Published in : Journal of Pharmaceutical and Biomedical Analysis (2024), 249, p. 116373
DOI: 10.1016/j.jpba.2024.116373
Status : Postprint (Author's version)

UNCERTAINTY MANAGEMENT FOR IN SILICO SCREENING OF REVERSED-PHASE LIQUID CHROMATOGRAPHY METHODS FOR SMALL COMPOUNDS

Thomas Van Laethem [a,b], Priyanka Kumari [a,b], Bruno Boulanger [c], Philippe Hubert [b], Marianne Fillet [a], Pierre-Yves Sacré [b], Cédric Hubert [b,**]

[a] *Laboratory for the Analysis of Medicines, University of Liège (ULiège), CIRM, Liège 4000, Belgium*
[b] *Laboratory of Pharmaceutical Analytical Chemistry, University of Liège (ULiège), CIRM, Liège 4000, Belgium*
[c] *Pharmalex Belgium, Mont-Saint-Guibert 1435, Belgium*

*** Corresponding author: chubert@uliege.be (C. Hubert).*

Keywords:

Reversed-phase liquid chromatography

Small pharmaceutical compounds

Response surface methodology

Multi-criteria decision analysis

Method development

Abstract

The process of developing new reversed-phase liquid chromatography methods can be both time-consuming and challenging. To meet this challenge, statistics-based strategies have emerged as cost-effective, efficient and flexible solutions. In the present study, we use a Bayesian response surface methodology, which takes advantage of the knowledge of the pKa values of the compounds present in the analyzed sample to model their retention behavior. A multi-criteria decision analysis (MCDA) was then developed to exploit the uncertainty information inherent in the model distributions. This strategic approach is designed to integrate seamlessly with quantitative structure retention relationship (QSRR) models, forming an initial in-silico screening phase. Of the two methods presented for MCDA, one showed promising results. The method development process was carried out with the optimization phase, generating a design space that corroborates the results of the selection phase.

Introduction

Liquid Chromatography (LC) is a highly versatile analytical technique widely used in the field of analytical chemistry for the identification and quantification of individual compounds in complex mixtures. LC has found applications in diverse fields including pharmaceuticals, biochemistry, environmental analysis, and food science.

The significance of method development methodologies in LC lies in the fact that separating complex mixtures can be challenging due to the potential need for different chromatographic conditions to effectively address the issue. In recent years, there has been an increasing interest in developing new methodologies that improve method development efficiency, speed, and accuracy [1,2]. Method development involves the optimization of chromatographic parameters to achieve the intended separation of the analytes of interest. This is critical in ensuring the accuracy and reliability of analytical results, as well as in maximizing the efficiency of the chromatographic process. Additionally, method development allows for the customization of the chromatographic method to suit the specific needs of a particular analysis, such as peak shape, signal intensity, and sensitivity, making it an essential aspect of LC [3].

Method development can be time-consuming and expensive as the number of experiments needed increases rapidly with the number of chromatographic parameters to optimize. Therefore, reducing the number of laboratory experiments is an important goal in LC method development, especially when dealing with complex samples. One approach to achieve this goal is to use a systematic approach, such as the Design of Experiments (DoE), to screen many parameters and identify the most influential factors [4].

DoE is a powerful statistical method that allows for the simultaneous variation of multiple factors to determine their individual and combined effects on the separation. Using a factorial or response surface design makes it possible to obtain a matrix of experiments that spans a large parameter space but with a minimal number of runs [5].

Another approach to reduce the number of experiments is to use a simulation model. Models using the quantitative structure-retention relationship (QSRR) approach can predict the chromatographic behavior based on the physicochemical properties of the analytes and the chromatographic conditions. Using simulations makes it possible to optimize the separation conditions before running any experiments,

reducing the time and resources required [6].

A simulation model is a powerful tool for predicting retention times in liquid chromatography. However, there are some limitations to this approach that must be taken into consideration.

One of the main limitations is the accuracy of the simulation models. While simulation models can accurately predict retention times for many compounds, there are some compounds for which it may not be accurate. This is especially true for compounds that have very different properties from the compounds used for the training, which may not behave according to the assumptions made by the simulation model [7].

Another limitation is the need for extensive calibration. To use simulation models for retention time prediction, it is necessary to first calibrate the models for each experimental setup using experimental data. This can be time-consuming and may require a large amount of experimental data, especially if the method is complex or if the compounds being analyzed are not well-characterized [6].

Finally, simulation models may not be able to account for all variables that can affect retention times in liquid chromatography. For example, the given descriptors may not encompass enough information about the retention mechanism. A limitation that may be lifted by graph neural networks [8].

Despite these limitations, simulation models can be very useful tools for retention time prediction in liquid chromatography, especially when used in conjunction with other experimental and computational methods [9].

Both DoE and simulation software can be very effective at reducing the number of experiments required for method development, while also improving the efficiency and robustness of the method. Additionally, various column screening approaches can be used to quickly identify suitable stationary phases and selectivity for the separation of complex mixtures [10].

In summary, reducing the number of experiments in LC method development can be achieved through a combination of statistical methods, and computer simulations. By employing these approaches, it is possible to develop methods that are efficient, robust, and reliable for the separation of complex mixtures. The combination of those approaches is what motivated our research. The association of individual QSRR models, capable of predicting the retention time of new compounds in specific conditions [11], with Response Surface Methodology (RSM), capable of predictions of the compounds in new conditions [12] is the approach chosen for in silico screening of chromatographic method. This integration of different techniques would be used as an initial step in method development to guide the coming steps and reduce the number of experiments needed.

In the prior research, the RSM was developed in the frequentist statistical framework and needed the building of a distribution to account for the error of the models during the Multicriteria Decision Analysis (MCDA) [12]. In this study, this concept is replaced by developing the RSM in a Bayesian framework. This change is mainly motivated by the interest in working with prediction distributions which are inherent elements of Bayesian models and therefore do not require additional construction when working directly with them. Those distributions improve the confidence in the results by giving a better understanding of their dispersion.

