UPTEC X 07 027
APR 2007
ISSN 1401-2138

Daniel Zakrisson

# Process economy calculations of downstream processes

Master's degree project

UPPSALA UNIVERSITET

# Molecular Biotechnology Programme

Uppsala University School of Engineering

| **UPTEC X 07 027** | **Date of issue 2007-04** |
|---|---|


Author

**Daniel Zakrisson**


Title (English)

**Process economy calculations of downstream processes**

Title (Swedish)


Abstract

A general interface in Microsoft Excel capable of analyzing different biotech processes with little adaptation for each process has been developed. The interface is able to perform sensitivity and uncertainty studies using Monte Carlo simulations and performs automated parametric variations.
Studies performed using the interface show how the production cost shifts to the downstream section of the process as product titers go up and that there could be an optimum number of cycles for unit procedures, in this case a protein A capture step.



Keywords

Process economy, downstream processing, Monte Carlo simulation, sensitivity analysis, uncertainty analysis



Supervisors

**Karol Lacki**
**GE Healthcare, Uppsala**

Scientific reviewer

**Patrik Forssén**


| Project name | Sponsors |
|---|---|
| Language **English** | Security |
| **ISSN 1401-2138** | Classification |
| Supplementary bibliographical information | Pages **33** |


**Biology Education Centre** Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 555217

# Process economy calculations of downstream processes


## Daniel Zakrisson



**Sammanfattning**

Läkemedelsindustrin är just nu inne i en stor förändringsprocess där allt fler av de nya läkemedel som godkänns är baserade på proteiner. Fram tills nu har dessa läkemedel inte haft någon prispress och de produceras till väldigt goda marginaler. I takt med att läkemedelskostnaderna ökar ifrågasätts varför vissa av dessa proteinläkemedel är så dyra och med det ökar kraven att produktionen ska ske på ett processekonomiskt optimalt sätt.

I det här arbetet har en datorapplikation som underlättar processekonomiska analyser tagits fram. Denna applikation kopplar samman förmågan i Microsoft Excel att spara, analysera och presentera data med modellbygge och ekonomiska beräkningar av biotekniska processer i SuperPro Designer från Intelligen och möjligheten att via ett program för Monte Carlo simuleringar, Crystal Ball från Decisioneering, utföra bland annat osäkerhetsanalyser och känslighetsanalyser.

Denna applikation har sedan använts för att genomföra analyser på en bioprocess för framställning av monoklonala antikroppar. Exempel på analyser som gjorts är känslighetsanalys och osäkerhetsanalys. Vidare så har det studerats om det finns ett optimalt antal cykler vissa steg i reningsprocessen, hur stor inverkan kostnaden för arbetskraft har och hur fördelningen av olika kostnader i produktionen ser ut.



**Examensarbete 20 p i Molekylär bioteknikprogrammet**
**Uppsala Universitet April 2007**

1. Introduction ..... 3
1.1 Literature overview ..... 3
1.2 GE Healthcare ..... 4
1.3 Aim ..... 4
2. Theory and background ..... 4
2.1 SuperPro Designer ..... 4
2.2 Crystal Ball ..... 4
2.3 Excel and add-ins ..... 5
2.4 Uncertainty analysis ..... 5
2.5 Sensitivity analysis ..... 5
2.6 Parametric variation ..... 6
3. Methods ..... 6
3.1 Description of the interface ..... 6
3.1.1 Setup section ..... 7
3.1.2 Input section ..... 7
3.1.3 Reporting/results section ..... 7
3.1.4 Dataflow in the application ..... 7
3.2 Bioprocesses ..... 8
4. Economic evaluation of a Mab bioprocess ..... 10
4.1 Evaluation of a Mab bioprocess project ..... 10
5. Drug development ..... 12
5.1 Expected clinical cost ..... 12
5.1.1 Clinical success rates ..... 12
5.1.2 Future success rates ..... 12
5.1.3 Future drug development cost ..... 13
6. Cost distribution ..... 14
6.1 Cost distribution for changing product titer concentration ..... 14
7. Number of chromatography cycles per batch ..... 15
8. Breakdown of costs ..... 18
8.1 Capital investment and operating costs ..... 18
8.2 Sensitivity analysis ..... 20
8.3 Parametric studies ..... 22
8.3.1 Capto Q resin binding capacity ..... 23
8.3.2 Mabselect SURE resin binding capacity and replacement frequency ..... 23
8.3.4 Labor cost ..... 26
8.4 Uncertainty analysis ..... 26
9. Discussions ..... 28
9.1 Improvements ..... 28
9.2 Further development ..... 29
10. Conclusions ..... 30
11 Acknowledgements ..... 30

# 1. Introduction

A steadily increasing demand for more affordable yet highly effective biopharmaceuticals puts a lot of pressure on big bio-pharmaceutical companies to introduce more cost effective production processes. Up to date these production processes, including both upstream (cell culture/fermentation) and downstream (purification) parts are more than often operated in a suboptimal mode, especially from a process economy perspective. According to many experts this situation must change in the nearest future in order for bio-pharma to remain a profitable industry.

The first part of this study will be an overview of the situation for pharmaceutical companies today and specifically for antibody based drugs. Based on the present debate over the high and increasing prices for pharmaceuticals a process for manufacturing monoclonal antibody based drugs will be financially evaluated. This will give a starting point and a reference for the present situation of pharmaceutical companies when developing drugs.

In healthcare systems around the world the expenditures on pharmaceuticals have grown faster than other components since the late 1990s. It has been said that the healthcare systems cannot take much higher increases in drug cost and today not all healthcare systems approve of administering all available drugs (McClelland, 2004). So, in order for the industry to remain profitable it must lower the cost of developing and producing the new drugs.

The second part of the study will be an analysis of the production process itself disregarding the R&D costs. Studies of cost distribution, single parameters, sensitivity and uncertainty analysis will be performed with the same process used in the first part.

## *1.1 Literature overview*

Previous work in the area includes papers from Sven Sommerfeld and Jochen Strube (Sommerfeld & Strube, 2005) regarding the development of generic processes for production of bio-pharmaceuticals since the quality is determined by the process and production processes are often scaled up from the pilot process. Joseph DiMasi and his group at Tufts University (DiMasi, Hansen & Grabowski, 2003) has studied development costs and clinical success rates for modern pharmaceuticals. When it comes to the use of process simulators and model building it has been described in several papers, the use of models is described by Levine and Latham (Levine & Latham, n.d.), the use of process simulators in pharmaceutical process development is described by Demetri Petrides (Petrides, Koulouris & Lagonikos, 2002). Petrides has also described how to use process simulators for debottlenecking and throughput analysis (Petrides, Koulouris & Siletti, 2002).

### *1.2 GE Healthcare*

The GE Healthcare's Life Sciences division in Uppsala is delivering systems and equipment for biopharmaceutical purification, drug discovery, biopharmaceutical manufacturing and cellular technology.

### *1.3 Aim*

The aim of this project was to perform economical analysis of a monoclonal antibody production process. The analysis required development of a user interface between the software packages Intelligen SuperPro Designer, Decisioneering Crystal Ball and Microsoft Excel capable of performing sensitivity and uncertainty analysis and capable of automatic parametric variation of variables.

# 2. Theory and background

### *2.1 SuperPro Designer*

SuperPro Designer (www.intelligen.com) is a software tool for modelling, evaluation and optimization of integrated processes in a range of development and production industries. Batch processes are modelled as a flow sheet design with a number of unit procedures interconnected with streams. A unit procedure represents a step in the process, for example a chromatography column purification step, and the unit procedure is then subdivided into operations, for example load, elute and wash operations.

In SuperPro Designer capability of acting as an OLE (Object Linking and Embedding) automation server using COM (Component Object Module) technology is included. The base for this is the Designer Type Library which includes methods and functions that can be called from Windows applications, such as Excel, Word or Visual Basic using the programming language VBA (Visual Basic for Applications).

### *2.2 Crystal Ball*

There are several different commercial Monte Carlo solutions available and in this project Crystal Ball (www.decisioneering.com) from Decisioneering, which is an add-on for Microsoft Excel, is used.

Monte Carlo simulation, named after its use of randomness and analogy to the repetitive processes used at a casino, is a technique to analyze the behaviour of large and complex models where analytical solutions are hard to find. The technique is based on randomization of input variables and monitoring of output variables for the model.

In traditional spreadsheet based analysis the problem of uncertainty in the model is usually tackled the in one of three ways. 1) Use of point estimates where the most likely value (the mode) is used as input; 2) Use of range estimates where usually three cases are covered (best, normal and worse); 3) Use what-if scenarios were range estimates are used in different combinations.

With Monte Carlo simulation this weak approach is not needed. Uncertain values are instead given as distributions either based on the variability of the variable or based on the uncertainty of the variable. The results in the output variables are then given as statistical distributions instead of fixed values.

### *2.3 Excel and add-ins*

With later versions of Microsoft Excel the possibility to write your own functions and control the interface is fairly well developed. The programming language used is VBA, Visual Basic for Applications, which is based on Microsoft's Visual Basic with some additions to easy the development of applications that is based on and used on top of other products from Microsoft. In Excel for example, all functions available by clicking and writing on the spreadsheet are available as functions that can be called by code by user-defined functions.

For each project in Excel there is a choice to connect the spreadsheet to OLE servers. Since SuperPro Designer has a capability of acting as an OLE server a connection between the software packages can be made.

### *2.4 Uncertainty analysis*

In a Monte Carlo simulation the output from the Crystal Ball software package contains all results and the probability for each result. If these probabilities are normalized it is possible to calculate the certainty for each outcome. The certainty is then defined as the percent chance that a particular forecast value will fall within a specified range (Crystal Ball manual, 2001).

Uncertainty (or risk) analysis is the process of analyzing the forecast values and how they correspond to the project goals. Questions that can be asked and answered are for example: How big is the certainty that the project will generate a profit? What selling price does a compound require for a project to reach a certain profit? What should the target production capacity be to have a 90 % certainty to reach the required production?

In model analysis the goal is often to find the certainty of achieving a particular result. The most noticeable benefits of risk analysis are the assistance in decision-making by the possibility of quickly examining all possible scenarios and in the exposure of risk in the model.

