Handling of uncertainty in hierarchical production planning

Chermin, B.

Award date:
2013

Eindhoven, September 2013

Handling of uncertainty in hierarchical production planning

by

Bob Chermin

BSc Industrial Engineering — INPG 2005

Student identity number 0615554

in partial fulfilment of the requirements for the degree of

Master of Science

in Operations Management and Logistics

Supervisors:

Prof.dr. J.C. Fransoo, TU/e, OPAC

Prof.dr. J.S.H. Van Leeuwaarden, TU/e, SPOR

Bob Chermin

TUE. School of Industrial Engineering.

Series Master Theses Operations Management and Logistics

Subject headings: production planning, uncertainty planning, stochastic control, pharmaceutical industry, hierarchical planning, design of planning concept

ACKNOWLEDGEMENTS

During the last 6 months I have had the privilege of doing a masters thesis project at Merck. This has been a challenging process, my first experience in an operational environment. The challenges had all kinds of dimensions, from motivational to organizational (an organization in transit, with some surprising management mechanisms) and of course theoretical.

From the motivational part I would like to thank the Hot Jalapeños and my group of friends from Roermond, for rubbing it in my face that the summer was exceptional this year. Furthermore I would like to thank Rasa Raoufi, David Brandstädter, Maria-Alexandra Bujor, Boudewijn Rosenmöller, Bart Ripperda, Viktor Tielen, Hugo Driesen and Joris Olsthoorn for showing that a shared problem is half the problem. Besides the group of old friends a word of thanks to the new people I met at MSD and especially the people in CO 1332. I might have complained about the noisiness a bit, but the spirit and ethos was very nice and relaxed. Besides that a word of thank to my family, that is Wiek Chermin, Jola Nelissen, Yvonne van Uden, Jan Platier, Geertje Stelten-Chermin and John Stelten, for supporting me at any point in time they were needed.

From the organizational perspective the many talks I have had with the different organizational departments, including the quality and off course production department, were very insightfull. A special word of thanks to (in arbitrary order) Adrienne Oerlemans, Wenny Raaymakers, Leon Verrijt and Hans Maasakkers. All of them helped, not only by giving insight in the process at hand, but also by answering and discussing numerous other organizational surprises I faced during my time at MSD. I enjoyed the discussion and interaction and was positively surprised by the eagerness to really make the project work for all parties. This last is probably best to be seen by my direct company supervisor, Ruud Grotenhuis. Despite having a busy job as is, he made time for me whenever I needed it and was very helpful in keeping the project relevant for the company.

For the more theoretical challenges a strong word of thanks towards both of my university supervisors, Jan Fransoo and Johan van Leeuwaarden. They have proven again to me that they are top of the bill at our university. The interactions during the project were both fun and challenging and I hope that we can continue to stay in contact after I have graduated.

A final word of thanks for the people that have helped me to develop to the person I am today, without that it would not be possible to graduate right now. Here my thanks go to the people I met during my board year at Integrand Nederland, all of my Mexican friends, Ingmar Frey in particular, Bart Jacobsz Rosier, and René Willemars, who have been a supportive friends for many years.

A closing word to you as a reader, you are probably one of the people mentioned above, or a person that will do a thesis in a similar subject or environment. In the latter case, make sure you enjoy the process of the thesis and take every possibility to learn. Not just from a theoretical perspective, but also look to become a little bit wiser and a little bit less naïve, as I would like to think I have become in these last couple of months.

MANAGEMENT SUMMARY

This case study studies planning hierarchies in a highly uncertain environment. In this case study a planning hierarchy was designed for one of the subsets of a production facility of pharmaceutical producer MSD in Oss, the Netherlands. For this planning the question was whether it was possible to remove uncertainty buffers in the system without worsening the service in terms of service while improving the performance in terms of costs. This proves to be possible, given the correct *design* choices and the use of the correct *models* for the different decisions. That means, applying both in such a way that they together effectively reduce the risk that comes from uncertainty in an efficient way.

Central in the planning hierarchy was the uncertainty that exists in a production environment. To determine whether this uncertainty should be covered for in design or be incorporated in the models a framework was developed that groups uncertainty based on the *influence* a planning can have on this uncertainty. In essence one recognizes uncertainty that can be influenced to a certain extend and uncertainty that cannot be influenced. Furthermore, one has to assess the impact of the uncertainty to determine the appropriate response in a planning hierarchy. Based on these two dimensions a certain action is followed as shown in Table 1.

	Possible to have influence on uncertainty	**Not possible to have influence on uncertainty**
High impact	Stochastic control function (modeling)	Isolate and buffer through smart design
Low impact	Deterministic control function (modeling)	No action

TABLE 1, DESIGN VS MODEL, WHERE UNCERTAINTY SHOULD BE HANDLED

The logic is simple, if it is not possible for the planning to influence the uncertainty in any way, it makes no sense to control it in a model. Therefore, it should be isolated to its core and be buffered for through smart design. If it is possible, one should do it in the way that introduces sufficient complexity given the impact of the decision.

In case it is possible to have influence on the uncertainty, one should strive to use stochastic control when there is a high impact on the objective. In case of low impact on the objective one still wants to make the appropriate decision, but it is not necessary to take the uncertainty into account, hence a deterministic control structure should be sufficient.

DESIGN

Smart design, in this case this incorporates isolation of product groups based on throughput time/quality type and demand size. By isolating certain groups in the system that create high uncertainty, total cost can be reduced and flexibility can be gained over the whole process.

Furthermore, smart design includes prioritizing cases in a production system such that possible delays earlier in the production process can be nullified. This again greatly reduces the standard deviation of throughput times in the system, making the performance better.

When looking at the impact for uncertainty, one rule should be to apply stochastic control for the bottleneck resource in the system. This way one ensures performance in terms of throughput times, without making the control model needlessly complex.

MODELING

During this study, several models were built. The results that are the most interesting is the model that is used for stochastic control of the bottleneck resource and its interaction with the model that is used to make the optimal batch production decision.

The stochastic control model is used for the bottleneck resource to determine the maximum load or minimal number of servers necessary not to exceed the planned lead time for the resource. This was done with the quality and efficiency driven regime (QED). This model is used for waiting time approximation based on the M/M/C waiting queue and shows a surprisingly good fit with the behavior in the actual waiting time, even for very unstable utilizations (rho) on the system.

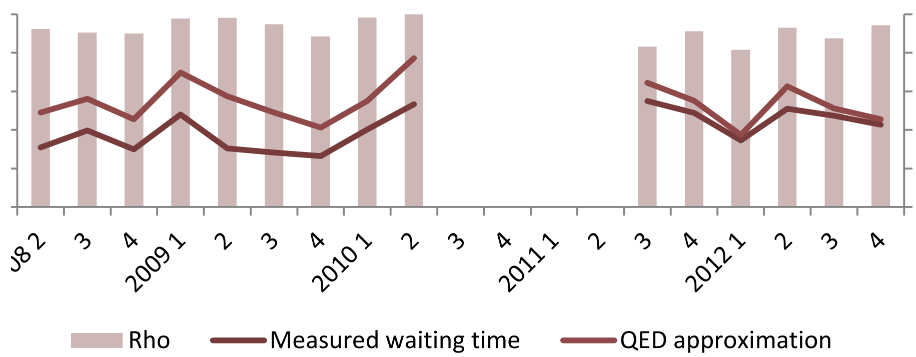

FIGURE 1, QED WAITING TIME VS ACTUAL WAITING TIME

Based on this QED it is possible to determine the amount of servers that minimizes the cost. The current behavior of the IPT WH seems to only minimize the cost (quite effectively) while not taking into account the throughput time that is associated with these costs.

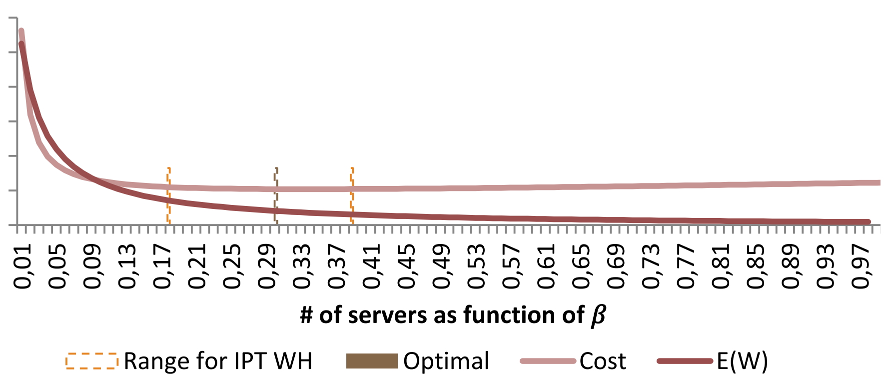

FIGURE 2, COST FOR DIFFERENT AMOUNT OF SERVERS

Furthermore, it is interesting to see that the optimal costs of the system are approximately linear in the load on the system the cost of a server for the optimal β. This makes the model interesting to apply in a deterministic optimization model.

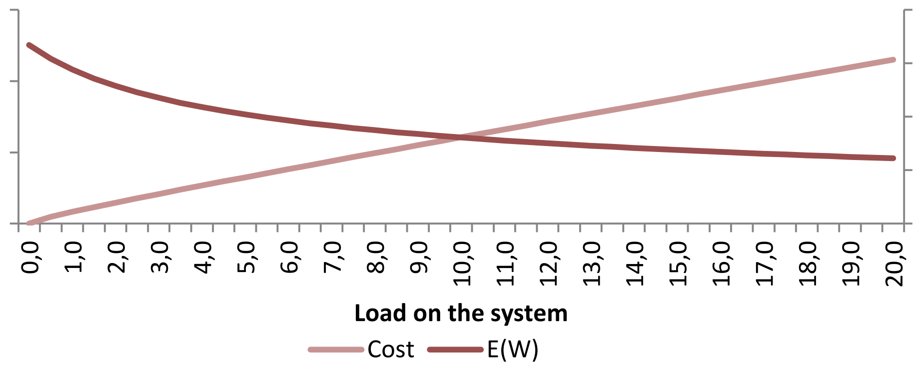

FIGURE 3, COST DIFFERENCE FOR DIFFERENT LOADS ON THE SYSTEM WITH OPTIMAL CAPACITY

Knowing the cost curve is flat around the optimum and linear in the load, and knowing that the waiting time is decreasing in β and the load it is possible to determine a maximum load on the system given throughput times and an amount of servers, or an amount of servers given a load, which can be used as a restriction in a linear model.

These results can be generalized in case the fraction $\frac{holding\ cost}{server\ cost}$ is small and also hold for small systems with relatively low number of servers.

For the production plan a mixed integer linear program (MILP) was constructed. This MILP takes into account backorder cost, holding cost and startup cost for the different stages in the process. Furthermore it takes into account a capacity restriction or a load restriction for the different production units.

RESULTS

Applying the smart design and relatively simple control models in the production hierarchy proves to be beneficial in terms of throughput time average and standard deviation, increasing production plan reliability and ultimately costs in the production process. Having all of these in place greatly reduces costs in the system without worsening service performance.

CONTENTS

Acknowledgements ... III
Management Summary ... IV
List of tables ... 0
List of figures ... 0
List of variables ... 1
List of abbreviations ... 2
1. Introduction; IPT women's Health ... 1
1.1 MSD Pharmaceutical operations Oss ... 1
1.2 IPT Women's Health Production ... 1
1.2.1 Bulk production ... 2
1.2.2 Bulk release ... 2
1.2.3 Packaging ... 3
1.2.4 Package release ... 3
1.2.5 Unit flow ... 4
1.2.6 Production lead time ... 4
1.2.7 Utilization of bottleneck resources ... 5
1.2.8 Arrival of demand ... 5
1.3 Organization and Planning structure ... 6
1.3.1 The OP tool ... 7
1.4 Concluding ... 7
2. Literature review on hierarchical planning and uncertainty ... 8
2.1 Hierarchical production planning ... 8
2.2 Effectuation lead time and anticipation ... 8
2.3 Modeling of uncertainty in hierarchical planning ... 9
2.4 Gaps in literature ... 10
2.5 Concluding ... 10
3. Research contribution ... 11
3.1 The case study ... 11
3.1.1 Case scope ... 11
3.2 Research questions ... 12
3.3 Research methodology ... 14
3.4 Concluding ... 14
4. *Design* of the planning concept of IPT WH ... 15

4.1 Overall planning concept 15
4.2 Uncertainty that can not be influenced 16
4.2.1 Uncertainty in demand 16
4.2.2 Uncertainty in production deviations 17
4.2.3 Uncertainty in production time 17
4.2.4 Uncertainty in API 17
4.3 Uncertainty that can be influenced 18
4.4 Decision sequence in the planning hierarchy for IPT WH 18
4.5 Concluding 19
5. Stochastic models in decision support 20
5.1 Timing of planned capacity constraining activities 20
5.1.1 Model requirements 20
5.1.2 PCCAs model 20
5.2 Stochastic bottleneck production unit control 21
5.2.1 Model requirements 21
5.2.2 The quality and efficiency driven regime 22
5.2.3 Comparison of QED with M/M/C for low number of servers 22
5.2.4 Model fit on real data 24
5.2.5 Control through the QED 26
5.3 Concluding 28
6. Deterministic decision support 29
6.1 Aggregate capacity scaling 29
6.1.1 uncertainty in demand 29
6.1.2 Uncertainty in planned capacity available 30
6.2 Order release decision bulk 30
6.2.1 mixed integer linear program 31
6.2.2 Uncertainty in demand 32
6.2.3 Uncertainty in throughput time 32
6.2.4 Use of the ORDB 33
6.3 concluding 34
7. Testing the new unceraintyhandling 35
7.1 Production units 35
7.2 Additional assumptions 36
7.3 Validation of the simulation model 36
7.4 Planned lead times Bulk production quantity model 37

7.5 New vs old uncertainty handling 39
7.6 Concluding 39
8. New Bulk production plan 40
8.1 Production plans 40
8.2 Different scenarios 40
8.3 Sensitivity analysis 41
8.4 concluding 41
9. Generalization of the QED model 42
9.1 Minimizing the cost in QED 42
9.2 Planned lead time 44
9.3 Implications of results 45
9.4 Concluding 45
10. Conclusion 46
10.1 Case assignment 46
10.2 How can the *design* of a production planning hierarchy assure the correct handling of uncertainty 46
10.2.1 Where to handle uncertainty 46
10.2.2 Smart design 47
10.3 How can *modeling* in a production planning hierarchy assure correct handling of uncertainty 47
11. Further research and limitations 49
11.1 Further research 49
11.2 IPT WH 49
Bibliography 0
Appendix I Bulk order production model, production plans 4
Sensistivity analysis; backorder costs 6
Sensitivity analysis; Server costs lab 9
Sensitivity analysis; Startup bulk production 11

LIST OF TABLES

Table 1, design vs model, where uncertainty should be handled ... IV
Table 2, Design vs Model, where uncertainty should be handled ... 13
Table 3, decision structure IPT WH ... 19
Table 4, Bulk release model ... 32
Table 5, Planning actions ... 34
Table 6, additional assumptions simulation ... 36
Table 7, Planned lead time per production unit ... 38
Table 8, Design vs model, where uncertainty should be handled ... 46
Table 9, production plan batch sizes free ... 4
Table 10, Production plan campaign sizes free ... 5
Table 11, Production plan Campaign sizes fixed ... 5
Table 12, Production plan free campaigns backorder costs 6 % ... 6
Table 13, Production plan Free campaign backorder costs 8 % ... 7
Table 14, Production plan free campaign backorder costs 10 % ... 7
Table 15, Production plan free campaign backorder costs 12 % ... 8
Table 16, Production plan free campaign server costs at 50 % ... 9
Table 17, production plan free campaign server costs at 100% ... 10
Table 18, Production plan free campaign server costs at 200 % ... 10
Table 19, Production plan free campaign startup cost bulk 2,5 % ... 11
Table 20, Production plan free campaign startup cost bulk 5 % ... 12
Table 21, Production plan free campaign startup cost bulk 10 % ... 12

LIST OF FIGURES

Figure 1, QED waiting time vs actual waiting time ... V
Figure 2, cost for different amount of servers ... V
Figure 3, Cost difference for different loads on the system with optimal capacity ... VI
Figure 4, IPT Women's Health poduction process ... 2
Figure 5, details of bulk production ... 2
Figure 6, Details in bulk release ... 3
Figure 7, details of packaging ... 3
Figure 8, Average throughput time development in working days ... 4
Figure 9, Standard deviation of throughput time development in working days ... 4
Figure 10, utilization of bottleneck resources per production unit ... 5
Figure 11, regular and tender demand 2012 ... 5
Figure 12, Eindhoven production framework (Jansen M. , 2011) ... 8
Figure 13, hierarchical planning concept ipt wh ... 15
Figure 14, Relationship waiting time - utilization QED and M/M/C model ... 24
Figure 15, relationship waiting time - number of servers QED and M/M/c model ... 24
Figure 16, Reduce in waiting time when adding an extra server for analytical model and QED ... 24
Figure 17, Measured waiting time vs QED approximation ... 25
Figure 18, ratio of measured waiting time/expected waiting time based on QED ... 26
Figure 19, Cost and E(W) as function of β ... 27
Figure 20, Cost and E(W) as function of a ... 27

Figure 21, decision sequence ... 28
Figure 22, Use of historic information ... 29
Figure 23, planning steps ... 33
Figure 24, average throughput and waiting times in working days ... 37
Figure 25, Standard deviations of throughput and waiting times in working days ... 37
Figure 26, Average production time bulk production + bulk release in working days ... 38
Figure 27, stdev production time bulk production + bulk release in working days ... 38
Figure 28, Old vs New uncertainty handling ... 39
Figure 29. Production plans, old - new - demand ... 40
Figure 30, Comparison of fixed production quantities ... 41
Figure 31, Relationship between βopt and $\omega\beta$ ... 43
Figure 32, Cost function with $\omega\delta$= 0.1 ... 43
Figure 33, Cost function with $\omega\delta$= 0,5 ... 43
Figure 34, Cost function with $\omega\delta$ = 1 ... 44
Figure 35, Cost function for optimal region + and – 0.2 ... 44

LIST OF VARIABLES

Variable	Definition
$\boldsymbol{I}$	Set of controlled items
$\boldsymbol{\varepsilon}$	Set of PUs
$\boldsymbol{\tau_e}$	Planned lead time for PU e
T	Planning horizon
$\boldsymbol{E}$	Bills of materials
$\widehat{\boldsymbol{R}_{i,t}}$	Planned order release quantity time t in number of batches
$\boldsymbol{A_e}$	Set of lead time feasible order release schedules
$\boldsymbol{I^{+}_{i,t}}$	Planned surplus of item i at time t in batches
$\boldsymbol{I^{-}_{i,t}}$	Planned shortage of item i at time t in batches
$\widehat{\boldsymbol{C}}_{i,t}$	Planned order release of item i at time t in campaigns
$\boldsymbol{CS_i}$	Campaign size of product i
$\widehat{\boldsymbol{P}_{i,t}}$	Binary variable of production of item i at time t
$\boldsymbol{h_i}$	Holding cost for product i for time period
$\boldsymbol{b_i}$	Backorder cost for product i for time period t
$\boldsymbol{S_{i,p}}$	Startup cost production
$\boldsymbol{S_{i,c}}$	Startup cost campaign
t	Time
$\boldsymbol{TEF}$	Set of tender forecasts tef
$\boldsymbol{d_{tef;t}}$	Demand for tender forecast tef at time t
$\boldsymbol{p_{tef}} \in \{\boldsymbol{0,25; 0,5; 0,75; 1}\}$	Success chance for tender forecast tef
$\boldsymbol{D_{TEF;t}}$	Sum of $d_{tef;t}$
$\boldsymbol{PCCA_t}$	Planned capacity constraining activities at over t
$\widehat{\boldsymbol{D}}_{t}$	Demand during time t
$\widehat{\boldsymbol{D}}_{RF;t}$	Regular demand during t
$\widehat{\boldsymbol{C^{\tau}}}$	Net (cumulative) planned capacity over t

$\widehat{GC}_t$	Gross planned capacity available over t
$PCCA_t$	Planned capacity constraining activities at over t
$OA_{t,m}$	Operational availability of the machines at time t of machine m
$CR_{i,m}$	Capacity requirement of product I at machine m

QED variables	Definition
W_q	Waiting time in queue
μ	Production rate per server
λ	Arrival rate
c	Number of servers
$\rho = \frac{\lambda}{\mu c}$	Utilization of the system
$K(\beta)$	Cost function of Beta
ω	Cost of holding inventory for one time unit in QED model
δ	Cost of adding one server
β_{opt}	Optimal β that minimizes the cost

LIST OF ABBREVIATIONS

Abbreviation	Definition
POO	Pharmaceutical Operations Oss
IPT	Integral Production Team
WH	Women´s Health
QED	Quality and Efficiency driven regime
MILP	Mixed Integer Linear Program
SCOP	Supply Chain Operations Planning
ORDB	Order release decision bulk
PCCAs	Planned capacity constraining activities
ADI	Advanced demand information

1. INTRODUCTION; IPT WOMEN'S HEALTH

As this masters thesis was conducted in the context of hierarchical planning at MSD at Pharmaceutical Operations Oss (POO), it is necessary to introduce this company shortly. This chapter introduces both the company and the production process. It is followed by a quick overview of relevant literature in the field in chapter 2. Chapter 3 discusses the research contributions. After this the thesis continues with the design of the planning hierarchy in chapter 4. Chapter 5 focuses on the stochastic control models while chapter 6 introduces the deterministic control models. Chapters 7 and 8 are used to determine the consequences for IPT WH based on each of these models. After this, chapter 9 gives a further detailed explanation of one of the stochastic control models to make the findings more general. The thesis is concluded with chapters 10 and 11 which give a conclusion and possibilities for further research.

