0.1% [v/v] acetone. An initial time of 4 min was used followed by a gradient from 0 to 100% B in 8 min. These were performed in triplicate and 210 nm was used as detection wavelength. The gradient delay was determined at 50% of the gradient.

The dead volume ($V_\text{d}$) of the analytical column and the guard column was determined by injecting uracil ($V_{\text{inj}} = 2\ \mu$L) on the column at a 50/50 [v/v] composition of mobile-phase components A and B and at a flow rate of 0.5 mL/min. The extra-column volume ($V_{\text{ex}}$) was determined by repeating the previous experiments while replacing the column with a union.

4.2.2.1. Isocratic measurements

Isocratic measurements were performed on an analytical column and two trapping columns (trap A and B). Seven different organic-modifier fractions ($\phi$) were performed at a flow rate of 0.5 mL·min$^{-1}$ with an injection volume of 2 µL. The chosen $\phi$ levels were 0.025, 0.05, 0.075, 0.1, 0.2, 0.3, and 0.4. The analysis time of the isocratic measurements was 10 min. After this period, mobile phase B was increased to 95% in 1 min and maintained for 0.5 min to elute potentially retained compounds. Afterwards, the mobile phase was brought back to the starting conditions of the subsequent method in 0.01 min and maintained for 1.49 min for
re-equilibration. All measurements for both stock solutions with all three columns (i.e. one analytical column and two trapping columns) were repeated 5 times using a single batch of buffer.

4.2.2.2. Scanning-gradient experiments
For the scanning gradients on the analytical column, the following gradients were performed. All gradients started with an initial time of 0.25 min at 1% B, followed by a linear gradient to 95% B in either 1, 2, 3, 6, 9, 12, 15, 24, or 48 min. 95% B was maintained for 0.5 min and brought back to 1% B in 0.01 min, followed by a 1.24-min re-equilibration step at 1% B. These scanning gradients were performed at a flow rate of 0.5 mL·min⁻¹. All measurements were performed in triplicate.

4.2.2.3. Dilution-flow experiments
Four different dilution-flow series (DF 1, 2, 3, and 4) were applied on the trap columns. Here, the initial $\phi$ of the 1D was varied between the four DF series. The initial 1D $\phi$ for DF 1, 2, 3 and 4 were 0.75, 0.5, 0.25, and 0.1, respectively. The initial 1D flow rate remained the same (50 μL·min⁻¹), which was diluted 1:0, 1:1, 1:2, 1:3, 1:4, 1:6.5, 1:9, 14:1 and 19:1, corresponding to total flow rates of 50, 100, 150, 200, 250, 375, 500, 750 and 1000 L·min⁻¹, respectively. This yielded 36 different methods, with varying flow rate and $\phi$, as can be seen in Appendix B-2. Every measurement was performed in triplicate.

Each of these 36 methods is essentially a 10-min isocratic measurement. For these measurements, the first 10 min consisted of the isocratic part at the described $\phi$ level and corresponding flow rate. After this part, the flow rate and %B were adjusted to quickly elute potentially retained compounds and re-equilibrate the trapping column.

4.2.4 Data processing
The in-house built MATLAB-based user interface MOREPEAKS [38,39] was used to fit the retention models and determine the retention parameters from the experimental data. The boundaries for the LSS model were -10 to 50 for $\ln k_0$ and 0 to 100 for $S_{\text{LSS}}$. For the QUA model the boundaries were -10 to 50 for $\ln k_0$ and 0 to 100 for $S_1$, and 0 to 100 for $S_2$. Expert fitting was used and the multistart function was used at 40 for both models, unless stated otherwise. Microsoft Excel and MATLAB R2020b were used for all other data processing.
4.3. Results & discussion

The goal of this work was to develop a tool to assess the feasibility of using SPAM in 2DLC by predicting retention on SPAM (trap) columns with sufficient accuracy using scanning-gradient experiments on analytical columns. This concept essentially comprised method transfer from one column (analytical) to another (trap) and prediction of isocratic retention at variable flow rate with data from gradient-elution experiments. To achieve this, three steps were outlined, namely i) establishing correct system parameters, such as $V_0$ and $V_{\text{ex}}$, ii) selection of appropriate method parameters, such as the number of scanning gradients, the retention model, and the gradient steepness, and iii) the accurate description of the effect of a dilution flow on retention. These points were envisaged to predict SPAM retention with a new tool developed in this work. This workflow is shown in Fig. 4.1. These three steps will be covered in Sections 4.3.1, 4.3.2, and 4.3.3, respectively.

