UPPSALA
UNIVERSITET

UPTEC F 23004
Examensarbete 30 hp
januari 2023

Detecting chromatography unit degradation

Comparison of single- and multi-point techniques implemented in system control and monitoring software

Max Markensten

Civilingenjörsprogrammet i teknisk fysik

UPPSALA
UNIVERSITET

Detecting chromatography unit degradation

Max Markensten

Abstract

Chromatography units, used in the production of pharmaceuticals, degrade with use and need to be changed or repackaged. This study investigates the effectiveness of two statistical methods, principal component analysis and simple and one-point multiparameter technique, for determining degradation in the Fibro chromatography unit. The methods have been shown to be effective on resin chromatography columns but not before tested on the relatively new Fibro chromatography unit. The statistical methods are implemented in an unreleased version of the monitoring and control software Unicorn. This implementation aims to be a proof of concept for including more complex methods for monitoring runs directly in the software, easing the workflow of operators by removing the need to export measurements to a third-party program. The methods were tested on measurements of absorbance, conductivity, and pressure from two series of chromatograms performed on two Fibro chromatography units. One of the units was defective and broke down halfway through the series. Principle component analysis could clearly visualize a difference between early and late runs on the defective unit. The same could only be achieved for the non-defective unit by excluding measurements of pressure. Simple and one-point multiparameter technique visualized trends from early to late in the series which were much clearer for the defective unit. Both methods showed signs of predicting degradation in a Fibro chromatograpy unit but require validation on chromatogram series with more direct measurements of performance and a wider range of failure causes.

Teknisk-naturvetenskapliga fakulteten

Uppsala universitet, Utgivningsort Uppsala/Visby

Handledare: Henrik Nässén Ämnesgranskare: Mikael Sternad

Examinator: Tomas Nyberg

1 Populärvetenskaplig sammanfattning

Modern läkemedelstillverkning använder sig av genmodifierade bakterier för att producera till exempel insulin åt diabetiker. Bakteriernas gener ändras så att de bildar insulin. De genmodifierade bakterierna odlas i stor skala i bioreaktorer - samma princip som när jäst förökar sig och producerar alkohol på ett bryggeri. Insulinet finns i bakterierna men behöver helt separeras från dem före användning. Bakterierna löses därför upp till en sörja och pumpas genom en kromatografienhet som låter allt passera förutom insulin. I enheten är porösa gelkulor packade och på gelkulorna sitter en molekyl (ligand) som binder till just insulinet. Efter att insulin har fastnat spolas resten av bakteriesörjan ut. Därefter ändras surheten för att insulinet ska släppa genom att en lösning pumpas in i kromatografienheten. Insulinet kommer ut ur kromatografienheten och kan användas i läkemedel.

En kromatografienhet kan inte användas hur många gånger som helst. Hålrummen i de porösa kulorna sätts igen så att vissa ligander aldrig får kontakt med det de ska fånga upp. Kanaler bildas mellan gelkulorna så att bakteriesörja passerar utan att få kontakt med lika många ligander. Ligander slits ut av surhetsförändringar och blir oanvändbara för att de inte längre kan binda. Efter ett tag behöver enheten bytas ut eller gelen packas om.

Den här studien undersöker två olika metoder för att upptäcka problem i en kromatografienhet. Metoderna har tidigare testats på kromatografienheter med gelkulor och fungerat bra. Den här studien undersöker istället hur de fungerar på en ny sorts kromatografienhet som innehåller överlappande cellulosasträngar där ligander fästs, istället för på gelkulor.

För att vara så smidiga att använda som möjligt för operatörerna i läkemedelstillverkningen så programmeras metoderna in i mjukvaran som styr de pumpar och mätinstrument som används för att genomföra kromatografin. Mätinstrumenten mäter bland annat tryckskillnaden över kromatografienheten och absorbans och konduktivitet i vätskan som kommer ur den. Dessa mätningar görs ungefär tio gånger i sekunden.

Båda metoderna analyserar mätningarna som görs av tryck, absorbans och konduktivitet när insulinet, eller annan läkemedelsmolekyl, kommer ut ur en kromatografienhet. Båda metoderna antar också att förändringar i kolonnen som gör att den fungerar sämre kommer ge upphov till förändringar i mätningarna. Den ena använder principal komponentanalys för att koka ner hundratals mätningar under en körning till en prick på en tvådimensionell graf som kan jämföras med prickar från körningar man vet fungerat bra. Den andra metoden tittar bara på förändringar i mätningen i en enda punkt mellan körningar.

Båda metoderna kunde visualisera anomalier i tryckmätningar när en kromatografienhet höll på att gå sönder. Att skilja körningar som gjordes tidigt från de som gjordes sent på samma enhet var dock mycket lättare om tryckmätningar uteslöts. Studien tyder på att de här metoderna skulle kunna fungera på ett liknande sätt för fiberkromatografi som för gelkromatografi. Det som saknas är bekräftelse på att de överrensstämmer med mer direkta mått på dålig funktion än om enheten varit i bruk länge eller snart helhavererar.

Contents

2 Introduction

The production of many pharmaceutical products requires a concentrated sample of a single molecule. Chromatography is a group of techniques for separating parts of a mixture. It can be used to produce concentrated samples by separating the parts of a mixture based on their size, charge or affinity for binding to another molecule. One way of separating an antibody from a mixture is with a liquid chromatography unit containing a resin coupled with a ligand. The ligand is a protein which binds well to the antibody, keeping it in the resin. A liquid sample containing both the desired antibodies and unwanted byproducts passes through the unit filled with resin. The antibodies bind to the ligand in the resin and are thereby separated from the original mixture.

Using statistical techniques to determine when to change chromatography unit can reduce waste and costs in the production of pharmaceuticals. The aim of this study is to explore the possibility of implementing principal component analysis for use with chromatograms, directly in the control and monitoring software of Cytiva chromatography systems. The effectiveness of the multivariate principal component analysis technique is compared to the less data-demanding simple and one-point multiparameter technique. The comparisons are performed on the Fibro chromatography unit (which replaces the normal porose resin with a cellulose matrix), on which they had not yet been tested.

The performance of chromatography units decrease over time and they need to be replaced to avoid waste of the molecule being separated or producing contaminated end products. If the unit is evaluated based on produced batches, valuable ingredients and production time will be wasted on batches that have to be dicarded. Resin is also a substantial part of the production cost, according to a 2015 article on re-use of Protein A resin [1], which means that changing unit or repacking the unit with new resin prematurely should be avoided if possible. It would be preferable to change the resin after it has been used as much as possible but before it produces a batch of low quality. This study investigates 2 methods of detecting signs of column degradation using only the built in sensors in the chromatography instruments used in production, thereby saving time and costs compared to separate tests such as analytical assays.