Indeed, Bayesian statistics can be a useful approach to enhance the Design of Experiments (DoE) methodology. Some benefits are that Bayesian methods provide a way to incorporate prior knowledge about the parameters of the model, which can help to improve the accuracy of the model. They can handle small sample sizes and missing data more effectively than traditional frequentist methods. Finally, they allow for the quantification of uncertainty in the estimates of the model parameters and the predictions, which can provide valuable information for decision-making. Indeed, working with mean responses, such as those provided by frequentist methodologies, may produce a good method on average but the variability in the results may fail to produce robust results. In opposition, some drawbacks are that Bayesian methods can be computationally intensive and require a higher level of statistical expertise compared to traditional frequentist methods. However, the selection of an appropriate prior distribution can present a considerable challenge and greatly influence the resulting outcomes. The latter drawback will be addressed in the study [13–15].

In case multiple criteria are important for the user, and a transparent and comprehensive decision-making process is desired, then MCDA could be a useful addition to the DoE methodology. Adjoining an MCDA to the DoE methodology can help to integrate multiple criteria, such as separation efficiency, speed of analysis, and robustness, into the experimental design, which can lead to a more comprehensive optimization of the method. It can facilitate the identification of the best compromise among different criteria and help to select a method that meets the needs of the user. It can provide a framework for transparent decision-making and communication of results to stakeholders. However, the MCDA framework can be complex and require a higher level of expertise than the DoE methodology alone. The identification and selection of relevant criteria and weights can be challenging and requires careful consideration of the needs of the user. The incorporation of MCDA can add additional steps to the method development process, which can increase the time and resources required for the optimization.

While some of the drawbacks have already been undertaken in our previous study, such as the choice of the criteria and the inconvenience of picking the weights. One of the aims of this study is to simplify the MCDA step of the proposed methodology.

The final stage of the strategy presented in this study will be devoted to presenting the succession of the screening and the optimization phase of the method development with the application of the design space principles [16]. This will allow the demonstration of the integration of the different steps in method development and the comparison of the results of the newly developed approach with a similar one.

The switch to the Bayesian framework and the optimization of the selected region are done in accordance with the Analytical quality by design (AQbD) principles. The AQbD is a systematic approach of method development and validation that aims to ensure the quality of analytical results by considering the entire lifecycle of the method, from development to routine use. The goal of AQbD is to design a method that is fit for its purpose and that meets the needs of the user, while also minimizing the risk of method failure or error by considering the process as a whole rather than as a succession of independent steps. This approach ensures the quality and reliability of the analytical results [17,18].

2. Materials and methods

2.1. METHODOLOGY

The improved methodology presented in this study extends upon the prior research. It applies a response surface methodology to formulate equations that are specific to various cases in which the effect of pH on a compound's retention time is analyzed. The predictions generated by the methodology in each condition are evaluated using different criteria, which are then processed through a multicriteria decision analysis, leading to the calculation of a desirability index to determine the optimal conditions.

The first significant improvement of the methodology is its development within the Bayesian statistical framework. Unlike the previous study, which used a frequentist approach and relied on fitting a Students' t distribution to each model to estimate the distribution of errors, the Bayesian framework inherently works with distributions and provides access to them through the posterior distribution with the chains from the sampler. Consequently, by relying on fewer mathematical constructs, the new methodology is simpler.

The second improvement involves the calculation of the criteria and the desirability index. Rather than being a single value, the criteria are now distributions. This modification takes advantage of the similarity between desirability functions and probabilities, resulting in the output of the desirability functions being replaced by the probability of reaching a specified threshold.

2.2. DATA

The data used for this study came from two preceding publications. They have in common that they were acquired on C18 columns Waters Corp. (Milford, MA, USA) XSelect HSS T3 2.1×100 mm 3.5 µm maintained at 25°C at a flow rate of 0.3 mL/min with Waters Corp. Alliance 2695 HPLC systems coupled with a UV–visible photodiode array detector 2996 module.

The training data of the models are composed of experimental retention times acquired for the ten conditions presented in a previous publication [19]. Those conditions are a combination of five pH values (2.7, 3.5, 5.0, 6.5 and 8.0) and two gradient times (20 and 60 minutes). The testing data were acquired at two conditions presented in the former research publication [12]. Those conditions were a combination of two pH values (3.0 and 6.0) and an intermediate gradient time of 40 minutes. The pH of buffers was measured using a SevenEasy S20 pH-meter (Mettler Toledo, Columbus, OH, USA).

The performance of the modeling part of the strategy was assessed with a first test set that is composed of 4-nitrophenol, ibuprofen, papaverine, and pindolol. 2,2′-bipyridine, metoclopramide, and verapamil were added to compose a more complex second test set, used for the application of the strategy presented in this study. All the compounds were selected for their diverse chromatographic behaviors. Molecular structures and physicochemical properties of the test set molecules are presented on Table S1 and Figure S1.

2.3. CHEMICALS AND REAGENT

Ammonium formate, and formic acid 99 % were purchased from VWR Chemicals (Leuven, Belgium). Ultrapure water was obtained from a Milli-Q Plus 185 water purification system from Millipore (Billerica, MA, USA). Methanol HPLC gradient grade was purchased from J.T. Baker (Deventer, Netherlands).

Standard compounds used during this study were 2,2′-bipyridine, 4nitrophenol, ibuprofen, metoclopramide, papaverine, verapamil hydrochloride acquired from TCI Europe (Anvers, Belgique); and pindolol from Abcam (Rozenburg, Netherlands). 2.4. Instrumentation and chromatographic conditions

The conditions of the optimization design were realized experimentally to check the prediction performance of new conditions and assess the correctness of the strategy. The samples were analyzed on the same system and in the same conditions as the data already acquired except for the parameters specified hereunder. The analytical method was a gradient from 0 % to 95 % methanol executed in different time duration (35, 48, and 60 min) followed by a 5 min hold. Different buffers at 0.1 mol•l^{-1} were used to set the pH of the mobile phase for the new conditions. Formic acid for a pH of 2.7, and ammonium formate with formic acid for pH of 3.2 and 3.5 The PDA was set to acquire data from 209 to 395 nm. Experimental retention times of the compounds in those conditions are available in Table S2.