![Figure 4.1. Workflow for the development of a retention-prediction tool on trap columns with retention data from analytical columns. Every step corresponds to a result section.](image)

4.3.1. Matching system parameters between analytical columns and trap columns

In order to assess the feasibility of using retention data from analytical columns to predict those on trap columns, it is necessary to compare isocratic retention factors obtained from both columns individually. To compare retention between columns with different dimensions, the dead volume ($V_0$) of the columns should be determined. Values for $V_0$ were experimentally determined by injection of a $V_0$ marker (i.e. uracil) in a mobile-phase composition of 50/50 buffer/ACN. Values of 12 μL and 88 μL were obtained for the trap- and analytical columns, respectively. Next to that, the extra-column volume between the injector and the detector ($V_{\text{ex}}$) needs to be determined. By replacing the column with a union, $V_{\text{ex}}$ was determined to be 29 μL. The retention factor can then be calculated with Eq. 4.1.
Modelling SPAM

\[ k = \frac{V_R - V_0}{V_0 - V_{ex}} \] (4.1)

In Fig. 4.2A, the logarithm of the retention factor is shown for nine compounds at different \( \varphi \) for the analytical column and the trap column using the two experimentally determined values for \( V_0 \).

Figure 4.2. Comparison of retention factors of nine compounds on a trap column and an analytical column at different \( \varphi \). The dashed lines depict trap data, whereas the solid lines depict analytical columns. Lines are drawn for visualization purposes if more than one data point is found (i.e. not for crystal violet and propranolol). A) retention factors calculated using experimental \( V_0 \) values for both the trap and the analytical column; B, retention factors calculated using experimental \( V_0 \) value for the analytical column and a proportional \( V_0 \) value for the trap. Note the non-continuous \( x \)-axis.

The retention factors at different \( \varphi \) are seen to deviate systematically between the trap and analytical columns. A possible explanation may be found in the values used for calculating the retention factors. The experimental dead volumes (12 \( \mu \)L and 88 \( \mu \)L) do not seem to vary in proportional to the empty column volumes (17.3 \( \mu \)L and 173 \( \mu \)L, respectively). The columns are packed with a similar stationary phase and have an identical internal diameter. It is reasonable to assume the packing and the resulting porosity to be similar. Thus, the actual dead volume of the trap is expected to be one-tenth of that of the analytical column (8.8 \( \mu \)L). A possible explanation of the higher measured dead volume of the trap could be that there is a significant volume in the trap hardware that is not filled with stationary phase, such as the frit volume. While this is only 3 \( \mu \)L, it amounts to more than 30% of the actual \( V_0 \).
The corrected dead volume shows better agreement (Fig. 4.2B), with the retention factors from the trap overlaying almost exactly with the retention on the column. This demonstrates that if traps and columns are packed with identical stationary phases, there is no need to measure neither analyte retention factors, nor the dead volume on the trap. This is in agreement with our other work [40]. From measurements performed on the column, retention on the trap can be accurately predicted.

SPAM trap columns are typically employed in pairs in fast-paced LC×LC experiments. Consequently, for our protocol to be applicable, good trap-to-trap repeatability is required. To investigate this, retention data was predicted from and for different column-trap or trap-trap combinations. The results are shown in Fig. 4.3. On the left in Fig. 4.3, the experimental values for \(V_0\) are used for each column. There are large deviations when comparing retention values on either of the traps with those obtained on the column (compare Fig. 4.2A). The trap-to-trap repeatability is good, but extrapolation to column retention factors is error prone. Much smaller errors are observed when the corrected dead volume is used for the traps (Fig. 4.3, right).

