ESCAPE-34 PSE-2024
European Symposium on Computer Aided Process Engineering
&
Process Systems Engineering

Flavio Manenti, Gintaras V. Reklaitis (Eds.), Book of Abstract of the 34th European Symposium on Computer Aided Process Engineering / 15th International Symposium on Process Systems Engineering (ESCAPE34/PSE24), June 2-6, 2024, Florence, Italy.

Global system analysis of on-line comprehensive two-dimensional liquid chromatography in gPROMS

Monica Tirapelle*, Dian Ning Chia, Fanyi Duanmu, Maximilian O. Besenhard, Luca Mazzei and Eva Sorensen

Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK
**m.tirapelle@ucl.ac.uk*

Abstract

Process system analysis and consideration of model uncertainty are important toolboxes used by chemical engineers to improve process knowledge, exploit the effect of the most important factors on several output variables, and investigate the inherent uncertainty of model predictions that may arise from uncertainty in input variables and model parameters. These analyses are even more important in regulated environments, where it is crucial to monitor process performance and product quality. In this work, we perform system analysis and model uncertainty analysis of comprehensive two-dimensional liquid chromatography systems that are commonly employed in pharmaceutical analysis. It is found that ^{1}D flow rate and ^{2}D column length are the main design parameters influencing the feasibility of a design. But most importantly, deviations from the optimal solution due to model uncertainty are shown to be minimal for the case studies considered.

Keywords: Two-dimensional liquid chromatography, in-silico method development, global system analysis, uncertainty analysis, hydrophobic-subtraction model.

1. Introduction

On-line comprehensive two-dimensional liquid chromatography (LCxLC) combines automated coupling of two liquid chromatography columns, ^{1}D (column 1) followed by ^{2}D (column 2), with different selectivities and different designs (Stoll, 2017). In the LCxLC mode, all the fractions of the ^{1}D effluent are transferred through a dual-loop modulation valve to the ^{2}D column such that the components co-eluting from the first dimension (^{1}D) are separated in the second dimension (^{2}D). LCxLC systems are employed in a wide field of applications, including biopharmaceuticals, environmental, food and synthetic polymers due to their high peak capacity (Pirok, et al., 2018b); however, LCxLC design and method development remains a challenging and complex task due to the large number of design variables and the interactions between the columns.

Many methods have been proposed seeking to develop a systematic procedure to replace trial-and-error approaches and one-variable-at-a-time strategies (Bedani, et al., 2012). Among these methods, some are based on the Poppe plot of plate time vs. plate height

(Poppe, 1997). These methods have been used to develop most of the protocols available today (Wang, et al., 2006; Schoenmakers, et al., 2006); however, as they are specific to the system considered, they lack generality. Other models use the Pareto-optimality method to find the best combination of parameters given two or more objective functions (Vivó-Truyols, et al., 2010; Sarrut, et al., 2015). In these studies, the objective functions are *sample-independent*, i.e., the authors focus on optimizing column efficiency and number of plates, but they do not consider parameters such as retention and selectivity. The only *sample-dependent* optimization strategies that have been proposed are by Pirok et al. (2018a) and Tirapelle et al. (2023).

This work applies our shortcut framework (Tirapelle, et al., 2023) for scenario analysis and for evaluation of model uncertainty. The scenario analysis will allow us to improve LCxLC process knowledge and understanding, but more importantly, to explore the design space *systematically*. The uncertainty evaluation allows us to investigate the impact of model inaccuracy on model predictions *quantitatively*. Both investigations are performed with the Global Systems Analysis (GSA) functionality of gPROMS ModelBuilder (Process Systems Enterprise, 2022).