### *2.5 Sensitivity analysis*

The sensitivity of a parameter on a particular forecast is the result two things, the sensitivity in the model for the forecast and the uncertainty in the assumption. For simple models and direct relations the sensitivity can be algebraically calculated by hand. With increasing model complexity the work needed can become too large or impossible to do by hand.

In Monte Carlo simulations the correlation between changes in assumption variables and forecast values can be monitored as the simulation progresses. If these

correlations are then ranked according to their covariance the sensitivity for each assumption value on the model can be given directly from the simulation.

Two different techniques to obtain the results are available, the first option is to simultaneously vary all parameters, track the result and calculate covariance. The second option is to vary one variable at a time while keeping the rest fixed for a more exact calculation of the specific variable's contribution to variation in a specific forecast, this generates a so called Tornado chart and is used for the studies in this paper.

The Tornado chart is used since it gives the best results when increasing the number of assumptions and it requires fewer trials to give as good or better results. It has been hard to analyze how to set up simulations with the standard sensitivity analysis to give correct results, often assumptions that are known to have a positive covariance are reported as negative due to that the number of assumptions is too large. The main reason for using the standard sensitivity method is that it is calculated at the same time as uncertainty calculations, and thus it does not require additional runs.

Sensitivity analysis has three main benefits. The first is the possibility to find out which assumptions influence the forecast the most; this reduces the time needed to refine estimates and helps in the selection of which variables to study further in parametric variation studies. The second is to find out which variables influence the least so they can be ignored or discarded. And third, with knowledge about how each of the assumptions influences the model better spreadsheets with higher performance can be built.

### *2.6 Parametric variation*

To get more data and to study the influence of certain parameters more than the information generated in the sensitivity analysis, parametric variation of the parameters can be done. In parametric variation one or two parameters (called decision variables) are chosen and studied over a specified range. A specific forecast is chosen and the effect of each decision value and the corresponding forecast value is saved. This represents a more advanced version of the scenario analysis covered under uncertainty analysis since uncertainty can be included in all other input parameters and the process of choosing the range of study is dynamic.

# 3. Methods

### *3.1 Description of the interface*

The interface is done as a VBA application in Microsoft Excel. It is written to be as general as possible to be able to handle as many processes as possible without need for modifications. In order to do that the processes will also need to be built so that as much as possible is automatically scaled and calculated. The interface itself is built around three parts, a setup section, an input section and a reporting/output section.

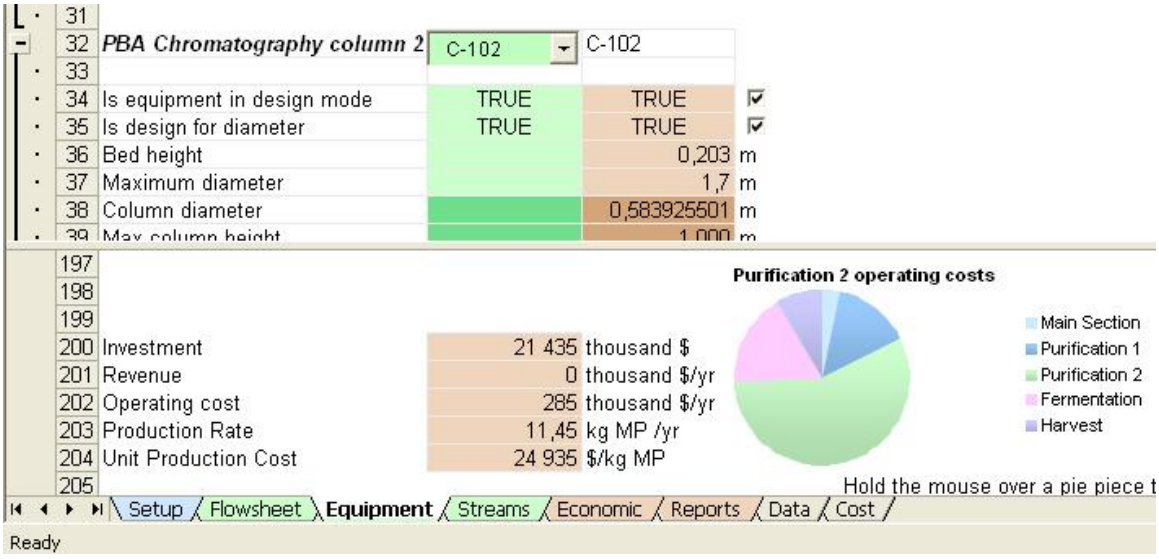


*Figure 3-1: Screenshot of the interface. The blue tab is the Setup section, green tabs are the Input section and brown tabs are the Output section.*

### *3.1.1 Setup section*

To facilitate the use as much as possible the first worksheet is a step by step instruction on how to set up the interface and Crystal Ball.

### *3.1.2 Input section*

The following three worksheets are where all settings of variables are done. The variables are either set to a new fixed value and the effect of the new setting can be monitored in the bottom part of the screen. If the aim is to perform Monte Carlo simulations assumption values are set by entering a value and then using Crystal Ball to set Cell > Assumption cell… and then defining the distribution of that variable.

### *3.1.3 Reporting/results section*

To have fast and easy access to the economic indices and results from the process and from the simulations a reporting feature is included. The report worksheet automatically collects much of the key information and formats it in a printable way.

### *3.1.4 Dataflow in the application*

To use the application, a bioprocess modelled in SuperPro Designer is required. When this process is done it is saved and then the Excel interface is loaded to work as a central hub between SuperPro, Crystal Ball and the capabilities of Excel itself.

To see how data flows between the different applications and the interface, a model of data flow is outlined in figure 3-2. Changes of variables are done in the Input section of the interface where a direct link to SuperPro sets the variable and then the interface reads the new value. When the interface performs a calculation of material and mass balances or an economic evaluation of the process the Output/Reporting section will read the new data from SuperPro.

When a Monte Carlo simulation is done, Crystal Ball will change a variable in the Input section. The Input section will then change and read the new value from SuperPro and automatically perform evaluations of the process. When the output variables change, Crystal Ball will monitor the changes.

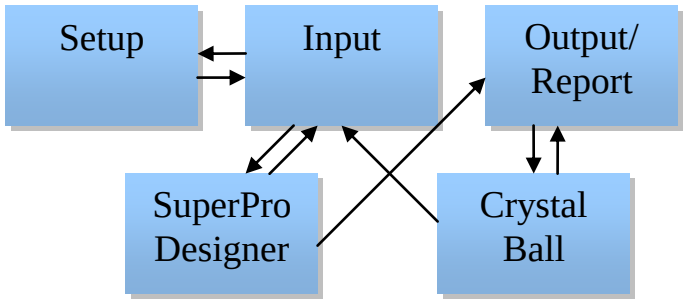


*Figure3-2: Information flow between the interface and the applications.*

*3.2 Bioprocesses*

For the purpose of presenting capability of the interface a typical monoclonal antibody purification process is used. The process (fig. 3-3) includes the following steps typically found in biopharmaceutical production processes which are:

- Cell removal
- Affinity chromatography
- Virus inactivation
- Cation exchange chromatography
- Anion exchange chromatography
- Virus clearance
- Sterile filtration

The process shown in figure 3-3 is studied using the interface by performing uncertainty and sensitivity analysis, and parametric studies of some variables. In each case considered, economic parameters are used as process descriptors.

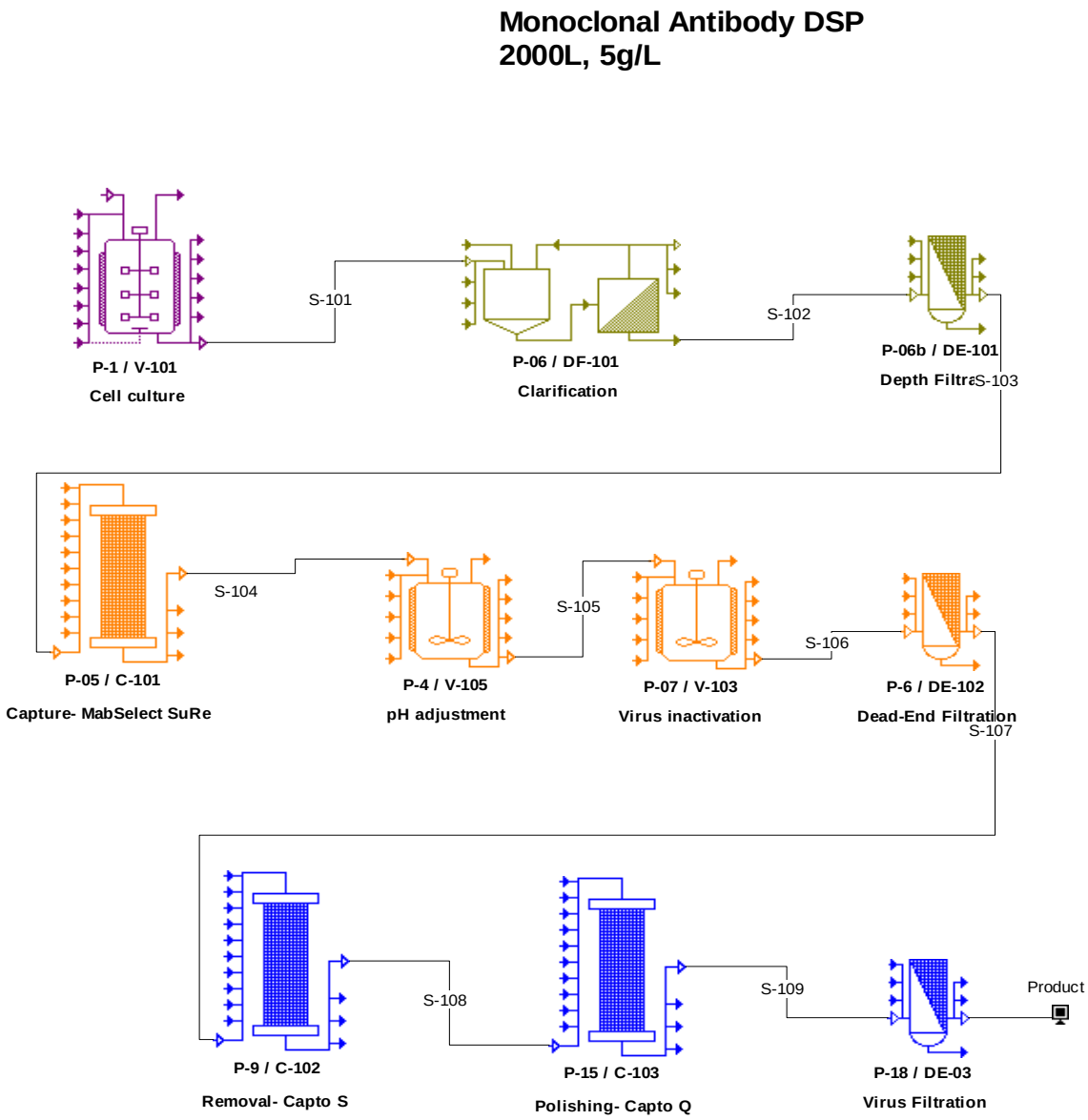


*Figure3-3: A simplified view of the process used for the analysis.*

In the analysis of a production process two capital budgeting terms are used, the internal rate of return and the net present value. They are both similar in nature but results in slightly different measures of the investment.