Only at the extremes of the composition range studied (scarce data with very high or very low retention factors), the errors are seen to be significant. Deviations can be seen between trap A and trap B, likely due to the fact that a somewhat different dead volume (12 and 11 \(\mu\)L) was measured, respectively, while after correction an identical value of 8.8 \(\mu\)L was used.

After correction, the average deviations in retention factors between the column and trap are similar to the deviations between traps. In general, using uncorrected values yields significantly larger errors.

**Figure 4.3.** Percentual difference in retention between traps and columns at different organic-modifier concentrations with the measured \(V_0\) and the corrected \(V_0\).
4.3.2. Establishing parameters for retention prediction on trap columns

4.3.2.1. Model selection

Retention modelling can be used to predict isocratic retention on trap columns based on scanning gradients conducted on analytical columns. In principle, models constructed using gradient data can be used to predict isocratic retention times, although significant errors may be obtained [41]. Our earlier work suggested that the most-useful models to fit scanning-gradient data for RPLC were the log-linear or linear-solvent strength (LSS) model and the adsorption (ADS) model. Also, we concluded that the range of slopes of the scanning-gradients set should encompass or approach that of the optimized gradient [35]. The previous research, however, focused on method optimization for gradients on the same column, while in the present study we attempt to predict isocratic retention on much shorter columns than that used to construct the model.

To explain the relevance of any prediction errors, Fig. 4.4 shows the isocratic retention of acetaminophen on a trap column plotted with both the LSS and ADS model fitted to the data.

Overestimation of the retention, i.e. predicting that the analyte is retained whereas it is not, is a false positive in the context of the present study, as it results in a loss of analyte. This is visualized in Fig. 4.4 by the ADS model, which predicts an ln \( k \) of about 5 for \( \phi = 0 \), whereas the actual retention factor may be much lower. In contrast, underestimation of retention, i.e. predicting that the analyte is not retained, whereas in fact it is, can be regarded as a false negative. This would not lead to a loss of analyte since it is trapped better than anticipated.

The LSS model, however, cannot deal with any curvature in the data. For this reason, the quadratic model (QUA)\(^1\) was also investigated. For the obtained fitting parameters, Fig. 4.4 shows that the QUA model yields similar retention factors at \( \phi = 0 \) as the LSS model, while accounting for some curvature across the range of \( \phi \). For this reason, both the LSS and the QUA model were investigated in the remainder of this research. The appropriate equations for these two models are

\[
\ln k = \ln k_0 - S_{\text{LSS}}\phi \quad (4.2)
\]

\[
\ln k = \ln k_0 - S_1\phi + S_2\phi^2 \quad (4.3)
\]

\(^1\) Model described as Q model in Chapter 2 and 3. To be consistent in this chapter, three-letter abbreviations were chosen for all models.
where $\ln k_0$ represents the logarithm of the retention factor at an imaginary organic-modifier concentration of 0, the $S_{LSS}$ value or the $S_1$ and $S_2$ values describe how the retention changes with changing modifier concentration for the LSS model and QUA model, respectively.\(^2\)

![Graph showing fit of the LSS (blue), QUA (violet), and ADS (yellow) models to the isocratic retention data of acetaminophen on a C18 trap column against the volume fraction of ACN.]

When coefficients in the retention equation ("retention parameters") are to be obtained from gradient-elution data, we cannot directly fit the retention model. Instead, we have to use the more-complex equations that relate the retention time under gradient conditions to the retention parameters and the parameters that describe the gradient program. There is no analytical solution for these equations and an iterative numerical approach must be followed, starting from an initial estimate of the retention parameters. Such an approach may lead to a local optimum and multiple combinations of model coefficients can yield similar sum-of-squares errors [40]. For that reason, the model fit was optimized with a “multistart” function, i.e. multiple starting points in the parameter space are used to find a global optimum of the fit. The number of multistart positions was examined. With a lower and upper boundary of 0 and 100, respectively, for both the $S_{LSS}$, $S_1$ and $S_2$ parameters of the model and a -10 to 50 range for the value of $\ln k_0$, multistarts with 20, 40, and 100 positions were investigated.