1.1 MSD PHARMACEUTICAL OPERATIONS OSS

With US$ 48 billion revenue and US$ 7,3 billion profit in 2012 MSD is one of the world largest pharmaceutical companies. Over time it has known both autonomous growth and growth through takeovers. One of these more recent takeovers was the takeover of Schering Plough, which at that point in time had just acquired Organon, a former Dutch pharmaceutical company. With this acquisition, MSD also acquired a large production facility in Oss, the Netherlands. Within this production facility, several product groups with similar production characteristics were chosen to form integral production teams. One of these production teams is IPT Women's Health, which produces solid medicines (tablets) for use by women. These medicines are mainly anti-conceptives, but also have other purposes.

Like most other pharmaceutical company, MSD is experiencing challenging times as patents are expiring, research productivity is declining and market buying power is increasing. This puts pressure on production facilities to reduce the costs, reduce the lead time and retain a high service level. Furthermore like all other producers, MSD tries to differentiate products based on looks and marketing, hence the number of product options is increasing.

Because of this, there is pressure on costs, lead time and service levels. Therefore the planning concept at the IPT WH is open for discussion.

1.2 IPT WOMEN'S HEALTH PRODUCTION

The production process of IPT Women's Health is depicted in Figure 4. The production process consists of four steps. It starts with the production of bulk tablets, this is a production step where batches of active granulate are pressed into tablets. These tablets are then sent to the bulk release, where they are subject to several tests to assure the quality of the product. If the quality requirements are met, the products are ready for packaging. If the quality requirements are not met, there are two options. Either extra tests are required with possible acceptance or rejection of the batch, or the batch is rejected right away. To start the packaging both tablets and packaging material have to be available. In the current situation, this is assured by keeping stock between bulk release and packaging. After the products are packaged, there is another quality check and there are outbound logistics to assure the right delivery with the client.

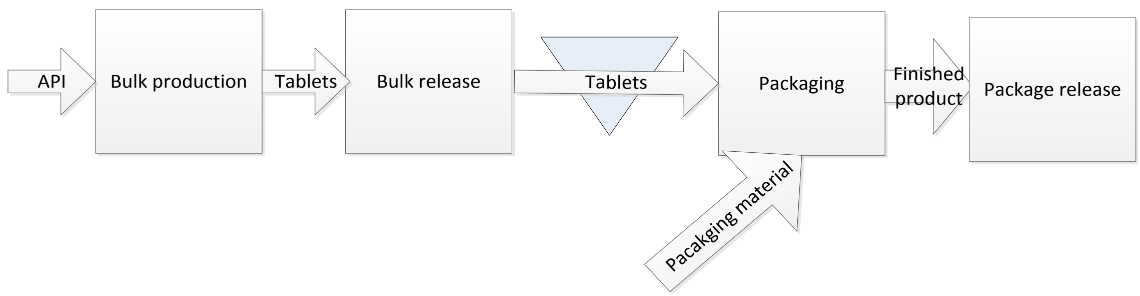

FIGURE 4, IPT WOMEN'S HEALTH PODUCTION PROCESS

1.2.1 BULK PRODUCTION

Bulk production consists of five consecutive steps, as is shown in Figure 5. The process starts with the weighing of ingredients; here all ingredients necessary for the production of batches are weighed. After the weighing of ingredients, the production process starts by production of basic granulate, followed by the adding of active material. Thereafter the activated granulate is pressed into tablets by presses, after which the batch either leaves the process or the tablets are provided with a coating. Throughout the process there are in process quality checks, which are reported in forms of deviations if there are problems in production. Furthermore the process has fairly lengthy sequence dependent setup times, as machines have to be thoroughly cleaned before new products can be produced.

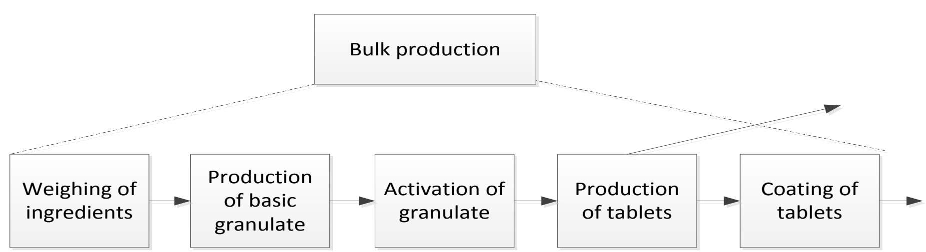

FIGURE 5, DETAILS OF BULK PRODUCTION

1.2.2 BULK RELEASE

Bulk release consists of two parallel steps followed by two serial steps, as is shown in Figure 6. This process starts with the lab and deviation analysis. The lab analysis consists of a series of tests to confirm the quality of the batch. The deviation analysis is only done when the in process checks resulted in a reported deviation. When this happens, an expert looks into the documentation and if necessary deviations. This expert then determines correctness of the paperwork and if necessary the impact of the deviation on the quality of the product. After both lab analysis and documentation analysis have been finalized, the quality release officer does an extra check on all paperwork is ok. If this is also true, there is a final check by a quality person, who then releases the batch for production. In case any of the analysis in either lab, deviation or in the paperwork shows anything extraordinary, further analysis can be done and in the end a batch might even be rejected. If that happens, the batch will also be scrapped.

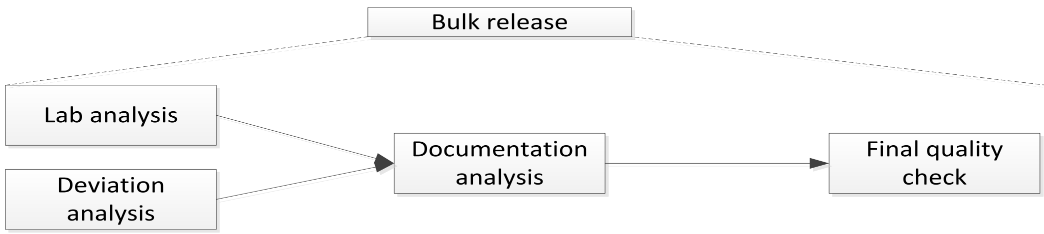

FIGURE 6, DETAILS IN BULK RELEASE

1.2.3 PACKAGING

Packaging consists of three serial steps, as is shown in Figure 7. The process starts with a blistering step, where the released bulk tablets are put into a blister. After this, products that are sensitive to moisture get sealed via a vacuum protecting sachet. Finally a cartoning line puts the blisters/sachets into a preprinted carton and adds the leaflet to this carton. Again, during this production process there are in process checks, to assure the quality of the different production steps, which are again documented and reported as deviations if necessary. Each of the steps in the process has the input of both the bulk tablets and a type of auxiliary material (foil, leaflet or carton).

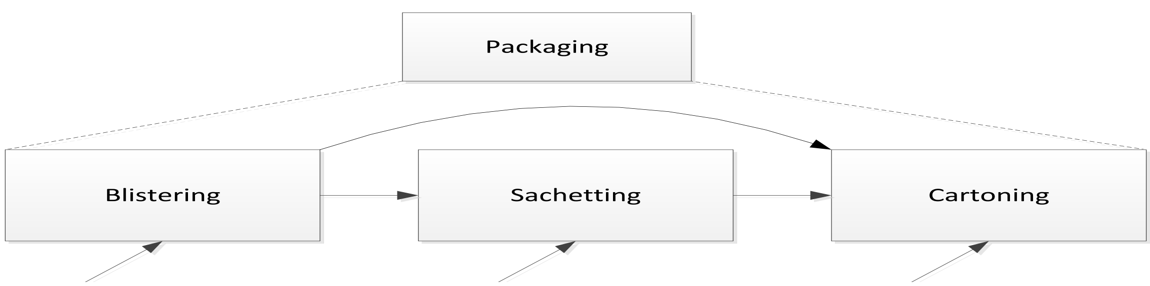

FIGURE 7, DETAILS OF PACKAGING

1.2.4 PACKAGE RELEASE

The final step in the production process is the package release. This step consists of the serial steps of a check on the final product and the outbound logistics. In the first step, possible deviations in production are checked on impact for the product. Furthermore all paperwork regarding both bulk batch and order is checked for completeness and correctness. Once this is finished, the products are released for transportation. This step of outbound logistics is done by another department. If the products are not released, rework in production can be required, or the production order is rejected as a total.

1.2.5 UNIT FLOW

Throughout the production process there are three different units of measurement that are of importance. First of all production starts with production of a campaign. A campaign consists of several batches of the same product. Each batch again exists out of a fixed number of tablets of the same product. Furthermore demand arrives in orders, which can be anything from a small part of a batch to a magnitude of batches. In bulk production, the unit of measurement that is of importance is the number of batches. For quality review this is the number of campaigns. After this, packaging starts production based on an allocation of orders per batch. Outbound logistics then operates on an order level, where each order has a set throughput time.

The BOM for the production process can be classified as divergent with product options. This implies that there are relatively few starting components which have a high number of end products. With MSD the factor of starting active materials versus end products is approximately 1:60.

1.2.6 PRODUCTION LEAD TIME

Figure 8 and Figure 9 show the development of throughput times for 2010, 2011 and 2012 for the different production units. As can be seen, the bulk release has the highest average and standard deviation in throughput time. This is followed by the bulk production, packaging and the package release.

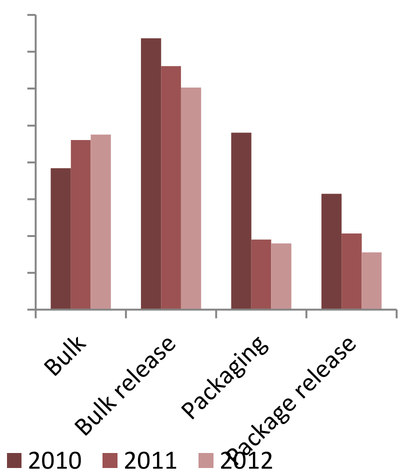

FIGURE 8, AVERAGE THROUGHPUT TIME DEVELOPMENT IN WORKING DAYS

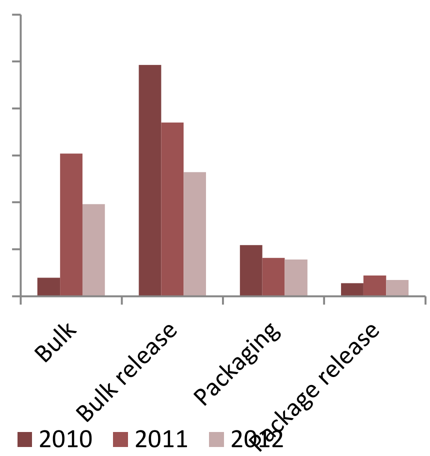

FIGURE 9, STANDARD DEVIATION OF THROUGHPUT TIME DEVELOPMENT IN WORKING DAYS

High throughput times can have several causes, the most important among them are uncertainty in available capacity, uncertainty in arrivals and uncertainty in production time per batch/order. All different types of uncertainty are present in the aforementioned process. Note that the total production process currently takes an average of ~80 days, or 16 working weeks of five days a week.

1.2.7 UTILIZATION OF BOTTLENECK RESOURCES

Each of the production units has a bottleneck resource that determines the pace of the production unit. As can be seen in Figure 10 the utilization for both the bulk production and the bulk release is very high, where it is relatively low for packaging. For the final quality check the utilization is unknown, as this is one individual person combining several jobs, where packaging release has priority if necessary, this is also irrelevant.

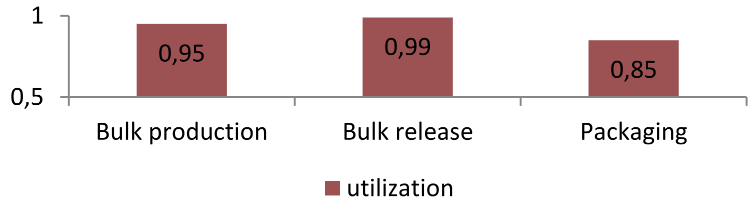

FIGURE 10, UTILIZATION OF BOTTLENECK RESOURCES PER PRODUCTION UNIT

The bottleneck for different steps is the granulate mixer and some of the presses in bulk production, the lab in bulk release and the blistering or cartoning step in packaging (depending on the production sequence).

1.2.8 ARRIVAL OF DEMAND

Demand for IPT WH can be split in two groups, regular demand and tender demand. Regular demand is the demand that comes from the different country organizations. Tender demand is demand that is normally sold to governments or for governments via third parties. Both are shown in Figure 11, which shows the arrival of the full demand.

Regular demand is always for internal parties and is generally make to stock demand. It is therefore flexible. Both timing and size of the demand are negotiable to some extent. Furthermore the level of the demand is relatively stable. Regular demand has a customer order lead time of 12 weeks.

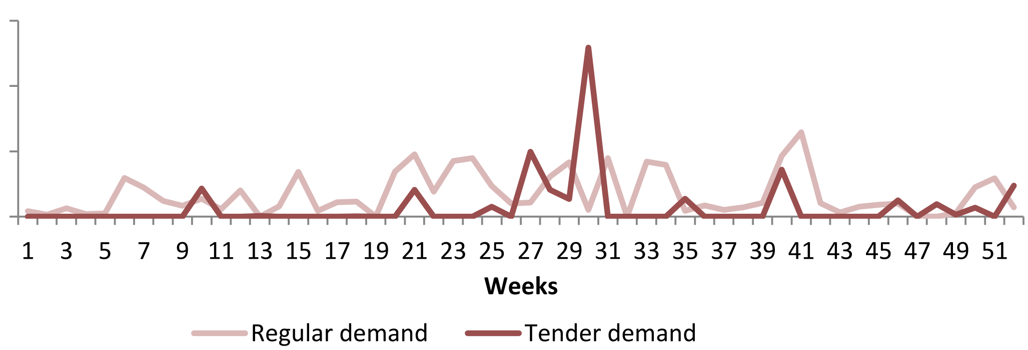

FIGURE 11, REGULAR AND TENDER DEMAND 2012

Tender demand on the other hand is always for external clients. These are typically large orders that are not flexible in size due to regulations. Furthermore one either gets awarded a full tender, or nothing at all. Hence there is no flexibility in the total size of the order. There is however flexibility in the delivery structure of the products. They can be delivered over longer periods of time. Furthermore tenders have a high monetary fee for late delivery. Finally, the global funding structure drives tender orders to arrive during the last two quarters of the year.

All demand is forecasted in some way. For the regular demand, forecasts are based on sales data and are available for the upcoming two years in monthly buckets. For forecasting the tender demand, all tenders that MSD applies for are recorded in terms of size, delivery structure and chance of success.

1.3 ORGANIZATION AND PLANNING STRUCTURE

The production of tablets crosses departmental boundaries twice. Both the bulk release and the packaging release are executed by a quality department that is strictly separated from the production department. Where the production IPT has resources that are virtually always exclusively used for the IPT, the quality department is in principal organized to handle the IPTs work, but also has a significant part of the load that is not related to the production process (~25 %). Examples of this are stability analysis, external lab demands or validation of lab equipment. In case of the stability analysis and external lab demands, these analyses can have preference over release analysis.

Looking at the current planning hierarchy, the decisions can be split into time dimensions with the long term (1-5 years) the mid-term (3 – 12 months) and the short term (0 – 3 month) as different components. The decisions for the long term regarding capacity increase in terms of new machinery or plants as well as decisions regarding the division of products over different IPTs are taken outside of the IPT.

For the mid-term, main decisions are capacity availability and stock levels. The scaling of capacity has an effectuation time of approximately 3 – 4 months. The decision to scale up or down is taken by a planner for the IPT in both bulk and packaging departments. The planner does this with the help of a deterministic model that accommodates capacity uncertainty for the bulk and packaging. Furthermore the planner takes a decision on the safety stock levels at three different echelons in the chain; the API level, the released bulk level and the finished product level. Currently only the API and released bulk hold safety stock. The level of this safety stock is determined based on a rule of thumb, where the expected throughput time of the quality department is the basis of the safety stock. Furthermore decisions are made for planned capacity constraining activities (PCCAs), based on the forecasted demand for a period.

The quality department determines capacity on a yearly basis, based on the total production that year. Furthermore the PCCAs are planned throughout the year, with little to no coordination with production between the quality and production department regarding the timing of these PCCAs.

For the short term production planning, a demand signal arrives in the form of a forecast or actual demand from several order hubs around the world. After this demand is received, a person within the production department determines the required production of both bulk products and packaged products. This production plan is based on mixed integer linear program, which takes into account the bulk production and packaging as full resources, including their capacity, and the bulk and package release as offsets in time for the production. The outcome of the system is a production plan determining the production based on biweekly production quantities. This is manually checked versus the production constraints of bulk production and packaging, hence there is no check with either of the quality departments. Afterwards some demand can be shifted or constraints can be changed, followed by another run of the integer programming, which gives the final production plan.

After this production plan, planners determine the sequencing and execute the order releases for the different products that have to be produced for the two week periods. This is done by different planners for the production of bulk and packaging. Furthermore the planner of the quality department concurs with the planner of packages on a weekly basis to prioritize the planned work.

1.3.1 THE OP TOOL

In the current situation the planning tool that is central to the planning operations is called the OP. This tool is an integer programming tool that determines the production planning for the upcoming two years based on forecast information. This forecast information is however limited to regular demand and is only updated once every month. Forecast regarding tender demand is only incorporated as a fixed capacity reservation per time unit. The OP operates with time intervals of two weeks and determines production amounts and inventory levels for both the bulk production and the packaging steps. It does so based on cost allocation for different points in time. The model is purely deterministic; it does however take into account a linearized capacity restriction where the expected net capacity is used instead of the gross capacity.