The 2 implemented techniques are principal component analysis and simple and one-point multiparameter technique applied to measurements of absorbance, conductivity, and pressure from the elution phase of a series of chromatography runs. The elution phase is when the desired molecules are released from the chromatography unit. Principle component analysis is a multivariate statistical technique for dimension reduction. In this study it is used to represent a chromatography run as a point in 2 dimensions while retaining information from hundreds of measurements of 3 variables (absorbance, conductivity, and pressure). The visualization of runs in 2 dimensions allows operators to notice changes occuring in the chromatography unit. Simple and one-point multiparameter technique is a visual tool for detecting changes in function of a chromatography unit by comparing measurements of absorbance, conductivity and pressure at a single point of the elution phase. This point is chosen where the highest absorbance value is measured. If drastic changes occur in the measurements at this point from one chromatogram to the next, it is a sign that degradation is occuring in the unit.

The techniques are meant to be used in a production setting, alerting operators early to problems. Performing the calculations for the techniques in the chromatography systems software keeps it quick and practical compared to exporting chromatograms to other software for analysis.

3 Background and Theory

3.1 Chromatography

3.1.1 Resin chromatography

Resin chromatography is a way of separating molecules with different properties. Separating molecules is an important step in the production of pharmaceuticals that uses bacteria, yeast or other single cells. One example is the protein hormone insulin used to treat diabetes which has been produced by genetically modified bacteria. One challenge faced in the production process is that bacteria produce other molecules which need to be removed before insulin can be injected into a patient. The bacteria are subjected to lysis, breaking down the cell membrane, and releasing its contents into a liquid. The molecule required for the medicine then needs to be separated from the other molecules from the lysated bacteria.

The resin is a collection of porous beads which the lysated bacteria molecules can travel through. The resin is tightly packed in a column where liquid enters in one end and leaves in the other. A chromatography column attached to a chromatography system is pictured in Figure 1. There are many ways that resins packed in a column can separate molecules. Size exclusion chromatography uses a resin which molecules of different sizes move through at different speeds. By timing when the output is collected relative to when a sample entered the column, molecules of a certain size can be extracted from a sample. Ion exchange chromatography separates molecules based on charge. Protein affinity chromatography separates molecules based on how well they bind to a specific molecule, a ligand, attached to the resin beads. Protein affinity chromatography is used in the chromatography runs considered in this study. However, in this study the ligands are not attached to resin beads but to the fibers in a Fibro chromatography unit discussed in the next section.

Figure 1 shows an Äkta Pure system with a resin filled chromatography column attached. The system includes pumps and valves to control what flows through the column and at what rate. Measurements of physical properties, like pressure in different parts of the system and conductivity of the liquid passing through the column, are made automatically. The behaviour of the system can be controlled and the measurements accessed through a system control and monitoring software.

Figure 1: Äkta pure system with resin chromatography column attached. Image belongs to Cytiva.

3.1.2 Fibro chromatography

Fibro chromatography units are a new range of products performing a similar function as protein affinity resin chromatography columns. Instead of resin beads packed in a column, a Fibro chromatography unit contains electrospun cellulose based strands forming a matrix. These strands are coupled with a ligand that binds to the desired protein in certain conditions, such as in a certain pH-range, and can be induced to let go of it in other. Instead of the liquid sample moving in and out of porous beads, it is pushed through the matrix. The desired proteins in the sample attach to the ligand on the strands. A Fibro chromatography run is faster than a column chromatography run but Fibro units are currently limited to volumes smaller than 4ml and therefore not suitable for large scale production [2, 3]. Figure 2 illustrates what a Fibro unit looks like on the outside and inside.

(a) Computer rendering of a HiTrap Fibro unit. Tubing from the chromatography system is connected to the top and bottom.

(b) The matrix of cellulose fibers contained inside a Fibro unit.

Figure 2: Images belong to Cytiva.

Equilibration	Sample application	Wash 1	Wash 2	Elution	Re-equilibration

Table 1: The six phases of the chromatography runs used in this study

In the beginning of a chromatography run, a solution is passed through the unit which makes the conditions favorable for the desired protein to bind to the ligand. The sample which the protein is to be extracted from is then pumped through the unit. The protein binds to the ligand and everything else from the sample is washed out in a washing phase. Another solution is passed through the unit, changing the conditions such that the ligand releases the protein in what is called the elution phase. The unit is then cleaned and prepared to bind the desired protein from the next sample. The order and names of the phases of the chromatography runs on Fibro units included in this study can be seen in Table 1.

3.1.3 Chromatograms

A chromatogram is a series of measurements automatically made while chromatography is being performed. For chromatograms used in this study, measurements are made with a frequency of roughly 12Hz throughout the entire chromatography run. The analysis in this thesis focuses on the elution phase where measurements of absorbance, conductivity, and pressure are used. Measurements of these variables in the elution phase have been shown to correlate with degradation in resin chromatography columns in a 2020 study by Feidl et al [4]. The absorbance of light with a wavelength of 280nm and conductivity are measured in the solution leaving the chromatography unit. The pressure measurment used is the difference in pressure over the chromatography unit.

Figure 3 shows the approximate behaviour of the absorbance, conductivity, and pressure values during the elution phase of a run on a Fibro chromatography unit.

Typical peaks and troughs in the Elution phase

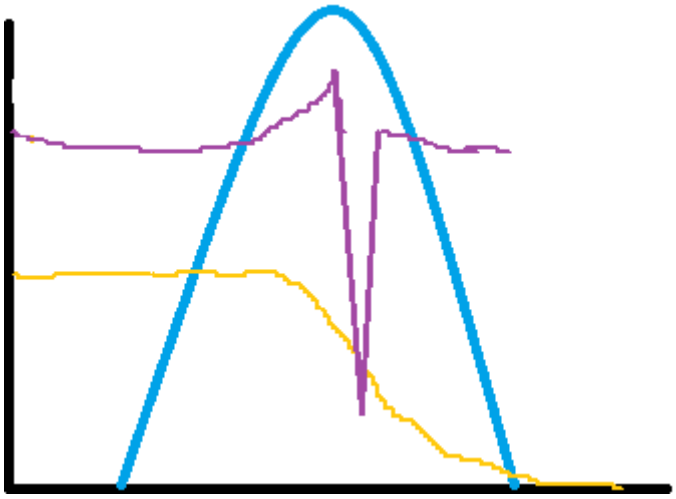

Figure 3: During an elusion phase absorbance (blue) shows one clear peak. Pressure (purple) shows a momenatry trough at the same time as absorbance peaks. Conductivity (yellow) decreases while absorbance peaks.

3.1.4 Chromatography unit degradation

In chromatography columns, performance decline can be caused by ligands not binding to the target protein or by the changes in the resin which limits contact between the ligands and target proteins. An article by Nweke et. al from 2018 lists the leading causes of chromatography unit degradation. The number of ligands available to bind can decrease because of "ligand leaching off of the matrix", the ligands being blocked by a non-target molecule, or chemical inactivation of the ligand. Deformation of the matrix holding the ligands can prevent some ligands from being reached by the target protein [5]. Molecules from the sample getting stuck in the pores of the resin has also been identified as a cause of resin degradation [6].