\(^2\) Both $S_1$ and $S_2$ parameters are described as $S_{1,0}$ and $S_{2,0}$ parameter in Chapter 2 and 3. Since QUA is the only 3-parameter models in this chapter, both parameters are described as $S_1$ and $S_2$. 

124
The model was fitted five times and the standard deviation in the AIC value was used as indication of the correct number of starting points. The results are summarized in Fig. 4.5. For some compounds, such as crystal violet, phenol, toluene, and trimethoprim, the multistart function does not seem to have an effect on the deviation. However, for other compounds the multistart 20 gives rise to a higher AIC value (i.e. a worse fit). There is very little difference between 40 and 100, although the latter requires significant computational resources. Thus, a multistart with 40 positions was selected.

![Figure 4.5. Average AIC values for multistart 20 (blue), 40 (violet), and 100 (yellow) for all analysed compounds. Note that the AIC values for uracil exceed the y-axis.](image)

**4.3.2.2. Validity of using gradient data to predict isocratic retention**

To assess the validity of using gradient data to predict isocratic retention, isocratic retention was predicted from gradient data on an analytical column. Five different sets of gradient-scanning experiments were used throughout this study, as specified in Appendix B-3. Nine data points were used, comprising of either i) 3 repeat experiments of 3-, 6- and 9-min gradients (set 2), or ii) one of each 1-, 2-, 3-, 6-, 9-, 12-, 15-, 24-, and 48-min gradients (set 5). The prediction error ($\varepsilon_{\text{pred}}$) was calculated from

$$\varepsilon_{\text{pred}} = t_{R,\text{pred}} - \overline{t_{R,\text{exp}}}$$  (4.4)

where $t_{R,\text{pred}}$ is the predicted retention time and $\overline{t_{R,\text{exp}}}$ is the measured retention time (average of triplicate measurements). The resulting values are plotted for all compounds and a series of $\phi$ values in Fig. 4.6A.
Chapter 4

4.3.2.3. Effect of the dilution flow

The experimental setup used to record the data of Fig. 4.6B resembles that encountered with SPAM experiments in two-dimensional liquid chromatography, but it does not account for a possible dilution flow, which would result in altered flow rates and mobile-phase compositions. The purpose of diluting the 1D effluent is a reduction of the organic-modifier fraction (\( \phi \)) at the cost of an increase in flow rate (\( F \)). Both parameters exert opposite effects on the retention volume. By decreasing \( \phi \) in RPLC the retention is increased, as prescribed by the \( S_{LSS} \) parameter in the LSS model, whereas the increasing flow rate results in a reduced \( t_0 \) and thus a reduced \( t_R \). This net result of the two effects is displayed in Fig. 4.7.

**Figure 4.6.** Prediction error for the LSS and QUA model with gradient sets of 3, 6, and 9 min (set 2) or all gradients (1, 2, 3, 6, 9, 12, 15, 24, and 48 min; set 5) on the analytical column (A) and trap column (B). The colours indicate the different \( \phi \) levels (increasing from left to right within each cluster). The crosses indicate the average, while the points indicate outliers. The complete graph with all outliers is shown in Appendix B-4, Fig. B-2.