Based on this model decisions for the mid-term regarding capacity as well as decisions for the mid-term regarding inventory of bulk and packaging are taken. The safety stocks, demand and capacity are external for the model. Hence all three are deterministic, but can be determined in a different way.

1.4 CONCLUDING

This chapter gave an introduction to the IPT WH, the production process and the planning at the IPT. The production process is one with lots of uncertainty. This uncertainty can be seen in both demand uncertainty and throughput time uncertainty. Furthermore the resources are utilized to quite a high extend, which implies that waiting times play an important role in the process.

2. LITERATURE REVIEW ON HIERARCHICAL PLANNING AND UNCERTAINTY

This chapter gives a brief overview of the current positioning in literature regarding uncertainty in hierarchical planning concepts. It starts with an introduction of hierarchical production planning, followed by the introduction of both effectuation lead time and anticipation, an introduction to uncertainty in hierarchical planning and finalized by the gaps that are found in literature.

2.1 HIERARCHICAL PRODUCTION PLANNING

The concept of hierarchy in planning was introduced by Anthony (1965) and was later extended by Max & Heal (1973) into hierarchical production planning. It was found that the planning problem was too complex to solve from a holistic perspective. The problem was not computationally tractable. To simplify the decisions, the concept of hierarchy was introduced. The general idea consists of a sequence of decisions, where decisions higher in the hierarchy restrict the solution space for decisions lower in the hierarchy. Hence, restrictions in solution space can be split into hard restrictions that are imposed by the environment, and soft restrictions that are imposed by previous decisions in the decision hierarchy. The first dimension that was taken into account for the division into hierarchies was a time dimension, splitting decisions in long, mid and short term decisions. This framework was later refined into an approach that incorporated the organization of a company in the framework. This resulted in a hierarchy with aggregate capacity planning, a supply chain operations planning (SCOP) and different production units, where each production unit again can consist of a goods flow control and production steps. This framework is shown in Figure 12 (Bertrand, Wortmann, & Wijngaard, 1990; Kok & Fransoo, 2002; Jansen M. , 2011).

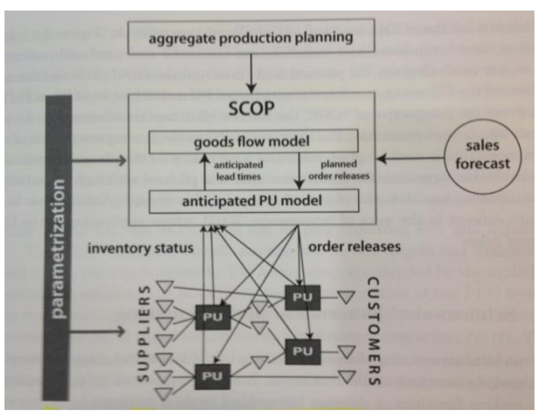

FIGURE 12, EINDHOVEN PRODUCTION FRAMEWORK (JANSEN M. , 2011)

2.2 EFFECTUATION LEAD TIME AND ANTICIPATION

The concepts of effectuation time and anticipation are an important addition the hierarchical planning. To fully understand these two concepts, it is important to make a distinction between a physical flow through a process and an information flow through a process. The physical flow represents goods that are produced and/or transported in the process, while the information flow represents the information that influences the physical flow or is received as a consequence of the physical flow. The effectuation lead time is the time it takes for a change in the information flow to be seen in the behavior of the physical flow. For the pharmaceutical industry, these effectuation lead times are typically long. Or to put it differently, the pharmaceutical industry's supply chain has slow responsiveness to changing circumstances (Kok & Fransoo, 2002; Altrichter & Caillet, 2005).

The concept of anticipation looks at the relation between different hierarchy levels. The concept is introduced by (Schneewiess, 2003) and makes a distinction between an upper level and a lower level. For each decision to be made, the upper level anticipates the behavior of the lower level, hence the decisions that the lower level makes. A distinction is made between four different types (Jansen M. , 2011; Schneewiess, 2003);

- explicit exact anticipation; all information at the lower level is available exactly to the higher decision level
- explicit approximate anticipation; all information is available, however some information is approximated
- implicit anticipation; only the relevant information is available
- no anticipation; there is effectively no information available

2.3 MODELING OF UNCERTAINTY IN HIERARCHICAL PLANNING

In a general modeling context uncertainty is referred to as variation or variability and a distinction can be made between explained variability and unexplained variability. In a production context the explained variation is the variation in a production process that is predicted by a model, where unexplained variation is not predicted by the model. This can either be because this variation is unknown to a modeler or because a modeler has chosen not to consider the variation in the specific model. As modeling is always an exercise to simplify reality, the explained variation in a model is almost always lower than the variation that in principle could be explained (Bertrand & Fransoo, 2002; Buzacott & Shantikumar, 1993).

The literature on hierarchical planning is mainly focuses on the different decisions that have to be taken in the planning concept. For each of the decisions there are models that have a stochastic nature or models that have a deterministic nature.

For any decision level, the criteria for incorporating any type of uncertainty into the model seems to be the existence of uncertainty at the aggregation level of information relevant for the decision. For long term decisions, like machine investments or fixed product allocation decisions, this implies that demand uncertainty and production lead time uncertainty are not taken into account. This is done because at the level of aggregation of information this uncertainty is negligible (Gatica, Papageorgiou, & Shah, 2003; Levis & Papageorgiou, 2004).

For the mid-term decisions, with the SCOP as the main decision, literature has shown that models that incorporate uncertainty in decisions for a supply chain planning outperform purely deterministic models. Models range from stochastic models that take into account the inventory holding decisions to, more recently, different types of lead time anticipation for the production units. The basis in both cases being a model based on fixed lead time between production units and stock at certain points in the supply chain (Kok & Fransoo, 2002; Jansen, Kok, & Fransoo, 2013).

For the very short term, models that are typically used are of a deterministic nature. The reasoning here being that the uncertainty is no longer present in the system, as for the short term all parameters are known.

2.4 GAPS IN LITERATURE

The identified gaps in literature can be split into two components. First, gaps regarding hierarchical concepts concerning uncertainty, and second gaps regarding the modeling for individual decisions. Regarding the hierarchical concepts there is little literature discussing the incorporation of uncertainty in hierarchical concepts at the different levels of uncertainty. Although the decomposition of the problem is very useful, the strong focus on the models that look at individual decisions fails to take into account the effects that previous decisions have had on the current one and the effect that a current decision has on lower hierarchy ones. Or stated differently, when looking at the soft constraints that are created at higher hierarchy levels, these are only looked from a point of view that does not take into account uncertainty. Although the anticipation of lower levels somewhat reduces this problem, there seems to be little literature regarding both the anticipation and other interaction effects that exist in the hierarchical planning. Furthermore, it is not clear from literature when what type of anticipation is required.

Regarding the individual decisions, there seem to be ways to improve the SCOP by incorporating different types of anticipation that are based on different queueing networks or queueing models. The planned lead time that is a part of the SCOP does not take into account the relationship between load on resources and throughput time of a resource. This could be covered by using queueing models as a ways to determine capacity or load restrictions. Development in queueing theory shows interesting progress regarding different types of queueing models. Although these models are still quite complex for application in an optimization context, it should be possible to take the general behavior that can be deducted from these models into account.

2.5 CONCLUDING

Looking at the research and the gaps, they are easily linked to the case study in three ways. The planning at IPT WH can be seen as a hierarchical planning with several existing layers. As mentioned there is a lot of uncertainty in the production process, but the current planning hierarchy only considers limited amounts of this uncertainty. Furthermore, there is limited to no realization of lower level interaction in terms of uncertainty buffering in the current situation. Finally, the current type of anticipation chosen for the different production units strongly differs per production unit. Where the production units that are a part of the production department have a more detailed anticipation in the SCOP model, the production units that are part of the quality department have a very shallow one. This seems to be done based solely on the existing management structure rather than a well-considered decision.

3. RESEARCH CONTRIBUTION

This thesis contributes to research in three different ways; a complete planning concept for the production in a highly uncertain situation, best practices regarding uncertainty when designing a hierarchical production concept, and the actual way of modeling of the uncertainty to support different decision functions.

3.1 THE CASE STUDY

The pressure on cost and lead time within MSD pushes IPT WH to reduce cost without losing reliability of service level performance.

When analyzing the cost structure of the IPT, one can see that the highest supply chain costs are caused by the stock held after bulk release, which can be either released or unreleased bulk tablets. These tablets have to be kept in conditioned rooms and represent relatively high value. Furthermore, holding these safety stocks represents a risk regarding the expiry date of the product. Cost for expired product are generally very high. Therefore finding a way to reduce this stock would be very beneficial for the costs in the IPT.

Besides pressure on costs, there is also pressure on customer order lead time to be reduced from twelve to six weeks by the end of 2013. Looking at the current process, this seems feasible, as the throughput time of the latter two stages combined is four weeks on average with a standard deviation of less than one week. This is achieved however, with the current process and slack.

Both of the aforementioned points lead to the assignment central to this case study.

CASE ASSIGNMENT:

DESIGN A PLANNING CONCEPT WITH WHICH MSD IPT WH CAN ACHIEVE A FLEXIBLE CUSTOMER LEAD TIME RANGING FROM 6-12 WEEKS WITH A CUSTOMER SERVICE LEVEL OF 98% FOR ALL PRODUCTS AT MINIMAL COSTS.

3.1.1 CASE SCOPE

The case study is conducted at the level of the production process. The scope of the project is therefore the production process within the IPT and the quality department. This implies that decisions that are taken outside the IPT WH, at a higher hierarchy level are considered outside of the scope of this research. This is done for two reasons.

First of all, the scope of the project should be restricted, because the time span of the project is limited. This case was done as a master's thesis and should therefore represent a workload of approximately six months.

Second, the choice is made to ease implementation. By assuring only decisions that are taken at the level of the IPT or the quality department, the complexity of implementation is greatly reduced, as there is autonomy with regard to these decisions.

This implies that the capacity decisions that are taken into account are restricted to some aggregate capacity, the decisions that are taken at the moment in the IPT. Hence all decisions regarding API availability, product allocation, and transportation are assumed to be out of scope. Decisions regarding workforce planning, maintenance planning, order release, order allocation, order sequence and internal stock level are in scope.

3.2 RESEARCH QUESTIONS

Looking at the assignment at hand, both the removal of the stock and the reduction of the lead time, have significant consequences for the operation. In essence, both operations remove slack from the operational process. In a case with less slack available in the process, it is more difficult to act on sudden events, which is bad for both schedule adherence and ultimately service level performance. Therefore designing aforementioned planning system implies making a decision about the uncertainty that arises with such a system.

Knowing both the gaps in literature regarding decomposition of hierarchical production planning and the assignment given in the case the first research question can be defined as follows.

RESEARCH QUESTION 1:

HOW CAN THE *DESIGN* OF A PRODUCTION PLANNING HIERARCHY ASSURE THE CORRECT HANDLING OF UNCERTAINTY

Handling uncertainty in a design can be achieved by implementing control structures or by introducing buffers in the system. To determine what type of uncertainty should be handled by the design, two types of uncertainty are recognized; uncertainty that is influenced by the decision and uncertainty that is not influenced by the decision.

For the uncertainty that can be influenced a control structure is applied. Depending on the impact of the uncertainty this is a stochastic or a deterministic control function. Impact has to be seen in terms of the optimal decision given the objective (cost or service performance). In case of high impact there is a stochastic function, while with lower impact there is deterministic control.

In case of uncertainty that cannot be influenced one again looks at the impact. This time the impact has to be seen solely in terms of the objective and the parameters of the system. If the impact is high the first step is to isolate the source of this uncertainty. The next step is to buffer solely for this source, hence create a design that has different rules for certain groups of products. In case of low impact it makes no sense to buffer, as this will only bring extra cost into the system.

	Possible to have influence on uncertainty	Not possible to have influence on uncertainty
High impact	Stochastic control function (modeling)	Isolate and buffer through smart design
Low impact	Deterministic control function (modeling)	No action

TABLE 2, DESIGN VS MODEL, WHERE UNCERTAINTY SHOULD BE HANDLED

This leads to the design rules as are shown in Table 2, which should ensure a system where all uncertainty that can be reduced in any possible way is modeled as control function, while all non-reducible uncertainty is buffered for by smart design.

After one has the design, including the way that different uncertainties in the system are covered, the next question that should be raised of how different decisions in the planning hierarchy should be modeled such that they incorporate the uncertainty sufficiently.

RESEARCH QUESTION 2:

HOW CAN *MODELING* IN A PRODUCTION PLANNING HIERARCHY ASSURE CORRECT HANDLING OF UNCERTAINTY

The main contribution in terms of modeling is made through a stochastic anticipation in the control function for the bulk order release. This control function is based on the work of Borst et al (2004) regarding a quality and efficiency driven (QED) control mechanism that exists for queueing systems.

For this model it is shown that for low holding cost versus capacity cost the optimal cost in the system are approximately linear in its load. This makes it possible to apply this control mechanism in a general mixed integer linear programming (MILP) such as a planning algorithm.

Furthermore, the algorithm shows that the cost for different waiting times in a system are practically the same given linear holding and server cost. Rewriting this regime one can determine a maximum load for a workstation or a minimum number of servers for the system, such that the costs are close to minimal and the maximum throughput time will not be exceeded.

Besides the QED a different decision support model is proposed to make a decision regarding the timing of PCCAs. This model takes into account uncertainty of demand to determine the optimal timing of these PCCAs.

3.3 RESEARCH METHODOLOGY

To answer both of the research questions a set of design rules, two stochastic and two deterministic control models were constructed. The design rules and the stochastic control models were tested in a monte carlo simulation to determine the effects of these models on the performance of the system in terms of cost and throughput times. The deterministic control models were analyzed for different scenarios to determine the effects of these scenarios on the production.

3.4 CONCLUDING

This chapter introduced the case assignment, research contribution and research methodology. The research contributions can be split in three parts, a planning concept for a highly uncertain production environment, best practices concerning design of planning concepts and ways of modeling in planning concepts.

The remainder of the thesis is built around these questions. First the new planning *design* is introduced. After this the different decision *models* are introduced. It starts with the models that are central in correctly handling the uncertainty, the stochastic control models. This is followed by the deterministic control models. Ultimately the simulation and scenarios for production were tested to determine the effect of the changes.

4. *DESIGN* OF THE PLANNING CONCEPT OF IPT WH

This chapter introduces the design and design decisions in the new planning concept of IPT WH. It handles the different types of uncertainty that can be faced in the production process and determines how these uncertainties should be incorporated in the hierarchical planning. The chapter starts with the planning concept as it is proposed, followed by the uncertainty that cannot be influenced, uncertainty that can be influenced, and ultimately gives an overview of the decision sequence and the decision support models for each of these decisions.

4.1 OVERALL PLANNING CONCEPT

Looking at the overall planning concept it is important to distinct the different decisions that have to be made in the hierarchy. First of all, a decision regarding the capacity in different production units has to be made. Second, decisions have to be made about the goods flow throughout the system and the short term capacity (overtime). Finally decisions have to be made regarding the sequencing and releases to production.

For each of these different decisions different information is required. This information changes in terms of level of aggregation and in level of uncertainty. Given that these decisions are done at different departments in the company, it is likely that not all relevant information is known at the same level of detail to all actors in the system. Therefore the decisions are organized based on the information relevant for the decision, the effectuation time of the decision and the actors that have to take the decision.

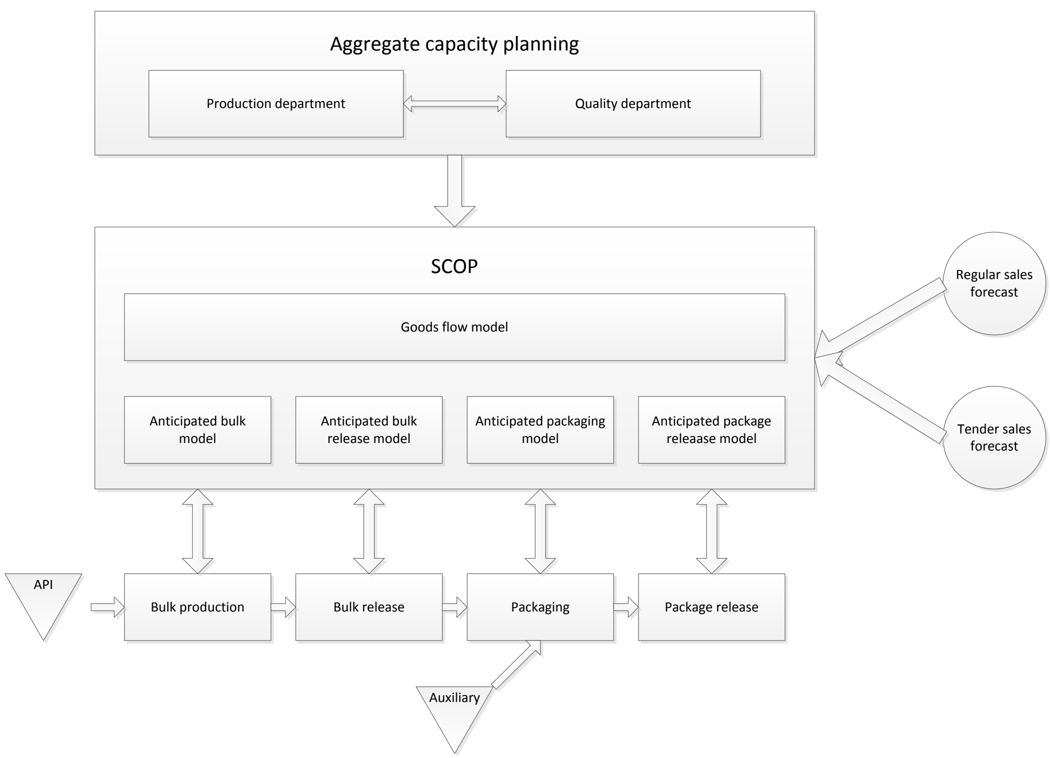

FIGURE 13, HIERARCHICAL PLANNING CONCEPT IPT WH

This results in the structure that is shown in Figure 14. As can be seen a distinction is made between an aggregate capacity planning that is made for both the quality and the production department. At this level information aggregated to the level of several months is sufficient. Furthermore, the effectuation time of this decision is quite long due to training of personnel. As the information is taken from and the decisions taken for two different departments, the collaboration is of importance. This decision is taken at a higher management level in the organization.

The next decision level is the SCOP level. At this level the decisions regarding the goods flow, overtime and releases are taken for the different production unit. As two of the production units are of a different department close collaboration is necessary to assure that all information is available. Putting all these decisions at one level assures an overview at this level. Furthermore having all these decisions at the same level increases the chance of the correct response given uncertain events. All medium to short term decisions regarding planning are taken at this level. Additionally, all decisions are taken by lower level of management.

4.2 UNCERTAINTY THAT CAN NOT BE INFLUENCED

Looking at the entire process there are four different types of uncertainty that are not influenced by the planning decisions. These are uncertainty in demand, uncertainty in production deviations, uncertainty in production time, and uncertainty in API delivery. In general, all of these uncertainties are covered to some extend by operating a system with rolling schedules and updated information. Because information is not always available in real timeand the rolling schedule decisions cannot be taken every instant it is necessary to give extra attention to each of these points.

4.2.1 UNCERTAINTY IN DEMAND

First of all there is the uncertainty in demand. No matter how one makes a planning, it does not change the uncertainty in demand. As the uncertainty in demand has a very high effect on the optimal production plan, it is important to isolate the uncertainty and buffer for the left over uncertainty in a way that is as cost efficient as possible

Analyzing the uncertainty in demand shows that this uncertainty comes mainly from tender orders. Therefore, a rule was deducted to create slack in the system for the production of tender orders based on this demand, which uses the fact that there are many quotations for tenders, which all have a different chance of success and a different delivery plan.