Internal rate of return (IRR) is a capital budgeting term used to evaluate long term investments. The IRR is the return rate which can be earned on the invested capital and should include an appropriate risk premium. The rate of return is defined as any discount rate that results in a net present value of zero of a series of cashflows. To find the rate of return the IRR that satisfies the following formula is used:

$$\text{Initial investment} = \sum_{t=1}^{N} \frac{C_t}{(+ IRR)^t}$$

Where 1…N is the time periods of interest and C is the net cash flow for each t (time period).

Net present value (NPV) is another standard method for financial evaluation of long term projects. By definition the NPV = Present value of cash inflows – Present value of cash outflows. In practice every cash inflow and outflow are discontinued back to its present value and then summed according to:

$$\text{NPV} = \sum_{t=0}^{n} \frac{C_t}{(+ r)^t}$$

Where n is the time periods, C is the net cash flow for each t and r is the discount rate used in the company for investments.

# 4. Economic evaluation of a Mab bioprocess

*4.1 Evaluation of a Mab bioprocess project*

Economic analysis of a monoclonal production process will be done by analyzing the following questions; 1) What would happen if the cost of developing new drugs would drop?; and 2) How the production cost is distributed and how it can be lowered.

Given the high prices of monoclonal antibody based drugs it is interesting to see how profitable a Mab production process could be today. As a measure of profitability the internal rate of return (IRR) of a typical process is calculated.

Protein pharmaceuticals can be divided into several categories. Today there are generally two types of Mab drugs. Medium dose Mab drugs such as Enbrel and high dose Mab drugs such as Herceptin and Rituxan. In this study Herceptin which is typical high dose Mab drug will be used as an example. Last year the production level of Herceptin was 251 kg and it was selling for around $6700 per gram (Sofer, Hagel & Jagschies, in press).

Production cost of a drug similar to Herceptin can be calculated using the SuperPro Designer software. A generic production process (see section 3.2) for monoclonal antibody production similar to the production process of Herceptin (Sommerfeld & Strube, 2005) was used. The process was setup using the parameters in Table 4-I.

The R&D cost is calculated using an industry average out-of-pocket development cost of $402 million per new drug (DiMasi, Hansen & Grabowski, 2003). This cost is based on a clinical success rate of 15% and includes the cost of failed projects that does not make it to market.

When setting the selling price of the product it is assumed that 50% of the final selling price to consumer of the drug will be markups to cover marketing, distribution,

advertising and royalty expenses. Consequently, for this analysis the selling price is set to $3 350/g.

The cost of capital in the pharmaceutical industry is on average 11% per year (DiMasi, Hansen & Grabowski, 2003). Thus all capital requirements (direct fixed capital, working capital and R&D expenses) is to be fully financed by loans with 11% interest rate and by setting the loan period to be as long as the project operating time, which is thirteen years.

With a scheduling requirement to run 45 batches per year, it is necessary to have two separate upstream trains since each batch require two weeks of cell culture time.

Table 4-I
*Setup of Mab production process (DSP).*

| | |
|---|---|
| Up front R&D | $402 million |
| Direct fixed capital | $21.2 million |
| Working capital | $5.3 million |
| Startup and validation cost | $1.1 million |
| Cost of capital | 11 % |
| Production of main product | 256 kg |
| Selling price | $3 350/g |
| Project lifetime | 15 years |
| Construction period | 30 months |
| Startup period | 4 months |
| Depreciation time | 5 years |
| Upstream volume | 10 000 L/batch |
| Main product titer | 1 g/L |
| Batches per year | 45 |

The resulting economic evaluation yields an internal rate of return of 98% before taxes. For more results from the economic evaluation, see Table 4-II.

Table 4-II
*Economic summary of Mab production process.*

| | |
|---|---|
| Investment | $430 million |
| Revenue | $858 million/year |
| Operating cost | $12.1 million/year |
| Production rate | 256.041 kg/year |
| Unit production cost | $47 179/kg |
| IRR before tax | 98% |
| Net present value | $2 593 million |

# 5. Drug development

The first scenario considered in this work will show how drug prices could change if the clinical success rate increases. It is assumed that pharmaceutical companies are content with the present capital budgeting measures and will value future projects in the same way. Based on an increasing success rate and that the present development decision factors are constant, such as the internal rate of return, the possible impact on drug prices will be investigated.

## *5.1 Expected clinical cost*

As it has been described before by DiMasi, Hansen & Grabowski (2003) the expected clinical cost of a random new drug is $E(c) = p_I\mu_I + p_{II}\mu_{II} + p_{III}\mu_{III} + p_A\mu_A$ where $p_{I\text{-}III}$ is the probability of entering the clinical test phase and $p_A$ is the probability for long term animal testing during clinical trials. The $\mu$'s are conditional expectations, here $\mu_I$, $\mu_{II}$ and $\mu_{III}$ are the mean costs for drugs that enter phases I-III and $\mu_A$ are the mean cost for long term animal testing. If this expected cost is divided by the clinical success rate for drugs this yields the average estimate cost of approved products.

### *5.1.1 Clinical success rates*

The clinical success rates of drug development have been studied before. In the DiMasi (2003) study the clinical success rate was 15%. This is one of the lower success rates published and other studies give success rates up to 67%, see Table 5-III.

For 2003 the expected cost was found to be $60.6 million for a random drug and if this is divided with the clinical success rate of 15% this gives the development cost of an approved drug which is $402 million (DiMasi et al., 2003).

Table 5-III
*Clinical success rates from different studies.*

| Study | Year | Success rate |
|---|---|---|
| DiMasi, Hansen & Grabowski | 1991 | 17% |
| Struck | 1994 | 67% |
| Breggar | 1996 | 23% |
| Mackler and Gammerman | 1996 | 34% |
| Sofer and Hagel | 1997 | 23% |
| Avgerinos | 1999 | 15% |
| DiMasi, Hansen & Grabowski | 2003 | 15% |

### *5.1.2 Future success rates*

Comparing the 2003 study by DiMasi et al. to an earlier study performed by the same group in 1991 there was an decrease in the number of drugs that made it to phase II (71.0% versus 75.0%) and an increase in the number of drugs that made it from phase III to market approval (68.5% versus 63.5%).

This indicates that pharmaceutical companies are getting better at sorting out drug candidates that don't hold up for market approval at earlier stages in the development. If this trend continues with even better drug discovery technologies the development cost per approved drug will decrease as there will be fewer drugs in late development that will not make it to market.

The future success rate of drug development can be expected to rise again with the trend of better project selection together with technology improvements, mapping of the human proteome and better drug discovery pipelines to at least get back to the levels reported in the mid 90s which were around 25%.

### *5.1.3 Future drug development cost*

If the success rate increases it could reflect in lower overall drug development costs since there are more successful projects to share the cost of the failed ones. The cost per clinical phase can be approximated to stay the same since the cost per project per phase is likely to increase with prices in general but the trend of better project selection of drugs that make it into late clinical trials to market increases at the same time as projects that have a high risk of failing are aborted at an earlier stage.


*Figure 5-1: Effect of increasing success rates on development cost and predicted selling price.*


*Figure 5-2: The increasing IRR with increasing clinical success rate if the selling price is constant at $3350/g.*

As stated earlier a success rate of 25% will be considered as an example which in turn would lead to a drug development cost of $242 million as seen in figure 5-1. In order to find out how this would translate into drug selling prices, a series of economic evaluations of the same process as in section 4 is performed with the difference that the up front R&D cost is changed according to figure 5-1 and the selling price of the product is changed in order to find the solution where the IRR before taxes is the same as before, which is 98%. The price where this happens for a 25 % success rate is at a selling price of $2 062 per gram.

This means that there is room to lower the selling prices of pharmaceuticals if the clinical success rate would increase and the same project investment decisions is used. For the example of an increase in success rate to 25% this could lead to a price drop of $1 288 or 19% of the present base selling price before marketing markup etc.

If instead the analysis is done from the eyes of the pharmaceutical industry and the cheaper development costs is used to increase profit from projects, then the IRR will increase with the increase in success rate (fig. 5-2). With a success rate of 40% the IRR will be as high as 160%.

To conclude this part of the study it is beneficial to spend money in R&D to find new methods and processes that can improve the success rate of the drug development process as the drug discovery success rate has a very big impact on the profitability for the biopharmaceutical industry.

# 6. Cost distribution

The other option for pharmaceutical companies to lower costs in order to either decrease prices or increase profitability is to lower the production costs. The production is divided into two parts, the upstream part where mammalian cell lines are grown in bioreactors and the downstream part where the biomolecules produced are processed to meet purity and quality requirements. The cost distributions between the upstream and downstream sections are around 50/50 for product titers of 0.2 g/L (Sommerfeld & Strube, 2005).