3.2 Statistical Techniques

3.2.1 Principal component analysis

Principal component analysis is a technique for reducing the number of dimensions in a dataset. A new basis is found for the dataset. The new basis vectors are ordered by how much of the variance present in the dataset that they explain. When removing dimensions from the dataset, more information is often preserved by keeping the dimensions with the most variance. With fewer dimensions the dataset is easier to visualize and make use of for decision making.

The first principal component of a dataset is a unit vector in the direction which contains the most variance. Subsequent principal components are also unit vectors in the direction which contains the most variance but with the added restriction that they are orthogonal to all earlier principal components. When reducing the dataset to fewer dimensions, earlier principal components are kept to retain as much variance as possible.

In this study each chromatogram is reduced to two dimensions to be visualized as a point on a two dimensional graph. Before being reduced it consists of measurements of three different variables at hundreds of points in time. There are many ways to approximate the principal components of a dataset. In this study it is done with singular value decomposition of a matrix implemented in the linear algebra library MATH.NET. [7]

3.2.2 Simple and one-point multiparameter technique

Simple and one-point multiparameter technique is a technique for visual detection of emerging problems in chromatography columns which has been used to monitor commercial scale production at the pharmaceutical company Bayer, presented in an article in BioProcess International. [8] The technique aims to replace more time and labor intensive techniques such as laboratory assays and expert visual review of whole chromatograms. Instead, only values from a single point in time are visually compared between chromatograms, which allows anyone involved in the production process to quickly be taught to do it.

The first step of the simple and one-point multiparameter technique is to find the highest measured absorbance in the elution phase of a chromatogram. The conductivity and pressure values at that point in time are also noted. The same is done to chromatograms from runs performed consecutively on the same chromatography unit. When all three noted values from a chromatogram differ more than usual from the preceding chromatograms, it is a warning to production personnel that there are issues with the column that need to be addressed. Case studies referenced in article Development and Application of a Simple and One-Point Multiparameter Technique show that rapid changes in one or two of the values may not be indicative of column degradation but could be caused by a number of things like change in sample concentration. [8]

The simple and one-point multiparameter technique has low requirements for frequency of measurements, computing power, and adjustment to small changes in the process compared to models based on more complex multivariate analysis like principal component analysis or partial least square regression. On the other hand it relies on enough infor-

mation to identify an emerging problem to be available in a single measurement point.

3.2.3 Evaluation of the statistical techniques

The usefulness of the techniques as visual aids for detecting problems in a chromatography unit is assessed. Making it easy to distinguish between early and late runs in a series is a sign that the technique can detect gradual decline in the unit. This assumes that the unit functions well in the beginning of the series, that a change in the chromatogram indicates a change in the unit, and that any change in the unit decreases performance or durability. The choice of the group of early runs being the first 20 is based on Feidl et al. result of identifying three groups of chromatograms in their series, the first of which was the first 20 cycles. [4] The techniques should indicate changes in the chromatograms before acute failure of the chromatography unit to perform another run occurs.

3.3 Software

3.3.1 Control and monitoring software

The control and monitoring software, Unicorn, allows the user to control the chromatography system and access the measurements it has collected. Settings for different phases in the run are programmed by the user. This includes flow rate of liquid through the chromatography unit, volume of sample, and volume of buffer to be used, among others. The software runs on .NET and imports third party libraries through NuGet. There are tools available in the software to draw plots of chromatograms but not general purpose graphing tools. The software includes a graphical user interface where the user can select saved chromatograms to perform calculations on. As a developer it is possible to take inputs from the user through pop-ups with drop-down menus, checkboxes, and free text boxes. As a developer it is also possible to print results from calculations to comma-separated values files on the users computer. The Unicorn version used is in development and is not commercially available.

4 Method

4.1 Chromatogram series

Two series of chromatograms from two different chromatography units were investigated. Both came from chromatography runs using HiTrap Fibro chromatography units connected to an Äkta pure 25 system. Figure 4 shows a Fibro chromatography unit being attached to an Äkta pure system.

Figure 4: A HiTrap Fibro unit is being attached to an Äkta pure system. Property of Cytiva

The two units were prepared differently and therefore not expected to behave the same way. One series contained 200 chromatograms from runs performed consecutively on the same system, with the same settings, the same unit attached, and the same type of sample being passed through. The other series contained 103 chromatograms, also, from runs performed consecutively on the same system, with the same settings, the same unit attached and the same type of sample being passed through. During run 103, in the series of 103 chromatograms, the column failed and the run could not be completed. The corresponding chromatogram was excluded from the study since it did not contain an elution phase and was therefore not comparable with the techniques used. Every completed run contained an equilibration phase, a sample application phase, 2 Fibro Unit Wash phases, an elution phase, and a re-equilibriation phase. Table 1, in 3.1.2 Fibro chromatography, shows the phases in order. No cleaning in place phase was performed between runs. The protein being purified was an anti-body.

4.2 Principal component analysis

4.2.1 Pre-processing

The measurements of absorbance of 280nm light, pressure differential over the chromatography unit, and measures of conductivity performed during the elution phase were used. The automatic labeling of start and end of phases by the software was used to decide which measurements were made during the elution phase. The smallest measured value

for each of the variables was noted. The chromatogram with the least volume passing through the unit during the elution phase was identified. Measurements were removed from the end of all other elution phases to make all elution phases the same length.

To line up the measurements volume-wise, spline interpolations were made and resampled from. The linearly interpolated splines were created with the function available in the linear algebra library Math.NET [9]. The resampling was performed every 0.1ml starting from the beginning of the elution phase. The corresponding lowest measured value was subtracted from all resampled absorbance and conductivity values. Since the baseline for the pressure curve was neither the smallest, nor the largest value, no baseline subtraction was performed.

The resampled values from the chromatogram elution phases are placed in a matrix where every row contains the resampled measurements from one chromatography run. The absorbance values are denoted by a, the conductivity values by c, and the pressure values by p,

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,n} & c_{1,1} & \dots & c_{1,n} & p_{1,1} & \dots & p_{1,n} \\ . & & . & . & & . & . & & . \\ . & & . & . & & . & . & & . \\ . & & . & . & & . & . & & . \\ a_{m,1} & \dots & a_{m,n} & c_{m,1} & \dots & c_{m,n} & p_{m,1} & \dots & p_{m,n} \end{pmatrix}$$

,while m is the number of chromatograms in the series and n the number of resampled measurement points.

Every row begins with absorbance values, followed by the conductivity values, and ends with the pressure values. Each column contains the resampled measurements of one variable at a certain time after the start of the elution phase in every run. To remove the impact of units, the first 20 columns were standardized, ensuring that they have a mean of zero and standard deviation of one. The standardized value z is calculated from the original value x by

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

, where μ is the mean of the first 20 values in that column and σ is their variance. The same formula was applied to the rest of the values in each column, except that the μ and σ used were the ones calculated for the first 20 for that column.

An additional matrix was created where the pressure values were left out, to investigate the impact of pressure on the result.