2.5. SAMPLE PREPARATION

The stock solutions were prepared in water, methanol, or a mixture of both depending on the solubility properties of each compound. The subsequent mixture solutions were prepared by diluting the stock solutions using water or a mixture of water and methanol to reach the target concentration of 20 µg•mL^{-1}. A more concentrated solution (40 µg•mL^{-1}) was necessary to detect the ibuprofen. For detailed information about each compound preparation, see Van Laethem et al., 2022 [12].

2.6. SOFTWARE

R version 4.1.1 (2021–08–10) has been used for all the data preparation, model development and plotting. The Bayesian models were developed using rstanarm version 2.21.1 [20].

3. Results and discussion

3.1. BAYESIAN MODEL

3.1.1. METHOD DEVELOPMENT

As developed in a former work [12], one model is built for each compound individually to predict its retention behavior. Since the factors have more than two levels, the response surface methodology is applied. A set of three polynomial equations with different factors or power of factors has been defined: Eq. 1, Eq. 2, and Eq. 3. Those equations are defined to model the possible behaviors of different compounds regarding their pKa. Indeed, the effect of the pH on the retention time of compounds follows a sigmoidal curve. Depending on the range of pH investigated and the pKa of the analyzed compound, the whole or only a part of the sigmoidal curve can be observed. Firstly, if the pKa value is not inside the pH range investigated with a margin of 2 to account for the effect of pH values close to the pKa, Eq. 1 is used. In this study for a pH range from 2.7 to 8, the range with margins is from 0.7 to 10. Secondly, if the pKa value is inside the pH range and is close to the center of the range such as part of the two plateaus of the curve are inside the investigated domain, Eq. 3 is used. In this study, this corresponds to a pKa in the range from 4.35 to 6.35. Thirdly, if the pKa value is inside the pH range and not in the center, Eq. 2 is used.

$$\log(tR) = \beta_0 + \beta_1 \times tG + \epsilon \quad (1)$$

$$\log(tR) = \beta_0 + \beta_1 \times pH + \beta_2 \times tG + \beta_{11} \times pH^2 + \beta_{12} \times pH \times tG + \epsilon \quad (2)$$

$$\log(t_R) = \beta_0 + \beta_1 \times pH + \beta_2 \times t_G + \beta_{11} \times pH2 + \beta_{111} \times pH3 + \beta_{12} \times pH \times t_G + \epsilon \quad (3)$$

Comparatively to the frequentist models, Bayesian models require some prior information about the model's parameters. In this study, it has been defined following the recommendations of [21]. The prior function used is R^2, the proportion of variance in the dependent variable that is explained by the independent variables. It has the benefits of being a single prior function instead of defining priors for each coefficient independently of each other and it can have penalizing properties comparable to a ridge regression. It is set as R^2 with a location at 0.5. This is recommended as a "never wrong" value for a weakly informative prior.

Regarding the performance of the models, a root mean square error of prediction (RMSEP) lower than or equal to 2 minutes and a mean absolute percentage error of prediction (MAPEP) lower than or equal to 10 % were defined as targets.

3.1.2. CASE STUDY

The models in this study were developed using the first set of compounds. Table 1 presents the performance of these models. The high values of the coefficient of determination (R^2) indicate that the equation parameters explain most of the variance in the retention time. The RMSE and MAPE values for the calibration and prediction performance are comparable, indicating a lack of overfitting.

When comparing the performance of the Bayesian model in this study to the frequentist models in the previous study, it can be concluded that both strategies provide similar performance and that neither outperforms the other.

Table 2 provides a performance evaluation focused specifically on external testing conditions. The results demonstrate that the predictions are reasonably accurate. Although the table shows some positive values for the Bayesian models, suggesting a lower bias compared to the frequentist models which has all negative error values, the values are nonetheless quite similar and do not reveal a clear advantage for one strategy over the other. This negative bias is reflected on the global regression equation between predicted and observed values: $\hat{t}_R = 0.937\ t_R + 1.046$ where $\hat{t}_R$ represents the predicted retention times and t_R represents the observed retention times, it can also be seen on Figure S2. The standard deviation of this global regression is 0.730 min. In this case, the frequentist and Bayesian models have similar performances. It can also be noted that the prediction performances are close or within the target ranges that were chosen.

Table 1
Individual performance of models in calibration and prediction of the previous and the currently presented strategy.

	Compound	R_2	RMSEC	MAPEC	RMSEP	MAPEP	t_R [min]
Previous	**4-nitrophenol**	0.984	0.65	3.91	0.73	3.64	13.35
	Ibuprofen	0.999	0.35	0.75	0.83	2.25	20.22
	Papaverine	0.998	0.48	1.61	2.06	8.38	16.97
	Pindolol	0.996	0.27	1.39	0.75	5.47	10.83
Current	**4-nitrophenol**	0.984	0.80	4.07	0.80	3.76	13.35
	Ibuprofen	0.999	1.51	4.11	0.24	0.78	20.22
	Papaverine	0.997	1.36	4.35	1.87	5.56	16.97
	Pindolol	0.996	0.30	1.65	0.69	4.91	10.83

R^2: coefficient of determination; RMSEC: root mean square error of calibration; MAPEC: mean absolute percentage error of calibration; RMSEP: root mean square error of prediction; MAPEP: mean absolute percentage error of prediction. t_R is the retention time at pH 5.0 with a gradient time of 20 min, given for information.

Table 2
Absolute and relative errors of the compound of the first test set in the two external conditions. Condition 1: pH 3.0 and 40 min gradient time. Condition 2: pH 6.0 and 40 min of gradient time.