In the upstream part the major optimization potential is in increasing the product titer since that is the limiting factor in how much product that can be produced in every batch. The mammalian cell lines used today for monoclonal antibody production reach titers up to 1g/L. Within the next 10 years the product titer is expected to reach approximately 10g/L which is thought to be a theoretical limit (Sommerfeld & Strube, 2005). If bacterial cell lines, such as *E. coli*, are successfully modified to produce humanized monoclonal antibodies the theoretical limit is expected to be around 40g/L.

### *6.1 Cost distribution for changing product titer concentration*

To see where the biggest possibilities for cost optimization are it is important to know the cost distribution between the upstream and downstream sections.

With low product titers the costs of upstream fermentation is bigger than cost of downstream processing, but with increasing product titers the upstream cost represent a smaller part of total production costs. Upstream, cost of goods decline due to economy of

scale, downstream this effect is not present. Therefore the downstream costs make up a bigger part of total production costs with increasing titers.

In figure 6-1 the trend of declining upstream cost can be seen. Specific costs of the upstream section are inversely proportional to the increased product concentration. This is true if the high secreting cell lines use the same medium and use the same culture length so that costs for labor and utilities are the same. This graph also show the declining overall production cost per kg of product since it is normalized against the product unit price. A development cost of increasing product titers is not accounted for here, but if included it should level out the difference between the upstream and downstream sections.

The decrease in production cost levels out with increasing product titers which means that in order to lower production cost at higher titers it is necessary to optimize downstream costs. At a product titer of 7 g/L over 75% of production costs are in the downstream section and it is therefore the place where further improvements have a bigger impact.

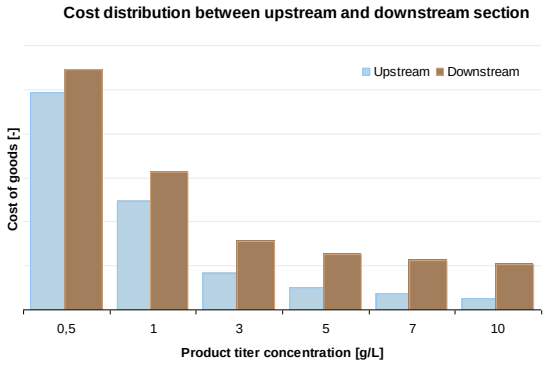


*Figure 6-1: Cost distribution of cost of goods between upstream and downstream section, normalized against unit production cost at 0.1 g/L.*

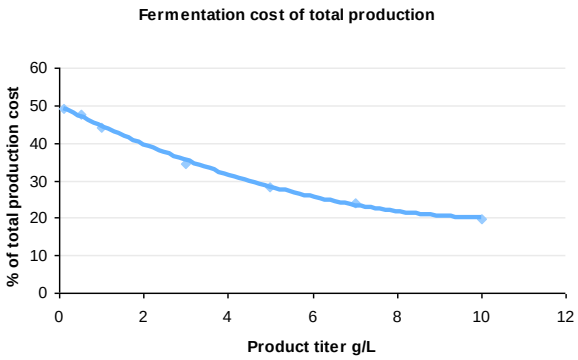


*Figure 6-2: Declining upstream costs as a function of higher product titer.*

# 7. Number of chromatography cycles per batch

One interesting study is to see if there is a point where the production costs can be optimized depending on whether the protein A chromatography step is run for one or for multiple cycles every batch. If the step is only run one cycle per batch that leads to shorter production times for every batch that is run and that way the production facility can be used for more batches every year or it can be used in other processes in the free time space.

But in order to run the chromatography step few times, bigger capital and up front investments have to be made since a larger and more expensive column and more chromatography media is required.

For example let's consider the scenario with a 5 kg production process per batch. In a one cycle per batch process the capital investment is $24.1 million but the fifty cycles per batch process only requires a capital investment of $21.2 million. When this difference is capitalized over a thirteen year payback time and an interest rate of 11% the cost of capital (interest plus amortization) every year will differ with $464 thousand in favor for the multi cycle process.

However, the time required for every batch is in this case 24 hours for a one cycle batch and 82 hours for every multi cycle batch. This leads to the possibility to run up to ten times as many single cycle batches than fifty cycle batches if the scheduling is

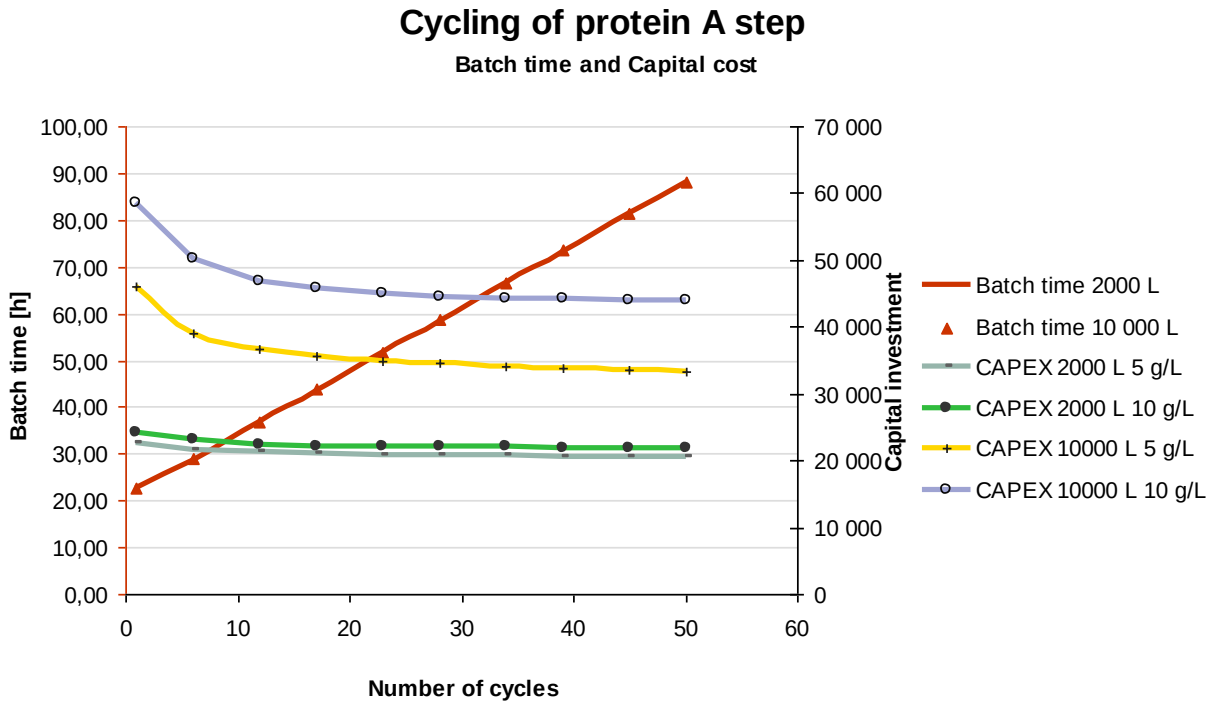


*Figure 7-1: The effect of changing the number of protein A chromatography cycles on batch time and on the capital investment for two different titers and for two different batch sizes. The capital investment is given in thousand dollars.*

optimized and the maximum number of batches is run.

When considering which approach is the most valuable one must value what the free time in the facility is worth and what the value of the increasing total production capacity is. In the case of producing monoclonal antibodies with selling prices in the range of $1000-$10 000 per gram it is clear that every possible batch of production is highly valuable and the opportunity cost of blocking the production facility could be much higher than the money saved when running the capture step for many cycles. This is visualized by plotting the batch time against the capital investment in fig. 7-1. The capital investment is given in thousand dollars. In this way it is possible to see the trend of quickly rising batch times versus the smaller decrease in capital investment when

increasing the number of cycles. The four different cases given as examples show the bigger effect of cycling as the batch size increase and it shows the small difference for a 2000 L batch with different titers.

There is a sharp drop in the capital investment when increasing the number of cycles at the lower end of the scale. This drop reflects the lower cost of chromatography columns. But the cost of columns is relatively cheaper with increasing size and the price does not follow the volume linearly. So when increasing the number of cycles economy of scale is lost and the batch time increases.

There is also a difference in the relative amount of capital investment decrease between different batch sizes. Larger batches have a higher potential of decreasing investment by running more cycles, and the difference between different product titers for a given batch increases as the batch volume increase.

If the value of the project is expressed in terms of its NPV (net present value) the values at different number of cycling steps can be investigated. Of particular interest is the area from one cycle to around fifteen cycles where the capital cost declines the most.

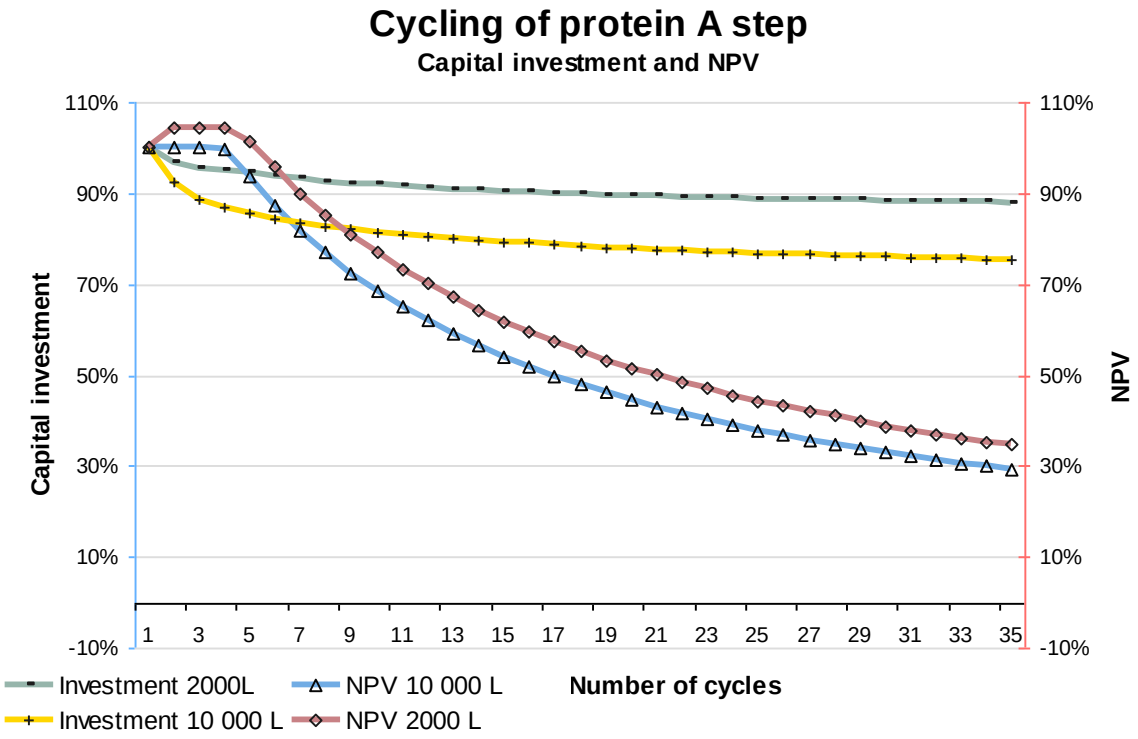


*Figure 7-2: The effect of changing the number of protein A chromatography cycles per batch on the capital investment and the net present value (NPV). NPV peaks at two to four cycles per batch. Both values are normalized against the batch time and the capital investment of running one cycle per batch for each case and use a 5 g/L product titer.*