4.2.2 Calculation

Principle component analysis is performed to reduce every row of the matrix, representing the elution phase of one chromatography run, to a single point on a two dimensional graph for visual inspection. The linear algebra library Math.NETs function for Singular Value Decomposition was used to calculate the first two principal components of the first 20 runs of the series. The first 20 runs are assumed to be functioning normally, based on the study mentioned in 4.2.3 Evaluation of the statistical techniques. The percentage of variance accounted for by each principal component was calculated [7]. All rows in the

matrix were then projected onto the first two principal components of the first 20 runs and plotted.

4.3 Simple and one-point multiparameter technique

Performing the simple and one-point multiparameter technique requires the identification of the absorbance elution peak and the conductivity and pressure values at the same point in time. The highest measured value of absorbance of 280nm light in the elution phase was identified. The measurements of pressure and conductivity performed closest to the peak were also identified. Absorbance values obtained this way from each chromatogram were plotted against the number of that run in the sequence of the series. The values for conductivity and pressure were plotted in the same way. The baseline values described in 4.2.1 Pre-processing for absorbance and conductivity were subtracted from values corresponding to that run.

5 Results

5.1 Principal component analysis

The graph resulting from the principal component analysis of the series of 102 chromatograms can be seen in Figure 5. The first and second principal component, which the series's chromatograms are projected onto, are calculated from the first 20 chromatograms in the series and marked with green. The last 20 chromatograms in the series are marked with red. The run number of the last two chromatograms has been noted next to their corresponding dots. This has been done for all graphs of principal components in the Results section. Notable in the figure is the four outliers and that there is just a little overlap between early and late runs.

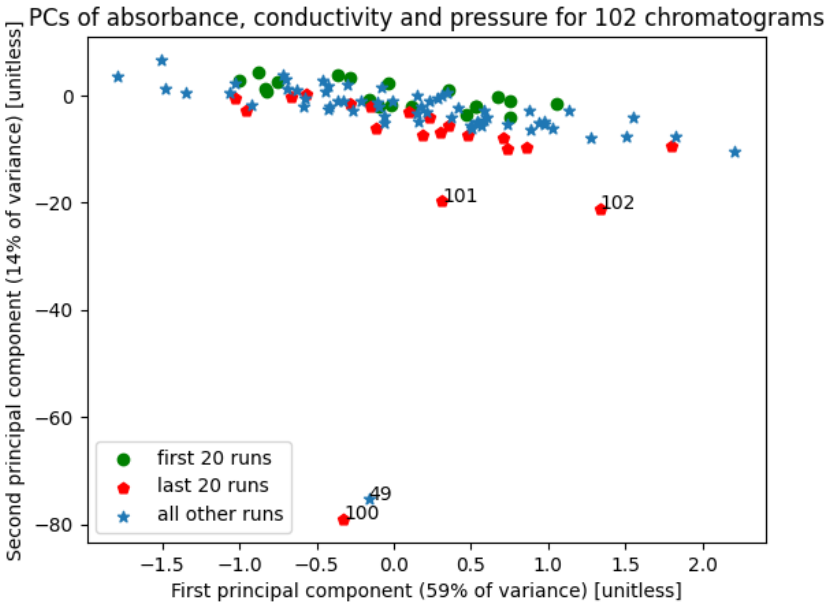

Figure 5: 102 consecutive chromatograms mapped onto principal components 1 and 2 of the first 20 chromatograms. The chromatograms are represented by sampled measurements of absorbance, conductivity and pressure from the elution phase.

To investigate the role of pressure, the principal component analysis was performed with all pressure values excluded. The resulting graph of the series of 102 chromatograms can be seen in Figure 6. The groups of the first 20 and last 20 runs are very clearly separated. None of the outlier runs in Figure 5 are outliers in Figure 6.

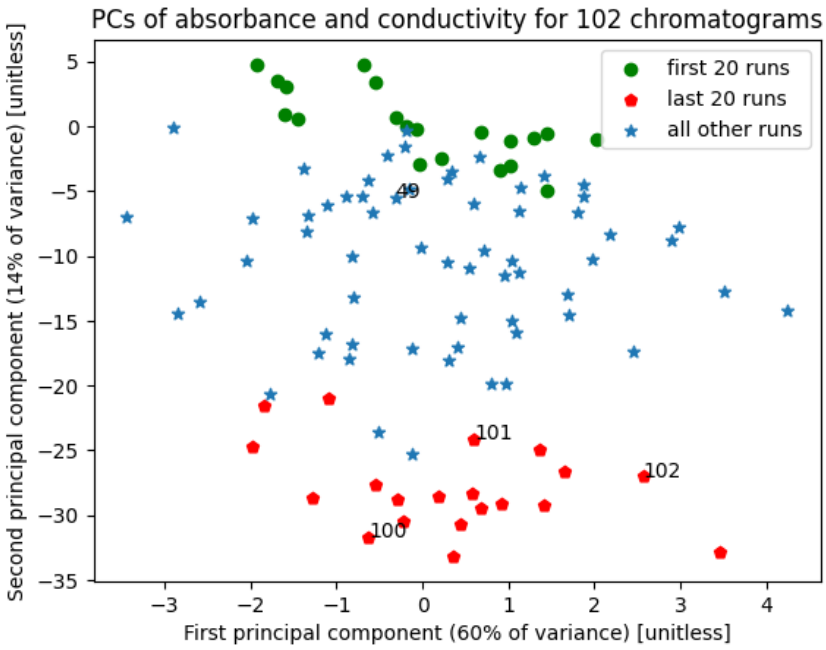

Figure 6: 102 consecutive chromatograms mapped onto principal components 1 and 2 of the first 20 chromatograms. The chromatograms are represented by sampled measurements of absorbance and conductivity from the elution phase. Measurements of pressure have been excluded.

Performing the principal component analysis on the series of 200 chromatograms yields Figure 7. Most notable in this figure is the separation between two groups of runs which does not seem to be related to the chromatography unit age.

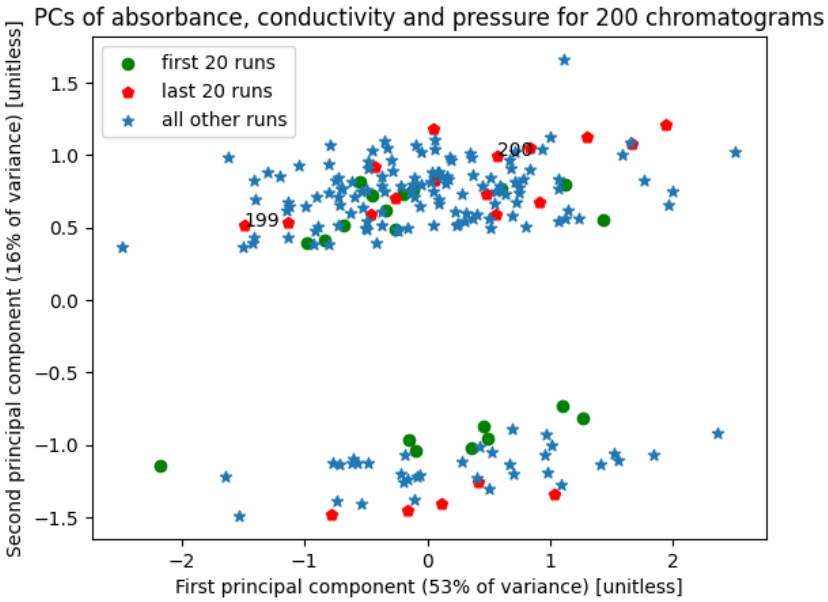

Figure 7: 200 consecutive chromatograms mapped onto principal components 1 and 2 of the first 20 chromatograms. The chromatograms are represented by sampled measurements of absorbance, conductivity and pressure from the elution phase.