In Fig. 4.6A, all models and sets consistently show underestimation (\( t_{R,pred} < t_{R,exp} \)) for the lowest \( \phi \) values. This error is more severe for set 2 and slightly larger for the LSS model. When results for set 5 are compared, a better prediction is obtained using the QUA model. This confirms the impression of Fig. 4.4 that the QUA model yields better predictions in the low-\( \phi \) range. When the same experiment was repeated for retention prediction on trap columns, slightly larger prediction errors were observed (Fig. 4.6B). The quadratic model seems to yield more-useful predictions relative to the LSS model.
In Fig. 4.7 the retention time is depicted along the $\phi$ range for phenol and orange G. While the retention time of orange G increases with decreasing $\phi$ (corresponding with an increasing $F$), phenol shows a maximum in retention time. When the flow surpasses 0.250 mL·min$^{-1}$ retention starts to decrease with increasing flow, even though $\phi$ is still decreasing. Clearly, a dilution flow can have a counter-intuitive effect on the retention of an analyte on a SPAM trap. Therefore, it is dangerous to dilute samples limitlessly. In the LSS and QUA model, the parameters that describe the effect of the organic modifier are $S_{LSS}$ and the $S_1$ and $S_2$, respectively. When the retention parameters of these two compounds, calculated from the scanning-gradient data of the analytical column, are compared, they are found to be much lower for phenol ($S_{LSS}$: 8.60, $S_1$: 11.42, $S_2$: 8.53) than for orange G ($S_{LSS}$: 28.90, $S_1$: 41.47, $S_2$: 48.31). This result indicates that the effect of the dilution flow is related to the magnitude of the retention parameters. Compounds with high $S$-values will benefit from higher dilution flows, whereas for those with low $S$-values the effect may be opposite. The effects of a dilution flow can be described by the following formulae for the LSS model (Eq. 4.5) and the QUA model (Eq. 4.6).

\[
\frac{t_{R,q}}{t_{R,p}} = \frac{1 + \exp\left(\ln k_0 - S_{LSS} \cdot \frac{\phi}{1 + q}\right)}{1 + \exp\left(\ln k_0 - S_{LSS} \cdot \frac{\phi}{1 + p}\right)} \cdot 1 + p
\]  

(Eq. 4.5)

\[
\frac{t_{R,q}}{t_{R,p}} = \frac{1 + \exp\left(\ln k_0 - S_1 \cdot \frac{\phi}{1 + q} + S_2 \cdot \left(\frac{\phi}{1 + q}\right)^2\right)}{1 + \exp\left(\ln k_0 - S_1 \cdot \frac{\phi}{1 + p} + S_2 \cdot \left(\frac{\phi}{1 + p}\right)^2\right)} \cdot 1 + p
\]  

(Eq. 4.6)
Where \( t_{R,p} \) and \( t_{R,q} \) are the retention times of an analyte at dimensionless dilution flows (actual flow divided by \( F \) \( p \)) and \( q \). At a \( p \) value of zero the dilution factor \((1+p)\) equals 1 and the effluent is not diluted. Eqs. 4.5 and 4.6 are derived in Appendix B-5. The first (large) factor on the right-hand side of Eqs. 4.5 and 4.6 represents the effect of the mobile-phase composition on the retention time, whereas the factor \((1+p)/(1+q)\) represents the effect of the flow rate.

### 4.3.2.4. Model evaluation and comparison

Now that the effect of the dilution flow has been described, scanning gradients will be used to predict retention for four series of dilution-flow experiments. These series started with \( F = 50 \mu\text{L min}^{-1} \) and \( \varphi \) values of 0.75, 0.5, 0.25, and 0.1. All effluents were diluted 1:0, 1:1, 1:2, 1:3, 1:4, 1:6.5, 1:9, 1:14, and 1:19 (\( F:p \)), corresponding to total flow rates through the trap column of 50, 100, 150, 200, 250, 375, 500, 750, and 1000 \( \mu\text{L min}^{-1} \). The results for all compounds are shown in Fig. 4.8A. Similar results are found regarding underestimation of retention at low \( \varphi \) values with the LSS model and gradient set 2. The smallest errors were again found for set 5 in combination with the QUA model. The latter was thus selected for the remaining studies. It should be noted here that underestimation of retention (a negative value of \( \varepsilon_{\text{pred}} \)) at low \( \varphi \) values may have no consequences for the SPAM process, since retention at these values is already much higher than the minimum value needed for successful trapping. In Figs. 4.8B and 4.8C, the predicted and measured retention factors for series DF2 are shown for riboflavin and toluene, respectively. Riboflavin was hard to model (i.e. high AIC values). It can be seen in Fig. 4.8B that set 2 led to underestimation of the retention with both models. The most-accurate retention prediction was found for set 5 with the QUA model. In Fig. 4.8C the underestimation of the LSS model with set 2 is also evident for toluene, while the other predictions seem to be close to the measured values.