I. *The aggregate capacity should incorporate tender advanced demand information*

II. *The sum of the deliveries during the total throughput time of the process must not be greater than the expected demand in the tender*

III. *The forecast for the production decision should represent all tenders with a single delivery of with the expected demand as the size at the moment of the first delivery*

These three rules assure that the system will always produce sufficient slack to provide all short term tender demand with adequate capacity, even in the worst (or best) case scenario that all tenders are won. If this happens, there is still at least a full production cycle time to produce the remainder of the tender orders. This set of rules introduces sufficient buffers to assume deterministic demand.

4.2.2 UNCERTAINTY IN PRODUCTION DEVIATIONS

Another point where there is no influence of the planning decision on the uncertainty is the chance of having a deviation in production. Deviations arise with a certain chance, which is independent of the different planning decisions.

Looking at the impact of the uncertainty in deviations, it can be seen that the average throughput time for a production batch with deviations is much larger than a production batch without deviations. With changing lead times, the reliability and ultimately the performance of the production plan is reduced. Hence, the optimal plan changes strongly if this source of uncertainty is not isolated. This is covered in the design by a rule in production.

IV. *In case of a deviation during bulk production, one extra batch of that product will be produced in the first available campaign. From a planning perspective, this batch can skip the bulk release.*

Here the fact that deviations are known early in the process and the possibility to package unreleased bulk products is used to isolate the problematic uncertainty. Also realize that the extra batch is in fact a buffer, which could also be put in stock in case such a predictor would not be available.

4.2.3 UNCERTAINTY IN PRODUCTION TIME

Due to various reasons the production times of the different production steps are not deterministic. Although this might seem to be a consequence of the planning decision, it is not. This is a matter of definition, here production time is defined as the time a unit starts production at a server until the moment the production is finished, hence change over and waiting time are not incorporated. The uncertainty in this setting is not caused by the planning.

The impact of this uncertainty on the optimal decision, it is high. As there is a service criterion, delays are problematic for the production. Therefore flexibility or a buffer has to be added to the system. This is done in two ways

V. *Hold inventory at the level of bulk products only for low volume products*

VI. *Produce First Due First Serve in all flexible resources*

Both rules introduce great flexibility to the system. Rule IV earns flexibility at the cost of holding inventory for low volume products. As inventory levels scale with volume, this is a relatively cheap way to assure that products with delay can be given preference if necessary. Rule number V assures that the preference is actually given to the correct product. Together these two rules assure that the uncertainty in production time should not be problematic.

4.2.4 UNCERTAINTY IN API

The final type of uncertainty that cannot be influenced is the uncertainty in API. As the delivery of API is done by an external partner and is not always reliable, there is no influence on this decision. Therefore, one should buffer against this uncertainty.

VII. Hold API on stock for all products

It is easily seen that VII buffers against uncertainty of the API supply, in the most classical way possible, namely by holding stock at this level.

4.3 UNCERTAINTY THAT CAN BE INFLUENCED

The second type of uncertainty identified is the uncertainty that can be influenced by the decisions in the planning model. The only uncertainty that can be influenced by the planning decision is the throughput time uncertainty. At all different levels in the production sequence, the throughput time of different products depends on the net capacity available, the load on the system and in some cases the production sequence. All of these are decisions that are made by the planner.

Looking in terms of impact, the impact on the optimal decision given throughput time is strongest for the bottleneck resource in the process. This is true because of the strong relationship between waiting time and utilization of a resource. Hence, given the possibilities to influence uncertainty and the high impact for the bottleneck resource, this resource should be represented and controlled in a *stochastic* matter in the modeling.

Furthermore, the impact is strong when there are strong temporary capacity restrictions on a production unit, as this will make any production resource the bottleneck production unit during that period of time.

Looking at all other resources, assuring that the bottleneck will not be problematic assures that the load on these resources will be lower as well. Furthermore, the buffers that apply for production time also apply for general throughput time, assuring sufficient slack for a *deterministic* control of the throughput time.

4.4 DECISION SEQUENCE IN THE PLANNING HIERARCHY FOR IPT WH

Having determined the structure of the planning, the different rules in the planning hierarchy, the final part of the design is the decision sequence in the hierarchy. This decision sequence, including the type of decision support model that will be used for this decision is shown in Table 2.

#	Decision level	Decision	Model type
1	Aggregate capacity	Aggregate capacity scaling decision	MILP with ADI for tender demand
2	Aggregate capacity	Planned Capacity Constraining Activities (PCCAs)	Stochastic model
3	SCOP	Order release decision bulk (ORDB)	MILP
3.1		ORDB Production anticipation	Linear representation with capacity buffer
3.2		ORDB Bulk Release anticipation (bottleneck resource)	QED approximation (stochastic)
3.3		ORDB Packaging anticipation	Linear representation with capacity buffer
3.4		ORDB Packaging release anticipation	Offset in time
4	SCOP	Sequencing decision bulk production	Fixed order

5	SCOP	Sequencing decision bulk release	FDFS
6	SCOP	Allocation of sales orders to bulk batches	Greedy algorithm
7	SCOP	Order release packaging	Greedy algorithm

TABLE 3, DECISION STRUCTURE IPT WH

As can be seen, the model incorporates one stochastic decision support model for the PCCA (model 2) and a stochastic decision support model for the bottleneck resource (model 3.2). Furthermore there are two deterministic models (models 1 and 3). Hence it can be seen that with relatively few stochastic model the largest impact in the production process will be covered.

As the greedy algorithms have little added value from a research perspective, these two (6 and 7) will not receive further attention in this thesis. The focus will be on the two decisions at the level of aggregate capacity and the one decision regarding the order releases at the bulk level.

4.5 CONCLUDING

This chapter introduced the design decisions of the planning hierarchy in the IPT WH. The process of determining whether demand can be influenced has led to a set of VII production rules that isolate and buffer against all impactful uncertainties that cannot be influenced by the decisions made by the planner. Furthermore, for the demand that can be influenced a clear guideline has been determined which requires a stochastic representation for the bottleneck resource in the planning and a deterministic representation of the other resources in the planning. Finally the different decisions and their support models were determined.

5. STOCHASTIC MODELS IN DECISION SUPPORT

As stochastic decision support models are used for the places where the uncertainty has the highest impact they play a central role in the handling of uncertainty in the planning hierarchy. The two models of this type are the timing of PCCAs and the control of the bottleneck resource. This chapter introduces both models, tests whether the model is applicable for the given situation and prescribes how the models should be used in the planning context.

5.1 TIMING OF PLANNED CAPACITY CONSTRAINING ACTIVITIES

The first stochastic model is the model that is used for the determination of the timing of PCCAs. The second decision that has high impact is the timing of high capacity constraining activities. As these activities shift the bottleneck resource in the system, it is important to plan to ensure the timing of PCCAs is optimal and assure the impact of these PCCAs given the uncertainty is minimal. Hence, the decision should determine when a large capacity restriction has the least impact on the optimal performance.

5.1.1 MODEL REQUIREMENTS

Looking at the uncertainty the relevant uncertainty for this decision is the uncertainty in demand. As PCCAs reduce the capacity buffer that is instated for the demand, it is necessary to determine when this demand is placed.

Regarding the PCCAs there are a couple of things that are important to know.

1. PCCAs take a predefined amount of time for different resources, differing per PCCA and per resource
2. During PCCAs the resource is not available for other activities
3. There is a minimal amount of times the PCCAs should be done, but these can be planned throughout the year

5.1.2 PCCAS MODEL

To determine the ideal moment to have these PCCAs, a distinction is made between high uncertainty demand and low uncertainty demand. Recognize that for this case this is the distinction between tender demand and regular demand. This distinction also implies that forecasts for the high uncertainty demand will be less accurate than those for low uncertainty demand. Hence the forecast for each demand type is an indicator for the uncertainty in demand that is to be expected in a time period.

Based on advanced demand information (ADI) the uncertainty in the demand can be predetermined. This ADI takes into account the size of tender orders, the timing of the tender order and the success chance of the tender order. With this the uncertainty in demand and ultimately the uncertainty in capacity requirement can be determined.

Define:

Variable	Definition
$\widehat{D}_t$	Demand during time t
$\widehat{D}_{RF;t}$	Regular demand during t
$\widehat{GCA_t, m}$	Gross planned capacity available over t
$PCCA_{t,m}$	Planned capacity constraining activities at over t
TEF	Set of tender forecasts tef
$d_{tef;t}$	Demand for tender forecast tef at time t
$p_{tef} \in \{0,25; 0,5; 0,75; 1\}$	Success chance for tender forecast tef
$D_{TEF;t}$	Sum of $d_{tef;t}$
$CR_{t,m}$	Capacity required during time t at machine m

$$E(D_{TEF;t}) = \sum_{tef \in TEF} p_{tef} * cr_{tef;t}$$

$$Var(D_{TEF;t}) = \sum_{tef \in TEF} \left(d_{tef;t} * p_{tef}(1 - p_{tef})\right)$$

For the regular demand it can be assumed the forecasts are reliable and are therefore deterministic. Hence by having the expected and standard deviation in demand, it is easy to calculate the confidence interval covered during a period if that period would incorporate the capacity below. Approximating the tender demand distribution by a normal distribution gives:

$$CR_{t,m} \sim N\left(CR_{RF;t} + E(CR_{TEF;t,m}); var(CR_{TEF;t,m})\right)$$

The ideal moment for planning of the PCCAs is when $P(CR_{t,m} > NCA_t | NCA_t = GCA_t - PCCA_t)$ is as small as possible, assuring maximum capacity buffer is available when it is needed the most. Having a stochastic decision support model ensures that the capacity reduction arrives at the moment in time when it has the least possible impact.

5.2 STOCHASTIC BOTTLENECK PRODUCTION UNIT CONTROL

The second type of stochastic control in the planning hierarchy is the control of the bottleneck production unit. For IPT WH, the bottleneck resource in the system is the lab, a part of the bulk release production unit. Control over this production unit can be exerted through the load released into or the capacity available in the production. The control should strive to minimize the cost while assuring the delivery performance criteria are met.

5.2.1 MODEL REQUIREMENTS

Looking at the uncertainty that one wants to influence by the model, the most important behavior that one has to capture is the non-linear growth of waiting time when utilization of the production unit grows. As explained before, this is more important when there is high utilization in the production unit. Besides this there is also a cost aspect to waiting time in the lab, with inventory cost being high for this resource. Hence this behavior has to be captured via the anticipation that is present in the ORDB model.

Looking at the bottleneck resource in the quality release, this is the lab. There are a couple of important characteristics of this lab.

1. The lab operates with analysts where each analyst handles one job at a time; hence it can be seen as a multiserver parallel production system
2. Due to vacations, PCCAs and load balancing with different IPTs, the number of analysts available for the job is unstable over time, it is however relatively plannable; hence there is some control opportunity for the capacity, in which case the cost per analyst can be seen as linear
3. There are several streams of work for the lab, including rework on batches with quality problems, external analysis requests and validation projects. All of these have uncertain arrivals and uncertain production times at the resource. Priority in this context is not given to bulk releases; hence there is little control over use of buffers in place
4. Like any other production unit, there is a fixed planned lead time

Based on these four characteristics, the model that should be applied has to take the waiting time into account given the capacity. Furthermore it should be able to translate a certain capacity to a maximum load, or the other way around a certain load to a required capacity. This also implies that it should take into account the scaling behavior of multiserver parallel networks. Finally, the model should be able to be used in a context of control where cost exist for servers and work in process.

5.2.2 THE QUALITY AND EFFICIENCY DRIVEN REGIME

A model that can do this is used in the telecommunication for the scaling of networks. This is an approximation method developed in 1981 (Halfin & Whitt, 1981) which approximates the M/M/C waiting chance and expected waiting time, based on an amount of servers. Hence it looks at a system with exponential inter arrival times, exponential processing times and C servers. This model therefore covers uncertainty in arrivals at the resource and uncertainty in production times at the resource. This regime is later referred as the quality and efficiency driven (QED) regime and is used in telephone networks for the approximation of several probabilities. This regime is used in an optimization context to determine the number of servers necessary in a call station, or to determine the chance of delay given a completely stochastic system with higher number of servers, places that typically have large amounts of servers (Janssen & Van Leeuwaarden, 2011).

Using this method therefore implies doing two checks.

1. Is the QED an accurate approximation of the M/M/C for low amount of servers
2. Is the labs waiting time behavior approximated by an M/M/C model

If 1 and 2 are both true, this implies that the QED is a good approximation of the labs waiting time behavior as well. If this is the case the an optimization method for linear costs for waiting time and fixed cost per server can be utilized.

5.2.3 COMPARISON OF QED WITH M/M/C FOR LOW NUMBER OF SERVERS

As mentioned, the first comparison is that between the M/M/C and the QED model for low number of servers. To test this the approximation is checked versus an analytical M/M/8 model to see if the approximation also works for lower numbers of servers. This is done by comparing the expected waiting time via both methods. Note that for the analytical waiting time the discrete version of the formulas for waiting time were used. Later on in this thesis the continuous variant of these formulas are used.

Variable	Definition
W_q	Waiting time in queue
μ	Production rate per server
λ	Arrival rate
c	Number of servers
$\rho = \frac{\lambda}{\mu c}$	Utilization of the system

$$P(W_q > 0) \approx \alpha(\beta)$$

$$\alpha(\beta) = \frac{1}{1 + \frac{\beta}{h(-\beta)}} \tag{2.1}$$

$$h(-\beta) = \frac{\varphi(-\beta)}{\phi(-\beta)}$$

$$E(W_q | W_q > 0) \approx \frac{1}{\sqrt{c}} \frac{1}{\mu\beta} \tag{2.2}$$

$$c = \frac{\lambda}{\mu} + \beta \sqrt{\frac{\lambda}{\mu}} \tag{2.3}$$

$$E(W_{q;analytical}) = \Pi_W \frac{1}{1 - \frac{\lambda}{c\mu}} \frac{1}{c\mu} \tag{2.4a}$$

$$\Pi_W = \frac{\left(c\frac{\lambda}{c\mu}\right)^c}{c!} \left(\left(1 - \frac{\lambda}{c\mu}\right) \sum_{n=0}^{c-1} \frac{\left(c\frac{\lambda}{c\mu}\right)^n}{n!} + \frac{\left(c\frac{\lambda}{c\mu}\right)^c}{c!} \right)^{-1} \tag{2.5}$$

Figure 16 and Figure 17 graphically show the relationship between the QED and the analytical model for the region of interest. As can be seen in Figure 16, the scaling of the waiting time as a consequence of the utilization is correctly captured in by the QED regime.

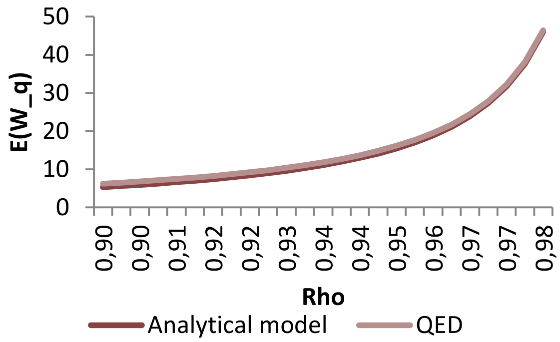

FIGURE 14, RELATIONSHIP WAITING TIME - UTILIZATION QED AND M/M/C MODEL

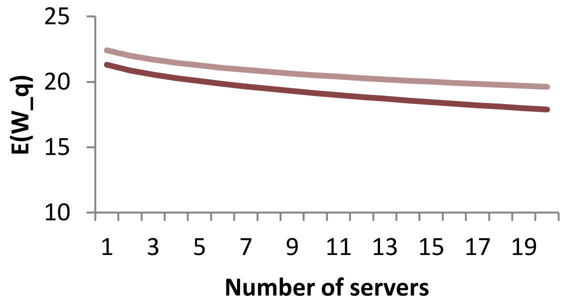

FIGURE 15, RELATIONSHIP WAITING TIME - NUMBER OF SERVERS QED AND M/M/C MODEL

Looking at Figure 17, it can be seen that the QED overestimates the waiting time of the analytical model slightly. This happens because both the chance of waiting and the conditional waiting time ($P(W_q > 0)$ and $E(W_q|W_q > 0)$) are upper bounds that converge to the actual number when n → ∞.

To isolate the scaling effects (the behavior that is relevant for the anticipation), one looks at the relation $\frac{E(W_q|c=n;\rho=x)}{E(W_q|c=n+1;\rho=x)}$. This relation is shown in Figure 18. As can be seen, the QED now underestimates the reduction in waiting time and underestimates the effects of scaling. The size of the underestimation is limited (0.5%), and will turn out to be negligible in comparison to further assumptions.

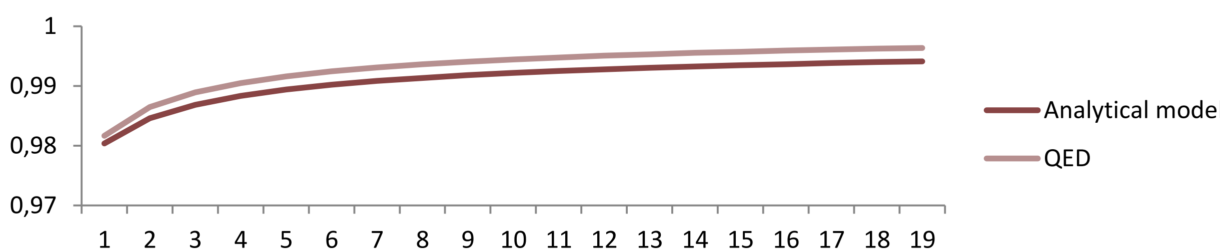

FIGURE 16, REDUCE IN WAITING TIME WHEN ADDING AN EXTRA SERVER FOR ANALYTICAL MODEL AND QED

Hence both the non-linear relationship between throughput time and utilization and the non-linear decrease in effectiveness of adding one extra server is represented correctly for small numbers of servers. From a theoretical perspective, the approximation has sufficient accuracy for the approximation of an M/M/c queue and captures all the behavior that is necessary to make the right decision regarding the uncertainty that one wants to control.

5.2.4 MODEL FIT ON REAL DATA

Knowing that the model is a good theoretical representation of the M/M/c behavior, the question becomes whether the behavior of the production unit is captured by the behavior of the M/M/C. Again, when talking about behavior, we are interested in the non-linear relationships between waiting time, utilization and number of servers available. To determine this fit a comparison is made between the average waiting time given the arrival rate, production speed and capacity over periods of three months. This model is than compared to the $E(W_q)$ and the continuous version of the $E(W_{q;analytical})$.

For the QED approximation rewriting 2.3 obtains the β for the period.

$$c = \frac{\lambda}{\mu} + \beta\sqrt{\frac{\lambda}{\mu}} \quad \rightarrow \beta = \frac{c - \frac{\lambda}{\mu}}{\sqrt{\frac{\lambda}{\mu}}}$$

Based on this beta we determine the expected waiting time per period by combining equation 2.1 and 2.2 to obtain.

$$E(W_q) = E(W_q | W_q > 0)P(W_q > 0) \approx \frac{1}{\sqrt{c}} \frac{1}{\mu\beta} \frac{1}{1 + \frac{\beta}{h(-\beta)}}$$

Figure 19 shows the measured, the QED approximation, the analytical M/M/C value of the waiting times and the load on the system. As can be seen the QED approximation overestimates the average waiting time for the lab. It does however represent the behavior of the waiting time at the lab. As can be seen the QED and M/M/C are practically identical, supporting the previous section. As can be seen there seems to be a remarkable good fit to the expected waiting time behavior, even though the rho is relatively instable over time. Based on the spread of waiting times and rho over time it is clear that there is currently no control function that correctly captures the behavior in this system.