Principal component analysis was also performed with pressure values excluded for the series of 200 chromatograms. This yielded Figure 8 where the groups of the first 20 and last 20 runs do not overlap, contrary to the results with pressure included, while the last two runs are still not separated from the rest.

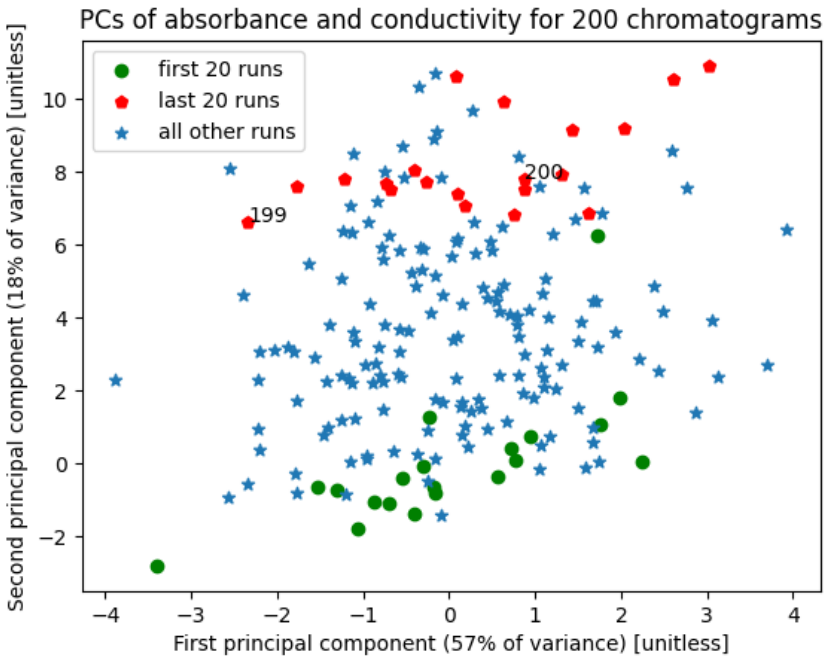

Figure 8: 200 consecutive chromatograms mapped onto principal components 1 and 2 of the first 20 chromatograms. The chromatograms are represented by sampled measurements of absorbance and conductivity from the elution phase. Measurements of pressure have been excluded.

5.2 Simple and One-Point Multiparameter Technique

All absorbance peaks in the series of 102 chromatograms are plotted in order from left to right in Figure 9. Visually identified outliers have been marked with their place in the series.

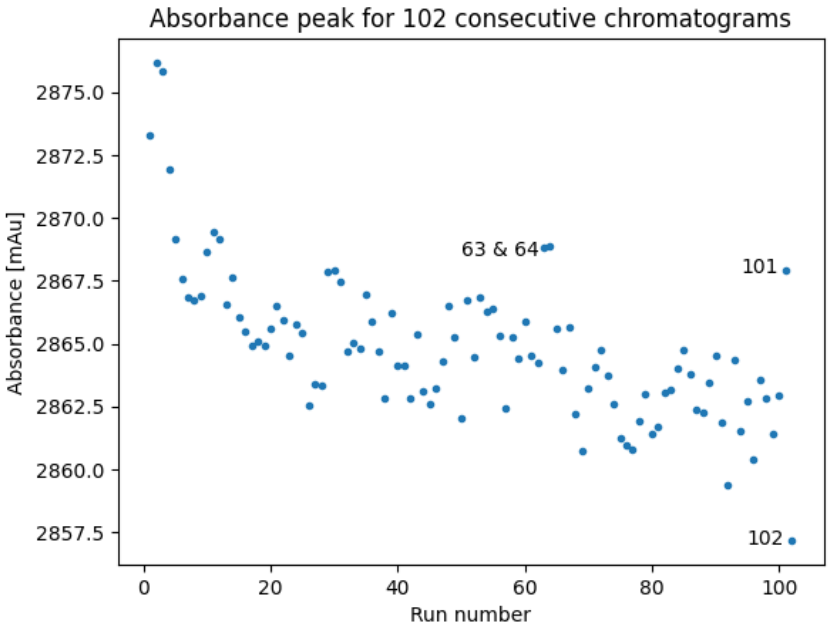

Figure 9: Peak absorbance as measured by UV 280mm light in the elution phase of 102 consecutive Fibro chromatograms.

The measurement of conductivity performed closest to the absorbance peak in every chromatogram from the series of 102 is shown in Figure 10. The measurements of conductivity from run number 100 and 101 were visually identified as outliers. Run number 65 was marked to clarify that it is not 63 or 64 which were marked as outliers among the absorbance peak values.

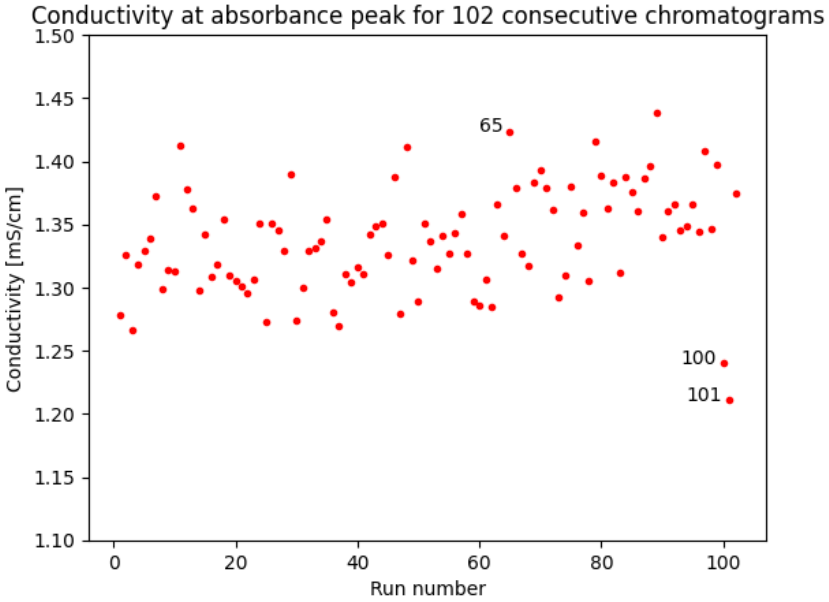

Figure 10: Conductivity measured at peak absorbance for 102 consecutive Fibro chromatograms.