**Figure 4.8.** A) Prediction errors for all compounds calculated with the LSS and QUA models with a gradient set of 3, 6, and 9 min or with all gradients (1, 2, 3, 6, 9, 12, 15, 24, and 48 min); C18 trap column; four dilution-flow series (overlapping in the figure). The final composition is depicted on the horizontal axis for each cluster; the darkness of the bars also decreases with increasing \( \varphi \) values. The crosses indicate the average, the points indicate outliers. Fig. B-3 in Appendix B-6 shows the complete y-axis. B-C) Retention curves for riboflavin and toluene, respectively, in the DF2 series (starting value for \( \varphi = 0.5 \)).
4.3.2.5. Optimal gradient set

Now that the most-accurate model has been established, we focus on the selection of the most-accurate gradient set. Five sets were designed, that either consisted of three repeats of three gradients (1, 2, and 3 min, set 1; 3, 6, and 9 min, set 2; 9, 12, and 15 min, set 3; and 15, 24, and 48 min, set 4) or one repeat of nine different gradients (1, 2, 3, 6, 9, 12, 15, 24, and 48 min, set 5), always yielding a total input of nine datapoints. The retention was predicted for all analytes for all four dilution-flow series and the results are shown in Fig. 4.9A.

![Figure 4.9 A) Prediction error for all compounds calculated with five different gradient sets (three repeats of 1-, 2-, and 3- min; 3-, 6-, and 9-min; 9-, 12-, and 15-min; and 15-, 24-, and 48-min gradients or a single repeat of all nine gradients, set 1-5 respectively) for the trap columns subjected to four different dilution-flow series. The final composition is depicted on the horizontal axis for each cluster; the darkness of the bars also decreases with increasing values. The crosses indicate the average, the points indicate outliers. Fig. B-4 in Appendix B-7 shows the complete y-axis. B-C) Retention curves for orange G and 4-hydroxy benzoic acid, respectively, in the DF2 series (starting value for $\varphi = 0.5$).]

The results show that the steep gradients (set 1) yield unstable predictions. Furthermore, both set 2 and 5 yield underestimations of retention in the low-$\varphi$ range (i.e. $\varphi < 0.02$). Set 3 and 4 both seem to mainly overpredict retention on trap columns. It is possible that the underestimation in the low-$\varphi$ range with set 5 is due to the inclusion of the steep gradients, which are not part of set 3. The inaccurate prediction from the steep-gradient set (set 1, 2, 3 min) is confirmed by the compound-specific retention plots shown for orange G and 4-hydroxy benzoic acid in Fig. 4.9B and Fig. 4.9C, respectively. Using the steep-gradient set, the retention is either overestimated (Fig. 4.9B) or underestimated (Fig. 4.9C), while the other gradient sets all seem to yield a similar and more accurate prediction of the retention.

4.3.3. Development of a SPAM-prediction tool

Prediction of successful trapping on SPAM columns requires calculation of the minimal retention factor needed. Suppose we have a trap that, for a given modulation time and at a given flow rate, must retain an analyte. The retention factor ($k$) is related to the void volume ($V_0$) and the retention volume ($V_R$) as shown by Eq. 4.7.
\[ V_R = V_0 (1 + k) \tag{4.7} \]

For our trap, there will be a maximum volume \( V_{\text{max}} \) or time \( t_{\text{max}} \) at a given flow rate for which the analyte can be retained. There is a corresponding minimal retention factor \( k_{\text{min}} \).