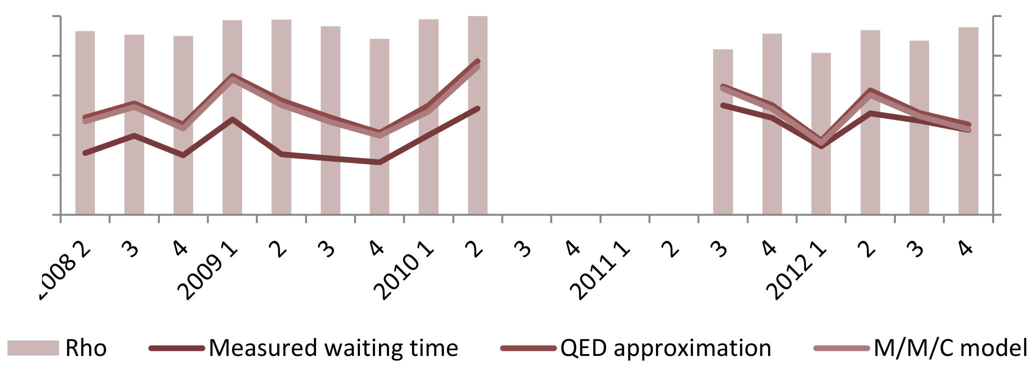

FIGURE 17, MEASURED WAITING TIME VS QED APPROXIMATION

Figure 20 compares the ratio of measured waiting time and QED approximation. It can be seen that this is relatively stable over time. There is however a strong shift in the data gap. This is caused by a restructuring of both lab and bulk production. This restructuring caused the variability in arrivals for the lab to increase strongly. Furthermore, the lab analysis was structured for smaller workloads, reducing the pooling effects that were present in the lab. Therefore the QED becomes more accurate after 2011, with an accuracy of approximately 0.88.

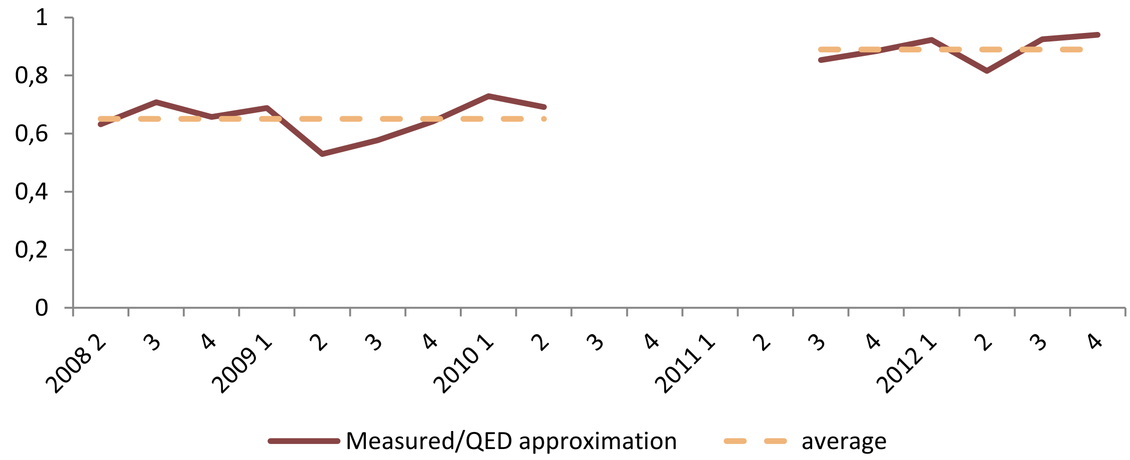

Figure 18, ratio of measured waiting time/expected waiting time based on QED

5.2.5 CONTROL THROUGH THE QED

Having determined that the QED approximation shows an accurate prediction of the waiting time behavior in the lab resource, the question becomes how this can be changed into a control function for the resource. Looking at the resource, the two things that are controlled by the planner are the load on the system and the capacity in the system. The aim of the control function is to minimize the cost given a certain load (1), while assuring the service performance by not exceeding the planned lead time (2).

The cost function of the lab resource can be seen as one with fixed cost per capacity server added and fixed cost (holding cost) per unit time spent in the system. As has been shown by Borst (2004) and later by Janssen and Van Leeuwaarden (2011) the optimal number of servers based on β can be determined. Define the cost function as is shown in (2.6). With fixed cost per server and linear cost over time for waiting.

Variable	Definition
$K(\beta)$	Cost function of Beta
ω	Cost of holding inventory for one time unit in QED model
δ	Cost of adding one server
β_{opt}	Optimal β that minimizes the cost

$$K(\beta) = \delta a + \sqrt{a}\left(\delta\beta + \frac{\omega}{\beta}C_*(\beta)\right) \tag{2.6}$$

$$a = \frac{\lambda}{\mu}$$

$$C_*(\beta) = \frac{1}{1 + \beta h(\beta)} \tag{2.7}$$

Based on the two conditions one is interested in the cost in the system and the waiting time in the system. Figure 21 depicts the costs and waiting time as a function of β for the IPT WH. Besides this β_{opt} and the historical range of β of IPT WH are displayed. Recall that the E(W(β)) is an overestimation of the actual waiting time.

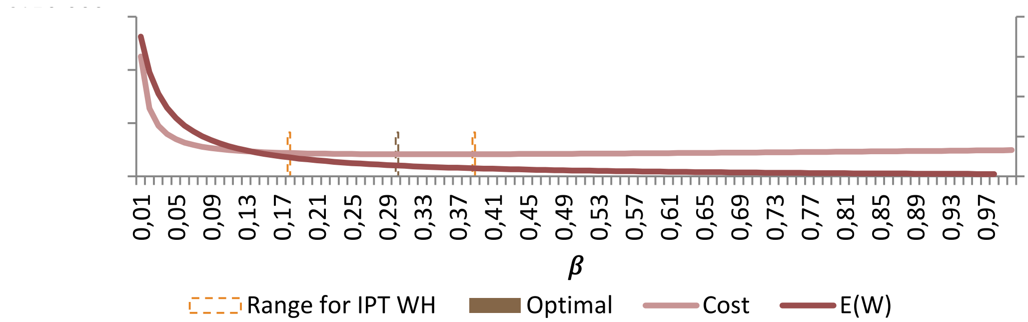

FIGURE 19, COST AND E(W) AS FUNCTION OF β

Looking at this figure there are a couple of interesting findings. First of all, concerning at the β range in the cost function, the IPT manages to operate a cost level that is close to the optimal cost. This implies that the current lab mechanism manages the cost aspect of the target well. Looking at the range of $E(W_q)$, it can be seen that this is way bigger. This is not surprising, given that there is currently no management control structure for throughput times. Even though this is the uncertainty that one wants to capture by this control function.

Figure 22 shows the relationship between a and $E(W_q)$ given β_{opt}. As can be seen, the optimal costs seem to be linear in a with $\frac{\delta}{\mu}$, while the $E(W)$ is decreasing in a. This linearity can be seen as a surprising result and will be explained in more depth in chapter 8. Having this linear cost result makes it possible to apply the stochastic model in combination with a linear optimization model, the type of model that is often used in planning control. In which case the choice for $\frac{\delta}{\mu}$ as the cost factor for the load for this production unit is a logical one. will ensure the minimal total cost in this production unit. Hence the $S_{i,c} = \frac{\delta}{\mu}$ for all i.

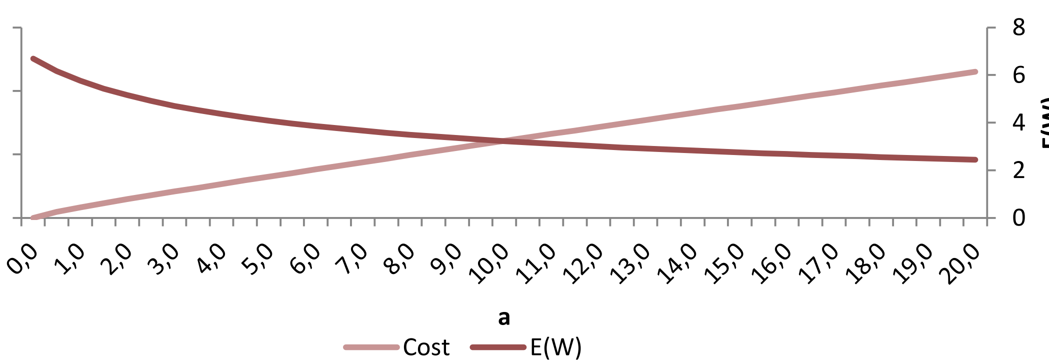

FIGURE 20, COST AND E(W) AS FUNCTION OF A

Knowing that the cost curve is flat and the costs are linear in a, the minimization of costs can be done in a broader context by a linear optimization model. This leaves the control of the throughput time for the production unit. Recall that the QED overestimates the $E(W_q)$ by $\frac{1}{0{,}88}$ the target $E(TW_q) = 0.88E(W_q)$. Realize furthermore that knowing β_{opt} and a means knowing c through $c = \frac{\lambda}{\mu} + \beta_{opt}\sqrt{\frac{\lambda}{\mu}}$. Figure 23 shows the decision sequence for the limitations on the MILP.

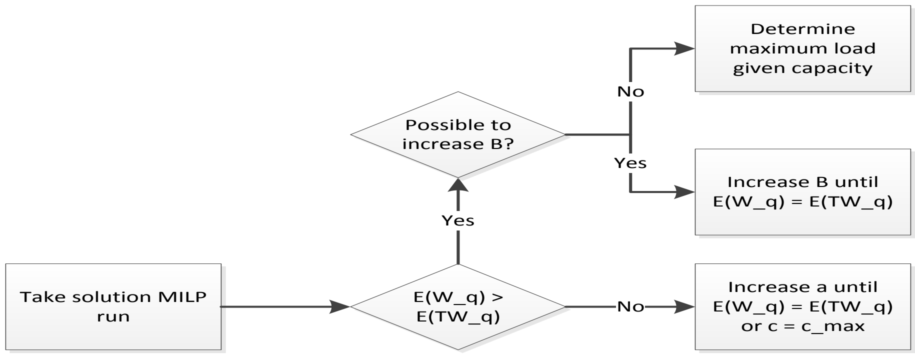

FIGURE 21, DECISION SEQUENCE

To determine the maximum load acceptable

1. Use $E(TW_q) \approx \frac{1}{\sqrt{c}}\frac{1}{\mu\beta}\frac{1}{1+\frac{\beta}{h(\beta)}}$ to numerically determine β given c[1]
2. Rewrite $c = \frac{\lambda}{\mu} + \beta\sqrt{\frac{\lambda}{\mu}}$ as $\lambda^2 + \beta^2\mu\lambda - c^2\mu^2 = 0$ such that $\lambda_{max} = \frac{-\beta^2\mu+\sqrt{(\beta^2\mu)^2+4c^2\mu^2}}{2}$

The resulting arrival rate multiplied by the time bucket is the maximum input rate for the Bulk production quantity decision, taking into account the non-linear relationship. Calculating this for each period of time gives a restriction to the load, such that the required waiting time is not exceeded.

5.3 CONCLUDING

This chapter introduces both the stochastic decision support models that are used in the planning hierarchy. These two relatively simple models assure the correct control decision such that the impact of the uncertainty in the system is minimized or kept at an acceptable level. The first model for PCCAs does so by determining the moment in time when the impact of the capacity reduction is lowest, while the second model does so by controlling the load or the capacity in the bottleneck resource in the system.

[1] Note that this is not β_{opt}, here the flat cost curve is applied to assure approximately minimal cost

6. DETERMINISTIC DECISION SUPPORT

Besides the stochastic models, there are two important deterministic decision support models. These are the models that take a decision for the aggregate capacity, and the order releases for the bulk production. For these models, there are still uncertainties in the system. This chapter introduces both models, explains how in each of the models the uncertainty is incorporated, and why this has little impact on the optimal decision. Both of these models should support control of either capacity or load that is released into the system in such a way that costs are minimal while performance is assured.

6.1 AGGREGATE CAPACITY SCALING

The first of these decisions is the aggregate capacity decision. This model should support in making a decision whether to scale up or down the workforce. The decision is taken for long periods of time due to the long effectuation time. For this decision there are two uncertainty types of importance. First of all there is uncertainty concerning demand, second there is uncertainty concerning the availability of planned capacity.

6.1.1 UNCERTAINTY IN DEMAND

To make the decision whether to expand capacity the current tool is used to determine the costs and is determined with the different available scenarios. Recall that this tool only has a limited representation of tender demand. This should be adjusted by adding the tender demand to the model by ways of ADI such as defined in the previous section regarding the size of this tender demand.

As at the moment ADI as such is not accurate yet a first approximation based on historical data is already sufficient to show that the current handling of this capacity is open for improvement. Figure 15 shows the demand pattern of tender demand with the current method of equal demand per quarter and using the demand pattern of 2011 on the expected demand for tenders in 2012 versus the actual demand in 2012. This demand will translate to capacity restrictions in the OP, which can help the planner to better determine the ideal moment for scaling up demand.

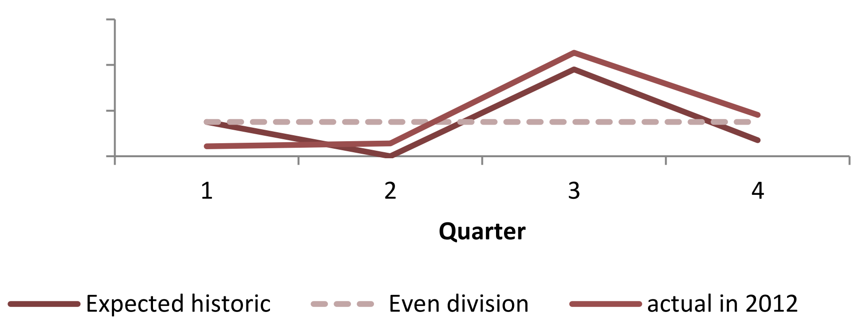

FIGURE 22, USE OF HISTORIC INFORMATION

In a future state the system should incorporate the tender demand by using.

$$E(D_{TEF;t}) = \sum_{tef \in TEF} p_{tef} * cr_{tef;t}$$

Hence, now the quotations are used as a predictor of the demand by taking the number of quotations and their expected size to determine the expected tender sales. By applying the rules regarding tender delivery structure and tender demand that were set in the design, sufficient slack should be in the system to consider the demand deterministic.

6.1.2 UNCERTAINTY IN PLANNED CAPACITY AVAILABLE

The second type of uncertainty is uncertainty in the availability of planned capacity. It is beforehand not known when machines will be available. To incorporate this the capacity should be translated from the planned available to the expected available capacity in the model.

Variable	Definition
$\boldsymbol{OA_{t,m}}$	Operational availability of the machines at time t of machine m
$\widehat{\boldsymbol{GC}}_{t,m}$	Gross capacity at time t of machine me

$$\widehat{GCA}_{t,m} = \boldsymbol{OA_{t,m}}\widehat{GC}_{t,m}$$

Here the $\boldsymbol{OA_{t,m}}$ is the expected availability of the resource. By having the flexibility and options in the design it is sufficient to only use the expected available capacity instead of incorporating a buffer through this variable.

6.2 ORDER RELEASE DECISION BULK

The second deterministic control model is the model that makes a decision of the order releases at the bulk. This decision should incorporate the stochastic control that was determined for the bottleneck production unit. The decision at hand is a decision regarding the load on the system, given the capacity in the system. Hence, the decision support model should determine the releases into bulk production such that (1) the costs are minimal and (2) the performance rate of the system is at least 98 %.

To achieve this an mixed integer linear program (MILP) is applied which determines the production decision based on fixed time intervals and with a planned lead time per production unit. There will however be some stochastically determined constraints based on the QED in this MILP. Please recall that the OP tool (the current planning tool) is also an MILP that makes the same decision.

In this context there are two different sets of uncertainty. First of all, the demand is uncertain, in a similar way to the previous decision. Second, the throughput times for the different production units are uncertain and finally the available capacity in the different production units is uncertain.

6.2.1 MIXED INTEGER LINEAR PROGRAM

By making a decision about the order releases in bulk, a decision is effectively taken about the different startup, inventory holding and backorder costs in the system. Therefore, if the objective is to minimize the costs, these three types of costs should be represented in the objective function.

Furthermore it is known that each of the production units has a planned lead time. After this time a release becomes available for a next production unit. Additionally, there is a maximum number of batches that can be produced in a campaign. Define:

Variable	Definition
I	Set of controlled items
ε	Set of PUs
τ_e	Planned lead time for PU e
T	Planning horizon
E	Bills of materials
$\widehat{R_{\iota,t}}$	Planned order release quantity time t in number of batches
A_e	Set of lead time feasible order release schedules
$I^+_{i,t}$	Planned surplus of item i at time t in batches
$I^-_{i,t}$	Planned shortage of item i at time t in batches
$\widehat{C}_{i,t}$	Planned order release of item i at time t in campaigns
CS_i	Campaign size of product i
$\widehat{P_{\iota,t}}$	Binary variable of production of item i at time t
h_i	Holding cost for product i for time period
b_i	Backorder cost for product i for time period t
$S_{i,p}$	Startup cost production
t	Time
$\widehat{PC}_{t,m}$	Planned (cumulative) capacity (quantity that can be processed) over the lead time of machine m
$OA_{t,m}$	Operational availability of the machines at time t of machine m
$CR_{i,m}$	Capacity requirement of product I at machine m

$$\min \sum_{t=1}^{T} \sum_{i\ \in I} \left(h_i I^+_{i,t} + b_i I^-_{i,t} + \widehat{P_{\iota,t}} S_{i,p} + \widehat{C_{\iota,t}} \frac{\delta}{\mu} \right) \tag{1.1}$$

S.T.

$$\hat{I}_{i,t+1} = \hat{I}_{i,t} + \hat{R}_{i,t-\vec{\tau}} - E\hat{R}_{i,t} - \widehat{D}_{i.t} \tag{1.2}$$

$$\hat{I}_{i,t+1}{}^{-} - \hat{I}^-_{i,t} \leq \widehat{D}_{i,t} \tag{1.3}$$

$$(\hat{P}_{i,t} - 1) \ \leq \hat{C}_{i,t} \leq \hat{P}_{i,t} \tag{1.4}$$

$$(\hat{C}_{\iota,t} - \widehat{1)CS}_\iota \leq \ \hat{R}_{i,t} \leq \ \hat{C}_{i,t} CS_i \tag{1.5}$$

$$\hat{C}_{i,t} \in N^0, \qquad \hat{R}_{i,t} \in N^0, \qquad \widehat{P}_{i,t} \in N^0 \tag{1.6}$$

$$R \in A_e \quad (1.7)$$

$$t = 0,1, \dots, T-1 \quad (1.8)$$

$$e \in \varepsilon \quad (1.9)$$

$$\sum_{\iota \in I} \hat{R}_{\iota,t} \widehat{CR_{\iota,m}} \leq PC_{t,m} OA_{t,m} \quad (1.10)$$

$$C_{i,t-\tau_1} \leq \lambda_{max,t} \quad (1.11)$$

TABLE 4, BULK RELEASE MODEL

Recognize the two different cost types of the objective function, the inventory costs $h_i I_{i,t}^{+} + b_i I_{i,t}^{-}$ and the startup costs $\widehat{P_{\iota,t}} S_{i,p} + \widehat{C_{\iota,t}} \frac{\delta}{\mu}$. Recognize furthermore the two elements that are a part of the objective function (1.1) and the set of restrictions (1.4), (1.5) and (1.6) that ensure the correct relationship between batches and campaigns. Also note that restriction (1.3) assures that items that do not have external demand are not allowed to have negative stock.