The NPV of the 2000 L 5g/L process peaks at two to four cycles per batch (fig 7-2) which reflects that at more cycles per batch the money saved on columns and media does not make up the lower economy of scale and the time saved that can be used for running more production batches or to use the facility in other production. For a 10 000 L 5 g/L production process the pattern is similar with declining NPV when the protein A step is run for more than 4 cycles each batch.

# 8. Breakdown of costs

This section will be an example of the capabilities and an example workflow for the application. In order to study the economy of processes this section will study an example process with a 10 kg production target. First the capital investment and operating cost will be investigated to find the biggest costs in the process. After that a sensitivity analysis will be performed to calculate the input parameters of the process with the most influence on a target output parameter. On the most sensitive parameters parametric studies will be performed to see the effect if that parameter would change. To measure the level of risk in processes uncertainty analysis will be performed on the key output parameters.

*8.1 Capital investment and operating costs*

Looking at operating costs these can be divided into the sections of the process. This process consists of five sections. The first purification section includes the protein A capture step and the second purification section includes the purification and polishing steps as well as a virus filtration step. The operating cost distribution of the process can be seen in table 8-I and 8-II where it is divided per section and each section is divided into the main cost categories.

Table 8-I
*The distribution of cost over the process sections and cost categories for two batches per year and 4 cycles at the protein A step. For two batches the depreciation cost is not included as it represents 88% of total costs and would dwarf the other costs.*

| | *Main Section* | *Purification 1* | *Purification 2* | *Harvest* | **Total** |
|---|---|---|---|---|---|
| Raw materials cost | 0,5% | 11,6% | 4,3% | 5,2% | **21,7%** |
| Labor cost | 3,4% | 3,4% | 4,6% | 2,5% | **13,8%** |
| Consumables cost | 0,2% | 13,3% | 47,2% | 3,8% | **64,5%** |
| **Total** | **4,1%** | **28,3%** | **56,1%** | **11,5%** | **100%** |

Table 8-II
*The annual cost for 40 batches per year including depreciation cost.*

| | *Main Section* | *Purification 1* | *Purification 2* | *Harvest* | **Total** |
|---|---|---|---|---|---|
| Raw materials cost | 0,4% | 9,6% | 3,6% | 4,4% | **18,0%** |
| Labor cost | 2,8% | 2,8% | 3,8% | 2,0% | **11,5%** |
| Consumables cost | 0,1% | 11,0% | 39,2% | 3,2% | **53,5%** |
| Depreciation | 2,1% | 6,7% | 4,3% | 4,0% | **17,1%** |
| **Total** | **5,5%** | **30,1%** | **50,8%** | **13,5%** | **100%** |

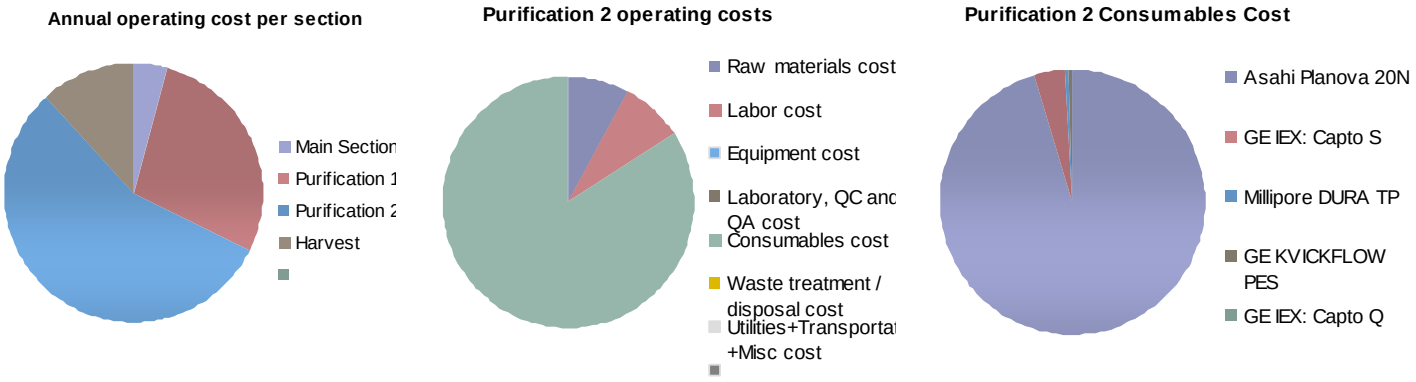


*Figure 8-1: Breakdown of operating costs for the two batches per year case.*

Of the operating costs for a two batch process, the biggest costs are in the second purification section. This section of the process represents more than 50% of the total cost in downstream processing. Of these costs the major part is from consumables (84%) and of these 95% is from the filters used in virus filtration. Another major cost in downstream processing of Mabs is often thought to be the protein A media. In this process the protein A media accounts for 13% of the total operating cost which is a considerably smaller amount than the total 45% for the virus filtration. The labor cost of the process is low, around 14 % of operating costs. Raw materials account for 20% of costs and of which the biggest part is from the first purification section.

If operating costs for 40 batches per year and the depreciation costs of equipment are included in the study, the cost pattern is similar but depreciation account for around a fifth of total costs. In this case the capture section of the process will account for a slightly bigger part of costs and the cost of the purification section will not be as dominant (down to 50% of total costs).

*8.2 Sensitivity analysis*

To find the parameters with the biggest impact on downstream production costs sensitivity analysis is performed using Monte Carlo simulations. In this test the parameters distributions are set according to table 8-III. Buffer costs are set to vary within a 10 % interval from their original cost. The monoclonal product titer is set to vary within a normal distribution reflecting the natural batch to batch variation in concentration level. Resin binding capacities are set to vary from moderate expectations of their performance up to their proposed max capacities. The resin lifetime is set to vary from a very short lifetime used in some pilot scale processes, to long lifetimes used by others for blockbuster drug production. The labor cost is also included to see if it would be useful to locate a production facility in a low cost country based on the operator's salaries.

Table 8-III
*Distributions used for input parameters in sensitivity analysis.*

| **Parameter** | **Name in fig. 8-2** | **Range** | **Distribution type** |
|---|---|---|---|
| Labor cost | Operator | 13-200 $/h | Uniform |
| Affinity column resin binding capacity | SURE RBC | 28-70 g/L | Uniform |
| Affinity column resin lifetime | SURE RF | 20-400 cycles | Uniform, discrete |
| Cation column resin binding capacity | Capto S | 40-250 g/L | Uniform |
| Cation column resin lifetime | CS RF | 20-200 cycles | Uniform, discrete |
| Anion column resin binding capacity | Capto Q | 30-250 g/L | Uniform |
| Anion column resin lifetime | CQ RF | 20-200 cycles | Uniform, discrete |
| Mab product titer | IgG | Mean 5 g/L Std dev 0,2 | Normal |
| **Buffers** | | | |
| CIEX wash1 | | 1,90-2,10 $/kg | Uniform |
| CIEX B1 | | 1,90-2,10 $/kg | Uniform |
| CIEX Strip | | 1,90-2,10 $/kg | Uniform |
| CIP1 | | 1,90-2,10 $/kg | Uniform |
| Prot A Elute | | 1,90-2,10 $/kg | Uniform |
| CIP2 | | 1,90-2,10 $/kg | Uniform |
| Prot A Wash | | 1,90-2,10 $/kg | Uniform |
| NaOH (1 M) | | 1,90-2,10 $/kg | Uniform |
| AIEX pH | | 1,90-2,10 $/kg | Uniform |
| AIEX eq | | 1,90-2,10 $/kg | Uniform |
| Prot A Load | | 1,90-2,10 $/kg | Uniform |

Using these input distributions a Tornado chart is generated since this gives the best accuracy when analyzing a large number of assumptions. All the variables used in the sensitivity analysis were ranked based on their contribution to variance in the selected forecast, which in this case was the unit production cost of the monoclonal antibody.