However, \( V_{\text{max}} \) represents the apex of the analyte band or the chromatographic peak. We want the front section of the peak to not prematurely elute either. By taking into account the efficiency of the trap (standard deviation \( \sigma \)), this can be corrected for (Eqs. 4.8-4.10).

\[ k_{\text{min}} = \frac{V_{\text{max}}}{V_0} + 2\sigma - 1 \tag{4.8} \]

If the peak is Gaussian, this leads to

\[ k_{\text{min}} = \frac{V_{\text{max}}}{V_0} + \frac{2V_{\text{max}}}{\sqrt{N}} - 1 \tag{4.9} \]

\[ k_{\text{min}} = \frac{V_{\text{max}}}{V_0} \left( 1 + \frac{2 V_0}{\sqrt{N}} \right) - 1 \tag{4.10} \]

where \( N \) is the plate number. With a given modulation time \( t_{\text{mod}} \) and total flow rate \( F_{\text{tot}} \), defined as the sum of the 1D flow rate \( F \), and a dilution flow rate, \( F^D = p F \), Eq. 4.11 is obtained.

\[ k_{\text{min}} = \frac{t_{\text{mod}} \cdot (1 + p F)}{V_0} \left( 1 + \frac{2 V_0}{\sqrt{N}} \right) - 1 \tag{4.11} \]

Depending on which model is used, the retention of a compound can be described according to the LSS model (Eq. 4.12A) or the QUA model (Eq. 4.13A) if the dilution flow is 100% aqueous \( (p = 0) \) or for a different composition \( (p \neq 0, \text{Eq. 4.12B and 4.13B}) \).

\[ k_{\text{LSS},p} = \exp \left( \ln k_0 - S_{\text{LSS}} \cdot \frac{i \varphi}{p + 1} \right) \tag{4.12A} \]

\[ k_{\text{LSS},p} = \exp \left( \ln k_0 - S_{\text{LSS}} \cdot \frac{i \varphi + p D \varphi}{p + 1} \right) \tag{4.12B} \]

\[ k_{\text{QUA},p} = \exp \left( \ln k_0 + S_1 \cdot \left( \frac{i \varphi}{p + 1} \right)^2 + S_2 \left( \frac{i \varphi + p D \varphi}{p + 1} \right)^2 \right) \tag{4.13A} \]

\[ k_{\text{QUA},p} = \exp \left( \ln k_0 + S_1 \cdot \left( \frac{i \varphi + p D \varphi}{p + 1} \right)^2 + S_2 \left( \frac{i \varphi + p D \varphi}{p + 1} \right)^2 \right) \tag{4.13B} \]

We developed a tool to determine whether SPAM can be applied to a sample and what would be the optimal dilution flow \( F^D \) and modulation time \( t_{\text{mod}} \) [42]. For this tool both
the equation for the minimal $k$ is used, as well as the formula for the compound-specific retention (Eqs. 4.11-4.13). The user is asked to provide a list of compounds with corresponding retention parameters. Next to that, the user is asked to provide values for $^1F$, $^1\phi$, $F_{\text{tot,min}}$, $F_{\text{tot,max}}$, $t_{\text{mod,min}}$, and $t_{\text{mod,max}}$. Besides these parameters, the user needs to estimate the plate number ($N$) and the $V_0$ of the trap column. This volume should be estimated from the $V_0$ of the analytical column, as described in Section 4.3.1. The $\phi$ of the dilution flow can be adjusted by the user. The user is advised to use scanning gradients that cover a wide range of scanning gradients without using very steep gradients (see Section 4.3.2.5).

To test the effectiveness of this tool, the retention was predicted for four dilution flow series, all starting with a flow rate of 50 $\mu$L$\cdot$min$^{-1}$ and with $\phi$ values of 0.75, 0.5, 0.25, and 0.1. All effluents were diluted 1:0, 1:1, 1:2, 1:3, 1:4, 1:6.5, 1:9, 1:14, and 1:19, corresponding to total flow rates of 50, 100, 150, 200, 250, 375, 500, 750, and 1000 $\mu$L$\cdot$min$^{-1}$. The modulation time was set to 0.5, 1, and 2 min, and the retention of all compounds was predicted with the QUA model. Sets 1 to 5 were used to estimate the coefficients (retention parameters) of this model. Four possible outcomes were distinguished.

- Predicted and measured retention above the modulation time (true positive)
- Predicted and measured retention below the modulation time (true negative)
- Predicted retention above the modulation time and measured retention below the modulation time (false positive)
- Predicted retention below the modulation time and the measured retention above the modulation time (false negative).