Furthermore, note that the decision is restricted to the production of batches, largely reducing the amount of variables and reducing the computational complexity.

6.2.2 UNCERTAINTY IN DEMAND

Similar to the decision regarding capacity the demand in the system can again be determined via the ADI and the forecast for regular demand. Recall that $\widehat{D}_{i,t} = E\left(D_{TEF,i,t}\right) + D_{RF,i,t}$, assuming full delivery at the initial tender for uncertain demand. This assures that at any moment there is at least sufficient bulk released to cover the first delivery. While in case of all tenders sold, there is no uncertainty anymore regarding the size and timing of these orders and the design assures sufficient time to react to these orders.

6.2.3 UNCERTAINTY IN THROUGHPUT TIME

The uncertainty in throughput time has to be covered through anticipation of the production units in the MILP. Recognize that the τ_e prescribes a fixed throughput time per production unit. To assure the actual throughput times do not exceed the prescribed throughput times the MILP controls the load on each of the production units through constraints (1.10) and (1.11).

As can be seen (1.11) incorporates the maximal load that is allowed through the stochastic control that was introduced in section 5.2.

For the other production units, the load in the system is controlled through (1.10). As can be seen the model again uses an approximation for the for the expected net capacity available, similar to the aggregate capacity decision.

Note furthermore that this restriction relies on a deterministic model of the whole production process, that makes an assumption for available capacity on the expected capacity available. Hence, there is a buffer against the uncertainty that cannot be influenced.

Taking this deterministic view does not take into account the relationship between the load on the resource and the throughput time that is to be expected. This is in accordance with the earlier defined anticipation for throughput time in this production unit and is acceptable for three main reasons.

First of all the production unit does not incorporate the bottleneck resource. This implies that the production unit will be operating under relatively low utilization in comparison to other production units. Therefore, the effects of not incorporating this uncertainty are limited.

Second, the production unit is under direct control of the planner, implying that all future decisions can be taken to both minimize uncertainty and assure successful use of the buffers that are in place in the deterministic model. Hence the impact of not taking the uncertainty into account is minimized.

Third, the main problem with uncertainty is uncertainty regarding the throughput time. When the product is late out of planning, this implies that there is higher variation in arrivals in the next production unit. Looking at the next production unit, this has explicitly modeled high uncertainty in arrivals through the QED control. Therefore increased variation that is a consequence of the first production step, will be handled by the next production step and should not be incorporated in the modeling.

6.2.4 USE OF THE ORDB

Actually using the ORDB to its fullest is somewhat more complex than simply running the model. The planning of the Bulk production quantity has to be done in several steps. Like any MILP, solving the MILP implies not knowing what decisions the model made. As decisions regarding capacity are not made at this decision level, short term capacity changes could be possible, while only running the MILP would not tolerate that. The same is true for possible shifting in demand timing or demand size. Therefore, planning consists of four steps, which are shown in Figure 24.

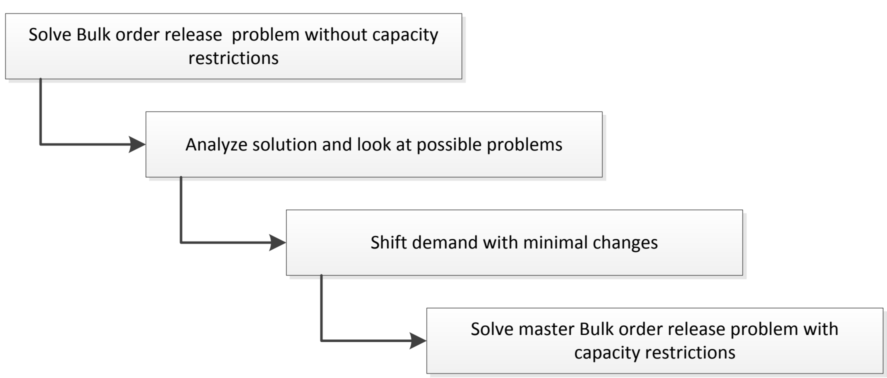

FIGURE 23, PLANNING STEPS

As can be seen, the first solution is one that does not take into account capacity restrictions. This solution basically determines a capacity requirement for the model. Having this capacity requirement makes it the planner´s job to then determine the feasibility of this requirement, by comparison to actual capacity or QED determined capacity.

If it is not possible to fulfill the capacity requirements, it becomes necessary to check whether there is flexibility in the delivery of demand. If this is true, the demand should be shifted such that the costs can be minimized as much as possible. In case this happens one or a combination of the actions in Table 8 have to be executed.

Situation	Action
Bulk production has capacity shortage	Shift production for product on safety stock forward or backward and fixate this
Bulk release has capacity shortage	Shift production for product on safety stock forward or backward and fixate this
Packaging has capacity shortage	Shift demand forward or backward

TABLE 5, PLANNING ACTIONS

Although the MILP will also shift the demand with the lowest impact on the optimal cost, the MILP does not know whether there is flexibility in the delivery and will therefore only deliver late when the backorder costs are lower. As the backorder cost are a general parameter to the MILP, this will not assure that the right demand is shifted. This can be assured by the planner, who is assumed to have access to the information necessary to make this decision.

After this is done, rerunning the Bulk production quantity model to determine the final production plan is necessary. This plan will assure production plan that is both reliable and assures a minimal cost.

6.3 CONCLUDING

This chapter introduces the two deterministic decision support models of the planning hierarchy. In these decision support models, there are uncertainties regarding the demand and the actual available capacity. Because the design assures sufficient buffering of demand a deterministic representation of this demand is sufficient for these models. Furthermore, because the capacity availability is not an issue for the non-bottleneck production units, it is sufficient to incorporate the expected value to be accurate in the deterministic representation.

7. TESTING THE NEW UNCERAINTY HANDLING

To determine the planned lead times and the effect of the new rules and bottleneck control of uncertainty handling, a simulation model was build using excel VBA. This chapter introduces this simulation model and the results of this model. This simulation model was built as a monte carlo simulation and incorporates behavior of the different production units. A distinction is made between the current model for validation (A), a model implementing the new production rules (B). This final model is used to support the case assignment.

The chapter is organized as follows; it starts with a short description of the detail and assumptions per production unit. After that other relevant assumptions are covered. Next is an explanation of the simulation period. Finally the results of the first validation model are shown.

7.1 PRODUCTION UNITS

The simulation model covers the bulk production, bulk release and packaging with different amount of detail. The packaging release is only represented as a throughput time, as this is the most reliable resource in the production process.

For the bulk production (Figure 5) all workstations are modeled, including a net availability of these workstations per time unit. Hence, capacity for the unit is the maximum capacity available multiplied by a stochastic variable. Next, all production orders are produced based on their bill of materials. Once they are in the system they are produced first come first served at the next workstation. Production time for each workstation is equal to the planned production time for the workstation. Changeovers have been divided into two groups, large and small changeovers. Large changeovers take place between batches of different products, while small changeovers take place between batches of the same product.

Bulk release (Figure 6) is modeled in two steps, a first step for the lab and deviation check and a second step for the QP. A distinction is made between batches with deviations (possible quality issues) and without deviations. For products with deviations, the first step consists of both the deviation check and the lab, while for products without deviations only the lab is used.

Production times for the lab are drawn from two production time distributions, one for regular and one for problematic cases. Net capacity in the lab is modeled as a variable, but deterministic amount per week. Throughput times for the deviation check are drawn from a throughput time distribution while the throughput times for the QP are drawn from two distributions (problematic and regular cases) for the validation while they work under a different rule for the new situation.

The packaging (Figure 7) is modeled as two parallel production lines with stochastically drawn net capacity per workstation per time unit. Again, all production orders are produced first come first served once they have entered the production unit. Changeovers are again split in large changeovers and small changeovers and take the amount of time that is planned per type. Production times are determined based on order size and production speed of the different machines.

7.2 ADDITIONAL ASSUMPTIONS

Besides the behavior of the production units, assumptions were made regarding the demand arrival and some secondary processes in the production system. Some of these processes have not been discussed in the thesis, because they are out of scope. It is nonetheless important to mention the assumptions done with regard to these products. The additional assumptions and their impact are listed in Table 9.

Subject	Assumption	Implication
Tender demand delivery plan	All tender demand has a first delivery of 50% of the goods in 10 weeks and a second delivery of 50% of the goods after 20 weeks	Tender demand is split differently over time, therefore capacity problems could arise at different points in time, as this is true for both the validation and the new situation, this is not problematic.
Artwork changes	Artwork changes are handled in time and do not affect production planning as such	There is a possibility that artwork changes interrupt the production plan and packaging has to be re planned, as artwork changes have to be announced 10 weeks before final delivery and most artwork has 4 weeks delivery time, this is not problematic
API and raw material inventory	All API and raw materials are kept on inventory and are available for use in all products	In line with the design, the API and raw material will be available from inventory
Country specific quality checks	All released batches are released for production in all countries	Some products have specific quality issues per country, this is however relatively rare and assumed to be solved in the assignment of orders to batches
Packaging materials	Packaging materials are available for production when needed	It can happen that there is no stock and delivery of packaging materials is late, due to the stock on packaging materials this is however quite rare
Load leveling	Load leveling products are in place for both bulk and packaging	Load leveling options ensure the correct use of the capacity buffers in place for all products, not having these in place would work beneficial for some groups and not beneficial for others

TABLE 6, ADDITIONAL ASSUMPTIONS SIMULATION

Furthermore, for the validation of the model, production and release data for 2012 was simulated. All orders and releases during the period were replicated. For the planned lead time determination, the again the release decisions of 2012 were used to determine a first insight of the behavior based on the new rules.

7.3 VALIDATION OF THE SIMULATION MODEL

Validation of the model has to be done to assure that the model captures the changes necessary. For this particular simulation, this implies determining the throughput times of the different production units. Furthermore, as there is a large expected change in the waiting time before the lab resource in the bulk release production unit, this waiting time is measured explicitly as well.

One could argue that the performance of the simulation in terms of service level should also be validated. The problem that arises is that there is no detailed information about the requested delivery date of orders, nor planned finishing date of batches or batch release. Therefore this check could not be done.

To determine the throughput times the arrivals at the production units were kept to the known level, while the rest was drawn from the distribution. Figure 25 and Figure 26 show the results for both the standard deviation and the average, including the 5% confidence interval of the monte carlo simulation after extensive parameterization.

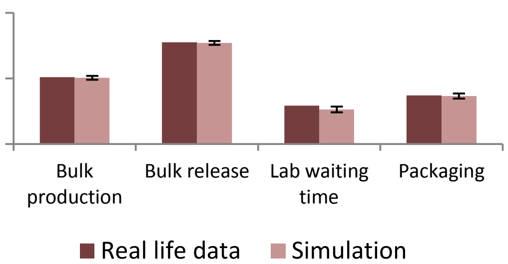

FIGURE 24, AVERAGE THROUGHPUT AND WAITING TIMES IN WORKING DAYS

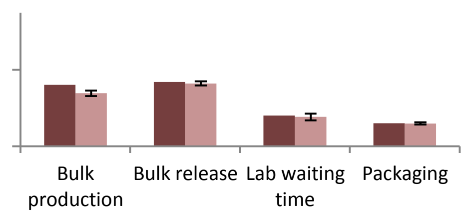

FIGURE 25, STANDARD DEVIATIONS OF THROUGHPUT AND WAITING TIMES IN WORKING DAYS

As can be seen the simulation in general approximates the behavior of the production units well. The most surprising result is the bad real life performance of the bulk production department. Even when extensively changing the available capacity for this production unit, the behavior does not approximate the standard deviation well.

This can be explained by observations. First of all, the simulation takes into account the sharing of resources, which is not fully operational yet. Second the simulation takes into account machine failure as a percentage each day, while real failures are quite lengthy. Please note that both points amplify each other.

Please note that therefore the results in the new situation might also underestimate the new standard deviation in the bulk production step.

7.4 PLANNED LEAD TIMES BULK PRODUCTION QUANTITY MODEL

Knowing that the simulation model represents the production process well, the next step is to determine the planned lead times and the behavior of the production process given the new way of handling uncertainty. To do this, the following behavior has been adjusted in the simulation.

1. Bulk release is processed via first due first served
2. Quality Person works with a maximum throughput time of 2 days
3. Isolation of possible problematic quality batches
4. Adjust the capacity of quality to the required capacity based on QED

As the goal is to find the planned lead times that are feasible for the bulk release planning, the throughput time of interest is not the throughput times of individual production units anymore. At this point the interest lies in the throughput time until a given production unit. To be specific, the throughput time until packaging. Hence the time that is of interest is the average time until the batch is *availabl*e for packaging.

Three different product groups are distinguished, first of all the regular batches, second the "isolation batches" that are released as a consequence of rule 3 and last the problematic batches. Where the regular and isolation batches are the ones that are of importance to determine the planned lead time.

Figure 27 and Figure 28 show the results for the different groups. A at first glance surprising result is that the isolated production group shows a higher standard deviation than the regular group. This is caused by the first due first served rule which reduces the variability strongly.

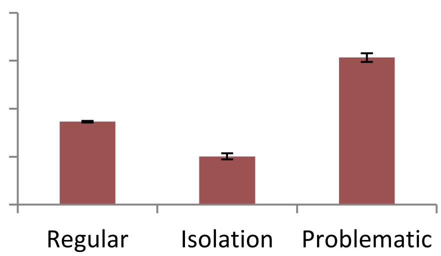

FIGURE 26, AVERAGE PRODUCTION TIME BULK PRODUCTION + BULK RELEASE IN WORKING DAYS

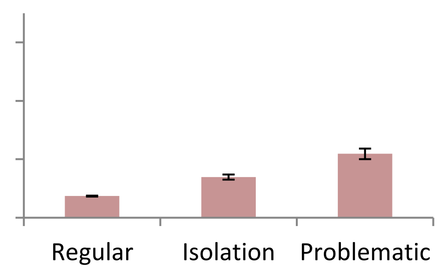

FIGURE 27, STDEV PRODUCTION TIME BULK PRODUCTION + BULK RELEASE IN WORKING DAYS

Besides, it can be seen that the negative effect of adding extra load on both the bulk production and bulk release by creating extra batches in terms of extra load and therefore extra waiting time is smaller than the positive effects of the isolation and the other rules.

Hence based on the information at hand the planned production time should be anywhere between 6,8 weeks (average) and 9,5 weeks (98% confidence). Based on the information out of the simulation and the biweekly planning intervals for the ORDB, this leads to the following configuration for the planned lead times in the ORDB.

Production unit	Planned lead time	Explanation
Bulk production	4 weeks	Because the next production step actively buffers the uncertainty, it is acceptable to have some late deliveries from the bulk production
Bulk release	4 - 6 weeks	Dependent on the safety buffer one wants to have this should be either 4 or 6 weeks throughput time.
Packaging	2 weeks	With no changes to the production this can stay the same
Packaging release	2 weeks	With no changes to production, this can stay the same

TABLE 7, PLANNED LEAD TIME PER PRODUCTION UNIT

7.5 NEW VS OLD UNCERTAINTY HANDLING

Besides determining the planned lead time, the simulation was also built to compare the old uncertainty handling design to the new uncertainty handling design. Figure 28 shows the average and standard deviation of the average and standard deviation of the throughput times of all regular products until they become available for packaging. As can be seen the new way of handling uncertainty in the system outperforms the old method in both average and standard deviation.

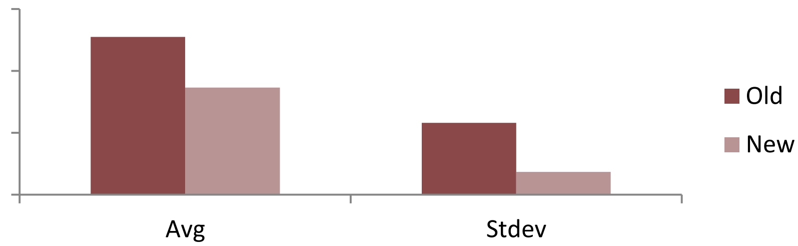

FIGURE 28, OLD VS NEW UNCERTAINTY HANDLING

Realizing that the average throughput time in the system determines the work in progress in the system makes one realize the reduction in costs that is eminent with this reduction. Taking into account that the average throughput time is reduced by almost one third of its original value implies an important decrease in cost.

Furthermore, the strong reduction in standard deviation makes it possible to make reliable production plans until the packaging step, removing an important part of the necessity for the stocks after the bulk release. Again this results in a significant reduction in costs.

7.6 CONCLUDING

This chapter introduced the simulation model and determines the planned lead time for the production unit. Based on this simulation model, it can be concluded that the new production rules show a large reduction in standard deviation until packaging. Making this step more reliable makes it possible to remove the stock that is currently kept after the bulk release without compromising the performance criteria that is at stake. Furthermore, the strong reduction in average throughput time reduces the work in process costs of both bulk production and bulk release.

8. NEW BULK PRODUCTION PLAN

This chapter examines the production plans that are a result of the MILP including the QED control function. The chapter consists of two parts. First of all a comparison is made between the production decision for the current decision and the current production plan and the production decision of the new ORDB. Second a comparison was made to see the impact of making campaign and possible batch size variable. Finally sensitivity analysis was done to see the robustness of the production plan for changing costs in the objective function. All production plans that are used for this analysis can be found in the appendix.

8.1 PRODUCTION PLANS

The first comparison to be made is a comparison regarding the four different products that are in scope for the production decision. Figure 29 shows the planned production amount in number of batches for these different products.

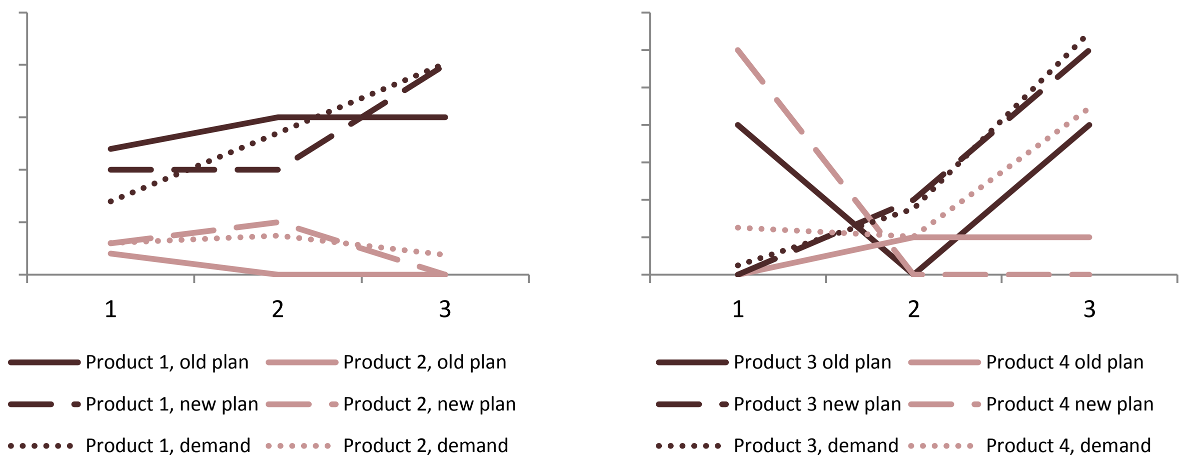

FIGURE 29. PRODUCTION PLANS, OLD - NEW - DEMAND

As can be seen the production amounts in the new case follow the actual demand better and therefore reduce inventory cost. Furthermore, it can be seen that the production quantity in batches is not as constant as it was with the former production plan. This indicates that it is beneficial to produce different be more flexible in the number of batches in a campaign.