2. Materials and methods

The shortcut model proposed by Tirapelle et al. (2023) consists of a set of analytical equations and relies on the Hydrophobic-Subtraction Model (HSM) (Wilson, et al., 2002; Snyder, et al., 2004) for the prediction of retention factors of different components in different reversed-phase (RP) columns. Due to its simplicity, the shortcut model can predict, in a matter of seconds, the position and band broadening of chromatographic peaks within the two-dimensional separation space of RPLCxRPLC systems. Combined with constraints on modulation time, maximum pressure drops, and minimum number of cuts per peak, the model has been embedded in a two-step framework for in-silico method development and optimization, and has been validated against rigorous numerical simulations based on the equilibrium dispersive model (EDM). However, the impact of model uncertainty and a proper exploration of the design space has yet to be considered and is therefore the focus of this work.

Here, we implement the shortcut model in gPROMS ModelBuilder, version 7.1.1 (Process Systems Enterprise, 2022), and we use the GSA functionality of gPROMS for: 1) *scenario analysis*, to fully exploit the design space and gain insights into RPLCxRPLC performance; and 2) *uncertainty analysis*, to investigate the effect of the main source of model uncertainty on the separation quality. For the uncertainty analysis, we consider the quasi-Monte Carlo method, with quasi-random (Sobol) sequences for sample generation (Process Systems Enterprise, 2022).

Throughout this work, we will consider four key performance indicators (KPIs), namely feasibility, total analysis time, overall separation quality, and number of components overlapping in the 2D chromatogram. For each design, feasibility suggests whether the design is off-spec (feasibility=0) or on-spec (feasibility=1) given the constraints on modulation time and pressure drop. The total analysis time, approximated by the analysis time of the ^{1}D column (Vivó-Truyols, et al., 2010), indicates how long the separation process lasts. Finally, the overall separation quality (Pirok, et al., 2016) and the number of overlaps (Tirapelle, et al., 2023), being both functions of the ^{2}D resolution (Schure, 1997), indicate how good the separation is.

3. Results and discussion

In the following sections, we will discuss the results of scenario (or parametric) analysis and model uncertainty analysis. As case studies, we will consider two of the mixtures used in Tirapelle et al. (2023), consisting of 8 and 16 components, all of which are to be separated. For the 8-component mixture, the columns considered are the Flare C18 (Diamond Analytic) and the Vydac 218MS (Grace/Vydac). For the 16-component mixture, the ZirChrom-PBD (ZirChrom) and the Primesep B (SIELC) columns are used (see Tirapelle et al., 2023 for further information).

3.1. Scenario analysis

As input factors for the scenario analysis, we consider the internal diameter, length, and flow rate of both ^{1}D and ^{2}D columns, as well as sample loop volume and pH pair. The input factors are assumed to have a uniform probability distribution, except for the pH pair, which has a discrete probability distribution. This is because only four combinations of pH pairs are possible for the case studies considered, since the column-selectivity database (Stoll, 2020) provides the column cation-exchange activity at just pH 2.8 and pH 7. The probability of occurrence of each pH pair is 25%. More information on input factors is summarized in Table 1.

Table 1. Input factors considered in the scenario analysis, with their relative symbol, probability distribution, and range of variability (bounds). Note that the pH pair indicates four possible combinations of pH, in order: (2.8, 2.8)=1, (7.0, 2.8)=2, (2.8, 7.0)=3, and (7.0, 7.0)=4.

Input factor	Symbol	Distribution	Bounds
^{1}D diameter (cm)	d1	Uniform	[0.1, 3.0]
^{2}D diameter (cm)	d2	Uniform	[0.1, 3.0]
^{1}D length (cm)	L1	Uniform	[2.0, 15.0]
^{2}D length (cm)	L2	Uniform	[2.0, 15.0]
^{1}D flow rate (mL/min)	FR1	Uniform	[0.05, 5.00]
^{2}D flow rate (mL/min)	FR2	Uniform	[0.05, 5.00]
Sample loop volume (mL)	LV	Uniform	[0.01, 1.00]
pH pair (-)	pair	Discrete	[1, 4]