In this simulation the parameter that was found to influence the target forecast variable the most was the Mabselect SURE replacement frequency followed by the Mabselect SURE resin binding capacity and the Capto S resin binding capacity as the third most important factor, see figure 6-2. Of the buffers it is the protein A buffers that

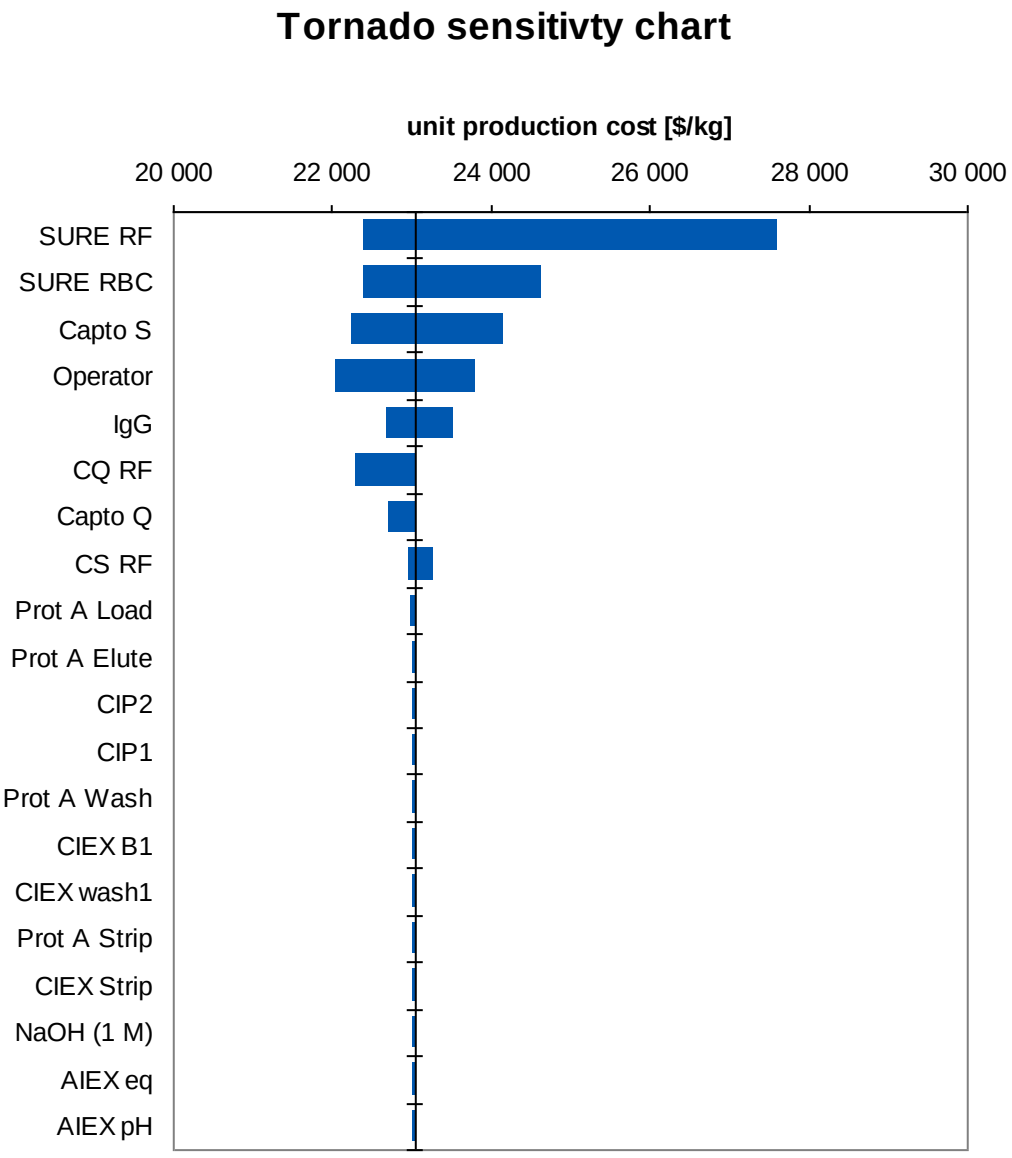


*Figure 8-2: Tornado sensitivity chart from a Monte Carlo simulation with parameter distribution according to table 8-III.*

are generally the most sensitive since they are used in the biggest volumes but all buffers have a low sensitivity in this model. The parameters with the least contribution to the production cost are the buffers and the replacement frequency of the Capto S media.

When a sensitivity analysis is performed the parameters with low impact on the target forecast can be omitted and new analyses can be performed with a lower number of assumption variables. This will increase the accuracy of the simulation while the number of trials needed to reach significant results will be lower. Also, the simulation will give information about which parameters might be interesting to study closer in a parametric study where the effect of changing a single or two parameters in isolation can be studied.

*8.3 Parametric studies*

The choice of which parameters to perform parametric studies on can be based on the results from the sensitivity analysis and any other variables that are thought to be of interest can be studied. In this application variables will be studied either by themselves or together in pairs to generate a matrix of all possible combinations between the two. This will generate either a line graph that shows the variation in a chosen forecast value or in the case of a pair of variables it will generate a surface chart of the results.

From section 8.2 the most sensitive parameters are clearly interesting to study further and also the effect of for example product titer and labor cost is interesting to include in the studies.

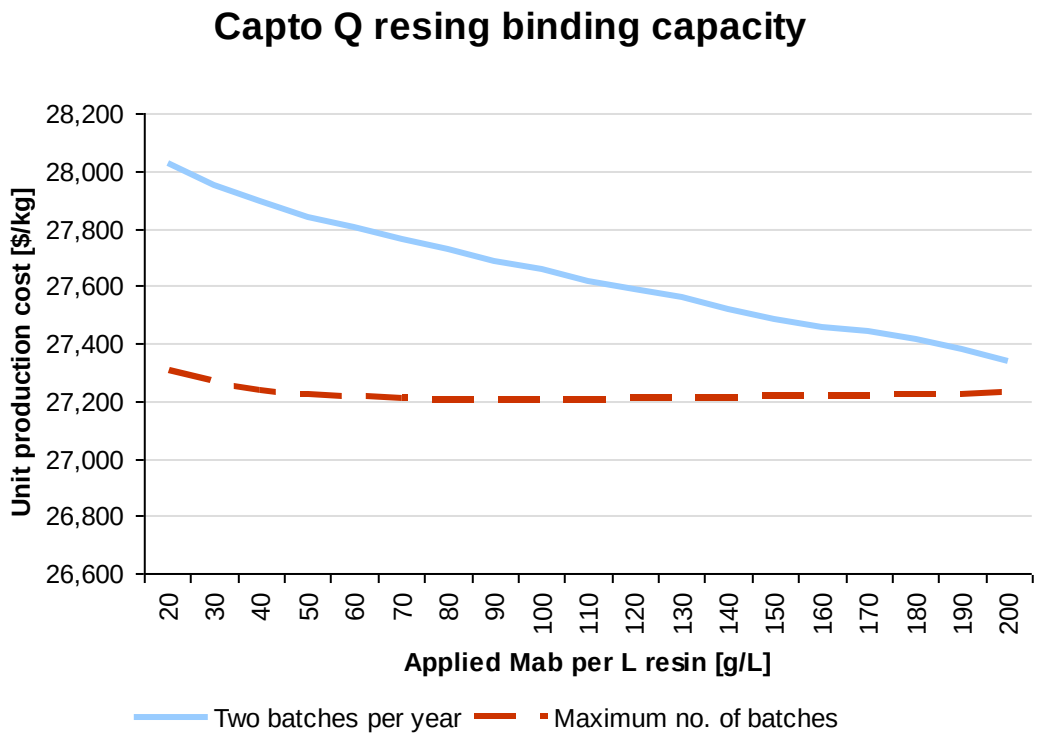


*Figure 8-3: The unit production cost with changing resin binding capacity when limiting the number of batches per year and when the number of batches is calculated to its maximum value.*

*8.3.1 Capto Q resin binding capacity*

In the sensitivity study the Capto Q applied Mab per L resin is changed from 20 to 200 g/L and the same interval can be used in a parametric study to see the effect of the amount of loaded Mab per L resin on the unit production cost of the antibody under production in this case.

As seen in figure 8-3 the decrease in unit production cost is more or less constant over the entire range if the number of batches or campaigns per year is low. This is the effect that a higher resin binding capacity will require a smaller column size and smaller amounts of chromatography media.

If the process is calculated with the maximum number of batches in mind, which reflects the situation where a production facility and equipment is used in the most efficient way, the binding capacity will reach an optimum around 100 g/L. The slight increase in production cost at the upper end of the scale is due to the fact that with increasing binding capacities the loading time of the chromatography step increases and the cost of longer cycle time cancel out the money saved on a smaller bed volume.

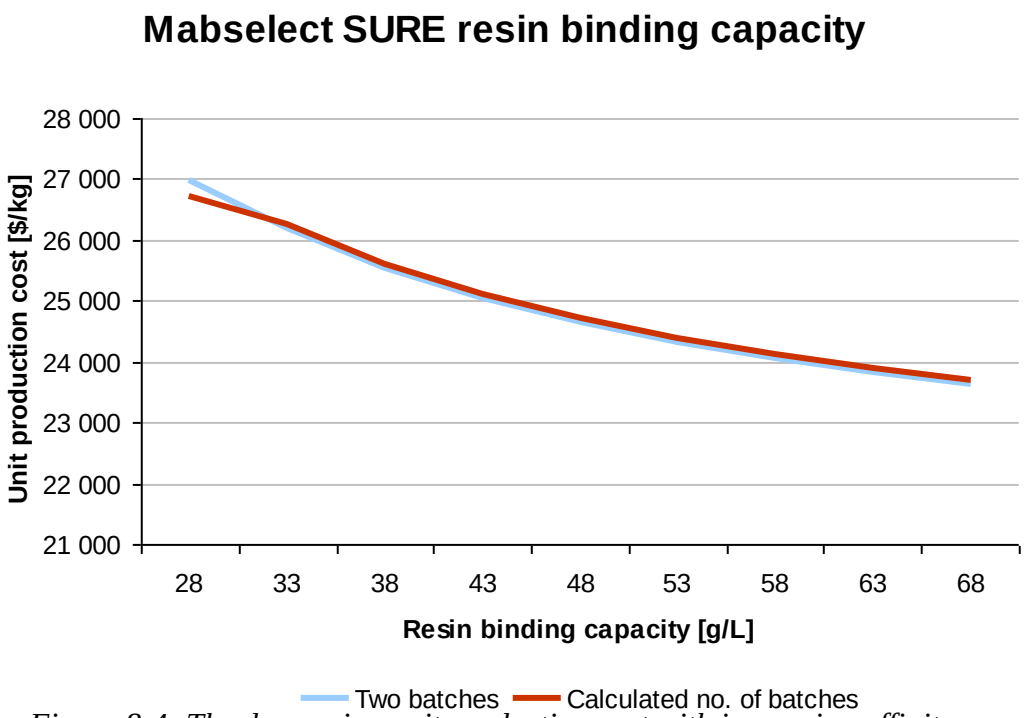


*Figure 8-4: The decreasing unit production cost with increasing affinity chromatography resin binding capacity.*

*8.3.2 Mabselect SURE resin binding capacity and replacement frequency*

The protein A chromatography step only shows a decreasing cost behavior when increasing the resin binding capacity (fig 8-4), both for two batches per year and for the calculated maximum number of batches per year. This difference to the ion exchange step

is due to the fact that under the interval of study the increasing loading time does not get big enough to cancel out the effect of the smaller volumes of media used.