For this tool to work, the ratios of the numbers of false positives and false negatives to the total number of measurements should be as low as possible. Both types of errors are undesirable. Therefore, the total fraction of false positives plus false negatives was considered in this work. This ratio was found to be 8.64%, 4.01%, 3.24%, 3.10%, and 2.94% for set 1, set 2, set 3, set 4, and set 5, respectively. However, this error was due to the compounds that were harder to fit the models to. This is illustrated in Fig. 4.10. There were no false positives or false negatives observed for phenol, toluene, uracil, and acetaminophen. Also, for crystal violet, only one false positive was found. However, because of the total number of measurements (6), this contributed significantly to the total percentage of false positives. There is no clear link between the goodness-of-fit and the rate
of false positive plus false negatives. The error rate along the different sets is similar for all compounds, except for those of set 1. This set consists of very steep scanning gradients. Set 2 yields better results than set 1, but the other three sets all perform better. It is surprising that set 5, with a single run of all nine gradients, yields the lowest error percentage on average. This could be because many points are available along the $\phi$ axis, so that the QUA model can fit the curve better, even though somewhat worse AIC values were observed than for all but the fastest gradient set (Fig. 4.10, right-hand scale).

![Figure 4.10. Percentage of false positives plus false negatives relative to the total number of measurements (bars). Retention times are predicted on the trap for four dilution-flow series at three modulation times, using retention parameters established from five different gradient sets (three repeats of 1-, 2- and 3- min, set 1; 3-, 6- and 9-min, set 2; 9-, 12-, and 15-min, set 3; and 15-, 24-, and 48-min gradients, set 4; or a single repeat of all nine gradients, set 5). The lines indicate the AIC values (right axis) for the specific compounds with the different gradient sets. Connecting lines between points are for visualization purposes only.](image)

In the present version of our tool, the composition of the 1D effluent is assumed constant during the run. In some LC×LC setups, for example with ion exchange or size-exclusion in the first dimension, this may be realistic [9,43]. However, in the more-common case where the 1D separation is a reversed-phase or hydrophilic-interaction liquid chromatography (HILIC) gradient, the organic-modifier concentration will vary with time. In that case, the correct 1D-effluent composition to use in Eqs. 4.12 and 4.13 is the elution composition of the specific analyte, which can readily be determined from the retention parameters of the analyte on the 1D column. A final difference between our current experimental setup and contemporary implementation of comprehensive two-dimensional liquid chromatography may be the possible incorporation of a mixer. A mixer may promote effective retention of the analyte on the trap, but it will contribute to the band broadening. Some of the experiments reported in this paper were repeated with a mixer incorporated. The results are documented in Appendix B-8, Fig. B-5. The early eluting peaks experience additional band broadening,
but the retention volumes are essentially unaffected. The additional band broadening can be accounted for in the model through the experimental value of σ or N. For later eluting compounds, the effects were more severe, but the consequences for predicting the success of trapping were minimal, because these compounds typically eluted much later than the modulation time.

4.4. Concluding remarks

In this research, a tool was developed that allows chromatographers to rapidly develop two-dimensional liquid chromatography methods with stationary-phase-assisted modulation (SPAM). The tool is publicly available on-line [42]. Retention modelling was applied to predict the feasibility of SPAM for a variety of analytes under a range of conditions. Experimentally obtained values for the dead volume of very small (trap) columns showed large errors. More-accurate values were obtained by extrapolating the values obtained on larger (analytical) columns. Among the retention models studied, the quadratic model yielded better predictions than the log-linear (“linear-solvent-strength”) model. Guidelines were formulated for scanning gradients. Short gradients gave rise to large prediction errors. Introducing variations in both flow rate and organic-modifier concentration, i.e. simulating dilution flows, caused a significant increase in prediction errors for all models and gradient sets analysed. Dilution with a weaker eluent is usually assumed to promote trapping, but it was shown that dilution may have an adverse effect for compounds that show modest retention. A tool was developed to help the analyst decide whether a specific analyte can or cannot be trapped. The tool yielded correct decisions in more than 95% of cases, for each set of gradients with either the quadratic or log-linear model.
Chapter 4

References