8.2 DIFFERENT SCENARIOS

The second analysis was done with regard to the freedom in production in bulk, which makes a comparison between total freedom in production quantity $\left(\hat{C}_{i,t} \in Q^0,\ \hat{R}_{i,t} \in Q^0\right)$, only multiples of batches as production quantity $\left(\hat{C}_{i,t} \in Q^0,\ \hat{R}_{i,t} \in N^0\right)$ and third only multiples of campaigns as production quantity $\left(\hat{C}_{i,t} \in N^0,\ \hat{R}_{i,t} \in N^0\right)$. This is done to determine the impact of relaxing the campaign decision on the production plan and the cost in the system. One would expect the costs to be lower with more freedom in the system.

This expectation is confirmed by the outcome of the model, as can be seen in Figure 30. It is clear that the step from fixed campaign size to free campaign size incorporates the most benefit in the optimal cost. Given the complexity of changing the batch sizes it should be sufficient to make the campaign sizes variable.

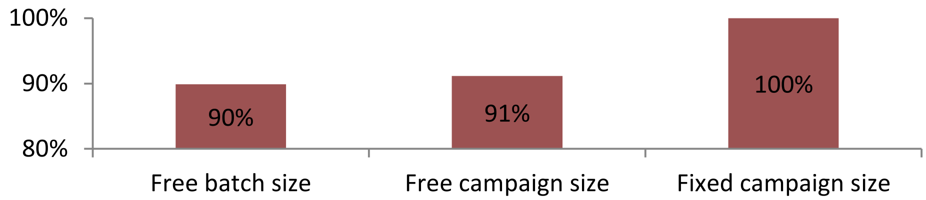

FIGURE 30, COMPARISON OF FIXED PRODUCTION QUANTITIES

Looking at the structural changes in the campaign sizes it can be seen that the flexibility is mostly used for the products the larger volume products. It is however not true for the largest volume product. As these are the products that are mostly shifted when flexibility has to be achieved, the advantage of the added flexibility becomes clearer.

8.3 SENSITIVITY ANALYSIS

Looking at the sensitivity to the different costs in the system there are two ratios that are of special interest. First of all the ratio between backorder cost and holding/production cost. If the backorder cost are too low, there is no incentive in the system to produce. The second ratio that is of importance is the ratio between holding cost and startup cost. If the startup cost get higher the optimal behavior regarding campaign sizes could change.

For the backorder costs a range from 8 – 12 % of the batch value was considered to determine the critical value of the backorder costs[2]. The production plan shows to be robust in the backorder costs. This can be explained by the structure of the MILP. Because the order release decision is not directly part of the objective function, the model becomes more robust.

Looking at the startup costs, for the bulk production startup a range of 2,5 – 10 % was considered, while a range 50% - 150% of current costs was considered for the server cost per campaign in bulk release. For the bulk production costs there are small changes in the structure. It turns out to be more beneficial to make larger orders, although the effect can only be seen with one product. For the quality the plans are identical for the different costs. Larger changes than the ones mentioned are unrealistic given the nature of these costs.

8.4 CONCLUDING

This chapter looks at the newly found production plans for the IPT WH. In comparison to the previous production plan, there is a stronger connection to the demand forecast. Furthermore the planning shows clear cost advantages in flexible campaign sizes. Besides that, the plan is robust against changes in relevant cost parameters.

[2] Based on ~5% startup cost for production and , ~1% holding cost per month and ~2% server cost per campaign

9. GENERALIZATION OF THE QED MODEL

One of the things that is new in this case is the use of the QED to determine a maximum load or a minimum capacity for the production decision. This chapter tries to generalize the findings in this model and tries to find different places for application of it. Recall that the model is used for multi-server parallel production system, has linear server and holding cost and has a fixed planned lead time. This chapter first handles the cost function, followed by an analysis in combination with the planned lead time concept. The chapter ends with a brief discussion on implications and possibilities of the model. Although this chapter might be somewhat out of tone in the remainder of the report, the findings from it are interesting. Therefore the chapter has been included in the thesis.

9.1 MINIMIZING THE COST IN QED

Recall the cost function with server cost δ and linear holding cost ω from paragraph 5.2.2:

$$K(\beta) = \delta a + \sqrt{a}\left(\delta\beta + \frac{\omega}{\beta}C_*(\beta)\right) \qquad (2.6)$$

Rewriting (2.6) like $K(\beta) = \delta\left(a + \sqrt{a}\beta\right) + \omega\left(\sqrt{a}\frac{1}{\beta}C_*(\beta)\right)$ shows the two parts in the cost function. First of all there are the costs that are necessary for the severs, where a can be interpreted as the minimal number of servers necessary to have a stable system, while $\sqrt{a}\beta$ can be seen as the capacity buffer. Second, there are costs for the *expected* number of units per time unit in the system, which can be approximated by $\sqrt{a}\frac{1}{\beta}C_*(\beta)$. Note that the costs for holding inventory ω are linear, which is in line with the MILP.

This function neglects the error term as it was introduced by Janssen and van Leeuwaarden (2011), which makes the function reliable for systems with high server cost and low inventory cost (the efficiency driven regime). For large values of $\frac{\omega}{\delta}$ this correction term should be incorporated in the function[3]. This results in an underestimation of the optimal server amount by less than 0.1 for any $\frac{\omega}{\delta} < 1$.

Note furthermore that this optimization does not have a restriction on an expected waiting time, hence it only determines the optimal time that a product should stay in the system given the different cost aspects. The waiting time that results from that is a result of this function. As there is a planned lead time in the deterministic model, this waiting time is of relevance. This aspect will be regarded in section 9.2.

Looking at the function please note that β_{opt} is insensitive to the load on the system. This can be seen by setting $K_*(\beta) = \delta\beta + \frac{\omega}{\beta}C_*(\beta)$ and therefore $K(\beta) = \delta a + \sqrt{a}\left(\delta\beta + \frac{\omega}{\beta}C_*(\beta)\right)$. This function will be optimal if $K_*'(\beta) = 0$ and is therefore insensitive to a. This implies that the only relevant

[3] An example of this could be a setting with capital goods

parameter for the determination of the β_{opt} is $\frac{\omega}{\delta}$. Keep in mind that although there is no relationship between β_{opt} and a, there is a relationship between $E(W_q)$ and a.

As can be seen in Figure 31 β_{opt} is increasing in $\frac{\omega}{\delta}$ for the region that is approximated to an acceptable range by the QED. Furthermore, β_{opt} is decreasingly sensitive to $\frac{\omega}{\delta}$ as $\frac{\omega}{\delta}$ gets larger.

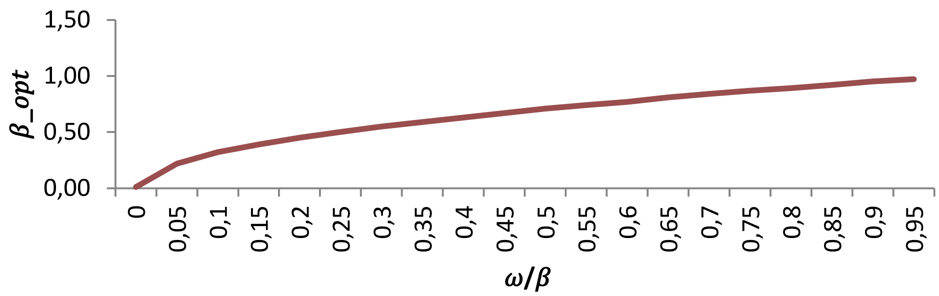

FIGURE 31, RELATIONSHIP BETWEEN β_{opt} AND $\frac{\omega}{\beta}$

As the main point of interest is the load that can be offered to the system, and it is known that β_{opt} is insensitive to this load given $\frac{\omega}{\beta} < 1$, it is interesting to look at the derivative towards load for $K(a|\beta_{opt})$. As can be seen this derivative becomes $\delta - \frac{1}{\sqrt{a}}\left(\delta\beta_{opt} + \frac{\omega}{\beta}C_*(\beta_{opt})\right)$. Hence, for any value with $\beta_{opt} < 1$ or $a > 10$, the minimal cost act approximately linear in its load. It is easy to see the linearity is approximately δ. As has been shown in the Figure 31, low values of $\frac{\omega}{\delta}$ imply low values of β_{opt}. Therefore systems with relatively low holding versus capacity cost can be approximated neatly with a linear optimal cost function. This was also confirmed for the cost function for IPT WH.

Besides the change in β_{opt} for $\frac{\omega}{\delta}$ it is also interesting to look at the shape of the cost function to understand the criticality of the optimum. Figure 32 to Figure 34 show $K(\beta)$, furthermore Figure 35 zooms in on the optimal region of each of these functions.

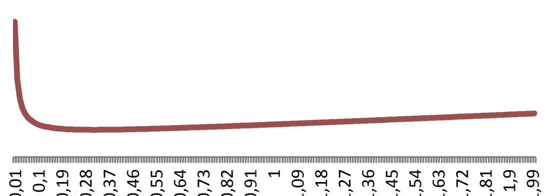

FIGURE 32, COST FUNCTION WITH $\frac{\omega}{\delta}$= 0.1

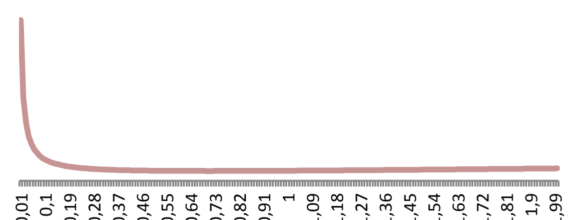

FIGURE 33, COST FUNCTION WITH $\frac{\omega}{\delta}$= 0,5

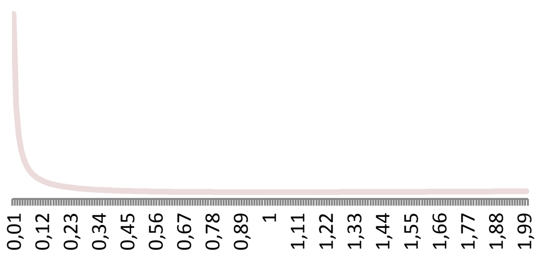

FIGURE 34, COST FUNCTION WITH $\frac{\omega}{\delta}$ = 1

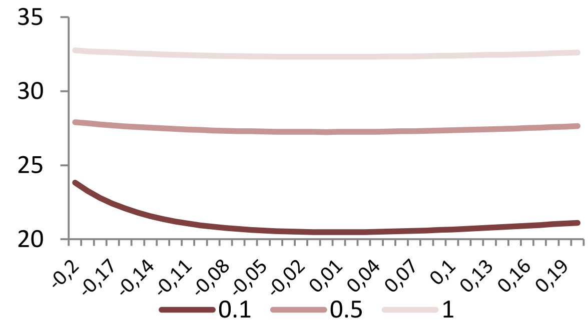

FIGURE 35, COST FUNCTION FOR OPTIMAL REGION + AND – 0.2

Based on these four graphs and their underlying functions three things can be concluded:

1) for any value of $\frac{\omega}{\delta}$ the case with the highest cost is the case with too little capacity in the system
2) with increasing $\frac{\omega}{\delta}$ the decision for the optimal point becomes less critical
3) relatively large changes in $\frac{\omega}{\delta}$ have a very limited effect on the optimal cost

Hence, as long as one does not have too little servers in the system, optimal or near optimal costs will be achieved. As the costs vary relatively little around β given $0.1 < \frac{\omega}{\delta} < 1$, the linear cost approximation also holds for $\beta \pm 0.2$. It is easy to see that such a linear relationship can be used in the MILP.

9.2 PLANNED LEAD TIME

Just like the model that was used for the IPT WH most planning models assume a planned lead time per production unit. Therefore use of the QED in a planning context implies taking a maximum lead time into account.

Working with a planned lead time has consequences on the generic outcome of the model. The most important consequence is the change of the time a unit spends in the system and therefore the ω portion of the cost. As one is typically looking to include a safety buffer in the planned lead time, $\omega\left(\tau - E\left(S|W_q\right)\right)$ can be seen as the cost of having this buffer. Furthermore, the size of this buffer should be a predetermined value and is considered to be known.

Recall from section 5.4.3.2 that $E\left(W_q\right) \approx \frac{1}{\sqrt{a+\beta\sqrt{a}}} \frac{1}{\mu\beta} \frac{1}{1+\frac{\beta}{h(-\beta)}}$ which implies that there are two options to change $E\left(W_q\right)$, either changing the load a or changing β. A distinction is made into two options;

A. $E\left(W_q\right) + buffer > planned\ lead\ time$; Increase buffer capacity $\beta_{opt}\sqrt{a}$ or decrease load a

B. $E\left(W_q\right) + buffer < planned\ lead\ time$; Decrease buffer capacity $\beta_{opt}\sqrt{a}$ or increase load a

As was clear from the previous section, as long as the changes are somewhat restricted, the changes should not change the optimality of the solution.

On a side note, the representation of the time buffer implies a fixed time buffer, while this could be influenced by the number of servers in the system as well. Although this is a likely effect, this effect has not been investigated throughout this case and will therefore not be taken into account. This might however be an interesting further extension.

9.3 IMPLICATIONS OF RESULTS

There are several interesting implications of the results. These are split in the flatness of the cost curve in β, the linearity of optimal and near optimal cost in the load, and the stability of the result in $\frac{\omega}{\delta}$.

Starting with the flatness of the cost curve in β, this basically implies that one can choose any waiting time, given the conditions for this flatness are met. As was shown in section 8.1, the cost do not change drastically with increasing β. Furthermore, as the amount of servers determines the waiting time in the system, reduction of waiting time is relatively cheap by increasing the amount of servers. This makes it possible to determine the waiting time to some extent.

Looking at the linearity of both the optimum and the region around the optimum (due to the flat cost curve) in the load on the system also brings some interesting opportunities. First of all, incorporation in a linear model as is presented in this thesis becomes a possibility. Second, knowing that this linearity exists supports the decoupling of the decisions into an aggregate capacity and a SCOP level. As long as the capacity has a small buffer in place, the decision at the capacity will not remove the optimal decision from the SCOP level.

Besides the support for the split in the two levels, the linear costs in the load and the reduction of waiting time in the load could also be used to support a scaling decision. As the costs for different systems are linear, it must be true that costs are linear for two smaller systems of the same type as well. Knowing what the load would be on each of these smaller systems, the waiting time given these systems would also be known. Therefore the decision on whether to scale is solely a decision on what waiting time one wants to achieve. The QED can give a first approximation of this waiting time based on the loads that could be achieved.

9.4 CONCLUDING

In this chapter it is tried to generalize the results of the QED model. The QED control function can be used in general planning context with is relatively insensitive to β. This makes the cost stable even if the amount of servers differs slightly. Furthermore it becomes possible to pick a desired throughput time at stable cost. Second, the costs of the system are approximately linear in the load on the system. This supports the choice for an aggregate capacity level and a SCOP level in the planning hierarchy, and makes it possible to answer scaling decisions. And finally the cost are relatively insensitive to $\frac{\omega}{\delta}$, making the aforementioned statements hold in situations where $0.1 < \frac{\omega}{\delta} < 1$ and making it possible to apply the QED stochastic control for situations where this statement holds

10. CONCLUSION

Based on the case study conclusions can be drawn in terms of the case assignment and the two research questions as they were defined in chapter 3.

10.1 CASE ASSIGNMENT

The simulation shows the new planning concept is strongly beneficial in terms of throughput time average and variability to introduce the proposed set of design rules and uncertainty incorporating models that were used in this study. Having this set of rules reduces the average throughput time until packaging by 33 %. Furthermore it greatly increases reliability of this throughput time, making it possible to package based on these orders. This impliest it becomes possible to remove the stock after the packaging for the large products.

Furthermore, introducing the cost for the bulk release in the ORDB and relaxing the campaign size restriction results in a significant reduction in costs for the production unit. This increases the flexibility in the system without rendering higher total cost in the system.

To achieve this new situation IPT WH should incorporate the VII rules that were introduced in chapter 4 and incorporate the stochastic decision support models that were introduced in chapter 5.

10.2 HOW CAN THE *DESIGN* OF A PRODUCTION PLANNING HIERARCHY ASSURE THE CORRECT HANDLING OF UNCERTAINTY

Regarding the design of the system there are two main learning from the study. First, a decision model to determine whether to incorporate uncertainty in the design of a planning structure or in the modeling of the structure. Second, different ways of smart design in the system can have strong positive effects on the outcome.

10.2.1 WHERE TO HANDLE UNCERTAINTY

The first learning in the design is a set of design rules regarding the handling of uncertainty in production planning. To determine whether the uncertainty should be handled via models or via the design one should answer two questions; does the planning decision influence the uncertainty, and, what is the impact of said uncertainty on the optimal solution. Based on the answers to these two questions different actions should be taken, as is shown in

	Possible to have influence on uncertainty	**Not possible to have influence on uncertainty**
High impact	Stochastic control function (model)	Isolate and buffer through smart design
Low impact	Deterministic control function (model)	No action

TABLE 8, DESIGN VS MODEL, WHERE UNCERTAINTY SHOULD BE HANDLED

Based on these rules one can choose whether to incorporate the uncertainty in the design or in the modeling in the system. As can be seen it is necessary to buffer through smart design in the case of high impact and no influence uncertainty. This can be done via smart production rules and smart incorporation of advanced demand information.

10.2.2 SMART DESIGN

In this case study there are several ways of smart design that turn out to be beneficial. These are split in two groups, design that isolates uncertainty sources and design that assures buffer existence and optimal buffer use.

Looking at the isolation, it is clear that there are two types of isolation in this case. First, the isolation of a demand group that is highly uncertain based on demand type. Second, the isolation of a product group with throughput times based on quality. This shows that one never knows beforehand what type of isolation is necessary to effectively isolate the groups that should be buffered and advocates a strong understanding of the behavior of a particular production process before making the design. Hence although the type of isolation is not known, the use of isolation itself is very effective.

As for the rules that assure buffer existence and buffer use, the rules include a rule that restricts the delivery plan, a rule that dictates the order handling in production unit and a rule that determines buffer for low volume product to ensure flexibility in the system. This shows that the rules can be relatively simple without it being necessary to impose large restrictions on the surroundings and still be very effective. Again this advocates understanding of consequences to the exterior world when choosing these design rules, and taking these consequences into account when choosing one.

10.3 HOW CAN *MODELING* IN A PRODUCTION PLANNING HIERARCHY ASSURE CORRECT HANDLING OF UNCERTAINTY

As for the modeling, the case study shows that relatively simple stochastic control models at the right place can have a large effect on the performance of the system. By incorporating stochastic control on load for the bottleneck resource there it becomes possible to have deterministic control for the remainder of the production units.

From the specific perspective, the extent to which the M/M/C queue predicts the behavior of the waiting time in the lab was surprising. It is fully realized that this might not be a result that can be copied to other situations under any given circumstances. It is however likely that the behavior of the waiting time in terms of scaling with high utilization in a multi-server context can be captured by an M/G/C queue and thus by the QED regime.

Knowing this, the QED regime can be used to determine the upper bound of a production load for a production unit. As there is the assumption of exponential inter arrival times, it can also take into account part of the uncertainty in arrivals from previous steps in the production process. Especially when looking at larger time intervals, as is the case with IPT WH, the capacity buffer based on the QED proofs to perform as intended. This leads to suspect that the same will be true for a load restriction.

Furthermore, by acquiring an upper bound for a load, the deterministic optimization model can still be utilized to determine the optimal solution, while the lead time feasibility of the plan is assured.