	Compound	Condition 1		Condition 2	
		Absolute error [min]	Relative error [%]	Absolute error [min]	Relative error [%]
Previous	**4-nitrophenol**	− 0.82	− 4.50	− 0.16	− 1.12
	Ibuprofen	− 0.10	− 0.30	− 0.88	− 3.03
	Papaverine	− 1.07	− 5.28	− 2.00	− 6.92
	Pindolol	− 0.56	− 4.24	− 0.65	− 4.18
Current	**4-nitrophenol**	− 1.13	− 6.22	− 0.13	− 0.87
	Ibuprofen	0.20	0.62	− 0.27	− 0.94
	Papaverine	− 0.21	− 1.06	− 2.64	− 9.13
	Pindolol	− 0.56	− 4.20	− 0.80	− 5.16

3.2. MULTICRITERIA DECISION ANALYSIS

3.2.1. METHOD DEVELOPMENT

To select the best condition for continuing the development of an analytical method, different criteria may be considered: the separation, the sensitivity to experimental parameters variations (robustness) and the analysis time. They are calculated in a manner that preserves the distribution of each criterion, preventing the loss of information that may occur if the distribution is reduced to a single value. The great advantage of having a distribution for each criterion is that the probability to be lower or greater than a specific value can be estimated. This would give stronger confidence in the decision that is made. Regarding this, the previously developed strategy could be applied [12], and a desirability function could be fitted to the distribution of criteria and would result in a distribution of desirability values. The other advantage of obtaining the distribution is that one can benefit from a property of the probabilities. As a matter of fact, the probability of an event is always between 0 and 1. Formerly, to calculate the desirability index, different desirability functions were designed to transform the values of all criteria between 0 and 1. Here, this step can be replaced by calculating the probability of an event around a specific value. This is a much simpler process because only a single value of acceptable threshold should be provided by the user instead of two, the best and worst values. The results of applying both previous and old strategies to the distribution of criteria will be presented and discussed.

The separation criterion is calculated for each condition by first determining the difference between all the pairs of retention times for each row in the chain of predicted retention times from the Bayesian models. Secondly, the second half width of the first peak and the first half width of the second peak are subtracted from this difference. Thirdly, the minimum of those distances for each row of the chain is determined and stored (Eq. 4).

$$\begin{aligned} &\mathrm{S_{p,g}} = \min_c \left(\mathrm{t_{R,i,c}} - \mathrm{t_{R,i+1,c}} - \mathrm{w_{0,i,l}} - \mathrm{w_{0,i+1,r}}\right), \\ &\mathrm{i} = 1, \ldots, \mathrm{M} - 1, \\ &c = 1, \ldots, C, \end{aligned} \tag{4}$$

where $S_{p,g}$ is the distribution of separation at pH p and gradient time g; $\min_c$ is a function that computes the minimum for each row c of the chains; $t_{R,i,c}$ is the retention time of compound i when ordered by decreasing retention time at row c; $w_{0,i,l}$ is the left baseline half-width of compound i; $w_{0,i+1,r}$ is the right baseline half-width of compound i+1; M is the number of compounds; and C is the number of rows in the chains.

The sensitivity to experimental parameters is calculated based on the separation criterion. It is the mean of the derivatives of the separation criterion in both directions of the analytical parameters. The derivative is calculated with the difference of the separation criteria between two juxtaposed conditions divided by the difference of the analytical parameter values (Eq. 5).

$$\mathrm{R_{p,g}} = \frac{|\mathrm{S'_{p,c}}| + |\mathrm{S'_{g,c}}|}{2}, c = 1, \ldots, C, \tag{5}$$

where $R_{p,g}$ is the distribution of sensitivity at pH p and gradient time g; $S'_{p,c}$ is the rate of change of the separation criterion in the direction of the pH for the considered condition at row c; $S'_{g,c}$ is the rate of change of the separation criterion in the direction of the gradient time for the considered condition at row c; and C is the number of rows in the chains.

The analysis time is defined as the highest retention time per row of the chain of predicted retention times and per condition (Eq. 6).

$$\mathrm{A_{p,g}} = \max_c(\mathrm{t_{R,i,c}}),\ \mathrm{i} = 1, \ldots, \mathrm{M},\ c = 1, \ldots, C, \tag{6}$$

where $A_{p,g}$ is the distribution of analysis time at pH p and gradient time g; $\max_c$ is a function that computes the maximum for each row *c* of the chains; $t_{R,i,c}$ is the retention time of compound i at row *c*; M is the number of compounds; and *C* is the number of rows in the chains.

The criteria are then combined in a desirability index. This desirability index is calculated with the weighted geometric mean of the criteria. For the probability strategy, the desirability index of each condition is a single value calculated on the probability to reach each criterion. For the desirability strategy, the desirability index of each condition is calculated for each row of the chains of criterion desirability. This last strategy calculations result in a distribution of desirability index for each condition.

The weights of criteria can be sampled from a Dirichlet distribution. The use of those sampled weights results in a distribution of desirability indexes which is less sensitive to small variations of those weights. The Dirichlet distribution is defined as the ratio of the different weights multiplied by a parameter defining the scale of the distribution that could be interpreted as the confidence the user has in the given weights' values. Once the distribution of desirability is calculated, the user can choose a value of desirability index and calculate the probability to reach that value. Then, by selecting a threshold of probability to reach, one or more regions of conditions will be selected in the knowledge space. One region, for example, the largest, can be selected to fit an optimization design around to continue with the optimization step.

3.2.2. CASE STUDY

The presented MCDA strategy is applied to the second set of test compounds. All the peak half-widths are set at 0.5 minutes. Fig. 1 shows the two possible strategies that result from using the distribution of predictions. The two contour plots illustrated were drawn with 10 levels each to allow the comparison of the density of the lines. For the probability strategy, the value 0 is used as a threshold to calculate the probabilities. It has been chosen as the limit case that would be interesting for a chromatographic method: a separation of 0 minutes at the baseline of the peaks. For this first criterion, the separation, it can be observed that using the probabilities to select the condition is much more restrictive. There is a large region at the bottom with probabilities close to 0 which is not present in the graph of the desirability. Moreover, while it is the same region that is chosen as the best regarding this criterion, the contour lines are much closer to each other on the probability plot than the desirability one. This indicates that a decision based on the probabilities is much more selective on the condition which should give the best results. This would have the consequence that the user can be much more confident on the results. It can also be noted that, although they are not clearly represented, the darker bands characteristic of the coelution of peaks can still be recognized in this figure.