The measurement of pressure performed closest to the absorbance peak in every chromatogram from the series of 102 is shown in Figure 11. Three visually identified outliers were marked with the number of the run they were measured in.

The measurement values closest to the absorbance peak of absorbance, conductivity, and pressure from the series of 200 chromatograms are shown in Figures 12, 13, and 14 respectively. No clear outliers were identified in either figure.

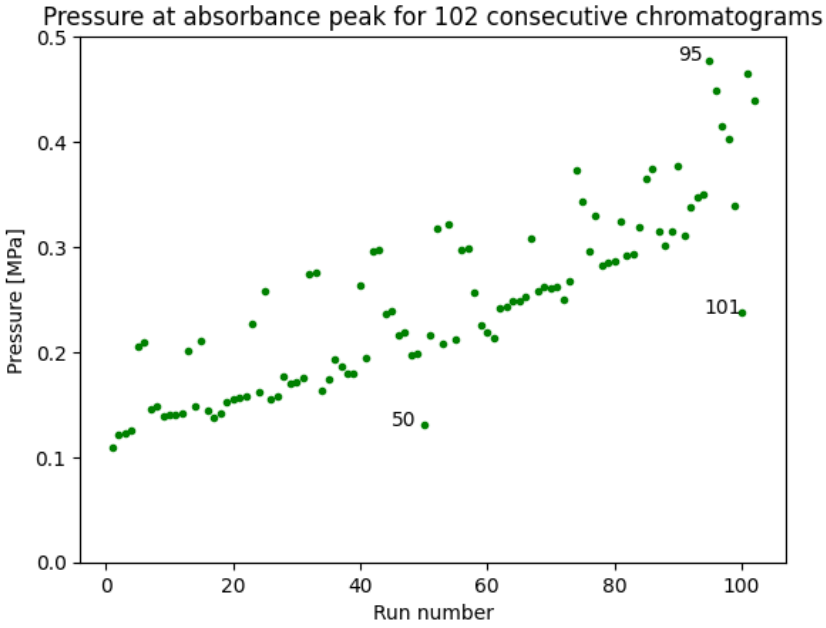

Figure 11: Pressure difference over the unit at peak absorbance for 102 consecutive Fibro chromatograms.

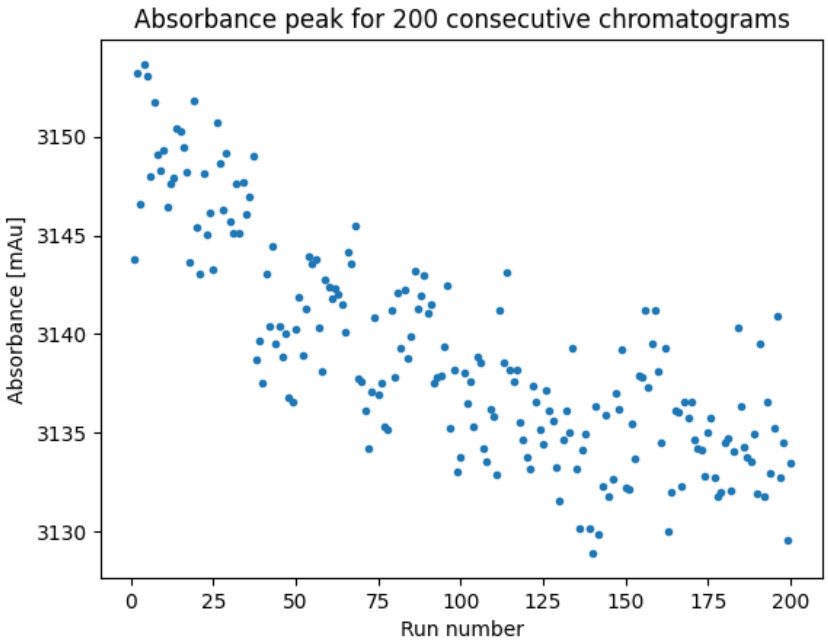

Figure 12: Peak absorbance as measured by UV 280mm light for 200 consecutive Fibro chromatograms.

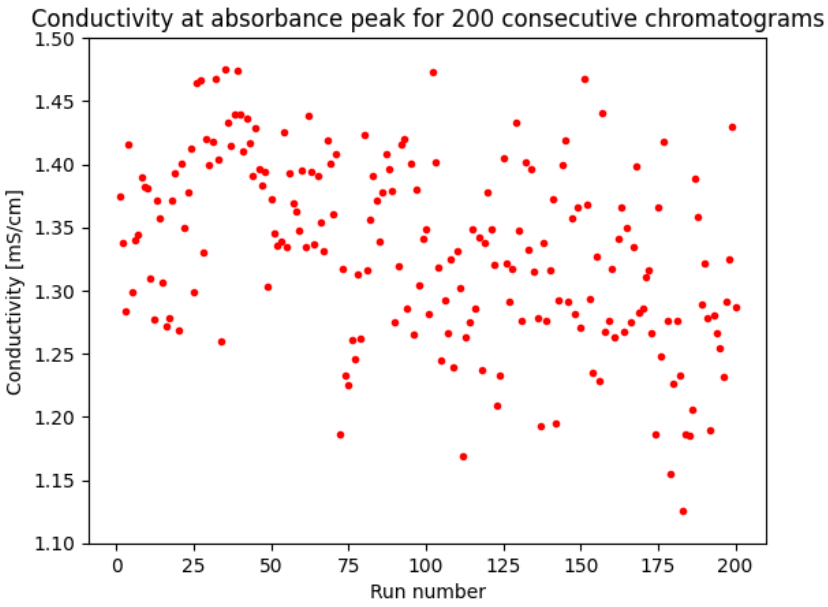

Figure 13: Conductivity measured at peak absorbance for 200 consecutive Fibro chromatograms.

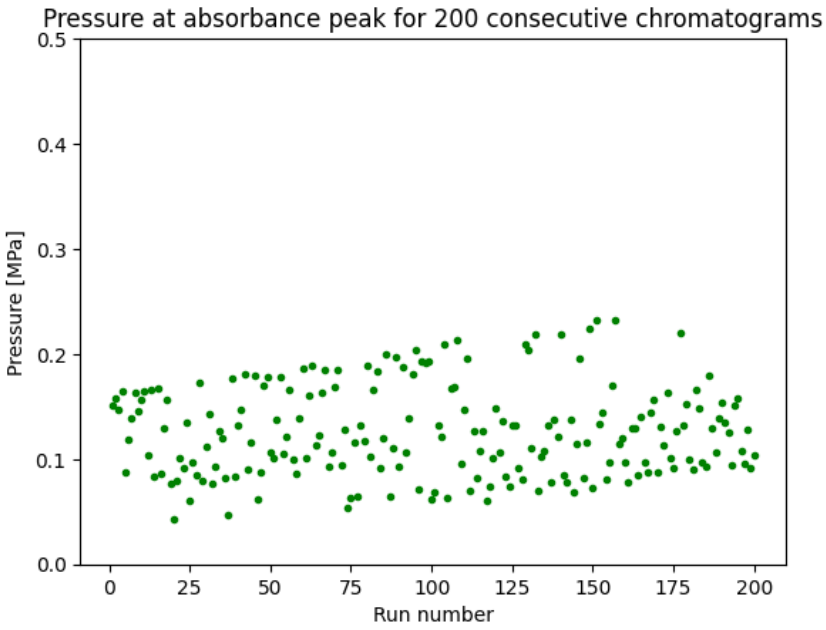

Figure 14: Pressure difference over the unit at peak absorbance for 200 consecutive Fibro chromatograms.