Figure 1 shows the resulting parallel plots for the 8-component mixture (top) and 16-component mixture (bottom). In both cases, 20,000 different scenarios are evaluated, of which only 1.34% and 0.55% are meeting the specifications (dark blue lines). (These numbers clearly indicate why experimental LCxLC method development is so challenging, if not impossible, and why in-silico procedures are needed.) Interestingly, the top three variables impacting (at 5% statistical significance) the feasibility of a given design are, in order, the ^{1}D flow rate (mainly characterized by a ^{1}D flow rate smaller than 2 mL/min), ^{2}D column length (smaller than 12 cm) and ^{2}D column internal diameter (smaller than 2 cm). Furthermore, on-spec designs favor larger columns and up to four-fold smaller flow rates in the ^{1}D column compared to the ^{2}D column. Also of interest is the fact that it is most often preferable to operate both columns at low pH. Higher pH values in the ^{1}D column result in longer overall analysis time (not shown here), while operating the ^{2}D column at higher pH reduces the number of feasible solutions (see pH pairs 3 and 4). These results can be used to reduce the range of variability of the parameters in the subsequent optimal design procedure, improving convergence significantly and reducing the solution time for optimization, and thereby improving the accuracy of the resulting designs significantly. This approach will allow users to develop accurate RPLCxRPLC methods systematically and quickly, and to choose the best settings and columns without labor- and material-intensive trial-and-error strategies.

Having considered what design options are practically available, we now discuss how different on-spec designs affect the separation performance. Figure 2 shows the impact of ^{1}D column (left) and ^{2}D column (right) design parameters on overall separation quality and analysis time. The size and intensity of the bubbles refer to the column internal diameter and length, respectively (blue: 8 components; orange: 16 components). With the increase in the number of components, there are fewer on-spec designs available, and there is a significant decline in separation performance. Furthermore, more analysis time is required to achieve good separation quality. If we compare the two panels, we can see that the design of the ^{2}D column is most critical (i.e., the design space is smaller). Note that the impact of each parameter on the separation performance may differ between different samples. The only significant parameters (at 5% statistical significance) impacting the overall separation quality are pH pair, d1, d2 and FR2 for the 8-component mixture and L2 for the 16-component mixture. This result shows that each parameter may impact the separation performance differently when different mixtures are considered, indicating that *sample-dependent* (or *targeted*) optimization strategies should be preferred over *sample-independent* (or *untargeted*) optimization strategies.

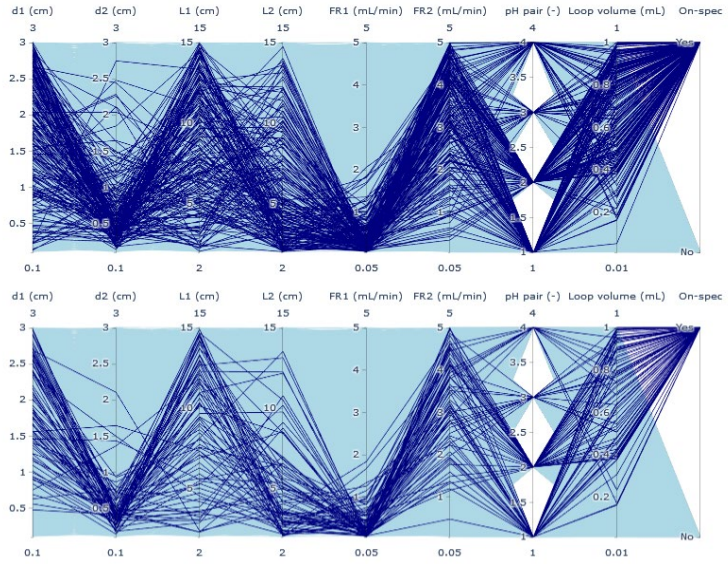

Figure 1. Parallel plots of the different scenarios tested for the 8-component mixture (top) and 16-component mixture (bottom). The dark blue lines refers to the feasible (on-spec) designs.