The results of both possible strategies applied to the sensitivity to experimental parameters can be seen in Fig. 2. For this criterion, the value of 30 was chosen as threshold. This unitless value has no physical meaning. It is selected among the values taken by the criterion. Any value chosen around the middle of the range available should work. As it can be seen in Fig. 2, an intermediate value will result in an ample range of probabilities. Once again, both contour plots are traced to show 10 levels to ease the comparison. From those graphs, the same interpretation can be drawn for the proximity of the lines. The probability plot displays contour lines closer to each other for the high-value region than the desirability plot. This highlights again that using the probability of the criterion leads to a more restrictive outcome.

The third criterion, the analysis time, is presented with the application of the two possible strategies in Fig. 3. A maximum of 40 minutes of analysis time was chosen for the probability strategy. One more time, an intermediate value is chosen with the objective of producing a gradient of probabilities on a large surface of the modeled space. Those graphs exhibit the same effect as the other criteria. The contour lines of the probability plot are closer to each other compared to the desirability plot, indicating the more restrictive behavior of this strategy. Similarly, as the first criterion, extreme cases are displayed by large regions, another effect resulting from the contour lines being tight. Regarding the effect on the selection of a condition, the strategy based on probability will not discriminate conditions as long as the analysis time is below 40 minutes, contrarily to the strategy based on desirability which will discriminate and favor shorter analysis time.

Once they are calculated, the three criteria are combined in a global desirability index. The separation criterion has been given a weight of 1, the sensitivity to analysis parameters criterion has been given a weight of 0.5 and the analysis time criterion has been given a weight of 0.1. The charts ensuing from those calculations are presented in Fig. 4.

The desirability plot implies that the best condition is at an acidic pH and a very short retention time, which does not corroborate the experimental results of the dataset.

This result might be explained by the repartition of the values used to calculate the criteria. A first hint of this uneven distribution of data can be observed by comparing the range of desirability covered by each criterion. Indeed, the analysis time desirability in Fig. 3 covers a larger range than the separation desirability in Fig. 1. This effect leads to the analysis time having a larger effect on the desirability index than expected even though the criteria desirability are weighted.

From the probability plot, it can be deduced that the best condition is at a pH of 2.7 with a gradient time between 50 and 55 minutes. The added value of this strategy compared to the previous one is that no other condition that is not in the direct vicinity of this maximum is presented as having a close value of desirability index. With the previous strategy, two regions were selected as interesting with very high desirability index values but only one of them was actually acceptable. The other condition leads to the coelution of some of the peaks. Although some coelution might not be avoidable and could be resolved by other means such as peak deconvolution with the UV–visible spectrum, if a better condition exists, it should be the one that is selected. In this case, the results were satisfying enough as the experimental testing of the selected conditions still lead to fewer experiments than a screening plan. Comparatively to the previously presented strategy, the strategy based on probability gives results with more confidence and only selects the good region.

The last steps of the strategy will only be presented with the desirability index calculated on the probabilities as it is demonstrated to be the best strategy. An optimization design is developed around the selected region for a specific equation. These conditions were selected to investigate further around the selected region. Eq. 7 is selected as appropriate for the optimization phase. Indeed, the selected area is narrower than one pH unit. Consequently, only the linear and squared factors will be considered for the pH. On the other hand, considering that the height of the selected region is quite large, and that the literature shows the non-linear effect of the gradient time on the retention time, a squared slope factor is added. The interaction factor between pH and slope is still relevant, so it is kept.

$$\log(t_R) = \beta_0+\beta_1\times pH+\beta_2\times t_G+\beta_{11}\times pH_2+\beta_{22}\times t_{G2}+\beta_{12}\times pH \times t_G+\epsilon \quad (7)$$

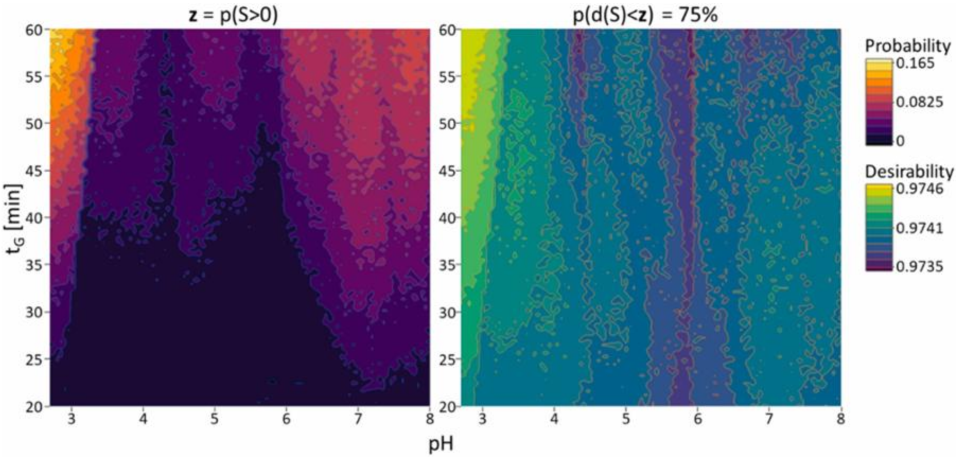

Fig. 1. The left graphic shows the probability that the minimum separation of a pair of peaks is greater than 0. The right graph shows the quantile 0.75 of the desirability of the separation criterion as it was calculated formerly.