Note that the absorbance axis in Figure 12 shows a different range than the absorbance axis in Figure 9. Note also that the run number axis extends to 102 for graphs showing measurements from the series of 102 chromatograms and 200 for those showing measurements from the series of 200 chromatograms. The pressure values measured closest to the absorbance peak in both the series are shown in the same graph in Figure 15.

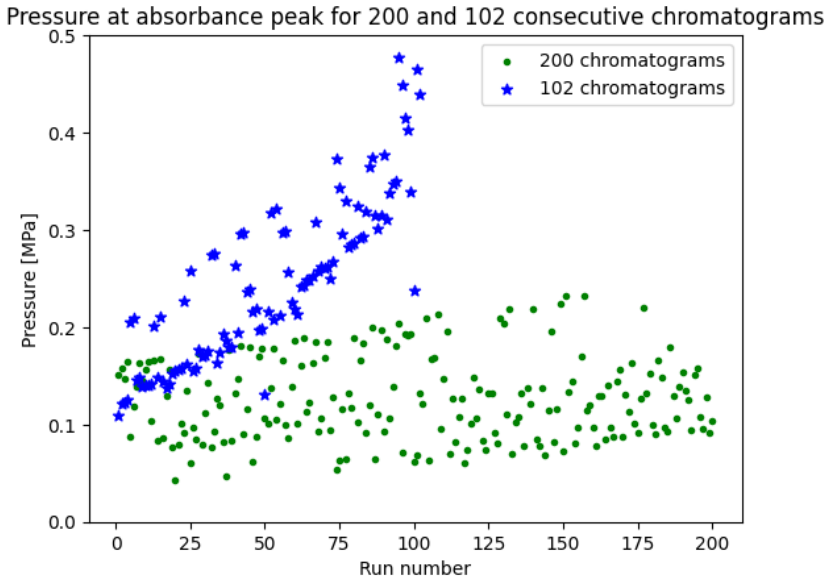

Figure 15

6 Discussion and conclusions

Principal component analysis and simple and one-point multiparameter technique both showed promise for detecting acute and gradual degradation of a Fibro chromatography unit. Given that both techniques have been shown to work on resin chromatography units, further tests on Fibro units are motivated. These tests should investigate performance on detecting several causes for failure and use more direct measurements of chromatography unit performance. The implementation of the techniques in the control and monitoring software enabled users to select chromatograms and statistical techniques in the same graphical user interface that they monitor and program chromatography runs in. No technique took more than 15 seconds to run on a laptop detailed in Appendix I. All graphs in the Results section were drawn in a separate program since the current control and monitoring software only has graphing tools specific to chromatograms.

6.1 Warning of imminent acute failure

For the series of chromatograms where the chromatography unit failed, both techniques gave an indication of arising problems before the complete failure to perform a chromatography run. Failure to perform another run is the most extreme form of performance decline. Preferably, operators of chromatography systems should be warned before less extreme performance declines too. Since no other performance indicator was measured with this series of chromatography runs, we do not know how long before the failure that the chromatography unit's performance decreased. More studies of the techniques on series where performance has been measured, through for example sample yield or runs, or dynamic binding capacity of the unit are required to find out what magnitudes of performance increase these techniques can predict in Fibro chromatography units.

6.1.1 Principal component analysis and acute failure

Principal component analysis separated the runs closest to failure from the rest and this was mostly due to information from the pressure measurements in the chromatogram. The results of principal component analysis shown in Figure 5 in the Results chapter place the last three runs before failure clearly outside of the larger group of chromatograms. Using proximity to other chromatograms as a mark of decreased performance or increased risk of failure assumes that the chromatography unit was functioning properly from the beginning and that changes in a chromatogram are indicative of changes in the chromatography unit which make it more likely to fail or perform worse.

The reliability can however be questioned since run 49 was a false positive. It was as far or farther from the main group as the last three runs before failure but unit continued functioning for another 53 runs after it.

To investigate the role of pressure measurements in the principal component analysis results, the same analysis was performed on the absorbance and conductivity values with pressure values excluded. When pressure was excluded, the results of which can be seen in figure 6, the last two runs before failure are not separated from the other late runs of the series. This shows that the pressure measurements of the chromatograms contain

the information necessary for principal component analysis to separate the runs closest to failure from the rest.

The results of principal component analysis on the series of chromatograms from the unit that did not fail is shown in figure 7. This figure does not show outliers clearly distanced from a tight group in the same way. The last two chromatograms are instead near the middle of the group. Since there were no measured performance problems it is difficult to know if the separation between groups of chromatograms is related. If the performance was good in all runs it is important to experimentally set boundaries that allow for the variations seen here.

In the 2020 Feidl et al. article on PCA as tool for column chromatography supervision, an area is drawn where chromatograms with verified good performance, at least 85% sample yield, are known to have placed [4]. A similiar approach could be taken in future studies on Fibro chromatography to determine how separated a run can be from the rest before it indicates poor performance or a higher risk of failure.

6.1.2 Simple and one-point multiparameter technique and acute failure

The simple and one point multiparameter technique issued a clear warning at run 101 in the series of chromatograms from the unit that failed. In figure 9 the 101st run has a higher absorbance peak than the preceding 40 runs which breaks the trend of decreasing peak. The 101st run also breaks the trend of increasing conductivity and pressure values, registering the lowest conductivity of all runs in the series in figure 10 and a sudden decrease in pressure from the preceding run in 11. That the technique warned us two runs before complete failure is promising but the conductivity and pressure values were back in line with the trend in the 102nd run. The fact that the run after the warning reverts back to the old trend does not match the case in the article presenting the simple and one-point multiparameter technique. In the examples presented there the values changed dramatically and persisted at their new levels [8]. The potential outlier runs 50, 63, 64, 95 and 100 were marked in figures 9, 10, and 11 but they did not match between the three figures and were therefore not false positives generated by the technique. As discussed regarding the principal component analysis, the more useful event to detect than failure to perform a run is a sudden decrease in performance. Studies of more series with accompanying performance measures are required to determine how well the simple and one point multi-parameter technique detects smaller acute changes in a Fibro chromatography unit.

6.2 Detecting gradual decline

Many problems with chromatography columns build up gradually. For example ligands being damaged, falling off, or build up occuring in pores of the resin as mentioned in section 3.1.4 Chromatography unit degradation. Since there are many similarities between Fibro chromatography and column chromatography some of these gradually worsening problems are expected to be the same. It is assumed that the Fibro units work well in the first 20 runs of the series and that any change in the chromatograms with time is a sign of changes in the unit which affect the function negatively or the remaining lifetime.