Besides this it was shown that the QED optimal cost are approximately linear in its load and are relatively insensitive to the β. This makes it possible to “choose” a required waiting time when designing the system by choosing the right number of servers without having high changes in cost.

11. FURTHER RESEARCH AND LIMITATIONS

Based on this case studies there are opportunities for further research that could be interesting from a theoretical perspective and for IPT WH.

11.1 FURTHER RESEARCH

From a theoretical perspective there are many options for additional research regarding uncertainty in a hierarchical planning environment.

A starter of this would be further extension and testing of the framework for handling uncertainty. Although this is a first step, there are still questions to be asked with this framework. One of the main questions is the positioning of the dimension time in the framework. It could be argued that time in general is not something that can be influenced and should therefore be buffered against.

Besides the design framework the QED results show interesting opportunities. If one would be able to replace the cost function in the QED with any other function, different types of decisions would become possible. Also an extension to determine a model for a scaling decision would be interesting.

11.2 IPT WH

For the IPT WH there are several interesting opportunities that for various reasons were left out of this masters thesis.

First of all, there is the bulk production. This production step seems to have a lot of planning and re planning done at the production unit. It is not exactly clear why this is done, but it is surprising that the actual throughput time is almost three times the production time in this production unit.

Second, there seems to be an interesting opportunity to determine a decision model whether to sign in on a tender. Right now, one only takes into account the possible value that the tender can bring, while any quote that the tender department signs into also brings extra cost to the production department. First thing that comes to mind in such a setting would be to see this quote as an option on the tender volume, where the value and cost of such an option could be determined.

BIBLIOGRAPHY

IMS. (2013, 1 24). Retrieved 1 2013, 24, from http://www.imshealth.com/deployedfiles/ims/Global/Content/Corporate/Press%20Room/Top-Line%20Market%20Data%20&%20Trends/2011%20Top-line%20Market%20Data/Global_Pharma_Market_by_Spending_2003-2011.pdf

US Pharmacist. (2013). Retrieved 1 25, 2013, from http://www.uspharmacist.com/content/s/216/c/35249/

Adan, I., & Resing, J. (2002). *Queueing Theory.* Eindhoven: Department of mathematics and computing science Eindhoven University of Technology.

Altrichter, F., & Caillet, T. (2005). Supply chain management in a pharmaceutical company, chapter 19. In *Supply chain management and advanced planning* (pp. 355-370).

Anthony, R. (1965). *Planning and control systems, A framework for analysis.* Boston: Graduate school of Business Administration, Harvard University.

Bertrand, J., & Fransoo, J. (2002). Operations management research methodologies using quantitative modeling. *International Journal of Operations & Production Management*, 241-264.

Bertrand, J., Wortmann, J., & Wijngaard, J. (1990). *Production Control: A Structural and Design Oriented Approach.* Amsterdam: Elsevier.

Blau, G. E., Pekny, J. F., Varma, V. A., & Bunch, P. R. (2004). Managing a Portfolio of Interdependent New Product Candidates in the Pharmaceutical Industry. *Journal of Product Innovation Management*, 227-245.

Bok, J.-K., Grossmann, I. E., & Park, S. (2000). Supply Chain Optimization in continuous Flexible Process Networks. 1279-1290.

Borst, S., Mandelbaum, A., & Rieman, M. (2004). Dimensioning large call centers. *Operations Research*, 17-34.

Boulaksil, Y. (2005). A procedure to set safety stock levels in multi echelon inventory systems. *Masters thesis.*

Buzacott, J. A., & Shantikumar, J. G. (1993). *Stochastic Models of Manufacturing systems.* New Jersey: Prentice Hall.

Cremer, P., Losch, M., & Schrader, U. (2009, April). Maximizing efficiency in pharma operations. *McKinsey Quarterly*, p. 4.

customization, E. f. (2005). Jack C.P. Su; Yih-Long Chang; Mark Ferguson. *Journal of Operations Management*, 305-318.

Dallery, Y., & Liberopoulos, G. (2000). Extended kanban control system: combining kanban and base stock. *IIE Transactions*, 369-386.

Donselaar, K. v., Woensel, T. v., Broekmeulen, R., & Fransoo, J. (2006). Inventory control of perishables in supermarkets. *international journal of Production Economics*, 462-472.

Drews, J. (2003). Strategic trends in the drug industry. *Drug discovery today*, 411-420.

Eliades, G., Retterath, M., Heultenschmidt, N., & Singh, K. (2012). *Healthcare 2020.* Bain&Company.

FDA. (n.d.). Retrieved 11 2012, from FDA: http://www.fda.gov/drugs/informationondrugs/ucm129689.htm

Fransoo, J. C., & Rutten, W. G. (1994). A Typology of Production Control Situations in Process Industries. *International Journal of Operations & Production Management*, 47-57.

Fransoo, J., & Wouters, M. (2000). Measuring the bullwhip effect in the supply chain. *Supply Chain management: An International Journal*, 78-89.

Fransoo, J., & Wouters, M. (2001). Multi-Echelon Multi-Company Inventory Planning with Limited Information Exchange. *The journal of the operational research society*, 830-838.

Gatica, G., Papageorgiou, L., & Shah, N. (2003). Capacity planning under uncertainty for the pharmaceutical industry. *trans IChemE*, 665-678.

Gershwin, S. B. (2005). Design and operation of manufacturing systems.

Gonce, A., & Schrader, U. (2010). Plantopia, a mandate for innovation in pharma manufacturing. *McKinsey Quarterly*.

Grabowski, H., & Kyle, M. (2007). Generic Competition and market exclusivity periods in Pharmaceuticals. *Managerial and decision economics*, 1-12.

Gupta, D., & Benjaafar, S. (2004). Make-to-order, make-to-stock, or delay product differentiation? A common framework for modeling and analysis. *IIE transactions*, 529-546.

Halfin, S., & Whitt, W. (1981). Heavt-Traffic Limits for Queues with many exponential servers. *Operations Research*, 567-588.

Hax, A. C., & Meal, H. C. (1973). Hierarchical Integration of Production planning and scheduling. 1-25.

Hennink, A. (2006, 11 20). A comparison of production planning strategies; a product vs a capacity oriented approach - an approach in a pharmaceutical company - Masters thesis. *A comparison of production planning strategies; a product vs a capacity oriented approach - an approach in a pharmaceutical company - Masters thesis*. Eindhoven.

Holland, K., & Batiz-Lazo, B. (2004). The global pharmaceutical industry. *General economics and teaching*.

Hopp, W., & Spearman, M. (2006). To Pull or Not to Pull: What is the Question. *Manufacturing & service Operations management*, 133-148.

Jansen, M. (2011). Anticipation in Supply Chain Operations Planning. Eindhoven.

Jansen, M. M., Kok, T. G., & Fransoo, J. (2013). Lead time anticipation in supply chain operations planning. *OR Spectrum*, 251-290.

Janssen, A., & Van Leeuwaarden, J. (2011). Refining Square-Root Safety Staffing by Expanding Erlang C. *Operations Research*, 1512-1522.

Karaesmen, F., Buzacott, J., & Dallery, Y. (2002). Integrating advance order information in make-to-stock production systems. *IIE transactions*, 649-662.

Kingsman, B. G. (2000). Modelling input-output workload control for dynamic capacity planning in production planning systems. *Production economics*, 73-93.

Kok, T. d., & Fransoo, J. (2002). Planning Supply Chain Operations: Definition and Comparson of Planning concepts (working paper). In J. Ayers, *Handbook of supply chain management.*

Krishnan, V., & Ulrich, K. T. (2001). Product Development Decisions: A Review of Literature. *Management Science*, 1-21.

Lee, Y., & O'Conner, G. C. (2003). New product launch strategy for network effects products. *Academy of marketing science Journal*, 241-254.

Leeuwaarden, J. V. (2006). Delay Analysis for the Fixed-cycle Traffic-Light Queue. *Transportation Science*, 189-199.

Leeuwaarden, J. v., Denteneer, D., & Resing, J. (2010). A discrete-time queueing model with periodically scheduled arrival and departure slots.

Levis, A. a., & Papageorgiou, L. G. (2004). A hierarchical solution approach for multi-site capacity planning under uncertainty in the pharmaceutical industry. *Computer and chemical engineering* , 707-725.

Lindgardt, Z. (2008, June). Waking the giant: business model innovation in the drug industry. *In Vivo: the business & medicine report*.

Maravelias, C. T., & Grossmann, I. E. (n.d.). Simultaneous Planning for New Product Development and Batch Manufacturing Facilities.

Martin-Moe, S., Lim, D. J., Wong, R. L., Sreedhara, A., Sundaram, J., & Sane, S. U. (2011). A new Roadmap for biopharmaceutical Drug Product development: integrating development, validation and quality by design. *Journal of pharmaceutical Sciences*, 3031-3043.

Masoumi, A. H., Yu, M., & Nagurney, A. (2012). A supply chain generalized network oligopoly model for pharmaceuticals under brand differentiation and perishability. *Transportation Research Part E*, 762-780.

Murphy, K., & Biesen, T. v. (2012). *Global healthcare Private equity report.* Bain&Company.

Nagurney, A., & Li, D. (2012). A dynamic network oligopoly model with transportation costs product differentiation and quality competition. 1-40.

Papageorgiou, L. G. (2009). Supply chain optimisation for the process industries: Advances and opportunities. *Computers and Chemical engineering*, 1931-1938.

Papageourgiou, L. G., Rotstein, G. E., & Shah, N. (2001). Strategic Supply Chain Optimization for the Pharmaceutical Industries. *Industrial Engineering and Chemistry research*, 275-286.

Rask-Andersen, M., Almen, M. S., & Schioth, H. B. (2011). Trends in the exploitation of novel drug targets. *Nature reviews drug discovery*, 570-590.

Reijers, H., & Van der Aalst, W. (2005). The effectiveness of workflow management systems: Predictions and lessons learned. *International Journal of Information Management*, 458-472.

Romero, J., Badell, M., Bagajewicz, M., & Puigjaner, L. (2003). Integrating Budgeting Models into Scheduling and Planning Models for the Chemical Batch Industry. *Industrial Engineering and Chemistry*, 6125-6134.

Rosetti, C. L., Handfield, R., & Dooley, v. J. (2011). Forces, trends, and decisions in the pharmaceutical supply chain management. *International journal of physical distribution and Logistics Management*, 601-622.

Russell, M., Gjaja, M., & Lawyer, P. (2008, April). The future of Pharma: ADjusting the Pharma R&D model. *BCG perspectives*.

Safaei, N., & Tavakkoli-Moghaddam, R. (2009). Integrated multi-period cell formation and subcontracting production planning in dynamic cellular manufacturing systems. *International journal of production economics*, 301-314.

Schmidt, C., & Grossmann, I. (1996). Optimizatino models for scheduling of testing tasks in new product development. *Industrial Engineering of Chemical Research*, 3498-3510.

Schneewiess, C. (2003). *Distributed Decision Making 2nd ed.* Berlin: Springer-Verlag.

Shah, N. (2004). Pharmaceutical supp chains: key issues and strategies for optimisation. *Computers and Chemical engineering*, 929-941.

Sousa, R. T., Liu, S., Papageourgiou, L. G., & Shah, N. (2011). Global supply chain planning for pharmaceuticals. *Chemical Engineering Research and Design*, 2396-2409.

Sundaramoorthy, A., Li, X., Evans, J. M., & Barton, P. I. (2012). Capacity planning under clinical trials uncertainty in continuous pharmaceutical manufacturing, 2: Solution Method. *Industrial and Engineering Chemistry research*, 13703-13710.

Vericourt, F. d., Karaesmen, F., & Dallery, Y. (2002). Optimal Stock Allocation for a Capacitated Supply System. *Management Science*, 1486-1501.

Vieira, G., Herrmann, J., & Lin, E. (2003). Rescheduling manufacturing systems: a framework of strategies policies and methods. *journal of scheduling*, 39-62.

You, F., & Grossmann, I. (n.d.). Optimal Design and Operational Planning of Responsive Process Supply Chains.

Yu, L. X. (2008). Pharmaceutical Quality by Design: Product and Process Development, Understanding, and Control. *Pharmaceutical Research*, 781-791.

APPENDIX I BULK ORDER PRODUCTION MODEL, PRODUCTION PLANS

All large products are yellow. All other products are not considered to be large products.

PRODUCTION PLAN FREE BATCH SIZE; # OF BULK BATCHES TO BE PRODUCED

Product\time period	1	2	3	4	5	6
1	2,811299	0	3,124216	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0,853876	0	0	0	0
5	0	0	0	0	0	0
6	0	1,486514	6	0	0	0
7	0	0	0	0	0	0
8	0	2,310327	0	0	0	0
9	0	0	0	0	0	0
10	9,676791	10	20	0	0	0
11	0	0	0	0	0	0
12	0	0	0,066778	0	0	0
13	0	0	0,066778	0	0	0
14	4,672933	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	7,645685	0	0	0	0	0
19	0	0	0,018563	0	0	0
20	5,733761	0	0	0	0	0

TABLE 9, PRODUCTION PLAN BATCH SIZES FREE

PRODUCTION PLAN FREE CAMPAIGN SIZE; # OF BULK BATCHES TO BE PRODUCED

Product\time period	1	2	3	4	5	6
1	3	0	3	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	0	0	0
5	0	0	0	0	0	0
6	0	2	6	0	0	0
7	0	0	0	0	0	0
8	0	3	0	0	0	0
9	0	0	0	0	0	0

10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	0	1	0	0	0
13	0	0	1	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	3	5	0	0	0	0
19	0	0	1	0	0	0
20	6	0	0	0	0	0

TABLE 10, PRODUCTION PLAN CAMPAIGN SIZES FREE

PRODUCTION PLAN FIXED CAMPAIGN SIZE; # OF BULK BATCHES TO BE PRODUCED

Product\time period	**1**	**2**	**3**	**4**	**5**	**6**
1	6	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	2	0	0	0	0
5	0	0	0	0	0	0
6	0	3	6	0	0	0
7	0	0	0	0	0	0
8	0	5	0	0	0	0
9	0	0	0	0	0	0
10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	0	5	0	0	0
13	0	0	5	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	10	0	0	0	0	0
19	0	0	2	0	0	0
20	6	0	0	0	0	0

TABLE 11, PRODUCTION PLAN CAMPAIGN SIZES FIXED

Product\time period	1	2	3	4	5	6
1	0	6	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	0	0	0
5	0	0	0	0	0	0
6	0	2	6	0	0	0
7	0	0	0	0	0	0
8	0	0	3	0	0	0
9	0	0	0	0	0	0
10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	0	1	0	0	0
13	0	0	1	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	3	5	0	0	0	0
19	0	0	1	0	0	0
20	6	0	0	0	0	0

TABLE 12, PRODUCTION PLAN FREE CAMPAIGNS BACKORDER COSTS 6 %

Product\time period	1	2	3	4	5	6
1	0	6	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	0	0	0
5	0	0	0	0	0	0
6	0	2	6	0	0	0
7	0	0	0	0	0	0
8	0	0	3	0	0	0
9	0	0	0	0	0	0
10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	0	1	0	0	0
13	0	0	1	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0

17	0	0	0	0	0	0
18	8	0	0	0	0	0
19	0	0	1	0	0	0
20	6	0	0	0	0	0

TABLE 13, PRODUCTION PLAN FREE CAMPAIGN BACKORDER COSTS 8 %

Product\time period	1	2	3	4	5	6
1	0	6	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	0	0	0
5	0	0	0	0	0	0
6	0	2	6	0	0	0
7	0	0	0	0	0	0
8	0	0	3	0	0	0
9	0	0	0	0	0	0
10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	0	1	0	0	0
13	0	0	1	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	8	0	0	0	0	0
19	0	0	1	0	0	0
20	6	0	0	0	0	0

TABLE 14, PRODUCTION PLAN FREE CAMPAIGN BACKORDER COSTS 10 %

Product\time period	1	2	3	4	5	6
1	3	0	3	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	0	0	0
5	0	0	0	0	0	0
6	0	2	6	0	0	0
7	0	0	0	0	0	0
8	0	3	0	0	0	0
9	0	0	0	0	0	0
10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	0	1	0	0	0

13	0	0	1	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	3	5	0	0	0	0
19	0	0	1	0	0	0
20	6	0	0	0	0	0

TABLE 15, PRODUCTION PLAN FREE CAMPAIGN BACKORDER COSTS 12 %

Product\time period	1	2	3	4	5	6
1	3	0	3	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	0	0	0
5	0	0	0	0	0	0
6	0	2	6	0	0	0
7	0	0	0	0	0	0
8	3	0	0	0	0	0
9	0	0	0	0	0	0
10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	1	0	0	0	0
13	0	0	1	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	8	0	0	0	0	0
19	0	1	0	0	0	0
20	6	0	0	0	0	0

TABLE 16, PRODUCTION PLAN FREE CAMPAIGN SERVER COSTS AT 50 %

Product\time period	1	2	3	4	5	6
1	3	0	3	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	0	0	0
5	0	0	0	0	0	0
6	0	2	6	0	0	0
7	0	0	0	0	0	0
8	0	3	0	0	0	0
9	0	0	0	0	0	0
10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	0	1	0	0	0
13	0	0	1	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0

17	0	0	0	0	0	0
18	3	5	0	0	0	0
19	0	0	1	0	0	0
20	6	0	0	0	0	0

TABLE 17, PRODUCTION PLAN FREE CAMPAIGN SERVER COSTS AT 100%

Product\time period	**1**	**2**	**3**	**4**	**5**	**6**
1	3	0	3	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	0	0	0
5	0	0	0	0	0	0
6	0	2	6	0	0	0
7	0	0	0	0	0	0
8	3	0	0	0	0	0
9	0	0	0	0	0	0
10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	1	0	0	0	0
13	0	0	1	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	8	0	0	0	0	0
19	0	1	0	0	0	0
20	6	0	0	0	0	0

TABLE 18, PRODUCTION PLAN FREE CAMPAIGN SERVER COSTS AT 200 %

Product\time period	1	2	3	4	5
1	3	0	3	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	1	0	0	0
5	0	0	0	0	0
6	0	2	6	0	0
7	0	0	0	0	0
8	3	0	0	0	0
9	0	0	0	0	0
10	10	10	20	0	0
11	0	0	0	0	0
12	0	1	0	0	0
13	0	0	1	0	0
14	5	0	0	0	0
15	0	0	0	0	0
16	0	0	0	0	0
17	0	0	0	0	0
18	3	5	0	0	0
19	1	0	0	0	0
20	6	0	0	0	0

TABLE 19, PRODUCTION PLAN FREE CAMPAIGN STARTUP COST BULK 2,5 %

Product\time period	1	2	3	4	5	6
1	3	0	3	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	0	0	0
5	0	0	0	0	0	0
6	0	2	6	0	0	0
7	0	0	0	0	0	0
8	0	3	0	0	0	0
9	0	0	0	0	0	0
10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	0	1	0	0	0
13	0	0	1	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0

17	0	0	0	0	0	0
18	3	5	0	0	0	0
19	0	0	1	0	0	0
20	6	0	0	0	0	0

TABLE 20, PRODUCTION PLAN FREE CAMPAIGN STARTUP COST BULK 5 %

Product\time period	1	2	3	4	5	6
1	6	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	0	0	0
5	0	0	0	0	0	0
6	0	8	0	0	0	0
7	0	0	0	0	0	0
8	0	0	3	0	0	0
9	0	0	0	0	0	0
10	10	10	20	0	0	0
11	0	0	0	0	0	0
12	0	0	1	0	0	0
13	0	0	1	0	0	0
14	5	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	8	0	0	0	0	0
19	0	0	1	0	0	0
20	6	0	0	0	0	0

TABLE 21, PRODUCTION PLAN FREE CAMPAIGN STARTUP COST BULK 10 %