If the effects of binding capacity and replacement frequency are studied together the importance of running the media for the specified number of cycles is clear. As seen in fig 8-5 the impact of running only a few cycles before changing the resin is big but the effect decreases gradually as the number of cycles get up to the specified limits.

The effect of replacement frequency on unit production cost is not linear (fig 8-6), this could be counterintuitive and is studied further. The cost distributions for a replacement frequency of 20, 40, 60 and 80 cycles and dynamic binding capacity of 40 g/L are plotted in figure 8-7. The cost of media is around $12 000/L and 53 L is used. If the total cost of media is then divided with the replacement frequency it gives the cost of protein A for each cycle.

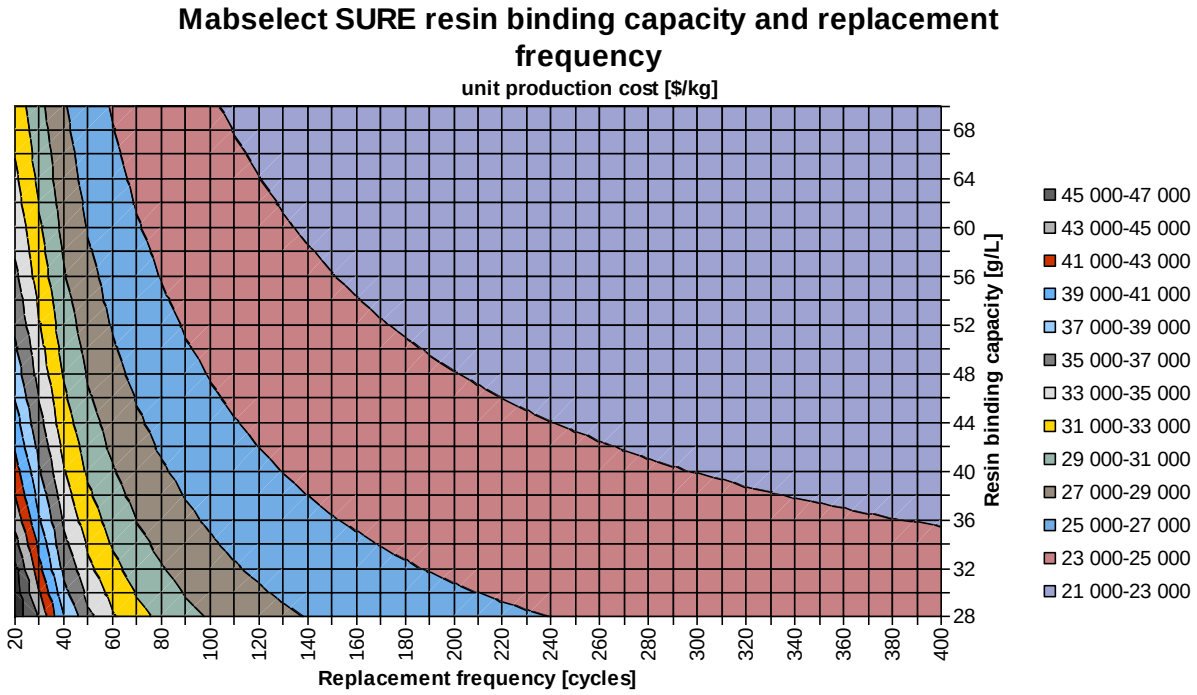


*Figure 8-5: A simultaneous look at Mabselect SURE resin binding capacity and replacement frequency that show the smaller decrease in unit production cost as the replacement frequency and resin binding capacity increases.*

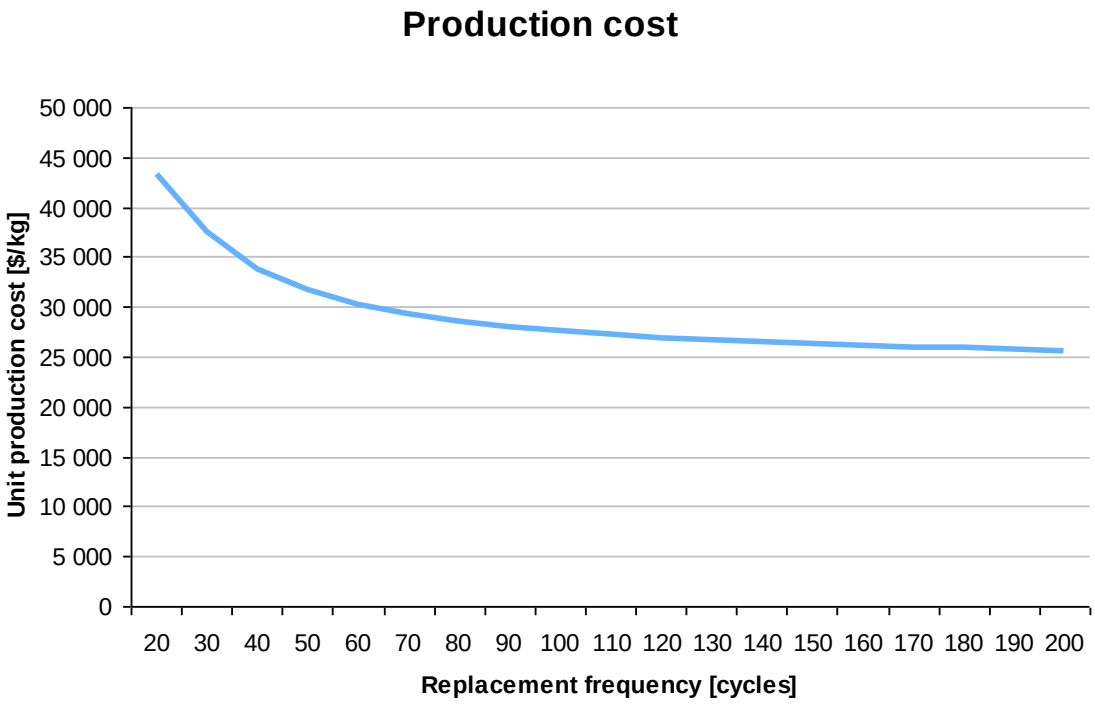


*Figure 8-6: The production cost at a resin binding capacity of 40 g/L. The decrease in cost is not linear is in accordance with fig 8-5.*

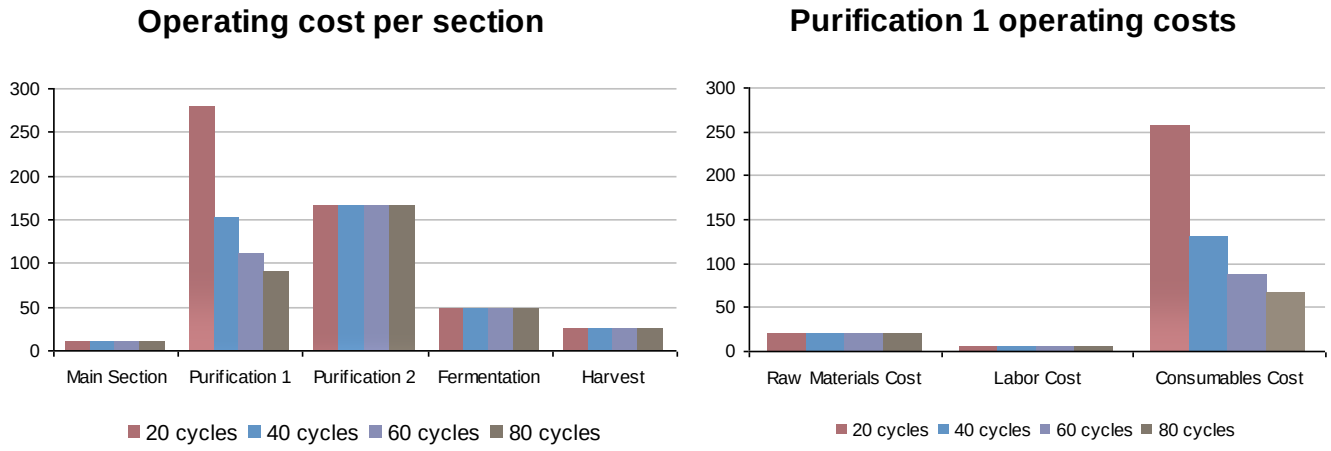


*Figure 8-7: Breakdown of costs with a resin binding capacity of 40 g/L reveals that it is only the cost of the resin that is decreasing.*

*8.3.4 Labor cost*

Depending on the location of a production facility the labor cost of its staff will vary according to the wages in that country. The impact of labor cost is as seen in figure 8-4 linear and the fairly small impact of the wages over this wide range makes it one of the parameters that can be omitted from further studies.

For every batch a total of 87 operator hours is used. The average number of operators in the process is 3


*Figure 8-4: The increasing unit production cost with increasing labor cost.*

*8.4 Uncertainty analysis*

Uncertainty analysis is done by running a Monte Carlo simulation of 21 000 trials with assumption distributions from table 8-II. By monitoring the distribution of a selected forecast parameter some conclusions can be made. For example, it can be informative to look at the production rate for an estimate of the planned production rate. In this case the $10^{th}$ and $90^{th}$ percentile is at 11.53 and 12.79 kg product per two batches and the mean production is 12.16 kg, which corresponds to a fairly large variation. As this difference solely dependent on the product titer in the input stream it is very important to measure the product yield and variability in the upstream process to minimize the uncertainty in the production as much as possible.