3.2. Uncertainty analysis

The shortcut model used in this work relies on the Hydrophobic Subtraction Model (HSM) for the prediction of the retention factors. Since the HSM is an empirical model with experimentally measured solute-specific parameters and derived column-specific parameters, it is subjected to inaccuracy. According to Wilson et al. (2002), the HSM model can predict the retention factor ${}^{x}k_i$ of component *i* in dimension *x* with a prediction accuracy of $\pm 0.7\%$. To evaluate how much this uncertainty in ${}^{x}k_i$ values impacts the response factors, and whether this impact jeopardizes the results of the shortcut model, we perform an uncertainty analysis.

For the uncertainty analysis, we consider the mixture of 8 components and the optimal design obtained by Tirapelle et al. (2023). All the ${}^{x}k_i$ values are assumed to follow a normal distribution with mean ${}^{x}\bar{k}_i$ and standard deviation ${}^{x}\sigma_i = 0.007$, while the number of uncertainty scenarios is set to 1000. Figure 3 shows the distributions of each KPI as a function of the uncertainty in the retention factors. The figure shows that the inaccuracy of the model does not impact the feasibility of the method (i.e., all scenarios are on-spec). The uncertainty results in a probability of occurrence of two- and four-component overlapping of 39.8% and 5.4%, respectively. However, the distribution of the overall separation quality suggests that the extent of the band overlap is minimal (expected value for the overall separation quality of 0.979); thus, components can still be separated but with a lower yield (i.e., overlapping parts will go to waste or will require reprocessing). This reveals that inaccuracies introduced by the underlying HSM model do not impact the performance of a given RPLCxRPLC system, thus demonstrating that the shortcut model can be safely used for in-silico RPLCxRPLC method development.

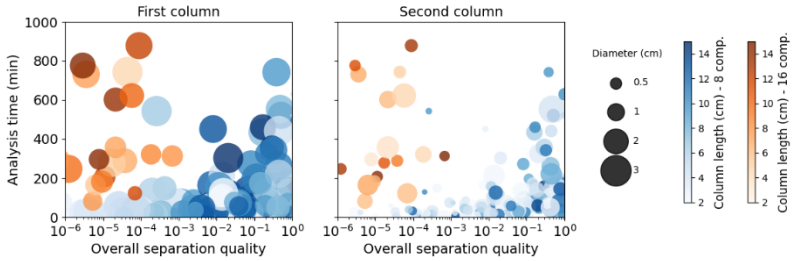

Figure 2. Bubble plot of the overall separation quality vs. analysis time of on-spec designs. The bubble size and color gradient refer to column internal diameter and column length, respectively, for the [1]D column (left) and [2]D column (right). The blue and orange color-maps refer to the 8-component and 16-component mixtures, respectively.

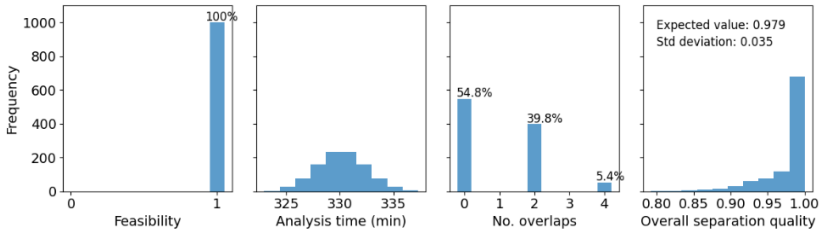

Figure 3. Distribution of feasibility, analysis time, number of overlaps and overall separation quality as a result of uncertainty in the retention factors of the Hydrophobic-Subtraction Model.

4. Conclusions

In this work, we have considered method development and optimization of on-line comprehensive two-dimensional liquid chromatography (LCxLC); in particular, we have demonstrated how global system analysis allows gaining insight into the performance of LCxLC systems and assessing the impact of the main sources of model uncertainty on the responses. Although model uncertainties impact the process performance, variation in the overall separation quality is limited, hence the proposed methodology is robust and permits developing fast and accurate two-dimensional designs.