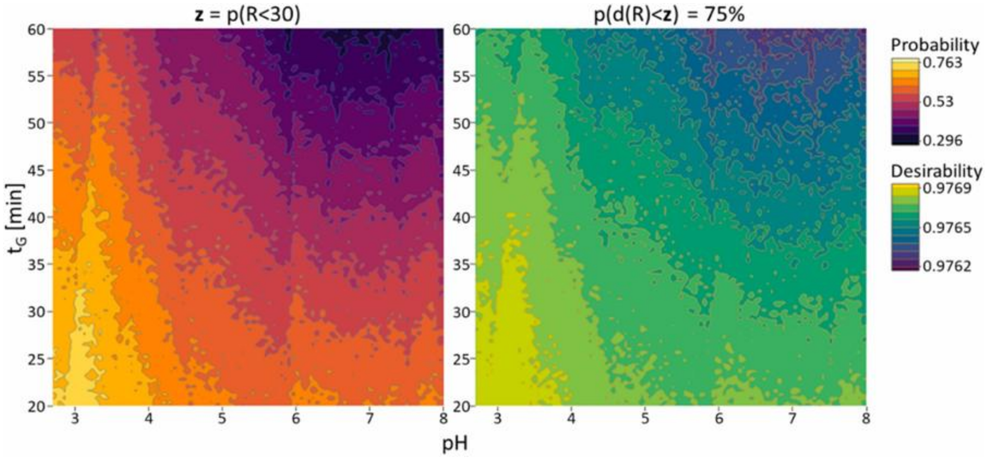

Fig. 2. The left graphic shows the probability that the sensitivity to experimental variations is lower than 30. The right graph shows the quantile 0.75 of the desirability of the sensitivity to experimental variations criterion as it was calculated formerly

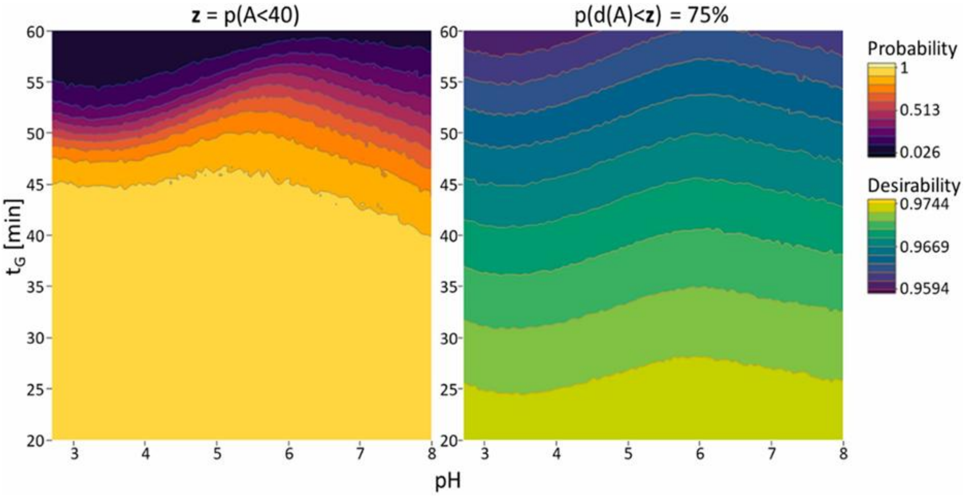

Fig. 3. The left graphic shows the probability that the analysis time is lower than 40. The right graph shows the quantile 0.75 of the desirability of the analysis time criterion as it was calculated formerly.

The design is limited to five points (conditions), the minimum needed to be able to fit the model. The results of the last part of this strategy are represented on Fig. 5.

The conditions of the optimization plan have been realized experimentally. The gradient time values have been rounded to the unit and those of the pH to the tenth of a unit.

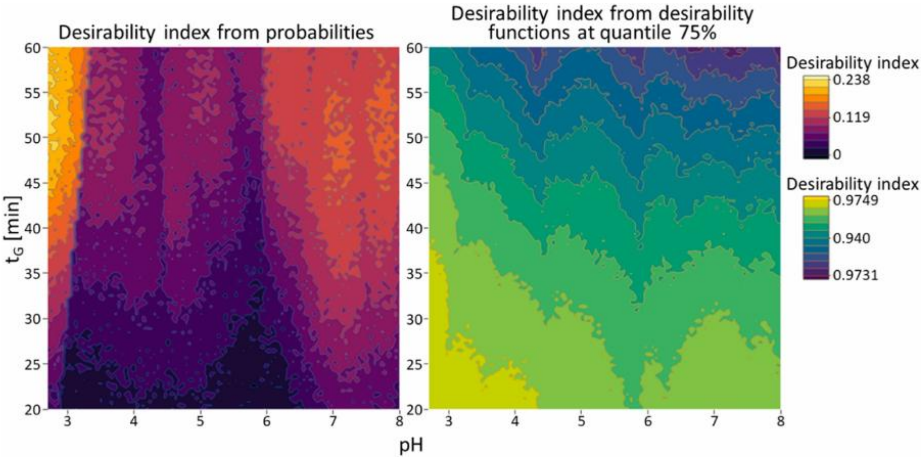

Fig. 4. The left graphic shows the desirability index calculated on the probabilities of the criteria. The right graph shows the quantile 0.75 of the desirability index calculated on the desirability functions of the criteria.

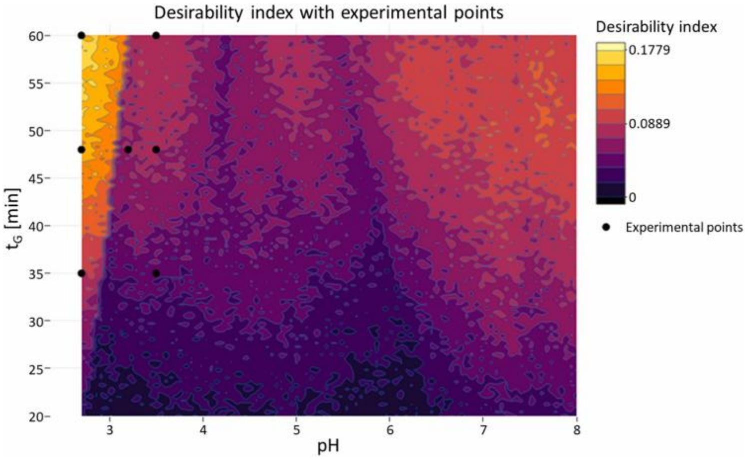

Fig. 5. Desirability index calculated on the probabilities of the criteria and the conditions selected for the optimization design (black discs).