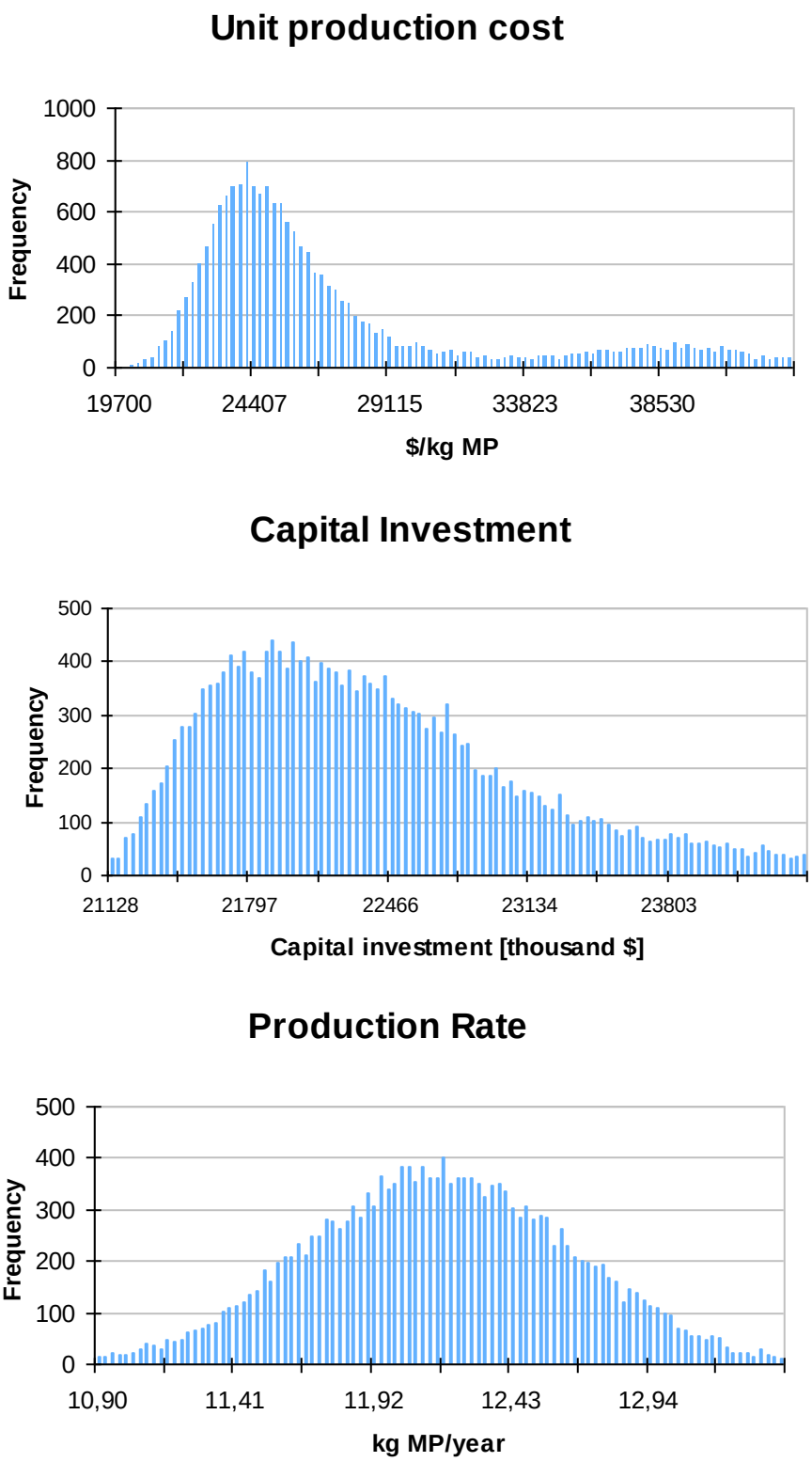


*Figure 8-8: Uncertainty distributions of forecast values from a simulation with 21 000 trials with assumption parameters according to table 8-II.*

The production cost is at a mean of 24 752 $/kg of monoclonal antibody. However, the cost distribution curve is not uniform and the cost interval reaches from 19 708 $/kg up to 33 987 $/kg. The probabilities for the higher costs are however low and the 90$^{th}$ percentile is at 28 869 $/kg. The impact of this wide distribution of cost on the design of the process is in the importance of reducing as much uncertainty as possible in each of the input variables and then rerunning the simulation with hopefully smaller uncertainty. In the case of study some inputs are given wide distributions that more resemble the development of the technology than the internal variation in a given process and this also results in less accuracy in the result.

The capital investment of the process is also not entirely uniform in its appearance and displays a big variation compared to its mean value.

If these forecast variables display an uncertainty too big for the projects target, attempts should be made to reduce the uncertainty. This is done by going back to the sensitivity analysis to perform additional studies of the most sensitive parameters to lower their uncertainty levels. After that, additional simulations with the new distributions are run and the new uncertainty of the forecasts is analyzed.

# 9. Discussions

The main part of the thesis has been the design and construction of the interface in Microsoft Excel that acts like a bridge between the functionality of the applications SuperPro Designer, Excel and Crystal Ball. As the analysis performed on a representational model of monoclonal production process was not the primarily target, some of the assumptions made may not be totally accurate. Also, since the simulation of a process requires a significant amount of CPU time only one process is analyzed, as an example the uncertainty analysis with 21 000 trials ran for more than 50 hours at a rate of around 400 trials per hour. This process does not reflect the data in any one single process but the variables and their distributions are chosen to represent both present and state of the art data together with data that cover many of the situations in use today. Varying the resin replacement frequency between 20 and 200 cycles is an example of this. The implication of this are that the sensitivity study and the uncertainty analysis are made as an example of a possible workflow rather than a case study of an actual process in design.

## *9.1 Improvements*

The interface in its present state represents a work in progress as all applications are. Since it is built around three other applications it will require a constant overview of the functionality as new versions of the other applications are released. With new versions of the other applications the possibility of adding new functionality to the interface should also not be forgotten.

In the present release there are a number of small features that were not possible to implement due to restrictions in the connected applications. Features that I would like to add when the possibility appears include:

- Protection of worksheets. Worksheets in Excel should be protected so that only cells marked green can be edited by the user. This will make it possible only to edit the right cells and impossible to edit cells containing formulas. The present limitation in Excel is that protection of worksheets cannot be combined with expandable rows.
- The number of forecasts displayed in the report worksheet should be dynamic. Presently any number of forecasts can be used, but only the three presets on the sheet will generate the preformatted diagrams and lists with data.

*9.2 Further development*

With the release of version 7.0 of the SuperPro software several new features are available and the use of these functions can further improve the functionality of the interface. Of special interest for this project is the extended availability to set the section specific capital cost adjustments. For example, is it now possible to set the amount of working capital and also R&D and marketing costs, which is very useful when evaluating the process economy. The possibility to change these factors should be added to the Flowsheet worksheet of the interface and below is a list (table 9-I) of functions that should be added to later versions of the interface. With version 7.0 it is also possible to have functions that dynamically makes lists of the sections, equipment etc. of the process which eliminates the need of some of the set-up functions in the present interface.

Table 9-I
Categories of functions to add to later versions of the interface.

| **Category** | **Example of settings** |
|---|---|
| Section Capital Cost - DFC | LANG factors etc. |
| Section Capital Cost - Misc | Working capital, start-up cost, R&D etc. |
| Section Operation Cost - Facility | Maintenance, depreciation etc. |
| Section Operation Cost - Utility | Electricity usage etc. |
| Section Operation Cost - Misc | Process validation, other costs etc. |
| | |
| Dynamic lists of: | |
| Unit procedures | |
| Operations | |
| Equipment | |
| Streams | Input, output and intermediate |
| Stock mixtures | |
| Sections | |
| Branches | |

# 10. Conclusions

The work on the interface has shown the possibility of writing a general interface capable of analyzing different processes with a small amount of adaptation for each process. The interface is able to perform sensitivity and uncertainty studies using Monte Carlo simulations and performs automated parametric variations.

The results from the studies have shown that as the upstream product titer goes up the main part of production costs shift to the downstream section. The next part showed that there are an optimum number of cycles for the chromatography steps where the net present value of the project reaches a maximum.

The sensitivity study showed as expected that the protein A resin binding frequency and replacement frequency were the most sensitive parameters in the system and cost of buffers the least sensitive. It also ranked the labor cost fairly low despite its very wide distribution which means that labor cost is not that important when projecting for a monoclonal antibody production facility.

# 11. Acknowledgements

I would like to thank Karol Lacki for his wonderful support in this project, both as a supervisor and as a friend. I have enjoyed working with the entire process economy project team, Matthias Bryntesson, Kjell Eriksson and Ingrid Long and thank you all for your support. Also I would like to thank all my new friends at the RULD section at GE Healthcare in Uppsala who has all been wonderful.

# 12. References


1. McClelland, M. *Global Biotechnology Report*. Ernst & Young LLP. **2004**, 39-40.
2. Sommerfeld, S., Strube, J., *Challenges in biotechnology production – generic processes and process optimization for monoclonal antibodies*. Chemical Engineering and Processing. **2005**, 44, 1123-1137.
3. Farid, S.S., et al., *Decision-Support Tool for Assessing Biomanufacturing Strategies under uncertainty: Stainless Steel versus Disposable Equipment for Clinical Trial Material Preparation.* Biotechnol. Prog. **2005**, 21, 486-497.
4. Levine, H.L., Latham, P., *The Use of Models to Estimate and Control the Cost of Biopharmaceutical Manufacturing.* Bioprocess Technology Consultants. No date.
5. Petrides, D.P., et al., *The Role of Process Simulation in Pharmaceutical Process Development and Product Commerialization.* Pharmaceutical Engineering. **2002,** 56-65.
6. Petrides, D.P., et al., *Throughput Analysis and Debottlenecking of Biomanufacturing Facilities.* BioPharm, **2002**, August.
7. DiMasi, J.A., et al., *The price of innovation: new estimates of drug development costs. Journal of Health Economics,* **2003**, 22, 151-185.
8. DiMasi, J.A., et al., *Cost of innovation in the pharmaceutical industry.* Journal of Health Economics, **1991**, 10, 107-142.
9. Sofer, Hagel & Jagschies, *Handbook of process chromatography*. **2007**, in press.
10. Crystal Ball manual. Intelligen. **2001**.
11. Genentech homepage, http://www.gene.com, **15 Jan. 2007**.