Acknowledgement

The authors wish to acknowledge the financial support given to this research project by Eli Lilly and Company, and the Engineering and Physical Sciences Research Council (EPSRC), grant code EP/T005556/1.

References

F. Bedani, P. J. Schoenmakers, H.G. Janssen, 2012. Theories to support method development in comprehensive two-dimensional liquid chromatography - A review. Journal of Separation Science, pp. 1697-1711.

B. W. J. Pirok, A. F. G. Gargano, P. J. Schoenmakers, 2018a. Optimizing separations in online comprehensive two-dimensional liquid chromatography. Journal of separation science, 68-98.

B. W. J. Pirok, S. Pous-Torres, C. Ortiz-Bolsico, G. Vivó-Truyols, P. J. Schoenmakers, 2016. Program for the interpretive optimization of two-dimensional resolution. Journal of Chromatography A, pp. 29-37.

B. W. J. Pirok, D. R. Stoll, , P. J. Schoenmakers, 2018b. Recent Developments in Two-Dimensional Liquid Chromatography: Fundamental Improvements for Practical Applications. Analytical Chemistry, pp. 240-263.

H. Poppe, 1997. Some reflections on speed and efficiency of modem chromatographic methods. Journal of Chromatography A, pp. 3-21.

Process Systems Enterprise, 2022. gPROMS ModelBuilder Version 7.1.

G. Qian, A. Mahdi, 2020. Sensitivity analysis methods in the biomedical sciences. Mathematical biosciences, 323, 108306.

M. Sarrut, A. D'Attoma, S. Heinisch, 2015. Optimization of conditions in on-line comprehensive two-dimensional reversed phase liquid chromatography. Experimental comparison with one-dimensional reversed phase liquid chromatography for the separation of peptides. Journal of Chromatography A, pp. 48-59.

P. J. Schoenmakers, G. Vivó-Truyols, W. M. C. Decrop, 2006. A protocol for designing comprehensive two-dimensional liquid chromatography separation systems. Journal of Chromatography A, pp. 282-290.

M. R. Schure, 1997. Quantification of resolution for two-dimensional separations. Journal of Microcolumn Separations, pp. 169-176.

L. R. Snyder, J. W. Dolan, P. W. Carr, 2004. The hydrophobic-subtraction model of reversed-phase column selectivity. Journal of Chromatography A, pp. 77-116.

D. Stoll, 2020. HPLC columns - Column Selectivity Database. [Online] Available at: https://www.hplccolumns.org/database/ [Accessed 17 October 2023].

D. R. Stoll, 2017. Introduction to two-dimensional liquid chromatography - theory and practice. Handbook of Advanced Chromatography/Mass Spectrometry Techniques, Elsevier, 227-286.

M. Tirapelle, D. N. Chia, F. Duanmu, M. O. Besenhard, L. Mazzei, E. Sorensen, 2023. In-silico method development and optimization of on-line comprehensive two-dimensional liquid chromatography via a shortcut model, submitted for publication.

G. Vivó-Truyols, S. J. Van Der Wal, P. J. Schoenmakers, 2010. Comprehensive Study on the Optimization of Online Two-Dimensional Liquid Chromatographic Systems Considering Losses in Theoretical Peak Capacity in First-and Second-Dimensions: A Pareto-Optimality Approach. Analytical chemistry, pp. 8525-8536.

X. Wang, D. R. Stoll, A. P. Schellinger, P. W. Carr, 2006. Peak capacity optimization of peptide separations in reversed-phase gradient elution chromatography: fixed column format. Analytical chemistry, pp. 3406-3416.

N. S. Wilson, M. D. Nelson, J. W. Dolan, L. R. Snyder, R. G. Wolcott, P. W. Carr, 2002. Column selectivity in reversed-phase liquid chromatography: I. A general quantitative relationship. Journal of chromatography A, pp. 171-193.