A technique being able to differentiate between early and late runs in the series is then a sign of it detecting gradual decline in the unit.

6.2.1 Principal component analysis and detecting gradual decline

Principle component analysis was able to separate the first and last 20 runs of the series into distinct groups when using only absorbance and conductivity measurements. Pressure measurements, on the other hand, caused the groups to overlap more. In Figure 7 pressure measurements also separated chromatograms very clearly unrelated to age.

Changes in the unit over time are expected to be greater in the series from the unit which failed after 102 runs. In this series principal component analysis separated the first and last 20 chromatograms with few overlaps in figure 5. The same thing is shown to be true in figure 6 where the same groups are much more visible when pressure is excluded. In the series from the unit which successfully performed 200 runs there is significant overlap between groups of early and late runs in figure 7. However, when the pressure values are excluded from the principal component analysis, as shown in figure 8, the first and last 20 chromatograms in the series are separated. This suggests that principal component analysis is able to detect changes happening over time but that noise in the pressure values interferes.

6.2.2 Simple and one-point multiparameter technique and gradual decline

Simple and one-point multiparameter technique performed on the series from the chromatography unit that failed shows a trend of decreasing absorbance peak, figure 9, and increasing pressure, figure 11, at the absorbance peak with increasing run number in the series. The same is true in the respective figures 12 and 14 from the series of 200 chromatograms. For this series where the unit did not fail the trend in the pressure values is much less clear, which is shown in the comparative figure 15, and variations caused by something else are much larger. Neither series showed any clear trend in their respective figures 10 and 13 plotting conductivity against run number in the series.

It should be noted that an increasing pressure value at the absorbance peak could be due to changes in the pressure of the unit or due to the absorbance peak occurring earlier. Since the pressure over the unit is decreasing when the absorbance peak occurs, a shift of the absorbance peak to earlier in the elution phase would cause a higher pressure value to be used in the simple and one-point multi parameter technique. This can be seen in figure 3 from the Chromatogram section of Background and Theory where the behavior of the three variables during the elution phase is shown. The only conclusion that can be drawn from the results mentioned in the last paragraph is that absorbance peak height was decreased with use of the units in both series, while conductivity at the absorbance peak was unaffected. Future research using the simple and one-point multiparameter technique should look at the degree of peak shift to determine if the pressure value at the absorbance peak contributes information about the state of gradual decline in the unit.

6.3 Practical application

A study of the techniques' performance on a chromatogram series with accompanying performance measurements should be performed before they can be used in a production setting with Fibro chromatography units. It can then be validated that the simple and one-point multiparameter technique warns when performance problems arise and not only right before complete failure to perform a run. A similar study can be used to determine how much a chromatography run can separate from the rest in principal component analysis before the performance has fallen beneath acceptable levels as was done in Model based strategies towards protein A resin lifetime optimization and supervision [4]. A stricter definition of what is an outlier in the simple and one-point multiparameter technique could improve consistency between different users and perhaps even allow automation of the graph inspection. Both techniques should be verified on different types of failure as this study has only investigated one.

A challenge when defining the rules for determining an outlier in either technique is that these may not be generalizeable between different ligands, samples, and unit volumes. Visually determining outliers works better for the simple and one-point multiparameter technique than for principal component analysis since the runs can be outliers in any direction on the two dimensional graph in principal component analysis.

To make the techniques easy and quick to use, a way to draw diagrams directly in the control and monitoring software is needed. The implementation in this study exported the results of the techniques from the program and drew diagrams with a separate Python script.

In conclusion, this study has shown that both techniques' have potential to be useful in Fibro chromatography. Both showed signs of being able to warn of imminent failure as well as gradual changes in the unit. It remains to be seen how well this translates to decline in performance of a unit.

6.4 Limitations

Standardizing the matrix column by column increases the effect on the results of a set point in time with small variance. An alternative could be calculating a common variance to divide by from all measurements of absorbance in the first 20 rows. This would favor a set point in time where the absolute variance is greater compared to the current choice which favors a set point in time where the relative variance is greater.

In this study, the pressure measurements were not baseline adjusted. This could lead to systematic errors if the baselines change between runs. However, the same baseline correction which was applied to the absorbance and conductivity values could not be applied to pressure. The pressure curve includes a sharp trough and if the depth of this trough were to change, then all other values would be affected by it due to the baseline shift.

6.5 Acknowledgements

The series of chromatograms was provided by Cytiva. Cytiva also supported the study financially, with expertise on chromatography, and by providing guidance on how to add features in the Unicorn software.

References

[1] A.S. Rathore, M. Pathak, G. Ma, D.G. Bracewell, "Re-use of Protein A Resin: fouling and economics", *BioPharm International 28 (2015)*

[2] L. Troeng, M. Townsend, O. Hardick, J. Zou, T. Frigard, "High-throughput mAb purification with Fibro chromatography", *Poster on Cytivas website* **https://cdn.cytivalifesciences.com/api/public/content/digi-33647-pdf (accessed December 10th, 2022)**.

[3] Cytiva, *cytivalifesciences web shop* **https://www.cytivalifesciences.com/en/us/shop/chromatography/fiber-chromatography/affinity-antibody/hitrap-hiscreen-fibro-prisma-units-for-research-and-process-development-p-23550 (accessed December 10th, 2022)**

[4] F. Feidl, M. F. Luna, M. Podobnik, S. Vogg, J. Angelo, K. Potter, E. Wiggin, X. Xu, S. Ghose, Z. Jian Li, M. Morbidelli, A. Butté, "Model based strategies towards protein A resin lifetime optimization" *Journal of Chromatography A 1625 (2020)*

[5] M. C.Nweke, A. S. Rathore, D. G. Bracewell, "Lifetime and Aging of Chromatography Resins during Biopharmaceutical Manufacture" *Trends in Biotechnology 36 (2018)*

[6] M. Pathak, A. S. Rathore, "Mechanistic understanding of fouling of protein A chromatography resin" *Journal of Chromatography A 1459 (2016)*

[7] Math.NET, *Math.NET documentation* **https://numerics.mathdotnet.com/api/MathNet.Numerics.LinearAlgebra.Factorization/Svd%601.htm (accessed December 10th 2022)**

[8] L. P. Pathange, L. Huang, J. Tsang, "Development and Application of a Simple and One-Point Multiparameter Technique (SOP-MPT)" *BioProcess International 16 (2018)*

[9] Math.NET, *Math.NET documentation* **https://numerics.mathdotnet.com/api/MathNet.Numerics.Interpolation/LinearSpline.html (accessed December 10th 2022)**

7 Appendix I - baseline shift of the pressure curve

All analysis was performed in an unreleased version of Unicorn running in debug mode of Visual Studio Professional. The computer had an 11th Gen Intel Core i7-11850H @ 2.50GHz, 64 GB installed RAM, and Windows 10 Enterprise.