It can be observed on the chromatograms (Fig. 6 that the elution order of the compounds does changes between the conditions at pH 2.7 and 3.2 and the conditions at pH 3.5, which illustrates well why the MCDA did not select this whole region as coelution-free. It can be noted that the worst condition regarding coelutions occurs at pH 3.5 and gradient time of 35 minutes where the strategy predicted that the results would not be the worst. Nonetheless, the experiments in less desirable conditions were needed to support the optimization models. Even if some conditions present some peaks coeluting, the optimal condition still needs to be defined in the optimization phase that follows.

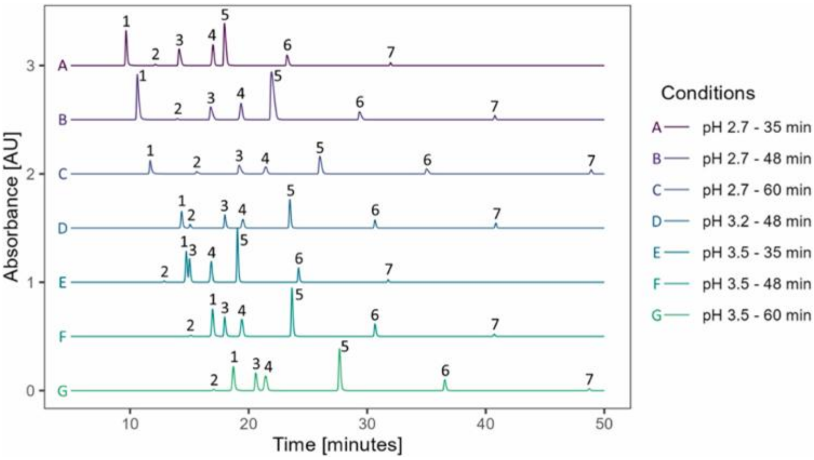

Fig. 6. The chromatograms extracted at 235 nm of the second test set analyzed in the different conditions of pH and gradient time. 1: 2,2'-bipyridine, 2: pindolol, 3: metoclopramide, 4: 4-nitrophenol, 5: papaverine, 6: verapamil, 7: ibuprofen

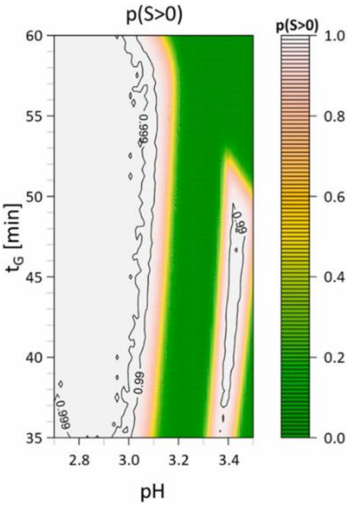

Fig. 7. The probability of the distance between the critical pair of peaks to be greater than 0 from the optimization design model with contour lines highlighting the regions where the probabilities are greater than 0.99 and 0.999.

Fig. 7 shows the design space as the probability of distance between the critical pair of peaks to be greater than 0 from the optimization models. Although the probabilities are different due to the models using more information (three retention times per peak), the resulting conditions selected to be optimal are in accordance with the results found after the screening but with more confidence, indicating the complementarity of the phases.

4. Conclusion

The previously developed approach has been improved and simplified regarding the number of steps to follow. Indeed, thanks to the Bayesian models, the model uncertainty is straightforwardly accessible, eliminating the construction of a distribution step. Moreover, the use of probabilities during the calculation of the criteria removes the need for a scaling step. Of the two models followed by the MCDA based on probabilities is better than the former strategy with frequentist models followed by an MCDA based on desirability functions. The new strategy is more restrictive on what constitutes a good condition. Compared to the initial strategy, which had two separated regions of conditions with similar desirability index values that had to be discriminated by experimental results, the new one clearly selects a single region as the best. Thus applying this strategy after a QSRR model could be used for in silico screening during method development. Using Bayesian models has one more advantage in that they are intrinsically capable of being updated. A new perspective could be to use the parameter distributions from the fitted models during this phase of development as prior information for the parameters of the models during the optimization phase during the experimental development of chromatographic method development.

CRediT authorship contribution statement

Thomas Van Laethem: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Cedric Hubert:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Pierre-Yves Sacre:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Conceptualization. **Marianne Fillet:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization. **Philippe Hubert:** Writing – original draft, Supervision, Resources, Methodology, Funding acquisition, Conceptualization. **Bruno Boulanger:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Priyanka Kumari:** Writing – review & editing, Data curation.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used Grammarly in order to comply with the publication standards for the language level. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Laboratory for the Analysis of Medicines & Laboratory of Pharmaceutical Analytical Chemistry reports financial support was provided by Fund for Scientific Research. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by FWO/FNRS Belgium EOS-programme, grant number 30897864 "Chemical Information Mining in a Complex World".

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jpba.2024.116373.

References

[1] I. Molnar,´ Computerized design of separation strategies by reversed-phase liquid chromatography: development of drylab software, J. Chromatogr. A. 965 (2002) 175–194, https://doi.org/10.1016/S0021-9673(02)00731-8.

[2] B. Debrus, P. Lebrun, J.M. Kindenge, F. Lecomte, A. Ceccato, G. Caliaro, J.M. T. Mbay, B. Boulanger, R. Marini, E. Rozet, P. Hubert, Innovative high-performance liquid chromatography method development for the screening of 19 antimalarial drugs based on a generic approach, using design of experiments, independent component analysis and design space, J. Chromatogr. A. 1218 (2011) 5205–5215, https://doi.org/10.1016/j.chroma.2011.05.102.

[3] F.T. Mattrey, A.A. Makarov, E.L. Regalado, F. Bernardoni, M. Figus, M.B. Hicks, J. Zheng, L. Wang, W. Schafer, V. Antonucci, S.E. Hamilton, K. Zawatzky, C. J. Welch, Current challenges and future prospects in chromatographic method development for pharmaceutical research, TrAC Trends Anal. Chem. 95 (2017) 36–46, https://doi.org/10.1016/j.trac.2017.07.021.

[4] L. Peng, X. Gao, L. Wang, A. Zhu, X. Cai, P. Li, W. Li, Design of experiment techniques for the optimization of chromatographic analysis conditions: a review, Electrophoresis 43 (2022) 1882–1898, https://doi.org/10.1002/elps.202200072.

[5] V. Czitrom, One-factor-at-a-time versus designed experiments, Am. Stat. 53 (1999) 126–131, https://doi.org/10.1080/00031305.1999.10474445.

[6] P.R. Haddad, M. Taraji, R. Szücs, Prediction of analyte retention time in liquid chromatography, Anal. Chem. 93 (2021) 228–256, https://doi.org/10.1021/acs. analchem.0c04190.

[7] E. Tyteca, M. Talebi, R.I.J. Amos, S.H. Park, M. Taraji, Y. Wen, R. Szucs, C.A. Pohl, J.W. Dolan, P.R. Haddad, Towards a chromatographic similarity index to establish localized quantitative structure-retention models for retention prediction: Use of retention factor ratio, J. Chromatogr. A. 1486 (2017) 50–58, https://doi.org/ 10.1016/j.chroma.2016.09.062.

[8] A. Kensert, R. Bouwmeester, K. Efthymiadis, P. Van Broeck, G. Desmet, D. Cabooter, Graph convolutional networks for improved prediction and interpretability of chromatographic retention data, Anal. Chem. 93 (2021) 15633–15641, https://doi.org/10.1021/acs.analchem.1c02988.

Published in : Journal of Pharmaceutical and Biomedical Analysis (2024), 249, p. 116373
DOI: 10.1016/j.jpba.2024.116373
Status : Postprint (Author's version)

[9] P. Wiczling, R. Kaliszan, How much can we learn from a single chromatographic experiment? A Bayesian perspective, Anal. Chem. 88 (2016) 997–1002, https://doi.org/10.1021/acs.analchem.5b03859.

[10] F. Gritti, Perspective on the future approaches to predict retention in liquid chromatography, Anal. Chem. 93 (2021) 5653–5664, https://doi.org/10.1021/acs. analchem.0c05078.

[11] P. Kumari, T. Van Laethem, D. Duroux, M. Fillet, P. Hubert, P.-Y. Sacre, C. Hubert,´ A multi-target QSRR approach to model retention times of small molecules in RPLC, J. Pharm. Biomed. Anal. 236 (2023) 115690, https://doi.org/10.1016/j. jpba.2023.115690.

[12] T. Van Laethem, P. Kumari, B. Boulanger, P. Hubert, M. Fillet, P. Sacr´e, C. Hubert, User-driven strategy for In Silico screening of reversed-phase liquid chromatography conditions for known pharmaceutical-related small molecules, Molecules 27 (2022) 8306, https://doi.org/10.3390/molecules27238306.

[13] P. Wiczling, Analyzing chromatographic data using multilevel modeling, Anal. Bioanal. Chem. 410 (2018) 3905–3915, https://doi.org/10.1007/s00216-0181061-3.

[14] A. Kamedulska, Ł. Kubik, J. Jacyna, W. Struck-Lewicka, M.J. Markuszewski, P. Wiczling, Toward the general mechanistic model of liquid chromatographic retention, Anal. Chem. 94 (2022) 11070–11080, https://doi.org/10.1021/acs. analchem.2c02034.

[15] J.J. Peterson, M. Yahyah, A Bayesian design space approach to robustness and system suitability for pharmaceutical assays and other processes, Stat. Biopharm. Res. 1 (2009) 441–449, https://doi.org/10.1198/sbr.2009.0037.

[16] P. Lebrun, B. Boulanger, B. Debrus, P. Lambert, P. Hubert, A bayesian design space for analytical methods based on multivariate models and predictions, J. Biopharm. Stat. 23 (2013) 1330–1351, https://doi.org/10.1080/10543406.2013.834922.

[17] T. Tome, N. Zigart,ˇ Z. Casar, A.ˇ Obreza, Development and optimization of liquid chromatography analytical methods by using AQbD principles: overview and recent advances, Org. Process Res. Dev. 23 (2019) 1784–1802, https://doi.org/ 10.1021/acs.oprd.9b00238.

[18] L.E. Blue, X. Guan, M.K. Joubert, S.T. Kuhns, S. Moore, D.J. Semin, M. Wikstrom,¨ J. Wypych, C.T. Goudar, State-of-the-art and emerging trends in analytical approaches to pharmaceutical-product commercialization, Curr. Opin. Biotechnol. 78 (2022) 102800, https://doi.org/10.1016/j.copbio.2022.102800.

[19] T. Van Laethem, P. Kumari, P. Hubert, M. Fillet, P.-Y. Sacr´e, C. Hubert, A pharmaceutical-related molecules dataset for reversed-phase chromatography retention time prediction built on combining pH and gradient time conditions, Data Br. 42 (2022) 108017, https://doi.org/10.1016/j.dib.2022.108017.

[20] B. Goodrich, J. Gabry, I. Ali, S. Brilleman, rstanarm: Bayesian applied regression modeling via Stan., (2020). 〈https://mc-stan.org/rstanarm〉 .

[21] J. Gabry, B. Goodrich, Estimating Regularized Linear Models with rstanarm, (2022) 1–7. 〈 https://cran.r-project.org/web/packages/rstanarm/vignettes/lm. html〉 (Accessed 15 November 2022).